Five-brane effective action in M-theory

Edward Witten [1]

School of Natural Sciences, Institute for Advanced Study, Olden Lane, Princeton, NJ 08540, USA

Received 11 November 1996; revised 28 November 1996

Abstract

On the world-volume of an M-theory five-brane propagates a two-form with self-dual field strength. As this field is non-Lagrangian, there is no obvious framework for determining its partition function. An analogous problem exists in Type IIB superstring theory for the self-dual five-form. The resolution of these problems and definition of the partition function is explained. A more complete analysis of perturbative anomaly cancellation for M-theory five-branes is also presented, uncovering some surprising details.

1. Introduction

M-theory is the conjectural quantum theory that under certain conditions looks at long-wavelengths like 11-dimensional supergravity [1] and reduces to various string theories in other limits. An important aspect of this theory is that the low energy description in 11 dimensions can be enriched with the addition of certain extended objects, namely membranes or two-branes [2] and five-branes. The five-brane, originally found as a solution of the low energy classical equations [3], is an extended object whose world-volume is a five-dimensional object that sweeps out a six-manifold in space–time.

In a recent paper [4], it was shown that the low energy effective action of M-theory on a closed 11-dimensional spin manifold Q is well-defined. To be more precise, let I_M be the Chern–Simons interaction of the long-wavelength limit of M-theory; schematically $I_M = CGG + CI_8(R)$, where C is the massless three-form, G the gauge-invariant field

---

strength of C, and $I_8(R)$ is a certain eight-form constructed from the Riemann tensor (first obtained in [5,6]). Let det D_{RS} be the path integral of the Rarita–Schwinger field;[2] it is real but not necessarily positive. The long-wavelength limit of the quantum measure of M-theory[3] is a product of manifestly well-defined factors times

$$\det D_{RS} \ e^{iI_M} \tag{1.1}$$

and it was shown in [4] that this product is well-defined (though neither factor is well-defined separately). The main novelty required for this result was a gravitational shift in the quantization law of G. The shifted quantization law says that

$$\left[\frac{G}{2\pi} \right] = \frac{\lambda}{2} + \text{integral cohomology class}, \tag{1.2}$$

where $[G/2\pi]$ is the cohomology class of $G/2\pi$, and $\lambda = p_1(X)/2$ (λ is integral for a spin-manifold Q)[7,8]. One would like to know whether the long-wavelength effective action is still well-defined in the presence of impurities. As partly anticipated above, there are three known kinds of impurities: boundaries of Q (where E_8 supermultiplets are believed to propagate); two-branes; and five-branes. Anomaly cancellation in the presence of boundaries has been demonstrated in [7,8].[4] Membrane world-volume anomalies are rather simple because the world-volume theory is non-chiral. The only issues concern the sign of the path integral of the world-volume fermions and the flux quantization law for G. There is potentially an anomaly affecting the sign of the fermion determinant, and it was shown in [4] that the $\lambda/2$ term in (1.2) gives an additional effect that cancels this anomaly [10–12]. The remaining case,[5] the five-brane, is the principal subject of the present paper. This case is particularly subtle because the five-brane world-volume theory is chiral. There are therefore potential perturbative anomalies, which have been partially analyzed in [6,10] and are considered more fully in Sections 3 and 5 below. One should also consider global anomalies, which can perhaps be treated by an extension of the methods in [9].

But a question more basic than those will be our main interest here: *What is the five-brane partition function?* This must be answered before the question of whether this partition function has anomalies can be formulated. The reason that there is a problem is simple. One of the fields on the five-brane world-volume is a self-dual three-form T; that is, on the world-volume there is a two-form, say β, which is constrained so that its field-strength T is self-dual. (This result was first obtained in the Type IIA context [13].) Such a field, which we will call a chiral two-form, has no Lagrangian formulation, and there is subtlety in defining, even formally, what its partition function should be.

A chiral two-form in six dimensions poses a rather similar problem to a chiral boson ϕ in two dimensions (whose field strength is a self-dual one-form $\partial\phi$). In fact, if the world-

---

[2] As in the case of any Majorana fermi field, this path integral might be more naturally called a Pfaffian rather than a determinant.

[3] By the quantum measure I mean the exponential of the effective action.

[4] To be more precise, perturbative anomaly cancelation was analyzed in those papers; it would be natural to analyze the global anomalies by an extension of ideas in [9].

[5] Still more generally, but beyond our scope here, one could consider membranes ending on boundaries [7] or five-branes [11,12].

volume W of the five-brane is of the form $W = \Sigma \times CP^2$,[6] with Σ a Riemann surface, then in the limit that Σ is much larger than the CP^2, the theory of the chiral boson on W reduces to the theory of one chiral boson on Σ. For one chiral boson, there is no modular-invariant partition function – a modular-invariant partition function for chiral bosons requires an even self-dual lattice [14], and in particular the net number of chiral bosons must be a multiple of 8. So this particular example shows that it is impossible to define for the chiral two-form a partition function that is "modular-invariant," that is, invariant under arbitrary diffeomorphisms of W.

We are not, however, dealing in M-theory with a bare six-manifold W. W is embedded in an 11-dimensional spin-manifold Q, on which there are fields obeying certain relations such as (1.2) and [11] $dT = G$. In Sections 2 and 3, we demonstrate that this additional data are exactly right to make it possible to define a partition function for the chiral two-form, which is invariant (up to an ordinary anomaly) not under all diffeomorphisms of W, but under those that preserve the physical data. (By "an ordinary anomaly", we mean an anomaly that only affects the phase of the partition function, and so can cancel anomalies coming from other interactions or fields.)

In Section 4, we make a digression from five-branes, and show that a similar problem arises, and can be treated similarly, for the chiral four-form of Type IIB superstrings – that is, the massless field whose field strength is a self-dual five-form. This problem was avoided in a previous analysis of Type IIB global anomalies [9] by considering only 10-manifolds of vanishing fifth Betti number; for such space–times the problem does not arise, as will be apparent.

Finally, in Section 5, we look at the ordinary anomalies. To be more precise, we give a more complete analysis of five-brane perturbative anomalies than has been done hitherto. The analysis proves to require some surprising novelty, and we actually get a complete answer only in the Type IIA case, not in M-theory.

My interest in these issues came originally from thinking about non-perturbative super-potentials generated by five-brane instantons [15]. Such an instanton contribution is roughly $e^{-\Phi} P$, where Φ is a chiral superfield whose real part is the volume of the instanton and P comes from quantum fluctuations. Roughly speaking, zeros of P determine the supersymmetric vacua. In [15], the zeros coming from fermion zero modes were analyzed. Additional zeroes can come from the behavior of the partition function of the chiral two-form. To determine the locations of the zeros requires the considerations of the present paper.

2. Review of chiral scalars in two dimensions

2.1. Basic framework

In this section, we review some aspects of the theory of a chiral scalar in two dimensions, and explain how some of the ideas generalize above two dimensions.

---

[6] This particular W is not a spin manifold, but it could be embedded in an 11-dimensional spin manifold Q, and so can arise in M-theory.

Since there is no effective Lagrangian formulation for a chiral scalar, there is no natural way to use path integrals to determine the partition function. One approach is to take a non-chiral boson, which does have a Lagrangian and a well-defined partition function Z, and try to write Z as the absolute value squared of the chiral boson partition function. As is well known, things do not work so simply. Z is not the square of a holomorphic function but a sum of such squares; the number of terms depends on the periodicity or radius of the non-chiral boson.

As will become clear, the radius of relevance to our problem is the free-fermion radius. This corresponds to momenta which take values in the unimodular, but not even, one-dimensional lattice $\mathbb{Z}$, endowed with the quadratic form $f(x) = x^2$ (that is, the unit element in $\mathbb{Z}$ is of length one). At the free-fermion radius, the partition function of the free boson on a Riemann surface Σ of genus g can be written as [7]

$$Z = \sum_\alpha |\Theta_\alpha|^2. \tag{2.1}$$

Here α runs over the spin structures on Σ, and Θ_α is the free-fermion partition function with spin structure α. In genus one, $\Theta_\alpha = \theta_\alpha/\eta$ where θ_α is a theta function and η is the Dedekind eta function. The functions Θ_α are the candidate partition functions of the chiral boson. Our problem is to understand – in a form suitable for generalization to five-branes or to Type IIB in 10 dimensions – the fact that a choice of spin structure enables one to pick out a particular one of the Θ_α's.

It is natural to couple the chiral boson to a background gauge field A, which we can think of as a connection on a line bundle $\mathcal{T}$. The field strength is then $\Lambda = d\phi + A$, and Λ is no longer closed; it obeys $d\Lambda = F$ with F the field strength of A. The existence of a gauge-invariant field Λ with $d\Lambda = F$ means that (unless we make operator insertions that modify that equation) $c_1(\mathcal{T})$ must vanish, so $\mathcal{T}$ is topologically trivial. Each choice of connection A gives (via the $\bar{\partial}$ operator) a complex structure to $\mathcal{T}$. The moduli space of such complex structures is the Jacobian J_Σ of Σ. One can identify J_Σ as $H^1(\Sigma, \mathbb{R})/H^1(\Sigma, \mathbb{Z})$.

The coupling to a background gauge field has an analog in the five-brane problem. The relevant background field here is the three-form C of 11-dimensional supergravity (restricted to the five-brane world-volume W). The field strength of C is $G = dC$. The self-dual three-form T on W obeys [11] $dT = G$, and this equation is obviously quite analogous to $d\Lambda = F$. The coupling to C plays for the self-dual three-form a role quite similar to the coupling to a background gauge field for the self-dual one-form. For reasons that will be explained, we will be able to think of C as defining a point in the "intermediate Jacobian" $H^3(W, \mathbb{R})/H^3(W, \mathbb{Z})$.

The equation $dT = G$ means that the restriction of G to the five-brane world-volume W must be zero cohomologically (just as above the relation $d\Lambda = F$ implied that $\mathcal{T}$ is trivial topologically). It follows, therefore, given (1.2), that *the restriction of λ to W is even*. This will be important in Section 3.

---

[7] A thorough and direct study of this is in [16]. Note that this problem is much simpler than the generic case of rational conformal field theory in that the space of conformal blocks has a distinguished basis, given by the Θ_α.

2.2. Embedding in non-chiral theory

We will look now more closely at the embedding of the chiral scalar in the theory of a non-chiral scalar. In doing so, we work on a two-dimensional surface Σ of *Euclidean* signature, so the Lagrangian will be complex, and the self-duality condition reads $d\phi = i * d\phi$.

We consider thus a scalar field ϕ, with a periodicity $\phi \to \phi + 2\pi$, and a Lagrangian (at the free-fermion radius)

$$L = \frac{1}{8\pi} \int d^2 x \sqrt{g} \, g^{ij} \partial_i \phi \partial_j \phi. \tag{2.2}$$

We introduce as explained above a $U(1)$ gauge field A with gauge transformation law $\delta A_i = -\partial_i a$, $\delta \phi = a$, and Lagrangian [8]

$$L = \frac{1}{8\pi} \int_{\Sigma} d^2 x \sqrt{g} \, g^{ij} (\partial_i \phi + A_i)(\partial_j \phi + A_j) + \frac{i}{4\pi} \int_{\Sigma} \phi \epsilon^{ij} \partial_i A_j. \tag{2.3}$$

(The transformation law of ϕ means that ϕ is a section of the circle bundle on which A is a connection, which therefore must be topologically trivial as explained earlier in a somewhat different way.) The point of this particular coupling is that A has been coupled only to the chiral part of ϕ. This can be made clear by introducing local complex coordinates $z, \bar{z}$ (with orientations so that $\epsilon_{z\bar{z}} = -\epsilon_{\bar{z}z} = i$) and expanding the Lagrangian to get

$$L = \frac{1}{4\pi} \int_{\Sigma} |dz \wedge d\bar{z}| (\partial_z \phi \partial_{\bar{z}} \phi + 2\partial_z \phi A_{\bar{z}} + A_z A_{\bar{z}}). \tag{2.4}$$

Thus, only $A_{\bar{z}}$ and not A_z couples to the quantum field ϕ. It follows that in a suitable sense the partition function

$$Z(A) = \int d\phi \, e^{-L} \tag{2.5}$$

depends holomorphically on $A_{\bar{z}}$.

To be more precise, introduce a complex structure on the space of gauge fields in which $A_{\bar{z}}$ is holomorphic and A_z is antiholomorphic. A holomorphic line bundle $\mathcal{L}$ over the space of gauge fields can be defined by taking $\mathcal{L}$ to be the trivial line bundle endowed with the covariant derivatives

$$\frac{D}{DA_i} = \frac{\delta}{\delta A_i} + \frac{i\epsilon_{ij}}{4\pi} A_j. \tag{2.6}$$

This means that

$$\frac{D}{DA_z} = \frac{\delta}{\delta A_z} + \frac{A_{\bar{z}}}{4\pi}. \tag{2.7}$$

---

[8] A more complete account of the point of view that follows, in the more general context of the WZW model, is in [17].

To show that this connection defines a *holomorphic* structure on $\mathcal{L}$, one must check from (2.6) that

$$\left[\frac{D}{DA_z(x)}, \frac{D}{DA_z(x')} \right] = 0. \tag{2.8}$$

This is straightforward. Then using the fact that A_z appears in L only in the $A_z A_{\bar{z}}$ term in (2.4), together with the explicit form of (2.7), one finds that $(D/DA_z)e^{-L} = 0$. The partition function $Z = \int d\phi \, e^{-L}$ therefore obeys the same equation:

$$\frac{D}{DA_z} Z = 0. \tag{2.9}$$

Thus, *the partition function is a holomorphic section of $\mathcal{L}$,* over the space of all connections.

The fact that the partition function is most naturally seen as a section of a line bundle rather than a function is related to the fact that the Lagrangian (2.3) is not gauge-invariant. Under gauge transformations $\delta\phi = a$, $\delta A = -\,da$, one has

$$\delta L = \frac{i}{4\pi} \int_{\Sigma} a \, \epsilon^{ij} \partial_i A_j. \tag{2.10}$$

The partition function thus obeys not standard gauge-invariance, which would read

$$\partial_i \frac{\delta}{\delta A_i} Z = 0, \tag{2.11}$$

but rather

$$\left(\partial_i \frac{D}{DA_i} - \frac{i\epsilon^{ij} F_{ij}}{4\pi} \right) Z = 0. \tag{2.12}$$

This means that the partition function Z is invariant under infinitesimal gauge transformations if interpreted as a section of $\mathcal{L}$ rather than as a function. The operators on the left-hand side of (2.12) are the ones that generate infinitesimal gauge transformations when acting on sections of $\mathcal{L}$. This is a special case of a more general assertion about the WZW model; see [17], Eq. (2.17).

So far we have coupled to arbitrary (topologically trivial) background gauge fields A. The partition function is constrained by the two conditions of gauge covariance (2.12) and holomorphy (2.9). Taken together, the two conditions determine how the partition function transforms under a *complex* gauge transformation $\delta A_{\bar{z}} = -\bar{\partial}\epsilon$, $\delta A_z = -\partial\bar{\epsilon}$. By a complex gauge transformation one can reduce to $F = 0$. So there is no loss of generality in considering only the coupling to background fields with $F = 0$.

The gauge field A, modulo infinitesimal gauge transformations and with $F = 0$, defines a point in $H^1(\Sigma, \mathbb{R})$. With infinitesimal gauge transformations acting as in (2.12), $\mathcal{L}$ descends to a line bundle – which we will also call $\mathcal{L}$ – on $H^1(\Sigma, \mathbb{R})$. Z, being gauge invariant in the sense of (2.12), descends to a section of $\mathcal{L}$ over $H^1(\Sigma, \mathbb{R})$. However, we want to divide by the "big gauge transformations", and interpret Z as a section of a line bundle over $H^1(\Sigma, \mathbb{R})/H^1(\Sigma, \mathbb{Z})$, which is the Jacobian J_Σ of Σ.

There is no natural choice of how the "big gauge transformations" should act on $\mathcal{L}$. Why this is so is explained below. There are in fact different and equally natural line bundles $\mathcal{L}_\alpha$, obtained by differing choices of how the big gauge transformations act on $\mathcal{L}$.

Related to this, the partition function Z considered so far is not really the object we want. It includes the contributions of the "wrong chirality" part of ϕ which though decoupled from A is still present in the partition function. If one tries to carry out holomorphic factorization to suppress the wrong chirality field, one finds as in (2.1) that Z is a sum of terms, each involving (as will become clear) a different $\mathcal{L}_\alpha$.

What has been said so far applies to a chiral $2k$-form β in $4k + 2$ dimensions for any k. The arguments have been presented in such a way that they carry over without any essential change. We repeat the story briefly. Letting C be a background $2k + 1$-form on a $4k + 2$-dimensional manifold W, and $G = dC$, consider a non-chiral $2k$-form β on W, with coupling to C given schematically by $L_\beta = \int_W (|\,d\beta + C|^2 + \beta \wedge G)$. Choose the coefficients in L_β chosen so that only the anti-self-dual part of C couples. The "wrong chirality" part of β is thus present, but decoupled. Under infinitesimal gauge transformations $\delta\beta = \alpha$, $\delta C = -\,d\alpha$, (with α a $2k$-form), L_β changes by

$$\delta L_\beta \sim \int_W \alpha \wedge G. \tag{2.13}$$

This means that the partition function Z should be understood as a section of a line bundle $\mathcal{L}$ over the space of C's. The fact that only the anti-self-dual part of C couples means that Z is actually a holomorphic section. Holomorphy plus gauge-invariance imply invariance under complexified gauge transformations, which can be used to reduce to the case that C is a harmonic $2k + 1$-form, defining a point in $H^{2k+1}(W, \mathbb{R})$. The partition function is thus naturally induced from a section of a line bundle over $H^{2k+1}(W, \mathbb{R})$. To define the theory of the chiral two-form, one must carry out holomorphic factorization, throw away the antichiral contribution, and divide by "big gauge transformations" so as to descend to $J_W = H^{2k+1}(W, \mathbb{R})/H^{2k+1}(W, \mathbb{Z})$, which is known as the "intermediate Jacobian" of W. Holomorphic factorization leads to a sum of different terms, each associated with a different line bundle on J_W. The key to finding the partition function of the self-dual scalar is to find a way to pick out a particular term from this sum, or equivalently a particular line bundle on J_W.

2.3. The line bundle

Now let us go back to the free-fermion approach to the non-chiral scalar in two dimensions. The factorization (2.1) still holds after coupling to A, but now the Θ_α are functions of A as well as of the complex structure τ of Σ.[9]

---

[9] If A is coupled chirally, as above, the $\overline{\Theta}_\alpha$ still depend on τ only; if one uses a vector-like coupling of A, the $\overline{\Theta}_\alpha$ would still be the complex conjugates of the Θ_α.

The Θ_α can be written as $\Theta_\alpha(\tau, A) = \theta_\alpha(\tau, A)/\tilde{\eta}(\tau)$, where the $\theta_\alpha(\tau, A)$ are theta functions on the Jacobian J_Σ (we recall that each A determines a point on J_Σ) and $\tilde{\eta}(\tau)$ depends only on τ.

Because the coupling of a chiral scalar to a gauge field violates gauge invariance, the partition functions $\Theta_\alpha(\tau, A)$, in their dependence on A, are not naturally understood as functions but as sections of appropriate line bundles $\mathcal{L}_\alpha$ over the Jacobian. We have essentially seen this already from the bosonic point of view. Moreover – as already indicated – the θ_α of different α are all sections of *different* line bundles over J_Σ. That is the key to our problem. It means that once one finds the line bundle $\mathcal{L}_\alpha$ on the Jacobian, a corresponding partition function Θ_α of the chiral scalar is naturally determined. In fact, each $\mathcal{L}_\alpha$ has (up to a complex multiple) only one holomorphic section, as we will see, and so automatically determines its own theta function Θ_α. [10] This approach, in which the partition function is defined by first finding the right line bundle, may at first sound esoteric, but can be implemented quite uniformly for our three cases: the chiral scalar in two dimensions; the self-dual three-form on the five-brane world-volume; and the self-dual five-form of Type IIB theory in 10 dimensions.

2.4. Line bundles on the Jacobian

Our problem, then, is to study line bundles on the Jacobian J_Σ of a Riemann surface Σ, or the intermediate Jacobian J_W of a $4k + 2$-dimensional manifold W.

To begin with, consider more generally a $2n$-dimensional torus $J = \mathbb{R}^{2n}/\Gamma$, where Γ is a rank $2n$ lattice in $\mathbb{R}^{2n}$. A "principal polarization" of J is an element $\omega \in H^2(J, \mathbb{Z})$ such that

$$\int_J \frac{\omega^n}{n!} = 1. \tag{2.14}$$

ω can be represented by a two-form on J which is uniquely determined if we require it to be invariant under translations of J; this will always be assumed.

An example of such an ω is as follows. Let x^i, y_j, $i, j = 1, \ldots, n$, be coordinates on $\mathbb{R}^{2n}$ such that Γ is spanned by unit vectors e_i and f^j in the x^i and y_j directions, respectively. Then $\omega = \sum_i dx^i \wedge dy_i$ defines a principal polarization. Conversely, any translation-invariant two-form ω representing a principal polarization can be put in such a form by a suitable choice of coordinates.

---

[10] It may appear that the uniqueness (up to an A-independent but possibly τ-dependent multiple) of the holomorphic section of $\mathcal{L}_\alpha$ determines only the A-dependence and not the τ-dependence of the chiral partition function Θ_α. But the partition function Θ_α obeys a heat equation (a special case of the KZB equation obeyed by the conformal blocks of the WZW model) that determines its τ-dependence when the A-dependence is known. The heat equation is a consequence of the Sugawara construction: the stress tensor of the chiral boson is the square of the current $\partial_z \phi + A_z$ that couples to $A_{\bar{z}}$. Likewise, the stress tensor of the chiral $2k$-form is a quadratic expression in the current $(d\beta + C)^+$ (the superscript $+$ refers to a projection on the self-dual part), as a result of which there is a heat equation that determines the dependence on the metric of W from the C dependence.

In the example just given, the pairings of the two-form ω with the vectors e_i and f^j are

$$\omega(e_i, f^j) = \delta_i^j, \qquad \omega(e_i, e_j) = \omega(f^i, f^j) = 0. \tag{2.15}$$

Thus, on the lattice Γ, ω defines an integer-valued antisymmetric pairing which is non-degenerate and "minimal" (the Pfaffian of ω is as small as possible – equivalent to assertion (2.14) that the volume of a unit cell is 1).

If J is the intermediate Jacobian of a $4k+2$-dimensional manifold W (so $J = H^{2k+1}(W, \mathbb{R})/H^{2k+1}(W, \mathbb{Z})$), then Γ corresponds to the lattice $H^{2k+1}(W, \mathbb{Z})$ (mod torsion). The intersection pairing or cup product $H^{2k+1}(W, \mathbb{Z}) \times H^{2k+1}(W, \mathbb{Z}) \to H^{4k+2}(W, \mathbb{Z}) \cong \mathbb{Z}$ defines an integer-valued antisymmetric pairing on Γ which is isomorphic to (2.15) in suitable coordinates. (For instance, if W is a Riemann surface, then the intersection pairing on H^1 can be put in the form (2.15) by a choice of A and B cycles.) Therefore, the intermediate Jacobians of interest to us are always naturally endowed with a principal polarization. The curvature form ω associated with this polarization actually can be seen in the curvature of the connection introduced in (2.6).

Given a metric on W, the intermediate Jacobian J_W has a natural metric defined as follows. The tangent space to J_W is the space $\mathcal{W}$ of harmonic $2k + 1$-forms on W. The Hodge $*$ operator maps $\mathcal{W} \to \mathcal{W}$ with $*^2 = -1$, so it defines a complex structure on J_W. A metric on J_W can be defined by saying that for $C \in \mathcal{W}$, $\|C\|^2 = \int_W C \wedge * C$. This metric on J_W is translation-invariant and Kähler. The associated Kähler form is our friend the polarization of J_W, $\omega(C, C') = \int_W C \wedge C'$. In particular, ω is of type $(1, 1)$ and positive in this Kähler metric.

In fact, the partition function of the chiral $2k$-form only depends on the metric on J_W and the line bundle $\mathcal{L}$, up to elementary factors determined by the anomalous Ward identities. This is a striking simplification, as the metric on W (on which the partition function might depend a priori) depends on infinitely many parameters, but the metric on J_W, being translation-invariant, depends on only finitely many parameters. This result can be seen using the heat equation described in the footnote at the end of Section 2.1: since the stress tensor is quadratic in the "currents", the response to a change in the metric W can be expressed in terms of the response to a change in the background C-field, and just as for a chiral scalar in two dimensions, the only non-elementary terms that arise are those that involve the change in metric of J_W under change in metric of W.

2.5. From polarization to line bundle

Since line bundles are classified topologically by their first Chern class, there is topologically up to isomorphism a unique line bundle $\mathcal{L}$ on J_W whose first Chern class is $c_1(\mathcal{L}) = \omega$. However, we need to describe $\mathcal{L}$ much more precisely. We want to find a $U(1)$ connection B on $\mathcal{L}$, whose curvature $F = dB$ equals $2\pi\omega$. This would lead as follows to a definition of the partition function of the chiral $2k$-form. Since ω is of type $(1, 1)$ in the complex

structure on J_W, the connection B determines a complex structure on $\mathcal{L}$. The index of the $\bar{\partial}$ operator on J, with values in $\mathcal{L}$, is

$$\sum_{i=0}^{\dim_C J_W} (-1)^i \dim H^i(J_W, \mathcal{L}) = \int_J e^{c_1(\mathcal{L})} \text{Td}(J_W) = \int_J e^\omega = 1. \tag{2.16}$$

(Here Td is the Todd genus; since J_W has a flat metric, $\text{Td}(J_W) = 1$. We also use (2.14).) Since ω is positive, the cohomology $H^i(J_W, \mathcal{L}) = 0$ for $i > 0$,[11] so the index formula actually asserts that $H^0(J_W, \mathcal{L})$ is one-dimensional. Thus, the line bundle $\mathcal{L}$ has (up to a complex multiple) a single holomorphic section. This section is the desired partition function of the chiral $2k$-form, at least as regards the C-dependence. But since the different terms in the holomorphic factorization of the non-chiral $2k$-form have different C-dependence, once we know the C-dependence we are essentially done. (More fundamentally, as remarked in a previous footnote, the τ-dependence can be determined once the C-dependence is known from the fact that the chiral partition function obeys an analog of the KZB equation for conformal blocks of the WZW model.)

So our basic problem is really to find a $U(1)$ gauge field on J_W whose curvature is $F = 2\pi\omega$. Now, on a simplyconnected manifold, a $U(1)$ gauge field B is determined up to isomorphism by its curvature. We are dealing instead with the torus $J = \mathbb{R}^{2n}/\Gamma$. (In what follows, J can be any torus with a principal polarization, as opposed to J_W which has additional structure such as a metric and complex structure. So we drop the subscript W for the time being.) To fix B we must give, in addition to the curvature, the holonomies around non-contractible cycles in J.

If a is a lattice point in $\Gamma \subset \mathbb{R}^{2n}$, then the straight line from the origin in $\mathbb{R}^{2n}$ to a determines a closed curve $C(a)$ in J. Let $H(a) = \exp(i \int_{C(a)} B)$ be the holonomy of B around $C(a)$. B will be completely fixed if the $H(a)$ are given. We would like to pick the $H(a)$'s to preserve, as much as possible, the invariance under the symplectic group $Sp(2n, \mathbb{Z})$ (which acts on Γ preserving ω). The most obvious choice would be $H(a) = 1$ for all a. This is impossible for the following reason.

The $H(a)$'s are constrained as follows. If a and b are any two lattice points, then

$$H(a + b) = H(a)H(b)(-1)^{\omega(a,b)}. \tag{2.17}$$

This is obtained as follows. The lattice points 0, a, b, and $a+b$ are vertices of a parallelogram $\Delta_{a,b}$ through which the magnetic flux is $2\pi\omega(a, b)$. The lattice points 0, a, and $a + b$ are vertices of a triangle $T_{a,b}$ which is just half of $\Delta_{a,b}$; the magnetic flux through $T_{a,b}$ is $\pi\omega(a, b)$. The sides of $T_{a,b}$ are C_a, C_b, and C_{a+b}. So (2.17) is the usual relation, following from Stokes's theorem, between the magnetic flux through a surface S – in this case $T_{a,b}$ – and the holonomy around the boundary of S.

A basic question is now whether it is possible to pick a line bundle in a way invariant under the $Sp(2n; \mathbb{Z})$ that acts on J preserving the polarization. Eq. (2.17) shows that this

---

[11] This is true by the Kodaira vanishing theorem, which uses the fact that $\bar{\partial}^*\bar{\partial} + \bar{\partial}\,\bar{\partial}^*$ is strictly positive, in this situation, for $i > 0$.

is impossible; $Sp(2n; \mathbb{Z})$ would require that $H(a) = H(b)$ for all primitive lattice vectors a and b, and this is incompatible with (2.17). [12] In particular, (2.17) does not permit us to take $H(a) = 1$ for all a. Eq. (2.17) does, however, allow $H(a)^2 = 1$ for all a. The numbers $H(a)^2$ would be the holonomies around C_a of the connection $2B$ on the line bundle $\mathcal{L}^2$, which we will call $\mathcal{M}$. Thus, there is a completely natural, $Sp(2n, \mathbb{Z})$-invariant line bundle $\mathcal{M}$ with holonomy $+1$ around each C_a and first Chern class 2ω. The factor of 2 means that (using the index formula as above) $H^0(J, \mathcal{M})$ is of dimension 2^g. We need a line bundle $\mathcal{L}$ of first Chern class $1 \cdot \omega$, with just one holomorphic section, which will be our partition function.

In searching for an $\mathcal{L}$ that is "as canonical as possible", we can require that $\mathcal{L}^2$ is isomorphic to $\mathcal{M}$, or equivalently that the holonomies $H(a)$ of $\mathcal{L}$ are all ± 1. The number of such $\mathcal{L}$'s is 2^{2n}; they differ by $H(a) \to H(a)(-1)^{\epsilon(a)}$ with $\epsilon \in H^1(J, \mathbb{Z}_2)$. Thus, the search for $\mathcal{L}$ is reduced to the selection among a finite set of possibilities.

In case $J = J_\Sigma$ is the Jacobian of a Riemann surface Σ, we already know from bose-fermi equivalence (that is, from the holomorphic factorization of the free boson at the free-fermion radius) that to pick an $\mathcal{L}$ out of the 2^{2n} possibilities, what we need is precisely a spin structure on Σ. In the remainder of this section, I will sketch three direct explanations of this fact. The first two, though not needed in the rest of the paper (and therefore not explained below in much detail), are included because they are short and illuminating. The third explanation, though not new [18], is perhaps less well-known. It is this third approach that we will later generalize above two dimensions.

2.6. The determinant bundle

The first approach is directly related to the free-fermion construction of the chiral boson. If we are given a spin structure α on Σ, then we can use the determinant of the Dirac operator to obtain a line bundle on J_Σ.

Thus, given a flat connection A on Σ representing a point in J_Σ, let $\mathcal{T}_A$ be the corresponding flat line bundle on Σ. Let $D_\alpha(A)$ be the Dirac operator with values in $\mathcal{T}_A$, using the spin structure α. The determinant line of $D_\alpha(A)$ is a complex line $\mathcal{L}_{\alpha,A}$, and these fit together as A varies to give the desired line bundle $\mathcal{L}_\alpha \to J_\Sigma$, which can be shown via index theory to have the right first Chern class. [13] It thus has a single holomorphic section, which is in fact the function Θ_α that appears in holomorphic factorization of the non-chiral boson.

---

[12] There is one significant exception; if J is two-dimensional then (2.17) allows $H(a) = -1$ for all primitive a. This corresponds to the fact that on (and only on) a genus one curve, there is a spin structure that is completely diffeormorphism-invariant, namely the "odd" one (the trivial spin bundle). For dim $J > 2$, the possibility $H(a) = -1$ for all primitive a is excluded as one can find vectors a, b with a, b, and $a + b$ all primitive and $\omega(a, b) = 0$.

[13] An important subtlety, which is the reason that one must here use the Dirac operator rather than the $\bar{\partial}$ operator (which does not require a choice of spin structure) is that because $D_\alpha(\mathcal{T}_A)$ has zero index, $\mathcal{L}_A$ depends in an appropriate sense only on the isomorphism class of $\mathcal{T}_A$. That is essential because without making any arbitrary choices, J_Σ parametrizes a family of isomorphism classes of line bundles on Σ, but non-modular-invariant choices are needed to get an actual family of line bundles.

So in particular, a choice of spin structure on Σ gives a choice of line bundle on J_Σ. Of course, the discussion has now brought us back to our starting point (2.1), and if we were interested only in the two-dimensional case we could have spared much of our effort.

2.7. The shifted Jacobian

Now we consider briefly another approach which is less obviously related to physics. We consider the shifted Jacobian $J_{\Sigma,n}$ of Σ, which parametrizes holomorphic line bundles on Σ of degree n. They are all non-canonically isomorphic to the ordinary Jacobian J_Σ. Fixing a line bundle S of degree n, the map $T \to T \otimes S$ (where T is a line bundle of degree zero, defining a point in J_Σ, and $T \otimes S$ therefore has degree n and defines a point in $J_{\Sigma,n}$) is an isomorphism between J_Σ and $J_{\Sigma,n}$. In particular, the existence of this isomorphism means that each $J_{\Sigma,n}$ is naturally endowed with a principal polarization ω.

On $J_{\Sigma,g-1}$, there is actually a completely natural ("modular-invariant") choice of line bundle with first Chern class ω; we call it $\mathcal{L}'$ to distinguish it from the desired line bundle $\mathcal{L}$ on Σ. In fact, a line bundle on a complex manifold can be given by specifying a divisor. On $J_{\Sigma,g-1}$, there is a natural divisor, the Θ divisor, that parametrizes line bundles with a holomorphic section. It can be shown that the associated line bundle $\mathcal{L}'$ on $J_{\Sigma,g-1}$ has first Chern class ω.

Now if S is *any* line bundle on Σ of degree $g-1$, then by using S as explained above to establish an isomorphism between J_Σ and $J_{\Sigma,g-1}$, we can interpret $\mathcal{L}'$ as a line bundle on J_Σ. Thus, to find the desired line bundle on J_Σ all we need to do is to pick an S. There is no natural choice of S. The closest one can come is to set S equal to one of the 2^{2g} spin structures of Σ. (Recall that a spin structure on Σ corresponds to a line bundle of degree $g-1$ whose square is isomorphic to the canonical line bundle; there are 2^{2g} of them, and they are on the smallest possible orbit of the modular group in $J_{\Sigma,g-1}$.) So we get again the expected result: any choice of spin structure gives a choice of line bundle $\mathcal{L}$ on J_Σ with first Chern class ω.

2.8. Chern–Simons theory

Finally, we come to the approach that we will actually use in the rest of this paper, in generalizing above two dimensions. To construct the desired line bundle on the Jacobian of a *two*-dimensional surface, we use the Chern–Simons functional of a gauge field in *three* dimensions.

Let M be a closed oriented three-manifold, and let A be a connection on a $U(1)$ bundle T over M. If T is topologically trivial, so that in a given gauge A is an ordinary one-form, the Chern–Simons functional is

$$I(A) = \frac{1}{2\pi} \int\limits_M \epsilon^{ijk} A_i \partial_j A_k. \tag{2.18}$$

If T is topologically non-trivial, a more powerful approach to defining I is needed. Let X be an oriented four-manifold with boundary M, over which A and T extend, and pick such

an extension. [14] (We assume that the orientation of X is related to that of M by a definite convention, for instance "outward normal first". We will abbreviate the statement that X has boundary M and $(A, \mathcal{T})$ have been extended over X by saying that X has boundary (M, A).) Then define

$$I_X(A) = \frac{1}{2\pi} \int\limits_X \epsilon^{ijkl} \partial_i A_j \partial_k A_l = \frac{1}{8\pi} \int\limits_X \epsilon^{ijkl} F_{ij} F_{kl}. \tag{2.19}$$

The point of this definition is, first of all, that if $\mathcal{T}$ is trivial and A is well-defined as a one-form, then by Stokes's theorem, $I_X(A)$ coincides with $I(A)$ as defined in (2.18). Furthermore, $I_X(A)$ is defined even if $\mathcal{T}$ is topologically non-trivial. What remains is to investigate the extent to which $I_X(A)$ depends on the choice of X (and the extension of A and $\mathcal{T}$). Given another oriented four-manifold X' with boundary (M, A), one can glue X and X' together along their common boundary to make a closed four-manifold Y with a $U(1)$ gauge field A; if we reverse the orientation of X', then the orientations of X and X' match along their common boundary, so that Y has a natural orientation. Then $I_X(A) - I_{X'}(A) = I_Y(A)$, where

$$I_Y(A) = \frac{1}{8\pi} \int\limits_Y \epsilon^{ijkl} F_{ij} F_{kl}. \tag{2.20}$$

The point is now that because the cohomology class $[F/2\pi]$ is integral, $I_Y(A) = 2\pi \cdot \text{integer}$. Hence $I_X(A)$ is independent of X modulo $2\pi \mathbb{Z}$; with the understanding that there is this $2\pi \mathbb{Z}$ ambiguity, we henceforth drop the subscript X and refer simply to $I(A)$.

The fact that $I(A)$ is well-defined modulo 2π means that

$$e^{iI(A)} \tag{2.21}$$

is well-defined. $U(1)$ Chern–Simons gauge theory at level one is in fact defined by the path integral

$$\int DA \; e^{iI(A)}. \tag{2.22}$$

If one considers M to be not a closed three-manifold, but rather $M = \Sigma \times \mathbb{R}$ where Σ is a Riemann surface and $\mathbb{R}$ parametrizes "time", then the moduli space of classical solutions of the Chern–Simons theory is the Jacobian J_Σ (since the equation for a critical point of $I(A)$ is $F = 0$). The quantum Hilbert space is a space of sections of a certain line bundle $\mathcal{M}$ over

---

[14] Since line bundles are classified by maps to CP^∞, the existence of such an X follows from the statement that the oriented bordism group $\Omega_3(CP^\infty)$ vanishes. In fact, a more precise statement (which we will need presently) also holds: the spin bordism group $\Omega_3(CP^\infty)$ vanishes. (This means that if M is a spin manifold with a given spin structure, one can choose X with boundary M so that A and the spin structure of M extends over X.) The following proof of this was sketched by P. Landweber. According to the proposition on p. 354 of [19], $\Omega_3(CP^\infty) = \Omega_5^{\text{Spin}^c}$. According to the theorem on p. 337 of the same book, since there are no rational characteristic numbers in odd dimensions, $\Omega_5^{\text{Spin}^c}$ is determined by Stieffel–Whitney numbers. A Spin^c manifold has $w_1 = w_3 = 0$. A five-dimensional manifold with $w_1 = 0$ has $w_5 = 0$. So a five-dimensional Spin^c manifold has no non-zero Stieffel–Whitney numbers, and hence $\Omega_5^{\text{Spin}^c} = 0$.

J_Σ. $\mathcal{M}$ is roughly the sort of object that we are looking for. However, since the construction is completely diffeomorphism-invariant and in particular no choice of spin structure has entered, $\mathcal{M}$ is modular-invariant and therefore cannot be the desired line bundle with first Chern class ω. $\mathcal{M}$ is in fact, as we will see, the modular-invariant line bundle with first Chern class 2ω (and all $H(a) = 1$) that was seen before.

To better understand $\mathcal{M}$, a first orientation is as follows. Let Σ be a Riemann surface and let N be a three-manifold with boundary Σ. Let A be a connection on a line bundle $\mathcal{T}$ over N, and consider the (level one) Chern–Simons Lagrangian

$$L_{CS} = -iI(A) = -\frac{i}{2\pi} \int_N \epsilon^{ijk} A_i \partial_j A_k. \tag{2.23}$$

In proving gauge-invariance, one must integrate by parts, and one picks up a non-zero surface term because N has a non-empty boundary. In fact, under $\delta A_i = -\partial_i a$, one finds

$$\delta L_{CS} = \frac{i}{2\pi} \int_\Sigma a\epsilon^{ij} \partial_i A_j. \tag{2.24}$$

Just as in our discussion of the chiral boson, this violation of gauge invariance means that $e^{-L_{CS}} = e^{iI(A)}$ is most naturally understood not as a function but as a section of a line bundle $\mathcal{M}$ over the space of gauge fields on Σ. In fact, the right-hand side of (2.24) has the same form as (2.10), but the coefficient is twice as large; the factor of two means that $\mathcal{M}$ is not the desired line bundle $\mathcal{L}$ for the theory of the chiral boson, but rather $\mathcal{M} = \mathcal{L}^2$, as we will show more fully later.

Let us describe more precisely the construction of $\mathcal{M}$. First I sketch a rather down-to-earth approach. To describe up to isomorphism a line bundle $\mathcal{M}$, with $U(1)$ connection, on any given manifold Z, it suffices to define the holonomies of the connection around an arbitrary loop in Z; these must obey certain axioms that will be discussed. In our case, Z is the space of $U(1)$ gauge fields on Σ. Suppose we are given a loop C in the space of gauge fields, that is to say a family of gauge fields on Σ depending on an extra parameter θ ($0 \leq \theta \leq 2\pi$); θ parametrizes the position on C. Making the θ dependence explicit, we write $A_i(x; \theta)$ for this family of gauge fields, where x is a point in Σ. Now on the three-manifold $\Sigma \times S^1$, we introduce the gauge field A_C whose component in the S^1 (or θ) direction is zero, and whose components along Σ are $A_i(x; \theta)$. We define the holonomy of $\mathcal{M}$ around the loop C to be

$$H(C) = e^{iI(A_C)}. \tag{2.25}$$

The property of the $H(C)$'s that is needed for them to be the holonomies of a connection on a line bundle $\mathcal{M}$ is the following. If C_1 and C_2 are two loops in Z that meet at a point $p \in Z$, and $C_1 * C_2$ is the loop made by "joining" C_1 and C_2 at p, one wishes

$$H(C_1 * C_2) = H(C_1)H(C_2). \tag{2.26}$$

In the present case, this is proved as follows. Let D be a "pair of pants" with three boundaries that we associate with C_1, C_2, and $-C_1 * C_2$. (The minus sign refers to a reversal of orientation.) Then the desired relation

$$e^{iI(A_{C_1 * C_2})} = e^{iI(A_{C_1})} e^{iI(A_{C_2})} \tag{2.27}$$

follows from definition (2.19) applied to the four-manifold $X = \Sigma \times D$ whose boundary is the union of $X \times C_1$, $X \times C_2$, and $-X \times C_1 * C_2$.

One can verify directly from definition (2.25) that if C is a straight line on the Jacobian from the origin to any point $a \in H^1(\Sigma, \mathbb{Z})$ (which we represent by a harmonic one-form of the same name), then $H(C) = 1$. This is done by first finding an oriented three-manifold B, of boundary Σ, over which a extends as a closed but no longer harmonic one-form with integral periods. (B can have very simple topology; it can be a "handlebody".) Over the four-manifold $X = B \times \mathbf{S}^1$, the gauge field A_C extends in a fairly obvious way (as a θ-dependent multiple of a, with again θ parametrizing the position on $\mathbf{S}^1$) so that $F \wedge F = 0$ pointwise, and therefore $I_X(A) = 0$. So as promised several times, $\mathcal{M}$ is our friend, the modular-invariant line bundle with all $H(a) = 1$ [20–22].

The foregoing constructs an isomorphism class of line bundles $\mathcal{M}$ but not an actual $\mathcal{M}$. If one is interested in "seeing" an actual $\mathcal{M}$ as precisely as possible, one may use the following somewhat abstract approach (for full details and a number of variants see [20–22]). Given a $U(1)$ gauge field A on Σ, we must construct a one-dimensional complex vector space $\mathcal{M}_A$, in such a way that the $\mathcal{M}_A$ vary nicely with A, as fibers of a complex line bundle $\mathcal{M}$. We simply declare that if N is any three-manifold with boundary Σ over which A extends – that is if N has boundary (Σ, A) [15] – then $\mathcal{M}_A$ has a basis vector $\psi_A(N)$. Given two possible N's, say N_1 and N_2, we must now exhibit a linear relation between $\psi_A(N_1)$ and $\psi_A(N_2)$. This linear relation is chosen to be as follows. Let P be the oriented three-manifold obtained by gluing N_1 and N_2 together (with opposite orientation for N_2) along their common boundary. Let A_P be the connection (on a line bundle over P) made by combining the chosen extensions of A over N_1 and N_2. The relation between $\psi_A(N_1)$ and $\psi_A(N_2)$ is then chosen to be

$$\psi_A(N_1) = e^{iI(A_P)} \psi_A(N_2). \tag{2.28}$$

Given many N_i's, one would construct three-manifolds P_{ij} by gluing N_i to N_j, and set $\psi_A(N_i) = e^{iI(A_{P_{ij}})} \psi_A(N_j)$; these relations can be shown to be compatible using the definition (2.19) of the Chern–Simons functional. Armed with these relations, the $\psi_A(N_i)$ can be interpreted as vectors in a common one-dimensional space $\mathcal{T}_A$, which is the desired fiber of $\mathcal{T}$ over the connection A. For a description of the connection on $\mathcal{T}$ and explanation of its properties, consult the references.

---

[15] If A is a connection on a topologically non-trivial line bundle $\mathcal{T}$ over Σ, then such an N does not exist. Instead one fixes a Riemann surface Σ_0 with a line bundle $\mathcal{T}_0$ of the same first Chern class as $\mathcal{T} \to \Sigma$, and takes N to be a bordism between Σ and Σ_0 (extending the line bundles and connection). The rest of the construction proceeds as in the text.

2.9. Chern–Simons theory at level one-half

Now it is clear what we need in order to get a line bundle whose first Chern class will be ω instead of 2ω. We must consider Chern–Simons theory at level one-half.

In other words, consider the functional

$$\frac{I(A)}{2} = \frac{1}{16\pi} \int\limits_X \epsilon^{ijkl} F_{ij} F_{kl}, \qquad (2.29)$$

where A is a gauge field on an oriented three-manifold M, and X is an oriented four-manifold, of boundary M, over which A has been extended. If $I(A)/2$ were well-defined (independent of the choice of X) modulo 2π, then by using everywhere $e^{iI(A)/2}$ instead of $e^{iI(A)}$ in the above construction, we would get the desired line bundle of first Chern class ω.

The problem is that, for a closed four-manifold Y, the integral

$$\frac{1}{16\pi} \int\limits_Y \epsilon^{ijkl} F_{ij} F_{kl} \qquad (2.30)$$

is an arbitrary integer multiple of π, but not necessarily an *even* integer multiple of π. This comes from the following. Let x be the cohomology class $[F/2\pi]$. Then the integral in (2.30) can be interpreted as πx^2. In general, the class x is integral, so x^2 is integral, but subject to no other general restrictions.

Suppose, however, that Y is a spin manifold. If so, the intersection form on $H^2(Y, \mathbb{Z})$ is even: for arbitrary $x \in H^2(Y, \mathbb{Z})$, x^2 is an *even* integer. This is just what we need. If we may assume that the four-manifolds Y in (2.30) are always spin, then (2.30) would indeed always be an integral multiple of 2π.

If one wants to encounter only four-manifolds with spin structure, one must begin by only considering three-manifolds with a chosen spin structure. Thus, let M be an oriented three-manifold with a chosen spin structure. Given a $U(1)$ gauge field A over M, we would like to define $I(A)/2$. (The definition will in general depend on the spin structure of M.) The definition is made by formula (2.29), where X is an arbitrary oriented four-manifold of boundary M, to which A and the spin structure of M extend. (A proof that such X's exist was sketched in a previous footnote.) Any two such X's will glue together to make a *spin*-manifold Y, for which (2.30) will be an integral multiple of 2π. Therefore, if $I(A)/2$ is defined using only X's of the indicated type, then the definition of $I(A)/2$ in (2.29) is independent of the choice of X modulo 2π.

Now let us tidy up a few details and solve our problem. We are given a Riemann surface Σ with a spin structure α. We want to find a line bundle $\mathcal{L}_\alpha$ over the Jacobian J_Σ, with a connection, compatible with the polarization of J_Σ. Given a loop C in the space of connections on Σ, we build (as above) the associated gauge field $A_C = A_i(x; \theta)$ on the three-manifold $M = \Sigma \times S^1$. We give M the spin structure which is the product of the spin structure α on Σ with the "Neveu–Schwarz" spin structure on S^1 (this is the "antiperiodic"

spin structure, the one which arises if S^1 is regarded as the boundary of a disk). We then characterize $\mathcal{L}$ by declaring its holonomy around C to be

$$H(C) = e^{iI(A_C)/2}. \tag{2.31}$$

The need to use the antiperiodic spin structure on S^1 emerges when one tries to prove that $H(C_1)H(C_2) = H(C_1 * C_2)$. The proof of this involves the four-manifold $\Sigma \times D$, where D is a "pair of pants" (with C_1, C_2, and $-C_1 * C_2$ on the boundary). The reason that one must use the antiperiodic spin structure on the boundary components is that (i) one needs to use the same spin structure on each boundary component, to treat them all symmetrically (or the $H(C)$'s will not obey the appropriate factorization); (ii) the spin structure on the boundary of D must extend over D. The two conditions together are satisfied precisely if the spin structure on the boundary is antiperiodic.

2.10. A note on bordism

Happily, this is the end of the story for the chiral boson, at least for the present paper. But with an eye to generalizations, we will say a word about how bordism really enters the construction.

The above recipe for defining the line bundles $\mathcal{L}_\alpha$ used the fact that if M is a spin three-manifold with line bundle $\mathcal{T}$, then there is a four-manifold X, with boundary M, over which the spin structure of M and the line bundle $\mathcal{T}$ extend. This is described mathematically by saying that $\Omega_3^{\mathrm{spin}}(CP^\infty) = 0$, a fact discussed in a previous footnote.

What would we say instead if $\Omega_3^{\mathrm{spin}}(CP^\infty)$ were non-zero? To make the discussion definite, suppose that $\Omega_3^{\mathrm{spin}}(CP^\infty)$ were $\mathbb{Z}_2$. Pick a three-manifold M_0 with a $U(1)$ gauge field A_0 such that the pair (M_0, A_0) represents the non-zero element of $\Omega_3^{\mathrm{spin}}(CP^\infty)$. There is then no natural way to define $I(A_0)/2$, since no appropriate X exists for the definition (2.29). However, if $\Omega_3^{\mathrm{spin}}(CP^\infty) = \mathbb{Z}_2$, then there exists an X whose boundary is *two* copies of (M_0, A_0). By using this X in (2.29), we can define $w = 2 \cdot I(A_0)/2$. There are now two candidates (namely $w/2$ and $\pi + w/2$) for what $I(A_0)/2$ should be. Pick one of them.

Once $I(A_0)/2$ has been defined, one can define $I(A)/2$ for any $U(1)$ gauge field A on a three-manifold M. In fact, the pair (M, A) is either the boundary of some X or is bordant to (M_0, A_0) via some X (that is, there is an X whose boundary is either (M, A) or $(M, A) - (M_0, A_0)$, where the minus sign represents a reversal of orientation). Taking the integral in (2.29) to define either $I(A)/2$ or $I(A)/2 - I(A_0)/2$ as the case may be, we get (since $I(A_0)/2$ has already been defined) a general definition of $I(A)/2$. This definition can be used to define the line bundle $\mathcal{L}_\alpha \to J_\Sigma$ for every spin surface (Σ, α). The definition is invariant under α-preserving diffeomorphisms of Σ, but will depend on the choice that was made in defining $I(A_0)/2$. Thus, we get a definition of the $\mathcal{L}_\alpha$'s that is diffeomorphism-invariant to the expected extent, but depends on a Σ-independent choice in the definition of $I(A_0)/2$. This choice is roughly analogous to the choice of a discrete theta angle.

The generalization from $\mathbb{Z}_2$ to an arbitrary group is easy to state. Let $H = \text{Hom}(\Omega_3^{\text{spin}}(CP^\infty), U(1))$. Then H classifies the possible definitions of $I(A)/2$, and thus the possible recipes for the association $(\Sigma, \alpha) \rightarrow \mathcal{L}_\alpha$. The sort of "discrete theta angle" that has just been described will turn out not to appear for five-branes – as the relevant bordism group is zero – but may appear in other, similar examples.

3. Chiral two-form on five-brane world-volume

In this section, we will use a differential form notation to avoid a proliferation of unilluminating constants. For instance, if M is a three-manifold with $U(1)$ gauge field A, and X is a four-manifold with boundary (M, A), we write $x = F/2\pi$, and we define the Chern–Simons functional simply as

$$I(A) = 2\pi \int_X x \wedge x. \tag{3.1}$$

This is aimed to make it clear that – as $\int_X x \wedge x$ would be an integer for X a *closed* four-manifold – $I(A)$ is well-defined with values in $\mathbb{R}/2\pi\mathbb{Z}$. We also will omit wedge products and write x^2 for $x \wedge x$.

Now, consider a chiral two-form β on a six-manifold W. First we consider W in isolation; then we consider a more general case – relevant to M-theory – in which W is embedded in an 11-dimensional spin manifold Q, satisfying certain physical conditions.

We know from Section 2 that to study the chiral two-form, we should introduce the intermediate Jacobian J_W of W. This carries a natural polarization $\omega \in H^2(J_W, \mathbb{Z})$, and we must find a line bundle on J_W whose first Chern class is ω. The partition function of the chiral two-form is then uniquely determined.

We also know that to find a line bundle on J_W, we should consider Chern–Simons theory in *seven* dimensions. Let then M be a seven-manifold with a three-form field C. Actually, the field C we want is not a three-form in the standard sense. The field strength $G = dC$ is allowed to have 2π periods (this statement will soon receive a slight modification), and the gauge transformations $C \rightarrow C + d\epsilon$ (ϵ a two-form) are to be supplemented by "big" gauge transformations adding to C a closed three-form with 2π periods. The relation of C to a conventional three-form is just like the relation of a $U(1)$ gauge field to a conventional one-form. Anyway, given such a C, we want to define the Chern–Simons functional $I(C)$. This is done in a fashion that should be familiar. Let X be an oriented eight-manifold with boundary M over which C extends. [16] We describe this in brief by saying that X has boundary (M, C). Let $x = G/2\pi$. The Chern–Simons functional is then

---

[16] Existence of such an X depends on whether the pair (M, G) vanishes in $\Omega_7(K(\mathbb{Z}, 4))$ (which classifies bordism classes of seven-manifolds with a four-dimensional cohomology class). I do not know if that group vanishes; if not, the considerations raised at the end of Section 2 will enter. But presently we will put a spin condition on M, and then the relevant bordism group becomes $\Omega_7^{\text{spin}}(K(\mathbb{Z}, 4))$, which does vanish by a result of Stong [23].

$$I(C) = 2\pi \int_X x^2. \tag{3.2}$$

This definition makes it clear that $I(C)$ is well-defined modulo 2π. Therefore, $I(C)$ can be used exactly as in Section 2 to define a line bundle $\mathcal{M}$ over the intermediate Jacobian J_W of a six-manifold. However, since the construction is completely diffeomorphism-invariant, $\mathcal{M}$ can hardly be the line bundle we want. In fact, rather as for the chiral scalar in two dimensions, $\mathcal{M}$ is the line bundle of first Chern class 2ω and holonomy one around every straight line in J_W.

To make progress, we must define $I(C)/2$ modulo 2π. Then the Chern–Simons construction of line bundles, starting with $I(C)/2$, will give the right line bundle.

To define $I(C)/2$ modulo 2π, some additional structure is needed. Recalling the case of the chiral scalar, and with an eye on physics, in which spinors are present, one's first thought is to assume that M is a spin manifold with a chosen spin structure (and restrict X so that the spin structure of M extends over X – such an X exists as noted in the last footnote). If it were the case that the intersection form on $H^4(X, \mathbb{Z})$, for an eight-dimensional spin manifold X, were always even, our problem would be solved. The even-ness of x^2 in the spin case would give a factor of two in (3.2), so that $I(C)/2$ would be well-defined modulo 2π.

It is, however, not true that the intersection form on the middle-dimensional cohomology of an eight-dimensional spin manifold is even. Indeed, the quaternionic projective space HP^2 is a simple counterexample. ($H^4(HP^2, \mathbb{Z}) = \mathbb{Z}$, and the intersection form is the unimodular, but not even, form with $f(x) = x^2$.) Instead there is the following relation. Let p_1 be the first Pontryagin class of X. In any dimension, there is a canonical way to divide the first Pontryagin class of a spin manifold by two to get an integral class that we will call λ. Then for any $v \in H^4(X, \mathbb{Z})$,

$$v^2 = v \cdot \lambda \quad \text{modulo } 2. \tag{3.3}$$

(A proof of this using E_8 index theory is at the end of Section 4 of [4].) This can be rewritten as the statement that

$$\frac{1}{2} \int_X \left(\left(v - \frac{\lambda}{2} \right)^2 - \frac{\lambda^2}{4} \right) \in \mathbb{Z}. \tag{3.4}$$

This suggests that we should be the following. Instead of asking that $G/2\pi$ should be integral, we require

$$\left[\frac{G}{2\pi} \right] = \frac{\lambda}{2} - v \tag{3.5}$$

with v an integral class. This is in fact the correct quantization law for G in M-theory [4]. Then, with x still denoting $G/2\pi$, we modify the definition of $I(C)$ slightly and take

$$\frac{\tilde{I}(C)}{2} = \pi \int_X \left(x^2 - \frac{\lambda^2}{4} \right). \tag{3.6}$$

Since $x = \lambda/2 - v$ with integral v, it follows from (3.4) that $\tilde{I}(C)/2$ is well-defined modulo 2π and so can be used to define a line bundle $\mathcal{L}$ (just as we would have done with $I(C)/2$ had that been well-defined modulo 2π on the original class of G's).

As we have divided by two, $\mathcal{L}$ has the desired first Chern class, $c_1(\mathcal{L}) = \omega$. We do have to be a little careful in describing where $\mathcal{L}$ is defined. Because the quantization law of G has been shifted, $G = 0$ may not be allowed. To be more precise, (3.5) permits $G = 0$ precisely if λ is even. If λ is not even, the above construction gives a line bundle $\mathcal{L}$ not on the intermediate Jacobian J_W but on a shifted version of it with $G/2\pi$ congruent to $\lambda/2$ modulo integral classes.

In the application to M-theory, with W understood as the world-volume of a five-brane, one can assume (because of the world-volume equation $\mathrm{d}T = G$) that the restriction of G to W has vanishing cohomology class. (Also (3.5) means therefore that the restriction of λ to W must be even.) In this situation, it is the ordinary, unshifted intermediate Jacobian J_W on which the above construction gives a line bundle.

The main point is that as the correction to quantization of G needed to define $\tilde{I}(C)/2$ is the same as the one that appears in the physics, the line bundle appears precisely in the right place.

3.1. Embedding in 11 dimensions

So far, we have considered the chiral two-form on a "bare" six-dimensional spin manifold W.

We are really interested in an M-theory application in which W is a six-dimensional submanifold of an 11-manifold Q. Moreover, W, though oriented, is not necessarily spin; it is Q that carries a spin structure.

We will assume for simplicity that W is compact. More general cases, in which for instance W is asymptotically flat, are also natural; in such cases, some knowledge of what is happening at infinity can serve as a substitute for compactness.

The discussion in the last subsection was adequate if the normal bundle to W in Q is trivial – that is if Q looks locally near W like $W \times \mathbb{R}^5$. If so, the normal directions can be decoupled from the discussion and what we are about to say reduces to what was said above.

The first question we might want to ask is how to achieve gauge-invariance even locally for the three-form field C of M-theory. We recall that C is coupled to the chiral two-form β on the five-brane, and that this coupling is not invariant under gauge transformations $\delta C = -\mathrm{d}\alpha$ of the C-field. The failure of gauge-invariance was described in (2.13): [17]

$$\delta L_{\mathrm{eff}} \sim \int_Q \alpha \wedge G. \tag{3.7}$$

---

[17] Of course, (2.13) is the violation of gauge invariance for a non-chiral two-form with chiral coupling to C, while we want a chiral two-form. The difference between the two is crucial in discussing subtle global issues such as those considered in this paper, but not for studying local perturbative anomalies.

The numerical coefficient multiplying the right-hand side can be most usefully described as follows. Suppose that W is the boundary of a seven-manifold M. Let $I(C)$ be the seven-dimensional Chern–Simons functional defined in (3.2). It is gauge-invariant on a closed seven-manifold, but not on a seven-manifold with boundary. In fact,

$$\delta L_{\text{eff}} = \tfrac{1}{2}\delta I(C). \tag{3.8}$$

This factor of one-half is the reason that in the discussion above it was necessary to define a version of $I(C)/2$. Upon taking $W = \Sigma \times CP^2$, with Σ a Riemann surface much larger than the CP^2, the chiral two-form on W reduces to a chiral scalar on Σ, and the factor of $1/2$ in (3.8) reduces to the factor of one-half difference between (2.10) and (2.24) which was essentially the subject of section two.

We must find another interaction in the theory that can cancel the anomaly (3.8). In the first instance, this interaction is simply the classical interaction $I_0 = \text{const} \cdot \int_Q C \wedge G \wedge G$ of 11-dimensional supergravity. The normalization of I_0 can be most usefully described by noting that if Q is the boundary of a 12-manifold Y over which G extends, and $x = G/2\pi$, then

$$I_0 = -\frac{2\pi}{6} \int_Y x \wedge x \wedge x. \tag{3.9}$$

The factor of $-1/6$, which was important in [4], is related to E_8 index theory as explained there.

To prove that I_0 is invariant under $\delta C = -d\alpha$, one must integrate by parts and use $dG = 0$. But in the field of a five-brane, dG is not zero; it is a delta function supported on the five-brane world-volume W. So instead of being zero, the gauge variation of I_0 is a multiple of $\int_W \alpha \wedge G$, and can potentially cancel (3.7). That the cancelation actually occurs is a consequence of the formula $(1/2) + 3(-1/6) = 0$, where the $1/2$ is present in (3.8), the $-1/6$ in (3.9), and the 3 reflects the fact that in comparing the 11-dimensional interaction $I_0 \sim CGG$ to the seven-dimensional interaction $I_1 \sim CG$, one takes one of the three fields in I_0 to be in the normal direction to W and two to be tangential; there are three ways to do this.

What we have just analyzed is a piece of perturbative anomaly cancelation for five-branes, namely the term involving C only. One should also consider perturbative gravitational anomalies for five-branes. Some such terms were studied in [6,10]; a more complete discussion is in Section 5 . Because of invariance under sign change of C together with reversal of orientation of the normal bundle, there are no "mixed" C-gravitational perturbative anomalies.

3.2. Definition of the line bundle

We now have the crucial clue for how to find the desired line bundle $\mathcal{L} \to J_W$ and therefore the desired partition function. In the previous examples (chiral boson and chiral two-form on a "bare" six-manifold), the key was to find a Chern–Simons interaction (in a

higher dimension) that is gauge-invariant on a closed manifold but not in the presence of a boundary. We now do exactly the same thing, but the higher dimension will be 11 (and not seven, as the previous experience might suggest), and the word "boundary" must be replaced by "five-brane".

Thus, in M-theory, there is a Chern–Simons interaction, schematically $I_M = CGG + CI_8(R)$, where the CGG term is the interaction I_0 considered above, and $I_8(R)$ is a certain quartic polynomial in the Riemann tensor [5,6]. The expression

$$W = \det D_{RS}\ e^{iI_M} \tag{3.10}$$

with $\det D_{RS}$ the Rarita–Schwinger path integral, is well-defined [4] on a closed 11-manifold Q.[18] The two factors are not separately well-defined, but as explained in [4], one can alternatively factor (3.10) as follows:

$$W = \{\det D_{RS}\ e^{iI_{RS}/2}\} \cdot e^{iI_{E_8}}. \tag{3.11}$$

Here I_{RS} is a properly normalized Chern–Simons term related to the Rarita–Schwinger operator, and I_{E_8} is a properly normalized Chern–Simons term related to E_8 index theory. The virtue of the factorization in (3.11) is that the C-dependence has been put entirely in $J = e^{iI_{E_8}}$, which is a conventional Chern–Simons term, so our general framework will apply. The factor in curly brackets in (3.11) will play practically no role in the discussion.

On an 11-manifold Q that has a boundary R, J is not gauge-invariant in the usual sense, but rather must be interpreted as a section of a line bundle, which we will now call $\mathcal{L}^{-1}$, over the space of fields on R. This line bundle $\mathcal{L}^{-1}$ can be described somewhat more concretely by arguments given in Section 2. Those arguments, after all, had a purely formal character: given *any* Chern–Simons interaction – such as I_{E_8} – which is well-defined on a closed manifold but not on a manifold with boundary, one always produces a line bundle over a suitable space of fields on the boundary. Once $\mathcal{L}^{-1}$ is found, the chiral two-form partition function must be a section of the inverse line bundle $\mathcal{L}$ (to cancel global as well as local anomalies), and therefore its partition function is determined.

Let us now go over some of that ground in a little more fully. Five-branes may sound different from boundaries, but for the present purposes, the two are quite similar. If W is a five-brane world-volume in a space–time Q, then the G field has a singularity along W. Since singularities are awkward at best and the low energy field theory description is really not valid near the singularity, one might want to cut out of space–time a small tubular neighborhood of W. The boundary of that neighborhood is a 10-manifold R which is an S^4 bundle over W. The fact that W was a five-brane world-volume is now captured by saying that is S is a fiber of $R \to W$, then

$$\int\limits_R \frac{G}{2\pi} = 1. \tag{3.12}$$

---

[18] The bordism statement used here is of course Stong's theorem [23] that $\Omega_{11}^{spin}(K(\mathbb{Z}, 4)) = 0$.

A sensible definition of the chiral two-form partition function should depend only on the local geometry near W – which means, apart from a knowledge of W itself, only a knowledge of the normal bundle to W in Q and a choice of spin structure on a neighborhood of W in Q. All these informations can be summarized by giving the 10-manifold R, together with the map $R \to W$ obeying (3.12), and a spin structure on R. (As the normal bundle to R in Q is a trivial real line bundle, the spin structure on Q induces one on R.) To obtain a definition of the chiral two-form partition function that only depends on the geometry near Q, we should show that we can make the definition given only R.

The first step is to define a line bundle $\mathcal{L}^{-1}$ over the space of C-fields on R, with an action of the local and global gauge transformations. For this we need a well-defined Chern–Simons action on a closed 11-manifold; we choose for this the object I_{E_8}. With this Chern–Simons action, the construction in Section 2 now produces a line bundle $\mathcal{L}^{-1}$ over the space of C-fields on R. We recall that this is done as follows: given a loop L in the space of C fields on R, one builds a C_L field on $R \times \mathbf{S}^1$, and declares the holonomy of $\mathcal{L}^{-1}$ around L to be the value of $e^{iI_{E_8}}$ on $R \times \mathbf{S}^1$ with $C = C_L$, the antiperiodic spin structure on $\mathbf{S}^1$, and the product metric on $R \times \mathbf{S}^1$.[19]

To finish, then, all we need is a map i from C-fields on W to C-fields on R that obey (3.12) (commuting with local and global gauge transformations). Given such a map, we use i to "pull back" the line bundle $\mathcal{L}^{-1}$ from the space of C-fields on R to the space of C-fields on W, and this gives finally the gauge-invariant line bundle we need on the space of C-fields on W.

The desired i is found as follows. Let $\pi : R \to W$ be the projection, and let $\tau : R \to R$ be the map that commutes with π and acts as -1 on each $\mathbf{S}^4$ fiber of π. A solution C_0 of (3.12) can be described uniquely up to gauge transformation by saying that $G_0 = dC_0$ is harmonic (remember that C_0 can have Dirac string singularities, so that this is possible) and that the field is odd under τ.[20] The desired map from C-fields on W to C-fields on R is $C \to \pi^* C + C_0$.

Via i, we pull back $\mathcal{L}^{-1}$ to the space of C-fields on W. Restricting to C-fields with $G = 0$ and dividing by local and global gauge transformations, we get a line bundle $\mathcal{L}^{-1}$ over J_W. The chiral two-form partition function is determined by the fact that it is a section of the dual line bundle $\mathcal{L}$.

---

[19] An important detail must be checked here. The factorization in (3.11), by which we eliminated the determinant and reduced to Chern–Simons theory, is unique only up to $I_{E_8} \to I_{E_8} + I'$, where I' is a properly normalized Chern–Simons interaction constructed from the metric only. To really have a unique construction of $\mathcal{L}^{-1}$, the holonomy around the loop L should be independent of the choice of factorization. This is true since $R \times \mathbf{S}^1$ with product metric is the boundary of $R \times D$ with product metric (D being a two-disk), and the relevant curvature polynomials are all zero pointwise on $R \times D$ with product metric.

[20] For existence of such a field, take any solution C_0 of (3.12) with harmonic G_0, and replace it by $(C_0 - \tau^* C_0)/2$. For uniqueness, note that if C_0 and C_0' both obey the conditions, then $C' = C_0 - C_0'$ is a closed three-form odd under τ and therefore (since τ acts as $+1$ on $H^3(R, \mathbb{Z})$, which is isomorphic to $H^3(W, \mathbb{Z})$) vanishes up to gauge transformation.

4. Chiral four-form in 10 dimensions

Now we move on to the other somewhat similar example: a chiral four-form γ in 10 dimensions, which appears in Type IIB superstring theory.

The field strength of γ is a self-dual five-form L. It does not obey $dL = 0$, as one might have guessed. Rather, if B^i, $i = 1, 2$, are the two two-forms of Type IIB supergravity (which of course transform in the two-dimensional representation of $SO(2, \mathbb{Z})$), and $H^i = dB^i$, then the relation is

$$dL = \epsilon_{ij} H^i \wedge H^j. \tag{4.1}$$

This means that if we set $E = \epsilon_{ij} B^i \wedge H^j$, then E behaves as a sort of "composite five-form gauge field" that is coupled to γ, just as a $U(1)$ gauge field can be coupled to a chiral scalar in two dimensions, and the C-field of 11-dimensional supergravity is coupled to the chiral two-form on a five-brane world-volume.

From Section 2, we know that the partition function of the chiral four-form γ on a 10-dimensional spin manifold W will be a section of a line bundle $\mathcal{L}$ over the intermediate Jacobian J_W, and that finding the partition function is equivalent to finding a line bundle $\mathcal{L}$ whose first Chern class equals the polarization ω of J_W. We also know that we can always use Chern–Simons theory to find a line bundle of first Chern class 2ω, and that we can use Chern–Simons theory to find a line bundle of first Chern class ω provided that, for a closed 12-dimensional spin manifold X, the intersection form on $H^6(X, \mathbb{Z})$ is always even. This last statement is, happily, true (unlike its counterpart in eight dimensions, whose falsehood made the last section more complicated).

More generally, in fact, the intersection form on the middle dimensional cohomology of a closed spin manifold is always even in $8k + 4$ dimensions. A proof of this using the Adem relations for the Steenrod algebra has been described by J. Morgan.[21]

One point that should be made about the chiral four-form in 10 dimensions is that in using Chern–Simons theory to define the line bundles, the bordism group that one meets is $\Omega_{11}^{\text{spin}}(K(\mathbb{Z}, 6))$. (This is the bordism group of an 11-manifold endowed with a six-dimensional class; it enters in the same way that $\Omega_3^{\text{spin}}(K(\mathbb{Z}, 2)) = \Omega_3^{\text{spin}}(CP^\infty)$ entered in Section 2.) As far as I know, this group has not been computed, and might conceivably be non-zero. If so, then as described at the end of Section 2, a sort of exotic "theta angle" would appear in the theory, parametrized by the dual group $H = \text{Hom}(\Omega_{11}^{\text{spin}}(K(\mathbb{Z}, 6)), U(1))$. Such an extra parameter in Type IIB superstring theory seems unlikely, so one might conjecture that in fact $\Omega_{11}^{\text{spin}}(K(\mathbb{Z}, 6)) = 0$.

---

[21] The argument is as follows. In $8k + 4$ dimensions, one has a relation in the Steenrod algebra $Sq^{4k+2} = Sq^2 Sq^{4k} + Sq^1 Sq^{4k} Sq^1$. Let x be an element of $H^{4k+2}(X, \mathbb{Z}_2)$, with X an $8k+4$-dimensional spin manifold. Then modulo two, one has $x^2 = Sq^{4k+2}(x) = Sq^2(Sq^{4k}x) + Sq^1(Sq^{4k}Sq^1x)$. The right-hand side vanishes, since Sq^2 and Sq^1, as maps to the top dimension, are multiplication by w_2 and w_1, which vanish for a spin manifold.

4.1. An example

It may be helpful to present a concrete example, to show that if we are given a specific 10-manifold W (or similarly a specific six-manifold and normal bundle in the five-brane case) we actually can use these considerations to find the line bundle $\mathcal{L}$ and therefore the partition function of the chiral field.

The example I will consider is $W = \mathbf{S}^5 \times \mathbf{S}^5$. Call the two factors $\mathbf{S}_1^5$ and $\mathbf{S}_2^5$. The middle-dimensional cohomology is $H^5(W, \mathbb{Z}) = \mathbb{Z} \oplus \mathbb{Z}$, with the two $\mathbb{Z}$'s coming cohomology of the two $\mathbf{S}^5$'s. The intermediate Jacobian is $J_W = \mathbb{R}^2/(\mathbb{Z} \oplus \mathbb{Z})$. Let a and b be the lattice points $(1, 0)$ and $(0, 1)$ in $\mathbb{Z} \oplus \mathbb{Z}$. There are two distinguished straight lines C_a and C_b in J_W, induced by the straight lines from the origin to a or b. I claim that the line bundle $\mathcal{L}$ made from the Chern–Simons construction has holonomy 1 around C_a or C_b. According to (2.17), this determines the holonomy around all cycles and uniquely determines the line bundle $\mathcal{L}$. (For instance, the holonomy on a straight line from the origin to $(1, 1)$ is -1.)

The Chern–Simons construction involves the coupling of the chiral four-form to a background five-form E. (In Type IIB superstring theory, E is a composite field, found in the discussion of (4.1), but for the present purposes E might as well be elementary.) For instance, to compute the holonomy around C_b, we must consider the 11-manifold $M = \mathbf{S}_1^5 \times \mathbf{S}_2^5 \times \mathbf{S}^1$, with an E-field which is the pull back of an E-field on $\mathbf{S}_2^5 \times \mathbf{S}^1$, and such that if $K = \mathrm{d}E$ is the field strength, then $\int_{\mathbf{S}^5 \times \mathbf{S}^1} K = 1$. ($E$ may also be taken to vary linearly in the $\mathbf{S}^1$ direction, but we will not need this.) Now, we can regard $\mathbf{S}_1^5$ as the boundary of a ball B^6, over which the spin structure extends. So M is the boundary of $X = B \times \mathbf{S}_2^5 \times \mathbf{S}^1$. The E-field on X can be a pull back of an E-field on $\mathbf{S}_2^5 \times \mathbf{S}^1$. The holonomy around C_b is to be computed from $\int_X K \wedge K$. But this vanishes, since (E being a pull back from the last two factors), $K \wedge K$ actually vanishes pointwise.

So the chiral four-form partition function on $\mathbf{S}^5 \times \mathbf{S}^5$ is determined. As a mathematical corollary, note that it follows that, since $\mathbf{S}^5 \times \mathbf{S}^5$ has only one spin structure, which is therefore preserved by all orientation-preserving diffeomorphisms, the diffeomorphism group of $\mathbf{S}^5 \times \mathbf{S}^5$ does not induce the full $SL(2, \mathbb{Z})$ action on $H^5(\mathbf{S}^5 \times \mathbf{S}^5, \mathbb{Z})$, but at most the subgroup preserving the particular line bundle $\mathcal{L}$. (This subgroup can be shown to be of index three.) By contrast, in the superficially similar case of $\mathbf{S}^1 \times \mathbf{S}^1$, the diffeomorphism group does induce the full $SL(2, \mathbb{Z})$ action on $H^1(\mathbf{S}^1 \times \mathbf{S}^1, \mathbb{Z})$.

5. Perturbative anomaly cancelation for five-branes

In this section, we re-examine perturbative anomaly cancelation for M-theory five-branes. We have already verified anomaly cancelation for the C-field gauge invariance in Section 3. It remains to consider gravitational anomalies. It turns out that this leads to a surprisingly long story – nothing about five-branes seems to be straightforward! – and we will get a complete answer only for Type IIA, not for M-theory.

The five-brane world-volume is a six-manifold W in an 11-manifold Q. W is oriented, and (though this requirement can be relaxed, as M-theory conserves parity) we will consider

only the case that Q is oriented. Let TQ be the tangent bundle to Q and $TQ|_W$ its restriction to W. We have $TQ|_W = TW \oplus N$, where TW is the tangent bundle to W and N is the normal bundle to W in Q. N, in particular, is an $SO(5)$ bundle over W.

Note that a Riemannian metric on Q induces a Riemannian metric on W, and a metric and $SO(5)$ connection on N. The theory along W therefore has some features of gravity coupled to $SO(5)$ gauge theory.

In analyzing anomalies, it is enough to consider only diffeomorphisms of Q that map W to W, since the presence of the five-brane wrapped over W explicitly breaks other diffeomorphisms. A diffeomorphism of Q that maps W to W induces first of all a diffeomorphism of W and secondly an $SO(5)$ gauge transformation of the normal bundle. In fact, diffeomorphisms of W and $SO(5)$ gauge transformations are all that the the world-volume fields (and anomalous interactions) "see" (at least in the long-wavelength limit that suffices for computing anomalies) so the discussion will amount to an analysis of gravitational and $SO(5)$ gauge anomalies on W.

Perturbative anomalies in $2n$ dimensions are always related to characteristic classes in $2n + 2$ dimensions, so in the present case – as W is six-dimensional – the anomalies will involve eight-dimensional characteristic classes. It turns out to be rather helpful to write the anomalies in terms of Pontryagin classes $p_i(TW)$ and $p_i(N)$. The anomaly eight-form is then a priori a linear combination of $p_2(TW)$, $p_1(TW)^2$, $p_1(N)p_1(TW)$, $p_1(N)^2$, and $p_2(N)$. The terms involving $p_i(TW)$ only have been analyzed before [6,10]; we will extend the analysis to include the other terms. As we will see, the discussion is surprisingly unstraightforward. The known and expected contributions to the $p_1(N)p_1(TW)$ and $p_1(N)^2$ anomalies will cancel, but something new is involved in the cancellation of the $p_2(N)$ term.

There are three known sources of anomalies: (1) world-volume fermions; (2) the chiral two-form; (3) the Chern–Simons couplings of the bulk theory, whose gauge invariance fails in the presence of the five-brane. Their contributions can be determined as follows.

5.1. World-volume fermions

The world-volume fermions are (four-component) chiral spinors on W with values in the (rank four) bundle $S(N)$ constructed from N by using the spinor representation of $SO(5)$. According to standard anomaly formulas [24], the contribution of these fields to the anomaly is

$$I_D = \tfrac{1}{2}\mathrm{ch}S(N) \cdot \hat{A}(TW), \tag{5.1}$$

where the $1/2$ arises because of considering chiral spinors, ch is the Chern character, and up through dimension eight

$$\hat{A}(TW) = 1 - \frac{p_1(TW)}{24} + \frac{7p_1(TW)^2 - 4p_2(TW)}{5760}. \tag{5.2}$$

To compute $\mathrm{ch}S(N)$, we note that if the Chern roots of the $SO(5)$ bundle N are $\pm\lambda_1$, $\pm\lambda_2$, and 0, then the Chern roots of $S(N)$ are $\pm(\lambda_1 \pm \lambda_2)/2$. So up through terms quartic in the λ's,

$$\mathrm{ch}S(N) = \sum_{\epsilon_1,\epsilon_2=\pm 1} \exp[\tfrac{1}{2}(\epsilon_1\lambda_1 + \epsilon_2\lambda_2)]$$
$$= 4 + \tfrac{1}{2}(\lambda_1^2 + \lambda_2^2) + \tfrac{1}{96}(\lambda_1^4 + \lambda_2^4 + 6\lambda_1^2\lambda_2^2)$$
$$= 4 + \tfrac{1}{2}p_1(N) + \tfrac{1}{96}p_1(N)^2 + \tfrac{1}{24}p_2(N). \tag{5.3}$$

In the last step we used $p_1(N) = \lambda_1^2 + \lambda_2^2$, $p_2(N) = \lambda_1^2\lambda_2^2$. The detailed form of I_D can be obtained by combining the last three equations.

5.2. The chiral two-form

The chiral two-form propagates on W and does not "see" the normal bundle. The standard anomaly of such a field is

$$I_A = \tfrac{1}{5760}(16p_1(TW)^2 - 112p_2(TW)). \tag{5.4}$$

5.3. Anomaly inflow from the bulk

The third term to consider is the anomaly inflow from the bulk. The 11-dimensional bulk theory has an interaction proportional to $\tilde{I} = C \wedge (p_1(TQ)^2/4 - p_2(TQ))$, where $p_1(TQ)$ and $p_2(TQ)$ can be understood as certain polynomials in the Riemann tensor of Q. It is important that what appears in $\tilde{I}$ is the Riemann tensor (and therefore the Pontryagin classes) of Q, not those of W; after all, $\tilde{I}$ is an interaction defined in the 11-dimensional bulk theory. The relevance to perturbative anomalies of $\tilde{I}$ is that, although gauge-invariant in bulk, it is not gauge-invariant in the field of a five-brane. The anomaly I_B coming from this term is

$$I_B = -\tfrac{1}{48}(\tfrac{1}{4}p_1(TQ|_w)^2 - p_2(TQ|_w)). \tag{5.5}$$

As $TQ|_w = TW \oplus N$, we have $p_1(TQ|_w) = p_1(TW) + p_1(N)$, $p_2(TQ|_w) = p_2(TW) + p_2(N) + p_1(N)p_1(TW)$. (5.5) can thereby be rewritten

$$I_B = -\tfrac{1}{48}(\tfrac{1}{4}(p_1(TW)^2 + p_1(N)^2 - 2p_1(TW)p_1(N)) - p_2(TW) - p_2(N)). \tag{5.6}$$

Upon summing up (5.1), (1.2), and (5.6), one finds that all terms involving $p_i(TW)$ cancel (both the purely gravitational term and the "mixed" term $p_1(TW)p_1(N)$), as does the $p_1(N)^2$ term. The remaining anomaly is in fact

$$\delta = \tfrac{1}{24}p_2(N). \tag{5.7}$$

So something new is needed. It turns out that it is easier to understand the new ingredient if one considers the problem in Type IIA superstring theory rather than in M-theory. In other words, we take Q to be $M \times \mathbf{S}^1$, where M is a 10-manifold, and we consider the case that the five-brane world-volume W is a submanifold of M (times a point in $\mathbf{S}^1$). M-theory then becomes equivalent to Type IIA superstring theory, and the five-brane becomes the

solitonic five-brane of Type IIA, which couples magnetically to the string theory two-form B (of field strength $H = \mathrm{d}B$).

The normal bundle N then becomes $N = N' \oplus \mathcal{O}$ where $\mathcal{O}$ is a trivial one-dimensional bundle (representing the tangent bundle to $\mathbf{S}^1$) and N' (the tangent bundle to W in M) is an $SO(4)$ bundle. In particular, $p_i(N) = p_i(N')$ for $i = 1, 2$.

One thing that is special about $SO(4)$, however, is that for an $SO(4)$ bundle – such as N' – p_2 can be written in terms of a four-dimensional class, called the Euler class of the bundle, $\chi(N')$. This has its roots in the fact that at the Lie algebra level, $SO(4) = SU(2) \times SU(2)$, so that an $SO(4)$ bundle has two four-dimensional characteristic classes – p_1 and χ – related to the instanton numbers in the two $SU(2)$'s. These are the independent characteristic classes of an $SO(4)$ bundle (in general a Lie group of rank r has r independent such characteristic classes). So in particular p_2 can be written in terms of these. The relation is in fact $p_2(N') = \chi(N')^2$. p_2 can be derived from the function $\lambda_1^2 \lambda_2^2$ of the Chern roots and χ can be derived from the function $\lambda_1 \lambda_2$. An expression $p_2 = \chi^2$ can exist in $SO(4)$ but not in $SO(5)$ because the function $\lambda_1 \lambda_2$ is Weyl-invariant in $SO(4)$ – whose Weyl group acts on the λ_i with pairwise sign changes – but not in $SO(5)$, whose Weyl group generates independent sign changes of the λ_i.

Some notation concerning the Euler class will be helpful. We represent an object α transforming in the adjoint representation of $SO(4)$ by a 4×4 antisymmetric tensor α^{ij}, $i, j = 1, \ldots, 4$. In particular, the curvature F of an $SO(4)$ connection A is a two-form F^{ij} with values in that representation. The Euler class is represented by the four-form $\chi(F) = \epsilon_{ijkl} F^{ij} \wedge F^{kl} / 32\pi^2$, where ϵ_{ijkl} is the fourth rank invariant antisymmetric tensor of $SO(4)$. Locally, given a choice of gauge, $\chi(F)$ is the exterior derivative of a Chern–Simons three-form that we call $\Omega_\chi(A)$: $\chi(F) = \mathrm{d}\Omega_\chi(A)$. If α is an $SO(4)$ gauge generator, we write $\chi(\alpha, F) = \epsilon_{ijkl} \alpha^{ij} F^{kl} / 16\pi^2$. The gauge variation $\delta_\alpha \Omega_\chi$ of Ω_χ under a gauge transformation by α is

$$\delta_\alpha \Omega_\chi(A) = d\left(\chi(\alpha, F)\right). \tag{5.8}$$

This equation is part of the "descent" formalism familiar in the study of anomalies.

5.4. The Euler class

The anomaly of interest is thus – in the Type IIA context – $\chi(N')^2/24$, and it will be necessary to have some understanding of the particular meaning of the characteristic class $\chi(N')$.

The basic question we have to focus on is: what is the H-field (that is, the three-form field strength of Type IIA) produced by a five-brane world-volume W in a 10-manifold M? The following considerations only involve the behavior near W, and so only depend on the topology of the normal bundle N' to W in M. We can in fact replace M with the total space of N'.

The H-field is supposed to be a three-form such that $\mathrm{d}H = \delta_W$, where δ_W is a delta function supported on W. Such an H is not uniquely determined, as one could add any

smooth, closed H-field. However, there is an obstruction to existence of H: the obstruction is that $\chi(N')$ must vanish. This is explained in [25], beginning on p. 70.

The essence of the problem is to precisely formulate what δ_W is supposed to be. δ_W should be a four-form on M which is closed, is supported in a small neighborhood of W in M, and (if one identifies M with the total space of N') integrates to 1 over any fiber of $N' \to W$.

If B is a differential form on a manifold M, the restriction of B to a submanifold W, written $B|_W$, is obtained by considering only the values on W of components of B tangent to W. Thus $B|_W$ is a differential form on W (of the same degree as B). The theory of the Thom isomorphism and the Euler class, as described in [25], shows that any δ_W with the properties stated in the last paragraph has the further property that $\delta_W|_W$ is in the cohomology class of $\chi(N')$, the Euler class of the normal bundle N'. If a connection A on N' is picked, so that $\chi(N')$ is represented by a differential form $\chi(F)$ (as defined above), then one can in a very natural way pick δ_W such that

$$\delta_W|_W = \chi(F). \tag{5.9}$$

Now it is clear why there is a restriction on the possible topology of the normal bundle to a five-brane. The "magnetic" field of the five-brane is supposed to be a three-form H obeying

$$dH = \delta_W, \tag{5.10}$$

but this implies

$$d(H|_W) = \delta_W|_W = \chi(F). \tag{5.11}$$

Such an H can exist only if the differential form $\chi(F)$ is trivial cohomologically, that is only if the characteristic class $\chi(N')$ vanishes.

A form with the properties of δ_W could be constructed using *any* connection A on N', but when W is embedded in M as a five-brane world-volume, there is a distinguished connection – coming from the Riemannian connection of M – and therefore a distinguished four-form $\chi(F)$ representing the Euler class of the normal bundle. Eq. (5.10) is part of the naive idea of what a five-brane wrapped on W is supposed to be, but is not usually stated precisely enough to exhibit the "finite part" of the delta function along W that appears in (5.11). We will assume in the rest of this paper that (5.11) should be taken as part of the definition of a five-brane.

Since the anomaly we are trying to eliminate is proportional to $\chi(N')^2$, the vanishing of $\chi(N')$ means that the integrated anomaly vanishes in a suitable sense. That is not enough; we need to cancel the anomaly locally, since gauge transformations are local. But vanishing of the integrated anomaly at least means that there is no topological obstruction to finding a counterterm that would cancel the anomaly. In fact it is easy to see, using (5.11), that there is such a counterterm. It is

$$\tilde{L} = \int\limits_W H|_W \wedge \Omega_\chi(A), \tag{5.12}$$

where again A is the Riemannian connection on N' and Ω_χ was introduced above. Indeed, using (5.8) and (5.11) and integrating by parts, the variation of ΔL under gauge transformations of N' is

$$\delta_\alpha \tilde{L} = - \int_W \chi(F) \wedge \chi(\alpha, F), \tag{5.13}$$

and this is the six-form related by the usual "descent" procedure to the anomaly eight-form $\chi(F)^2/2$.

So a multiple of $\tilde{L}$ will cancel the $\chi(N')^2$ perturbative anomaly. Moreover, given the invariance of Type IIA superstrings under reversal of orientation of N' together with sign change of H, this is the only anomaly that can be canceled by such a term. It is thus gratifying that precisely this term is the one whose contributions from previously known interactions do not cancel.

5.5. Back to M-theory

For M-theory five-branes, we need a generalization of this, but how to do so is somewhat puzzling. The replacement of $SO(4)$ by $SO(5)$ and of $\chi(N')^2$ by $p_2(N)$ makes even the absence of a topological obstruction to canceling the anomaly mysterious. For a specific five-brane world-volume W, $p_2(N)$ vanishes, as W is six-dimensional. But the topological interpretation of anomalies involves considering certain two-parameter families of physical objects [26], and one could perfectly well have a two-parameter family of W's with non-zero $p_2(N)$ over the total space. It is not clear what sort of physical mechanism would suppress such families in M-theory.

Beyond canceling the topological obstruction, we need to actually find a local counterterm that cancels the $p_2(N)$ anomaly. The fact that this counterterm must reduce to (5.13) under the appropriate conditions makes it clear roughly what the desired counterterm must be, though the matter still seems rather obscure. For an $SO(5)$ bundle N, the closest analog of the characteristic class χ is a certain four-form with values in N, defined by

$$\chi_e = \frac{F^{ab} \wedge F^{cd} \epsilon_{abcde}}{32\pi^2}. \tag{5.14}$$

The factorization $p_2 = \chi^2$ that holds for $SO(4)$ bundles becomes for $SO(5)$ bundles

$$p_2(N) = \sum_e \chi_e \wedge \chi_e. \tag{5.15}$$

In M-theory, the fundamental differential form in space–time is not a three-form H but a four-form G. However, along W, one can define an N-valued three-form, namely the part of G with three indices tangent to W and one N-valued index. Let us write this part of G as H_e, where e is the N-valued index, and the indices tangent to W are not written explicitly. The analog of (5.12) must be something like

$$\int_W \sum_e H_e \wedge \Omega_e, \tag{5.16}$$

where Ω_e is an N-valued three-form related to χ_e. H_e must also obey an N-valued version of (5.11). It is not clear exactly what the right equations are. The fact that the story works so nicely for Type IIA nevertheless gives some faith that a satisfactory answer must exist in M-theory.

References

[1] E. Cremmer, B. Julia and J. Scherk, Supergravity theory in eleven-dimensions, Phys. Lett. B 76 (1978) 409.
[2] E. Bergshoeff, E. Sezgin and P.K. Townsend, Supermembranes and eleven-dimensional supergravity, Phys. Lett. B 189 (1987) 75.
[3] R. Guven, Black p-brane solutions of $D = 11$ supergravity, Phys. Lett. B 276 (1992) 49.
[4] E. Witten, Flux quantization in M-theory, hepth/9609122.
[5] C. Vafa and E. Witten, A one-loop test of string duality, Nuclear Phys. B 447 (1995) 261.
[6] M.J. Duff, J.-T. Liu and R. Minasian, Eleven-dimensional origin of string/string duality: A one loop test, hepth/9509084.
[7] P. Horava and E. Witten, Heterotic and type I string dynamics from eleven dimensions, hepth/9510209; Eleven-dimensional supergravity on a manifold with boundary, hepth/9603142.
[8] S. De Alwis, Anomaly cancellation in M-theory, hepth/9609211.
[9] E. Witten, Global gravitational anomalies, Commu. Math. Phys. 100 (1985) 197; Global anomalies in string theory, in: Anomalies, Geometry, and Topology, eds. W.A. Bardeen and A. White (World Scientific, Singapore, 1985).
[10] E. Witten, Five-branes and M-theory on an orbifold, hepth/9512219.
[11] P. Townsend, D-branes from M-branes, Phys. Lett. B 373 (1996) 68.
[12] A. Strominger, Open P-branes, Phys. Lett. B 383 (1996) 44.
[13] C.G. Callan, Jr., J.A. Harvey and A. Strominger, Worldbrane action for string solitons, Nucl. Phys. B 367 (1991) 60.
[14] P. Goddard and D. Olive, Algebras, lattices, and strings, in: Vertex Operators in Mathematics and Physics, Proc. of a Conf., 10–17 November, 1983, eds. J. Lepowsky, S. Mandelstam and I.M. Singer (Springer, New York, 1983) p. 51.
[15] E. Witten, Non-perturbative superpotentials in string theory, hepth/9604030
[16] L. Alvarez-Gaumé, J.-B. Bost, G. Moore, P. Nelson and C. Vafa, Bosonization on higher genus Riemann surfaces, Comm. Math. Phys. 112 (1987) 503.
[17] E. Witten, On holomorphic factorization of the WZW model, Comm. Math. Phys. 144 (1992) 189.
[18] R. Dijkgraaf and E. Witten, Topological gauge theories and group cohomology, Comm. Math. Phys. 129 (1990) 393.
[19] R.E. Stong, Notes on Cobordism Theory (Princeton University Press, Princeton, NJ, 1968).
[20] T.R. Ramadas, I.M. Singer and J. Weitsman, Some comments on Chern–Simons theory, Comm. Math. Phys. 126 (1989) 409.
[21] S. Axelrod, Geometric quantization of Chern–Simons theory, Ph.D. Thesis, Princeton University (1991).
[22] D. Freed, Classical Chern–Simons theory: Part 1, Adv. Math. 113 (1995) 237.
[23] R. Stong, Calculation of $\Omega_{11}^{\text{spin}}(K(\mathbb{Z}.4))$, in: Unified String Theories, eds. M.B. Green and D.J. Gross (World Scientific, Singapore, 1986).
[24] L. Alvarez-Gaumé and E. Witten, Gravitational anomalies, Nucl. Phys. B 234 (1983) 269.
[25] R. Bott and L.W. Tu, Differential Forms in Algebraic Topology (Springer, Berlin, 1982).
[26] M.F. Atiyah and I.M. Singer, Dirac operators coupled to vector potentials, Proc. Nat. Acad. Sci. US 81 (1984) 2597.