Max-Planck-Institut für Mathematik
Bonn

Diffeomorphisms of the 4-sphere, Cerf theory and Montesinos twins

by

David Gay

Max-Planck-Institut für Mathematik
Preprint Series 2021 (7)

Date of submission: February 26, 2021

Diffeomorphisms of the 4-sphere, Cerf theory and Montesinos twins

by

David Gay

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

160 Milledge Terrace
Athens, GA 30606
USA

MPIM 21-7

Diffeomorphisms of the 4-sphere, Cerf theory and Montesinos twins

DAVID T GAY

One way to better understand the smooth mapping class group of the 4–sphere would be to give a list of generators in the form of explicit diffeomorphisms supported in neighborhoods of submanifolds, in analogy with Dehn twists on surfaces. As a step in this direction, we describe a surjective homomorphism from a group associated to loops of 2–spheres in $S^2 \times S^2$'s onto this smooth mapping class group, discuss two natural and in some sense complementary subgroups of the domain of this homomomorphism, show that one is in the kernel, and give generators as above for the image of the other. These generators are described as twists along Montesinos twins, i.e. pairs of embedded 2–spheres in S^4 intersecting transversely at two points.

1 Introduction

Given a smooth oriented manifold X, let $\mathit{Diff}^+(X)$ be the space of orientation preserving diffeomorphisms of X (fixed on a collar neighborhood of ∂X if $\partial X \neq \emptyset$). Here, inspired heavily by Watanabe's work [10] on homotopy groups of $\mathit{Diff}^+(B^4)$ and Budney and Gabai's work [2] on knotted 3–balls in S^4, we initiate a study of $\pi_0(\mathit{Diff}^+(S^4))$, i.e. the smooth mapping class group of the 4–sphere. We know very little about this group except that it is abelian and that every orientation preserving diffeomorphism of S^4 is *pseudoisotopic* to the identity; the group could very well be trivial, like the topological mapping class group. Ideally we would like to find a generating set for this mapping class group defined explicitly and geometrically, for example as explicit diffeomorphisms supported in neighborhoods of explicit submanifolds of S^4, in analogy with Dehn twists as generators of the mapping class groups of surfaces. In this paper we construct a surjective homomorphism from a limit of fundamental groups of certain embedding spaces of 2–spheres in 4–manifolds onto $\pi_0(\mathit{Diff}^+(S^4))$, we describe one geometrically natural subgroup of the domain of this homomorphism which we show to be in its kernel, and we describe a "complementary" geometrically natural subgroup and give an explicit list of generators as above for its image in $\pi_0(\mathit{Diff}^+(S^4))$.

Given smooth manifolds X and Y, let $Emb(X,Y)$ denote the space of embeddings of X into Y. Let $\#^n(S^2\times S^2)^\dagger$ refer to a punctured $\#^n(S^2\times S^2)$. (The puncture is not important until the next paragraph.) Let $\mathcal{S}_n = Emb(\amalg^n S^2, \#^n(S^2\times S^2)^\dagger)$. This $\mathcal{S}$ stands for "spheres", as in "space of embeddings of collections of spheres". Fix a point $p\in S^2$ and let $\amalg^n(S^2\times\{p\})\subset \#^n(S^2\times S^2)^\dagger$ denote the union of one copy of $S^2\times\{p\}$ in each $S^2\times S^2$ summand of $\#^n(S^2\times S^2)^\dagger$. This will be our basepoint in $\mathcal{S}_n$, which we will often suppress from our notation, with the understanding that $\mathcal{S}_n$ is a pointed space. We will also be interested in two subspaces of $\mathcal{S}_n$: Let $\mathcal{S}_n^0$ denote the subspace of embeddings with the property that for each i and j the i'th component of $\amalg^n S^2$ intersects the $\{p\}\times S^2$ in the j'th summand of $\#^n(S^2\times S^2)^\dagger$ transversely at δ_{ij} points. Let $\widehat{\mathcal{S}}_n$ denote the subspace of embeddings with the property that the image of $\amalg^n S^2$ is disjoint from $\amalg^n(S^2\times\{p'\})$ for some fixed $p'\neq p\in S^2$. Note that our basepoint lies in both of these subspaces.

There is a natural homomorphism $\mathcal{H}_n:\pi_1(\mathcal{S}_n)\to\pi_0(Diff^+(S^4))$ defined as follows, and with more care in the next section: Given a loop of embeddings $\alpha_t:\amalg^n S^2\hookrightarrow \#^n(S^2\times S^2)^\dagger\subset \#^n(S^2\times S^2)$ representing some $a\in\pi_1(\mathcal{S}_n)$, for each t build a 5–dimensional cobordism by attaching first n 5–dimensional 2–handles to $[0,1]\times S^4$ with some standard fixed attaching data in $\{1\}\times S^4$, so that the top boundary is canonically identified with $\#^n(S^2\times S^2)$, and then attach n 5–dimensional 3–handles using α_t as the attaching map. Varying t, we get a bundle over S^1 of 5–dimensional cobordisms; since attaching 3–handles along the basepoint embedding $\amalg^n(S^2\times\{p\})$ canonically cancels the 2–handles, the bundle of cobordisms is a cobordism from $S^1\times S^4$ to an interesting S^4–bundle over S^1. We define $\mathcal{H}_n(a)$ to be the monodromy of this bundle.

Thanks to the puncture, we have a natural basepoint-preserving inclusion $\jmath:\mathcal{S}_n\hookrightarrow\mathcal{S}_{n+1}$, respecting the inclusions of $\mathcal{S}_n^0$ and $\widehat{\mathcal{S}}_n$, and thus inclusion-induced homomorphisms $\jmath_*:\pi_1(\mathcal{S}_n)\to\pi_1(\mathcal{S}_{n+1})$ which commute with $\mathcal{H}_n$ and $\mathcal{H}_{n+1}$ and with the inclusion-induced homomorphisms $\imath_*:\pi_1(\mathcal{S}_n^0)\to\pi_1(\mathcal{S}_n)$ and $\imath_*:\pi_1(\widehat{\mathcal{S}}_n)\to\pi_1(\mathcal{S}_n)$. As a consequence, we have limit groups which we will denote $\pi_1(\mathcal{S}_\infty)$, $\pi_1(\mathcal{S}_\infty^0)$, $\pi_1(\widehat{\mathcal{S}}_\infty)$ (it is not important for us to think about the limiting spaces, just the groups, but this notation is convenient), limiting inclusion-induced homomorphisms between them, and a limit homomorphism $\mathcal{H}_\infty:\pi_1(\mathcal{S}_\infty)\to\pi_0(Diff^+(S^4))$. Our first result is:

Theorem 1 *The homomorphism $\mathcal{H}_\infty:\pi_1(\mathcal{S}_\infty)\to\pi_0(Diff^+(S^4))$ is surjective and the kernel of $\mathcal{H}_\infty$ contains $\imath_*(\pi_1(\mathcal{S}_\infty^0))$.*

Our second result characterizes $\mathcal{H}_\infty(\imath_*(\pi_1(\widehat{\mathcal{S}}_\infty)))$ in terms of a countable list of explicit generators which we will now describe. This could in principle be all of $\pi_0(\mathit{Diff}^+(S^4))$, although we have no evidence for or against that possibility.

A *Montesinos twin* in S^4 is a pair $W = (R, S)$ of embedded 2–spheres $R, S \subset S^4$ intersecting transversely at two points. For us, the 2–spheres are both oriented. Montesinos shows [8] that the boundary of a neighborhood $\nu(W)$ of $R \cup S$ is diffeomorphic to $S^1 \times S^1 \times S^1$ and that in fact there is a canonical parametrization $S^1_l \times S^1_R \times S^1_S \cong \partial\nu(W)$, canonical up to postcomposing with diffeomorphisms of $\partial\nu(W)$ which are isotopic to the identity and precomposing with independent diffeomorphisms of S^1_l, S^1_R and S^1_S. This parametrization is characterized by $S^1_l \times \{b\} \times \{c\}$ being homologically trivial in $H_1(S^4 \setminus (R \cup S))$, i.e. a "longitude", $\{a\} \times S^1_R \times \{c\}$ being a meridian for R, and $\{a\} \times \{b\} \times S^1_S$ being a meridian for S. This then parametrizes a neighborhood of $\partial\nu(W)$ as $[-1,1] \times S^1_l \times S^1_R \times S^1_S$. We adopt the orientation conventions that S^1_R and S^1_S have the standard meridian orientations coming from the orientations of R and S, that $[-1,1]$ is oriented in the outward direction from $\nu(W)$, and that S^1_l is oriented so that the orientation of $[-1,1] \times S^1_l \times S^1_R \times S^1_S$ agrees with the standard orientation of S^4.

Definition 2 Given a Montesinos twin W in S^4, parametrize a neighborhood of $\partial\nu(W)$ as $[-1,1] \times S^1_l \times S^1_R \times S^1_S$ as above. Let $\tau_l : [-1,1] \times S^1_l \to [-1,1] \times S^1_l$ denote a right-handed Dehn twist. The *twin twist along* W, denoted τ_W, is the diffeomorphism of S^4 which is the identity outside this neighborhood of $\partial\nu(W)$ and is equal to $\tau_l \times \mathrm{id}_{S^1_R} \times \mathrm{id}_{S^1_S}$ inside this neighborhood.

By the canonicity of our parametrization, τ_W is well-defined up to isotopy, i.e. $[\tau_W]$ is a well-defined class in $\pi_0(\mathit{Diff}^+(S^4))$. Incidentally, we have the following as a consequence of our orientation conventions:

Lemma 3 *If $W = (R, S)$ is a Montesinos twin, then $[\tau_W]^{-1} = [\tau_{(S,R)}] = [\tau_{(\overline{R},S)}] = [\tau_{(R,\overline{S})}]$.*

Proof Either switching the spheres S and R, or reversing the orientation of one of them, reverses the orientation of $S^1_R \times S^1_S$, which then forces the reversal of the orientation of $[-1,1] \times S^1_l$, changing a positive Dehn twist to a negative Dehn twist. □

We now describe a family of twins $W(i) = (R(i), S)$, for $i \in \mathbb{N} \cup \{0\}$. Figure 1 illustrates $W(3)$; $W(i)$ is the same but with i turns in the spiral rather than 3 turns. Figure 2 and 3 give two alternate descriptions of this twin. Orientations are not made explicit here since our main claim is simply that the twists invoved generate a certain

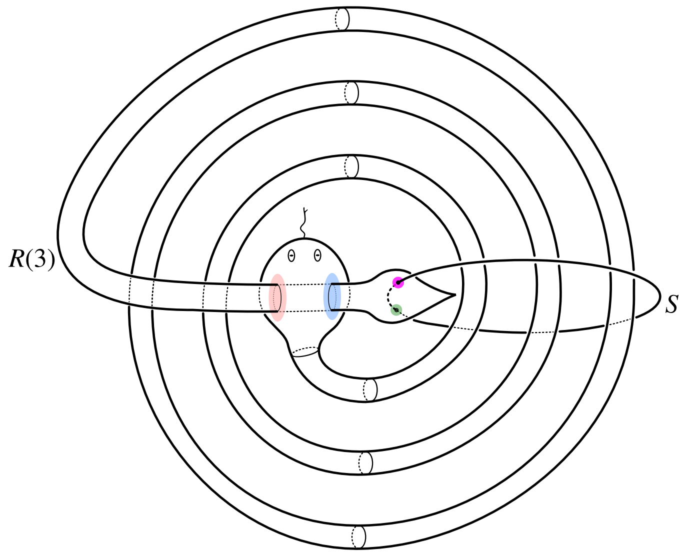

Figure 1: An illustration of $W(3) = (R(3), S)$. The picture mostly happens in the slice $\{t = 0\} \subset \mathbb{R}^4 = \{(x, y, z, t)\} \subset \mathbb{R}^4 \cup \{\infty\} = S^4$. The ring labelled S is a slice through S, which shrinks to a point as we move forwards and backwards in the "time" coordinate t. The "snake whose tail passes through his head" is $R(3)$, which is projected onto $\{t = 0\}$, intersecting itself along one circle in the middle of the red disk (the "snake's left ear hole") and along another circle in the middle of the blue disk (the "right ear hole"). Blue and red indicate that these disks are pushed slightly forwards (blue) and backwards (red) in time to resolve these intersections; otherwise $R(3)$ lies in the slice $\{t = 0\}$. The bright pink dot where the ring pierces the tail is the positive point of intersection of $R(3)$ and S; the faint green dot on the back side of the tail is the negative intersection point.

group, and the inverse of a generator is still a generator. Note that both $R(i)$ and S are individually unknotted 2–spheres, and that the twin $W(0)$ is the trivial "unknotted twin".

Our second result is:

Theorem 4 *The subgroup $\mathcal{H}_\infty(\imath_*(\pi_1(\widehat{\mathcal{S}}_\infty)))$ of $\pi_0(Diff^+(S^4))$ is generated by the twin twists $\{\tau_{W(i)} \mid i \in \mathbb{N}\}$.*

(In fact these automorphisms $\tau_{W(i)}$ are also examples of the "barbell maps" discussed in [2]; readers familiar with barbell maps should be able to use the description of $W(i)$ in Figure 3 to see the connection.)

Question 5 Is $[\tau_W] \in \mathcal{H}_\infty(\imath_*(\pi_1(\widehat{\mathcal{S}}_\infty)))$ for an arbitrary Montesinos twin? More generally, given any embedding of $S^1 \times \Sigma \hookrightarrow S^4$, for closed surface Σ, a tubular

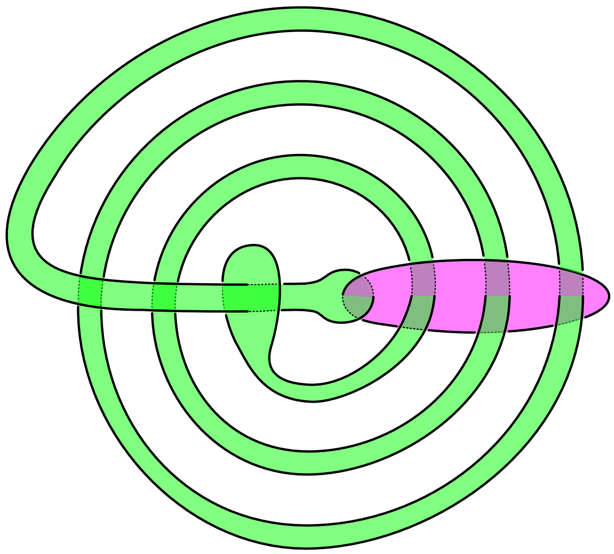

Figure 2: Another illustration of $W(3) = (R(3), S)$. Here we have drawn an immersed pair of disks, one green and one pink, with mostly ribbon intersections except for one nonribbon arc. Pushing these two disks into $\mathbb{R}^4_+$ or $\mathbb{R}^4_-$ and resolving ribbon intersections in the usual way gives two embedded disks intersecting each other transversely once, and then taking one copy in $\mathbb{R}^4_+$ and one in $\mathbb{R}^4_-$ glued along their common boundary, i.e. doubling the ribbon disks, gives $R(3)$ (green) and S (pink) in $\mathbb{R}^4 \subset S^4$.

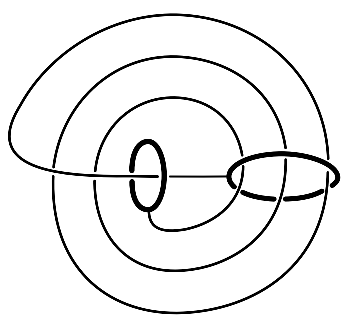

Figure 3: Yet another illustration of $W(3)$. Here we have drawn two disjoint, embedded 2–spheres in S^4 (the two thick circles, becoming 2–spheres when shrunk to points forwards and backwards in time) and an arc connecting them. Push a finger from one of the spheres out along this arc and then do a finger move when you encounter the other sphere, creating a pair of transverse intersections, and the result is $W(3)$.

neighborhood gives an embedding of $[-1,1]\times S^1\times\Sigma\hookrightarrow S^4$ which gives a diffeomorphism $\tau\times\mathrm{id}_\Sigma$, where τ is the Dehn twist on $[-1,1]\times S^1$. Are such diffeomorphisms in $\mathcal{H}_\infty(\imath_*(\pi_1(\widehat{\mathcal{S}}_\infty)))$?

One could try to answer these questions either through Cerf theory, by explicitly identifying a pseudoisotopy from a given diffeomorphism of S^4 to the identity, and then extracting a loop of attaching spheres for 5–dimensional 2–handles, or one could try to work explicitly with the diffeomorphisms in S^4 and try to find relationships amongst such twists, to relate them to twists along our standard Montesinos twins $W(i)$.

The bigger questions are the following, with affirmative answers to both showing that the smooth mapping class group of S^4 is trivial:

Question 6 Is $\mathcal{H}_\infty(\imath_*(\pi_1(\widehat{\mathcal{S}}_\infty)))$ trivial?

Theorem 4 could help prove this if one can exhibit explicit isotopies from $\tau_{W(i)}$ to id_{S^4}.

Question 7 Is $\mathcal{H}_\infty(\imath_*(\pi_1(\widehat{\mathcal{S}}_\infty))) = \pi_0(\mathit{Diff}^+(S^4))$?

Since $\imath_*(\pi_1(\mathcal{S}^0_\infty))$ is in the kernel of $\mathcal{H}_\infty$, we know that $\mathcal{H}_\infty$ factors through the quotient map $\pi:\pi_1(\mathcal{S}_\infty)\to\pi_1(\mathcal{S}_\infty)/\langle\imath_*(\pi_1(\mathcal{S}^0_\infty))\rangle$, where $\langle H\rangle$ denotes the normal closure of a subgroup H. Thus one way to show that $\mathcal{H}_\infty(\imath_*(\pi_1(\widehat{\mathcal{S}}_\infty))) = \pi_0(\mathit{Diff}^+(S^4))$ would be to show that $\pi\circ\imath:\pi_1(\widehat{\mathcal{S}}_\infty)\to\pi_1(\mathcal{S}_\infty)/\langle\imath_*(\pi_1(\mathcal{S}^0_\infty))\rangle$ is surjective. On the other hand this does not need to be true for the answer to this question to be "yes", since the kernel of $\mathcal{H}_\infty$ could presumably be much larger than $\langle\imath_*(\pi_1(\mathcal{S}^0_\infty))\rangle$.

In the next section we elaborate on the connection between loops of embeddings of certain spheres and self-diffeomorphisms of other spheres, setting up the general theory in various dimensions and codimensions. After that, we devote one section to the proof of Theorem 1, and we break the proof of Theorem 4 into the three remaining sections.

The author would like to thank Bruce Bartlett, Sarah Blackwell, Mike Freedman, Robert Gompf, Jason Joseph, Danica Kosanovic, Peter Lambert-Cole, Gordana Matic, Benjamin Ruppik, Rob Schneiderman, Peter Teichner and Jeremy Van Horn-Morris for helpful conversations along the way, and most especially Hannah Schwartz for pointing out Montesinos's work and Dave Gabai for initial inspiration and for pointing out the mistake in the first version that claimed far too much, and both Hannah and Dave for many clarifying conversations about loops of spheres.

Much of this work was carried out during the author's year at the Max Planck Institute for Mathematics in Bonn, and thus the author gratefully acknowledges the institute's support in the form of the 2019-2020 Hirzebruch Research Chair. This work was also supported by individual NSF grant DMS-2005554 and NSF Focused Research Group grant DMS-1664567.

2 From loops of spheres to diffeomorphisms

Here we describe the homomorphism $\mathcal{H}_\infty : \pi_1(\mathcal{S}_\infty) \to \pi_0(\mathit{Diff}^+(S^4))$ as a special case of a more general family of homomorphisms turning loops of framed embedded spheres in various dimensions into bundles of cobordisms and hence into self-diffeomorphisms of smooth manifolds. This should mostly be standard "Cerf theory". In the introduction above, we had 2–spheres embedded in 4–manifolds, but we did not mention framings of these 2–spheres. Below, we will work with framed spheres and then later when we relate this back to the terminology of the introduction, we will see where the framings come from.

Fix an m–manifold X, for some $m \geq 2$, fix integers $0 < k < m$ and $n \geq 0$, and let $\#^n(S^k \times S^{m-k})^\dagger$ denote a punctured $\#^n(S^k \times S^{m-k})$. As in the introduction, the puncture is needed so that we can view $\#^n(S^k \times S^{m-k})^\dagger$ as a subspace of $\#^{n+1}(S^k \times S^{m-k})^\dagger$. Now consider the following space of embeddings of collections of *framed* spheres:

$$\mathcal{FS}_n(X,k) = \mathit{Emb}(\amalg^n(S^k \times B^{m-k}), X\#(\#^n(S^k \times S^{m-k})^\dagger))$$

Picking a fixed point $p \in S^{m-k}$ and a disk neighborhood U of p parametrized as B^{m-k}, we get a natural basepoint $\amalg^n(S^k \times U) \subset X\#(\#^n(S^k \times S^{m-k})^\dagger)$ for $\mathcal{FS}_n(X,k)$, which we will again generally suppress in our notation, understanding that $\mathcal{FS}_n(X,k)$ is a pointed space.

Then we get a homomorphism

$$\mathcal{FH}_n : \pi_1(\mathcal{FS}_n(X,k)) \to \pi_0(\mathit{Diff}^+(X))$$

defined as follows. Represent an element of $\pi_1(\mathcal{FS}_n(X,k))$ by a 1–parameter family of embeddings $\beta_t : \amalg^n(S^k \times B^{m-k}) \hookrightarrow X\#(\#^n(S^k \times S^{m-k})^\dagger)$ with $\beta_0 = \beta_1 = \amalg^n(S^k \times U)$. We will use this to build a $(m+2)$–manifold Z_β which is a cobordism from $S^1 \times X$ to some X bundle over S^1, and the monodromy of this bundle will be $\mathcal{H}_n([\beta_t])$. To build Z_β, first let Y be $[0,1] \times X$ with n $(m+1)$–dimensional k–handles attached along n unlinked 0–framed unknotted S^{k-1}'s in a ball in X. Thus Y is a cobordism from X to $X\#(\#^n(S^k \times S^{m-k}))$. Now consider $S^1 \times Y$, which is a cobordism from $S^1 \times X$ to $S^1 \times X\#(\#^n(S^k \times S^{m-k}))$. Identifying S^1 with $[0,1]/1 \sim 0$ and using t as the S^1–coordinate, we can now use each β_t as the attaching map for n $(m+1)$–dimensional $(k+1)$–handles attached to $\{t\} \times X\#(\#^n(S^k \times S^{m-k})) \subset \{t\} \times Y$; the result is our construction of Z_β. Thus Z_β is a bundle over S^1, each fiber of which is a $(m+1)$–dimensional cobordism from X to an m–manifold X_t, built with n k–handles and n $(k+1)$–handles. Note that the $(k+1)$–handle attached along $\beta_0 = \beta_1$ is the canonical cancelling handle for the corresponding k–handle, and thus X_0 is canonically diffeomorphic to X. Therefore the full $(m+2)$–dimensional cobordism Z_β is a cobordism from $S^1 \times X$ to some X bundle over S^1 with a monodromy which is well-defined by the homotopy class of the loop of embeddings β_t. This defines $\mathcal{FH}_n$ and it is straightforward to see that $\mathcal{FH}_n$ is a group homomorphism.

Thanks to the punctures, we have basepoint preserving inclusions

$$\ldots \subset \mathcal{FS}_n(X,k) \subset \mathcal{FS}_{n+1}(X,k) \subset \ldots$$

and thus induced maps on π_1 and thus a direct limit

$$\ldots \to \pi_1(\mathcal{FS}_n(X,k)) \to \pi_1(\mathcal{FS}_{n+1}(X,k)) \to \ldots \to \pi_1(\mathcal{FS}_\infty(X,k))$$

Again, we do not really care about the limiting spaces, just the groups. Thus one should think of an element of $\pi_1(\mathcal{FS}_\infty(X,k))$ as an equivalence class of loops in some $\mathcal{FS}_n(X,k)$, where two such loops are equivalent if they become homotopic after including into some $\mathcal{FS}_N(X,k)$ for some $N \geq n$. It is not hard to see that these induced maps on π_1 commute with the $\mathcal{FH}_n$ homomorphisms, so that finally we get the limit homomorphism

$$\mathcal{FH}_\infty : \pi_1(\mathcal{FS}_\infty(X,k)) \to \pi_0(\mathit{Diff}^+(X))$$

As in the unframed setting of the introduction, we have two natural subspaces of $\mathcal{FS}_n(X,k)$: Let $\mathcal{FS}^0_n(X,k)$ denote those embeddings of $\amalg^n(S^k \times B^{m-k})$ into $X\#(\#^n(S^k \times S^{m-k}))$ for which the $S^k \times \{0\}$ in the i'th $S^k \times B^{m-k}$ transversely intersects the $\{p\} \times S^{m-k}$ in the j'th $S^k \times S^{m-k}$ summand transversely at δ_{ij} points. Let $\widehat{\mathcal{FS}}_n$ denote the subspace of embeddings with the property that the image of $\amalg^n(S^k \times B^{m-k})$ is disjoint from $\amalg^n(S^k \times \{p'\})$ for some fixed $p' \in S^{m-k} \setminus U$. Note that our basepoint lies in both of these subspaces.

It is standard that $\imath_*(\pi_1(\mathcal{FS}^0_n(X,k)))$ is in the kernel of $\mathcal{FH}_n$, because when the $S^k \times \{0\}$'s are dual to the $\{p\} \times S^{m-k}$'s for all t in a loop of embeddings α_t, then for all t the k–handles and $(k+1)$–handles cancel "uniquely" (this is Cerf's *l'unicité de mort* [3]). Thus the cobordism Z_α becomes a bundle over S^1 with each fiber supporting a Morse function without critical points, so that the monodromy at the top is id_X.

Now we discuss the relationship to spaces of spheres without framings. Let

$$\mathcal{S}_n(X,k) = \mathit{Emb}(\amalg^n S^k, X\#(\#^n(S^k \times S^{m-k})^\dagger))$$

There is an obvious "framing forgetting" map of pointed spaces

$$\mathcal{F} : \mathcal{FS}_n(X,k) \to \mathcal{S}_n(X,k)$$

given by restricting an embedding of $S^k \times B^{m-k}$ to $S^k = S^k \times \{0\}$. Palais [9] shows that such maps are locally trivial and thus satisfy the homotopy lifting property. The fiber of $\mathcal{F}$ is the space of framings of a fixed S^k, i.e. (up to homotopy) maps from S^k to $SO(m-k)$. Note that the fiber over the basepoint is actually a subspace of $\mathcal{FS}^0_n(X,k)$ and thus π_1 of the fiber lands in the kernel of $\mathcal{FH}_n$. As a consequence, even though

$$\mathcal{F}_* : \pi_1(\mathcal{FS}_n(X,k)) \to \pi_1(\mathcal{S}_n(X,k))$$

may not be injective, if the fiber is not simply connected, we still have that $\mathcal{FH}_n$ induces a well-defined homomorphism $\mathcal{H}_n$ from the image $\mathcal{F}_*(\pi_1(\mathcal{FS}_n(X,k)))$ of $\mathcal{F}_*$ in $\pi_1(\mathcal{S}_n(X,k))$ to $\pi_0(\mathit{Diff}^+(S^m))$. All of this also commutes with the inclusion maps from n to $n+1$ giving $\mathcal{H}_\infty$.

When $m=4$ and $k=2$, the fibers of $\mathcal{F}$ are connected, i.e. a sphere in a 4–manifold with self-intersection 0 has only one framing up to isotopy (because $\pi_2(SO(2)) = 0$). Therefore

$$\mathcal{F}_* : \pi_1(\mathcal{FS}_n(S^4,2)) \to \pi_1(\mathcal{S}_n(S^4,2) = \mathcal{S})$$

is surjective and so $\mathcal{FH}_n$ and $\mathcal{FH}_\infty$ induce homomorphisms $\mathcal{H}_n$ and

$$\mathcal{H}_\infty : \pi_1(\mathcal{S}_\infty) \to \pi_0(\mathit{Diff}^+(S^m))$$

3 Surjectivity of $\mathcal{H}_\infty$

In this section we will prove Theorem 1, i.e. that $\mathcal{H}_\infty : \pi_1(\mathcal{S}_\infty) \to \pi_0(\mathit{Diff}^+(S^4))$ is surjective. (We have already seen in the previous section that the kernel of $\mathcal{H}_\infty$ contains $\imath_*(\pi_1(\mathcal{S}^0_\infty))$.) Since $\mathcal{H}_\infty$ is a limit of maps $\mathcal{H}_n$, what we are really trying to prove is the following:

Theorem 8 *For any orientation preserving diffeomorphism $\phi : S^4 \to S^4$, there is some $n \in \mathbb{N}$ and some loop α_t in $\mathcal{S}_n$ based at the basepoint $\amalg^n(S^2 \times \{p\})$, such that $\mathcal{H}_n([\alpha_t]) = [\phi] \in \pi_0(\mathit{Diff}^+(S^4)$.*

In other words, once we prove this theorem then we have:

Proof of Theorem 1 This is an immediate corollary to Theorem 8 and the discussion in the previous section. □

In fact, to make things more concrete, we will prove this:

Theorem 9 *For any orientation preserving diffeomorphism $\phi : S^4 \to S^4$, there is a 6–manifold Z with the following properties:*

(1) *Z is a bundle over S^1 and is a cobordism from $S^1 \times S^4$ to $S^1 \times_\phi S^4 = ([0,1] \times S^4)/(1,p) \sim (0, \phi(p))$, with the three fibrations over S^1, of Z, $S^1 \times S^4$ and $S^1 \times_\phi S^4$, being compatible.*

(2) *For each $t \in S^1$, the fiber Y_t over t in Z, being a 5–dimensional cobordism from S^4 to a 4–manifold X_t which is diffeomorphic to S^4, comes with a handle decomposition with n 2–handles and n 3–handles.*

(3) *The n 2–handles for each Y_t are attached along a fixed standard collection of n standardly framed S^1's in S^4, so that the 4–manifold immediately above these 2–handles is canonically identified with $\#^n(S^2 \times S^2)$*

(4) *The n 3–handles for each Y_t are attached along a moving family of n framed S^2's in $\#^n(S^2 \times S^2)$ such that, ignoring the framings, these S^2's are given by a loop of embeddings of $\amalg^n S^2$ into $\#^n(S^2 \times S^2)$ all missing a single fixed point and starting and ending at the standard embedding $\amalg^n(S^2 \times \{p\})$, i.e. a based loop α_t in $\mathcal{S}_n$.*

Then:

Proof of Theorem 8 This is an immediate corollary of Theorem 9, since the homomorphism $\mathcal{H}_n$ comes precisely from constructions of cobordisms as in Theorem 9. □

We will use Cerf theoretic techniques, beginning with a pseudoisotopy.

Lemma 10 *Every orientation preserving self-diffeomorphism of S^4 is pseudoisotopic to the identity.*

Proof Consider an orientation preserving diffeomorphism $\phi : S^4 \to S^4$. Let $f : S^5 \to \mathbb{R}$ be projection onto the last coordinate in $\mathbb{R}^6$, and for any interval $I \subset \mathbb{R}$, let $S^5_I = f^{-1}(I)$. Let V be a smooth vector field on $S^5 \setminus \{(0,0,0,0,0,\pm 1)\}$ which is orthogonal to all level sets of f and scaled so that $df(V) = 1$. Let $X = S^5_{[-1,0]} \cup_\phi S^5_{[0,1]}$, where $\phi : \partial S^5_{[0,1]} = S^4 \to -S^4 = \partial S^5_{[-1,0]}$ is now seen as an orientation reversing gluing diffeomorphism. Arrange the gluing (i.e. the smooth structure on X) so that the vector field V on the two halves of X is still a smooth vector field on X, which we call V_X. Note that X also inherits the Morse function f, which we label $f_X : X \to \mathbb{R}$, and we can use the same notation $X_I = f_X^{-1}(I) \subset X$. The point is that if $I \subset (-\infty, 0]$ or if $I \subset [0, \infty)$ then $X_I = S^5_I$, i.e. they are actually equal sets, not just diffeomorphic manifolds.

Now note that X is homotopy equivalent to S^5 and therefore [6, 7] diffeomorphic to S^5. For some small $\epsilon > 0$ we can assume that the diffemorphism $\Phi : S^5 \to X$ is the identity on $S^5_{[-1,-1+\epsilon]} = X_{[-1,-1+\epsilon]}$ and on $S^5_{[1-\epsilon,1]} = X_{[1-\epsilon,1]}$. Using flow along V and V_X, respectively, and the standard identification of $\partial S^5_{[-1,-1+\epsilon]}$ with S^4, we can parametrize both $S^5_{[-1+\epsilon,1-\epsilon]}$ and $X_{[-1+\epsilon,1-\epsilon]}$ as $[-1+\epsilon, 1-\epsilon] \times S^4$ and then Φ restricts to give a diffeomorphism from $[-1+\epsilon, 1-\epsilon] \times S^4$ to itself. Furthermore, this map is the identity on $\{-1+\epsilon\} \times S^4$ and by continuity must equal ϕ on $\{1-\epsilon\} \times S^4$. After reparametrizing $[-1+\epsilon, 1-\epsilon]$ as $[0, 1]$ we get the desired pseudoisotopy.

□

Now we begin the Cerf theory.

Proof of Theorem 9 Given the orientation preserving diffeomorphism $\phi : S^4 \to S^4$, let $\Phi : [0,1] \times S^4 \to [0,1] \times S^4$ be a pseudoisotopy from the identity to S^4, i.e. $\Phi(0,p) = (0,p)$ and $\Phi(1,p) = (1,\phi(p))$. Let $f_0 : [0,1] \times S^4 \to [0,1]$ be projection onto the first factor, let V_0 be the unit vector field on $[0,1] \times S^4$ in the $[0,1]$ direction, let $(f_1, V_1) = \Phi^*(f_0, V_0) = (f_0 \circ \Phi, D\Phi^{-1}(V_0))$, and let (f_t, V_t) be a generic homotopy of Morse functions with gradient-like vector fields from (f_0, V_0) to (f_1, V_1). Hatcher and Wagoner (Chapter VI, Proposition 3, page 214 of [5]) show that this family of functions with gradient-like vector fields can be homotoped rel $t \in \{0, 1\}$ so as to arrange the following properties (see Figure 4):

- The only Morse critical points of any f_t are critical points of index 2 and 3.
- None of the functions f_t have two critical points mapping to the same critical value.
- All critical points of index 2 stay below all critical points of index 3. In other words, for every t, if p is a critical point of index 2 for f_t and q is a critical point of index 3 for f_t then $f_t(p) < f_t(q)$.
- There are no handle slides. In other words, none of the vector fields V_t have flow lines between critical points of the same index.
- All births of cancelling pairs of critical points happen before all deaths of cancelling pairs, and no two births or deaths happen at the same time.

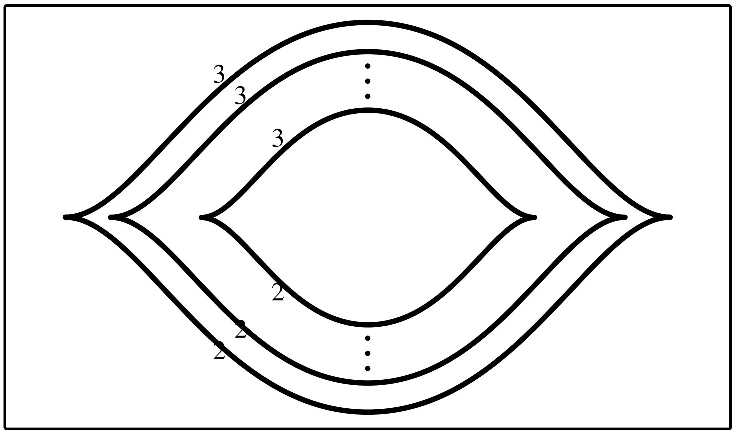

Figure 4: A nice Cerf graphic with only critical points of index 2 and 3.

- At the moments of birth and death, the handle pair dying or being born does not run over any other handles. In other words, there are no V_t flow lines between a non-Morse birth/death critical point and any other critical point.

Now consider the 6–manifold $Z = [0,1] \times [0,1] \times S^4/(1,s,p) \sim (0,\Phi(s,p))$. We will show that this has the properties advertised in the statement of Theorem 9. Currently we have the first property, namely that Z is a bundle over S^1 and is a cobordism from $S^1 \times S^4$ to $S^1 \times_\phi S^4 = ([0,1] \times S^4)/(1,p) \sim (0,\phi(p))$, with the three fibrations over S^1, of Z, $S^1 \times S^4$ and $S^1 \times_\phi S^4$, being compatible. Each fiber Y_t of $Z \to S^1$ gets a handle decomposition from (f_t, V_t) (allowing for birth and death handle decompositions) but it will take some work to arrange that these satisfy the remaining properties as stated in Theorem 9. First let us characterize each fiber Y_t, with its given handle decomposition, as much as possible in terms of handle attaching data in $\{t\} \times S^4 \subset S^1 \times S^4$.

Suppose that the births in our Morse functions f_t occur at times $0 < t_1 < t_2 < \ldots < t_n < 1/4$ and that the deaths occur at times $3/4 < t'_n < \ldots < t'_2 < t'_1 < 1$. We will show how each cobordism Y_t can be described as built from $\{t\} \times [0,1] \times S^4$ according to handle attaching data, and to do this we establish a few conventions. First, all of our embeddings of spheres and disks of various dimensions are framed embeddings, but to keep the terminology minimal we will sometimes suppress mention of framings. Later we will make sure to carefully relate this back to the case of loops of unframed spheres. Second, given a framed embedding ϵ of a sphere in a manifold X, let $X(\epsilon)$ denote the result of surgering X along ϵ. Third, a framed embedding δ of a disk D^k into an n–manifold X gives us several auxiliary pieces of information: We get a framed embedding α of $S^{k-1} = \partial D^k$ into X. We get the surgered manifold $X(\alpha)$. And, in $X(\alpha)$, there is a natural framed embedding β of S^k into $X(\alpha)$ which coincides with δ on the part of $X(\alpha)$ away from the surgery, and which coincides with the co-core of the surgery inside the surgered region. This is the basic model for the result of attaching a $(n+1)$–dimensional cancelling k–$(k+1)$ handle pair to $[0,1] \times X$, with the attaching data for the pair being described entirely in X by the framed disk δ. Thus we also see that $X(\alpha)(\beta)$ is canonically identified with X. Here we will work with the case $k = 2$ and $k+1 = 3$, but later in the paper we will be interested in the case $k = 1$ and $k+1 = 2$.

The notation introduced above also makes sense when the δ's, α's or β's are embeddings of disjoint unions of disks or spheres. Using this, we now describe the form of the explicit t–varying handle attaching data that gives the fiberwise construction of Z, i.e. the data that shows how to construct each Y_t starting with $\{t\} \times [0,1] \times S^4$ and attaching various 5–dimensional handles.

- For $0 \leq t < t_1$, no handles are attached, i.e. $Y_t = \{t\} \times [0,1] \times S^4$.
- For $t = t_1$, there is a framed embedding δ_1 of a disk D^2 into S^4 such that Y_{t_1} is built from $\{t_1\} \times [0,1] \times S^4$ by attaching a cancelling pair of a 5–dimensional 2–handle and 3–handle, with the 2–handle attached along the framed embedding $\alpha^1_{t_1} = \delta_1|_{S^1}$, and with the 3–handle attached along the resulting framed embedding $\beta^1_{t_1}$ of S^2 into $S^4(\alpha^1_{t_1})$.
- For $t_1 \leq t \leq t_2$, we have a t–varying framed embedding $\alpha_t = \alpha^1_t$ of S^1 into S^4 and a t–varying framed embedding $\beta_t = \beta^1_t$ of S^2 into $S^4(\alpha_t)$. For $t_1 \leq t < t_2$, Y_t is built from $\{t\} \times [0,1] \times S^4$ by attaching a 2–handle along α_t and then a 3–handle along β_t. At $t = t_1$, the α_t and β_t agree with the α^1_t and β^1_t from the previous point.
- For $t = t_2$, there is a framed embedding δ_2 of D^2 into S^4 disjoint from the images of $\alpha^1_{t_2}$ and $\beta^1_{t_2}$, such that Y_{t_2} is built from $\{t_2\} \times [0,1] \times S^4$ by attaching
  - first a 5–dimensional 2–handle along $\alpha^1_{t_2}$,
  - then a cancelling 2–3–pair in which the 2–handle is attached along $\alpha^2_{t_2} = \delta_2|_{S^1}$ and the 3–handle is attached along the resulting framed 2–sphere $\beta^2_{t_2}$ in $S^4(\alpha^1_{t_2})(\alpha^2_{t_2})$, and
  - then a 3–handle attached along $\beta^1_{t_2}$, which can be seen as a framed 2–sphere in $S^4(\alpha^1_{t_2})$, in $S^4(\alpha^1_{t_2})(\alpha^2_{t_2})$ or in $S^4(\alpha^1_{t_2})(\alpha^2_{t_2})(\beta^2_{t_2}) \cong S^4(\alpha^1_{t_2})$.
- For $t_2 \leq t \leq t_3$, we have a t–varying framed embedding $\alpha_t = \alpha^1_t \amalg \alpha^2_t$ of $S^1 \amalg S^1$ into S^4 and a t–varying framed embedding $\beta_t = \beta^1_t \amalg \beta^2_t$ of $S^2 \amalg S^2$

into $S^4(\alpha_t)$, agreeing with the $\alpha^1_{t_2}$, $\alpha^2_{t_2}$, $\beta^2_{t_2}$ and $\beta^1_{t_2}$ of the preceding point when $t = t_2$, so that Y_t is built from $\{t\} \times [0,1] \times S^4$ by attaching 2–handles along α_t and then 3–handles along β_t.

- This process continues with each birth at time t_i governed by a new framed disk δ_i, generating a new framed S^1 $\alpha^i_{t_i}$ and a new framed S^2 $\beta^i_{t_i}$, which then join the previous framed spheres to create $\alpha_t = \alpha^1_t \amalg \ldots \amalg \alpha^i_t$ and $\beta_t = \beta^1_t \amalg \ldots \amalg \beta^i_t$, which are the attaching spheres for 2– and 3–handles for $t_t \leq t < t_{i+1}$.
- Reversing time we see the deaths governed by (most likely quite different) disks $\delta'_n, \ldots, \delta'_1$ and the same pattern of framed S^1's and S^2's in between these times.
- For $t_n \leq t \leq t'_n$, there is a t–parametrized family α_t of framed embeddings of $\amalg^n S^1$ into S^4 and a t–parametrized family β_t of framed embeddings of $\amalg^n S^2$ into $S^4(\alpha_t)$ which constitute the attaching data for the n 2–handles and n 3–handles used to construct each Y_t in this range.

We will now improve the format of this data somewhat. First, we can arrange that for some small $\epsilon > 0$, on the time interval $t_1 \leq t \leq t_1 + \epsilon$ the embeddings α_t and β_t are independent of t, i.e. the first framed circle and sphere do not move for a short time after their birth. Next, since a birth happens at a point, we can make the second birth happen at an earlier time so that in fact $t_1 < t_2 < t_1 + \epsilon$. Repeating this, and doing the same in reverse with the deaths, we can assume that on the whole interval $t_1 \leq t \leq t_n$ and on the whole interval $t'_n \leq t \leq t'_1$, the α_t and β_t are independent of t except for the fact that at each t_i a new α^i_t and β^i_t is added to the mix (and ditto for the deaths).

Thus now the governing data for the constructions of each Y_t can be more succinctly described by the following data:

- A framed embedding $\delta = \delta_1 \amalg \ldots \amalg \delta_n$ of $\amalg^n D^2$ into S^4 (for the births), defining framed embeddings α of $\amalg^n S^1$ into S^4 and β of $\amalg^n S^2$ into $S^4(\alpha)$.
- A framed embedding $\delta' = \delta'_1 \amalg \ldots \amalg \delta'_n$ of $\amalg^n D^2$ into S^4 (for the deaths), defining framed embeddings α' of $\amalg^n S^1$ into S^4 and β' of $\amalg^n S^2$ into $S^4(\alpha')$
- A t–parameterized family, for $t \in [1/4, 3/4]$, of framed embeddings $\alpha_t = \alpha^1_t \amalg \ldots \amalg \alpha^n_t$ of $\amalg^n S^1$ into S^4, with $\alpha_{1/4} = \alpha$ and $\alpha_{3/4} = \alpha'$.
- A t–parameterized family, for $t \in [1/4, 3/4]$, of framed embeddings $\beta_t = \beta^1_t \amalg \ldots \amalg \beta^n_t$ of $\amalg^n S^2$ into $S^4(\alpha_t)$, with $\beta_{1/4} = \beta$ and $\beta_{3/4} = \beta'$.

We would now like to modify this data (by a homotopy of the family (f_t, V_t)) so that $\delta = \delta'$ is a fixed standard embedding and so that the family α_t is actually invariant with respect to t, i.e. $\alpha_t = \alpha = \alpha'$ for all t. Since any collection of n disks is isotopic to any other, it is easy to arrange that $\delta = \delta'$ (we tack the necessary isotopies on at the beginning and end of the t–parameterized family of handle attaching data). Thus we now assume $\delta = \delta'$ is standard.

Next focus on α_t^1, which is a loop of embeddings of S^1 starting and ending at a standard embedding. By an isotopy rel endpoints we can arrange that α_t^1 is independent of t at a fixed point on S^1 (since $\pi_1(S^4) = 0$), and then on a fixed interval neighborhood of that point (since $\pi_1(S^3) = 0$). These are achieved by homotopies of the family of embeddings of S^1, i.e. isotopies of isotopies, and can be extended to all of S^4 by the parametrized isotopy extension theorem, so that the other α_t^i's move out of the way when we move α_t^1. Also, when we move the α_t^i's, we can realize this as a homotopy of (f_t, V_t) by just modifying the gradient-like vector field V_t in $f_t^{-1}[a, b]$ for some small $0 < a < b$ with b below the lowest critical value of all the f_t's. Thus we do not need to worry about how this affects the framed 2–spheres β_t, they will still be a family as above of framed embeddings of $\amalg^n S^2$ into $S^4(\alpha_t)$. Now we use the fact that the space of embeddings of B^1 in B^4 with fixed endpoints on ∂B^4 is simply connected [1] to finally arrange that α_t^1 is independent of t, i.e. that $\alpha_t^1 = \alpha_{1/4}^1 = \alpha_{3/4}^1$.

We could proceed to do the same thing for α_t^2 if we knew that the parametrized isotopy extensions involved when fixing α_t^2 (and after that, α_t^3 up to α_t^n) would not mess up the work we have already done to fix α_t^1. Recall that δ_1 is a fixed disk bounded by the (now constant in t) circle α_t^1; if the other circles $\alpha_t^2, \ldots, \alpha_t^n$ do not pass through δ_1, then we can arrange that the isotopy extensions do not need to move δ_1 and hence do not need to move α_t^1. (This is because the isotopies of isotopies of circles involved are ultimately 3–dimensional objects and can be arranged to miss points, and hence disks. They cannot be assumed in general to miss circles.) To deal with this issue, let δ_t^1 be a family of framed embeddings of D^2 into S^4, for $t \in [1/4, 3/4]$, with $\partial \delta_t^1 = \alpha_t^1$ and $\delta_{1/4}^1 = \delta_1$, such that δ_t^1 is disjoint from $\alpha_t^2 \amalg \ldots \amalg \alpha_t^n$ for all $t \in [1/4, 3/4]$. This is also achieved by an isotopy extension, simply pushing δ_1 out of the way as the S^1's $\alpha_t^2 \amalg \ldots \amalg \alpha_t^n$ move around and pass through δ_1. Note that there is no reason to expect that $\delta_{3/4}^1 = \delta_1$. However, in the time interval $[t_2', t_1']$, after the death of all of the cancelling handle pairs involving $\alpha_t^2, \ldots, \alpha_t^n$, we can reverse the isotopy of the δ_t^1's and get back to $\delta_{t_1}^1 = \delta_1$. Finally, looking at the entire family of disk embeddings δ_t^1 from $t = 1/4$ to $t = t_1'$, this is homotopic rel endpoints to a constant family of embeddings $\delta_t^1 = \delta_1 = \delta_1'$, and this homotopy, i.e. isotopy of isotopies, extends to a parametrized ambient isotopy which again moves $\alpha_t^2, \ldots, \alpha_t^n$ for $t \in [1/4, 3/4]$. Note that we can assume the disks $\delta_t^2, \ldots, \delta_t^n$ are not moved by this ambient isotopy again because they are disks and hence can be avoided by the isotopy of the family of disks δ_t^1.

Thus we get $\alpha_t^1 = \alpha_{1/4}^1 = \alpha_{3/4}^1 = \partial \delta_1$, and that $\alpha_t^2 \amalg \ldots \alpha_t^n$ is disjoint from δ_1

for all t. Now we can repeat this with α_t^2 up to α_t^n. This gets us to our desired situation where $\delta = \delta'$, $\alpha_t = \alpha_{1/4} = \alpha_{3/4} = \partial\delta$, and δ is a standard framed embedding of $\amalg^n D^2$ into S^4. With this setup, we can now move the birth of the first 2–3–pair closer to $t = 0$ and the death closer to $t = 1$ until they merge in $Z = [0,1] \times [0,1] \times S^4/(1,s,p) \sim (0,\Phi(s,p))$, and repeat with the second pair and so on. After this, Y_0 is built from $[0,1]\times S^4$ by attaching n standard cancelling 2–3–pairs and Y_t is built by attaching n 2–handles along a standardly embedded collection of n framed circles α and n 3–handles along a loop of framed embeddings β_t of $\amalg^n S^2$. This is precisely the conclusion of the theorem.

□

4 Turning 2–3–handle pairs into 1–2–handle pairs

We now need to work toward the connection with Montesinos twins and the proof of Theorem 4, that twists along Montesinos twins generate the subgroup of $\pi_0(\mathit{Diff}^+(S^4))$ corresponding, via the 2–3–handle pair construction above, to loops of 2–spheres which remain disjoint from parallel copies of the basepoint 2–spheres. More precisely, recall that $\widehat{\mathcal{S}}_n$ is the space of embeddings of $\amalg^n S^2$ into $\#^n(S^2 \times S^2)^\dagger$ which remain disjoint from $\amalg^n(S^2 \times \{p'\}$ for some fixed $p' \in S^2$, with basepoint being $\amalg^n(S^2 \times \{p\})$ for $p \neq p'$. We want to study the subgroup $\mathcal{H}_\infty(\imath_*(\pi_1(\widehat{\mathcal{S}}_\infty)))$. To relate this to Montesinos twins we will first need to relate it to families of 1–2–handle pairs coming from loops of framed circles in $\#^n(S^1 \times S^3)$. In particular, using the notation from Section 2, in this section we will prove:

Theorem 11 *For any n, $\mathcal{H}_n(\imath_*(\pi_1(\widehat{\mathcal{S}}_n))) = \mathcal{F}\mathcal{H}_n(\pi_1(\mathcal{F}\mathcal{S}_n(S^4,1)))$.*

This means that any isotopy class of diffeomorphisms of S^4 that can be realized by a family of cobordisms built with n 2–3–handle pairs governed by a loop in $\widehat{\mathcal{S}}_n$ can also be realized by a family of cobordisms built with n 1–2–pairs governed by a loop in $\mathcal{F}\mathcal{S}_n(S^4,1)$, which is the space of framed embeddings of a disjoint union of n circles in $\#^n(S^1 \times S^3)^\dagger$.

Proof Consider a cobordism Z from $S^1 \times S^4$ to $S^1 \times_\phi S^4$ built as before as a family Y_t of cobordisms, such that each Y_t is built by attaching n 2–handles to $[0,1] \times S^4$ and then n 3–handles to the result. The 2–handles are attached along a family α_t of framed embeddings of $\amalg^n S^1$ into S^4, except that in fact $\alpha_t = \alpha_0$ does not vary with t and is a standard embedding, so that $S^4(\alpha_t)$ is canonically identified with $\#^n(S^2 \times S^2)$. The 3–handles are attached along a family β_t of framed embeddings of $\amalg^n S^2$ into $\#^n(S^2 \times S^2)$, with $\beta_0 = \beta_1 = \amalg^n(S^2 \times \{p\})$, and with each β_t disjoint from $\amalg^n(S^2 \times \{p'\})$.

We will now introduce two 5–dimensional analogues of the "dotted circle" notation traditionally used for 4–dimensional 1–handles. A 5–dimensional 1–handle attached to the boundary of a 5–manifold in such a way that the 1–handle *could be cancelled* (which here means that its attaching S^0 bounds a B^1 and that the framing of the S^0 comes from a framing of the B^1) can be represented by a "dotted unknotted 2–sphere" *together with a choice of* 3*–ball bounded by the* 2*–sphere* in the 4–manifold, which means we push the interior of the 3–ball into the 5–manifold and carve out its neighborhood. The B^3 bounded by the dotted S^2 intersects the B^1 bounded by the 1–handle's attaching S^0 transversely once. Likewise, a 5–dimensional 2–handle attached to the boundary of a 5–manifold in such a way that it *could be cancelled* (which here means that its attaching framed S^1 bounds a framed B^2) can be represented by a "dotted circle" *together with a choice of* 2*–disk bounded by the circle* in the 4–manifold. Here the B^2 bounded by the dotted circle intersects the B^2 bounded by the 2–handle's attaching circle transversely once.

We emphasize in both these 5–dimensional cases that, using "dotted" notation, in principle we need to know what the disks/balls are which we are going to push in and carve out, with the 4–dimensional case being perhaps exceptional because an unknot in a S^3 bounds a unique disk. Budney and Gabai have shown [2] that unknotted 2–spheres in S^4 can bound "knotted" 3–balls, and of course, although all S^1's in S^4 are unknotted, 2–knots can be tied into any spanning disk for such an S^1. What really matters is whether these spanning disks and balls are isotopic in dimension 5, and we leave this as an interesting question; the Budney-Gabai examples are in fact isotopic in B^5, but there might in principle be more complicated examples that remain nonisotopic even when pushed into B^5. However, here we will bypass this issue by working with "dotted disks" instead of "dotted circles" and "dotted balls" instead of "dotted spheres".

Now, given our handle attaching data α_t and β_t used to build Z, since the α_t's are invariant in t, and $\alpha_t = \alpha_0$ bounds a fixed collection of framed disks δ_0, we can instead represent the 2–handles by a (for now, t–invariant) t–parametrized family of n dotted disks $\alpha_t^\bullet : \amalg^n B^2 \hookrightarrow S^4$. Note that $\alpha_t^\bullet$ is not an extension of α_t, but is rather an embedding of dual disks to the fixed disks bounded by α_t. Also, we insist on maintaining the subscript t even though these are t–invariant because we will shortly modify the family so as to lose t–invariance. Now the instructions for building the cobordism Z are to build each Y_t from $[0,1] \times S^4$ by pushing the interior of each of these disks from $\{1\} \times S^4$ into the interior of $[0,1] \times S^4$ and removing their neighborhoods, and then attaching 3–handles along β_t. Note that, after carving out the disks but before attaching the 3–handles, the upper boundary of this cobordism Y_t is the surgered 4–manifold $S^4(\partial\alpha_t^\bullet)$. In other words, when looking at the 4–dimensional boundary, we cannot tell whether we carved out the dotted disks or attached 2–handles along their boundaries, because the resulting surgeries are the same.

The crucial point here is that, because each β_t is disjoint from $\amalg^n(S^2 \times \{p'\})$, we can isotope the family β_t so that it never goes over the surgered region of $S^4(\partial\alpha_t^\bullet)$, and thus the entire handle attaching data now lives in S^4. Thus we can now describe each Y_t, and thus Z, via data entirely lying in S^4, i.e. $\alpha_t^\bullet \amalg \beta_t : (\amalg^n B^2) \amalg (\amalg^n S^2) \to S^4$. This is not an embedding but is an embedding when restricted to $\amalg^n B^2$ and to $\amalg^n S^2$, and the only intersections occur between $\amalg^n S^2$ and the interiors of the disks $\amalg^n B^2$. At times $t = 0$ and $t = 1$, each S^2 intersects its corresponding B^2 transversely once and is disjoint from all the other B^2's, i.e. the spheres and disks are in "cancelling position"

Our goal is now to "switch the dots from the circles to the spheres", i.e. to think of β_t as being dotted spheres, thus corresponding to 5–dimensional 1–handles, and to think of $\partial\alpha_t^\bullet$ as attaching circles for 2–handles, rather than dotted circles describing 2–handles. However, as discussed above, before we can "put dots on" β_t we need to extend them to be embeddings of balls not just of spheres. In other words, we want to extend $\beta_t : \amalg^n S^2 \hookrightarrow S^4$ to $\beta_t^\bullet : \amalg^n B^3 \hookrightarrow S^4$. Unfortunately, if we could do this without moving β_t and also achieving the property that $\beta_0^\bullet = \beta_1^\bullet$, we would then have shown that β_t was trivial in $\pi_1(Emb(\amalg^n S^2, S^4))$, and we do not (at least this author does not) know enough about $\pi_1(Emb(\amalg^n S^2, S^4))$ to make such an assertion. So we proceed carefully, and in fact we will end up modifying both $\alpha_t^\bullet$ and β_t, abandoning $\alpha_1^\bullet = \alpha_0^\bullet$ and $\beta_1 = \beta_0$, but maintaining the property that $\alpha_1^\bullet$ and β_1 are in cancelling position.

Using the isotopy extension theorem we can easily extend β_t to $\beta_t^\bullet : \amalg^n B^3 \hookrightarrow S^4$ (pick B^3's bounded by β_0 and move them around by an ambient isotopy for β_t) but we should not expect that $\beta_1^\bullet = \beta_0^\bullet$. We can, however, assume that each component of $\partial\alpha_1^\bullet$ transversely intersects the corresponding component of $\beta_1^\bullet$ at one fixed point in the interior of each B^3, since β_1 and $\partial\alpha_1^\bullet$ are meridians of each other. Thus we know that $\partial\alpha_1^\bullet$ and $\beta_1^\bullet$ are now in cancelling position, and we still have that $\alpha_1^\bullet$ and $\partial\beta_1^\bullet$ are in cancelling position. Now assume that $\beta_t^\bullet = \beta_1^\bullet$ and $\alpha_t^\bullet = \alpha_1^\bullet$ for $t \in [1-\epsilon, 1]$. Then there is an isotopy from $\beta_{1-\epsilon}^\bullet$ back to $\beta_0^\bullet$ (one could use the original forward isotopy in reverse, or any other isotopy) which can be extended to an ambient isotopy and then used on the interval $[1-\epsilon, 1-\epsilon/2]$ to move both $\beta_t^\bullet$ and $\alpha_t^\bullet$, discarding completely the given $\beta_t^\bullet$ and $\alpha_t^\bullet$ on $[1-\epsilon, 1]$. However, although we have thrown away the original path of embeddings on this interval, we have maintained that fact that both the pair $\partial\beta_t^\bullet$

and $\alpha_t^\bullet$ and the pair $\beta_t^\bullet$ and $\partial\alpha_t^\bullet$ are in cancelling position. Now, when $t = 1 - \epsilon/2$, we have that $\beta_t^\bullet = \beta_0^\bullet$ but the only thing we know about $\partial\alpha_t^\bullet$ is that it is in cancelling position with respect to $\beta_t^\bullet = \beta_0^\bullet$, and also that $\partial\alpha_0^\bullet$ is *also* in cancelling position with respect to $\beta_t^\bullet = \beta_0^\bullet$. But now we can use Gabai's 4–dimensional lightbulb theorem to isotope $\partial\alpha_{1-\epsilon/2}^\bullet$ to $\alpha_0^\bullet$ remaining in cancelling position, and use this isotopy to get $\partial\alpha_t^\bullet$ for $t \in [1 - \epsilon/2, 1$. Note that we have lost track of, but do not really care about, the embeddings $\alpha_t^\bullet$ for $t \in [1 - \epsilon/2, 1]$, but care only about their boundaries.

To translate the above modifications into a modification of our cobordism Z built as a family of cobordisms Y_t, we first cancel the 2– and 3–handles for all $t \in [1 - \epsilon, 1]$, then switch the dots from the α's to the β's for all $t \in [0, 1 - \epsilon]$, thus getting a new cobordism Z' from $S^1 \times S^4$ to $S^1 \times_\phi S^4$. This Z' is built from cobordisms Y_t' which start with $[0, 1] \times S^4$, then experience the birth of n 1–2 pairs which then evolve in t, with 1–handles described by the dotted balls $\beta_t^\bullet$ and the 2–handles attached along the circles $\partial\alpha_t^\bullet$, and all end up cancelling at time $t = 1 - \epsilon$. Then the cancellation can be postponed until $t = 1 - \epsilon/2$ to give the circles $\partial\alpha_t^\bullet$ to return to their starting positions. Finally, wrapping from $t = 1 - \epsilon/2$ around to $t = 1 \sim 0$, we can merge the deaths and births so that in the end the cobordisms Y_t are described by the loop of maps $\partial\alpha_t^\bullet \amalg \beta_t^\bullet : (\amalg^n S^1) \amalg (\amalg^n B^3) \to S^4$, with $\beta_t^\bullet$ being dotted balls describing 1–handles and $\partial\alpha_t^\bullet$ being attaching circles for 2–handles.

Finally, a further ambient isotopy can now be used to arrange that $\beta_t^\bullet$ is independent of t, and thus the entire construction is governed by the loop of *framed* circles $\partial\alpha_t^\bullet$.

□

5 From many 1–2 pairs to a single 1–2 pair

We now know that diffeomorphisms of S^4 that can be realized by a family of cobordisms built with n 2–3–handle pairs can also be realized by a family of cobordisms built with n 1–2–handle pairs, as long as the original governing loop of embeddings of $\amalg^n S^2$ into $\#^n(S^2 \times S^2)$ lies in $\widehat{\mathcal{S}_n}$. Before we get to Montesinos twins, we need now to show that every diffeomorphism of S^4 that can be a realized by a family of cobordisms built with n 1–2–handle pairs can be realized by family built with a single 1–2–handle pair.

Theorem 12 *For any n, $\mathcal{FH}_n(\pi_1(\mathcal{FS}_n(S^4, 1))) = \mathcal{FH}_1(\pi_1(\mathcal{FS}_1(S^4, 1)))$.*

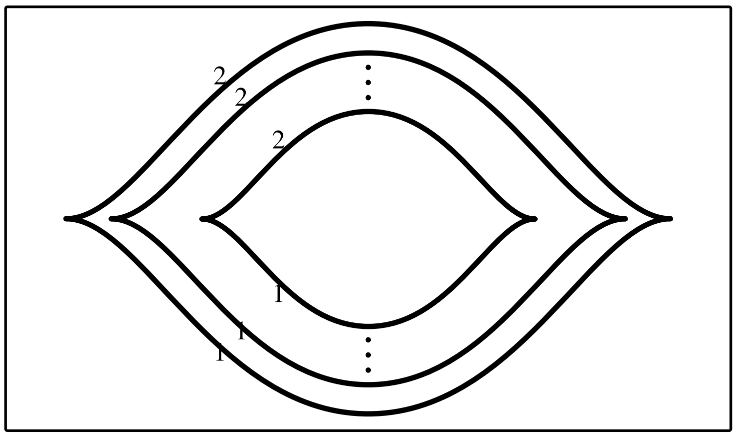

Figure 5: A nice Cerf graphic with only critical points of index 1 and 2.

Proof We begin again with a cobordism Z from $S^1 \times S^4$ to $S^1 \times_\phi S^4$ built as a family of cobordisms Y_t, each Y_t built by attaching n fixed standard 1–handles to $[0,1] \times S^4$ followed by n "moving" 2–handles governed by a loop of embeddings $\alpha_t : \amalg^n(S^1 \times B^3) \hookrightarrow \#^n(S^1 \times S^3)$. Cancelling the 1–2 pairs at time $t = 0 \sim 1$, we revert to the Cerf theoretic perspective to get a family (f_t, V_t) of Morse functions with gradient-like vector fields on $[0,1] \times S^4$ interpolating from f_0, which is projection onto $[0,1]$, to f_1, which is the pullback of f_0 via some pseudoisotopy $\Phi : [0,1] \times S^4 \to [0,1] \times S^4$ from id_S^4 to ϕ. The graphic now looks like Figure 5, exactly as in Figure 4 except that now the critical points are of index 1 and 2; there are still no handle slides.

Theorem 2.1.1 of Chenciner's thesis [4], restated as Hatcher and Wagoner's Proposition 1.4 on p.177 of [5], asserts that, given a 1–parameter family f_t of Morse functions on $[0,1] \times X$ where X is an m–manifold, if the Cerf graphic contains a swallowtail involving critical points of index i and $i+1$ as in the left of Figure 6, with $i \leq m-3$, then the swallowtail can be cancelled to give the graphic on the right in Figure 6. This applies in our setting because $m = 4$ and $i = 1 = 4 - 3$. We use this to reduce the number of 1–2 pairs using the main idea of Proposition 4 on p.217 of [5], as in the figure on the top of p.218 of [5]. We essentially reproduce this figure here in Figure 7 which shows how to reduce a nested pair of birth-deaths of 1–2 handles to a single pair. (The other elementary moves are introducing a swallowtail, which can always be done, and merging a death with a birth, which can always be done if level sets are connected, which they are in our case.)

Repeating this we can turn n nested 1–2 "eyes" into a single nested 1–2 "eye", and

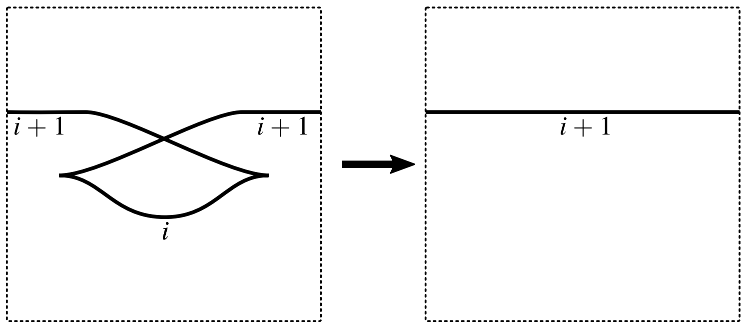

Figure 6: Eliminating a swallowtail.

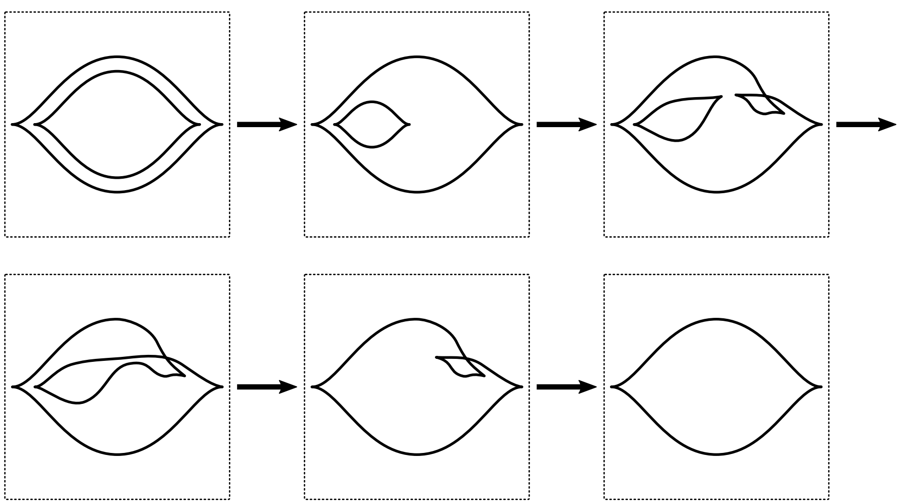

Figure 7: Using a swallowtail to turn a nested pair of "eyes" into a single eye.

then we can merge the birth again at $t = 0 \sim 1$. Note that in fact we could have left this last (bottom-most) index 1 critical point completely unchanged in this whole process, we even did not need to cancel it with its cancelling 2–handle at the beginning. Thus we can easily arrange that this last 1–handle is still stationary, i.e. its attaching map does not move with t. This shows that this cobordism can be built with a single fixed standard 1–handle followed by a single moving 2–handle whose attaching map is given by a loop of embeddings $S^1 \times B^3 \hookrightarrow S^1 \times S^3$. Therefore $[\phi] \in \mathcal{FH}_1(\pi_1(\mathcal{FS}_1(S^4, 1)))$.

□

Remark 13 In case it is needed in another context, the general version of this theorem is that, if $i \leq m - 3$ and X is an m–manifold, then for any n, $\mathcal{FH}_n(\pi_1(\mathcal{FS}_n(X, i))) = \mathcal{FH}_1(\pi_1(\mathcal{FS}_1(X, i)))$.

6 Twists along Montesinos twins

As a consequence of the preceding two theorems we now know that any diffeomorphism of S^4 arising as the monodromy of the top of a cobordism constructed as above from a loop of n 2–spheres in $\#^n(S^2 \times S^2)$ which remain disjoint from a parallel copy $\amalg^n(S^2 \times \{p'\})$ of the basepoint embedding $\amalg^n(S^2 \times \{p\})$ is isotopic to a diffeomorphism arising from a loop of embeddings of a single circle in $S^1 \times S^3$. This is summarized as:

Corollary 14

$$\mathcal{H}_\infty(\imath_*(\pi_1(\widehat{\mathcal{S}}_\infty))) = \mathcal{FH}_1(\pi_1(\mathcal{FS}_1(S^4, 1)))$$

Our next goal, which will complete the proof of Theorem 4, is to show that

$$\mathcal{FH}_1(\pi_1(\mathcal{FS}_1(S^4, 1)))$$

is generated by twists along Montesinos twins as advertised.

Proof of Theorem 4 Recall that $\mathcal{FS}_1(S^4, 1))$ is the space of framed embeddings of S^1 in $S^1 \times S^3$ while $\mathcal{S}_1(S^4, 1)) = Emb(S^1, S^1 \times S^3)$ is the space of unframed embeddings of S^1 in $S^1 \times S^3$. As noted at the end of Section 2, we might worry that the homomorphism $\pi_1(\mathcal{FS}_1(S^4, 1))) \to \pi_1(\mathcal{S}_1(S^4, 1)))$ is not surjective, since there are two possible framings of a circle in a 4–manifold. However, Budney and Gabai [2] give explicit representatives of generators for $\pi_1(Emb(S^1, S^1 \times S^3))$ all of which can be seen to lift to framed loops of embeddings, and thus the map is surjective so we do not need to worry about framings anymore. (There is probably some other more direct way to see this, the point being that there is no loop of embeddings of S^1 in $S^1 \times S^3$ which switches the two framings of S^1.)

For the remainder of this proof, we will use the less obscure notation $Emb(S^1, S^1 \times S^3)$ to refer to the space $\mathcal{S}_1(S^4, 1))$, the latter more complicated notation only being helpful when placing things in the much more general context of Section 2. Also note that in Section 2 we punctured the target space of our embeddings, but here we drop the puncture for simplicity. The point is that, at the level of π_1, the puncture is irrelevant since loops of circles are 2–dimensional while homotopies of loops of circles are 3–dimensional, so since the ambient space is 4–dimensional everything can be assumed to miss a point.

In fact [2] shows that every class $g \in \pi_1(Emb(S^1, S^1 \times S^3))$ can be represented by a loop $\gamma_t : S^1 \hookrightarrow S^1 \times S^3$ of embeddings such that the associated map $\Gamma : S^1 \times S^1 \to S^1 \times S^3$ given by $\Gamma(t, s) = \gamma_t(s)$ is itself an embedding. Thus we have an embedded torus $\Gamma : S^1 \times S^1 \hookrightarrow S^1 \times S^3$ such that $\Gamma(\{0\} \times S^1)$ is the basepoint embedding $C = S^1 \times \{p\}$. Surgery along C applied to the pair $(S^1 \times S^3, \Gamma)$ yields (S^4, R) for some embedded 2–sphere $R \subset S^4$, and the 2–sphere S dual to the surgery circle is an unknotted sphere $S \subset S^4$ such that (R, S) forms a Montesinos twin. Furthermore, the boundary $\partial\nu(R \cup S)$ of a neighborhood of this twin in S^4 is the same as the boundary of a tubular neighborhood of $\Gamma(S^1 \times S^1)$ in $S^1 \times S^3$. When this 3–torus is parametrized as $S^1_l \times S^1_R \times S^1_S$ as in the introduction, we see that the S^1_l parameter corresponds to the t parameter in $\Gamma(t, s) = \gamma_t(s)$, that the S^1_R direction corresponds to the s–parameter, and that the S^1_S direction corresponds to the boundary of the disk factor in the tubular neighborhood $\nu(\Gamma(S^1 \times S^1)) \cong D^2 \times S^1 \times S^1$.

Because Γ is embedded, it is relatively easy to see what $\mathcal{H}_1([\gamma_t])$ looks like. We need an ambient isotopy ϕ_t of $S^1 \times S^3$ with $\phi_0 = \mathrm{id}$, $\phi_t \circ \gamma_0 = \gamma_t$ and ϕ_1 equal to the identity on a neighborhood of C. (This is the "circle pushing" map we get by dragging the circle around the embedded torus and back to its starting position.) This can happen entirely in a tubular neighborhood $D^2 \times S^1 \times S^1$ of Γ, by spinning in the t direction more and more as we move towards the center of D^2, which we state explicitly as follows: Let (r, θ) be polar coordinates on D^2, and let (t, s) be coordinates as before on $S^1 \times S^1$. Choose a smooth non-increasing function $T : [0, 1] \to [0, 2\pi]$ which is 1 on $[0, 1/4]$, 0 on $[3/4, 1]$, and let $\phi_t(r, \theta, t, s) = (r, \theta, t + T(r), s)$. From this it is clear that ϕ_1 is the identity on $r \in [0, 1/4]$ and $r \in [3/4, 1]$, and on the intermediate $[1/4, 3/4] \times S^1 \times S^1 \times S^1$ is equal to a Dehn twist on $[1/4, 3/4] \times S^1$ crossed with the identity in the remaining $S^1 \times S^1$ direction. Back in S^4 this is exactly the twist τ_W along the twin $W = (R, S)$.

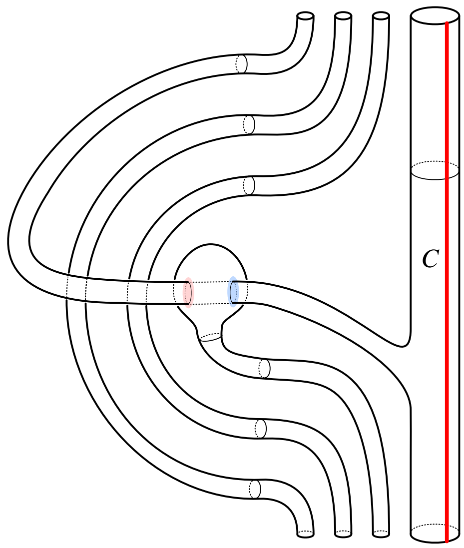

Figure 8: The embedded torus $T(3)$ in $S^1 \times S^3$, the obvious next member of the family of tori described in Figure 4 of [2]. The top is glued to the bottom, and horizontal slices are S^3's, with the "time" coordinate indicated in red/blue shading, as in Figure 1. Here we have exaggerated certain features of this torus and deformed somewhat from the drawings in Figure 4 of [2] so that the connection with the Montesinos twin $W(3) = (R(3), S)$ in Figure 1 is visually apparent. Surgering along the red circle C collapses the vertical cylinder on the right into a ball (the tail of the snake), with the dual sphere to C becoming the tail-piercing sphere S.

In fact [2] establishes an isomorphism

$$W_1 \times W_2 : \pi_1(Emb(S^1, S^1 \times S^3), C) \to \mathbb{Z} \oplus \Lambda^1_3$$

where Λ^1_3 is a free abelian group on a countably infinite generating set. The $\mathbb{Z}$ factor in $\mathbb{Z} \oplus \Lambda^1_3$ is given by the loops of S^1–reparametrizing embeddings $\gamma_t(s) = \gamma_0(s + nt)$, and it is easy to see that $\mathcal{H}_1$ applied to such a loop of embeddings is isotopic to id_{S^4}, i.e. this $\mathbb{Z}$ factor is in the kernel of $\mathcal{H}_1$. Modulo this $\mathbb{Z}$ factor, Figure 4 in [2] gives the first two tori $T(1)$ and $T(2)$ in an obvious family $T(i)$ of tori in $S^1 \times S^3$ which give the countably infinite generating set corresponding to Λ^1_3. We draw $T(3)$ in Figure 8. In this figure, the circle $C \subset S^1 \times S^3$ is represented as a red vertical line on the far right side of the torus. The torus $T(n)$ is just like this but wraps n times around the S^1 direction. Surgering along C yields our Montesinos twins $W(i) = (R(i), S)$.

□

References

[1] Ryan Budney. A family of embedding spaces. In *Groups, homotopy and configuration spaces*, volume 13 of *Geom. Topol. Monogr.*, pages 41–83. Geom. Topol. Publ., Coventry, 2008.

[2] Ryan Budney and David Gabai. Knotted 3-balls in S^4, 2019. arXiv:1912.09029.

[3] Jean Cerf. La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie. *Inst. Hautes Études Sci. Publ. Math.*, (39):5–173, 1970.

[4] Alain Chenciner. *Sur la geéométrie des strates de petites codimensions de l'espace des fonctions différentiables rélles sur une variété*. PhD thesis, 1971.

[5] Allen Hatcher and John Wagoner. *Pseudo-isotopies of compact manifolds*. Société Mathématique de France, Paris, 1973. With English and French prefaces, Astérisque, No. 6.

[6] Michel A. Kervaire and John W. Milnor. Groups of homotopy spheres. I. *Ann. of Math. (2)*, 77:504–537, 1963.

[7] J. P. Levine. Lectures on groups of homotopy spheres. In *Algebraic and geometric topology (New Brunswick, N.J., 1983)*, volume 1126 of *Lecture Notes in Math.*, pages 62–95. Springer, Berlin, 1985.

[8] José M. Montesinos. On twins in the four-sphere. I. *Quart. J. Math. Oxford Ser. (2)*, 34(134):171–199, 1983.

[9] Richard S. Palais. Local triviality of the restriction map for embeddings. *Comment. Math. Helv.*, 34:305–312, 1960.

[10] Tadayuki Watanabe. Some exotic nontrivial elements of the rational homotopy groups of $\mathrm{Diff}(S^4)$. arXiv:1812.02448.