THE 4-DIMENSIONAL LIGHT BULB THEOREM

DAVID GABAI

Abstract. For embedded 2-spheres in a 4-manifold sharing the same embedded transverse sphere homotopy implies isotopy, provided the ambient 4-manifold has no $\mathbb{Z}_2$-torsion in the fundamental group. Among other things, this leads to a generalization of the classical light bulb trick to 4-dimensions, the uniqueness of spanning discs for simple closed curves in S^4 and $\pi_0(\mathrm{Diff}_0(S^2 \times D^2)/\,\mathrm{Diff}_0(B^4)) = 1$.

1. Introduction

In his seminal work on immersions[Sm1] Steven Smale classified regular homotopy classes of immersions of 2-spheres into Euclidean space and more generally into orientable smooth manifolds. In [Sm2] he gave the regular homotopy classification of immersed spheres in $\mathbb{R}^n$ and asked:

Question 1.1. *(Smale, P. 329* [Sm2]*) Develop an analogous theory for imbeddings. Presumably this will be quite hard. However, even partial results in this direction would be interesting.*

This paper works in the smooth category and addresses the question of isotopy of spheres in 4-manifolds. In that context Smale's results [Sm1] show that two embeddings are homotopic if and only if they are regularly homotopic. Given that 2-spheres can knot in 4-space, isotopy is a much more restrictive condition than homotopy. Indeed, the author is aware of only one unconditional positive result and that was proved more than 50 years ago: A 2-sphere in a 4-manifold that bounds a 3-ball is isotopic to a standard inessential 2-sphere, [Ce1], [Pa].

Recall that a *transverse sphere* G to a surface R in a 4-manifold is a sphere with trivial normal bundle that intersects R exactly once and transversely. The following is the main result in this paper.

Theorem 1.2. *Let M be an orientable 4-manifold such that $\pi_1(M)$ has no 2-torsion. Two embedded 2-spheres with common transverse sphere are homotopic if and only if they are ambiently isotopic via an isotopy that fixes the transverse sphere pointwise.*

The generalization to multiple pairs of spheres is given in §10.

Here are some applications.

Theorem 1.3. *A properly embedded disc in $S^2 \times D^2$ is properly isotopic to a fiber if and only if its boundary is standard.*

Version 1.01, August 7, 2017

Partially supported by NSF grants DMS-1006553, 1607374.

Theorem 1.4. *Two properly embedded discs D_0 and D_1 in $S^2 \times D^2$ that coincide near their standard boundaries are properly isotopic rel boundary if and only if they are homologous in $H_2(S^2 \times D^2, \partial D_0)$.*

Let $\mathrm{Diff}_0(X)$ denote the group of diffeomorphisms of the compact manifold X that are properly homotopic to the identity.

Corollary 1.5. $\pi_0(\mathrm{Diff}_0(S^2 \times D^2)/\mathrm{Diff}_0(B^4)) = 1$.

Remark 1.6. In words, modulo diffeomorphisms of the 4-ball, homotopy implies isotopy for diffeomorphisms of $S^2 \times D^2$.

The classical *light bulb* theorem states that a knot in $S^2 \times S^1$ that intersects a $S^2 \times y$ transversely and in exactly one point is isotopic to the standard vertical curve, i.e. a $x \times S^1$. The next result is the 4-dimensional version.

Theorem 1.7. *(4D-Lightbulb Theorem) If R is an embedded 2-sphere in $S^2 \times S^2$, homologous to $x_0 \times S^2$, that intersects $S^2 \times y_0$ transversely and only at the point (x_0, y_0), then R is isotopic to $x_0 \times S^2$ via an isotopy fixing $S^2 \times y_0$ pointwise.*

In 1985, under the above hypotheses, Litherland [Li] proved that there exists a diffeomorphism *pseudo-isotopic* to the identity that takes R to $x_0 \times S^2$ and proved the full light bulb theorem for smooth m-spheres in $S^2 \times S^m$ for $m > 2$. Another version of the light bulb theorem was proven in 1986 by Marumoto [Ma]. He showed that two locally flat PL m-discs in an n-sphere, $n > m$ with the same boundary are topologically isotopic rel boundary. Here we prove that theorem for discs in S^4 in the smooth isotopy category.

Theorem 1.8. *(Uniqueness of Spanning Discs) If D_0 and D_1 are discs in S^4 such that $\partial D_0 = \partial D_1 = \gamma$, then there exists an isotopy of S^4 taking D_0 to D_1 that fixes γ pointwise.*

Remark 1.9. The analogous result for 1-discs in S^4 is well known using general position. The result for 3-discs in S^4 implies the smooth 4D-Schoenflies conjecture.

This paper gives two proofs of the 4D-Light Bulb Theorem. The first proof has two steps. First we give a direct argument showing that R is isotopic to a *vertical sphere*, i.e. viewing $S^2 \times S^2$ as $S^2 \times S^1 \times [-\infty, \infty]$ where each $z \times S^1 \times \infty$ and each $z \times S^1 \times -\infty$ is identified with points, then R is transverse to each $S^2 \times S^1 \times t$ and intersects each such space in a single component. This involves an analogue of the normal form theorem of [KSS] and repeated use of $S^2 \times 0$ as a transverse 2-sphere. The second step invokes Hatcher's [Ha] theorem (the Smale conjecture) to straighten out these intersections.

The proof of Theorem 1.2, and hence a somewhat different one for $S^2 \times S^2$ makes use of Smale's results on regular homotopy of 2-spheres in 4-manifolds [Sm1]. We show that if R_0 is homotopic to R_1 and both are embedded surfaces, then the homotopy from R_0 to R_1 is *shadowed by tubed surfaces*, i.e. there is an isotopy taking R_0 to something that looks like R_1 embroidered with a complicated system of tubes. Through various geometric arguments we show that these tubes can be reorganized and eventually isotoped away. The proof formally relies on the first proof of the Light Bulb theorem at the very last step, though we outline how to eliminate the dependence in Remarks 8.3.

Both arguments make extended use of the 4D-Light Bulb Lemma, which is the direct analogue of the 3D-version where one can do crossing a change using the transverse sphere.

More is known in other settings. In the topological category a locally flat 2-sphere in S^4 is topologically equivalent to the trivial 2-knot if and only if its complement has fundamental group $\mathbb{Z}$ [Fr], [FQ]. There are topologically isotopic smooth 2-spheres in 4-manifolds that are not smoothly isotopic, yet become smoothly isotopic after a stabilization with a single $S^2 \times S^2$ [AKMR], [Ak]. Topologically isotopic smooth 2-spheres in simply connected 4-manifolds are smoothly pseudo-isotopic by [Kr] and after finitely many stabilizations with $S^2 \times S^2$'s are smoothly isotopic by [Qu].

The paper is organized as follows. §2 recalls some classical uses of transverse spheres and proves the Light Bulb Lemma. The Light Bulb theorem is proven in §3 as well as a generalization. Basic facts about regular homotopy are recalled in §4. The definition of tubed surface, basic operations on tubed surfaces and notions of shadowing a homotopy by such are given in §5. Tubed surfaces may have *single* and *double* tubes. In §6 it is shown how to transform pairs of double tubes into pairs of single tubes such that in the end all but at most one of the double tubes remains and that one is homotopically inessential. This uses the no $\mathbb{Z}_2$-torsion condition. The reader is cautioned that *tubes* are used in two contexts here, as tubes that follow curves lying in the surface and tubes that follow arcs with endpoints in the surface. The latter are dealt with in §6. A crossing change lemma for the former is proven in §7 enabling distinct tubes to be disentangled. The proof of Theorem 1.2 is completed in §8 by showing how to unknot a given tube. An extension to higher genus surfaces is given in §9. In particular it is shown that a closed surface in $S^2 \times S^2$ homologous to $0 \times S^2$, that intersects $0 \times S^2$ transversely in one point, is standard. Applications and questions are given in §10.

Acknowledgements 1.10. This work was carried out in part while attending many annual Dublin topology workshops and during preceding visits to Trinity College, Dublin. We thank Martin Bridgeman for making this possible. Also, this work was partially carried out while the author was a member of the Institute for Advanced Study. We thank Bob Edwards for his many constructive comments and for his interest. We thank Abby Thompson for asking about higher genus surfaces and Remark 3.3.

2. The 4-dimensional light bulb lemma

Unless said otherwise, all manifolds in this paper are smooth and orientable and immersions are self-transverse.

Definition 2.1. A *transverse sphere* G to the immersed surface R is a sphere with trivial normal bundle that intersects R transversely in a single point.

All transverse spheres in this paper are embedded. The following is well known. We give the proof as a warm up to the light bulb trick.

Lemma 2.2. *If R is an immersed surface with embedded transverse sphere G in the 4-manifold M, then the induced map $\pi_1(M \setminus R) \to \pi_1(M)$ is an isomorphism. If R is a sphere, then the induced maps $\pi_1(M \setminus R \cup G) \to \pi_1(M)$ and $\pi_1(M \setminus G) \to \pi_1(M)$ are isomorphisms.*

Proof. Surjectivity is immediate by general position. If γ is a loop in $M \setminus R$ bounding the singular disc $D \subset M$, then after a small perturbation we can assume that D is transverse to R. Tubing off intersections with copies of G shows that the map is also injective. For the second and third cases, we can assume that D is transverse to $R \cup G$. First use R to tube off intersections of D with G. This proves injectivity for the third case. (If R does not have a trivial normal bundle or is not embedded, then the resulting disc may have extra intersections with R.) Tubing with G eliminates all the $D \cap R$ intersections and so the induced map in the second case is also injective. $\square$

Lemma 2.3. *(4D-Light Bulb Lemma) Let R be an embedded surface with transverse sphere G in the 4-manifold M and let $z = R \cap G$. Let α_0 and α_1 be two smooth compact arcs that coincide near their endpoints and bound the pinched embedded disc E that is transverse to R with $R \cap E = y$ and $E \cap G = \emptyset$. Let f_t be an ambient isotopy of M taking α_0 to α_1 that corresponds to sweeping α_0 across E. Here f_t is fixed near $\partial\alpha_0$ and is supported in a small neighborhood of E. Suppose that $N(\alpha_0)$ is parametrized as $B^3 \times I$ and $R \cap N(E) = C \cup B$ where C is the disc containing y and $B \subset \mathrm{int}(B^3) \times I$. If y and z lie in the same component of $R \setminus B$, then R is ambiently isotopic to $g(R)$ where $g|R \setminus B = \mathrm{id}$ and $g|B = f_1|B$. The ambient isotopy fixes G pointwise and the isotopy restricted to R is supported in B.*

If G has a non trivial normal bundle with even Euler class, then the conclusion holds except for the assertion that the ambient isotopy fixes G.

If the Euler class is odd, then under the additional hypothesis that B is a union of unknotted and unlinked annuli parallel to α_0, the above conclusion holds with the additional modification $g(B) = f_1(B)$.

Remarks 2.4. i) After an initial isotopy of R supported near $N(\alpha_0)$ we can assume that it is of the form $L \times I$ where L is a link in $\mathrm{int}(B^3) \times 0$. In applications in this paper, L is the unlink.

ii) The hypothesis does not hold if B separates y from z in R.

iii) In the Euler class odd case, under the original hypothesis, our argument will conclude that $g|N(\alpha_0)$ is the composition of the standard isotopy taking $N(\alpha_0)$ to $N(\alpha_1)$ followed by the non trivial element of $SO(3)$ along $N(\alpha_1)$.

Proof. Since y lies in the same component of z we can tube off E with a copy of G to obtain a disc D that coincides with E near ∂E and $(D \setminus E) \cap (R \cup G) = \emptyset$ and $D \cap C = \emptyset$. Since G has a trivial normal bundle, there exists a framing of the normal bundle of D that coincides with that of E near ∂E. See Figure 2.1. Therefore, we can isotope B to $f_1(B)$ by sweeping across D rather than E.

When G has a nontrivial normal bundle with Euler class n, then $|D \cap G| = n$ and so the ambient isotopy taking $N(\alpha_0)$ to $N(\alpha_1)$ does not fix G. This isotopy is the composition of the standard one followed by n full twists along $N(\alpha_1)$. Since $\pi_1(SO(3)) = \mathbb{Z}_2$, the twisting can be isotopically undone when n is even. When n is odd the twisting can be isotoped to a single full twist. If the tubes in B are unknotted and unlinked, then they can be isotoped so that $g(B) = f_1(B)$ where g differs from f_1 by a Dehn twist along each of the tubes. $\square$

Remark 2.5. The ambient isotopy is supported in a neighborhood of D.

3. Proof of the light bulb theorem for $S^2 \times S^2$

Definition 3.1. A *light bulb* in $S^2 \times S^2$ is a smooth 2-sphere transverse to a $S^2 \times y_0$ and intersects $S^2 \times y_0$ in a single point. View $S^2 \times S^2$ as a quotient of $S^2 \times (S^1 \times [-\infty, \infty])$ where each $x \times S^1 \times -\infty$ and $x \times S^1 \times \infty$ are identified with points and y_0 is identified with $(z_0, 0) \in S^1 \times [-\infty, \infty]$. We say that the light bulb R is *vertical* if it is transverse to each

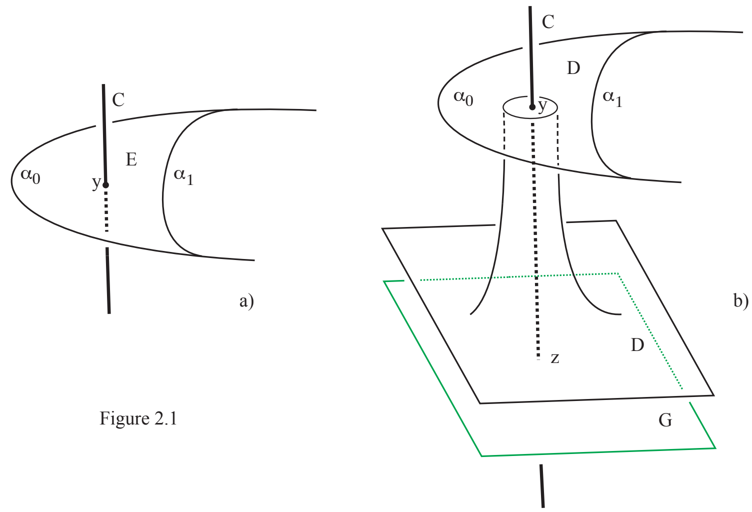


Figure 2.1

$S^2 \times S^1 \times u$, for $u \in [-\infty \times \infty]$. Let G^{std} denote the sphere $S^2 \times z_0 \times 0$ and R^{std} denote the sphere $x_0 \times S^1 \times [-\infty, \infty] \subset S^2 \times S^2$.

To prove the light bulb theorem it suffices to assume that R and R^{std} coincide in some neighborhood U of $(x_0, z_0, 0)$.

Step 1. The light bulb R is isotopic to a vertical light bulb by an isotopy fixing a neighborhood of G^{std} pointwise.

Step 1A. We can assume that R coincides with R^{std} within $S^2 \times (z_0 - \epsilon, z_0 + \epsilon) \times [-\infty, \infty] \cup (S^2 \times S^1 \times [-\infty, 10)) \cup (S^2 \times S^1 \times (10, \infty]$.

Proof. This follows from the fact that R intersects a neighborhood of $S^2 \times y_0 \times 0$ as does R^{std} and a small regular neighborhood of G^{std} is naturally ambiently isotopic to $S^2 \times (z_0 - \epsilon, z_0 + \epsilon) \times [-\infty, \infty] \cup (S^2 \times S^1 \times [-\infty, 10)) \cup S^2 \times S^1 \times (10, \infty])$. $\square$

From now on we will take U to be the neighborhood of $(S^2, z_0, 0)$ given in the statement of Step 1A. Note that U is the complement of $S^2 \times [z_0 + \epsilon, z_0 - \epsilon] \times [-10, 10]$, where $S^1 = [z_0 + \epsilon, z_0 - \epsilon] \cup (z_0 - \epsilon, z_0 + \epsilon)$.

Step 1B. Via an isotopy fixing $R \cap U$, R can be isotoped to be transverse to each $S^2 \times S^1 \times u$ except for $u = -9, -6, 6, 9$. As u increases, p local minimal (with respect to u) appear at $u = -9$, p saddles appear at $u = -6$, $R \cap S^2 \times S^1 \times u$ is connected for $u \in (-6, 6)$, q saddles appear when $u = 6$ and q local maxima appear when $u = 9$.

Proof. This is essentially the normal form of [KSS]. Here is a brief outline. In the usual manner R can be isotoped so that it is transverse to each $S^2 \times S^1 \times u$ except for $u = -9, 0, 9$ where local minimal, saddles, local maxima respectively appear. Up to smoothing of corners, the local minimal (resp. maxima) correspond to the appearance of discs and the saddles correspond to the appearance of bands. After further isotopy we can assume that the bands are disjoint from each other, so for δ small, $R \cap S^2 \times S^1 \times \delta$ is the result of doing band sums to $R \cap S^2 \times S^1 \times -\delta$.

If p (resp. q) is the number of local minima (resp. maxima), then since $\chi(R) = 2$ the total number of saddles is $p + q$. Since R is connected there exist p bands such that the result of only doing band sums along these bands yields a connected curve. Push these bands to $S^2 \times S^1 \times -6$ and push the remaining bands to level to $S^2 \times S^1 \times 6$. $\square$

In what follows we let C_u denote the *core curve* i.e. the component of $R \cap S^2 \times S^1 \times u$ which is transverse to $S^2 \times z_0 \times u$. We abuse terminology by calling a *core curve* such a curve C without specifying u. After a further isotopy we can assume that all the bands at $u = -6$ have one end that attaches to the core curve.

In summary, up to smoothing corners, we can assume that $R \cap S^2 \times S^1 \times [-10, -5]$ appears as follows. For $u \in [-10, -9)$, $R \cap S^2 \times S^1 \times u$ is the standard core curve $x_0 \times S^1 \times u$. At $u = -9$, discs $D_1, \cdots, D_p$ appear. Let $c_1, \cdots, c_p$ denote their boundary curves. The surface $R \cap S^2 \times S^1 \times (-9, -6)$ is the product $(C \cup c_1 \cup \cdots \cup c_p) \times (-9, -6)$. Here we again abuse notation by denoting a c_i without specifying its u level. At $u = -6, p$ bands $b_1, \cdots, b_p$ appear where b_i connects C and c_i. Again $R \cap S^2 \times S^1 \times (-6, -5]$ is a product where each u section is parallel to $R \cap S^2 \times S^1 \times -6$ with the relative interiors of the bands removed.

By a vertical isotopy push the bands $b_2, \cdots, b_p$ up to level -5 and the disc D_1 to level -8. Let $\pi : S^2 \times S^1 \times [-\infty, \infty] \to S^2 \times S^1$ be the projection. To complete the proof of Step 1 we will show that after isotopy $\pi(b_1) \cap \pi(D_1) \subset \pi(\partial D_1)$. It follows that b_1 can be pushed to level -8 its critical point can be cancelled with the one corresponding to D_1. Step 1 then follows by induction and the usual turing upside down argument to cancel the saddles at $u = 6$ with the maxima at $u = 9$.

Step 1C. There exist pairwise disjoint discs $E_1, \cdots, E_p \subset S^2 \times S^1 \times -6$ spanning $c_1, \cdots c_p$ such that for all i, $\pi(\text{int}(E_i)) \cap \pi(b_1) = \emptyset$ and $E_i \cap C \cup U = \emptyset$.

Proof. To start with, for $i = 1, \cdots, p$, let $E_i = D_i$. A given E_i projects to one intersecting $\pi(b_1)$ in finitely many interior arcs. View b_1 as a band starting at c_1 and sequentially hitting the various E_i's before attaching to C. Again we abuse notation by suppressing the fact that we should be talking about projections. Starting at the last intersection of b_1 with an E_i, sequentially isotope the E_i's to remove arcs of intersection at the cost of creating two points of intersection of an E_i with C. Next by following C but avoiding the arc $b_1 \cap C$ tube off these intersections with parallel copies of $S^2 \times z_0$ to obtain the desired set of discs which we still call $E_1, \cdots E_p$. See Figure 3.1 $\square$

Remark 3.2. For the purposes of visualization one can ambiantly isotope R via level preserving isotopy supported in $S^2 \times S^1 \times [-9.5, -5.5]$ so that the discs E_i become small and round and b_1 becomes a straight band connecting C and c_1, which is disjoint from the interior of the E_i's. Furthermore, up to rounding corners, (possibly complicated) discs $D_2, \cdots, D_p$ appear at level -9, D_1 appears at level -8 and vertical annuli $c_1 \times [-8, -6], c_2 \times [-9, -6], \cdots, c_p \times [-9, -6]$ emanate from the ∂D_i's. At level -6 R appears as in the first sentence.

Step 1D. We can assume that $\pi(D_1) \cap \pi(E_i) = \emptyset$ for $i > 1$.

Proof. Let π_{-8} denote the projection of $S^2 \times S^1 \times -6$ to $S^2 \times S^1 \times -8$ fixing the first two factors. By construction $\partial D_1 = c_1$ and D_1 is disjoint from $C \cup U$ as well as the c_i's for $i > 1$. Assuming that D_1 is transverse to the $\pi_{-8}(\text{int}\, E_i)$'s, $i = 1, \cdots, p$, it follows that for all i, $D_1 \cap \pi_{-8}(\text{int}(E_i))$ is a union of pairwise disjoint circles. Starting at the innermost ones in the various $\pi_{-8}(E_i)$'s, compress D_1 to obtain the 2-spheres $S_1, \cdots, S_k$ and a single disc E with

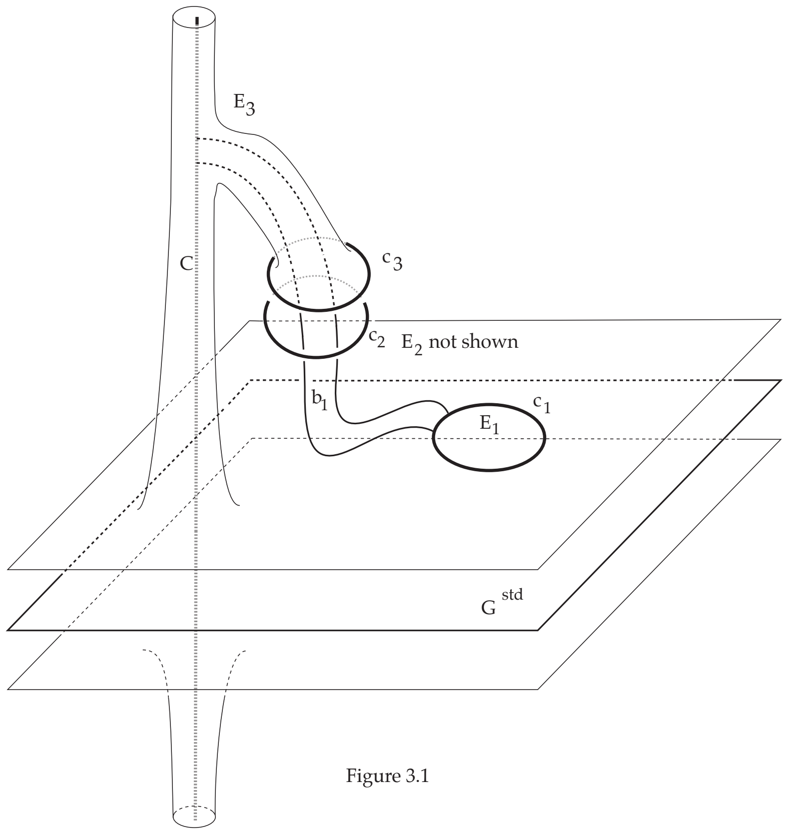


Figure 3.1

$\partial E = c_1$. After possibly sliding tubes that connect distinct S_i's, it follows that D_1 is isotopic, in $S^2 \times S^1 \times -8 \setminus (C \cup U)$, to the surface obtained by starting with the disc E and tubing it to the spheres $S_1, \cdots, S_k$. By construction $\pi(E \cup S_1 \cup \cdots \cup S_k) \cap \pi(E_1 \cup \cdots \cup E_p) = \pi(c_1)$, $(D_1 \cup E \cup S_1 \cup \cdots \cup S_k) \cap (C \cup U) = \emptyset$ and E and the S_i's are pairwise disjoint.

Using the fact that $R \setminus D_1$ is connected and intersects G^{std} transversely once it follows from the Light Bulb Lemma 2.3 that D_1 is isotopic to the surface obtained by homotoping the tubes that intersect the various $\pi_{-8}(E_i)$'s, straight off of these projections when $i > 1$. See Figure 3.2. $\square$

To summarize the situation at the moment: At level -9 discs $D_2, \cdots, D_p$ appear, at level -8 disc D_1 appears, vertical annuli $c_1 \times [-8, -6], c_2 \times [-9, -6], \cdots, c_p \times [-9, -6]$ emanate from the ∂D_i's, a band connects the core to $c_1 \times -6$ and by Step 1D, for $i > 1$, $\pi(E_i) \cap \pi(D_1) = \emptyset$

We can therefore isotope $D_2, \cdots, D_p$ so that $c_2, \cdots, c_p$ are *far away* from D_1, E_1 and b_1. This means that $\pi(c_2), \cdots, \pi(c_p)$ lie in a 3-ball $B \subset S^2 \times S^1$ that intersects C in a connected unknotted arc and B is disjoint from $\pi(D_1), \pi(E_1)$ and $\pi(b_1)$.

Step 1E. Cancel the critical points corresponding to D_1 and b_1 without introducing new ones, thereby completing Step 1.

Proof. Note that $(S^2 \times S^1 \times -6) \setminus (C \cup B \cup S^2 \times z_0 \times -6)$ is diffeomorpic to $\mathbb{R}^3$. Therefore, the discs $\pi(D_1)$ and $\pi(E_1)$ are isotopic rel c_1 via an isotopy disjoint from $\pi(C \cup B) \cup (S^2 \times z_0)$ since spanning discs for the unknot in $\mathbb{R}^3$ are unique up to isotopy. After the corresponding isotopy of D_1, supported in $S^2 \times S^1 \times -8$ it follows from Remark 3.2 that $\pi(b_1) \cap \pi(D_1) \subset \pi(\partial D_1)$.

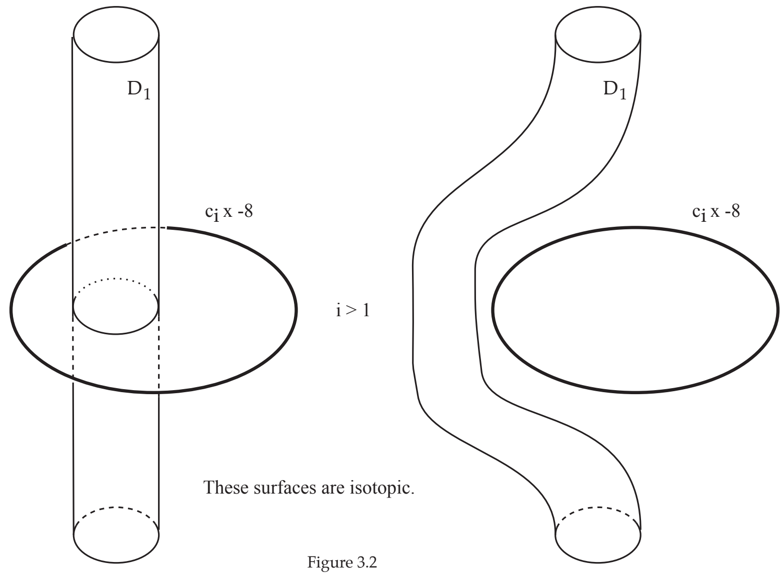


Figure 3.2

Therefore, b_1 can be pushed down to level -8, hence the critical points corresponding to b_1 and D_1 can be cancelled. $\square$.

Remark 3.3. Abby Thompson pointed out that the above arguments work for higher genus surfaces to eliminate critical points of index 0 and 2. So if genus$(R) = g$, then R can be isotoped so that $2g$ bands appear at $u = -6$ and there are no other critical levels.

From the point of view of $S^2 \times -6$ these bands can be twisted and linked. See §9.

Step 2. A vertical light bulb R homologous to R^{std} that agrees with R^{std} near G^{std} is isotopic to R^{std} via an isotopy fixing a neighborhood of G^{std} pointwise.

Remarks 3.4. 1) It is easy to construct a vertical lightbulb homologous to $[R^{std}] + n[S^2 \times z_0 \times 0]$ by first starting with R^{std}, removing a neighborhood of $(x_0, z_1, 0)$ and replacing it by one that *sweeps across* $S^2 \times z_1$ n times while $u \in (-\epsilon, \epsilon)$ where $z_1 \neq z_0$.

Proof The proof of Step 1 shows that we can assume that R coincides with R^{std} away from $S^2 \times S^1 \times [-10, 10]$ further it is coincides with R^{std} near $S^2 \times z_0 \times [-\infty, \infty]$. Thus R is standard outside a submanifold W of the form $S^2 \times [0, 1] \times [-10, 10]$ and within W corresponds to a smooth path of embedded smooth paths $\rho_t : D^1 \to S^2 \times I$ for $t \in [-10, 10]$, where $\rho_{-10}(D^1) = \rho_{10}(D^1) = (x_0, z_0, I)$ and ρ_t is fixed near the endpoints of D^1. By identifying D^1 with (x_0, z_0, I) we can assume that $\rho_{-10} = \rho_{10} = \text{id}$. Note that R^{std} corresponds to the identity path.

By the covering isotopy theorem, ρ_t, extends to a path $\phi_t \in \text{Diff}(S^2 \times I, \text{rel}(\partial S^2 \times I))$ with $\phi_0 = \text{id}$. We first show that such a path can be chosen so that $\phi_{10} = \text{id}$. By uniqueness of regular neighborhoods we can first assume that restricted to some D^2 neighborhood of x_0, in polar coordinates, $\rho_{10}(r, \theta, s) = (r, \theta + h(s)2\pi, s)$ for some $h : [0, 1] \to \mathbb{R}$ with $h(0) = 0$. Since $[R] = [R^{std}] \in H_2(S^2 \times S^2)$ it follows that $h(1) = 0$ and hence after further isotopy, that $\rho_{10}|D^2 \times I = \text{id}$. Since $S^2 \times I \setminus (\text{int}(D^2) \times I) = B^3$, we can assume that $\rho_{10} = \text{id}$, by [Ce2] or [Ha].

Thus ρ_t is a closed loop in $\text{Diff}(S^2 \times I, \text{rel}(\partial S^2 \times I))$ which by Hatcher [Ha] is homotopically trivial since $\pi_1(\Omega(O(3)) = \pi_2(O(3)) = \pi_2(\mathbb{R}(P^3)) = 0$. Here we are using formulation (8) (see the appendix of [Ha]) of Hatcher's theorem which asserts that $\text{Diff}(D^1 \times S^2 \operatorname{rel} \partial)$ is homotopy equivalent to $\Omega(O(3))$. Restricting this homotopy to ρ_t gives the desired isotopy of ρ_t to id. $\square$

Conjecture 3.5. *The space of light bulbs is not simply connected.*

We now extend Theorem 1.7 to multi-light bulbs in $\#_k S^2 \times S^2$

Definition 3.6. Let W_k denote $\#_k S^2 \times S^2$and let $(S^2 \times S^2)_i$ denote the i'th summand. Within $(S^2 \times S^2)_i$, let G_i^{std} (resp. R_i^{std}) denote $S^2 \times y_0$ (resp. $x_0 \times S^2$). Define a *multi-light bulb* $\mathcal{R} = \{R_1, \cdots, R_k\}$ to be a set of 2-spheres in Y_k such that $|R_i \cap G_j^{std}| = \delta_{ij}$ where the intersections are transverse.

Theorem 3.7. *If $\mathcal{R} = \{R_1, \cdots, R_k\}$ is a multi-light bulb in $W_k = \#_k S^2 \times S^2$ such that for each i, $[R_i] = [R_i^{std}] \in H_2(\#_k S^2 \times S^2)$, then $\mathcal{R}$ is isotopic to $\{R_1^{std}, \cdots, R_k^{std}\}$ via an isotopy fixing $G_1^{std} \cup \cdots \cup G_k^{std}$ pointwise.*

Proof. Let $\mathcal{G}^{std} = G_1^{std} \cup \cdots \cup G_k^{std}$ and $\mathcal{R}^{std} = R_1^{std} \cup \cdots \cup R_k^{std}$. We will show that R_1 can be simultaneously isotoped off of each R_j^{std}, $j > 1$ via an isotopy fixing $\mathcal{G}^{std}$. If so, then since $(S^2 \times S^2)_i \setminus R_i^{std} \cup G_i^{std} = \mathbb{R}^4$, it follows that R_1 can be isotoped into $(S^2 \times S^2)_1$ again by an isotopy fixing $\mathcal{G}^{std}$. Next apply Theorem 1.7 to isotope R_1 to R_1^{std}. Note that this theorem applies since the S^3 which splits off $(S^2 \times S^2)_1$ can readily be avoided. By isotopy extension, this isotopy can be done ambiently. Finally, isotope $\mathcal{R} \setminus R_1^{std}$ out of $(S^2 \times S^2)_1$ via an isotopy fixing $R_1^{std} \cup \mathcal{G}^{std}$ pointwise. The result then follows by induction on k.

Observe that $Z = W_k \setminus \mathcal{G}^{std} \cup R_1 \cup R_2^{std} \cup \cdots \cup R_k^{std}$ is simply connected. Indeed if $\gamma \subset$ is a closed curve in Z, then since $W_k \setminus \mathcal{G}^{std} \cup R_2^{std} \cup \cdots \cup R_k^{std}$ is simply connected, $\gamma = \partial E$ where E is an immersed disc with $\gamma \subset Z \cap R_1$. Since G_1^{std} is a transverse sphere to R_1 we can also assume that $E \cap R_1 = \emptyset$.

We now show that R_1 can be isotoped off of the R_j^{std}'s, $j > 0$. Since $[R_1] = [R_1^{std}]$, R_1 intersects each R_i^{std} algebraically zero times. Suppose $R_1 \cap Rs_j \neq \emptyset$. In the usual manner e.g. see [FQ] or [E] we can find a Whitney disc w between them such that $\text{int}(w) \subset Z$. Note that any excess intersections of w with R_j^{std} created in the process of fixing the framing or desingularizing w can be eliminated by tubing with parallel copies of G_j^{std}. Use w to eliminate a pair of intersections between R_1 and R_j^{std}. Thus all the intersections between R_1 and the R_j^{std}'s can be eliminated using Whitney discs. $\square$

4. Regular homotopy of embedded spheres in 4-manifolds

The main result of this section is essentially Theorem D of [Sm1].

Theorem 4.1. *(Smale (1957)) Two smooth embedded spheres in an orientable 4-manifold are regularly homotopic if and only if they are homotopic.*

Definition 4.2. Let S be a smooth immersed self transverse surface in the smooth 4-manifold Z. A *finger move* is the operation of regularly homotoping a disc in S along an embedded arc to create a pair of new transverse self intersections. A *Whitney move* is a regular homotopy supported in a neighborhood of a Whitney disc to eliminate a pair of self intersections. By an *isotopy* of S we mean a regular homotopy through self transverse surfaces. In particular, no new self intersections are either created or cancelled.

The next well-known proposition follows by considering a generic regular homotopy and the usual ordering by index argument.

Proposition 4.3. *Let A and B be smooth embedded surfaces in the smooth 4-manifold Z. If A is regularly homotopic to B, then up to isotopy, the regular homotopy can be expressed as the composition of finitely many* finger *moves followed by finitely many* Whitney *moves.* □

Remarks 4.4. i) Upon reversing the points of view of B and A the Whitney moves become finger moves and vice versa.

ii) If $\pi_1(Z) = 1$, then any two finger paths are homotopic and hence isotopic.

Corollary 4.5. *If A and B are regularly homotopic smooth embedded surfaces in the smooth 4-manifold Z, then there exists an immersed self transverse surface C and systems $\mathcal{F}$ of finger discs and $\mathcal{W}$ of Whitney discs so that Whitney moves applied to the finger (resp. Whitney) discs transforms C to A (resp. B).*

5. Shadowing regular homotopies by tubed surfaces

In this section we show that if $f_0 : A_0 \to M$ is an embedding of a smooth surface with embedded transverse sphere G into a smooth 4-manifold and $f_t : A_0 \to M$ is a generic regular homotopy supported away from G, then f_t can be *shadowed* by a *tubed surface*. Roughly speaking there is a smooth isotopy g_t with $g_0(A_0) = f_0(A_0)$ such that for $t > 0$, $g_t(A_0)$ is approximately $f_t(A_0)$ with tubes connecting to copies of G.

This section is motivated by the following lemma. While we have yet to define terms used in its statement the proof should make clear what they mean. The formal definitions of *shadowed* and *tubed surface* are given after the proof and comprise much of this long section.

Lemma 5.1. *Let R be a connected embedded smooth surface in the smooth 4-manifold M. If R has an embedded transverse sphere G, then a finger move on R disjoint from G can be shadowed by tubed surfaces disjoint from G.*

Proof. Let $z = R \cap G$, $x, y \in R \setminus G$ and κ a path from y to x with $\mathrm{int}(\kappa) \cap (R \cup G) = \emptyset$. The finger move associated to κ can be shadowed as follows. Let $\sigma \subset R \setminus y$ be an embedded path from x to z. Let D be a small disc transverse to R with $D \cap R = x$. Let T be the disc disjoint from R which is the union of a tube which starts at ∂D and follows along σ and then attaches to a parallel copy of $G \setminus \mathrm{int}(N(z))$ disjoint from G. Let $T \times I$ be a product neighborhood. The shadow isotopy starts off with the finger approaching x, but instead of crashing through and creating the immersed surface R_1, it isotopes through $T \times I$ to become the surface R_2 which is R_1 with neighborhoods of two points replaced by $T \times 0$ and $T \times 1$ as in Figure 5.1. □

Definition 5.2. A *framed embedded path* is a smooth embedded path $\tau(t), t \in [0,1]$ in the 4-manifold M with a framing $\mathcal{F}(t) = (v_1(t), v_2(t), v_3(t))$ of its normal bundle. Let $(C(0), x(0))$ consist of a smooth embedded circle $C(0)$ with base point $x(0)$ lying in the normal disc to τ through $\tau(0)$ that is spanned by the vectors $(v_1(0), v_2(0))$ with $x(0)$ lying in direction $v_1(0)$. Define $(C(t), x(t))$ a smoothly varying family having similar properties for each $t \in [0,1]$. Call the annulus $(C(t), x(t))$, $t \in [0,1]$ the *cylinder connecting* $C(0)$ and $C(1)$. It should be thought of as lying very close to τ.

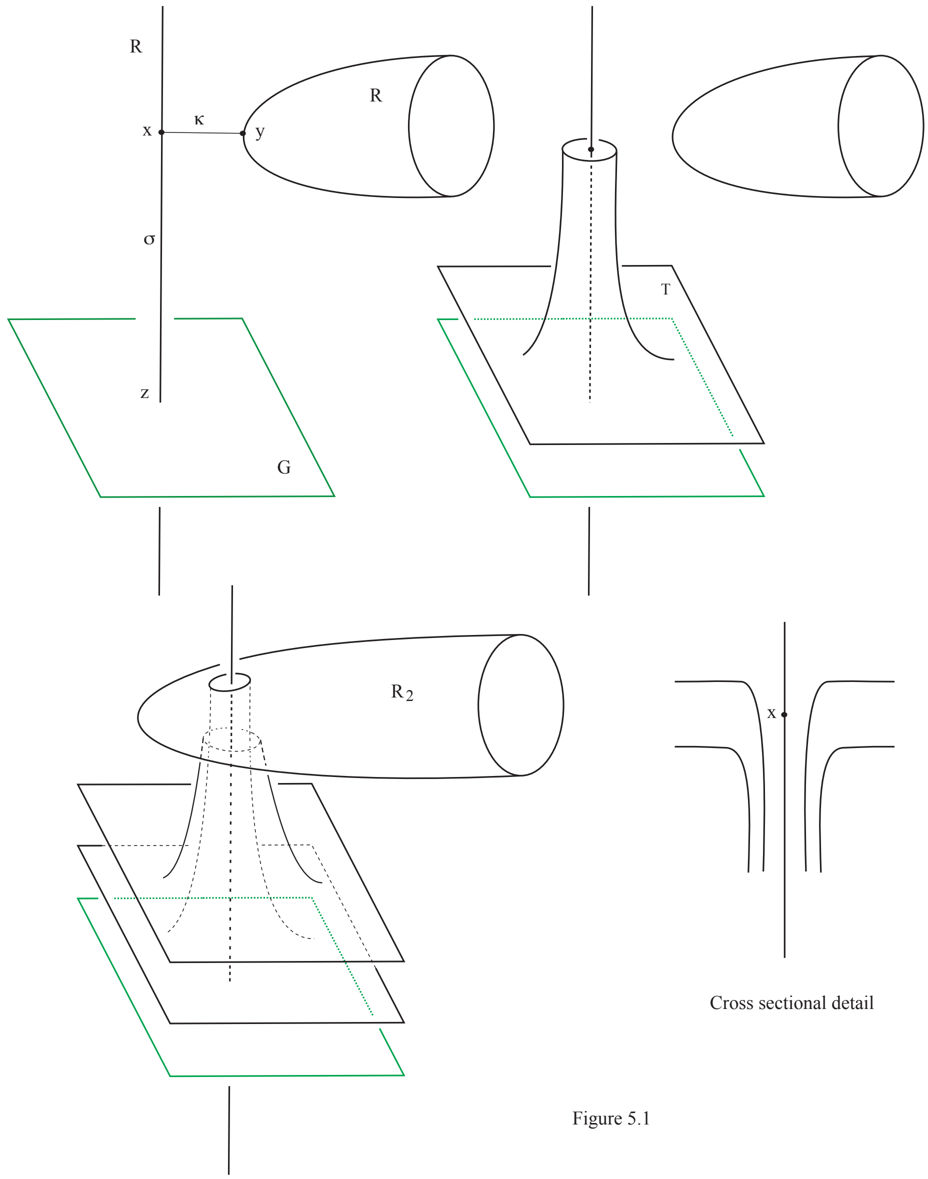


Figure 5.1

Definition 5.3. A *tubed surface* $\mathcal{A}$ in the 4-manifold M consists of

i) a self transverse immersion $f : A_0 \to M$, where A_0 is a closed surface based at z_0 with A_1 denoting $f(A_0)$. The preimages $(x_1, y_1), \cdots, (x_n, y_n)$ of the double points are pairwise ordered. A_0 is called the *underlying* surface and A_1 the *associated* surface to $\mathcal{A}$.

ii) An embedded transverse 2-sphere G to A_1, with $A_1 \cap G = z = f(z_0)$.

iii) For each i, an immersed path $\sigma_i \subset A_0$ from x_i to z_0.

iv) immersed paths $\alpha_1, \cdots, \alpha_r$ in A_0 with both endpoints at z_0 and for each i, pairs of points (p_i, q_i) with $p_i \in \alpha_i$ and $q_i \in A_0$ and a framed embedded path $\tau_i \subset M$ from $f(p_i)$ to $f(q_i)$ with $\mathrm{int}(\tau_i) \cap (G \cup A_1) = \emptyset$.

v) pairs of immersed paths $(\beta_1, \gamma_1), \cdots, (\beta_s, \gamma_s)$ in A_0 where β_i goes from z_0 to b_i and γ_i goes from g_i to z_0 and framed embedded paths $\lambda_i \subset M$ from $f(b_i)$ to $f(g_i)$ with $\mathrm{int}(\lambda_i) \cap (G \cup A_1) = \emptyset$.

Curves of the form $\sigma_i, \alpha_j, \beta_k, \gamma_l$ are called *tube guide* curves and the τ_p and λ_q curves are called *framed tube guide* curves. All such curves are required to be self transverse and transverse to each other with interiors disjoint from the z_0, x_i, q_j, b_k, g_l points and disjoint from the p_j points except where required in iv). At points of intersection and self intersection of these curves, except z_0, one curve is determined to be *above* or *below* the other curve. The various points $z_0, x_i, y_j, p_k, q_l, b_m, g_n$ are all distinct.

The curves τ_i and λ_j are pairwise disjoint, disjoint from G and intersect A_1 only at their endpoints. There are conditions on the framings of these curves that will be given in Definition 5.4. This ends the definition of tubed surface.

Conditions i) - iii) are what's needed to create a tubed surface arising from a finger move as in Lemma 5.1, though in that case σ is embedded. Tubed surfaces will undergo various operations in the course of shadowing a regular homotopy. Conditions iv) and v) are needed to describe tubed surfaces arising from Whitney moves. Crossings of tube guide curves may occur in preparation for Whitney moves and in the process of transforming pairs of double tubes to pairs of single tubes in §6.

We now show how a tubed surface gives rise to an embedded surface.

Definition 5.4. Associated to the tubed surface $\mathcal{A}$ construct an embedded surface A, called the *realization* of $\mathcal{A}$ as follows. For each i, remove from A_1 a small D^2 neighborhood of y_i. Attach to $f(\partial D^2)$ a disc consisting of a tube t_i that follows $f(\sigma_i)$ and connects to a slightly pushed off copy of $G \setminus \mathrm{int}(N(z))$. See Figure 5.2. If $u \in \sigma_i \cap \sigma_j$, $u \neq z_0$, and σ_i lies above σ_j at u, then near $f(u)$, construct t_j to lie closer to A_1 than does t_i. With abuse of notation, this allows for the case $i = j$. See Figure 5.3. Let $\hat{A}$ be the embedded surface thus far constructed.

In similar manner associated to the path α_i is a 2-sphere P_i with $P_i \cap A_1 = \emptyset$, consisting of two pushed off copies of $G \setminus \mathrm{int}(N(z))$ tubed together along the path $f(\alpha_i)$. Next attach a tube $T(\tau_i)$ following the framed embedded path τ_i from $C(0) = P_i \cap \partial N(\tau_i)$ to $C(1) = \hat{A} \cap \partial N(f(q_i))$. Here we assume that τ_i approaches $f(p_i)$ normally to $\hat{A}$ and is parametrized by $[-1/4, 1]$ and framed so that restricting to $[0, 1]$, $C(0)$ (resp. $C(1)$) is in the plane spanned by $v_1(0)$ and $v_2(0)$ (resp. $v_1(1)$ and $v_2(1)$) as in Definition 5.2. This assumption is the condition on the framing on τ_i that is required but not explicitly stated at the end of Definition 5.3. The tube $T(\tau_i)$ is called a *single tube*. See Figure 5.4. Let $\hat{A}'$ the embedded surface constructed at this stage.

Next for each i, construct discs $D(\beta_i)$ and $D(\gamma_i)$ consisting of pushed off copies of $G \setminus \mathrm{int}(N(z))$ tubed very close to and respectively along $f(\beta_i)$ and $f(\gamma_i)$ with boundary lying in discs normal to $\hat{A}'$ at $f(b_i)$ and $f(g_i)$. Roughly speaking the rest of the construction of A from $\hat{A}'$ proceeds as follows. Appropriately sized 4-balls $N(f(b_i))$ and $N(f(g_i))$ have the property that $\partial N(f(b_i))) \cap (\hat{A}' \cup D(\beta_i))$ and $\partial N(f(g_i)) \cap (\hat{A}' \cup D(\gamma_i))$ are Hopf links. Connect these links by tubes that parallel λ_i using the normal framing. Here $\partial N(f(g_i)) \cap \hat{A}'$ (resp. $\partial N(f(b_i)) \cap \hat{A}'$) connects to $\partial N(f(b_i)) \cap D(\beta_i)$ (resp. $\partial N(f(g_i)) \cap D(\gamma_i)$). See Figure 5.5.

More precisely, delete $\mathrm{int}(N(f(b_i)))$ from $D(\beta_i)$ and continue to call $D(\beta_i)$ the disc that remains. Next remove $\mathrm{int}(1/2N(f(b_i)))$ from $\hat{A}'$ and let $C(0) = \partial((1/2N(f(b_i))) \cap \hat{A}')$. Also remove $\mathrm{int}(N(f(g_i)))$ from $\hat{A}'$ and $\mathrm{int}(1/2N(f(g_i)))$ from $D(\gamma_i)$ and call $D(\gamma_i)$ what remains with $C(1) = \partial D(\gamma_i)$. Suppressing the epsilonics, $1/2N(f(g_i))$ and $1/2N(f(b_i))$ are *half*

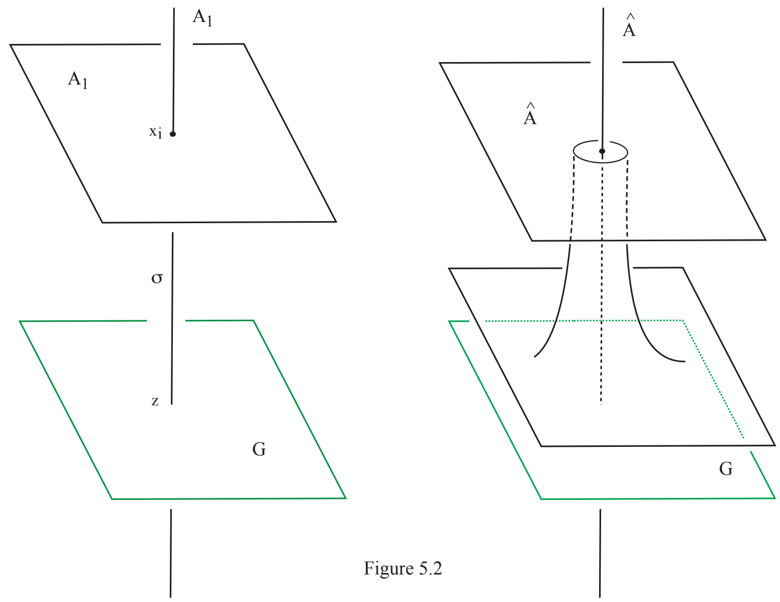
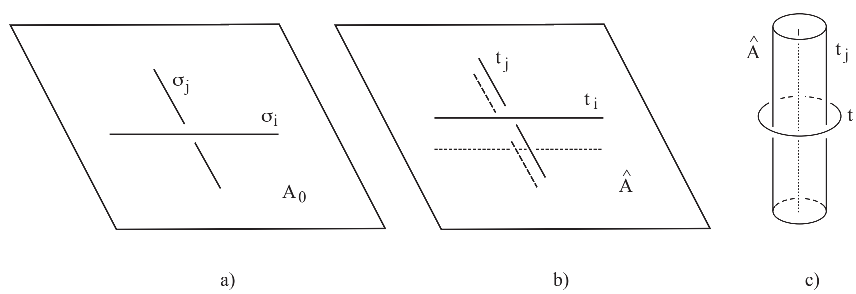


Figure 5.2



Creating tubes associated to σ_i and σ_j. Figures b) and c) are local 3-dimensional slices of $\hat{A}$.

Figure 5.3

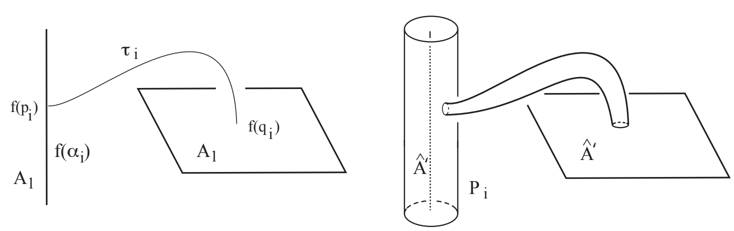


Figure 5.4

radius 4-balls about $f(b_i)$ and $f(g_i)$ and with respect to that scale, the tubes of $D(\beta_i)$ and $D(\gamma_i)$ are very close to $f(\beta_i)$ and $f(\gamma_i)$. See Figure 5.5 b).

We assume that λ_i approaches $\hat{A}'$ in geodesic arcs near $f(g_i)$ and $f(b_i)$, and in the two 3-planes spanned by these arcs and $\hat{A}'$, it approaches $\hat{A}'$ at angle $\pi/3$ and has angle $2\pi/3$ to $f(\gamma_i)$ and to $f(\beta_i)$. We assume that λ_i is parametrized by $[0,1]$ and framed so that $C(0)$ is in the plane spanned by $v_1(0)$ and $v_2(0)$ as in Definition 5.2. Also $x(0) = C(0) \cap \beta_i$ and

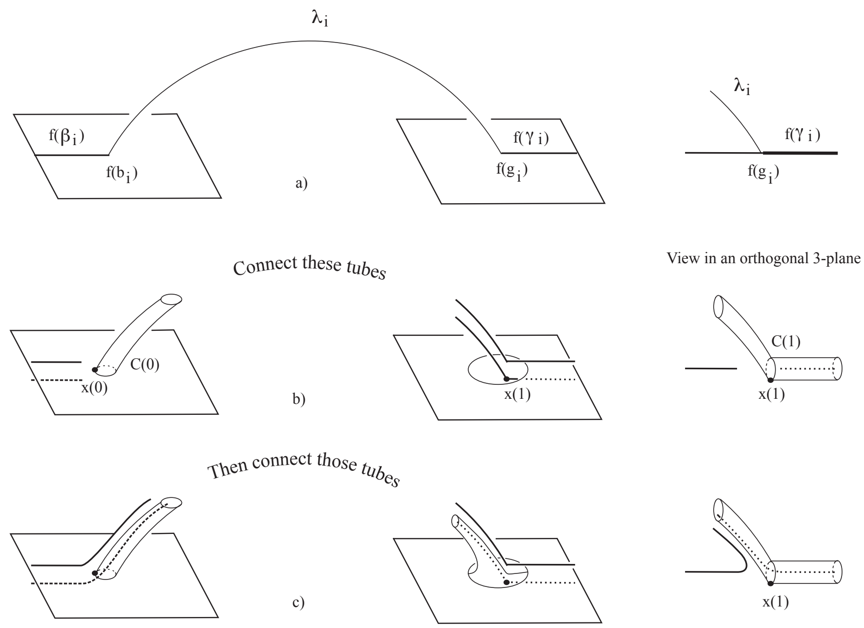


Figure 5.5

$v_1(0)$ points towards $x(0)$. Let $C(1) = \partial D(\gamma_i)$ with $x(1)$ the point indicated in Figure 5.5 b) and assume that $C(1)$ lies in the plane spanned by $v_1(1)$ and $v_2(1)$ with $x(1)$ lying in the arc spanned by $v_1(1)$.

As in Definition 5.2 use λ_i to build a tube connecting $C(0)$ and $C(1)$. Using a tube about the path $x(t), t \in [0, 1]$ connect $\partial D(\beta_i)$ to $\partial N(g_i) \cap \hat{A}'$ as in Figure 5.5 c) to complete the construction of the realization A of $\mathcal{A}$.

We now describe operations on a tubed surface $\mathcal{A}$ that correspond to isotopies of the realizations.

Definition 5.5. We enumerate *tube sliding moves* on a tubed surface $\mathcal{A}$ corresponding to redefining the location and crossing information of tube guide curves in the underlying surface A_0.

i) Type 2), 3) Reidemeister moves. See Figure 5.6 a).

ii) Reordering near z. See Figure 5.6 b).

iii) Sliding across a double point. See Figure 5.6 c). There are two cases depending on whether or not the tube guide κ lies in the sheet through y_i or the sheet through x_i. In the former case we require that $\kappa \neq \sigma_i$.

iv) Sliding across a double tube. See Figure 5.6 d).

v) Sliding across a single tube. Here a tube guide curve $\kappa \neq \alpha_i$ can slide across q_i and over p_i. Any tube guide curve can slide under p_i. See Figure 5.6 e).

Remark 5.6. Sliding σ_i across y_i is analogous to a handle sliding over itself. Similarly for sliding β_i (resp. γ_i, resp. α_i) across g_i (resp. b_i, resp. q_i).

Lemma 5.7. *If $\mathcal{A}$ and $\mathcal{A}'$ are tubed surfaces that differ by tube sliding, then their realizations A and A' are isotopic.*

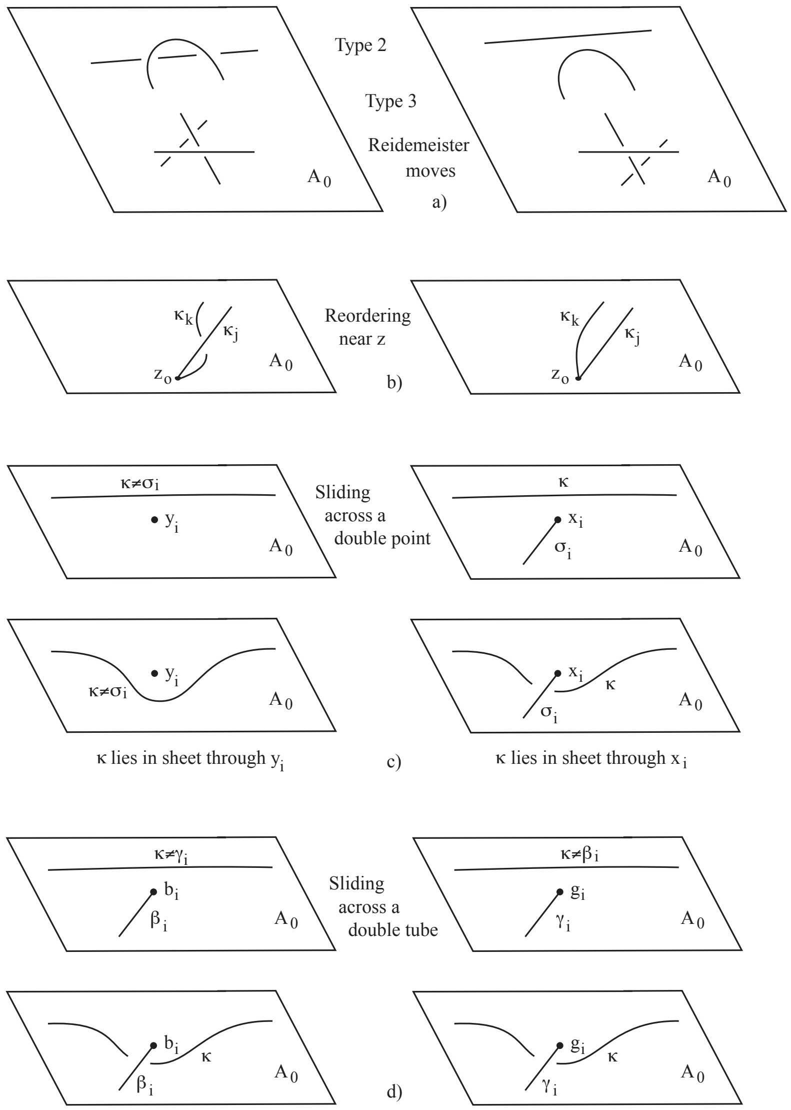
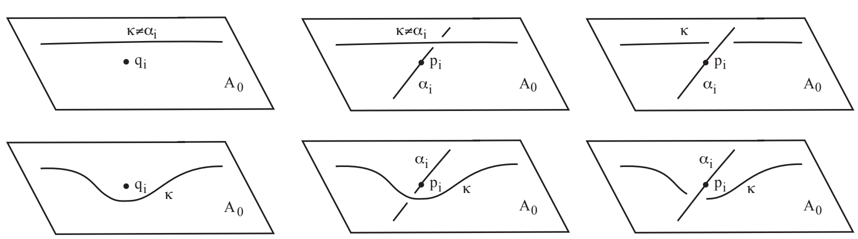


Figure 5.6



Sliding across a single tube

Figure 5.6 e)

Proof. We consider the effect on the realization of $\mathcal{A}$ by the various tube sliding moves. The under/over crossing data in A_0 reflects how close one tube is to A_1 compared with the another. As the Reidemeister 2), 3) moves respect this closeness it follows that they induce an isotopy from A to A'.

Next we consider reordering near z_0. Since G has a trivial normal bundle, there is an S^1 worth of directions that it can push off itself. These directions correspond to the directions that the image of tube guide curve $f(\kappa) \subset A_1$ can approach z. We can assume that the various perturbed copies of $G \setminus \mathrm{int}(N(z))$ are equidistant from G at angle that of the angle of approach of the various $f(\kappa)$'s. Let $D \subset A_1$ denote a disc which is a small neighborhood of the bigon that defines the reordering. Let K_i denote the disc consisting of a perturbed copy G_i of $G \setminus \mathrm{int}(N(z))$ together with its tube that follows the arc $f(\kappa_i) \cap D$. If κ_j is above κ_k as in Figure 5.6 b) and K_j' and K_k' are the discs resulting from the reordering, then there is an isotopy of A to A' supported on K_k where G_k is first pushed radially close to G, then rotated to the angle defined by $f(\kappa_k')$ and then pushed out.

Next, consider sliding across a double point $(f(x_i), f(y_i))$. The corresponding isotopy involves sliding a tube across an embedded disc D. This is a local operation if κ lies in the sheet containing x_i as in Figure 5.6 c). If the sheet contains y_i, then the disc D consists of copy of $G \setminus \mathrm{int}(N(z))$ tubed along $f(\sigma_i)$. If $\kappa \neq \sigma_i$, then the tube about $f(\kappa)$ is disjoint from D. It follows that A and A' are isotopic.

Sliding across a double tube follows similarly. E.g. sliding κ across b_i involves sliding the tube about $f(\kappa)$ across the disc D consisting of a copy of $G \setminus \mathrm{int}(N(z))$ that connects to a tube that first follows $f(\gamma_i)$ and then λ_i. This can be done if $\kappa \neq \gamma_i$.

The same argument shows that a tube guide κ can slide across q_i provided $\kappa \neq \alpha_i$. Sliding under p_i is a local operation. Sliding over p_i requires the light bulb trick to disentangle λ_i from κ which in turn requires that $\kappa \neq \alpha_i$. $\square$

We now define operations on tubed surfaces corresponding to finger and Whitney moves.

Definition 5.8. Let A_1 be the associated surface to the tubed surface $\mathcal{A}$. To a generic finger move from A_1 to A_1' with corresponding regular homotopy of f to f' we obtain a new tubed surface $\mathcal{A}'$ said to be obtained from $\mathcal{A}$ by a *finger move*. By *generic* we mean that the support of the homotopy is away from all the framed tube guide curves and images of tube guide curves of $\mathcal{A}$. $\mathcal{A}'$ will have the same underlying surface A_0 as $\mathcal{A}$ and A_1' will be its associated surface. Let (x_1, y_1) and (x_2, y_2) be the new pairs of f' preimages of double points in A_0, where both x_1 and x_2 lie in the same local sheet of A_1'. Let σ_1 and σ_2 be parallel embedded paths from x_1 and x_2 to z_0 transverse to the existing tube guide paths. The tube guide locus of $\mathcal{A}'$ to be that of $\mathcal{A}$ together with σ_1 and σ_2 where all crossings of these σ_i's with existing tube guide curves are under crossings. See Figure 5.7.

Remark 5.9. There is flexibility in the construction of $\mathcal{A}'$ from $\mathcal{A}$ in the choice of which pair of points are called x_i points and in the choice of the σ_i paths.

Lemma 5.10. *If $\mathcal{A}'$ is obtained from $\mathcal{A}$ by a finger move, then their associated realizations are isotopic.*

Proof. This lemma is a restatement of Lemma 5.1. The proof is the same after recognizing that if we identify A with the surface R in that proof, then A' is isotopic to the surface R_2. $\square$

Definition 5.11. Let A_1 be the associated surface to the tubed surface $\mathcal{A}$. A Whitney move from A_1 to A_1' corresponding to the regular homotopy of f to f' with Whitney disc w is said to be *clean* if $\mathrm{int}(w)$ is disjoint from the framed tube guide curves of $\mathcal{A}$ and ∂w intersects

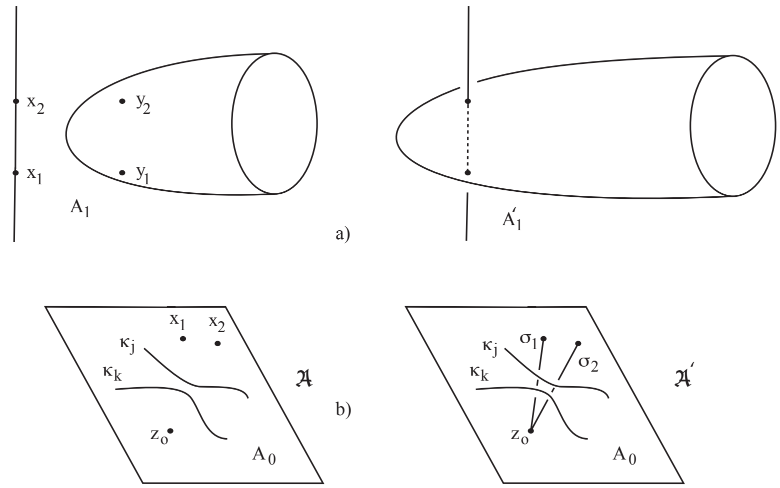


Figure 5.7

the images of tube guide curves of $\mathcal{A}$ only at double points of A_1. Let $(x_1, y_1), (x_2, y_2)$ denote the pairs of points in A_0 corresponding to these double points with notation consistent with that of Definition 5.3. We say that the Whitney move is *uncrossed* if both $f(x_1)$ and $f(x_2)$ lie in the same local sheet of A_1 and *crossed* otherwise. If w is an uncrossed Whitney disc, then we obtain the tubed surface $\mathcal{A}'$ as indicated in Figure 5.8 and if w is crossed, then $\mathcal{A}'$ is obtained as in Figure 5.9. Accordingly $\mathcal{A}'$ is said to be obtained from $\mathcal{A}$ by an uncrossed or crossed Whitney move.

Remark 5.12. An uncrossed Whitney moves gives to a single tube while a crossed Whitney move gives rise to a double tube. In the former case two σ curves become an α curve. In the latter case, the σ curves become β and γ curves.

Lemma 5.13. *If $\mathcal{A}'$ is obtained from $\mathcal{A}$ by a clean Whitney move, then its realization is isotopic to that of $\mathcal{A}$.* □

Definition 5.14. We define an *elementary tubed surface isotopy*, or *elementary isotopy* for short, on the tubed surface $\mathcal{A}$ as any of the following operations on $\mathcal{A}$.

a) The defining data changes smoothly without combinatorial change. In particular, at no time are there new tangencies or new intersections among the various objects.

b) tube sliding moves.

c) finger moves

d) clean Whitney moves

Lemma 5.15. *If $\mathcal{A}$ and $\mathcal{A}'$ are tubed surfaces that differ by an elementary isotopy, then their realizations are isotopic.* □

Definition 5.16. Let R_0 be an immersed surface in the smooth 4-manifold M with embedded transverse sphere G. Let $f_t : R \to M^4$ be a generic regular homotopy supported away from G which is an immersion except at times $\{t_i\}$ where $0 < t_1 < \cdots < t_m < 1$. Let $R_0 = f_0(R), R_1, \cdots, R_m = f_1(R)$ be such that for $i = 1, \cdots, m-1$, R_i is a surface $f_s(R)$ for some $s \in (t_i, t_{i+1})$. We say that the regular homotopy f_t is *shadowed by tubed surfaces* if there exists a sequence $\mathcal{R}_0, \mathcal{R}_1, \cdots, \mathcal{R}_m$ of tubed surfaces such that for all i, R_i is the

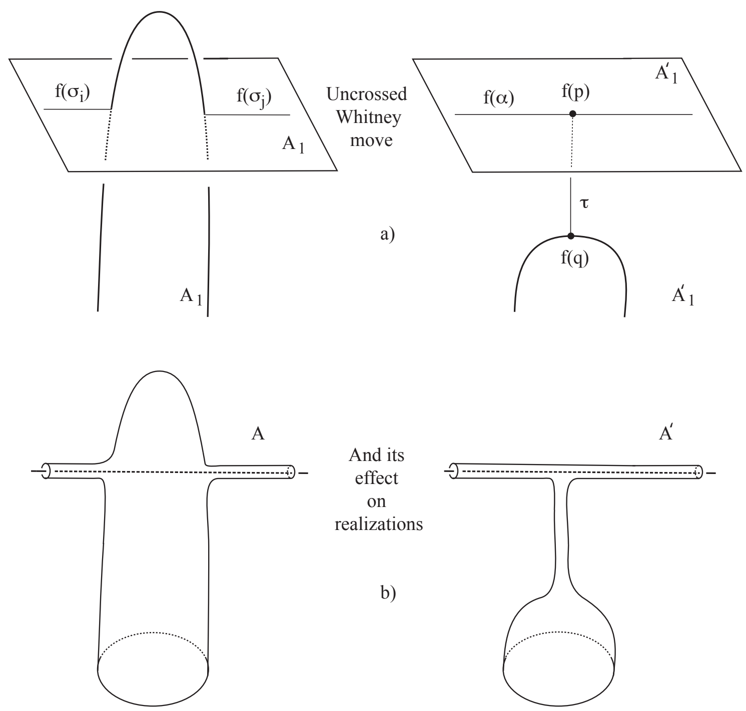


Figure 5.8

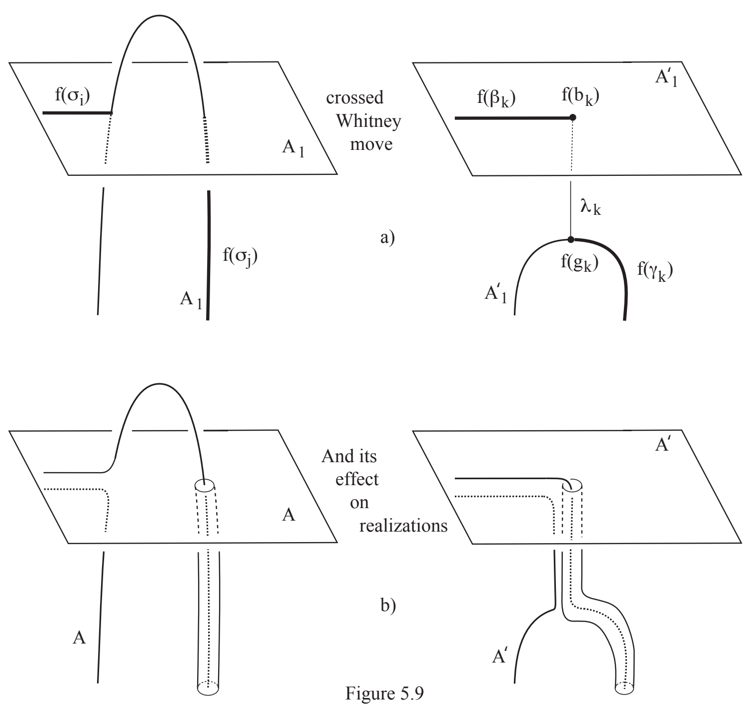


Figure 5.9

associated surface to $\mathcal{R}_i$ and for $i \neq m$, $\mathcal{R}_{i+1}$ is obtained from $\mathcal{R}_i$ by elementary isotopies. The tubed surfaces $\mathcal{R}_0, \mathcal{R}_1, \cdots, \mathcal{R}_m$ are called *f_t-shadow tubed surfaces.*

Theorem 5.17. *If $f_t : A_0 \to M$ is a generic regular homotopy with $f_0(A_0)$ an embedded surface, M a smooth 4-manifold, G a transverse embedded sphere to $f_0(A_0)$ and f_t is supported away from G, then f_t is shadowed by tubed surfaces.*

Proof. Let $0 < t_1 < \cdots < t_m < 1$ be the singular times of f_t. If there are no singular times in $[s, s']$ and a tubed surface $\mathcal{A}_s$ has been constructed whose associated surface $A_1^s = f_s(A_0)$, then an elementary isotopy of type a) transforms it to one with $A_1^{s'} = f_{s'}(A_0)$. Thus we need only show how to shadow regular homotopies near singular times. Now each singular time corresponds to either a finger or Whitney move. Since $f_0(A_0)$ is embedded, t_1 is the time of a finger move. The shadowing of the initial finger move is given in the proof of Lemma 5.10. More generally, that lemma shows how to shadow any generic finger move. By induction we assume that the conclusion holds through time t, where $t_{k-1} < t < t_k$ and t_k is the time of a Whitney move.

Let $\mathcal{A}_k$ denote the tubed surface with underlying surface $A_1^k = f_t(A_0)$. By Lemma 5.15 we can assume that t is a time just before the Whitney move. Let w be a Whitney disc for the Whitney move. Being 2-dimensional we can assume that w is disjoint from the framed tube guide curves to $\mathcal{A}_k$. Suppose that w cancels the f_t images of the points $\{u, v, u', v'\} \subset A_0$, where A_0 is the underlying surface to $\mathcal{A}_k$ and $f_t(u) = f_t(v)$ and $f_t(u') = f_t(v')$. Here u and u' (resp. v and v') are the endpoints of disjoint arcs ϕ and ϕ' in A_0 which map to ∂w under f_t and $(u, v) = (x_i, y_i)$ and $(u', v') = (x_j, y_j)$, notation as in Definition 5.3. By switching ϕ and ϕ' and/or i and j if necessary, we can assume that the first equality is that of an ordered pair and the second is setwise.

Next we show that after tube sliding moves w becomes a clean Whitney disc, i.e. no tube guide curve crosses $\operatorname{int} \phi$ or $\operatorname{int} \phi'$. Now $x_i \in \partial\phi$, so all the tube guide crossings with $\operatorname{int}(\phi)$ can be eliminated by sliding across the double point x_i. If $x_j \in \partial\phi'$, then we can clear $\operatorname{int}(\phi')$ of tube guide curves in a similar manner. If $\partial\phi' = (y_i, y_j)$ and the tube guide κ crosses $\operatorname{int}(\phi')$ we clear it from ϕ' as follows. Since $i \neq j$, it follows that $\kappa \neq \sigma_k$ for some $k \in \{i, j\}$. Apply a sequence Reidemeister 2 moves supported in a small neighborhood of ϕ' to make κ adjacent to y_k and then slide it across the double point y_k.

Since w is now a clean Whitney disc we can shadow the Whitney move by an uncrossed (if $u' = x_j$) or crossed (if $u' = y_j$) Whitney move. Thus $\mathcal{A}_{k+1}$ is obtained from $\mathcal{A}_k$ by a sequence of tube sliding moves and a clean Whitney move. $\square$

Remarks 5.18. i) If the regular homotopy f_t is of the form finger moves followed by Whitney moves, then the tubed surface following the finger moves can be chosen so that the curves σ_i are embedded and pairwise disjoint away from z_0.

ii) If $f_1(A_0)$ is embedded, then the final tubed surface has no σ_i curves.

iii) There is no restriction on the surface A_0. In the next section we require that A_0 be a 2-sphere.

6. From double tubes to single tubes

In what follows $\mathcal{A}$ is a tubed surface in M with realization A whose underlying surface A_1 is an embedded 2-sphere. The goal of the next three sections is to show that if in addition $\pi_1(M)$ has no 2-torsion and A_1 is homotopic to A, then A_1 is isotopic to A.

In this section we show that $\mathcal{A}$ can be transformed to $\mathcal{A}'$ with isotopic realizations without changing A_1 so that $\mathcal{A}'$ has at most one double tube, which is homotopically inessential. This is done by appropriately replacing pairs of double tubes with pairs of single tubes.

Definition 6.1. Let $\mathcal{A}$ be a tubed surface with realization A. Let κ denote one of σ_i, β_j or γ_k and y the corresponding x_i, b_j or g_k. If we compress the tube in A that follows $f(\kappa)$ near $f(y)$ we obtain an immersed surface, one component of which is an embedded 2-sphere $T = T(\kappa)$, that is homotopic to the transverse sphere G. This sphere has an induced orientation that coincides with A away from the compressing disc. Define $\epsilon(T) = \epsilon(\kappa) = +1$ if $[T] = [G]$ and -1 otherwise. Here M, A and G are oriented so that $< A, G >= +1$.

Similarly compressing A near a point of τ_i gives rise to an embedded surface one component of which is an embedded 2-sphere $T(\alpha_i)$ isotopic to two oppositely oriented copies of G tubed together along α_i.

Lemma 6.2. *If $T = T(\kappa)$ is constructed as above and D is the compressing disc that splits off T and oriented to coincide with that of T, then $\epsilon(T) =< D', A >$ Here D' is D shrunk slightly to have boundary disjoint from A.* □

Lemma 6.3. *If $\mathcal{A}$ is a tubed surface, then for all i, $[T(\beta_i)] = -[T(\gamma_i)] = \pm[G] \in H_2(M)$ and for all j, $[T(\alpha_j)] = 0$.* □

Sign Convention: By switching β_i and γ_i, if necessary, we can assume that $\epsilon(\beta_i) = -1$ and $\epsilon(\gamma_i) = +1$. Orient β_i to point from z to b_i, λ_i to point from $f(b_i)$ to $f(g_i)$ and γ_i to point from g_i to z.

We next calculate $[A] \in \pi_2(M)$. Since A and G are 2-spheres, each element of $\pi_1(M)$ gives rise to a distinct pair of geometrically dual spheres in $\tilde{M}$ that projects to the pair (A, G). Thus, $\pi_2(M) = \pi_2(\tilde{M}) = H_2(\tilde{M})$ which contains the group ring $\mathcal{H} = H_2(G)\pi_1(M)$ as a submodule. Since $\mathcal{A}$ is a tubed surface, $[A]$ lies in the coset $\mathcal{H} + [A_1]$. Since $\pi_1(A_1) = 0$, each λ_i determines a well defined element, $[\lambda_i] \in \pi_1(M)$. With the above conventions the triple $(\beta_i, \lambda_i, \gamma_i)$ gives rise to the element $[G][\lambda_i] - [G][\lambda_i]^{-1} \in \mathcal{H}$. On the other hand, each α_j, τ_j gives rise to the trivial element. We therefore have:

Lemma 6.4. $[A] = [A_1] + \sum_{i=1}^{s}[G][\lambda_i] - [G][\lambda_i]^{-1} \in \pi_2(M)$. □

Remark 6.5. If A is homotopic to A_1, then $\sum_{i=1}^{s}[G][\lambda_i] - [G][\lambda_i]^{-1} = 0$. Therefore, if $[\lambda_i] = 1$ whenever $[\lambda_i]^2 = 1$ holds, then we can reorder the λ_i's so that $[\lambda_1] = [\lambda_2]^{-1}, \cdots, [\lambda_{2p-1}] = [\lambda_{2p}]^{-1}$ and $[\lambda_s] = 1$ if $s = 2p + 1$.

The following is the main result of this section.

Proposition 6.6. *Let $\mathcal{A}$ be a tubed surface in the 4-manifold M whose underlying surface A_1 is an embedded sphere homotopic to the realization A of $\mathcal{A}$. Assume that $\pi_1(M)$ has no 2-torsion and G denotes the transverse sphere to A_1. Then via an isotopy supported away from G, A is isotopic to the realization A' of a tubed surface $\mathcal{A}'$ with underlying surface A_1 such that $\mathcal{A}'$ has at most one double tube.*

Remark 6.7. By Lemma 6.4 the tube path λ corresponding to the unique double tube is homotopically trivial.

Proof. By Remark 6.5 we can reorder the λ_i's so that $[\lambda_1] = [\lambda_2]^{-1}, \cdots, [\lambda_{2p-1}] = [\lambda_{2p}]^{-1}$ and $[\lambda_s] = 1$ if $s = 2p+1$.

We will show that an isotopy of A transforms $\mathcal{A}$ to $\mathcal{A}'$ with the same A_1 but with two fewer double tubes. The lemma then follows by induction. To start with we consider another model for a double tube as shown in Figure 6.1.

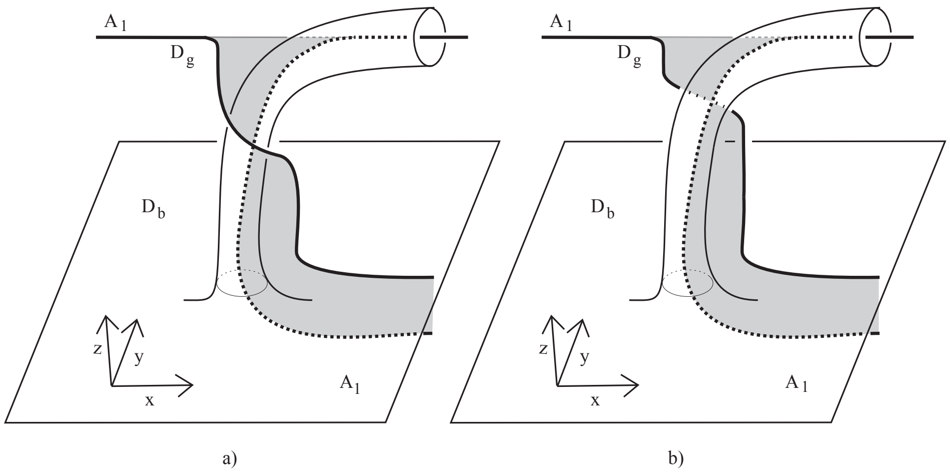


Figure 6.1

Remarks 6.8. i) Figures 6.1 a) and b) each show the projection of a neighborhood of a double tube associated to a path λ into the x, y, z plane. This consists of tubes emanating from two discs D_b and D_g lying in A_1 that are respectively neighborhoods of $f(b)$ and $f(g)$, where λ connects $f(b)$ to $f(g)$. In the figure, D_b lies in the x, y plane and the other lies in the x, t plane. The shaded regions are projections of the tube from D_g into the x, y, z plane. Except where it is twisted, this tube lies in the x, z, t plane. Its intersection with the x, y, z plane is the union of the thick solid and dashed lines.

ii) Since $\pi_1(SO(3)) = \mathbb{Z}_2$, up to isotopy, there are two ways of constructing a double tube associated to the path λ with given germs of double tubes near its endpoints. Representatives are shown in Figures 6.1 a) and b).

Since A_1 has the transverse 2-sphere G it follows that the induced map $\pi_1(M \setminus (A_1 \cup G)) \to \pi_1(M)$ is an isomorphism. Since homotopy implies isotopy we can isotope λ_1 and λ_2 to be anti-parallel. I.e. there exists an embedded square D with opposite edges respectively on A_1 and λ_1 and λ_2. See Figure 6.2 a). Here E and F denote the components of $A_1 \cap N(D)$. Figure 6.2 b) shows how A *might* intersect $N(D)$. Figure 6.3 shows how to isotope the surface to effect a change of the framed embedded path corresponding to the non trivial element of $SO(3)$. Thus we can assume that A appears near λ_1 and λ_2 as depicted in Figure 6.2 b). Now $A \cap N(D)$ may fail to appear as in Figure 6.2 b) because the images of tube guide paths may cross the interior of $D \cap A_1$, however by doing tube sliding moves in a manner similar to those in the proof of Theorem 5.17 we can clear such curves from the neighborhood. Thus we can assume that $A \cap N(D)$ appears as in Figure 6.2 b).

Figure 6.4 shows an isotopy of A, supported in $N(D)$, that transforms a pair of double tubes as in Figure 6.2 b) into a pair of single tubes as shown in Figure 6.4 c). Call the *E-component* of $A \cap N(D)$ the one that intersects E and call the other the *F-component.* Again in Figure 6.4 the intersection of the E-component with the x, y, z-plane is drawn in thick,

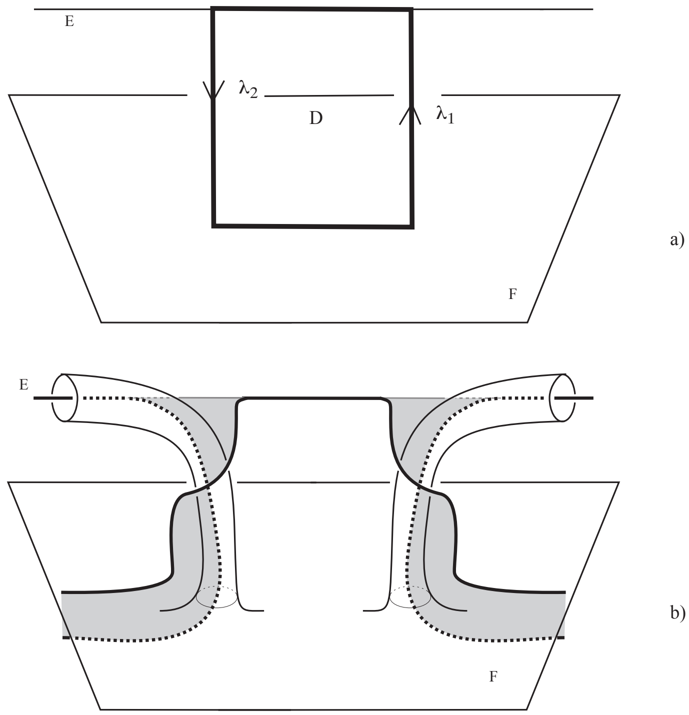


Figure 6.2

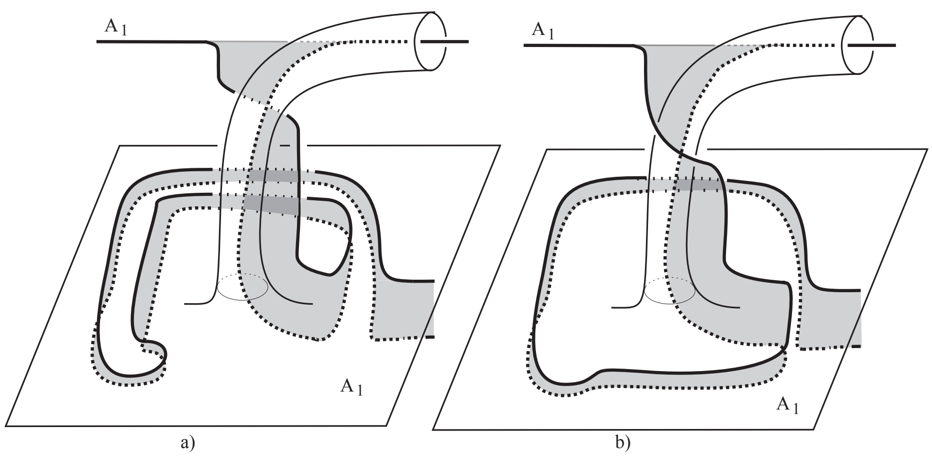


Figure 6.3

possibly dashed, lines. The isotopy from Figure 6.2 b) to Figure 6.4 a) is the composition of three isotopies, the first and third supported in the F-component and the second supported in the E component. The isotopy to Figure 6.4 b) is supported in the E component. The isotopy from Figure 6.4 b) to Figure 6.4 c) is as follows. Without changing the projection into the x, y, z plane, first push the tube emanating from E into the future and then isotope the tube emanating from F as indicated. □

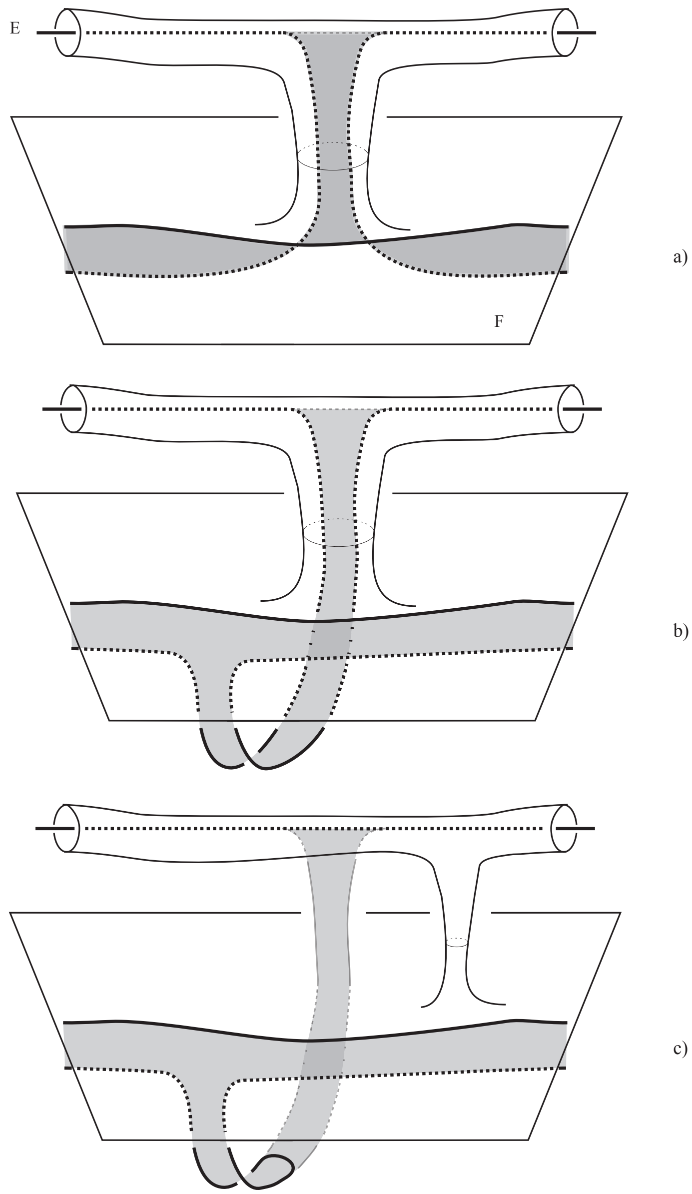


Figure 6.4

7. Crossing Changes

In this section we show that crossing changes involving distinct tube guide curves do not change the isotopy class of the realization of a tubed surface.

Lemma 7.1. *(Crossing Change Lemma) If the tubed surface $\mathcal{A}'$ is obtained from the tubed surface $\mathcal{A}$ by a crossing change involving either distinct tube guide paths or distinct components of $\alpha_i \setminus p_i$, then the corresponding realizations A' and A are isotopic via an isotopy supported away from the transverse surface G. See Figure 7.1 a).*

Proof. Since by Lemma 5.7 changing a tubed surface by type 2) and 3) Reidemeister moves does not change the isotopy class of the realization, it suffices to assume that the crossing is adjacent to z_0 as in Figure 7.1 b), e.g. see Figure 7.1 c).

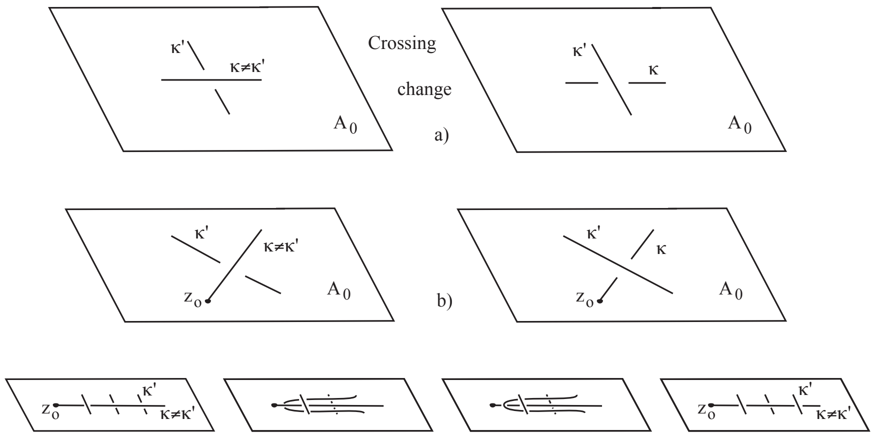


Figure 7.1

The isotopy from A to A' is demonstrated in Figure 7.2. Think of $D^2 \times S^2$ as $D^2 \times S^1 \times [-\infty, +\infty]/\sim$, where each $x \times S^1 \times -\infty$ and each $x \times S^1 \times +\infty$ is identified to a point and with G being the S^2 fiber through z, the origin of D^2. The various subfigures of Figure 7.2 show a neighborhood of $z \in D^2 \times S^1 \times 0$. Let $N(z)$ denote a 4-dimensional regular neighborhood of z. The sets K' and K seen in this figure are projections of the components of $A \cap N(z)$ that contain the local tubes about $f(\kappa')$ and $f(\kappa)$. Also $A_1 \cap N(z)$ is a $D^2 \times \text{pt} \times 0$. Notice that only K' moves during the isotopy. The isotopy corresponding to Figures 7.2 a) and 7.2 b) can be also viewed in Figure 7.3. The dark lines in these figures coincide. That figure shows part of the isotopy seen from the 3-plane orthogonal to A_1 that contains K'. Each subfigure intersects A_1 in a line and K in a circle. The isotopy is supported in a neighborhood of a $S^2 \times I$ where the S^2 is parallel to G. The passage from Figure 7.2 c) to Figure 7.2 d) is the light bulb move, whereby a tube T appears to be crossing A at the point y. This requires that there be a path σ in A from y to z disjoint from T which in turn requires that $\kappa \neq \kappa'$. The isotopy corresponding to Figures 7.2 d) and e) is essentially the reverse of that from 7.2 a) and 7.2 b). There, the projection of the $S^2 \times I$ to A_1 is an arc disjoint from the projection of K, so K is not in the way during that isotopy. $\square$

8. Proof of Theorem 1.2

Lemma 8.1. *Suppose that R_0 and R_1 are spheres with common transverse sphere G in the 4-manifold M. If R_0 and R_1 coincide near G and are homotopic in M, then they are homotopic with via a homotopy whose support in M is disjoint from G.*

Proof. It suffices to show that if a sphere R_3 in $M \setminus G$ is homotopically trivial in M it is homotopically trivial in $M \setminus G$. Let $\tilde{M}$ denote the universal covering of M and $\tilde{G}$ the preimages of G. By Lemma 2.2 the universal cover $\hat{M}$ of $M \setminus G = \tilde{M} \setminus \tilde{G}$, hence if $\tilde{R}_3$ denotes a lift of R_3 to $\tilde{M}$, it suffices to show that $\tilde{R}_3$ is homologously trivial in $\hat{M}$. Since $\tilde{R}_3$ is homologously trivial in $\tilde{M}$ and $\tilde{G}$ is a union of 2-spheres there exists a bounding cycle disjoint from $\tilde{G}$. $\square$

Suppose that the embedded spheres R and A_1 are homotopic with the common transverse sphere G. After an initial isotopy and the previous lemma we can assume that R and

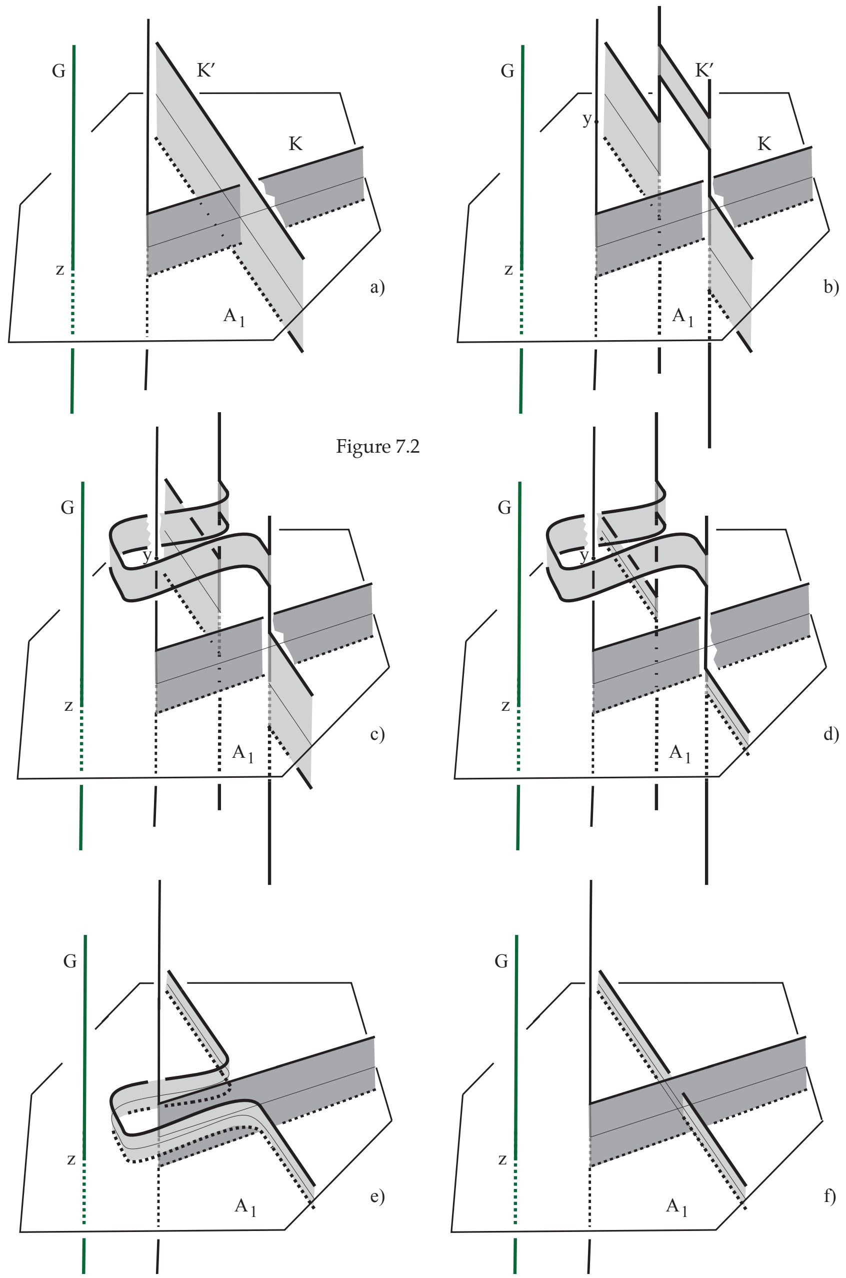


Figure 7.2

A_1 coincide near G and that the homotopy F from R to A_1 is supported away from a neighborhood of G. .

It follows from [Sm1] that R is regularly homotopic to A_1 via a homotopy that is also supported away from a neighborhood of G. See §4. By Theorem 5.17 the homotopy from R to A_1 is shadowed by tubed surfaces. Thus there exists a tubed surface $\mathcal{A}$ with realization A and underlying surface A_1 such that R is isotopic to A. It suffices to show that A is isotopic to A_1 via an isotopy supported away from a neighborhood of G.

By the tube cancellation lemma we can assume that A has at most one double tube and that that tube can be homotoped rel endpoints into A_1 via a disc disjoint from G whose interior is disjoint from A_1. In what follows we assume for consistency of notation that there

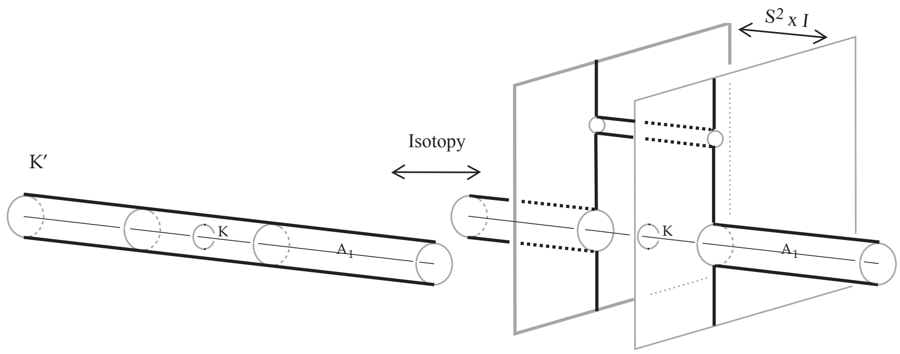


Figure 7.3

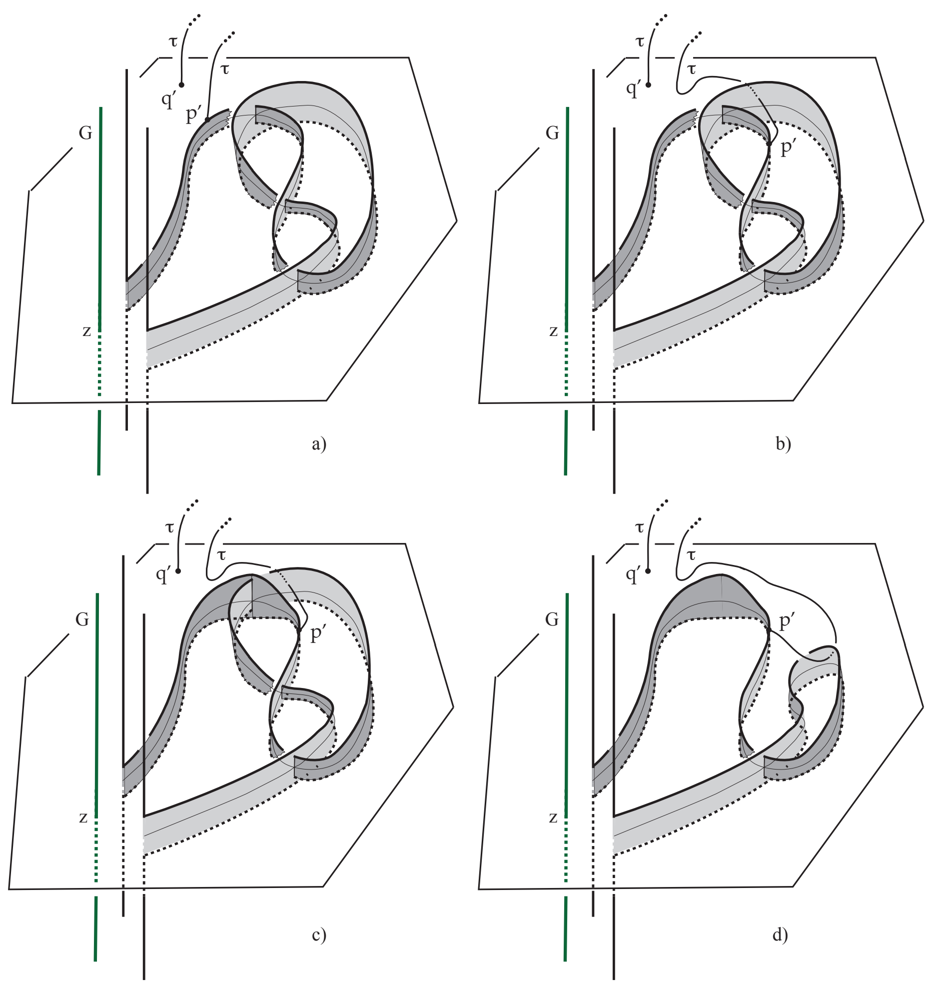


Figure 8.1

exists one such double tube. Using the crossing change and tube sliding Lemmas 7.1 and 5.7 we can assume that each pair $(\alpha_1, q_1), \cdots, (\alpha_r, q_r), (\beta_1, \gamma_1)$ lies in a distinct sector of A_0. This means there exists a neighborhood D of $z_0 \in A_0$ parametrized as the unit disc in polar coordinates so that (α_i, q_i) lies in the subset of D where $((i-1)/(r+1))2\pi < \theta < (i/(r+1))2\pi$ and (β_1, γ_1) lies in the region $(r/(r+1))2\pi < \theta < 2\pi$. We can further assume that $q_1 \in \partial D$. It

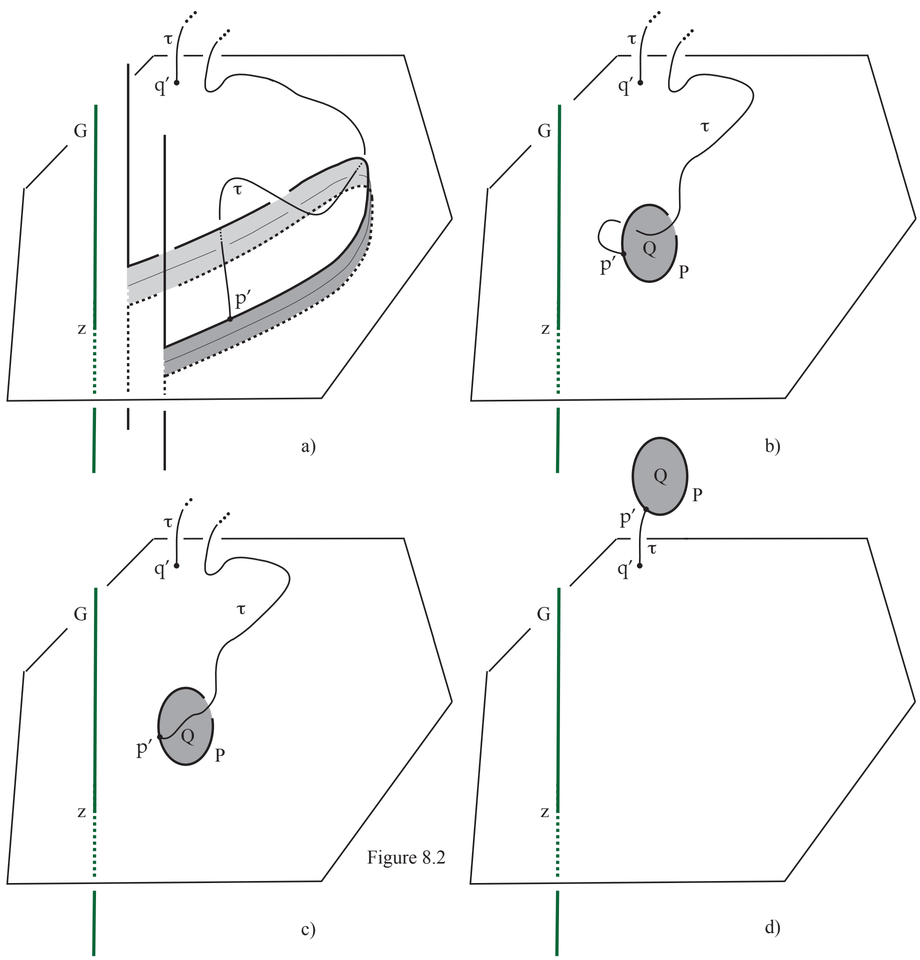


Figure 8.2

remains to prove Lemma 8.2 after which we can assume that $r = 0$. We then can homotope, hence isotope λ_1 and thus A to lie in the $S^2 \times S^2$ factor, hence the result follows from Theorem 1.7.

Lemma 8.2. *Let $\mathcal{A}$ be a tubed surface such that a neighborhood $D \subset A_0$ of z_0 is parametrized as the unit disc D in polar coordinates and $\alpha_1 \cup q_1$ is contained in the subset of D where $0 < \theta < \pi/2$ and that that region is devoid of all other tube guide curves and associated points. Assume also $q_1 \in \partial D$. Then the tubed surface $\mathcal{A}'$ whose data consists of that of $\mathcal{A}$ with $\alpha_1 \cup \tau_1$ deleted has realization A' isotopic to the realization A of $\mathcal{A}$.*

Proof. We first fix some terminology. To simplify notation $\alpha_1, p_1, q_1, \tau_1$ will be respectively denoted α, p, q, τ. $T(\alpha)$ will denote the tube about $f(\alpha)$ and P_α will denote the 2-sphere consisting of two parallel copies of G tubed together along $T(\alpha)$. $T(\tau)$ will denote the tube about τ. So A is the surface obtained from A' by connecting P_α to A' by the tube $T(\tau)$. We will let p' and q' denote the points respectively on P_α and A' so that the ends of $T(\tau)$ connect to $\partial N(p')$ and $\partial N(q')$. Here p' orthogonally projects to $f(p) \in A_1$, where $p \in \alpha \subset A_0$. We let α_L, and α_R denote the components of α separated by p.

The first observation is that by isotopy extension p' can be isotoped to any point in $T(\alpha)$ at the cost of seemingly *entangling* τ with $T(\alpha)$. See Figures 8.1 a) and b). One cannot obviously use the light bulb lemma to remove the intersection of $\mathrm{int}(\tau)$ with the projection of $T(\alpha)$ in Figure 8.1 b), since $T(\tau)$ separates $T(\alpha)$ from z.

By suitably moving p, the local arcs of α at a given crossing can lie in the different components α_L and α_R of $\alpha \setminus p$. Thus the proof of the crossing change lemma allows us to change this crossing as well as any other at the cost of entangling τ with $T(\alpha)$. This process is illustrated in Figures 8.1 a), b), c). Similarly, at the cost of further such entanglement we can perform the reordering move, Definition 5.5 ii) to arcs of α. Compare Figure 8.1 d) and Figure 8.2 a). Thus by crossing changes, Reidemeister 2), 3) moves and the tube sliding reordering move, we can assume that α is unknotted. It follows that P_α can be isotoped to an unknotted 2-sphere P, that bounds a 3-ball Q disjoint from $A' \cup G$. See Figure 8.2 b).

Further there exists a 4-ball B such that $Q \subset B$, $A' \cap B = \emptyset$ and $\tau \cap B$ is connected. Since $\pi_1(B \setminus P) = \mathbb{Z}$ it follows that via isotopy supported within B, A can be isotoped so that $Q \cap \mathrm{int}(T(\tau)) = \emptyset$ as in Figure 8.2 c). Use the fact that $T(\tau)$ can be rotated about P and is the frontier of a neighborhood of an arc. It follows that A can be isotoped to A', thereby completing the proof. $\square$

Remarks 8.3. With more work one can eliminate reliance on Theorem 1.7. The above argument was free of Theorem 1.7 when there were no double tubes. Otherwise, it reduced to the case of no single tubes and one homotopically trivial double tube where $\beta_1 \cap \gamma_1 = \emptyset$. If β_1 and γ_1 are unknotted, then a direct argument allows for the elimination of this data from $\mathcal{A}$, via isotopy of A and hence the result follows. If not, then reverse this procedure to create a second homotopically trivial double tube. The proof of Proposition 6.6 shows that these double tubes can be transformed into a pair of single tubes and hence the result follows.

9. Higher genus surfaces

In this section we give a partial generalization of our main result to higher genus surfaces, that is a full generalization for $S^2 \times S^2$.

Definition 9.1. Let S be an immersed surface in the 4-manifold M. We say that the embedded disc $D \subset M$ is a *compressing disc* for S if $\partial D \subset S$ and a section of the normal bundle to $\partial D \subset S$ extends to a section of the normal bundle of $D \subset M$.

Lemma 9.2. *If S is immersed in the 4-manifold M and $\alpha \subset S$ is an embedded curve with a trivial normal bundle in S and is homotopically trivial in M, then α bounds a compressing disc.*

Proof. First span α by an immersed disc D_0. Using boundary twisting [FQ] we can replace D_0 by D_1 that satisfies the normal bundle condition. Eliminate the self intersections of D_1 by applying finger moves. $\square$

Lemma 9.3. *Let S be an orientable embedded surface in the 4-manifold M whose components have pairwise disjoint transverse spheres. Let $\alpha_1, \cdots, \alpha_k \subset S$ be pairwise disjoint simple closed curves, disjoint from the transverse spheres, such that for each component S' of S, $S' \setminus \{\alpha_1, \cdots, \alpha_k\}$ is connected. Suppose that for each i, α_i is homotopically trivial in the complement of the transverse spheres. Then there exist pairwise disjoint compressing discs $D_1, \cdots, D_k$ such that for each i, $D_i \cap S = \alpha_i$.*

Proof. Construct compressing discs $A_1, \cdots, A_k$ for the α_i's as in Lemma 9.2. These discs can be chosen to be disjoint from the transverse spheres. Using finger moves they can be made disjoint from each other. Finally use the transverse spheres to tube off intersections of the A_i's with S to create the desired D_i's. $\square$

Definition 9.4. We say that the surface S_1 is obtained from S by compressing along D if $S_1 = S \setminus \text{int}(N(\partial D)) \cup D' \cup D''$ where D', D'' are two pairwise disjoint parallel copies of D.

Lemma 9.5. *Surfaces can be compressed along compressing discs. If S_1 is obtained by compressing the embedded surface S along the compressing disc D and $D \cap S = \partial D$, then S_1 is embedded.* $\square$

Definition 9.6. We say that the surface $S \subset M$ is *G-inessential* if the induced map $\pi_1(S \setminus G) \to \pi_1(M \setminus G)$ is trivial.

The following is a generalization to higher genus surfaces of Theorem 1.2.

Theorem 9.7. *Let M be an orientable 4-manifold such that $\pi_1(M)$ has no 2-torsion. Two homotopic, embedded, G-inessential surfaces S_1, S_2 with common transverse sphere G are ambiently isotopic, via an isotopy that fixes the transverse sphere pointwise.*

Proof. For each $i \in \{1, 2\}$ let $\alpha_1^i, \cdots, \alpha_k^i$ be a set of pairwise disjoint simple closed curves in S_i whose complement is a connected planar surface containing $S_i \cap G$. Let $D_1^i, \cdots, D_k^i$ be associated pairwise disjoint compressing discs with interiors disjoint from S_i and let T_i be the result of compressing S_i along these discs. Then T_i is a 2-sphere and S_i is obtained from T_i by attaching k tubes. Each tube $S^1 \times I$ extends to a solid tube $D^2 \times I$ which intersects T_i exactly at $D^2 \times 0$ and $D^2 \times 1$, which we call the *bases of the tube*. By construction, these tubes are pairwise disjoint. After a further isotopy we can assume there are k small pairwise disjoint 4-balls which intersect S_i in a single standard disc and each such disc contains the bases of a single solid tube.

Since S_i is G-inessential, it follows by the light bulb lemma that the solid tubes can be isotoped to 3-dimensional neighborhoods of tiny standard arcs with endpoints on T_i. Note that the induced ambient isotopy can be chosen to fix a neighborhood of T_i pointwise.

To complete the proof it suffices to show that T_1 and T_2 are homotopic and hence isotopic by Theorem 1.2. To see this, consider the lifts $\tilde{T}_1, \tilde{T}_2$ of T_1, T_2 to the universal covering $\tilde{M}$ of M which intersect a given lift $\tilde{G}$ of G. Since the S_i's are π_1-inessential and homotopic, their corresponding lifts $\tilde{S}_1, \tilde{S}_2$ are homotopic and hence homologous. It follows that $\tilde{T}_1$ and $\tilde{T}_2$ are homologous and hence homotopic and therefore so are T_1 and T_2. $\square$

Applying to the case of $S^2 \times S^2$ we obtain:

Theorem 9.8. *Let R be a connected embedded genus-g surface in $S^2 \times S^2$ such that $R \cap S^2 \times y_0 = 1$. Then R is isotopically standard. I.e. it is isotopic to the standard sphere in its homology class, with g standard handles attached, via an ambient isotopy that fixes $S^2 \times y_0$ pointwise.* $\square$

10. Applications and questions

We begin by stating the main result for multiple spheres.

Theorem 10.1. *Let M be an orientable 4-manifold such that $\pi_1(M)$ has no 2-torsion. Let $G_1, \cdots, G_n$ be pairwise disjoint embedded spheres with trivial normal bundles. Let $R_1, \cdots, R_n$ be pairwise disjoint embedded spheres transverse to the G_i's such that $|R_i \cap G_j| = \delta_{ij}$. Let $S_1, \cdots, S_n$ be another set of spheres with the same properties. If for each i, R_i is homotopic to S_i, then there exists an isotopy of M fixing the G_i's pointwise such that for all j, R_j is taken to S_j.*

Under corresponding hypotheses, the same conclusion holds when the R_i's are G-inessential connected surfaces, where $G = \{G_1, \cdots, G_n\}$.

Proof. The methods of §9 reduce the general case to the case that all the S_i's are spheres.

Proof by induction on n.

Step 1: R_1 is ambient isotopic to S_1 via an isotopy that fixes the G_i's pointwise.

Proof After a preliminary isotopy we can assume that R_1 and S_1 coincide near $R_1 \cap G$ and that the homotopy from R_1 to S_1 is supported away from a neighborhood of $\cup G_i$. Step 1 follows by applying Theorem 1.2 to the manifold $M \setminus \cup_{i=2}^n N(G_i)$. Note that the inclusion of $M \setminus \cup_{i=2}^n N(G_i) \to M$ induces a fundamental group isomorphism so the no 2-torsion condition is satisfied. □

Induction Step: Suppose that we have for $j < k$, $R_j = S_j$. There exists an isotopy of M fixing $\cup_{i=k}^n G_i$ pointwise and supported away from $\cup_{j=1}^{k-1}(G_j \cup S_j)$ such that R_k is taken to S_k.

Proof After a preliminary isotopy we can assume that R_k and S_k coincide near G_k and that R_k is homotopic to S_k via a homotopy supported away from $\cup_{j=1}^{k-1}(S_j \cup G_j)$. Next apply Step 1 to $R_k \subset M \setminus N(\cup_{j=1}^{k-1}(S_j \cup G_j))$. Again, the argument of Lemma 2.2 implies that the inclusion $M \setminus N(\cup_{j=1}^{k-1}(S_j \cup G_j)) \to M$ induces a fundamental group isomorphism, so the no 2-torsion condition is satisfied. □

Definition 10.2. An essential simple closed curve in $S^2 \times S^1$ is said to be *standard* if it is isotopic to $x \times S^1$ for some $x \in S^2$.

Theorem 10.3. *Two properly embedded discs D_0 and D_1 in $S^2 \times D^2$ that coincide near their standard boundaries are isotopic rel boundary if and only if they are homologous in $H_2(S^2 \times D^2, \partial D_0)$.*

Proof. Homologous is certainly a necessary condition. In the other direction, after reparameterizing, we can assume that $\partial D_0 = x_0 \times S^1 \subset S^2 \times S^1$. Let $M = S^2 \times D^2 \cup d(S^2 \times D^2) = S^2 \times S^2$ be obtained by doubling $S^2 \times D^2$ with $d(S^2 \times D^2)$ denoting the other $S^2 \times D^2$. This $d(S^2 \times D^2)$ can be viewed as a regular neighborhood $N(G)$ of $G = d(S^2 \times 0)$. Let R_i denote the sphere $D_i \cup d(x_0 \times D^2)$ which we can assume is smooth for $i = 0, 1$. G is a transverse sphere to the homologous spheres R_0 and R_1. By Theorem 1.2 there is an isotopy of M fixing a neighborhood of G pointwise taking R_0 to R_1. Since R_0 and R_1 coincide in a neighborhood of $N(G)$ there is an isotopy of $S^2 \times D^2$ taking D_0 to D_1 that fixes a neighborhood of $S^2 \times S^1$ pointwise. □

Theorem 10.4. *A properly embedded disc D in $S^2 \times D^2$ is properly isotopic to a fiber if and only if its boundary is standard.*

Proof. After a preliminary isotopy we can assume that ∂D is the standard vertical curve $x_0 \times S^1$ which we denote by J. Let F be a D^2 fiber of $S^2 \times D^2$. Now $0 \to H_2(S^2 \times D^2) \to H_2(S^2 \times D^2, J) \to H_1(J) \to 0$ is split and exact, so the subgroup H mapping to the generator $[\partial F]$ of $H_1(J)$ equals Z and is represented by the classes $[F] + n[S^2 \times y_0]$, where $y_0 \in \partial D^2$. By properly isotoping D to D' where $\partial D' = J$ and so that the track of the homotopy restricted to the boundary is approximately $J \cup S^2 \times y_0$ it follows that $[D'] = [D] + [S^2 \times y_0] \in H_2(S^2 \times D^2, J)$. Therefore any class in H is represented by a disc properly isotopic to D. In particular after proper isotopy we can assume that $[D] = [F]$. After a further isotopy we can assume that D coincides with F near ∂D. The result now follows by Theorem 10.3. The other direction is immediate. $\square$

Recall that $\operatorname{Diff}_0(X)$ denotes the group of diffeomorphisms properly homotopic to the identity.

Corollary 10.5. $\pi_0(\operatorname{Diff}_0(S^2 \times D^2)/\operatorname{Diff}_0(B^4)) = 1$.

Remark 10.6. This means that a diffeomorphism of $S^2 \times D^2$ properly homotopic to the identity is isotopic to one that coincides with the identity away from a compact 4-ball disjoint from $S^2 \times S^1$.

Proof. If $f : S^2 \times D^2 \to S^2 \times D^2$ is properly homotopic to the identity, then $\partial f : S^2 \times S^1 \to S^2 \times S^1$ is homotopic to the identity, hence isotopic to the identity by [La]. After another isotopy we can assume that $f|N(S^2 \times S^1) = \text{id}$. By Theorem 10.4 a further isotopy takes a fiber $z \times D^2$ to itself. By Smale [Sm3] we can additionally assume that $f|x \times D^2 = \text{id}$. After a further isotopy we can assume that $f|N(S^2 \times \partial D^2 \cup z \times D^2)$ is the identity. Since the closure of what's left is a B^4 the result follows. $\square$

The following is an immediate consequence of our main result.

Theorem 10.7. *(4D-Lightbulb Theorem) If R is an embedded 2-sphere in $S^2 \times S^2$, homologous to $x_0 \times S^2$, that intersects $S^2 \times y_0$ transversely and only at the point (x_0, y_0), then R is isotopic to $x_0 \times S^2$ via an isotopy fixing $S^2 \times y_0$ pointwise.* $\square$

Litherland [Li] proved that there exists a diffeomorphism pseudo-isotopic to the identity that takes R to $x_0 \times S^2$.

Another version of the light bulb theorem was obtained in 1986 for PL discs in S^4 by Marumoto [Ma] where the isotopy is topological. He makes essential use of Alexander's theorem that any homeomorphism of B^n that is the identity on S^{n-1} is (topologically) isotopic to the identity. Here we prove a general form of the smooth version.

Theorem 10.8. *(Uniqueness of Spanning Surfaces) If R_0 and R_1 are smooth embedded surfaces in S^4 of the same genus such that $\partial R_0 = \partial R_1 = \gamma$, where γ is connected, then there exists a smooth isotopy of S^4 taking R_0 to R_1 that fixes γ pointwise.*

Proof. First consider the case that R_0 and R_1 are discs. After a preliminary isotopy of S^4 that fixes γ pointwise, we can assume that R_0 and R_1 coincide in a neighborhood of their boundaries. Now $S^4 \setminus \operatorname{int}(N(\gamma)) = S^2 \times D^2$. Thus R_0 and R_1 restrict to properly embedded discs E_0 and E_1 in $S^2 \times D^2$ that coincide near their boundaries.

We can assume that after a second isotopy $[E_0] = [E_1] \in H_2(S^2 \times D^2, \gamma)$ also holds. Indeed, $N(\partial E_0)$ is determined by monotone maps $f_t : S^1 \to S^2 \times S^1$, where monotone means transverse to the S^2-factor and f_0 corresponds to ∂E_0. The second isotopy corresponds to one $f_t^s : S^1 \times I \to S^2 \times S^1$ with $f_1^s = f_1$ all s and f_0^s sweeps across the S^2-factor as many times as needed for $s \in [0,1]$. It follows by Theorem 10.3 that E_0 can be isotoped to E_1 via an isotopy supported away from a neighborhood of $S^2 \times S^1$.

The general case similarly follows using Theorem 9.8. □

Remark 10.9. By induction Marumoto [Ma] proved more generally that two locally flat PL m-discs in an n-sphere, $n > m$ with the same boundary are topologically isotopic rel boundary. Here is an outline of his argument for smooth discs in the n-sphere for the representative case $m = 2, n = 4$, where we use [Ce1], [Pa] to avoid his induction steps. Actually, the below argument works in all dimensions and codimensions since the same is true of [Ce1], [Pa] and the Alexander isotopy.

Start with D_0, D_1 where D_1 is the standard 2-disc in S^4 and $\partial D_0 = D_1$. Then by [Ce1], [Pa] there is a diffeomorphism $f : S^4 \to S^4$ taking D_0 to D_1 fixing ∂D_0. We can assume that f fixes pointwise a neighborhood of ∂D_0. Next remove a small ball about a point in ∂D_0. After restricting and reparametrizing we obtain a map $g : B^4 \to B^4$ such that $g(E_0) = E_1$ where the E_i's are the restricted reparametrized D_i's. Here B^4 is the unit ball in $\mathbb{R}^4, \partial E_0$ is a straight properly embedded arc connecting antipodal points of ∂B^4 and $g|\partial B^4 = \mathrm{id}$. Finally apply the Alexander isotopy to obtain a topological isotopy of g to the identity which fixes ∂E_0 pointwise. □

More generally we have the following uniqueness of spanning discs in simply connected 4-manifolds.

Theorem 10.10. *If D_0 and D_1 are smooth embedded discs in the simply connected 4-manifold M such that $\partial D_0 = \partial D_1 = \gamma$, then there exists a smooth isotopy of M taking D_0 to D_1 fixing γ pointwise if and only if the mapped sphere $S = D_0 \cup_\gamma D_1$ is inessential in M.*

Proof. If D_0 and D_1 are isotopic, then the isotopy sweeps out a contracting ball for S. Conversely, after an initial isotopy of D_1 we can assume that it coincides with D_0 near γ and that the interior of the mapped 3-ball B defining the contraction of S intersects γ algebraically zero. Indeed, the second isotopy in the proof of Theorem 10.8 enables modification of the intersection number. These intersections can be eliminated using immersed Whitney discs. Next surger γ to obtain the simply connected manifold N so that D_0 and D_1 give rise to homotopic spheres R_0 and R_1 with common transverse sphere G, that coincide near their intersection with G. By Theorem 1.2, R_0 and R_1 are isotopic via an isotopy fixing G pointwise and hence D_0 and D_1 are isotopic rel boundary. □

Remark 10.11. In a similar manner, using Theorems 10.1 and 9.7, one can obtain uniqueness theorems for certain surfaces spanning simple closed curves in closed 4-manifolds with no 2-torsion in their fundamental groups.

One can ask the following parametrized form in the smooth category.

Question 10.12. *For $i = 1, 2$ let $f_i : D^k \to S^4$ be smooth embeddings such that $f_1|\partial D^k = f_2|\partial D^k$. Is there a smooth isotopy $F : S^4 \times I \to S^4$ such that $F_0 = \mathrm{id}_{S^4}, F_t(f_1(x)) = f_1(x)$ for $x \in \partial D^k$ and $t \in [0,1]$ and for $y \in D^k, F_1(f_2(y)) = f_1(y)$?*

Remark 10.13. For $k \leq 3$ the unparametrized version implies the parametrized one by [Ce3] for k=3 and [Sm3] for $k = 2$ with the $k = 1$ case being elementary. The point of this question is to link various theorems, conjectures and questions.

Case k=1: This is the theorem *homotopy implies isotopy for curves in 4-manifolds.*

Case k=2: This is Theorem 10.8.

Case k=3: This implies the Schoenflies conjecture. Indeed the Schoenflies conjecture is equivalent to a positive resolution of the question after allowing lifting of the f_i's to some finite branched covering of S^4 over $\partial(f_i(D^3))$. See [Ga].

Case k=4: This is the question of connectivity of $\mathrm{Diff}_0(B^4, \partial)$.

Question 10.14. *Is the $\mathbb{Z}_2$-condition necessary for Theorem 1.2?*

Remark 10.15. In the context of the statement of Theorem 1.2, if R is an embedded sphere and $M_1 \to M$ is a finite cover such that $\pi_1(M_1)$ has no 2-torsion, then the preimages of R are simultaneously isotopically *standard*, though perhaps not equivariently.

Question 10.16. *Does Theorem 10.1 hold without the G-inessential condition? What if G-inessential is replaced by π_1-inessential?*

The following are special cases of the long standing questions of whether a sphere R in $\mathbb{CP}^2$ homologous to $\mathbb{CP}^1$ is equivalent up to isotopy or diffeomorphism to the standard $\mathbb{CP}^1$. See problem 4.23 [Ki].

Questions 10.17. *i) If R is a smooth sphere in $\mathbb{CP}^2$ that intersects $\mathbb{CP}^1$ once is R isotopically standard?*

ii) [Me] *Is $(\mathbb{CP}^2, R)$ diffeomorphic to $(\mathbb{CP}^2, \mathbb{CP}^1)$?*

Remark 10.18. In his unpublished 1977 thesis, Paul Melvin [Me] showed that blowing down $\mathbb{CP}^2$ along $\mathbb{CP}^1$ transforms R to a 2-knot T in S^4 and Gluck twisting S^4 along T yields S^4 if and only if $(\mathbb{CP}^2, R)$ is diffeomorphic to $(\mathbb{CP}^2, \mathbb{CP}^1)$. He gave a positive answer to ii) for 0-concordant knots.

References

[Ak] S. Akbulut, *Isotoping spheres in 4-manifolds*, Proc. Gokova Geom.-Top. Conf. 2014, 264–266.

[AKMR] D. Auckly, H. Kim, P. Melvin & D. Ruberman, *Stable isotopy in four dimensions*, J. Lond. Math. Soc., (2)**92** (2015), 439–463.

[Ce1] J. Cerf, Thesis.

[Ce2] J. Cerf, *Sur les Diffeomorphismes de la Sphere de Dimension Trois ($\Gamma_4 = 0$)*, Springer Lecture Notes, **53** (1968).

[Ce3] J. Cerf, *La stratication naturelle des espaces de fonctions differentiables reels et le theoreme de la pseudo-isotopie*, Inst. Hautes Etudes Sci. Publ. Math. **39** (1970), 5–173.

[E] R. D. Edwards, *The solution of the* 4*-dimensional annulus conjecture (after Frank Quinn)*, Contemp. Math., **35** (1982), 211–264.

[Fr] M. Freedman, *The disk theorem for four-dimensional mnaifolds*, Proc. Int. Cong. Math., (1983), 647–665.

[FQ] M. Freedman & F. Quinn, Topology of 4-manifolds, Princeton Mathematical Series, **39**, Princeton University Press, Princeton, NJ, 1990.

[Ga] D. Gabai, *3-spheres in S^4 and pseudoisotopies of $S^3 \times S^1$*, in preparation.

[Gl] H. Gluck, *The embedding of two-spheres in four-spheres*, Bull. Amer. Math. Soc., **67** (1961), 586–589.

[Ha] A. Hatcher *A proof of the Smale conjecture,* $\mathrm{Diff}(S^3) \simeq \mathrm{O}(4)$ Ann. of Math. (2) **117** (1983), 553–607.

[KSS] A Kawauchi, T. Shibuya & S. Suzuki, *Descriptions on surfaces in four-space. I Normal forms*, Math. Sem. Notes Kobe Univ. **10** (1982), 75–125.

[Ki] R. Kirby, *Problems in low dimensional manifold theory*, Proc. Sympos. Pure Math., **32** Part 2 (1978), 273–312.

[Kr] M. Kreck, *Isotopy classes of diffeomorphisms of (k-1)-connected almost-parallelizable 2k-manifolds*, Springer Lect. Notes. Math., **763** (1979), 643-661.

[La] F. Laudenbach, *Sur Les 2-spheres d'une variete de dimension 3*, Annals of Math., **(2) 97** (1973), 57–81.

[Li] R. Litherland, *A Generalization of the Lightbulb Theorem and PL I-Equivalence of Links*, Proc. AMS, **98** (1986), 353–358.

[Ma] Y. Marumoto, *On Higher Dimensional Light Bulb Theorem*, Kobe J. Math, **3** (1986), 71–75.

[Me] P. Melvin, *Blowing up and down in 4-manifolds*, 1977 Ph. D. thesis, UCBerkeley.

[Pa] R. Palais, *Extending diffeomorphisms*, Proc. Amer. Math. Soc. **11** (1960), 274–277.

[Qu] F. Quinn, *Isotopy of 4-manifolds*, J. Diff. Geom. **24** (1986), 343–372.

[Sm1] S. Smale, *A classification of immersions of the two-sphere*, Trans. AMS (1957), 281–290.

[Sm2] S. Smale, *The classification of immersions of spheres in Euclidean spaces*, Ann. of Math. (2) **69** (1959), 327–344.

[Sm3] S. Smale, *Diffeomorphisms of the 2-Sphere*, Proceedings of the American Mathematical Society, **10** (1959), 621–626.

Department of Mathematics, Princeton University, Princeton, NJ 08544
E-mail address: gabai@math.princeton.edu