J. DIFFERENTIAL GEOMETRY
35 (1992) 219–254

UNCOUNTABLY MANY EXOTIC $\mathbf{R}^4$'S IN STANDARD 4-SPACE

STEFANO DEMICHELIS & MICHAEL H. FREEDMAN

Abstract

It is known that the standard (Euclidean) smooth structure on 4-space when restricted to certain open subsets homeomorphic to $\mathbf{R}^4$ gives a smooth structure which is not diffeomorphic to the standard one. This behavior is a consequence of Donaldson's counterexample [5] to the smooth 5-dimensional h-cobordism theorem and was noticed (in anticipation of Donaldson's result) by A. Casson and the second named author (see [14, Theorem 3, Chapter 14]). Taubes [24] developed a technically demanding theory of the Yang-Mills equation on "asymptotically end periodic" 4-manifolds in part to verify that a known family of exotic $\mathbf{R}^4$'s were mutually distinct. That family lays smoothly in $S^2 \times S^2$ but not $\mathbf{R}^4$. We combine ideas from the above-mentioned papers to address a nested family of $\mathbf{R}^4$ homeomorphs called "ribbon $\mathbf{R}^4$'s" lying in $\mathbf{R}^4$ standard. There are continuum many pairwise distinct smooth structures represented within this family.

0. Introduction

Our philosophy is that any Donaldson-style invariant [5] can be defined on an "end periodic" manifold and these invariants commute with the passage between a compact manifold and such noncompact geometric limits. In principle the Γ-invariant or "polynomial-invariant" is suitable for this discussion; however, we carry out the analysis in detail only for D. Kotschick's "simpler" Φ-invariant [16]. Kotschick distinguishes a certain algebraic surface, the Barlow surface B, from the rational surface $Q = CP^2\#8\overline{CP}^2$ by showing that $|\Phi(B)| \geq 4$ and $\Phi(Q) = 0$. Taubes paper [24] on the self-dual Yang-Mills equation on end periodic 4-manifolds provides much of the technical foundation for our extension.

It is known that B and Q are smoothly h-cobordant (and therefore homeomorphic); that is, there exists $(W^5; B, Q)$ with $\partial W^5 = B \amalg -Q$, and the inclusions $B \hookrightarrow W^5$, $Q \hookrightarrow W^5$ are homotopy equivalences. It is, by now, a standard idea that W^5 should be analyzed with a mind toward

Received June 11, 1990 and, in revised form, March 20, 1991. The authors were supported in part by National Science Foundation Grant DMS-89-01412.

pushing the proof of the smooth h-cobordism theorem as far as possible. When this is done a "partial" diffeomorphism is found (cf. [13, Theorem 7.1c]). More precisely, an exotic $\mathbf{R}^4$, $\mathbf{R}^4_1$, is found smoothly imbedded in B and in Q, and a compactum $K \subset \mathbf{R}^4_1$ is produced so that $B \backslash \mathrm{inc.}\,1\,K$ and $Q \backslash \mathrm{inc.}\,2\,K$ are diffeomorphic. Starting at $\mathbf{R}^4_1 \backslash K$ and moving around the diagram below in a circle give a smooth involution

$$\iota\colon (\mathbf{R}^4_1 \backslash K) \hookleftarrow\colon$$

$$\begin{tikzcd} & B \hookleftarrow B \backslash \mathrm{inc.}\,1\,K \arrow[dd, "\wr\ \mathrm{diff}", equal] \\ \mathbf{R}^4_1 \backslash K \subset \mathbf{R}^4_1 \arrow[ur, hook, "\mathrm{inc.}\,1"'] \arrow[dr, hook, "\mathrm{inc.}\,2"] & \\ & Q \hookleftarrow Q \backslash \mathrm{inc.}\,2\,K \end{tikzcd}$$

Using a topological radius function ρ on $\mathbf{R}^4_1$, which is ι-invariant on $\mathbf{R}^4_1 \backslash K$, we may construct a continuous nested family $\mathbf{R}^4_t$, $0 \le t \le 1$, of sub-$\mathbf{R}^4_1$'s with $K \subset \mathbf{R}^4_0$ and $\mathbf{R}^4_s \subset \mathbf{R}^4_t$ for $s \le t$.

Suppose two pairs $(\mathbf{R}^4_s, K)$, $(\mathbf{R}^4_t, K)$, $s < t$, are diffeomorphic relative to the identity on K, $(d, \mathrm{id}_K)\colon (\mathbf{R}^4_t, K) \to (\mathbf{R}^4_s, K)$. Then $\mathbf{R}^4_t \backslash \bigcap_{i=1}^{\infty} d^i(\mathbf{R}^4_t)$ has a "covering" action by the monoid $\{d, d^2, d^3, \cdots\}$; the quotient Y is a smooth manifold. Giving Y some Riemannian metric, it is easy to alter the Riemannian metric on B and Q so that n consecutive rings $\mathbf{R}^4_t \backslash d^n(\mathbf{R}^4_t)$ cover Y isometrically. Call the results B_n and Q_n. In the limits as $n \searrow \infty$ we reach noncompact end periodic manifolds B_∞ and Q_∞ which are simply Riemannian structures on $B \backslash \mathrm{inc.}\,1\,(\bigcap_{i=1}^{\infty} d^i(\mathbf{R}^4_t))$ and $Q \backslash \iota \circ \mathrm{inc.}\,2\,(\bigcap_{i=1}^{\infty} d^i(\mathbf{R}^4_t))$. From our identifications and the inclusions, $B_\infty \subset B \backslash \mathrm{inc.}\,1\,K \overset{\mathrm{diff}}{\cong} Q \backslash \iota \circ \mathrm{inc.}\,2\,K \supset Q_\infty$; the restriction gives a diffeomorphism between B_∞ and Q_∞. We have the contradiction

$$\Phi(B) = \Phi(B_n) = \Phi(B_\infty) = \Phi(Q_\infty) = \Phi(Q_n) = \Phi(Q),$$

the key step being that $\Phi(B_\infty)$ is defined and equal to $\Phi(B_n)$ (and the corresponding assertions when B is replaced by Q).

This contradiction shows that all pairs $(\mathbf{R}^4_t, K)$, $0 \le t \le 1$, are smoothly distinct. The compactum K may be taken to be a smooth codimension zero submanifold. For such K, it is easily proved that there are only countably many embeddings up to isotopy in any $\mathbf{R}^4_t$, so the set of parameter values $[0, 1]$ is partitioned into equivalence classes, each of at most countable cardinality, according to the diffeomorphism type of the total space $\mathbf{R}^4_t$, $t \in [0, 1]$. In Zermelo-Frankel set theory with choice

(ZFC) it is easily argued that {distinct diffeomorphism types among $\mathbf{R}_t^4$, $t \in [0, 1]$} has the cardinality of the continuum.

The family $\{\mathbf{R}_t^4, t \in [0, 1]\}$ has a *nearly* explicit description (see Theorem 3.2) when t lies in the standard Cantor set $\mathrm{CS} \subset [0, 1]$. $\mathbf{R}_t^4$ is diffeomorphic to $\mathbf{R}_{\mathrm{std}}^4 \backslash X_t$, where the closed subset X_t is a "Cantor set of wild arcs" whose geometry depends on the parameter value t. These multiplicities could, in principle, be worked out by studying an explicit h-cobordism W and going through the proof (see [11] or [13]) of the reembedding theorem to isolate an upper bound on the increase in capped-group complexity required to achieve reembedding. This interesting project has not been carried out.

The paper is organized as follows. §1 gives a brief review of the Φ-invariant in Kotschick's compact setting, and the body of the argument that Φ commutes with geometric limits is presented in §2. In §3 the smooth structure of simply connected 5-dimensional h-cobordisms is examined with an eye toward isolating the difference between the ends. It is found that this difference is an involution defined near the end of an exotic $\mathbf{R}^4$; regluing by this involution permutes the smooth structures represented by the ends of the n-cobordism. In §4 the outline contained in the introduction is filled out to provide a complete statement and proof of the main theorem.

Appendices A and B supply two important analytical details needed to complete the arguments of §2. We expect these results to be useful elsewhere.

The first author would like to express his deep gratitude to the second author for unfailing advice and generous support during his stay at the University of California at San Diego.

1. A brief review of the Φ-invariant

The canonical bundle over CP^2 is a $\mathrm{U}(1) = \mathrm{SO}(2)$ bundle with Euler class χ the positive generator of $H_2(CP^2, Z)$ and $p_1 = \chi^2 = 1$. Forming connected sums there is an $\mathrm{SO}(2)$ bundle $E' \to Q$ with $\chi(E') = (1, \cdots, 1) \in H^2(Q, Z)$ and $p_1 = -7$. According to the classification of SO(3) bundles [4] there is an SO(3)-bundle E with "least negative charge" $w_2(E) = w_2(E' \oplus \varepsilon) = (1, \cdots, 1) \in H^2(Q, Z_2)$ and $p_1(E) = -7+4 = -3$. The bundle E cannot have its structure group reduced to SO(2) because any potential Euler class reduces to $w_2 = (1, \cdots, 1)$ and so must satisfy $\chi^2 \equiv -7 \pmod 8$, while $\chi^2 = p_1 = -3$. Thus E has no reducible connections. According to the deformation theory for anti-self-dual (ASD) connections on E, the virtual dimension of the moduli space is

$$\dim \mathscr{M} = -p_1(\operatorname{ad} \mathfrak{g} \otimes \mathbf{C}) - \dim \mathfrak{g}(1 + b_2^+ - b_1).$$

In the case of SO(3) bundles, $\operatorname{ad} \mathfrak{g} \cong E$ and, as a real bundle, $E \otimes \mathbf{C} \cong E \oplus E$ so

$$\dim \mathscr{M} = -2p_1 - 3(2) = 0.$$

Since there are no reducible connections, $\mathscr{M}$ will be a manifold for a generic metric on Q [10, Proposition 3.20], in this case a collection of points.

According to Uhlenbeck's fundamental compactness results [27], a sequence $[A_i]$ of gauge equivalence classes of ASD connections will always have a subsequence which converges—but possibly on a less twisted bundle. (The bundle might lose "charge" near isolated points.) For SO(3) bundles the "charge" lost at a point comes in positive multiples of -4. Since it is not possible to have an ASD connection on a bundle with $p_1 > 0$ and $p_1(E) = -3$, $\mathscr{M}$ is compact.

Kotschick [16] shows that $\mathscr{M}(Q) = \varnothing$ by degenerating the metric along an appropriate connected sum 3-sphere: an ASD connection over Q would limit to an ASD connection over a summand with a moduli space of negative formal dimension.

On the other hand, the entire bundle discussion goes through for any smooth manifold such as B which is homotopy equivalent to Q. The corresponding moduli space over the Barlow surface is defined and a computation in algebraic geometry yields exactly for ASC connections over B each with a possible (positive) multiplicity.[1]

The invariant Φ is the number of points in a generic $\mathscr{M}$ counted according to sign. To determine the sign [7] it is necessary to choose a Spin-c structure and an orientation on $H_+^2(\ ; R)$. Thus Φ is integer-valued (which are homotopy equivalent to Q) together with the above data. In any case, the data only affect the sign of Φ, so B and Q are not diffeomorphic; $|\Phi(Q)| = 0$ and $|\Phi(B)| \geq 4$.

2. Φ commutes with geometric limit

Let $M \simeq CP^2 \# 8\overline{CP}^2$ be homotopy equivalent to Q. Let $\mathbf{R}_1^4$ be an open subset of M homeomorphic to $\mathbf{R}^4$. Suppose there is an open submanifold $\mathbf{R}_2^4$ contained in a compact submanifold of $\mathbf{R}_1^4$, $\mathbf{R}_2^4 \subset C \subset \mathbf{R}_1^4$,

[1] After a more detailed analysis Kotschick has proved that the multiplicity is 2.

and a diffeomorphism $d: \mathbf{R}_1^4 \to \mathbf{R}_2^4$. Then define $\mathbf{R}_k^4 = d^{k-1}\mathbf{R}_1^4$. The quotient $Y = (\mathbf{R}_1^4 \backslash \bigcap_{i=1}^{\infty} \mathbf{R}_i^4)/\text{pt.} \equiv d^i(\text{pt.})$ has a natural smooth structure, which *can be* described as $C\backslash d(C)/\text{pt.} \equiv d(\text{pt.})$. Give Y a fixed Riemannian metric. Let M_n, $n = 1, 2, 3, \cdots$, be a sequence of Riemannian metrics on M so that:

(1) in the induced metric, the projection $\mathbf{R}_1^4 \backslash \mathbf{R}_{n+1}^4 \to Y$ is a local isometry and

(2) the identity $M_n \xrightarrow{\text{id}} M_m$ is an isometry in the complement of $\mathbf{R}^4_{\min(n+1, m+1)}$.

Let M_∞ be the unique metric on $M \backslash \bigcap_{i=1}^{\infty} \mathbf{R}_n^4$ so that the inclusions $M_\infty \hookrightarrow M_n$ are isometries on the complement of $\mathbf{R}_{n+1}^4$.

We investigate the behavior of Φ under the limit $M_n \to M_\infty$. We find $\Phi(M_\infty)$ is well defined and equal to $\Phi(M_n)$ for each n.

To do elliptic analysis on M_n, Taubes [24] suggests introducing a proper function $\tau: M_\infty \to \mathbf{R}^+$ such that $\tau(d(x)) = 1 + \tau(x)$ for all $x \in \mathbf{R}_1^4$. A section s of any Riemannian or Hermitian bundle with connection ∇ over M can be measured according to the norms

$$\|s\|_{L^p_{k,\delta}} = \left[\int_{M_\infty} e^{\tau\delta} \sum_{k=0}^{j} |\nabla^{(k)} s|^p \, d\,\text{vol}\right]^{1/p},$$

where δ is a positive real number. The weighted Sobolev spaces $L^p_{k,\delta}$ are the completions of $C_0^\infty(M_\infty)$ with respect to these norms. The fundamental theorem (Proposition 4.2) of [24] guarantees that the appropriate ASD Yang-Mills deformation complex

$$0 \to \Omega^0(\mathfrak{g}) \xrightarrow{d} \Omega^1(\mathfrak{g}) \xrightarrow{d_+} \Omega^2_+(\mathfrak{g}) \to 0$$

becomes Fredholm when completed with respect to the norm $\| \; \|_{L^p_{k,\delta}}$ unless δ belongs to a discrete set of exceptional values $D(p, k) \subset \mathbf{R}^+$. This permits the extension of many standard arguments to the end periodic case.

Over M_∞ the bundle which we are studying should clearly have $w_2 = (1, \cdots, 1)$. The condition $p_1 = -3$ can be interpreted by considering only those connections $A_\infty = d + a$, $a \in L^2_{1,\delta}(\Omega^1)$, where d is a trivialization near infinity with respect to which $p_1 = -3$, and δ is a sufficiently small positive number. In principle $\Phi(M_\infty)$ will be the number of ASD connections in the moduli space $\mathscr{M}_\infty$ consisting of equivalence classes of connections A_∞ as above. However, just as in the compact case, to count properly a perturbation of the metric on M_∞ is required and then generic solutions are counted according to sign; the perturbation can decay arbitrarily fast toward infinity.

Theorem 2.1. *Suppose* $M = M_0, M_1, M_2, \cdots, M_\infty$ *is a sequence of Riemannian* 4*-manifold as above. Define* $\Phi(M_\infty)$ *to be the number (counted according to sign) of equivalence classes of connections* A_∞ *as above for a generic perturbation of the metric on* M_∞. *Then* $\Phi(M) = \Phi(M_n)$, $0 \le n \le \infty$.

Proof. The proof is an immediate consequence of the following four points:

Point 1. Given a sequence of ASD connections A_n on Q_n, there is a subsequence A_{n_i} such that A_{n_i} converges (in the C^∞ topology) on compact subsets.

Point 2. Given any A_∞ on M_∞, an $\alpha > 0$, and a compact subset $J \subset M_\infty$, there is an integer $n(\alpha)$ such that for $n > n(\alpha)$ there exist an ASD A_n on M_n which is within α in the C^0 norm on J.

Point 3. Suppose the metric on M_∞ is perturbed so that H^2 (deformation complex) $\cong 0$. Let A_n and A_n' both converge in C^∞ on compact sets to A_∞. Then for n sufficiently large $A_n = A_n'$ on Q_n.

Point 4. The moduli space of connections on M_∞ has an orientation defined up to an overall sign. The latter can be chosen so that it is compatible, in the limit, with the canonical orientations on the moduli spaces of $\mathbf{Q}_n$ and $\mathbf{B}_n$ given in Donaldson [7].

Proof of Point 1. Exhaust M_∞ by an expanding union of compact sets $J_1 \subset J_2 \subset J_3 \subset \cdots$. If i is large enough so that $w_2|_{J_i} \not\equiv 0$, then Sedlacek's thesis [21] shows that $A_n|_{J_i}$ cannot have a subsequence converging to a trivial connection. Since $\pi_1(J_i) \to \pi_1(J_j)$ is zero for $j \gg i$, $A_n|_{J_i}$ cannot have a subsequence convering to any flat connection. Also, no subsequence can concentrate curvature near a point. If this happened cutting and gluing would result in a bundle over M_n, n large, with positive p_1 and a connection with arbitrarily small self-dual curvature contradicting

$$4\pi^2 p_1 = \int_{M_n} \|F_+\|^2 \, d\,\mathrm{vol} - \int_{M_n} \|F_-\|^2 \, d\,\mathrm{vol}.$$

According to Uhlenbeck's compactness results ([27], [28]) the only remaining alternative is that A_n have a subsequence converging uniformly to a nontrivial connection on J_i. By taking further subsequences for J_{i+1}, J_{i+2}, etc. and diagonalizing we find a subsequence converging on compact subsets to a connection on M_∞. By Fatou's lemma it is clear that the energy of the limit $\int_{M_\infty} |F_{A_{\text{limit}}}|^2 \le 12\pi^2$. Because the end is simply connected at infinity no fractional charge can be lost by limiting to a nontrivial flat connection, so equality actually holds.

In Appendix B it is proved that the energy along the tube must decay exponentially (this also uses the uniform simple connectivity at infinity) so that $\int_{M_\infty} |F_{A_{\text{limit}}}|^2 e^{\tau\delta} < \infty$ for sufficiently small $\delta > 0$. By the gauge fixing arguments of [27] and the bounded geometry of M_∞ it then follows that A_{limit} may be gauged to the desired form, $A_\infty = d + a$, $a \in L^2_{1,\delta}(\Omega^1)$. Modulo Appendix B, this explains Point 1.

Proof of Point 2. The proof uses ideas and techniques which are well known in the field. The reader can find similar treatments in [8], [25], and [19].

Let $A_\infty = d + a$, $a \in L^2_{1,\delta}(\Omega^1)$, be an ASD connection on M_∞ as before. We first define an approximately ASD A'_n on M_∞ for n sufficiently large.

Let $K(\nu)$, $\nu > 0$, be a compact subset of M_∞ containing all but a (topological) product neighborhood of $\operatorname{end}(M_\infty)$ and so large that $(12\pi^2)^{-1} \int \|F_{A_\infty}\|^2 \ge 1 - \nu$. By enlarging $K(\nu)$ we may also assume $\|a\|_{L^2_{1,\delta}}$ is arbitrarily small on $M_\infty \backslash K$. For n sufficiently large, we may define

$$A'_n = \begin{cases} A_\infty & \text{on } K(\nu), \\ d + \beta a & \text{on } M_n \backslash K(\nu), \end{cases}$$

where $\beta \colon Q_\infty \backslash K(\nu) \to [0, 1]$ is 1 near $K(\nu)$ and 0 on $\mathbf{R}^4_n$. Except on $\operatorname{supp}(\beta(1-\beta))$ the connection A'_n is anti-self-dual, and $\|F^+_{A'_n}\|_{C^k}$ is bounded by $C(\beta)\|a\|_{C^{k+1}}$.

Given A'_n our goal is to use Taubes' iteration (see [17, §§2 and 3, Chapter 6] for example) in order to reach an exact solution. The results of Taubes [23] tell us that $\forall \in \exists \nu$ such that $\|F^+_{A'_n}\| < \nu$, and $H^2(\operatorname{ad}\mathfrak{g}; A'_n) \cong 0$ implies the existence of an ASD A_n with $|A_n - A'_n| < \varepsilon$. Given $\varepsilon > 0$, ν depends only on the local geometry near $\operatorname{supp}(\beta(1-\beta))$ and the size of the first eigenvalue $\lambda_{1,n}$ of the composition

$$\underbrace{\Omega^2_+(\operatorname{ad}\mathfrak{g}; A'_n) \xrightarrow{\delta} \Omega^1(\operatorname{ad}\mathfrak{g}, A'_n) \xrightarrow{d^+} \Omega^2_+(\operatorname{ad}\mathfrak{g}, A'_n)}_{\Delta_n}, \tag{2.1}$$

where the first map is an adjoint of d^+. A definition of the adjoint metrics is required. To obtain uniformity of the $\lambda_{1,n}$ and to achieve the $\| \ \|_{L^2_{,\delta}}$ metric on the limit M_∞ we follow Mrowka's thesis [19] and define "tent"

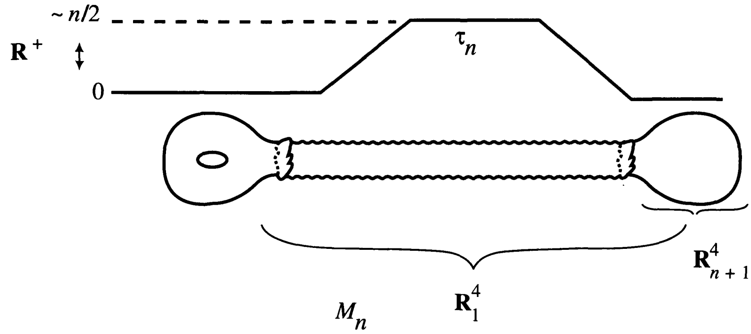


FIGURE 2.1

functions $\tau_n\colon M_n \to \mathbf{R}^+ \cup 0$ roughly as in Figure 2.1.

Norms $\| \ \|_{L^2_{,\delta}} = \int_{M_n} e^{+\tau_n\delta} \| \ \|_{L^2} d\mathrm{vol}$ give Hilbert space completions (which we continue to denote $\Omega^i(M)$) and the adjoint map δ above. The exact form of τ_n will be specified later.

Integration by parts allows us to express δ in terms of the unweighted adjoint $(d^+_{A'_n})^*$ as

$$\delta = e^{-r\delta}(d^+_{A'_n})^* e^{\tau\delta}. \tag{2.2}$$

The composition $\Delta_n = d^+\delta$ is a positive elliptic operator on a compact space. Thus $\mathrm{Spec}(\Delta_n)$ is a discrete subset of $[0, +\infty]$. Let $\lambda_{1,n}$ be the smallest eigenvalue.

Lemma 2.1. *For n large, $\lambda_{1,n}$ is uniformly bounded away from zero.*

Proof. If not, there is a sequence $\phi_n \in \Omega^2_+(\mathrm{ad}\,\mathfrak{g})_n$ such that

(1) $\|\phi_n\|_{L^2_{,\delta}} = 1$ and

(2) $\|\Delta_n\phi_n\|_{L^2_{,\delta}} \searrow 0$.

Now fix two (large) integers l and $\bar{l}$, and define $\beta_{l,\bar{l}}$ according to Figure 2.2.

We next prove that $\|(1-\beta_{l,\bar{l}})\phi_n\|_{L^2} \searrow 0$. Since $\mathrm{supp}(1-\beta_{l,\bar{l}})\phi_n$ is uniformly bounded, the exponential weight does not matter. In fact, the support of $(1-\beta_{l,\bar{l}})\phi_n$ is concentrated in $K_{l,\bar{l}}$ and $K'_{l,\bar{l}}$. The restrictions of Δ_n and $\Delta_\infty = d^+_{A_\infty} e^{-\tau\delta}(d^+_{A_\infty})^* e^{\tau\delta}$ to $K_{l,\bar{l}}$ agree provided n is large enough so that the cutoff (β) used to define A'_n occurs on $\beta^{-1}_{l,\bar{l}}(1)$.

Moreover, for a generic metric (see [24, §6]) Δ_∞ has lowest eigenvalue $\lambda_{1,\infty}$ bounded away from zero:

$$\|\Delta_\infty\phi\|_{L^2_{,\delta}} \geq (\lambda_{1,\infty})\|\phi\|_{L^2_{,\delta}}, \qquad \phi \text{ supported in } K_{l,\bar{l}}. \tag{2.3}$$

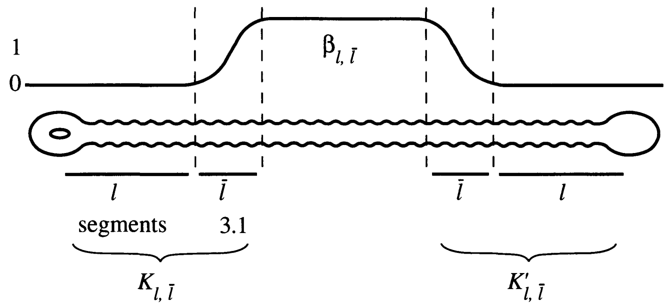


FIGURE 2.2

On $K'_{l,\bar{l}}$, Δ_n is the same as $\Delta = d^+_\Gamma e^{-\tau\delta}(d^+_\Gamma)^* e^{\tau\delta}$, where Γ is the trivial connection. By [24, Proposition 5.1], the kernel of Δ is empty and Δ is Fredholm. This implies that (2.3) holds for ϕ supported in $K'_{l,\bar{l}}$ (but with a different constant λ_1).

We conclude that

(2.4) $|(1-\beta)\phi|_{L^2_{,\delta}} \le (\Delta_n(1-\beta)\phi)/\lambda$ for some $\lambda > 0$ independent of n.

But

$$\Delta_n(1-\beta_{l,\bar{l}})\phi = (1-\beta_{l,\bar{l}})\Delta_n\phi + T, \tag{2.5}$$

where T stands for terms involving first and second derivatives of β, the value of ϕ and first derivatives of ϕ in the part in which $d\beta \neq 0$. The bounds on $d\phi$ follow from the Gårding inequality (see [26, Theorem 36.1 and Remark 36.1]) and the derivative of β can be controlled by choosing $\bar{l}$ large. The result is

$$\|(1-\beta_{l,\bar{l}})\phi\|_{L^2_{,\delta}} \le C\|(1-\beta_{l,\bar{l}})\Delta_n\phi_n\|_{L^2_{,\delta}}. \tag{2.6}$$

Now we need to estimate $\|\beta_{l,\bar{l}}\phi_n\|_{L^2_{,\delta}}$. This is essentially Lemma 6.7 of Mrowka's thesis [19]. We sketch his proof and indicate where we depart from it. The sequence $\beta_{l,\bar{l}}\phi_n$ can be thought of as defined on the universal cover $\widetilde{Y}$ of Y. Recall from Figure 2.1 the form of the tent function τ_n over this infinite tube $\widetilde{Y}$ (Figure 2.3, next page).

The operator Δ_n is unchanged by replacing τ_n by $\tau^0_n = \tau_n - n/2$. The τ^0_n converge to τ^0 on $\widetilde{Y}$ and Δ_n converges in the norm topology for operators on $\widetilde{Y}$ to $\Delta = d^+_\Gamma e^{-\tau^0\delta}(d^+_\Gamma)^* e^{\tau^0\delta}$. We must prove that, for a compactly supported section ϕ,

$$\|\Delta\phi\|_{L^2_{,\delta}} \ge C\|\phi\|_{L^2_{,\delta}} \quad \text{for some } C \ge 0. \tag{2.7}$$

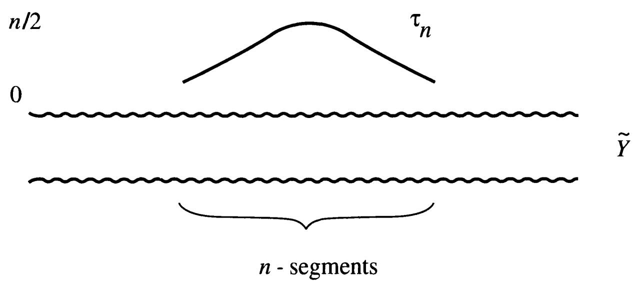


FIGURE 2.3

This would imply $|\beta_{l,\bar{l}}\phi_n| \searrow 0$ and combined with (2.6) would finish the proof of Lemma 2.1.

Suppose (2.7) does not hold; then $0 \in \operatorname{Spec}(\Delta)$. But we assumed $\delta \in \mathbf{R}^+ \backslash D$ so Δ is Fredholm [23] and 0 is actually on eigenvalue of Δ, i.e., there is an $L^2{}_{,\delta}$ section ϕ of norm one with $\Delta(\phi) = 0$. Taking the inner product with ϕ and using the definition of adjoint, we find $(d^+)^* e^{\tau^0 \delta}\phi = 0$, and since $*\phi = \phi$ we have

$$d\psi = d^*\psi = 0, \quad \text{where } \psi = e^{\tau_0 \delta}\phi. \tag{2.8}$$

In the case of a product end Mrowka proves $\psi \equiv 0$ by explicitly solving the equations. In our case we reason as follows.

The section ψ belongs to L^2 and since $\tau_0 \le 0$, $e^{\tau^0 \delta} \le 1$, we have

$$\int e^{-\tau^0 \delta} |\psi|^2 \, dv = \int e^{+\tau^0 \delta} |\phi|^2 \, dv < \infty \quad \text{by definition of } \psi. \tag{2.9}$$

Since ψ is closed we have $\psi = d\alpha$ with α a 1-form. Moreover,

$$\int_{\tau^0(x) \le \tau^0 \le \tau^0(x)+1} |\alpha(x)|^2 \le ct^k \int_{|\tau^0| < \tau^0(x)+1} |\psi|^2 \, d\mu.$$

The actual value of k does not matter and can be computed by the reader. The proof follows from the sheaf-theoretic proof of DeRham's theorem, which we sketch:

First there is a Leray covering $\{V_i\}$ with V_i convex and 1-forms α_i so that $d\alpha_i = \psi|_{V_i}$. An application of the Poincaré lemma shows that $\sup_{V_i} |\alpha_i| < \text{const} \sup_{V_i} |\psi|$. If $V_i \cap V_j \neq \varnothing$ we can write $\alpha_i - \alpha_j = df_i - df_j$ because $[\psi] = 0$ in H^2. We can choose the f_i so that they obey estimates in terms of $\int_{\tau < \tau(x)} |\psi|$. Writing $\tilde{\alpha}_i = \alpha_i - df_i$ gives α satisfying the estimate above. Since $|\psi|$ is in L^2, $|\alpha|$ is bounded by a polynomial in τ^0. Since τ^0 is negative, $e^{\tau^0 \delta}\alpha \in L^2$ $\forall \delta > 0$.

So

$$\int |\psi|^2 = \int d\alpha \wedge d\alpha = \int d(\alpha \wedge d\alpha) = \int d(e^{\tau\delta'} \alpha \wedge e^{-\tau\delta'} d\alpha) = 0.$$

If $0 < \delta' < \delta/2$, the last integral is equal to zero because $e^{-\tau\delta'} d\alpha = e^{-\tau\delta'} \psi \in L^2$ since

$$\int e^{\tau\delta} |\psi|^2 < +\infty.$$

Thus $\psi \equiv 0$, completing the proof of Lemma 2.1. q.e.d.

A uniform lower bound λ for $\lambda_{1,n}$ allows a Taubes' iteration to be applied to convert A'_n to an ASD A_n. As in the compact case threshold conditions (depending on λ) on A'_n must be satisfied. These are achieved by making the original cutoff β sufficient far out along the end where the a is small. This is possible since a is small in the $C_n^\infty \cap L^2_{,\delta}$ norm, implying that it is small in any norm needed in Taubes' iteration.

Proof of Point 3. Suppose that A_n and A'_n are sequences of connections on M_n which converge on compact sets to a connection A_∞ on M_∞ of the type described above. That is, $A_\infty = d + a$ near infinity, where d is standard differentiation with respect to a trivialization at infinity fixing $p_1 = -3$ and $a \in L^2_{1,\delta}$. By elliptic regularity compact convergence in L^2 implies compact convergence in C^∞, so the hypothesis may be taken in either sense.

If we write $A'_n = A_n + a_n$, then $a_n \searrow 0$ in the C^∞ norm on compact sets. Again we will consider a large compact (smooth) submanifold K_ε which contains all of M_∞, except a topological product neighborhood of infinity, and supports most of the energy of the connection:

$$\frac{1}{12\pi^2} \int_{K_\varepsilon} |F_{A_\infty}|^2 > 1 - \varepsilon \quad \text{for } \varepsilon > 0. \tag{2.10}$$

Since convergence $A_n \searrow A_\infty$ and $A'_n \searrow A_\infty$ is C^1 on K_ε for n sufficiently large, (2.10) will hold for A_n and A'_n with 2ε replacing ε on the right-hand side.

Uhlenbeck's [28] basic method for constructing radial gauges in the presence of small curvature may be applied inductively in patches along the periodic tube in M_n. Because the geometry is bounded there are functions $C(\varepsilon)$, $C'(\varepsilon)$, $C''(\varepsilon)$ which go to zero with ε so that, after a gauge transformation,

$$\int_{M_n \setminus K_\varepsilon} |a_n|^2 dv < C(\varepsilon) \tag{2.11}$$

and then by elliptic regularity

$$\sup_{M_n\setminus K_\varepsilon} |a_n| < C'(\varepsilon). \tag{2.12}$$

By Appendix B and using $\tau_n < \tau$ we get

$$\int_{Q_n} e^{\tau_n\delta}|a_n|^2 < C''(\varepsilon). \tag{2.13}$$

Adapting Uhlenbeck's radial gauges to periodic ends is the goal of Lemma 10.4 [23]. In Appendix A we give a more detailed proof of this lemma, in which the role of simple connectivity at infinity is clarified.

We wish to prove that $a_n = 0$ for n large. Our first step is to use the continuity method to improve our radial gauge to a Hodge gauge, i.e., to achieve

$$d^*_{A_n} e^{\tau_n\delta} a_n = 0 \tag{2.14}$$

and still preserve

$$|a_n|^2_{C^0} + \int e^{\tau_n\delta}|a_n|^2 \le \varepsilon. \tag{2.15}$$

Dropping the " n " from our notation, we consider the problem of finding g_t so that a_t defined by (2.16) will also satisfy (2.14) and (2.15):

$$a_t = tg_tag_t^{-1} + g_tdg_t^{-1}. \tag{2.16}$$

When $t = 0$, $g_0 = 0$ is a solution; we want a solution for $t = 1$.

First we check that the set t for which (2.14)–(2.16) are soluble is closed. Since g takes values in a compact Lie group (SO(3)), a sequence of solutions certainly will have a subsequence which converges pointwise on a dense collection of points. However, to obtain a reasonable limit some equicontinuity of the g's is necessary. This can be deduced from the equation $gA' - tAg = dg$. At this point elliptic regularity takes over and implies that a limiting g is well behaved.

On the other hand, if t has a solution, then an implicit function theorem [20] allows (2.14) and (2.16) to be solved for $t' = t + \nu$, for ν sufficiently small, but in (2.15) the right-hand side must be increased to, say, 2ε.

Some work is now required to lower 2ε to ε. We claim that $a_{t'}$ satisfies

$$d^+_{A_n} a'_t = (-a_{t'} \wedge a_{t'} + t(t-1)ga \wedge ag^{-1})^+. \tag{2.17}$$

To verify this we expand:

$$
\begin{aligned}
d_{A_n} a_{t'} + a_{t'} \wedge a_{t'} &= d_{A_n}(tgag^{-1} + gdg^{-1}) \\
&\quad + (tgag^{-1} + gdg^{-1}) \wedge (tgag^{-1} + gdg^{-1}) \\
&= td_{A_n} gag^{-1} + tgd_{A_n} ag^{-1} + tgad_{A_n} g^{-1} \\
&\quad + d_{A_n} g \wedge d_{A_n} g^{-1} + gd_{A_n} g^{-1} \wedge d_{A_n} g^{-1} \\
&\quad + t^2 ga \wedge ag^{-1} + tga \wedge d_{A_n} g^{-1} \\
&\quad - td_{A_n} g \wedge ag^{-1} + d_{A_n} g \wedge g^{-1} d_{A_n} gg^{-1}.
\end{aligned}
\tag{2.18}
$$

Since $A_n' = A_n + a$ is ASD, $d_{A_n} a^+ = (-a \wedge a)^+$ so upon taking $+$-self-dual parts terms 2 and 5 above survive as desired; the others drop out. Terms 1 and 8 cancel as do terms 3 and 7, and 4 and 9. Because A_n is ASD the $+$ part of term 5 vanishes.

We can write $a_{t'} = e^{-\tau\delta} d_{A_n}^* e^{\tau\delta} \omega = \tilde{d}^* \omega$ for some ω because $d_{A_n}^* e^{\tau\delta} a_{t'} = 0$ and the appropriate H^1 vanishes; so $\ker d^* e^{\tau\delta} = \operatorname{Im} e^{-\tau\delta} d_{A_n}^* e^{\tau\delta}$. Writing $d_{A_n}^+ = \tilde{d}^+$ we have

$$
\tilde{d}^+ \tilde{d}^* \omega = (-d^* \omega \wedge d^* \omega + t(t-1) ga \wedge ag^{-1})^+.
$$

Now if we integrate over M_n with a weight, we have

$$
\int (e^{\tau\delta} \omega, \tilde{d}^+ \tilde{d}^* \omega) = \int e^{\tau\delta} (\omega, d^* \omega \wedge d\omega) + t(t-1) \int e^{\tau\delta} (\omega, (ga \wedge ag^{-1}))^+.
$$

Integrating by parts and using the Schwarts inequality we get

$$
\begin{aligned}
\int e^{\tau\delta} |d^* \omega|^2 &\le \left(\int e^{\tau\delta} |\omega|^2 \right)^{\frac{1}{2}} \left(\int e^{\tau\delta} (d^* \omega \wedge d^* \omega)^2 \right)^{\frac{1}{2}} \\
&\quad + t(t-1) \left(\int e^{\tau\delta} |\omega|^2 \right)^{\frac{1}{2}} t(t-1) \left(\int e^{\tau\delta} |(ga \wedge ag^{-1})|^+ \right)^{\frac{1}{2}}.
\end{aligned}
\tag{2.19}
$$

But $\int e^{\tau\delta} |d^* \omega|^2 \ge \lambda_1 \int e^{\tau\delta} |\omega|^2$, where λ_1 is independent of n.

Dividing by $\int e^{\tau\delta} |\omega|^2$ yields

$$
|\omega|_{L^2_{,\delta}} < C(|d^* \omega \wedge d^* \omega|_{L^2_{,\delta}} + |a \wedge a|^+_{L^2_{,\delta}})
\tag{2.20}
$$

so $|\omega|_{L^2_{,\delta}} \le C(4\varepsilon^2) + \varepsilon^2 < C'\varepsilon^2$. Elliptic regularity now gives $|\omega|_{L^p_{k,\delta}} \le C''\varepsilon^2$. The desired bound on sup norm follows for small ε. Thus (2.14) and (2.15) are simultaneously achieved.

Let D_n denote the linear operator

$$D_n = d^+_{A_n} \oplus e^{-\tau\delta} d^*_{A_n} e^{\tau\delta} : \Omega^1_\delta \to \Omega^{2+}_\delta \oplus \Omega^0_\delta,$$

and let $H^1(\operatorname{ad}\mathfrak{g}; A_\infty) \cong 0$ by the index theorem and the choice of generic metric. As in the proof of Point 2,

$$|D_n a_n|_{L^2_{,\delta}} \geq \lambda |a_n|_{L^2_{,\delta}} \quad \text{for some } \lambda > 0. \tag{2.21}$$

The constant λ does not depend on n if n is large enough. But $\|a_n\|_{C_0} < \varepsilon$ so

$$\|a_n \wedge a_n\|^+_{L^2_{,\delta}} < \varepsilon \|a_n\|_{L^2_{,\delta}}. \tag{2.22}$$

If $\varepsilon < \lambda$, then (2.21) and (2.22) are contradictory. This completes the proof of Point 3.

Proof of Point 4. The proof of Point 2 implies that if

$$D_{A_l} = (e^{-\tau\delta} d^*_{A_l} e^{\tau\delta} ; d^+_{A_l}) : \Omega^1(\operatorname{ad} A_l) \to \Omega^0(\operatorname{ad} A_l) \oplus \Omega^2_+(\operatorname{ad} A_l),$$

then we have

$$|D_{A_l}\alpha|_{L^2_{,\delta}} > \lambda |\alpha|_{L^2_{,\delta}}, \qquad \lambda > 0,$$

provided that A is anti-self-dual and $l \gg 0$ or $l = \infty$.

Consider the spaces:

$\mathscr{B}_{l,n} = \{A | A$ connection on M_l with exponential decay and such that $\sup|A|$ on $M_l \backslash K(n) < \varepsilon\lambda\} / \{$Gauge transformations constant on $M_\infty \backslash K(n)\}$.

Here $l = 1, 2, \ldots$ or ∞ and $n < l$.

Remark that $B_{l,n}$ exhaust B, and also

$$\mathscr{B}_{c,n} = \frac{\{A | A \text{ connection on } M_\infty \equiv \text{trivial connection on } M_\infty \backslash K(n)\}}{\{\text{Gauge transformations constant on } M_\infty \backslash K(n)\}}.$$

A suitable cutoff gives maps $r_l : B_{l,n} \to B_{c,n}$. Moreover,

$$\sup|r_l A - A| \leq \varepsilon\lambda. \tag{$*$}$$

On $\mathscr{B}_{c,n}$ we have the bundles $\Omega^1(\operatorname{ad} A)$ and $\Omega^0(\operatorname{ad} A) \oplus \Omega^2_+(\operatorname{ad} A)$, and a real elliptic operator D_A between them. As in [7, Chapter 3] this defines a determinant line bundle $\Lambda_{c,n}$.

Similarly we define determinant bundles $\Lambda_{l,n}$ on $\mathscr{B}_{l,n}$, compatible with the inclusions $\mathscr{B}_{l,n} \hookrightarrow \mathscr{B}_{l,n+1}$.

Lemma 2.2. *$r_l^* \Lambda_{c,n}$ is isomorphic to $\Lambda_{l,n}$ for n, l, and $l - n$ large.*

Proof. There are maps from (co)ker D_A to (co)ker D_{A_c} and vice versa, defined by suitable cutoff and ortogonal projections.

If n, l, and $l-n$ are large, then all the cutoff functions β have small $d\beta$; this and condition $(*)$ imply that all these maps are isomorphisms.

As in [5] this gives a canonical isomorphisms $\Lambda_{c,n'} \simeq \Lambda_{c,n} \simeq \Lambda_{l,n}$ for n, n' large but much smaller than l. Assuming n meets these conditions, we drop it from the notation.

Lemma 2.3. *Λ_c is trivial.*

Proof. We will prove that $\mathscr{B}_c$ is simply connected.

First, $\mathscr{B}_c$ is homotopy equivalent to $\mathscr{B}$, the space of connections on $E \to M$ modulo the pointed gauge group $\mathscr{G}_0$. Since SO(3) is centerless, $\mathscr{B} \simeq B(\mathscr{G}_0)$.

Thus it is sufficient to check that $\mathscr{G}_0$ is connected. The pointed gauge group $\mathscr{G}_0$ may be described as the space of base point preserving sections of the adjoint bundle $\Gamma(\operatorname{Ad}\mathrm{SO}(3))$. The adjoint representation lifts to $\mathrm{SO}(3) \to \mathrm{Aut}(\mathrm{SU}(2))$ by $x \mapsto (g \to \pi^{-1}(x)g(\pi^{-1}(x))^{-1})$, with $\pi\colon \mathrm{SU}(2) \to \mathrm{SO}(3)$ the two-fold cover; then $\operatorname{Ad}\mathrm{Su}(2) \to \operatorname{Ad}\mathrm{SO}(3)$ is also a two-fold cover. Since the base M is simply connected, any section of $\operatorname{Ad}\mathrm{SO}(3)$ lifts to a section of $\operatorname{Ad}\mathrm{SU}(2)$. It is sufficient to check that $\Gamma(\operatorname{Ad}\mathrm{Su}(2)) = \widetilde{\mathscr{G}}_0$ is connected since an arc in $\widetilde{\mathscr{G}}_0$ may be projected to $\mathscr{G}_0$. Since SU(2) is locally connected, any section can be normalized to be the identity on a disk D containing the base point of M. But over the complement $\mathscr{M}\backslash D$ the bundle is trivial (since any map of a 3-complex in BSU(2) is null homotopic) and the sections trivial at ∂D. The space of such sections $\widetilde{\mathscr{G}}_0$ is clearly equal to maps (M, S^3). By [10, Proposition (5.12)] $\pi_0(\mathscr{G}_0) = \pi_0$ maps $(M, S^3) \cong 0$ so $\pi_0(\widetilde{\mathscr{G}}_0) \cong 0$, $\pi_0(\mathscr{G}_0) \cong 0$, and $\pi_1(\mathscr{B}) \cong 0$. q.e.d.

Now $\det A$ or $\det A_l$ are real sections of Λ_c and Λ_l nonvanishing on a neighborhood of self-dual connections in the B_l's and of their images in B_c.

Lemmas 2.2 and 2.3 give compatible identifications of the Λ's with R. Since the $r_l A_l$ converge to $r_l A_\infty$ and the sign of $\det A$ is locally constant, this gives compatible orientations for the points in the moduli space, up to an overall sign.

Thus

$$|\Phi(B)| = |\Phi(B_n)| = |\Phi(B_\infty)|, \qquad |\Phi(Q)| = |\Phi(Q_n)| = |\Phi(Q_\infty)|. \tag{2.23}$$

According to Donaldson [7], if the data described in §1 is fixed, then there is a canonical overall choice of orientation on the moduli spaces permitting the absolute value signs to be removed in (2.23).

3. Exotic $\mathbf{R}^4$'s and the smooth structure of 5-dimensional h-cobordisms

Let $(W; B, Q)$ be a compact simply connected smooth 5-dimensional h-cobordism between closed 4-manifolds. The ends are labeled to suggest the "Barlow" surface and the corresponding rational surface but this is not assumed. Another interesting example is an h-cobordism from the Dolgachev surface to a rational surface. Also, the assumption that the ends be closed is not important provided W already has a product structure over $\partial(B)$.

It is known that W has a topological product stucture. Although W, in general, cannot have a smooth product structure—this exists precisely when there is a diffeomorphism $b \to Q$ in the homotopy class determined by W—Smale's high dimensional proof of the h-cobordism theorem may be adapted to give a "partial product structure" on W which identifies B and Q except over an exotic $\mathbf{R}^4$ imbedded in both. The reader may wish to compare out analysis of W with the proof of Theorem 7.1c in [13] and Kirby's sketch of Theorem 3 in [14, Chapter 14]. All objects and morphisms will be smooth (C^∞) unless otherwise specified, e.g. $\cong$ means diffeomorphic.

Theorem 3.1. *W contains a compact 5-dimensional submanifold with boundary $(\overline{J}; J_0, J_1) \subset (W; B, Q)$ and a noncompact 5-dimensional proper h-cobordism $(U; V_0, V_1) \subset (W; B, Q)$ so that the following hold:*

(1) $J_0 \times I \cong \overline{J} \cong J_1 \times I$; *in particular,* $J_0 \cong J_1$.

(2) $\overline{J} \cup U = W$.

(3) V_0 *and* V_1 *are homeomorphic but not diffeomorphic to* $\mathbf{R}^4$ *(and therefore by the topological theory* U *is homeomorphic to* $\mathbf{R}^4 \times I$).

(4) $V_0 \cong V_1$.

(5) *Combining* (1), (2), *and* (4) *gives an reimbedding* $\operatorname{end} V_0 \hookrightarrow \operatorname{end} V_0$. *This can be chosen to be an involution.*

(6) V_0 *(and hence* V_1*) inherit their smooth structure as open subsets of* $\mathbf{R}^4_{\mathrm{std}}$.

Proof. Give W a handle body structure relative to B. Handles of index 0 and 5 are cancelled outright. In the appendix of [18] Milnor explains how handles of index 1 may be cancelled at the expense of introducing new 3-handles; similarly handles of index 4 can be removed, but this introduces new 2-handles. This leaves only 2- and 3-handles. They generate relative cellular chain groups. Since $H_*(W, B; Z) \cong 0$, the boundary map is an isomorphism $C_3 \xrightarrow{\partial} C_2$. The Euclidean algorithm is used to construct handle slides (among the 3-handles) so that the 3- and 2-handles pair algebraically:

(3.1) $\partial h_i^3 = \delta_{ij} h_j^2$

(given some indexing i, $j = 1, \cdots, n$ of the handles).

We interpret this formula in the middle level M of W, $M = \partial((B \times I) \cup \{h_j^2\}) \setminus B$. In M we see two collections of disjointly imbedded 2-spheres: $\{A_j\}$ = co-cores or "ascending" spheres of the 2-handles $\{h_j^2\}$ and $\{D_i\}$ = boundary of "descending" sphere of the 3-handles $\{h_j^3\}$. We may assume all intersections are transverse; equality (3.1) becomes

$$D_i \cdot A_j = \sum_{D_i \cap A_j} (\text{sign}) = \delta_{ij}. \tag{3.2}$$

A central observation of Casson's theory (see [2] or [14, Chapter 12]) is that there is an isotopy (called finger moves) of $\coprod_i D_i$ so that, after the isotopy,

$$\pi_1 \left(M \Big\backslash \left(\bigcup_{i=1}^{n} D_i \cup \bigcup_{j=1}^{n} A_j \right) \right) \cong \{e\}. \tag{3.3}$$

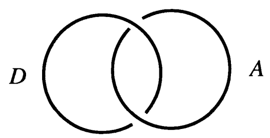
This enabled Casson to find Casson handles CH in M attaching to the boundary of a regular neighborhood $\mathcal{N} = \mathcal{N}(\bigcup_{i=1}^n D_i \cup \bigcup_{j=1}^n A_j)$ so that $Y^- = (\mathcal{N} \cup \text{CS's})^-$ is homeomorphic[2] to $(\#_{i=1}^n (S^2 \times S^2) \setminus \text{pt.})$. Casson handles are homeomorphic to *open* 2-handles $(D^2 \times \mathbf{R}^2, \partial D^2 \times \mathbf{R}^2)$. After attaching them to $\mathcal{N}$ we delete any remaining boundary, this is the meaning of the $^-$-superscript.

With an eye to achieving condition (5) we should do this part of the construction with some care so that Y admits a smooth involution exchanging $\bigcup_{i=1}^n D_i$ and $\bigcup_{j=1}^n A_j$. For this we use a handlebody description of Y (perhaps [11] and [14] are the best general references for drawing these pictures). Although these pictures are part of a rigorous theory, we will proceed by example. The geometrically inclined reader should enjoy arguing about the accuracy of our diagrams. All unlabeled 2-handle circles are "0-framed;" all 1-handles are indicated by a circle bearing a dot.

If $n = 1$ and $D \cap A$ = one point, then no Casson handles need be attached and the picture is as depicted in Figure 3.1 (next page).

If $n = 1$ and $D \cap A = 3$ points, then we attach two Casson handles as in Figure 3.2.

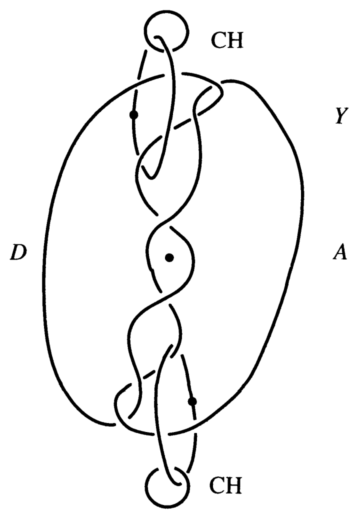
[2]In 1975 Casson would have said "proper homotopy equivalent;" a homeomorphism was not available for several more years.



$$Y \overset{\text{diff.}}{\cong} \overline{S^2 \times S^2 \setminus B^4}$$

$$Y^- \overset{\text{diff.}}{\cong} S^2 \times S^2 \setminus \text{pt.}$$

FIGURE 3.1



$$Y^- \overset{\text{homeo}}{\cong} S^2 \times S^2 \setminus \text{pt}$$

FIGURE 3.2

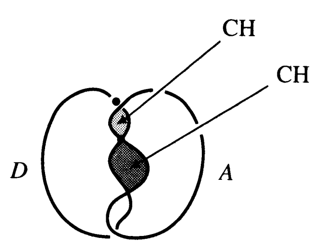


FIGURE 3.3

Figure 3.2 is an exact representation of the more intuitive schematic shown in Figure 3.3.

The Casson handles are labeled CH, and the central dot indicates a vertical axis of a 180° rotational symmetry. For this symmetry to ex-

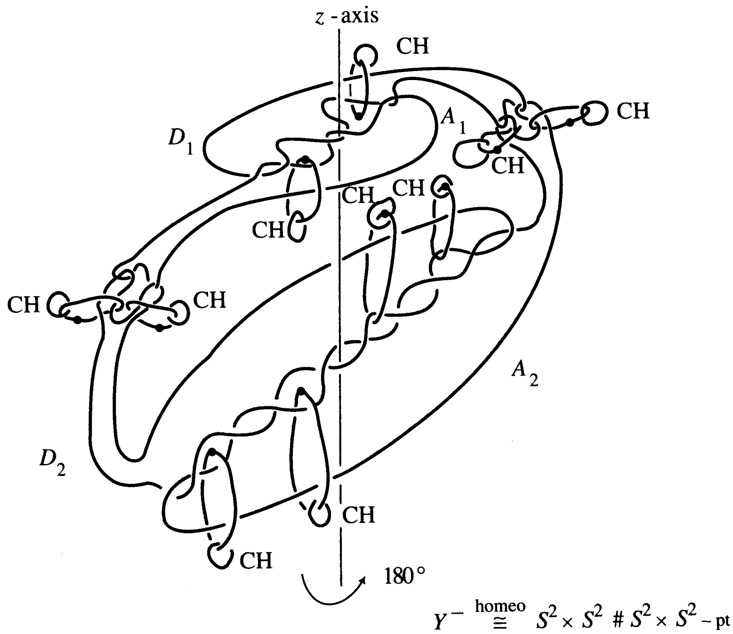


FIGURE 3.4

ist the two Casson handles should be diffeomorphic. Casson handles are parametrized by rooted, finitely branching trees. It is an easy observation that $\mathrm{CH}_{T_1} \subset \mathrm{GH}_{T_0}$ if T_1 is "more complicated" than T_0 in the obvious partial order or trees. Since any finite set of trees has a "least upper bound" more complicated than either, any finite set of Casson handles may be replaced by identical sub-Casson handles. We exploit this to assume that all Casson handles we encounter are diffeomorphic. To see that Y^- is homeomorphic to $S^2 \times S^2 \backslash \mathrm{pt}$. in Figure 3.2 we work topologically. Let the CH's cancel the 1-handles, then unravel the picture to get Figure 3.1.

If $n = 1$ and $D \cap A = 2k + 1$ points, then the picture is very similar to Figure 3.2—there will be $2k$ CH's—we do not draw it. Now suppose $n > 1$; the handlebody picture for Y should be thought of as drawn near n horizontal levels with vertical feelers corresponding to paired intersection $D_i \cap A_j$, $i \neq j$. Also to maintain symmetry extra intersections (via finger moves) must be introduced to ensure that $\mathrm{card}(D_i \cap A_j) = \mathrm{card}(D_j \cap A_i)$.

In Figure 3.4 we see the handles corresponding to $\{D_1, D_2, A_1, A_2, \mathrm{CH}_1, \cdots, \mathrm{CH}_{10}\}$. We have drawn the cases $D_i \cap A_j = 2$ points $i \neq j$,

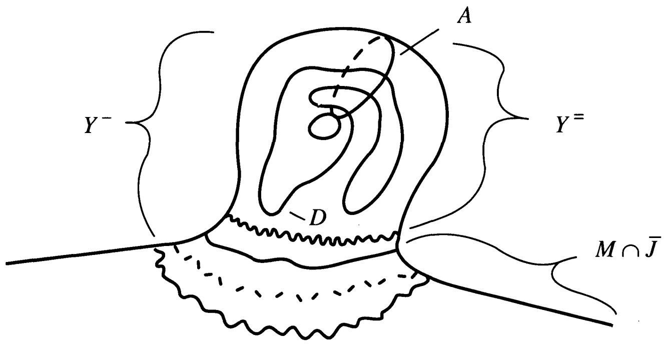


FIGURE 3.5. A PICTURE OF THE MIDDLE LEVEL M.

$D_1 \cap A_1 = 3$ points, and $D_2 \cap A_2 = 5$ points in such a way as to make the Z_2 rotational symmetry evident.

We are now ready to describe the decomposition $W = \overline{J} \cup U$. One begins by defining U and $\overline{J}$ where they intersect the middle level M. We set $M \cap U = Y^-$ and set $M \cap \overline{J}$ equal to a compact manifold neighborhood of $M \backslash Y^-$ which is small enough that $(M \cap \overline{J}) \cap Y^= = \varnothing$. Here $Y^= \subset Y^-$ is constructed by replacing $\mathscr{N}$ by a thinner regular neighborhood $\mathscr{N}_0$ of $(\bigcup_{i=1}^n D_i \cup \bigcup_{j=1}^n A_j)$ and forming $Y^= = (\mathscr{N}_0 \cup \mathrm{CH}_0\text{'s})^-$, where the CH_0's are sub-Casson handles compactly contained in the CH's and extended at the attaching region to meet $\mathscr{N}_0$. The subset $Y^=$ has the same formal handle description as Y^- (except the CH_0's are parametrized by more complicated trees) and therefore are also homeomorphic to $\#_n S^2 \times S^2 \backslash \mathrm{pt}$. (see Figure 3.5).

Given a subset of the middle level M there is a canonically determined subset of W. In the case of $M \cap \overline{J}$ the canonically determined subset $\overline{J}$ is simply a bicollar on $M \cap \overline{J}$, which in the smooth category would be described as the union of all gradient trajectories passing through $M \cap \overline{J}$. For $M \cap U$ the canonically determined subset U is a certain completion of the union of trajectories through M. U is the double trace of surgery on $\{D_i\}$ "from above" and surgery on $\{A_j\}$ "from below:" U may be identified with

$$(M \cap U) \times [0, 1] \bigcup_{\text{at level}=1} \left(\bigcup_{\text{at level}=0} h_i^3 \right) \bigcup_{\text{at level}=0} \left(\bigcup_{j=1}^{n} \hat{h}_j^2 \right),$$

where the $\widehat{\ }$ denotes dual attachment.

By construction $\overline{J}$ is a product connecting its ends $J_0 = \overline{J} \cap B$ and $J_1 = \overline{J} \cap Q$. Cancelling h_i^3 with h_i^2 (in the topological category), we see that U

is actually a topological product $(U; V_0, V_1) \overset{\text{homeo}}{\cong} (\mathbf{R}^4 \times I; \mathbf{R}^4 \times 0, \mathbf{R}^4 \times 1)$. The smooth involution σ on Y^- interchanges $\{D_i\}$ with $\{A_j\}$ and so after surgery descends to a diffeomorphism $V_0 \cong V_1$.

The smooth embedding mentioned in (5), $\operatorname{end} V_0 \to \operatorname{end} V_0$, may be identified with the restriction of σ to $\operatorname{end}(Y^-)$ and is therefore an involution.

Of the conclusions to Theorem 3.1 it remains only to check that V_0 is not diffeomorphic to $\mathbf{R}^4$. The starting point is the existence of two simply connected h-cobordant manifolds distinguished by a Donaldson type invariant (Γ, polynomial, or Φ). The original argument from this starting point was combinatorial and used only the fact that the manifold at the ends of the h-cobordism were different and did not inquire as to how they were ditinguished. This argument is well described in [14, p. 101] so we will not recapitulate it here. We will give a different argument, more in the spirit of §2 which actually provides a bit more information.

Claim. There is a compact set $K \subset V_0$ so that if $C \subset V_0$ is a smoothly embedded homology 4-ball containing K, then $\pi_1(\partial C)$ must admit nontrivial representations into SO(3).

Proof of Claim. Initially we assume B is the Barlow surface and Q is a rational surface. Let Σ_0 denote the homology 3-sphere $= \partial C$. If K is large enough, $\Sigma_0 \subset J_0$. Since Σ_0 is smoothly embedded there is a normal product collar which can be extended by deforming the Riemannian metric. This is actually a simple case of a "periodic end." As the metric is extended we create a sequence $B = B_0, B_1, B_2, \cdots \to B_\infty$. Using $\Sigma_1 \subset J_1$ we produce the analogous $Q = Q_0, Q_1, Q_2, \cdots \to Q_\infty$. Because B_∞ is isometric to Q_∞, we have $\Phi(B_\infty) = \Phi(Q_\infty)$, contradicting $|\Phi(B)| \geq 4$ and $|\Phi(Q)| = 0$ and the commutation of Φ with geometric limits (§2).

In §2 we worked with simply connected ends in order to avoid noncompactness resulting from the sequence $\{A_i\}$ limiting at infinity to nontrivial flat connections and the attendent discussions of Floer homology [9]. If all representation $\pi_1(\Sigma_0) \to \mathrm{SO}(3)$ are trivial, there are no such limits possible, so our discussion is applicable.

Now suppose that B and Q are arbitrary as in Theorem 3.1 (perhaps W is even a smooth product). Recall that Y is the interesting part of the middle level M of W. Just as any Casson handles have a "least common multiple" if Y is associated to W and Y' is associated to W', it is possible to find a Y'' (more complicated than either Y or Y') which is associated to both W and W'. We simply arrange to make our constructions so that V_0 ($V_0 = Y^-/\text{surgery}\{A_j\}$) comes from a Y which

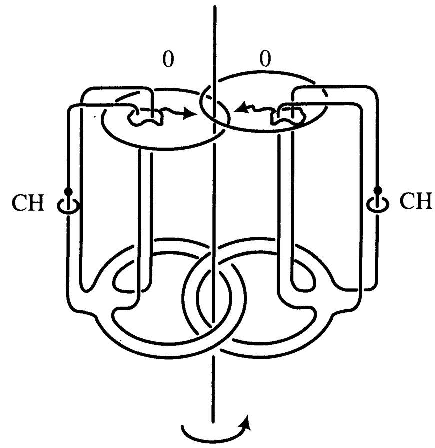


FIGURE 3.6

is complicated enough to be associated with an h-cobordism between the Barlow surface and a rational surface.

The preceding argument completes the proof of the claim and the theorem except for point (6). To justify point (6) it is sufficient to observe that Casson handles CH are realtively imbedded in standard open handles. Replacing the Casson handles in Figure 3.4 with standard handles we may cancel all 1-handles to obtain a standard diagram in which surgery on the D's yields $\mathbf{R}^4$. q.e.d.

Regarding the claim, we note that S^3 is the only known homology 3-sphere without a nontrivial representation into SO(3).

Next we describe V_0, which we referred to as $\mathbf{R}^4_1$ in the introduction, as an open subset of $\mathbf{R}^4_{\text{standard}}$ with the induced smooth structure.

We begin with a "model case" for $\mathbf{R}^4_1$ based of Figure 3.2. It is simplest but unfortunately it is not known to be associated with any nontrivial h-cobordisms and consequently not known to be exotic. Understanding the general case is only a small further step. Figure 3.2 is isotopic to Figure 3.6.

The dotted lassos in Figure 3.6 indicate two pairs of handle slides. Preform these and notice that the equivariant hyperbolic pair separates from the rest of the diagram and may be eliminated by passing from Y^- to V_0.

Passage from Y^- to V_0 is achieved by turning (by surgery) one of the smooth 2-handles into a 1-handle. This cancels the two simply linked circles in Figure 3.6 (see Figure 3.7).

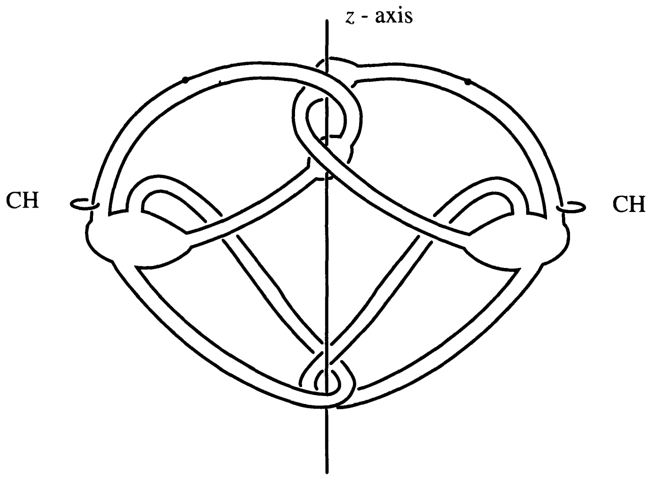


FIGURE 3.7

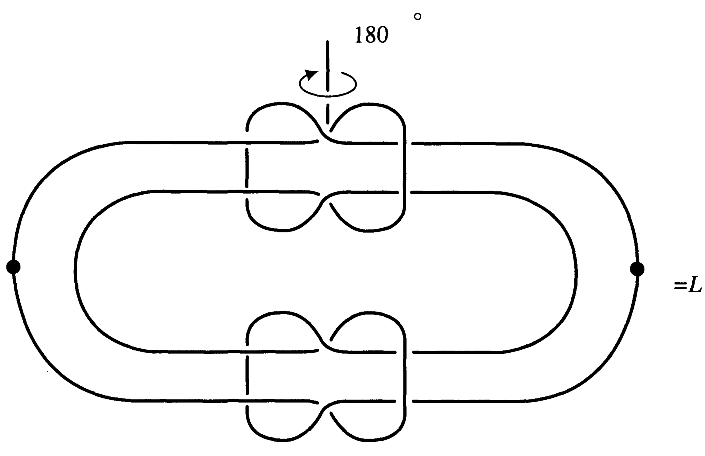


FIGURE 3.8

The reader should note how the axis of symmetry passes through the upper clasp in Figure 3.7. Our drawing contains a slight intentional break of symmetry near the clasp to clarify the location of the axis.

Bob Edwards observed that the 2-components link of 1-handle curves in Figure 3.7 has an equivalent picture (Figure 3.8) which is closely related to Bing's famous "house with two rooms." In this picture an unrelated symmetry is more apparent.

It is clear from either Figure 3.7 or Figure 3.8 that the 1-handle curves form a ribbon link L. Ribbons can be obtained from the Seifert surfaces in Figure 3.7 by surgering the upper-right and lower-left annuli (or the image of these under the involution). The reader should notice that the symmetry of the link in Figure 3.7 does not extend to symmetry of these ribbons. The ribbons visible in both Figure 3.7 and Figure 3.8 are isotopic. In a handle diagram 1-handles may be thought of as deleted trivial 2-handles. These ribbons are the image under surgery and handle cancellation of these missing 2-handle cores.

To complete our description of the "model case" recall that the simplest (unramified) Casson handle $\mathrm{CH}_1 \cong (D^2 \times \mathbf{R}^2, \partial D^2 \times \mathbf{R}^2) \backslash \mathrm{cone(Wh)}$, where $\mathrm{Wh} \subset D^2 \times S^1$ is a Whitehead continuum [11] contained in the "missing" boundary $D^2 \times$ (circle at infinity of $\mathbf{R}^2$) and the cone is a linear cone to $0 \times 0 \subset D^2 \times D^2$.

In Figure 3.7 the Casson handles topologically cancel or fill in the deleted 2-handles ("ribbons") except for a regular neighborhood of L in ∂B^4. In the smooth category the 2-handles are not entirely restored, rather a cone(Wh) is missing from each one. Again each cone is on a Whitehead continuum running along a component of L and the cone is linear in the standard smooth product structure for the restored 2-handle. To obtain V_0 $(= \mathbf{R}^4)$ in the model case remove the boundary points from the manifold described by Figure 3.7.

To summarize, the model picture for $\mathbf{R}^4_1$ is $\mathbf{R}^4$ with two closed sets deleted. Each closed set is of the form $\mathrm{Wh} \times [0, \infty)/(w, 0) \sim (w', 0)$ for all $w \in \mathrm{Wh}$. These open cones emulate from disjoint points and meet the 3-sphere S^3_∞ at infinity in two Whitehead continua imbedded along $L \subset S^3_\infty$ (Figure 3.9).

Next we describe how this model case can be modified to be general. Refer to Figure 3.4, the implied drawing of the general case, to describe $\mathbf{R}^4_t$, $t \in \mathrm{CS} \subset [0, 1]$. Notice that after surgery on the ascending spheres $\{A\text{'s}\}$ the link of 1-handle curves L' becomes a slice link. This is because L' bounds disjoint embedded disks in S^3 meeting $\bigcup_{i=1}^n D_i$ but disjoint from $\bigcup_{j=1}^n A_j$. After surgery on the $\{A\text{'s}\}$ there is a natural radial structure on the resulting B^4. Since there are no local maxima on these disks with respect to the radial structure, they are ribbon disks and L' becomes a ribbon link after surgery on the $\{A\text{'s}\}$. In fact pictures similar to Figures 3.7 and 3.8 may be drawn for a general L'.

As before, the Casson handles $\{\mathrm{CH}\}$ topologically restore the deleted ribbons neighborhoods or 2-handles (except for some boundary material).

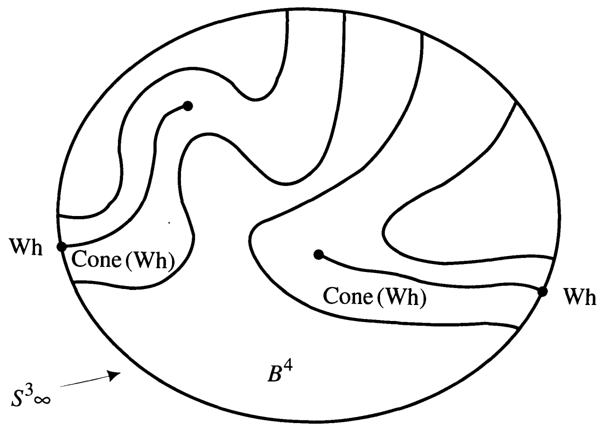


FIGURE 3.9. DELETE THE BOLD LINES $\sim S^3_\infty \cup 2(\text{cone(Wh)})$ TO GET $\mathbf{R}^4_1$.

Smoothly the restoration is only partial, a closed set X is missing from each 2-handle. In the model case we took the Casson handles to be unramified and $X = \text{cone(Wh)}$.

In the general case ramification seems inevitable, but some simplification is possible. In [13, Chapters 1–4] Casson handles are generalized (GCH) to "towers of capped groups." These differ from Casson handles in that many surface stages are interspersed between the immersed disks of Casson's construction. The GCH are also indexed by rooted finitely branching objects. (There is a slight additional structure beyond that of a tree derived from the symplectic pairing of H_1 (surface stages; Z).) If the number of surface branches grows "sufficiently fast" (the exact exponential rate is calculated in [1]) the GCH is described in [13, Chapter 4] as $(D^2 \times \mathbf{R}^2, \partial D^2 \times R^2) \backslash M(\text{WCS} \to \text{CS})$. The mapping cylinder M is an embedded family of arcs connecting a wild Cantor set $\text{WCS} \subset D^2 \times S^1$ to a tame Cantor set $\text{CS} \subset \text{interior}\, D^2 \times D^2$. Thus to form a general $\mathbf{R}^4_1$, begin with $\mathbf{R}^4_{\text{std}}$ and delete a finite number of closed sets X each of which is a Cantor set worth of rays embedded near (open) ribbon disks which emerge at S^3_∞ to form one of the links L' constructed above. The free parameters in the construction of $\mathbf{R}^4_1$ are the complexity of L' and the complexity of the GCH's. Let $\mathbf{R}^4_1$ denote any "ribbon 4-space" as above; that is, one obtained from a diagram such as Figure 3.3 where CH is replaced by GCH and the number of surface stages of each GCH grows sufficiently fast.

This completes our discussion of the final point (6) mentioned in Theorem 3.1.

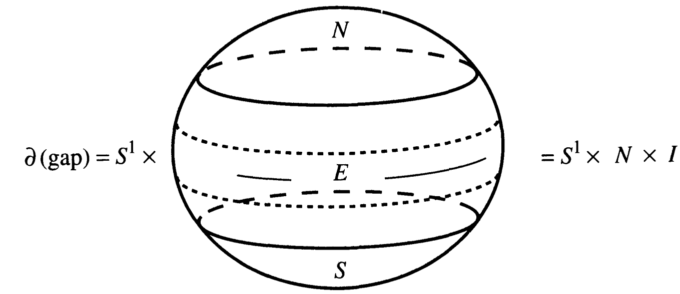


FIGURE 3.10

Theorem 3.2. *If* $\mathbf{R}_1^4$ *is a ribbon 4-space and* $K \subset \mathbf{R}_1^4$ *is a compactum, then there exists a topological radius function (polar coordinate)* $\rho\colon \mathbf{R}_1^4 \to [0, +\infty)$ *such that if we set* $t = 1 - 1/r$ *and* $\mathbf{R}_t^4 = \rho^{-1}([0, r))$ *then* $K \subset \mathbf{R}_0^4$ *and* $\mathbf{R}_t^4$ *is also a ribbon 4-space for* t *belonging to the standard Cantor set* $\mathrm{CS} \subset [0, 1]$.

Proof. The key players in this argument are the "gaps" in the original proof [11] of the Poincaré conjecture. These are not mistakes but rather a countable collection of closed sets in a CH which resisted direct description. Working with a GCH as above, the gaps still exist and $\overline{\mathrm{GCH}\backslash\{\mathrm{gap}\}}$ is called the design: it admits an essentially explicit foliation $\mathscr{F}$ by 3-manifolds and each gap is a compact 4-manifold homotopy equivalent to a circle with boundary $\partial\mathrm{gap} = S^1 \times S^2 = S^1 \times N \cup S^1 \times E \cup S^1 \times S$ (see Figure 3.10).

Leaves of $\mathscr{F}$ come in perpendicularly to $S^1 \times E$ along $S^1 \times$ longitudes; $S^1 \times N$ and $S^1 \times S$ are subsets of leaves.

Although $\partial(\mathrm{gap}) \cong S^1 \times S^2$ and $\mathrm{gap} \simeq S^1$, it was not known at the time of the proof that the gap was homeomorphic to $S^1 \times D^3$. Instead it was dealt with by decomposition methods.

In fact,

$$(\mathrm{gap}; S^1 \times N, S^1 \times E, S^1 \times S) \overset{\mathrm{top}}{\cong} (S^1 \times D^3; S^1 \times N, S^1 \times E, S^1 \times S).$$

This may be proved in two ways. One can quote [12], in which the topological surgery sequence is established for $\pi_1 \cong \mathbf{Z}$, to recognize the gap. Alternatively the fact that the gap imbeds in $\mathbf{R}^4$ and can itself be partially explored by a 3-dimensional foliation permits the original proof—including all the decomposition arguments—to be adapted to produce the desired homeomorphism.

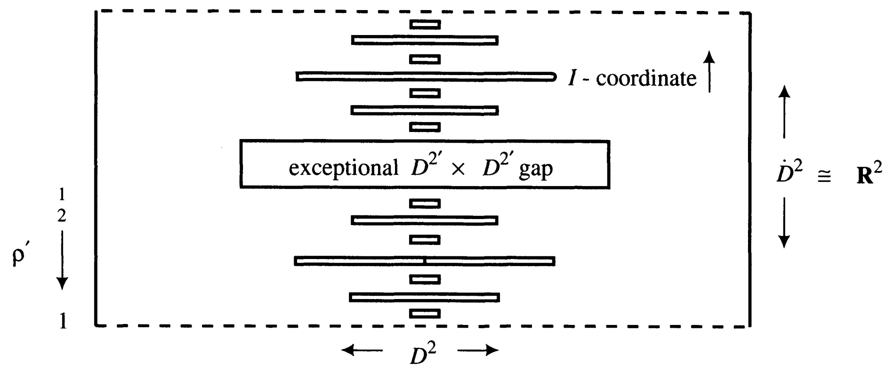


FIGURE 3.11. A TOPOLOGICAL COORDINATIZATION OF THE DESIGN, A FOLIATED SUBSET OF GCH. THE COORDINATES EMBED DESIGN $\subset D^2 \times D^2$. ALL SMALLER RECTANGLES ARE $S^1 \times D^3$-GAPS.

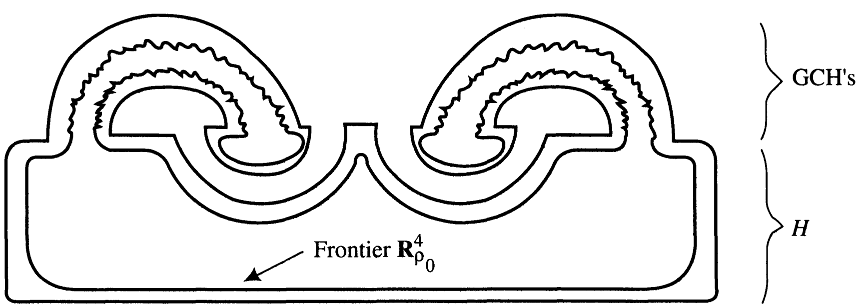


FIGURE 3.12. $\mathbf{R}^4_1 = $ FIGURE$\backslash$BOUNDARY.

Thinking of each gap as $(S^1 \times N \times I;\, S^1 \times N \times 1,\, S^1 \times \partial N \times I,\, S^2 \times N \times 0)$ we can reparametrize the interval coordinate to be compatible with the radial coordinate ρ' on the co-core direction of the design. The schematic in Figure 3.11 indicates the gaps fitting into the design with their I-product structure matching the ρ' levels.

If $2\rho'_0 - 1 \in \mathrm{CS} \subset [0, 1]$, then $D^2 \times D^1_{\rho'_0}$ is a "complete" $D^2 \times S^1$ leaf disjoint from the interior of any gap.

Figure 3.11 suggests the correct topological coordinatization of a GCH. The preimage of $[0, \rho'_0)$ is a sub-GCH for $2\rho'_0 - 1 \in \mathrm{CS} \subset [0, 1]$.

Since $\mathbf{R}^4_1$ arises from a smooth handle body H by attaching GCH's with radius $\rho' = 1$ (and then deleting the remaining boundary) and since ρ' is perfectly standard in a neighborhood of the attaching region, it is a simple matter to find a topological product structure near $\mathrm{end}(\mathbf{R}^4_1)$ whose levels are ρ' levels on the GCH and agree with a previously specified product structure near ∂H away from a neighborhood of the $\bigcup$GCH (see Figure 3.12). The Schoenflies theorem allows this product structure near infinity to be extended to a global radial coordinate. The further conditions of Theorem 3.2 are now just a question of normalization. q.e.d.

We emphasize that the difference among the $\mathbf{R}_t^4$'s in the family for $t \in \mathrm{CS} \subset [0, 1]$ lies only in the details of the wild Cantor set (and therefore the associated mapping cylinder). These wild Cantor sets are all defined by iterated-ramified Bing and Whitehead doubling and differ only with regard to various multiplicities.

Remark 3.1. If B and Q are as in the statement of Theorem 3.1, then the constructed involution (5) may be thought of as changing the smooth structure on B by cutting out $\mathbf{R}_1^4$ and gluing it back with a twist (of order 2). In some ways this is analogous to the classical construction of exotic structures, e.g., on S^n, $n = 7$, by cutting out a closed ball B^n and regluing it via an exotic diffeomorphism of $\partial B^n = S^{n-1}$.

Remark 3.2. Donaldson's Γ- and polynomial-invariants require consideration of pullbacks of (bundle, connection) to a singular surface. To compare the pullbacks corresponding to surfaces in B and Q (as in Theorem 3.1), notice that any singular surface $\Sigma_1 \looparrowright B$ may be pushed out along a topological radial coordinate on V_0 and so is homotopic to a singular surface Σ_0' contained in J_0 (similarly for $\Sigma_1 \looparrowright Q$). In fact, by immersion theory if $\Sigma_0 \looparrowright B$ and $\Sigma_0' \looparrowright J_0$ are immersed with isomorphic normal bundles, then they are regularly homotopic. Thus singular surfaces can always be assumed to lie in the region of diffeomorphy.

4. Main theorem

Any open subset of a Euclidean space acquires a natural smooth structure by restriction. By a topological radial function we mean a homeomorphism followed by the usual radius function $\mathbf{R}_{\mathrm{std}}^4 \to [0, \infty)$.

Theorem 4.1. *There is a subset $\mathbf{R}_1^4$ of Euclidean 4-space $\mathbf{R}_{\mathrm{std}}^4$ which is homeomorphic but not diffeomorphic to $\mathbf{R}_{\mathrm{std}}^4$. Moreover, $\mathbf{R}_1^4$ has a topological system of polar coordinates with radial function ρ so that the open balls of radius r are "ribbon 4-spaces" $\mathbf{R}_t^4$ in the sense of* §3 *whenever $t = 1 - 1/r$ belongs to the standard Cantor set* $\mathrm{CS} \subset [0, 1]$. *We say t and t' are equivalent if $\mathbf{R}_t^4$ and $\mathbf{R}_{t'}^4$ are diffeomorphic. The resulting equivalence classes are at most countable. Also, ρ restricted to $[0, r]$ for any r extends to a topological polar coordinate on all of $\mathbf{R}_{\mathrm{std}}^4$.*

Corollary 4.1. *In Zermelo-Frankel set theory with choice* (ZFC) *we may assert that there is a collection of parameter values* $\{t'\} \subset \mathrm{CS} \subset [0, 1]$ *with* $\mathrm{card}(\{t'\})$ = *"continuum"* = $\mathscr{C}$ *such that the subsets* $\mathbf{R}^4_{t'}$ *are pairwise nondiffeomorphic.*

Proof of Theorem 4.1. Referring to the outline in the introduction, we need only comment on a couple of points. According to Theorem 3.1(2), $W = \overline{J} \cup U$. This means that the end of U is contained in the (smooth) product region $\overline{J}$. The compact set $K \subset V_0 \subset B$ of the introduction should be chosen to be a smooth codimension zero submanifold large enough so that $V_0 \backslash K \subset J_0 \subset B$ and with ∂K invariant under the involution (Theorem 3.1(5)). (Essentially explicit examples of such K can be derived from Figure 3.4 and the handle description of the GCH's.) Thinking of Q as formed from B by regluing $\mathbf{R}^4_1$ (as in Remark 3.1) and remembering the invariance of ∂K, we see that: (1) K embeds in Q, and (2) the product structure on $\overline{J}$ induces a diffeomorphism $B\backslash K \cong Q\backslash K$.

Now consider the radial coordinate ρ of Theorem 3.2 where the compact set is specified to be this same K. We depart slightly from the introduction by considering where $\rho = r$, where $t = 1 - 1/r \subset \mathrm{CS} \subset [0, 1]$. Suppose there is a diffeomorphism

$$(d, \mathrm{id}_K)\colon (\mathbf{R}^4_s, K) \to (\mathbf{R}^4_t, K), \qquad s \neq t \in \mathrm{CS}.$$

As in the introduction, this allows the construction of end periodic metrics: B_∞ isometric to Q_∞. This leads via Theorem 2.1 to a contradiction and the conclusion that if there is a diffeomorphism $d\colon \mathbf{R}^4_s \to \mathbf{R}^4_t$, then d restricted to K is not the identity and, in fact, not isotopic to the identity.

It is a general fact that a smooth compact manifold admits only countably many smooth embeddings, up to isotopy, into any smooth metrizable manifold. In dimension four, the equivalence of the smooth and PL categories allows an effortless argument. A PL embedding means a simplicial embedding after k barycentric subdivisions of domain and range for some nonnegative integer k, but the number of simplicial embeddings of a finite complex into a countable one is certainly countable. Thus there are only countably many PL embeddings of a finite complex into a countable complex. The analogous statement—up to isotopy—now follows in the smooth category [15].

Combining the previous two paragraphs, we see that the subsets $\mathbf{R}^4_t$ could be mutually diffeomorphic for at most countably many parameter values $t \in \mathrm{CS}$.

The final assertion of Theorem 4.1 follows from the Schoenflies theorem.

Proof of Corollary 4.1. In the presence of choice, we may well-order the set of equivalence classes $\mathscr{E}$. If we do this with least order type we will have achieved a partition of $\mathscr{E}$ into a disjoint union of a set of cardinally $\aleph_0$: $\mathscr{E} = \bigcup_{i\in\mathscr{X}} \mathscr{E}_i$, $\operatorname{card}(\mathscr{E}_i) = \aleph_0$. A typical $\mathscr{E}_1$ begins with a limit ordinal and contains all its successors less than the next limit ordinal. The standard diagonal argument shows that the cardinality of $T_i = \{t \in \mathrm{CS}|$ equivalence class of $\mathbf{R}^4_t$ lies in $\mathscr{E}_i\}$ is $\aleph_0$ for each i. Thus for each i we have some bijection between T_i and $\mathscr{E}_i$. Taking the union over $i \in \mathscr{I}$ we have a bijection

$$\mathrm{CS} = \bigcup_{i\in\mathscr{X}} T_i \Leftrightarrow \bigcup_{i\in\mathscr{X}} \mathscr{E}_i = \mathscr{E}.$$

If the axiom of choice is not permitted, or is substantially weakened, a correspondingly weaker conclusion about the cardinally of $\mathscr{E}$ is obtained.

Appendix A: Gauges for low energy connections

Cliff Taubes informed us, early in our work, that a useful technical result [24, Lemma 10.4] omits a necessary hypothesis on simple connectivity at infinity. This appendix is one way to make Lemma 10.4 precise.

Lemma 10.4′ [24]. *Let $K_0 \subset \mathring{K}_1$ be an inclusion of compact Riemannian 4-manifolds with boundary with $\pi_1(K_0) \overset{\mathrm{inc}_\#}{\to} \pi_1(K_1)$ zero. Let A be a connection on K_1 such that $F_A^+ = 0$ and $\|F_A^-\|_{L^2(\text{on } K_1)} < \varepsilon$, where the exact choice of ε depends on the inclusion $K_0 \hookrightarrow K_1$. Then there exists a gauge transformation g such that $A^g = d + a$ and*

$$(\|a\|_{L^p_k})^2 < C(p, k, K_0 \hookrightarrow K_1) \int_{K_0} \|F_A\|^2. \tag{A.1}$$

Proof. By Uhlenbeck ([27], [28]) there are a finite convex covering $\{U_\alpha\}$ of K_1 and gauge transformations g_α on U_α such that $A^{g_\alpha} = d + a_\alpha$ on U_α and (A.1) holds when the integrals take over U_α. We want to glue the g_α into a global gauge transformation achieving (A.1).

If $g_{\alpha\beta} = g_\alpha^{-1} g_\beta$ on $U_\alpha \cap U_\beta$, then

$$a_\alpha = g_{\alpha\beta}^{-1} a_\beta g_{\alpha\beta} + g_{\alpha\beta}^{-1}\, d g_{\alpha\beta}. \tag{A.2}$$

The bounds on the a_α give bounds on $dg_{\alpha\beta}$, hence we can write $g_{\alpha\beta} = \overline{g}'_{\alpha\beta}(1 + h_{\alpha\beta})$ with $|h_{\alpha\beta}|^2_{L^p_k} \le c' \int |F_A|^2$ for some constant c' and $\overline{g}'_{\alpha\beta}$

constant functions. Think of $\overline{g}'_{\alpha\beta}$ as an "average" value for $g_{\alpha\beta}$. Unfortunately $\overline{g}'_{\alpha\beta}$ is not a cocycle, however we have

Claim. There exist constants $\overline{g}_\alpha$ such that

$$\|\overline{g}'_{\alpha\beta} - \overline{g}_{\alpha\beta}\|^2 < \text{const.} \int_{K_2} |F_A|^2, \quad \text{where } \overline{g}_{\alpha\beta} = \overline{g}_\alpha^{-1}\overline{g}_\beta. \tag{A.3}$$

By using (A.2) and the Sobolev embedding theorem the norm on the left can be any C^k norm.

Proof of Claim. We use log and exp to pass between small elements of SO(3) and its Lie algebra, this allows averages to be taken. We suppress notational distinctions between the group near id and the Lie algebra.

Let G^i_δ denote the constant i-cochains on K_1 with coefficients in the structure group. In our case, this means Čech cochains based on $\{U_\alpha\}$ with coefficients in $\mathrm{SO}(3)_\delta$; the δ indicates the discrete topology. The restrictions to K_0 will be denoted by r. For example, $rG^0_\delta = \{\overline{g}_\alpha$ on $U_\alpha \cap K_0\}$ and $rG^1_\delta = \{\overline{g}_{\alpha\beta}$ on $U_\alpha \cap U_\beta \cap K_0\}$, etc. There is an action of G^0_δ ($\cong \mathrm{SO}(3)^{\text{order cover}}$) on G^1_δ which is compatible with restriction and defined by $g_{\alpha\beta} \mapsto g_\alpha^{-1} g_{\alpha\beta} g_\beta$.

From the cocycle condition for $\{g_{\alpha\beta}\}$ and the smallness of $h_{\alpha\beta}$ we obtain

$$\|\overline{g}'_{\alpha\beta}\overline{g}'_{\beta\gamma}\overline{g}'^{-1}_{\alpha\gamma} - \mathrm{id}\| < C' \int |F_A|^2. \tag{A.4}$$

We are interested in the coboundary map $rG^1_\delta \xrightarrow{\partial} G^2_\delta$ given by $\delta(r\overline{g}'_{\alpha\beta}) = \bigcup_\gamma(\overline{g}^1_{\alpha\beta}\overline{g}'_{\beta\gamma}\overline{g}'^{-1}_{\alpha\gamma})$. The differential of the coboundary at the identity is well known to be

$$\{h_{\alpha\beta}\} \overset{d(\delta)}{\mapsto} \{h_{\alpha\beta} + h_{\beta\gamma} - h_{\alpha\gamma}\}. \tag{A.5}$$

(Think of h as identified by logarithms with an element of the Lie algebra so(3).)

The coboundary map above is a smooth map between finite dimensional manifolds and is equivariant with respect to the action of G^0_δ. Furthermore, $\delta^{-1}(0)/G^0$ is naturally isomorphic to the conjugacy classes of representations $\pi_1(K_0) \to \mathrm{SO}(3)$ induced by representations $\pi_1(K_1) \to \mathrm{SO}(3)$. By the hypothesis, $\pi_1(K_0) \overset{\text{zero}}{\to} \pi_1(K_1)$ so $\delta^{-1}(0)/G^0$ is a point. Put another way, $\check{H}^1(K_1; \mathrm{SO}(3)_\delta) \to \hat{H}^1(K_0; \mathrm{SO}(3)_\delta)$ is zero. Direct calculation with (A.5) shows that if T is an infinitesimal *slice* for the action of G^0_δ on

rG^1_δ at id, then the differential of the coboundary map $d(\delta)$ is one-to-one on T. Thus if $\overline{g}'_{\alpha\beta}$ is sufficiently close to a cocycle then it may be gauged (by $\overline{g}_\alpha$) to $\{t_{\alpha\beta}\} \in T$ such that

$$\text{(A.6)} \qquad \|\{t_{\alpha\beta} - \mathrm{id}\| < \text{const.}'' \|\delta\{t_{\alpha\beta}\} - \mathrm{id}\| < \text{const.}' \int |F_A|^2,$$

where

$$\text{(A.7)} \qquad t_{\alpha\beta} = \{\overline{g}_\alpha\} \circ \{\overline{g}^{-1}_{\alpha\beta}\} = \{\overline{g}_\alpha^{-1} \overline{g}^{-1}_{\alpha\beta} \overline{g}_\alpha\}.$$

But (A.6) and (A.7) imply $\|\overline{g}'_{\alpha\beta} - \overline{g}_\alpha^{-1} g_\beta\| < \text{const.} \int |F_A|^2$, proving the claim.

Now that $\overline{g}_\alpha$ has been constructed, consider a local gauge transformation

$$\text{(A.8)} \qquad \hat{g}_\alpha = g_\alpha \overline{g}_\alpha^{-1}.$$

Since $\overline{g}_\alpha$ is constant, the associated a for $\hat{g}_\alpha$ is the same as for g_α, thus (A.1) is satisfied over U_α for $\hat{g}_\alpha$. Moreover, by the claim,

$$\text{(A.9)} \qquad \|\hat{g}_{\alpha\beta} - \mathrm{id}\| \le \text{const.} \int \|F_A\|^2, \quad \text{where } \hat{g}_{\alpha\beta} = \hat{g}_\alpha^{-1} \hat{g}_\beta.$$

Thus if $\int \|F_A\|^2$ is sufficiently small, the 1-cocycle $\{\hat{g}_{\alpha\beta}\}$ is arbitrarily close to the identity. We may employ the following

Fact. Let P be an H-principal bundle over an n-dimensional manifold with a cover (U_α) corresponding to a slightly thickened handle decomposition. Suppose that a 1-cocycle $\{g_{\alpha\beta}\}$ for P maps entirely within a ball of radius ε about $\mathrm{id} \in H$ and suppose that the larger ball of radius $n\varepsilon$ is still small enough to be contractible; then P is trivial. Furthermore $\{g_{\alpha\beta}\}$ is the coboundary of $\{g_\alpha\}$ with values in the ball of radius $n\varepsilon$.

Proof of Fact. Set $g_\alpha = \mathrm{id}$ for U_α a thickened zero-handle. For U_α a thickened 1-handle, g_α is defined to be some homotopy in $\mathrm{Ball}_\varepsilon(\mathrm{id})$ between the $g_{\alpha\beta}$'s specified near the *attaching region.* Because this homotopy lies in $B_\varepsilon(\mathrm{id})$, the "new" $g_{\alpha\beta}$ between thickened 2-handles and thickened handles of lower index is now the product of functions into $B_\varepsilon(\mathrm{id})$ and therefore lies in $B_{2\varepsilon}$. Similarly g_α on thickened 2-handles is some null homotopy of these new $g_{\alpha\beta}$ in $B_{2\varepsilon}(\mathrm{id})$. Continuing this way, the thickened k-handles are gauged by maps to $B_{k\varepsilon}(\mathrm{id})$, proving the fact.

Since there is no loss of generality in choosing our $\{U_\alpha\}$ to derive from a handle decomposition, we can use this fact to write

(A.10)
$$\hat{g}_{\alpha\beta} = \hat{\hat{g}}_{\alpha}^{-1} \hat{\hat{g}}_{\beta}$$

with

$$\|\hat{\hat{g}}_{\alpha} - \mathrm{id}\|^2 < 4\,\mathrm{const.} \int \|F_A\|^2.$$

Finally if we set $f_\alpha = \hat{\hat{g}}_{\alpha}^{-1} \hat{g}_\alpha$, then f_α will satisfy (A.1) (with C replaced by a larger constant), and since $\{\hat{\hat{g}}_\alpha\}$ and $\{\hat{g}_\alpha\}$ have the same coboundary, $f_\alpha = f_\beta$ on $U_\alpha \cap U_\beta$. This is the desired global gauge transformation.

Appendix B: L^2 implies $L^2_{,\delta}$

The proof of this fact in the case of an end isometric go $S^3 \times [0; +\infty)$ can be found in the book of Donaldson [8].

Let M be a smooth 4-manifold with a single and F homeomorphic to $S^3 \times \mathbf{R}$. We consider a case somewhat more general than "asymptotically periodic." We say E has *bounded geometry* if there is a proper Morse function $\tau\colon E \to [0, \infty)$ such that the set of compact submanifolds $\{\tau^{-1}[x, y];\, x, y$ regular values of f and $|x - y| <$ const.$\}$ is compact in the sense that every sequence has a subsequence N_i admitting $(1 + \frac{1}{i})$-quasi-isometries between all N_j and N_k for $j, k > i$. (For i large, there is no loss of generality in assuming that these quasi-isometries are diffeomorphisms [22].)

Theorem. *Let A be an ASD connection on M with $A = d + a$ on F and $\int_E |a|^2 < \infty$. There exist a gauge transformation g, $A^g = d + a'$ on E, and a $\delta > 0$, so that $\int_E e^{\tau\delta}|a'|^2 < \infty$.*

Proof. Let $I_t = \int_{\tau > t} |F_A|^2$. As in [24, Proposition 10.5] (using our Appendix A to replace 10.4) the theorem will be proved if we show I_t has exponential decay. We do this.

The other hypotheses imply: (1) (After a linear rescaling of τ) a topological factor 3-sphere lies in each segment $\tau^{-1}[i, i + 1]$, and (2) there exist constants $0 < b < c$ and a closed subset $\Delta \subset [0, \infty)$ such that $b < |\mathrm{grad}\,\tau(x)| < c$ for $x \in \tau^{-1}(\mathbf{R}^+ - \Delta)$ and where the measure of Δ intersected with any interval of length one is smaller than $\frac{1}{2}$. We think of $\tau^{-1}(\mathbf{R}^+\backslash\Delta)$ as the "far from critical" subset.

Assuming (1) above and Appendix A, we find that for N large enough, $\int_{\tau^{-1}[N-1, N+2]} |F_A|^2 = E < \varepsilon$ is so small that after a gauge transformation over $\tau^{-1}[N-1, N+2]$ the C^k norm of a on $\tau^{-1}[N, N+1]$ is bounded by const.$\,E^{1/2}$. By "bounded geometry," the constant is uniform for N

large. In this local gauge we have for $t_0 \in [N, N+1]\backslash\Delta$:

$$\begin{aligned}\left|\frac{DI_t}{dt}\right|_{t=t_0'} &= \int_{\tau^{-1}(t_0)} |da + a\wedge a|^2 \operatorname{grad}\tau\, d\mathrm{vol}_3 \\ &\geq b\int_{\tau^{-1}(t_0)} |da + a\wedge a|^2\, d\mathrm{vol}_3 \\ &\geq b\int_{\tau^{-1}(t_0)} |da|^2\, d\mathrm{vol}_3 - b\int_{\tau^{-1}(t_0)} |a\wedge a|^2\, d\mathrm{vol}_3 \\ &\geq b\int_{\tau^{-1}(t_0)} |da|^2\, d\mathrm{vol}_3 - b\ \text{const.}\left(\int_{\tau^{-1}[N,N+1]} |F_A|^2\, d\mathrm{vol}\right)^2.\end{aligned} \tag{B.1}$$

Also if E is small enough, only the value of the final constant is affected by the additional requirement that a be a "Hodge gauge" relative to ordinary d, i.e., that the components of a are coclosed real valued 1-forms. Because the inclusion $\tau^{-1}[N-1, N] \subset \tau^{-1}[N-2, N+1]$ is trivial on π_1, the restriction map on $H^1_{\mathrm{DR}}(\ ;\mathbf{R})$ is zero. Thus the projection of the components of a into harmonic forms is trivial. Thus the components of a are perpendicular to the closed 1-forms. From this follows the estimate

$$\frac{1}{2}\int_{\tau^{-1}(t_0)} |da|^2\, d\mathrm{vol}_3 \geq \frac{\lambda_1}{3}\int_{\tau^{-1}(t_0)} \operatorname{trace}(a\wedge da)\, d\mathrm{vol}_3, \tag{B.2}$$

where λ_1 is the first nonzero eigenvalue of curl on $\tau^{-1}(t_0)$. Again by "bounded geometry" for N large there is a uniform lower bound $\lambda \leq \lambda_1(t_0)$. The factor $\frac{1}{2}$ comes from the fact that A is ASD so that the three curvature targets to $\tau^{-1}(t_0)$ are equal in pairs to three complementary curvatures.

Combining (B.1) and (B.2) we get, for some new constants C_1 and C_2, and $t_0 \in [N-1, N+2]\backslash\Delta$, N large enough:

$$\begin{aligned}\left|\frac{dI_t}{dt}\right|_{t=t_0} &\geq C_1\int_{\tau^{-1}(t_0)} \operatorname{trace}(a\wedge da)\, d\mathrm{vol}_3 \\ &\quad - C_2\left(\int_{\tau^{-1}[N,N+1]} |F_A|^2\, d\mathrm{vol}_3\right)^2.\end{aligned} \tag{B.3}$$

Since our bundle is trivialized over the end, the Chern-Simons invariant CS of $(\tau^{-1}(t_0), A|_{\tau^{-1}(t_0)})$ is a well-defined real number and has a 3-dimensional description very close to (B.3):

$$\mathrm{CS} = \int_{\tau^{-1}(t_0)} \operatorname{trace}(a\wedge da + \tfrac{2}{3}a\wedge a\wedge a)\, d\mathrm{vol}_3, \tag{B.4}$$

and also by Chern-Weyl theory a 4-dimensional description [3]

$$\text{(B.5)} \qquad \mathrm{CS} = \int_{\tau^{-1}(t_0,\infty)} |F_A|^2 \, d\mathrm{vol} = I_t.$$

From (B.4) we have (for N large and therefore $\int_{\tau^{-1}[N-1,N+2]} |F_A|^2 \, d\mathrm{vol}$ small),

$$\text{(B.6)} \qquad \mathrm{CS} \le \int_{\tau^{-1}(t_0)} \mathrm{trace}(a \wedge da) \, d\mathrm{vol}_3 + C_3 \left(\int_{\tau_{-1}[N,N+1]} |F_A|^2 \right)^3 .$$

Again, if the energy on $\tau^{-1}[N-1, N+2]$ is small enough, combining (B.4) with (B.3) and (B.6) gives

$$\text{(B.7)} \qquad \left| \frac{dI_t}{dt} \right| \ge C_1(\mathrm{CS}) - C_4 \left(\int_{\tau^{-1}[N-1,N+2]} |F_A|^2 \, d\mathrm{vol} \right)^{3/2}$$

$$\text{(B.8)} \qquad = C_1 I_t - C_4 \left(\int_{\tau^{-1}[N,N+1]} |F_A|^2 \, d\mathrm{vol} \right)^{3/2} \quad \text{(using (B.5))}.$$

Since I_t is strictly decreasing, (B.8) implies that on a set $\mathbf{R}^+ \backslash \Delta$ of density $\ge \frac{1}{2}$ either the rate of decrease is $\ge \frac{C_1}{2} I_t$ or $C_4(I_N - I_{N+1})^{3/2} \ge \frac{C_1}{2} I_{n+1}$. An elementary argument shows that any combination of these two alternatives on $\mathbf{R}^+ \backslash \Delta$ gives at least exponential decay, hence

$$\text{(B.9)} \qquad I_t \le C_5 e^{-\tau\delta} \quad \text{for some } \delta.$$

References

[1] R. Ancel & M. Starbird, *The shrinkability of Bing-Whitehead decompositions*, Topology **28** (1989) 291–304.

[2] A. Casson, *Three lectures on new infinite constructions in 4-dimensional manifolds*, Lecture notes by Lucian Guillon, Prepubl. Orsay 81T06, Orsay, France, 1974.

[3] S. S. Chern & J. Simons, *Characteristic forms and geometric invariants*, Ann. of Math. **99** (1972) 48–69.

[4] A. Dold & H. Whitney, *Classification of oriented sphere bundles over a 4-complex*, Ann. of Math. **69** (1959) 667–677.

[5] S. K. Donaldson, *Irrationality and the h-cobordism conjecture*, J. Differential Geometry **26** (1987) 141–168.

[6] ——, *Polynomial invariants for smooth four-manifolds*, preprint Oxford, 1987.

[7] ——, *Orientation of Yang-Mills moduli spaces and 4-manifold topology*, J. Differential Geom. **26** (1987) 397–428.

[8] S. K. Donaldson & P. Kronheimer, *The geometry of 4-manifolds*, Oxford University Press, London, 1990.

[9] A. Floer, *Witten's complex and infinite dimensional Morse theory*, J. Differential Geometry **30** (1989) 207–222.

[10] D. Freed & K. Uhlenbeck, *Instantons and 4-manifolds*, MSRI Series, Vol. 1, Springer, Berlin, 1984.

[11] M. Freedman, *The topology of* 4*-dimensional manifolds*, J. Differential Geom. **17** (1982) 357–453.
[12] ——, *The disk theorem for four-dimensional manifolds*, Proc. Internat. Congr. Math. (August 1983, Warsaw) 647–663.
[13] M. H. Freedman & F. Quinn, *The topology of four-dimensional manifolds*, Princeton University Press, Princeton, NJ, 1990.
[14] R. C. Kirby, *The topology of* 4*-manifolds*, Lecture Notes in Math., Vol. 1374, Springer, Berlin, 1989.
[15] R. Kirby & L. Siebenmann, *Foundational essays on topological manifolds, smoothing, and triangulations*, Annals of Math. Studies, No. 88, Princeton University Press, Princeton, NJ, 1977.
[16] D. Kotschick, *On manifolds homeomorphic to* $\mathrm{CP}^2\#8\overline{\mathrm{CP}}^2$, Invent. Math. **95** (1989) 591–600.
[17] H. B. Lawson, *The theory of gauge fields in four dimensions*, Regional Conf. Ser., No. 58 Amer. Math. Soc., Providence, RI, 1985.
[18] J. Milnor, *Lectures on the h-cobordism theorem*, Princeton Math. Notes, Princeton University, Princeton, NJ, 1965.
[19] T. Mrowka, *A local Mayer-Vietoris principle for Yang-Mills moduli spaces*, Ph.D. Thesis, University of California, Berkeley, 1989.
[20] W. Rudin, *Functional analysis*, McGraw-Hill, New York, 1973.
[21] S. Sedlacek, *The Yang-Mills functional on* 4*-manifolds*, Ph.D. Thesis, University of Chicago, 1983.
[22] Y. Shikata, *On a distance function on the set of differentiable structures*, Osaka J. Math. **3** (1966) 65–79.
[23] C. Taubes, *Self-dual connections on non-self-dual* 4*-manifolds*, J. Differential Geometry **17** (1982) 134–170.
[24] ——, *Gauge theory on asymptotically periodic* 4*-manifolds*, J. Differential Geometry **25** (1987) 363–430.
[25] ——, *Moduli spaces of* L^2 *connections on open manifolds*, preprint.
[26] F. Treves, *Basic linear partial differential equations*, Academic Press, New York, 1975.
[27] K. K. Uhlenbeck, *Removable singularities for Yang-Mills fields*, Comm. Math. Phys. **83** (1982) 11–30.
[28] ——, *Connections with* L^p*-bounds on curvature*, Comm. Math. Phys. **83** (1982) 31–43.

HARVARD UNIVERSITY
UNIVERSITY OF CALIFORNIA AT SAN DIEGO