SYMMETRIES OF 4-MANIFOLDS || WARSAW LECTURES

MARK POWELL

Contents

1. Introduction 1
Acknowledgements 3
2. Examples 3
3. Applications 4
3.1. A Dehn twist 4
3.2. Topological unknotting 4
4. Pseudo-isotopy 5
5. Bundles 7
6. Modified surgery 8
7. Proof of the existence of a pseudo-isotopy 9
8. Pseudo-isotopy implies isotopy 13
9. Cerf theory 14
10. The middle-middle level 17
11. Geometric moves and Quinn's embedded arc criterion 20
11.1. Dual spheres 20
11.2. Embedded arc criterion 20
11.3. The sum square move 22
11.4. Whitney spheres 24
12. Algebraically cancelling finger-Whitney intersections 24
13. Completion of the pseudo-isotopy implies isotopy proof 26
14. A Cork Theorem for diffeomorphisms 28
15. Mapping class groups of 1-connected 4-manifolds with boundary 29
16. Simple spines for knot traces 31
References 31

1. Introduction

These are the notes for a series of lectures I delivered on 26th, 27th, and 28th March 2024 in Warsaw, at the Department of Mathematics of the University of Warsaw and at IMPAN. The aim of these lectures is to explain the computation of the (topological)

The lecture series and my stay in Warsaw was funded by the Thematic Research Program "Topological field theories and algebraic invariants of links" within the IDUB Program at the University of Warsaw. I was partially supported by EPSRC New Investigator grant EP/T028335/2 and EPSRC New Horizons grant EP/V04821X/2.

mapping class groups of compact simply-connected 4-manifolds, focussing for simplicity on the closed case.

I will also consider the case of nonempty boundary, and I will compare with the smooth category in various ways. There will be some applications to surfaces.

We have a good understanding of the mapping class groups of compact, simply-connected 4-manifolds, which I will present in detail. I will discuss what is known beyond that, in less detail. I will mention open problems and future research directions, of which there are many. I hope that this motivates understanding the simply-connected case in detail: the ideas I try to impart have been generalised and applied to many problems on mapping class groups, and I believe there are many more such possibilities.

Definition 1.1. Let X be a closed, topological 4-manifold. We write $\mathrm{Homeo}(X)$ for the homeomorphism group of X, the topological group of homeomorphisms $f\colon X \xrightarrow{\cong} X$ and we write $\mathrm{Homeo}^+(X)$ for the topological group of orientation preserving (o.p.) homeomorphisms $f\colon X \xrightarrow{\cong} X$

Multiplication is via composition, and the topology is the compact-open topology. The connected components of this group, $\pi_0(\mathrm{Homeo}(X))$ is the *mapping class group* of X. The *o.p. mapping class group* of X is $\pi_0(\mathrm{Homeo}^+(X))$.

The main theorem I want to discuss is a computation of the mapping class groups of closed, simply-connected 4-manifolds, in the sense of reducing it to algebra. Henceforth for brevity I will write 1-connected for simply-connected. Note that this implies path connected as well.

Theorem 1.2 (Freedman, Kreck, Perron, Quinn)**.** *Let X be a closed, 1-connected topological 4-manifold. Then*

$$\begin{aligned} \pi_0(\mathrm{Homeo}^+(X) &\xrightarrow{\cong} \mathrm{Aut}(H_2(X), \lambda_X) \\ f &\mapsto f_* \end{aligned}$$

is an isomorphism of groups.

The references are [Fre82, Kre79, Per86, Qui86]. Freedman proved surjectivity. The precise attributions for the injectivity will be dealt with later; see also [GGH+23].

Here,

$$\begin{aligned} \lambda_X\colon H_2(X) \times H_2(X) &\to \mathbb{Z} \\ (x, y) &\mapsto \langle \mathrm{PD}^{-1}(y), x\rangle \end{aligned}$$

is the intersection pairing of X, and an automorphism in $\mathrm{Aut}(H_2(X), \lambda_X)$ is an isomorphism $\varphi\colon H_2(X) \to H_2(X)$ such that $\lambda_X(\varphi(x), \varphi(y)) = \lambda_X(x, y) \in \mathbb{Z}$ for all $x, y \in H_2(X)$.

Throughout, X will be assumed to be a topological manifold. If it is explicitly stated then we may, and often will, endow X with a smooth structure, but this is not the default assumption.

Here is a summary of what these notes will aim to cover.

(i) Explain the proof of injectivity in Theorem 1.2 in some detail. A key concept will be *pseudo-isotopies*. They will allow us to break the proof of injectivity into two distinct steps.
(ii) Generalise to the case of nonempty boundary.

(iii) Compare with smooth mapping class groups: exotic diffeomorphisms, exotic pseudo-isotopies, non-smoothable homeomorphisms, corks for diffeomorphisms.
(iv) Give applications to surfaces in 4-manifolds.
(v) Mention generalisations to non-simply connected 4-manifolds, both known and possible future extensions.

Acknowledgements. I am extremely grateful to the organisers of the Simons semester in Warsaw for the invitation and the opportunity to speak. I am equally grateful to the audience, who persisted with a marathon of 8 lectures in 3 days.

Some of the text and the pictures in this write up were lifted from [OP22] and [GGH$^+$23], and I thank my coauthors for their forbearance. It seemed inefficient to create new inferior versions.

Much of my understanding of this topic is due to conversations with many excellent collaborators and PhD students, and so I would like to take the opportunity to thank (in alphabetical order by surname) Anthony Conway, Michelle Daher, David Gabai, Daniel Galvin, David Gay, Daniel Hartman, Daniel Kasprowski, Slava Krushkal, Andrew Lobb, Anubhav Mukherjee, Weizhe Niu, Isacco Nonino, Patrick Orson, Brendan Owens, Arunima Ray, Oliver Singh, Peter Teichner, and Terrin Warren.

These notes are based on the work of many mathematicians, but most strongly on the visionary ideas of Matthias Kreck and Frank Quinn.

2. Examples

In the upcoming examples, we use Theorem 1.2 to compute $\pi_0(\mathrm{Homeo}^+(X))$, and then deduce $\pi_0(\mathrm{Homeo}(X))$ as a consequence. A unimodular, symmetric, bilinear form $(\mathbb{Z}^n, L)$ is called a lattice, and the groups of symmetries of a lattice is called its orthogonal group.

In general by considering the action of a homeomorphism on $H_4(X) \cong \mathbb{Z}$ to define a homomorphism to $\mathbb{Z}/2$, we have an exact sequence

$$0 \to \pi_0 \,\mathrm{Homeo}^+(X) \to \pi_0 \,\mathrm{Homeo}(X) \to \mathbb{Z}/2.$$

Example 2.1. Let $X = S^4$. Then since $H_2(S^4) = 0$, we have that $\mathrm{Aut}(H_2(S^4), \lambda_X) = \{\mathrm{Id}\}$, and so $\pi_0(\mathrm{Homeo}^+(S^4)) = \{[\mathrm{Id}]\}$ by Theorem 1.2. Since S^4 admits an orientation-reversing homeomorphism, we have that $\pi_0(\mathrm{Homeo}(S^4)) \cong \mathbb{Z}/2 = \{[\mathrm{Id}], [R]\}$, where $R\colon S^4 \to S^4$ is a reflection.

Example 2.2. Let $X = \#^n \mathbb{C}P^2$ be a connected sum of canonically oriented complex projective planes. This has $H_2(X) \cong \mathbb{Z}^n$ and λ_X represented by the size n identity matrix. The automorphism groups of this form, or in other words the orthogonal group of this lattice, which is isomorphic to $\pi_0 \,\mathrm{Homeo}^+(X)$, fits into a short exact sequence

$$\{0\} \to (\mathbb{Z}/2)^n \to \mathrm{Aut}(\mathbb{Z}^n, \mathrm{Id}) \to \Sigma_n \to \{1\}.$$

The sequence splits, so we have a semi-direct product, and the symmetric group Σ_n acts on $(\mathbb{Z}/2)^n$ by permuting the coordinates. It is known as the signed permutation group, or the Coxeter group of type B_n. Its order is $2^n \cdot n!$. Since the signature is nonzero, there is no orientation reversing homeomorphism, so $\pi_0 \,\mathrm{Homeo}(X) = \pi_0 \,\mathrm{Homeo}^+(X)$.

Example 2.3. Let $X = S^2 \times S^2$. Then $H_2(X) \cong \mathbb{Z}^2$, and the intersection form is hyperbolic, represented by $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. So $\pi_0 \,\mathrm{Homeo}^+(X) \cong \mathrm{Aut}(\mathbb{Z}^2, H) \cong \mathbb{Z}/2 \times \mathbb{Z}/2$. The isometries here are straightforward to compute by hand. Generators are given by $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Let $R\colon S^2 \to S^2$ be a reflection. Then $R \times \mathrm{Id}$ is an orientation reversing homeomorphism of order two. It follows that we have a split short exact sequence

$$0 \to \pi_0 \operatorname{Homeo}^+(X) \to \pi_0 \operatorname{Homeo}(X) \to \mathbb{Z}/2 \to 0.$$

The sequence splits, so we have that $\pi_0 \operatorname{Homeo}(X) \cong (\mathbb{Z}/2 \times \mathbb{Z}/2) \rtimes \mathbb{Z}/2$. By computing the orders of elements using the action on H_2 and H_4, I computed that the action in the semi-direct product is such that this group is isomorphic to D_8, the dihedral group of order 8.

Example 2.4. Let $X = E_8$ be the E_8 manifold. It is built by plumbing D^2-bundles over S^2 with Euler number 2 together according to the E_8 Dynkin diagram, and then capping off the boundary with a contractible 4-manifold, who existence was proven by Freedman. The intersection form is the E_8 lattice. Its automorphism group is the Weyl or Coxeter group of type E_8. This is a famous group, whose order is $4! \cdot 6! \cdot 8!$.

In general, Wall gave explicit generators for automorphism groups of unimodular lattices $\operatorname{Aut} \lambda_X$ [Wal63].

3. Applications

3.1. **A Dehn twist.** Recall the K_3 surface, which is a famous smooth, close, 1-connected 4-manifold. It generates the 4-dimensional smooth spin cobordism group, for example. It is given by

$$\{[x, y, z, w] \in \mathbb{C}P^3 \mid x^4 + y^4 + z^4 + w^4 = 0\}.$$

There are several other descriptions. We let $X := K_3 \# K_3$, and consider the connected sum sphere $S^3 \subseteq X$. Let $U \cong S^3 \times [0,1]$ be a neighbourhood of this S^3. Let $\rho_\theta\colon S^3 \to S^3$ be rotation of S^3 through an angle θ about a fixed axis. Define

$$\begin{aligned} \Phi\colon S^3 \times [0,1] &\to S^3 \times [0,1] \\ (x,t) &\mapsto (R_{2\pi t}(x), t). \end{aligned}$$

Define

$$\begin{aligned} f\colon X &\to X \\ x &\mapsto \begin{cases} \Phi(x) & x \in U \\ x & x \notin U. \end{cases} \end{aligned}$$

This is called Dehn twist on $S^3 \subseteq X$. Then $f_* = \mathrm{Id}_X\colon H_2(X) \to H_2(X)$, so f is topologically isotopic to the identity. However Kronheimer and Mrowka [KM20] proved that f is not smoothly isotopic to the identity of X.

3.2. **Topological unknotting.** Let $f\colon S^2 \to S^4$ be a 2-knot, and suppose that

$$\pi_1(S^4 \smallsetminus f(S^2)) \cong \mathbb{Z}.$$

That is, the knot group is the same as that of the un-2-knot U. The normal bundle νf is trivial, homeomorphic to $S^2 \times D^2$. We have a homeomorphism $g\colon \nu f \to \nu U$ such that $g|_{f(S^2)}$ sends $f(S^2)$ to U.

Theorem 3.1 (Freedman-Quinn [FQ90]). *The homeomorphism g can be extended to a homeomorphism $\widehat{g}\colon S^4 \to S^4$.*

The theorem above is a highly nontrivial input, of course. However there is still the question of whether we can obtain an ambient isotopy. By Example 2.1 we have that $\pi_0 \operatorname{Homeo}^+(S^4) = \{[\mathrm{Id}]\}$, so $\widehat{g}$ is isotopic to Id_{S^4}. This gives an ambient isotopy from $\widehat{g} \circ f(S^2) = U$ to $\mathrm{Id}_{S^4} \circ f(S^2) = f(S^2)$. We deduce that every 2-knot with knot group $\mathbb{Z}$ is topologically unknotted.

4. Pseudo-isotopy

How do we prove that $\pi_0 \operatorname{Homeo}^+(X) \cong \operatorname{Aut}(H_2(X), \operatorname{Aut} \lambda_X)$? A key concept in the proof will be *pseudo-isotopy*. This notion enables us to break down the problem into two distinct steps, whose proofs are very different in character.

An analogy with manifold classification is in order. To prove that two 1-connected n-manifolds, $n \geq 4$, are homeomorphic, the standard method is as follows: first prove that the manifolds are h-cobordant; this uses surgery theoretic methods. Recall that an h-cobordism between two n-manifolds M and N, with a homeomorphism $\partial M \cong \partial N$ is a cobordism $(W; M, N)$, namely an $(n+1)$-manifold with boundary $M \cup (\partial M \times I) \cup N$, such that the inclusion maps $M \to W$ and $N \to W$ are both homotopy equivalences.

Then, we apply the h-cobordism theorem (Smale, Kirby-Siebenmann, Freedman-Quinn), which says that h-cobordant n-manifolds are homeomorphic, provided $n \geq 4$. Its proof uses Morse theory.

To prove that two homeomorphisms are isotopic, there are two analogous steps. First, we prove that they are pseudo-isotopic. This uses surgery theory. Then we apply a result, proven using Cerf theory (1-parameter Morse theory) which says that pseudo-isotopic homeomorphisms are isotopic. So pseudo-isotopy is to classifying homeomorphisms of 4-manifolds up to isotopy as h-cobordism is to classifying 4-manifolds up to homeomorphism.

Let $\mathrm{CAT} \in \{\mathrm{Diff}, \mathrm{Top}\}$ and let X^4 be a compact CAT 4-manifold.

Definition 4.1 (Pseudo-Isotopy). A CAT pseudo-isotopy (PI) of X is a CAT isomorphism (i.e. a diffeomorphism or a homeomorphism) $F\colon X \times I \xrightarrow{\cong} X \times I$ with $F|_{\sqsubset} = \mathrm{Id}$, where $\sqsubset := (X \times \{0\}) \cup (\partial X \times I)$.

We say that $F|_{X \times \{1\}}$ is (CAT) pseudo-isotopic to Id_X.

Remark 4.2. For $f\colon X \xrightarrow{\cong} X$ a CAT isomorphism, we have that

$$\text{isotopic to Id} \Rightarrow \text{pseudo-isotopic to Id} \Rightarrow \text{homotopic to Id} \Rightarrow f_* = \mathrm{Id}_{H_2(X)}\,.$$

We will see that for X closed and 1-connected, all of these implications can be reversed.

We briefly justify the first two implications. Let $f_t\colon X \to X$ be an isotopy. Then $F\colon X \times I \to X \times I$ sending $(x,t) \mapsto (f_t(x), t)$ is a pseudo-isotopy. On the other hand, given a pseudo-isotopy $F\colon X \times I \to X \times I$ from $f = F|_{X \times \{1\}}$ to Id, we obtain a homotopy $\mathrm{pr}_1 \circ F\colon X \times I \to X$, which gives a homotopy from f to Id.

Definition 4.3. We say that $f, g\colon X \xrightarrow{\cong} X$ with $f|_{\partial X} = g|_{\partial X} = \mathrm{Id}_{\partial X}$ are CAT *pseudo-isotopic* if and only if $g^{-1} \circ f\colon X \to X$ is pseudo-isotopic to Id_X.

Note that if f and g are pseudo-isotopic then there is a CAT isomorphism $F\colon X \times I \to X \times I$ with $F|_{X \times \{0\}} = f$ and $F|_{X \times \{1\}} = g$.

Definition 4.4. We write

$$\widetilde{\pi}_0 \operatorname{Homeo}^+(X) := \{f \colon X \xrightarrow{\cong} X\}/\text{pseudo-isotopy}.$$

We can factor the map $f \mapsto f_*$ from Theorem 1.2 as

$$\pi_0 \operatorname{Homeo}^+(X) \xrightarrow{\varphi_1} \widetilde{\pi}_0 \operatorname{Homeo}^+(X) \xrightarrow{\varphi_2} \pi_0 \operatorname{hAut}^+(X) \xrightarrow{\varphi_3} \operatorname{Aut}(H_2(X), \lambda_X).$$

The plan for the proof of Theorem 1.2 will be to show that φ_1 and $\varphi_3 \circ \varphi_2$ are isomorphisms.

See Cochran-Habegger [CH90] for the computation of $\pi_0 \operatorname{hAut}^+(X)$ for X a closed, 1-connected 4-manifold.

Surjectivity of φ_1 is obvious. The refined version of Freedman's classification of closed, 1-connected 4-manifolds includes the following statement.

Theorem 4.5 (Freedman [Fre82])**.** *For every automorphism $\theta \in \operatorname{Aut}(H_2(X), \lambda_X)$, there exists a homeomorphism $f \colon X \to X$ with $f_* = \theta$.*

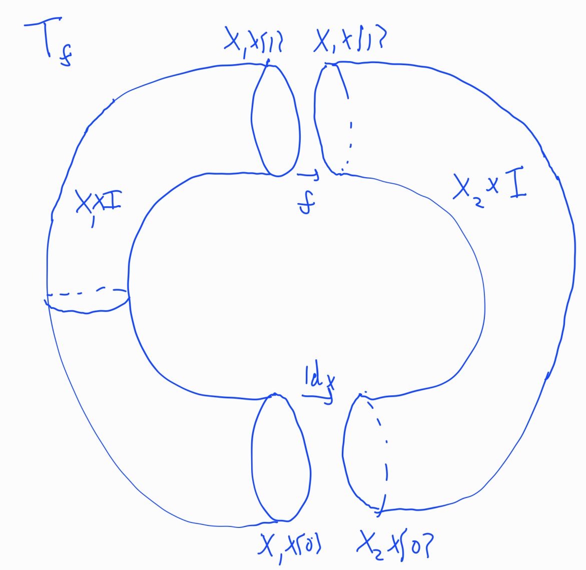
Thus $\varphi_3 \circ \varphi_2$ and indeed $\varphi_3 \circ \varphi_2 \circ \varphi_1$ is surjective. We will focus on injectivity, and indeed establishing injectivity will be the main goal of these notes. First, we will show that

$$\varphi_3 \circ \varphi_2 \colon \widetilde{\pi}_0 \operatorname{Homeo}^+(X) \to \operatorname{Aut}(H_2(X), \lambda_X)$$

is injective. This is due to Kreck [Kre79] and Quinn [Qui86]. Note that $\varphi_3 \circ \varphi_2$ is simply the map that sends a homeomorphism f to the induced map on homology $f_* \colon H_2(X) \to H_2(X)$.

So, let $f \colon X \xrightarrow{\cong} X$ be a homeomorphism with $f_* = \operatorname{Id}$. We consider the *mapping torus* $T_f := X \times [0,1]/(x,0) \sim (f(x),1)$. We think of a homeomorphic space to this, as follows. Let X_1 and X_2 be two copies of X, and consider f as a map $f \colon X_1 \to X_2$. Let $\operatorname{Id} \colon X_1 \to X_2$ be the identity map of X. Then

$$T_f \cong \frac{X_1 \times [0,1] \sqcup X_2 \times [0,1]}{(x_1, 1) \sim (f(x_1), 1), (x, 0) \sim (\operatorname{Id}(x), 0)}.$$



The aim is to find an h-cobordism $(V; X_1 \times I, X_2 \times I)$, relative to the boundary $X_1 \times \{0,1\}$, from $X_1 \times I$ to $X_2 \times I$, with the gluing indicated. That is V must be a 6-dimensional manifold with boundary $\partial V = T_f$, and being an h-cobordism means that the inclusion maps $X_1 \times I \to V$ and $X_2 \times I \to V$ are both homotopy equivalences. The h-cobordism theorem for 1-connected manifolds of dimension at least 5 is due to Smale and Kirby-Siebenmann [Sma62, KS77]; see also [Mil65]. It says that h-cobordism are homeomorphic to products, and we may use a given identification of one end of the two cobordisms.

Theorem 4.6 (h-cobordism theorem)**.** *There is a homeomorphism*

$$(G, \mathrm{Id}, g)\colon (V; X_1 \times I, X_2 \times I) \xrightarrow{\cong} (X_1 \times I \times I; X_1 \times I, X_1 \times I)$$

relative to the identity map on $X_1 \times I$ and some homeomorphism $g\colon X_2 \times I \to X_1 \times I$.

Then $g\colon X_2 \times I \xrightarrow{\cong} X_1 \times I$ is a pseudo-isotopy. Because of our initial choice of gluing, i.e. the fact that we took the mapping torus T_f, it is a pseudo-isotopy from f to Id. So if we can find the h-cobordism V, we will have proven that f is pseudo-isotopic to Id_X, and hence will have proven that $\varphi_3 \circ \varphi_2$ is injective. This will be the goal of the next three sections.

5. Bundles

In this section we recall some basic bundle theory for topological manifolds. Let X be a topological manifold. Recall that

$$\{\text{smooth vector bundles on } X \text{ of rank } n\}/\cong \leftrightarrow [X, \mathrm{BO}(n)].$$

Here $\mathrm{BO}(n)$ is the classifying space for smooth vector bundles, which can be modelled by the Grassmannian $\mathrm{Gr}_n(\mathbb{R}^\infty)$ of n-dimensional subspaces of $\mathbb{R}^\infty$. We have inclusions

$$\mathrm{BO}(n) \hookrightarrow \mathrm{BO}(n+1) \hookrightarrow \cdots$$

with colimit $\mathrm{BO} := \mathrm{colim}_{n\to\infty} \mathrm{BO}(n)$. Then $[X, \mathrm{BO}]$ is in one to one correspondence with stable vector bundles on X.

For the analogue of this in the topological category we define

$$\mathrm{Top}(n) = \{f\colon \mathbb{R}^n \xrightarrow{\cong} \mathbb{R}^n \mid f(\underline{0}) = \underline{0}\}.$$

There is an associated classifying space $\mathrm{BTop}(n)$, with $\mathrm{colim}_{n\to\infty} \mathrm{BTop}(n) =: \mathrm{BTop}$, such that

$$\{\mathbb{R}^n \text{ fibre bundles on } X\}/\cong \leftrightarrow [X, \mathrm{BTop}(n)]$$

and such that $[X, \mathrm{BTop}]$ is in one to one correspondence with stable $\mathbb{R}^n$-fibre bundles on X.

In particular, if $\dim X = d$, there is a *topological tangent bundle* $\tau_X \in [X, \mathrm{BTop}(d)]$. For N large we can embed X in $\mathbb{R}^{d+N}$, and there is a normal bundle bundle $\nu_X \in [X, \mathrm{BTop}(N)]$ such that $\tau_X \oplus \nu_X$ is null-homotopic, i.e. this is isomorphism to the trivial fibre bundle $\underline{\mathbb{R}}^{d+N}$ on X. We consider the *stable normal bundle of* X, $\nu_X \in [X, \mathrm{BTop}]$. This is independent of the choice of embedding of X into Euclidean space.

For oriented vector bundles, we can use the oriented Grassmannian $\mathrm{BSO}(n)$ and its stable version BSO. The topological analogue is $\mathrm{BSTop}(n)$ and BSTop. Finally, just as vector bundles have spin structures, topological $\mathbb{R}^n$-bundles have topological spin structures, and there is a fibration BTopSpin $\to$ BTop. A topological spin structure on a stable $\mathbb{R}^n$-bundle corresponds to a lift $X \to$ BTopSpin of the classifying map $X \to$ BTop.

The main references for this theory are Milnor's papers, where he defined microbundles [Mil64, Mil61], and Kister's paper, where he showed that microbundles are fibre bundles [Kis64]. In view of Kister's theorem, we have only talked about the conceptually easier notion of $\mathbb{R}^n$-fibre bundles.

6. Modified surgery

To obtain an h-cobordism $(V; X_1 \times I, X_2 \times I)$ with boundary T_f, I will use the method of modified surgery, which is due to Kreck [Kre99]. The proof that I am presenting is modelled on that in [Kre79], although that paper was written before the terminology of modified surgery had been solidified. Hence we use the terminology from [Kre99].

Theorem 6.1. *Let X be a compact n-manifold and let $\nu_X \colon X \to \mathrm{BTop}$ be the stable normal bundle. For any $k \geq 0$ there exist $(B, \xi, \overline{\nu}_X)$, with $\xi \colon B \to \mathrm{BTop}$ a fibration, such that ν_X factors as*

$$\begin{tikzcd} & B \arrow[d, "\xi"] \\ X \arrow[ur, "\overline{\nu}_X"] \arrow[r, "\nu_X"'] & \mathrm{BTop} \end{tikzcd}$$

where ξ is $(k+1)$-coconnected and $\overline{\nu}_X$ is $(k+1)$-connected.

Here ξ is $(k+1)$-coconnected means that $\pi_i(\xi)$ is an isomorphism for $i > k+1$ and $\pi_{k+1}(\xi)$ is injective. Dually, $\overline{\nu}_X$ being $(k+1)$-connected means that $\pi_i(\overline{\nu}_X)$ is an isomorphism for $i < k+1$, i.e. for $i \leq k$, and $\pi_{k+1}(\overline{\nu}_X)$ is surjective. The factorisation in the theorem is called the *Moore-Postnikov factorisation* of ν_X [Bau77, Rob72]

Definition 6.2. The fibration (B, ξ) is called the *normal k type* of X, and a lift of ν_X along ξ, $\overline{\nu}_X \colon X \to B$, is called a *normal k-smoothing.*

Given X, the normal k-type (B, ξ) is well-defined up to fibre homotopy equivalence. The relevance of the normal k-type to the classification of homeomorphisms is the following key theorem from [Kre99].

Theorem 6.3 (Kreck)**.** *Let M and M' be compact, oriented n-manifold with $\partial M \cong \partial M'$. Suppose that $n \geq 4$, and $n = 2k$ or $n = 2k+1$. Suppose that M and M' have the same normal k-types. Let $\overline{\nu} \colon M \to B$ and $\overline{\nu}' \colon M' \to B$ be normal k-smoothings that are bordant over (B, ξ). That is, we assume there is a bordism $(W^{n+1}; M, M')$ rel. $\partial M \cong \partial M'$ and a lift $\overline{\nu}_W \colon W \to B$ of $\nu_W \colon W \to \mathrm{BTop}$ that restricts to $\overline{\nu}$ and $\overline{\nu}'$ on M and M' respectively. There is a* surgery obstruction

$$\theta(W, \overline{\nu}_W) \in L^s_{n+1}(\mathbb{Z}\pi_1(B))$$

(or in $L^s_{n+1}(\mathbb{Z}\pi_1(B), S)$ if $n+1 = 6$ or 14) such that $(W, \overline{\nu}_W)$ is bordant rel. ∂W over (B, ξ) to an s-cobordism if and only if $\theta(W, \overline{\nu}_W) = 0$.

The strategy of the proof is to perform a sequence of surgeries on framed, embedded spheres, in order to kill the relative homotopy groups $\pi_i(B, W)$ for $i = 0, 1, \ldots, k$. This done, it can be proven using Poincaré duality that we have an s-cobordism. The bundle data is needed in order to make sure we can find spheres with trivial normal bundles for surgery when we need to. In the odd case $n = 2k+1$, there is an obstruction to killing π_{k+1}. This takes the form of the equivariant intersection form of the B-bordism. This is an element $\theta(W, \overline{\nu}_W) \in L^s_{n+1}(\mathbb{Z}\pi_1(B))$. It is represented by a nonsingular, $(-1)^{k+1}$-Hermitian, sesquilinear pairing on a based, f.g. free $\mathbb{Z}[\pi_1(B)]$-module, with values in $\mathbb{Z}[\pi_1(B)]$, together with a quadratic refinement $\mu\colon \mathbb{Z}[\pi_1(B)]/g \sim (-1)^{k+1}g^{-1}$. The S in $L^s_{n+1}(\mathbb{Z}\pi_1(B), S)$ indicates a modified quadratic refinement. For us, the key fact will be that $L^s_6(\mathbb{Z}) \cong \mathbb{Z}/2$, generated by a form with Arf invariant 1, and that $L^s_6(\mathbb{Z}, S) = 0$. As a result there is no surgery obstruction for the cobordism, and we have the following corollary.

Corollary 6.4. *Let $n = 5$, so $k = 2$, and suppose $(M, \overline{\nu})$ and $(M', \overline{\nu}')$ are 1-connected and are bordant over their common normal 2-type (B, ξ). Then M and M' are h-cobordant rel. boundary.*

7. Proof of the existence of a pseudo-isotopy

For simply-connected manifold, h- and s-cobordism agree, so it is enough to find an h-cobordism. We will apply Corollary 6.4 with

- X a closed, 1-connected, spin 4-manifold, with two copies X_1 and X_2 as above;
- $M := X_1 \times I$;
- $M' := X_2 \times I$;
- the identification $f \sqcup \mathrm{Id}$ between their boundaries;
- (B, ξ) the normal 2-type of X.

I will restrict to the spin case here. The argument in the non-spin case is similar, so one learns the main ideas from just considering the spin case, and we avoid too much repetition.

Example 7.1. We determine the normal 2-type for a closed, 1-connected, spin topological 4-manifold X. Recall that $H_2(X) \cong \mathbb{Z}^m$ for some m. Let $K := \prod^m \mathbb{C}P^\infty$. A choice of identification $H_2(X) \cong \mathbb{Z}^m$ determines (up to homotopy) a map $\eta\colon X \to K$. A choice of spin structure determines a lift $\mathfrak{s}\colon X \to \mathrm{BTopSpin}$ of ν_X. The combination of these two maps gives a diagram:

$$\begin{tikzcd} & \mathrm{BTopSpin} \times K \arrow[d, "\xi"] \\ X \arrow[ur, "\overline{\nu}_X"] \arrow[r, "\nu_X"'] & \mathrm{BTop}\,. \end{tikzcd}$$

Here $\xi\colon$ BTopSpin $\times K \to$ BTop is given by projection to the first factor then the standard map BTopSpin $\to$ BTop. We leave it to the reader to check that $\overline{\nu}_X$ is 3-connected and that ξ is 3-coconnected.

We have the same normal 2-type for $X \times I$:

$$\begin{tikzcd} & \mathrm{BTopSpin} \times K \arrow[d, "\xi"] \\ X \times I \arrow[ur, "\overline{\nu}_{X\times I}"] \arrow[r, "\nu_{X\times I}"'] & \mathrm{BTop}\,. \end{tikzcd}$$

We take two copies of $X \times I$, as above, and glue them using $f \sqcup \mathrm{Id}$ to obtain the mapping torus T_f. On each copy of $X \times I$ we have the normal 2-smoothing $\eta \times \mathfrak{s}\colon X \times I \to B = K \times \mathrm{BTopSpin}$. When we glue with Id on $X \times \{0\}$, the two maps are clearly compatible. However, we also have that $\mathfrak{s} \circ f\colon X \times \{1\} \to X \times \{1\}$ is homotopic to $\mathfrak{s}$, because X is 1-connected and so has a unique spin structure. In addition, $\eta \circ f \sim \eta\colon X \times \{1\} \to K$, because f induces the identity on $H_2(X)$. We can insert the resulting homotopy between $(\eta \times \mathfrak{s}) \circ f$ and $\eta \times \mathfrak{s}$ into a collar on $X \times \{1\}$ in one of the copies of $X \times I$, to obtain a map

$$\overline{\nu}_{T_f}\colon T_f \to B$$

that restricts to the given normal 2-smoothings on each copy of $X \times I$. It remains to prove that $(T_f, \overline{\nu}_{T_f})$ is B-null-bordant. Once this is done, we can apply Corollary 6.4 to obtain the desired h-cobordism. We have a bordism group $\Omega_5(B, \xi)$, which can be identified with $\Omega_5^{\mathrm{TopSpin}}(K)$.

Proposition 7.2. $\Omega_5^{\mathrm{TopSpin}}(K) = 0$.

Proof. To compute this, we recall that bordism groups give rise to a generalised homology theory. These satisfy the same axioms as ordinary homology, except that the generalised homology groups of a point need not be concentrated in degree zero. In fact, for $\Omega_*^{\mathrm{TopSpin}}$ we have $\Omega_q^{\mathrm{TopSpin}} \cong \mathbb{Z}$, $\mathbb{Z}/2$, $\mathbb{Z}/2$, 0, $\mathbb{Z}$, 0 for $q = 0, 1, 2, 3, 4, 5$ respectively. There is an Atiyah-Hirzebruch spectral sequence (AHSS) computing $\Omega_5^{\mathrm{TopSpin}}(K)$:

$$E^2_{p,q} = H_p(K; \Omega_q^{\mathrm{TopSpin}}) \;\Rightarrow\; \Omega_{p+q}^{\mathrm{TopSpin}}(K).$$

Here are the relevant terms on the E_2 page of the AHSS. Since the coefficients vanish for $q = 3, 5$, we have zeros in those rows. Since the homology of $\prod^m \mathbb{C}P^\infty$ vanishes in odd degree, the columns with p odd have only zeros. The remaining relevant terms for computing with the $p+q = 5$ anti-diagonal are shown. (The diagram comes from [OP22].)

q							
5	0	0	0	0	0	0	0
4		0		0		0	
3	0	0	0	0	0	0	0
2		0	$H_2(K;\mathbb{Z}/2)$	0 $\boldsymbol{d_2^{4,1}}$		0	
1		0		0	$H_4(K;\mathbb{Z}/2)$	0 $\boldsymbol{d_2^{6,0}}$	
0		0		0		0	$H_6(K;\mathbb{Z})$
	0	1	2	3	4	5	6 p

We have to show that homology of the sequence $H_6(K;\mathbb{Z}) \to H_4(K;\mathbb{Z}/2) \to H_2(K;\mathbb{Z}/2)$ vanishes at the central term. Then we will see that $E^3_{4,1} = 0$, and the proposition will follow.

The following computation come from [OP22]. Choose any basis $\langle y_1, \dots, y_m \rangle$ of $\pi_2(X)$. Recall that in the definition of (B, ξ) there is a reference map $\eta\colon X \to K$ that is an isomorphism on π_2. Thus $\langle y_1, \dots, y_m \rangle$ determines a basis $\langle \eta_*(y_1), \dots, \eta_*(y_m) \rangle$ of $\pi_2(K)$. Write $e^0_i \cup e^2_i \cup e^4_i \cup \cdots$ for the standard cell structure of the i^{th} $\mathbb{C}P^\infty$ factor of $\prod_{i=1}^m \mathbb{C}P^\infty$. Choose a homotopy equivalence $g\colon K \simeq \prod_{i=1}^m \mathbb{C}P^\infty$, such that $g_*(k_*(y_k)) = [e^2_i]$. Write $x_i \in H^2(\prod_{i=1}^m \mathbb{C}P^\infty; \mathbb{Z}/2)$ for the class that is $\mathbb{Z}/2$-dual of $[e^2_i]$. There are isomorphisms of graded algebras

$$(\mathbb{Z}/2)[x_1, \dots, x_m] \xrightarrow{\cong} H^*(\textstyle\prod_{i=1}^m \mathbb{C}P^\infty; \mathbb{Z}/2) \xrightarrow[f^*]{\cong} H^*(K; \mathbb{Z}/2),$$

where the first isomorphism is given by the Künneth theorem. In an abuse of notation, we will refer to the images of the x_i in $H^*(K;\mathbb{Z}/2)$ by the same name. Thus, we have bases

$$\begin{aligned} H^2(K;\mathbb{Z}/2) &\cong \langle x_i \,|\, 1 \le i \le m \rangle, \\ H^4(K;\mathbb{Z}/2) &\cong \langle x_i \cup x_j \,|\, 1 \le i \le j \le m \rangle, \\ H^6(K;\mathbb{Z}/2) &\cong \langle x_i \cup x_j \cup x_k \,|\, 1 \le i \le j \le k \le m \rangle. \end{aligned}$$

By a theorem of Teichner [Tei92, Tei97], the differentials $d_2^{4,1}$ is the Hom-dual to the map $x_i \mapsto \mathrm{Sq}^2(x_i) = x_i \cup x_i$. Writing $(-)^*$ for the dual as a $\mathbb{Z}/2$-vector space, we see that

$$d_2^{4,1}([x_i \cup x_j]^*)(x_k) = [x_i \cup x_j]^*(x_k \cup x_k) = \begin{cases} 1 & \text{if } i = j = k \\ 0 & \text{otherwise.} \end{cases}$$

In particular, this shows $\ker(d_2^{4,1})$ has basis $\langle [x_i \cup x_j]^* \,|\, 1 \le i < j \le m \rangle$. Teichner also showed that the differential $d_2^{6,0}$ is reduction modulo 2 followed by the dual to the map Sq^2. The Cartan formula shows

$$\begin{aligned} \mathrm{Sq}^2(x_i \cup x_j) &= \mathrm{Sq}^0(x_i) \cup \mathrm{Sq}^2(x_j) + \mathrm{Sq}^1(x_i) \cup \mathrm{Sq}^1(x_j) + \mathrm{Sq}^2(x_i) \cup \mathrm{Sq}^1(x_j) \\ &= x_i \cup x_j \cup x_j + x_i \cup x_i \cup x_j. \end{aligned}$$

Using this, we have

$$(\mathrm{Sq}^2)^*([x_i \cup x_j \cup x_k]^*)(x_r \cup x_s) = [x_i \cup x_j \cup x_k]^*(x_r \cup x_s \cup x_s + x_r \cup x_r \cup x_s).$$

By inspection, for $1 \le i < j \le m$ and $1 \le r \le s \le m$, we see that $(\mathrm{Sq}^2)^*([x_i \cup x_j \cup x_j]^*)$, projects nontrivially to the span of $[x_r \cup x_s]^*$ if and only if $(i,j) = (r,s)$. Thus $(\mathrm{Sq}^2)^*([x_i \cup x_j \cup x_j]^*) = [x_i \cup x_j]^*$. From this we can conclude that $\mathrm{Im}(d_2^{6,0}) = \ker(d_2^{4,1})$, so that $E_3^{4,1} = 0$. Hence the $r + s = 5$ line dies already on the E_3 page and it follows that $\Omega_5^{\mathrm{TopSpin}}(K) = 0$, as desired. □

This completes our discussion of the proof that if X is a closed, 1-connected, 4-manifold and $f\colon X \to X$ is an o.p. homeomorphism with $f_* = \mathrm{Id}_{H_2(X)}$, then f is pseudo-isotopic to Id_X.

Remark 7.3.

(1) The non-spin case is proved similarly. The main difference is that the d_2-differentials are twisted by the second Steifel-Whitney class, w_2. e.g. the differential $H_4(K;\mathbb{Z}/2) \to H_2(K;\mathbb{Z}/2)$ is dual to the map $H^2(K;\mathbb{Z}/2) \to H^4(K;\mathbb{Z}/2)$ given by $x \mapsto \mathrm{Sq}^2(x) + w_2 \cup x$. This changes the computation of the differentials, but the outcome, that the relevant bordism group vanishes, is the same. Details can be found in [Kre79] or [OP22].
(2) The entire proof works smoothly. Indeed, if X and $f\colon X \to X$ are smooth, then under the same hypotheses f is smoothly pseudo-isotopic to the identity. What emphatically does not work smoothly is the upcoming pseudo-isotopy implies isotopy theorem.
(3) In work with Orson [OP22], we extended the above method to the case of compact 1-connected 4-manifolds with nonempty boundary. I will outline the main result of this work at the end of these notes.
(4) In work with Kreck and Orson, in progress, we extend this to the case of fundamental group F_n, the free group on rank n.
(5) In both of the last two items, there are extra obstructions beyond the intersection form. For example, let $X = S^1 \times S^3$. Then we can perform a Dehn twist on the non-separating S^3. This acts trivially on π_1 and π_2, but is not pseudo-isotopic to the identity. So it is certainly not enough to consider the automorphisms of the intersection pairing. There is another very interesting invariant due to Stong and Wang [SW00].
(6) What is the classification for other fundamental groups? For example, for finite cyclic fundamental group $\mathbb{Z}/n$?
(7) Let $X = X' \# S^2 \times S^2$, and suppose that $H_1(X;\mathbb{Z}/2) \neq 0$. Then Daniel Galvin proved that there is a homeomorphism of X that is not pseudo-isotopic to any diffeomorphism. He did this by realising the *Casson-Sullivan invariant.*
(8) Let $X = (T^2 \times S^2) \#^k (S^2 \times S^2)$ (for some k that is hard to specify explicitly). Then in work with Orson, we found a diffeomorphism of X that is topologicially pseudo-isotopic to the identity but is not smoothly pseudo-isotopic. If we stabilise by more copies of $S^2 \times S^2$, it will never become smoothly isotopic to the identity.
(9) In the last two items, we implicitly considered the map

$$\widetilde{\pi}_0(\mathrm{Diff}^+(X)) \to \widetilde{\pi}_0(\mathrm{Homeo}^+(X)).$$

There exist choices for X such that this map is neither injective nor surjective. For the failure of both injectivity and surjectivity, the results use a version of the Kirby-Siebenmann obstruction. The classical Kirby-Siebenmann obstruction is an element of $H^4(X, \partial X;\mathbb{Z}/2)$, which for 4-manifolds is $\mathbb{Z}/2$, and it obstructs the existence of a smooth structure on X [FQ90]. The different versions are all smoothing obstructions, but the Kirby-Siebenmann class appears in different guises depending on whether it is obstructing the smoothing of manifolds, homeomorphisms, or isotopies. In all cases, because it is defined in a highly non-explicit manner, it is nontrivial to evaluate.
(10) Krannich-Kupers [KK24] showed that there are examples of closed 4-manifolds X, with a diffeomorphism $f\colon X \to X$, such that f is homotopic but not pseudo-isotopic to Id.

8. Pseudo-isotopy implies isotopy

Recall that we consider the composition

$$\pi_0 \operatorname{Homeo}^+(X) \to \widetilde{\pi}_0 \operatorname{Homeo}^+(X) \to \operatorname{Aut}(H_2(X), \lambda_X).$$

We have shown that the second map is an isomorphism. We know the first map is surjective. The next theorem implies that the first map is injective, so it will complete our main aim, to show that

$$\pi_0 \operatorname{Homeo}^+(X) \xrightarrow{\cong} \operatorname{Aut}(H_2(X), \lambda_X).$$

Theorem 8.1 (Perron [Per86], Quinn [Qui86], Gabai-Gay-Hartman-Krushkal-P [GGH+23], Gabai [Gab22])**.** *Let X be a 1-connected, compact 4-manifold, let $f\colon X \to X$ be a homeomorphism with $f|_{\partial X} = \operatorname{Id}_X$. Then f is topologically isotopic to the identity. In fact, every pseudo-isotopy $F\colon X \times I \to X \times I$ is isotopic to the identity rel. $\sqsubset$.*

Moreover, if X, f, and F are smooth, then f is stably smoothly isotopic to the identity.

The first paragraph is due to Perron [Per86] and Quinn [Qui86], with a correction to Quinn's proof in [GGH+23]. It is worth noting that Perron's proof uses Quinn's earlier work in [Qui82] heavily.

The second paragraph is due to Quinn [Qui86], again with a correction in [GGH+23]. Gabai [Gab22] gave an alternative proof of the second paragraph.

Here we say that f is smoothly stably isotopic to Id if there exists k such that, extending by the identity on $\#^k S^2 \times S^2$, $f \# \operatorname{Id}\colon X \#^k S^2 \times S^2 \to X \#^k S^2 \times S^2$ is smoothly isotopic to the identity.

In particular, the first conclusion implies the injectivity we seek. I will attempt to explain the proof that every pseudo-isotopy $F\colon X \times I \to X \times I$ is isotopic to the identity rel. $\sqsubset$. First, let me give some context for this result.

The first result along these lines was the following, in high dimensions.

Theorem 8.2 (Cerf [Cer70])**.** *Let X be a smooth compact n-manifold, with $n \geq 5$. Let $f\colon X \to X$ be a diffeomorphism with $f|_{\partial X} = \operatorname{Id}_{\partial X}$ and such that f is pseudo-isotopic to Id_X. Then f is smoothly isotopic to Id_X.*

Perron [Per86] and Quinn [Qui86] were attempting to adapt Cerf's result to dimension 4. The fact that the conclusion was a topological isotopy, and not a smooth one, was inevitable, given the following result.

Theorem 8.3 (Ruberman [Rub98])**.** *There exists a smooth, closed, 1-connected 4-manifold X with an infinitely generated subgroup of*

$$\ker\big(\pi_0(\operatorname{Diff}^+(X)) \to \pi_0(\operatorname{Homeo}^+(X))\big).$$

For non-trivial fundamental group, there is no analogue of Theorem 8.1, and the following result shows that this is not possible.

Theorem 8.4 (Budney-Gabai [BG19])**.** *There exists a subgroup $\mathbb{Z}^\infty \leq \pi_0(\operatorname{Homeo}^+(S^1 \times S^3))$, which maps trivially to $\widetilde{\pi}_0(\operatorname{Homeo}(S^1 \times S^3)) \cong \mathbb{Z}/2$.*

This is somewhat expected, given the prior work of Hatcher-Wagoner-Igusa [HW73, Igu84]. In dimensions at least 6, they obtained an exact sequence as follows, where $\mathcal{P}(X)$

denotes the space of pseudo-isotopies of X. So $\pi_0(\mathcal{P}(X))$ is pseudo-isotopies up to isotopy rel. $\sqsubset$. The Hatcher-Wagoner-Igusa sequence is:

$$K_3(\mathbb{Z}\pi_1(X)) \to \mathrm{Wh}_1(\pi_1(X); \mathbb{Z}/2 \times \pi_2(X)) \to \pi_0(\mathcal{P}(X)) \to \mathrm{Wh}_2(\pi_1(X)) \to 0.$$

I will not define the terms here. The terms other than $\pi_0(\mathcal{P}(X))$ are algebraic obstruction groups, related to K-theory.

Work of Singh [Sin21] and Igusa [Igu21] has made some progress towards understanding how much of this sequence holds in dimension 4, but there is a long way to go.

9. Cerf theory

Now we begin the proof of Theorem 8.1. Recall that this theorem is analogous to the h-cobordism theorem. Whereas the h-cobordism theorem uses Morse theory, this uses *Cerf theory*, which is essentially the 1-parameter Morse theory.

In this proof we are going to restrict to the case that X is closed and smooth, and $F\colon X \times I \to X \times I$ is smooth pseudo-isotopy. One still learns most of the exciting ideas behind the proof from the smooth input – topological output version. This is not a necessary assumption; the theorem is true as stated. It is just to make the exposition manageable for these notes.

So let X be a smooth, closed, 1-connected 4-manifold, and let $F\colon X \times I \to X \times I$ be a smooth pseudo-isotopy. We have two Morse functions on $X \times I$ without critical points. First,

$$g_0 := \mathrm{pr}_2 \colon X \times I \to I$$

given by projection to the second coordinate. Next,

$$g_1 := \mathrm{pr}_2 \circ F \colon X \times I \to I$$

also has no critical points.

Recall that a Morse function $g\colon X \times I \to \mathbb{R}$ is a smooth function such that every critical point $p \in X \times I$, the Hessian matrix of second partial derivatives is nondegenerate (this condition is coordinate independent). Near a critical point we have coordinates $(x_1, \dots, x_5$ and $h \in \{0, \dots, 5\}$ such that

$$g(\underline{x}) = g(p) - x_1^2 - \cdots - x_h^2 + x_{h+1}^2 + \cdots x_5^2.$$

Here h is the *index* of the critical point.

Next, any two Morse functions on a manifold, so in particular on $X \times I$, can be connected by a 1-parameter path of *generalised Morse functions*

$$g_t \colon X \times I \to \mathbb{R}.$$

This is a smooth path of functions, such that for each $t \in I$, either g_t is a Morse function, or g_t is a Morse function everywhere except possibly at one critical point p, which is a *birth/death* type singularity. At p we have coordinates $(x_1, \dots, x_5)$ such that in the coordinates

$$g_t(\underline{x}) = g(p) + x_1^3 - x_2^2 - \cdots - x_h^2 + x_{h+1}^2 + \cdots x_5^2.$$

In a 1-parameter family near p, we can assume that g_{t+s} has the form

$$g_{t+s}(\underline{x}) = g(p) + x_1^3 \pm s x_1 - x_2^2 - \cdots - x_h^2 + x_{h+1}^2 + \cdots x_5^2.$$

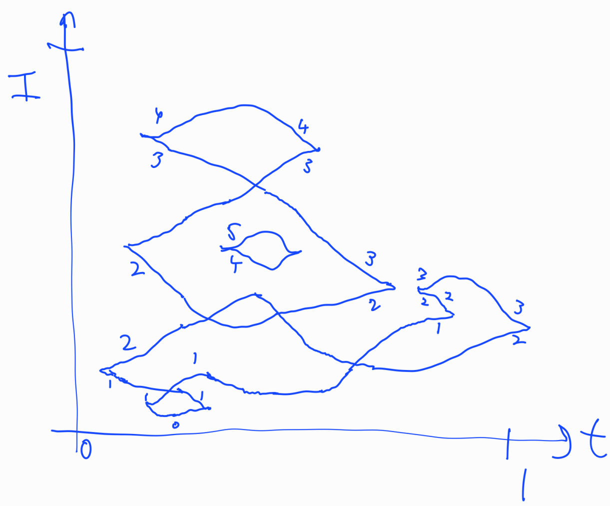


FIGURE 1. A generic Cerf graphic. I label the critical lines with the indices of the corresponding critical points.

Then $\pm s$ is negative we see two Morse critical points, and when $\pm s$ is positive there are no critical points in the coordinate neighbourhood.

There is a corresponding 1-parameter family of *gradient-like vector fields* (glvf) ξ_t on $X \times I$, for $t \in [0, 1]$, such that ξ_t is a glvf for g_t. Using the glvf, we obtain, for each t where g_t is Morse, a handle decomposition of $X \times I$, where each critical point of index h gives rise to an h-handle. Trajectories of ξ_t between handles of index $(h + 1)$ and h correspond to attaching data, namely intersections of the attaching sphere of the index $(h+1)$-handle with the belt sphere of the index h handle. Trajectories of ξ_t between two index h handles can occur at isolated t-values, and correspond to handle slides.

Some of the data of g_t can be presented in a *Cerf graphic*, as shown in Figure 1. Here we plot with two axes, $t \in [0, 1]$, and the interval I in which our generalised Morse functions g_t take values. For each $t \in [0, 1]$, we consider all the critical points of g_t, namely $P_t := \{p_i \mid Dg_t(p_i) = 0\}$. Then we plot the critical values $g_t(P_t)$ at t. Doing this for every $t \in [0, 1]$ gives rise to the Cerf graphic. One cannot recover the Morse function from the Cerf graphic, but it turns out to contain some useful information that makes it easier to describe the qualitative features of our family g_t, and to describe the key features deformations that we wish g_t to undergo.

We can assume, generically, that two critical points have the same critical value at isolated value of t, and that in this case the critical lines in the Cerf graphic intersect transversely. Important data that is not shown in the graphic is the trajectories between critical points. We will not indicate this data in the Cerf graphic, although one is free to invent schema to do so.

Whenever we change (g_t, ξ_t), we speak of a *deformation* of the family. Our aim will be to deform (g_t, ξ_t), without changing (g_i, ξ_i) for $i = 0, 1$, to a family with no critical points, i.e. one with empty Cerf graphic. Why is this our aim?

Given a glvf without critical points, we can integrate it in order to start with $X \times \{0\}$, and flow it along the integral curves to obtain a self-homeomorphism of $X \times I$. There is also a topological version of this, which we will not investigate.

If we integrate ξ_0, we obtain a homeomorphism of $X \times I$ isotopic to $\mathrm{Id}_{X\times I}$. If we integrate ξ_1, we obtain a homeomorphism of $X \times I$ isotopic to $F\colon X \times I \to X \times I$. Moreover, if ξ_t has no critical points, then for each t we obtain a homeomorphism $F_t\colon X \times I \to X \times I$, with $F_0 = \mathrm{Id}$ and $F_1 = F$. Since ξ_t depends continuously on t, we obtain an isotopy between Id and F, as desired. It follows that it suffices to deform (g_t, ξ_t) rel. $t = 0, 1$ to a family with no critical points, as asserted.

The first step is contained in the following proposition.

Proposition 9.1 (Cerf, Hatcher-Wagoner)**.** *For X with $\pi_1(X) = \{1\}$, there exists a deformation of (g_t, ξ_t) to a family with a nested eye graphic.*

Here, a *nested eye graphic* is one of the form shown in the next figure. All births come first, one at a time, and they are all births of cancelling index 2 and 3 pairs of critical points. There are no rearrangements of critical values, and no handle slides, i.e. there are no 2/2 and no 3/3 trajectories. The births and deaths are independent, meaning that at each birth time and each death time, there are no trajectories that go either from or to the birth or death point, from or to another critical point. Each circle in the figure is called an *eye*.

The procedure to arrange the nested eye graphic is somewhat complicated, involving a careful use of codimension 2 singularities. These deformation are to generalised Morse functions as generalised Morse functions are to ordinary Morse functions. The procedure works in all dimensions at least 4. Here we are going to quote it and not attempt to justify it.

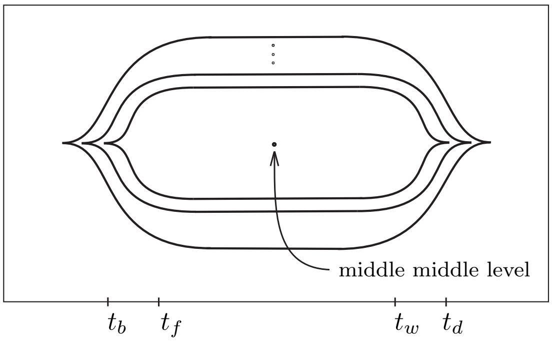


FIGURE 2. A nested eye Cerf graphic. Each loop is called an *eye*. The middle-middle level is also indicated, as well as the birth time t_b, the death time t_d, the finger move time t_f and the Whitney move time t_w. (Picture from [GGH+23].)

Figure 2 shows a nested eye Cerf graphic, and it indicates the times at which births and deaths appear. The additional labels will be explained carefully in the next section.

10. The middle-middle level

In the proof of the h-cobordism theorem, for 5-dimensional h-cobordisms, we put a Morse function $f\colon W \to I$ on our h-cobordism W, and then perform handle trading to arrange that there are only 2- and 3-handles. We then consider the *middle level* $f^{-1}(1/2)$, in which we see two sets of mutually disjointly embedded 2-spheres. The first, $A_1, \dots, A_k$, are the ascending spheres of the index 2 critical points. In handle language, they are the belt spheres of the 2-handles. The second set of mutually disjointly embedded 2-spheres, $B_1, \dots, B_k$, are the descending spheres of the index 3 critical points, or the attaching spheres of the 3-handles. After some handle sliding we may assume that the intersection numbers of these spheres are $A_i \cdot B_j = \delta_{ij} \in \mathbb{Z}$, since W is an h-cobordism. The goal of the proof is to arrange by an isotopy that the geometric intersection numbers of these spheres agrees with the algebraic intersection numbers. Then we can cancel the critical points in 2-3 pairs, to obtain a cobordism with no critical points, which is hence a product. The proof of the pseudo-isotopy theorem is analogous to this, but one level of complexity higher.

We also consider the middle level, $g_t^{-1}(1/2)$, for each t. It turns out that the data of the pseudo-isotopy can be captured in the *middle-middle level*, which is the inverse image

$$M := g_{1/2}^{-1}(1/2).$$

This is the inverse image of the central point shown in Figure 2. This is a 4-manifold diffeomorphic to $X\#^k(S^2 \times S^2)$, where k is the number of 2-3 pairs.

Right after the birth time t_b, in the middle level $g_{t_b+\varepsilon}^{-1}(1/2) \cong X\#^k(S^2 \times S^2)$, we see the ascending 2-spheres $A_1, \dots, A_k$ of the index 2-critical points, of the form $\{\mathrm{pt}\} \times S^2$ in each of the $S^2 \times S^2$ summands. We also see the descending spheres $B_1, \dots, B_k$ of the index 3 critical points, of the form $S^2 \times \{\mathrm{pt}\}$ in each of the $S^2 \times S^2$ summands. Note that A_i and B_j intersect in exactly δ_{ij} points, so the critical points are in cancelling position. The situation is the same just before the death time, in $g_{t_d-\varepsilon}^{-1}(1/2) \cong X\#^k(S^2 \times S^2)$.

In between, looking at the ascending and descending spheres in $g_t^{-1}(1/2)$, for $t \in [t_b + \varepsilon, t_d - \varepsilon]$, we can assume that the $\{A_i\}$ stay fixed, and the $\{B_j\}$ move around by an isotopy, while staying pairwise disjoint and embedded. During this motion extra intersections between the $\{A_i\}$ and the $\{B_j\}$ can appear, and later disappear. At the start and end we know that the spheres are in cancelling position.

We can also assume that all the extra intersections are created first, and at the same time t_f, by *finger moves*; see Figure 3.

Then the extra intersections are all removed simultaneously by Whitney moves, at the time t_w. A Whitney move is guided by a Whitney disc W; see Figure 4.

Also, after a finger move there is a finger-move Whitney disc V, with the property that performing a Whitney move on that finger-move disc V undoes the finger move. In the middle-middle level, we see two collections of discs, the Whitney discs $\{W_\ell\}$, that guide the Whitney moves that will happen at time $t = t_w$. After the Whitney move the $\{A_i\}$ and the $\{B_j\}$ are in cancelling position. We also see the finger-move discs $\{V_m\}$, which have the property that reversing time leads to Whitney moves using them. These Whitney moves also remove all excess intersections between the $\{A_i\}$ and $\{B_j\}$.

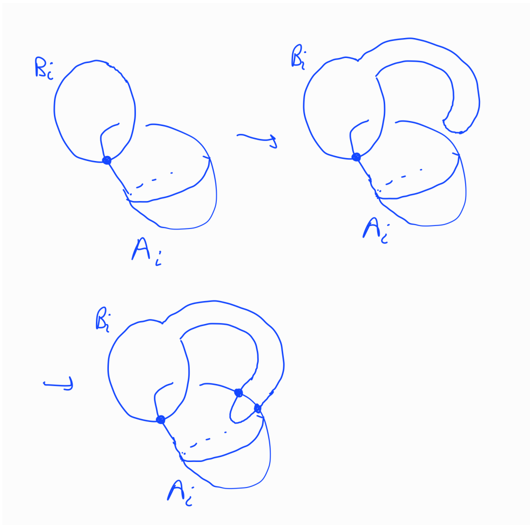


FIGURE 3. A finger move.

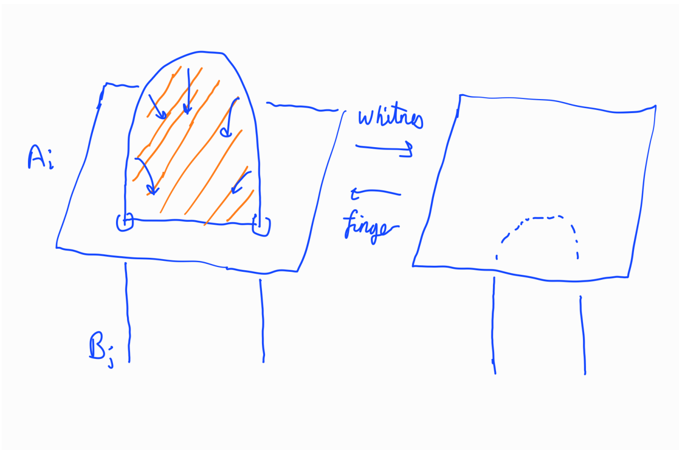


FIGURE 4. The Whitney move goes from left to right, and the finger move reverses it, going from right to left.

All of the important data about the pseudo-isotopy can now be captured by the spheres $\{A_i\}$ and $\{B_j\}$ in the middle-middle level, which intersect in δ_{ij} times algebraically, together with the two collections of Whitney discs $\{V_m\}$ and $\{W_\ell\}$. Each of these collections have mutually pairwise disjoint interiors, and has the property that using them for Whitney moves places the $\{A_i\}$ and $\{B_j\}$ in cancelling position.

We will prove the following proposition in the next three sections.

Proposition 10.1. *We can perform a deformation to remove/cancel the innermost eye.*

In order to be able to remove the innermost eye, it suffices that there is a unique trajectory between the index 3 and the index 2 critical point corresponding to that eye.

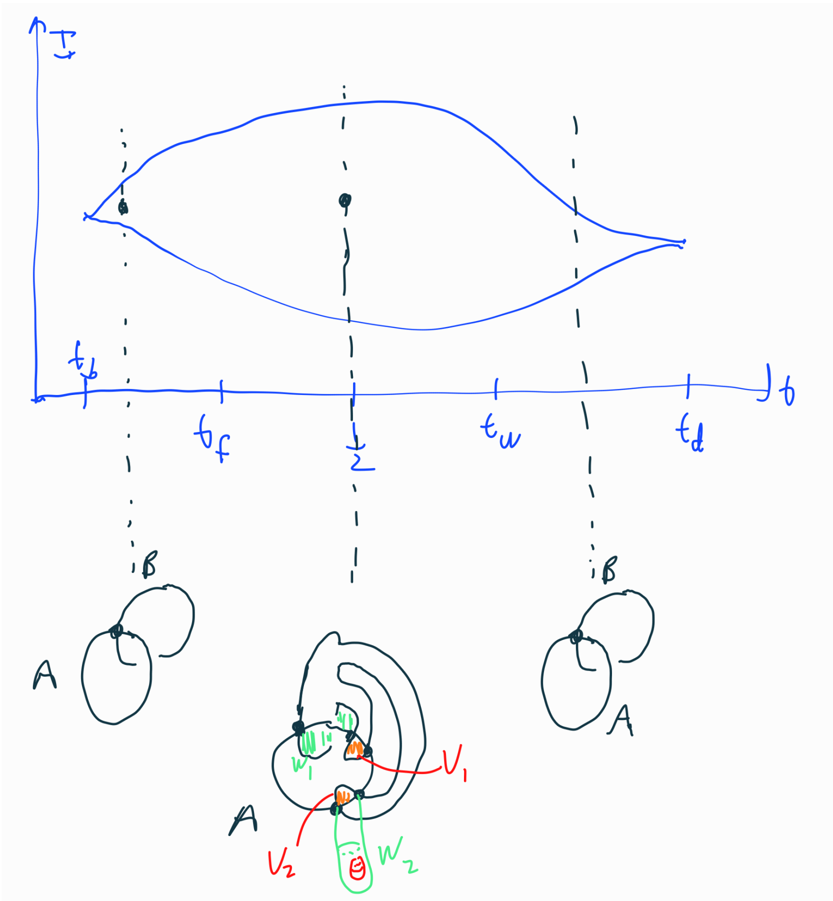


FIGURE 5. A Cerf graphic with one eye, together with schematics of the spheres in the middle levels at different time values. In the schematic, spheres are represented by circles. The key data is contained in the middle-middle level, where we see finger and Whitney discs that guide the pseudo-isotopy in the past and future respectively. The Whitney discs and the finger discs shown are distinct. First, W_1 and V_1 pair up the double points in a different way. Secondly, the Whitney discs V_2 and W_2 are assumed to have the same boundary (although this need not be the case, a priori), but thei union represent a nontrivial element in $\pi_2(M)$. This is supposed to be indicated by the small red sphere.

That is, the critical points must be in cancelling position the entire time interval for which they exist. Then we can perform a 1-parameter families worth of cancellations, to entirely remove the eye.

Note that the proposition implies the theorem, because we inductively close the eyes, always working on the innermost eye in the Cerf graphic. With this in mind, from now on to simplify the exposition we pretend that $k = 1$, i.e. that the Cerf graphic consists of a single eye. If we can remove this eye, we will be able to apply the same argument to remove the eyes one at a time, in the general case.

With $k = 1$, we set $A := A_1$ and $B := B_1$ to be the spheres in the middle-middle level M. We still have two families of discs $\{V_m\}$ and $\{W_\ell\}$. We write $V := \bigcup\{V_m\}$ and $W := \bigcup\{W_\ell\}$. The goal is to arrange, by a deformation of the family, for a particular configuration of the finger and Whitney discs, which will enable us to remove them, and whence arrange that the critical points be in cancelling position the entire time interval for which they exist.

11. Geometric moves and Quinn's embedded arc criterion

We need introduce some geometric moves, that we will use in the proof of Proposition 10.1, and we discuss Quinn's arc condition. First, we construct some dual spheres.

11.1. **Dual spheres.** Directly after the birth time, $A \cup B$ are in standard position. Their normal bundles are trivial. We may therefore consider another sphere A' pushed-off from A using a nonvanishing section of its normal bundle. Note that A' intersects B at precisely one point, is framed and embedded, and is disjoint from A.

As the family evolves, at t_f several finger moves of B through A occur. We may assume by general position that the arcs guiding all the finger moves are disjoint from A'. After the finger moves, allow A' to move, via isotopy extension, so that it remains geometrically dual to B, i.e. intersects B in exactly one point. We also keep A' disjoint from all finger discs V. In the middle-middle level M, call the resulting sphere $D^B_{A,V}$. The notation D stands for dual, the superscript is to remind us that it is a dual sphere to B, and the subscripts tell us that the dual sphere is disjoint from $A \cup V$. We do not control the intersections of $D^B_{A,V}$ with W, other than arranging that any such intersections in M are transverse.

By a similar construction we can obtain a dual sphere $D^A_{B,V} \subseteq M$. By starting near the death time and going backwards in time, we obtain similar dual spheres $D^B_{A,W}$ and $D^A B, W$, that are disjoint from W but may intersect V transversely. I will make copious use of these dual spheres later.

11.2. **Embedded arc criterion.** Recall that we consider a 1-parameter family whose Cerf graphic consists of a single eye, the data of which is captured in the middle level by two framed, embedded algebraically dual spheres A and B, together with two collections V and W of Whitney discs leading to geometrically dual A and B.

The next major simplification we seek to make to our 1-parameter family (g_t, ξ_t) concerns the boundary arcs of the discs V and W on A and on B. That is, we consider

$$\Gamma := \bigcup_m \partial V_m \cup \bigcup_\ell W_\ell \subseteq A \cup B.$$

The restrictions $\Gamma \cap A$ and $\Gamma \cap B$ consist of an immersed arc and a collection of immersed circles in a 2-sphere, A and B respectively.

Proposition 11.1 (Quinn's embedded arc condition)**.** *There exists a deformation to that $\Gamma \cap A$ and $\Gamma \cap B$ each consist of an embedded arc, with no circles.*

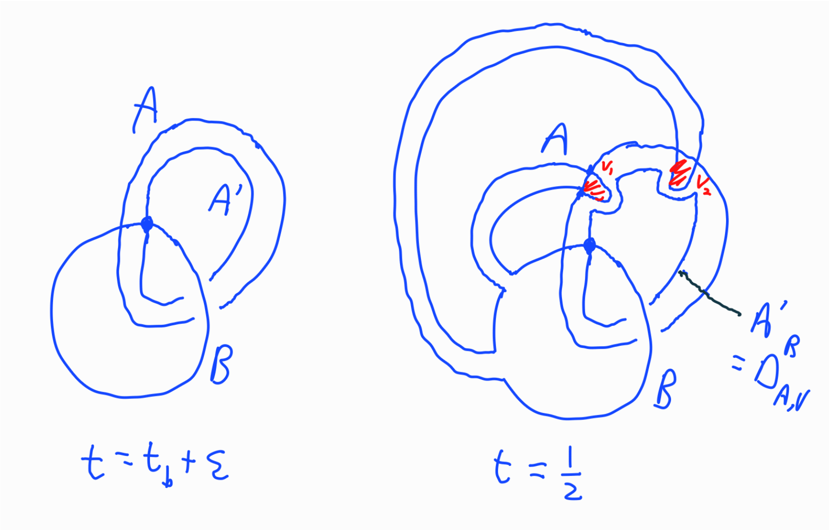


FIGURE 6. The construction of the dual sphere $D^B_{A,V} \subseteq M$. Start with a push off of A, and consider its image in M using the isotopy extension theorem to keep it dual to B and disjoint from $A \cup V$.

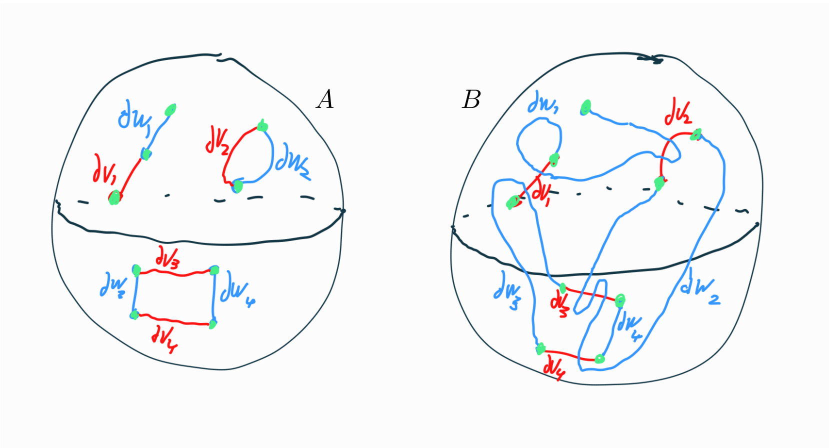
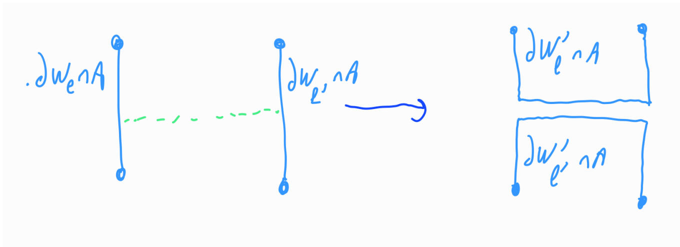


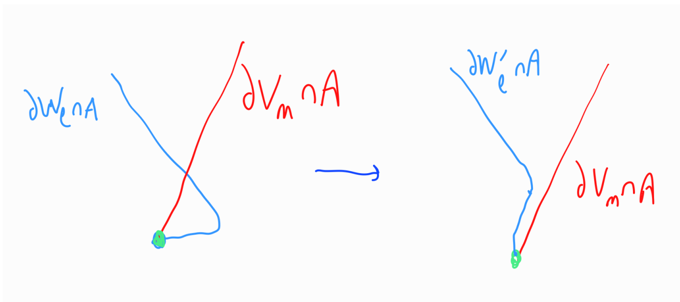
FIGURE 7. A somewhat generic picture of a possible configuration of boundary arcs of $V \cup W$, on the left on A, and on the right on B. What is not generic is that the arc and the circles on A are embedded. I drew it like this to show some basic configurations. In all likelihood, the situation will be more like that shown on B, with multiple intersections between the boundaries of the A and B discs. Note that $\partial V \cap A$ and $\partial W \cap B$ are both collections of embedded arcs. The failure to be embedded comes from considering $\partial V \cap B$ and $\partial W \cap B$ simultaneously (and in general, the same would hold when considering $\partial V \cap A$ and $\partial W \cap A$ simultaneously).

To achieve this, we will use the sum square move, which we introduce in the next section. It is a way of modifying either a pair of discs in W, or a pair of discs in V. The sum square move done to a pair of discs in W has the following effect on the intersection of the boundaries with A.



The effect on the intersection of ∂W with B is similar.

We can also perform a push of an arc as shown in the next picture, when an intersection between an arc of $\partial V \cap A$ and an arc of $\partial W \cap B$ is the closest such intersection to the preimage on A of an intersection point $A \cap B$. The same move can be done on B, of course.



Given the ability to make these somewhat drastic changes, although we will not elaborate, it should be plausible that some combinatorial argument exists to arrange the boundaries of the discs into a pleasant configuration. The details are in [Qui86].

11.3. **The sum square move.** In this section I explain Quinn's sum square move [Qui86, Section 4.2], which can be used to modify finger or Whitney disc configurations by a deformation of the pseudo-isotopy. The description I give is similar to that in [GGH+23]. The data for the move is a framed embedded square S in the middle-middle level, with interior disjoint from the spheres A and B, and from the discs V. The square has two edges on the V discs, denoted W_1 and W_2 in Figure 8, one edge on A, and one on B. New W discs are obtained by cutting W_1, W_2 along the boundary edges of the sum square S, and gluing in two parallel copies of S. The effect of the move on the boundaries of the discs, on A and B spheres, was illustrated above.

Figure 8 is a 3-dimensional model for the sum square. Here A, W_1, V_2, and a neighbourhood of the arc of ∂S in B are pictured in the "present", $\mathbb{R}^3 \times \{0\} \subseteq \mathbb{R}^3 \times \mathbb{R}$. The rest of B extends into the past and the future. The framing of S along its boundary is determined in the 3-dimensional model by a non-vanishing vector field on ∂S which is normal to S and tangent to A, B and the V discs; this framing has to admit an extension over S for the move to give rise to embedded V discs.

Given the boundary data ∂S, the challenge is to find a framed, embedded S with interior disjoint from $A \cup B \cup W$. (If the sum square is used to modify V discs then the interior of

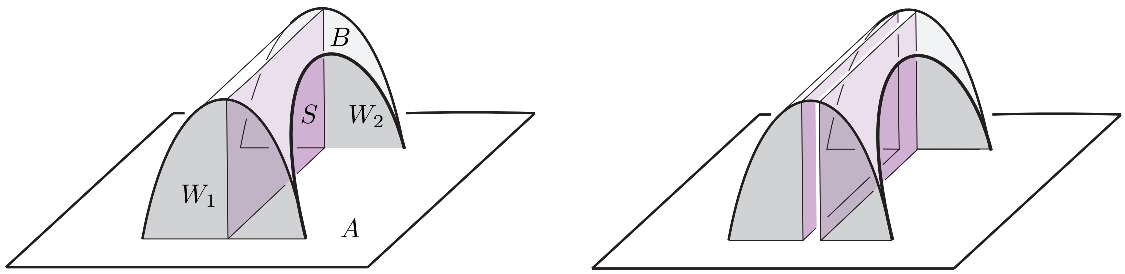


FIGURE 8. The sum square move. (Picture from [GGH+23]).

S must be disjoint from V.) For this we make use of the dual sphere $D^A_{B,W}$ and also the Whitney spheres that I will introduce in the next section.

To justify that the sum square move arises via a deformation of the pseudo-isotopy, we consider how the Whitney discs are deformed. The key intermediate step is shown in Figure 9.

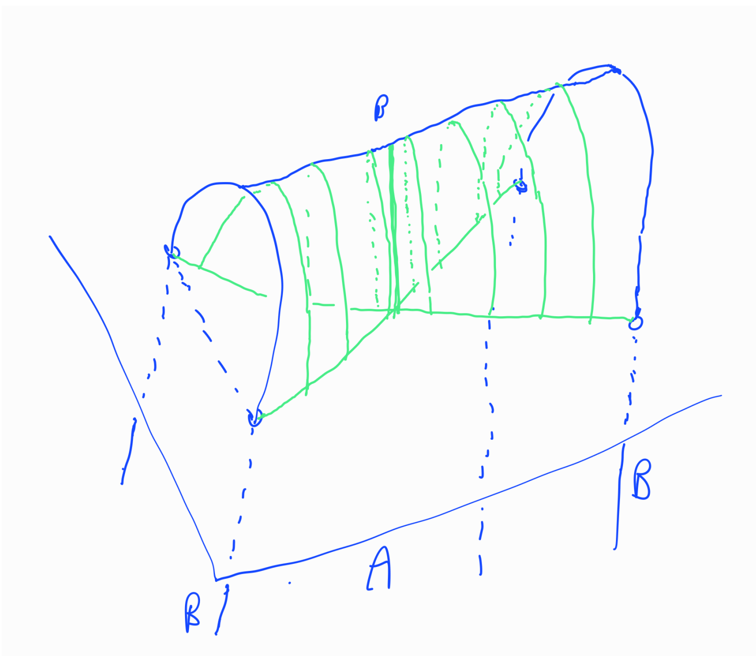


FIGURE 9. The intermediate step in the deformation associated to a sum square move. The two Whitney discs on the left of Figure 8 have been deformed into the green singular Whitney umbrella. Afterwards, the singularity resolves and they become the discs on the right of Figure 8.

Before and after this step, we have two Whitney moves, on one pair of discs beforehand shown on the left of Figure 8, which get deformed closer together, and on a different pair of discs shown on the right of Figure 8 afterwards. At the moment of transfer between the two discs, we have a singular Whitney umbrella, and the family performs a kind of singular double-Whitney move guided by the umbrella.

11.4. **Whitney spheres.** We recall a construction of the *Whitney sphere* S_{V_m} associated with a finger move disc V_m. These spheres have appeared in different guises in the literature, cf. [Qui86, Section 4.3], [FQ90, Section 3.1, Ex. (2)], [COP20, Section 4.2], [ST19, Section 2]. We use the terminology and the description from [ST19] and [GGH$^+$23].

The description is given in $\mathbb{R}^3 \times \mathbb{R}$ where A and the finger disc V_m are in $\mathbb{R}^3 \times \{0\}$, and B is represented as (arc $\subseteq \mathbb{R}^3$)$\times[-1,1]$; see Figure 10. The Whitney sphere is drawn red, and consists of two discs, $D_i \subseteq \mathbb{R}^3 \times \{i\}$, $i = -1, 1$, joined by an annulus (circle $\subseteq \mathbb{R}^3$)$\times[-1,1]$. Each D_i is constructed using two copies of the finger move disc V_m, so overall the Whitney sphere contains four pushed-off copies of V_m. The Whitney sphere is framed and embedded and can be assumed to lie in an arbitrarily small neighbourhood of V_m.

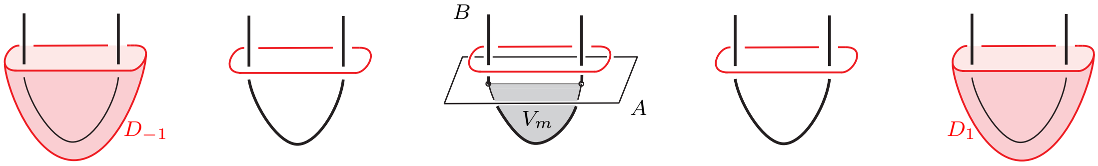


FIGURE 10. A description of the Whitney sphere S_{V_m} in $\mathbb{R}^3 \times \mathbb{R}$. (Picture from [GGH$^+$23]).

There is a similar construction of a sphere S_{W_ℓ} associated with each disc W_ℓ.

12. Algebraically cancelling finger-Whitney intersections

Let us assume that Quinn's embedded arc condition has been arranged. The configuration in M is as shown in Figure 11, for the case of 4 extra intersection points in M.

The next step is to arrange that for each m and for each ℓ, the following algebraic count of transverse intersections vanishes:

$$\mathring{V}_m \cdot \mathring{W}_\ell = 0.$$

This is not at all clear. In his original proof, Quinn employed an idea (that he credited to Igusa, based on a preprint apparently circulated by Igusa in the early 1980s, but which I

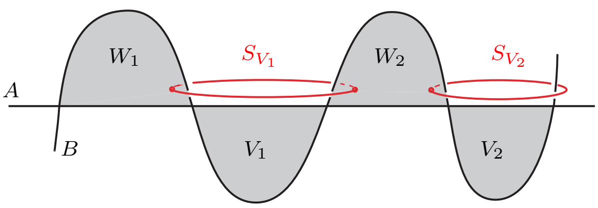


FIGURE 11. Finger and Whitney discs with boundaries forming an arc in A and in B. Whitney spheres corresponding to the finger move discs are also shown. (Picture from [GGH$^+$23]).

have not seen), called *disc replacement* [Qui86, Section 4.5]. Unfortunately the proposed proof of disc replacement is flawed. The proof does not use the simply-connected hypothesis, nor the fact that the embedded arc condition holds. Therefore the same argument, if correct, would lead to a contradiction with the Hatcher-Wagoner-Igusa sequence.

It is still possible that the disc replacement lemma holds, and it is a very interesting open question. In work with Gabai, Gay, Hartman, and Krushkal, we were able to provide an alternative argument, that I will explain here [GGH$^+$23]. In that paper we also explained the problem with the disc replacement proof carefully.

Example 12.1. It can certainly happen in examples that the algebraic intersection $V_m \cdot W_\ell$ is nonzero. Start with a trivial pseudo-isotopy of S^4 whose Cerf graphic is empty. Deform this pseudo-isotopy by creating a single $2,3$-handle eye with no finger or Whitney moves. Deform this family of generalised Morse functions further to one where the spheres A, B undergo a single finger move and a single Whitney move. The finger and Whitney discs are standard and satisfy Quinn's arc condition, as shown in the 3-dimensional slice in Figure 12.

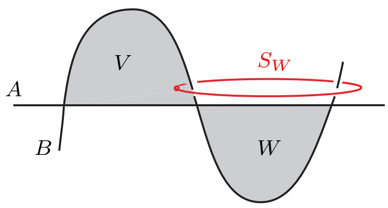


FIGURE 12. The data of finger and Whitney discs in the middle-middle level determining a potentially nontrivial pseudo-isotopy of S^4. (Picture from [GGH$^+$23].)

Recall the construction of Whitney spheres from Section 11.4. Consider the Whitney sphere S_W; it is disjoint from W and intersects V in a single point. Consider a disc $\widetilde{W}$ whose boundary is identical to ∂W and whose interior is a slight displacement of that of W, tubed into S_W. Now we consider a new pseudo-isotopy of S^4 determined by the pair $(V, \widetilde{W})$. It gives rise to a self-diffeomorphism f of S^4. Since V intersects $\widetilde{W}$ in a point, there is no immediate way to smoothly trivialise this pseudo-isotopy.

Conjecture 12.2. *The diffeomorphism $f\colon S^4 \to S^4$ is not smoothly isotopic to the identity.*

Hence it would be extremely interesting to know whether the disc replacement lemma holds. We suspect that a proof would need to use the 1-connected hypothesis, and the embedded arc condition, and possibly also could use, when trying to replace W with $\widetilde{W}$, that the interiors of W and $\widetilde{W}$ are disjoint and $W \cup \widetilde{W} = 0 \in \pi_2(M)$.

Now we start to show how to arrange $\mathring{V}_m \cdot \mathring{W}_\ell = 0$, for every m, ℓ. We will make use of the Alexander trick, so the proof becomes inherently topological at this stage. Given the previous example and conjecture, this is perhaps not surprising. The Whitney spheres S_{V_m} have algebraic intersection numbers with the Whitney discs W_ℓ of the pattern shown in Figure 11. By tubing the S_{V_m} together judiciously, we can construct an algebraically dual collection of embedded spheres E_ℓ for the discs $\{W_\ell\}$, i.e.

$$W_\ell \cdot E_q = \delta_{\ell q} \in \mathbb{Z}$$

for all ℓ, q. Note that the E_ℓ are disjoint from all the discs $V = \bigcup_m V_m$, and they all have trivial normal bundles.

Next, if $\mathring{V}_m \cdot \mathring{W}_\ell = a_{m\ell} \neq 0$, tube the disc V_m into $-a \cdot E_\ell$, where $a \cdot E_\ell$ means a disjoint parallel copies of the sphere. Doing this for all pairs m, ℓ yields a new collection of discs V'_m, with the property that $V'_m \cdot W_\ell = 0$ for all m, ℓ.

Now we want to replace V_m with V'_m for all m. At this point Quinn appealed to the replacement criterion. Instead, we use a method called *factorisation*. Shortly after the finger move time t_f, but before $t = 1/2$, we use the discs $\{V'_m\}$ as Whitney discs, and then shortley thereafter we undo these moves via finger moves. Then we proceed as before, using the original Whitney discs $\{W_\ell\}$ at time t_w. This is a new family that can be obtained from the original by a deformation, because the Whitney-then-finger move sequence using the same set of discs can be deformed to a constant family, with no moves occurring at all. We denote the new family by

$$V \cdot \overline{V'} \cdot V' \cdot \overline{W},$$

where Whitney discs have a bar on them. In the centre of this, between $\overline{V'}$ and V', the spheres A and B are in cancelling position. We can therefore deform to a new family where the spheres cancel, and then shortly thereafter reappear. We obtain two consecutive eyes. The first eye has finger-Whitney data $V \cdot \overline{V'}$. The second eye has finger-Whitney data $V' \cdot \overline{W}$.

Consecutive eyes correspond to a composition of homeomorphisms. Hence we can deal with the two eyes separately. First, we observe that the finger-Whitney data $V \cdot \overline{V'}$ in the left hand eye is entirely local to the extra $S^2 \times S^2$ summand. This corresponds to a homemorphism of X that is supported on a 4-ball $D^4 \subseteq X$. By the Alexander coning trick, such a homeomorphism is isotopic to the identity. It therefore remains to consider the family with data $V' \cdot \overline{W}$. We have effectively replaced the collection of finger discs V by the collection V', which by construction has the advantage that $\mathring{V}'_m \cdot \mathring{W}_\ell = 0$ for all m, ℓ.

As mentioned above, the use of the Alexander trick was one of the places where we made use of the topological category.

13. Completion of the pseudo-isotopy implies isotopy proof

The proof is nearly over at this stage. Given that $\mathring{V}'_m \cdot \mathring{W}_\ell = 0$ for each ℓ, m, we choose immersed Whitney discs in M pairing up the double points. A priori these discs can intersect $A \cup B \cup V' \cup W$. We use the dual spheres $D^A_{B,V'}$, etc, to clean them up so they miss these other surfaces, but potentially gain more self-intersections. Then we apply the *disc embedding theorem* of Freedman to obtain topologically embedded Whitney discs. The argument is from [Qui86, Section 4.6]; Quinn's argument here is rather subtle.

The disc embedding theorem was the main result behind Freedman's Fields medal work on the classification of topological 4-manifolds. It enables us, under favourable conditions, to find locally flat embedded discs to use for Whitney moves. I was part of a team to create a book on this incredible proof [BKK+21]. Interested readers wanting to know more

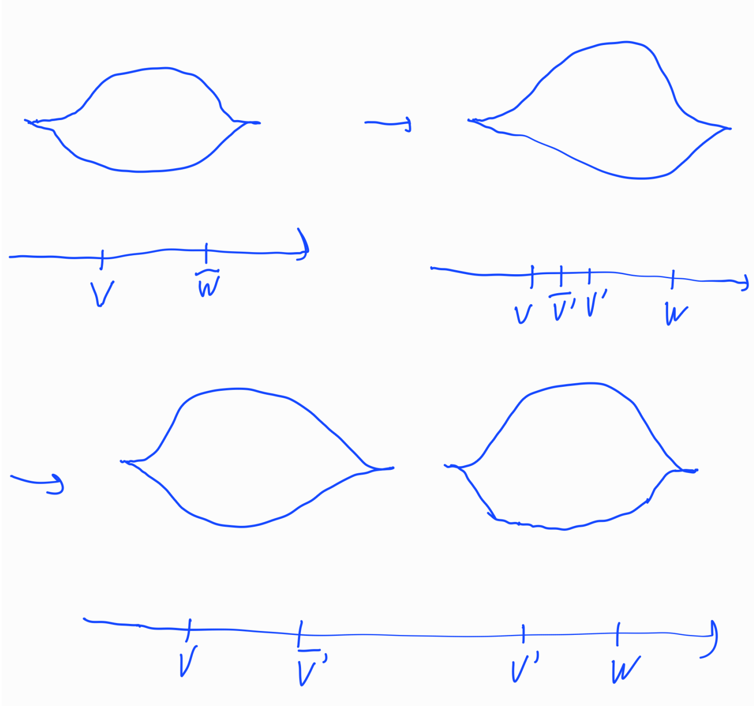


FIGURE 13. Factorising the family into two consecutive eyes, by introducing a cancelling sequence of operations given by a collection of Whitney moves using the discs $\overline{V'}$, that are straight away undone by finger moves using the same set of discs V'. Then the family is split into two eyes. In the result, the first eye has finger-Whitney data $V \cdot \overline{V'}$. The second eye has finger-Whitney data $V' \cdot \overline{W}$.

about the disc embedding theorem are recommended to consult it. See also the original sources [Fre82, FQ90], of course.

Using the Whitney discs arising from the disc embedding theorem, we can perform isotopies of the W discs, which correspond to deformations of the family (g_t, ξ_t), to arrange that the interiors of all V' and W discs are pairwise disjointly embedded. This done, we can deform the family to one where $A \pitchfork B \subseteq M$ is a single point. This indicates that the A and B sphere intersect in a single point for all time, which implies that one can perform cancellation consistently for all time t with $t_b \leq t \leq t_d$, which has the effect of removing the entire eye from the Cerf graphic.

Here are some details on the end of the proof.

Proposition 13.1. *There exists an isotopy of the discs V' in the complement in M of $A \cup B$ that makes the interiors of the discs V' and W disjoint.*

Proof. Start with spheres $\{G_m\}$ made from sums of the Whitney spheres S_{W_ℓ}, that are an algebraically dual collection to the V'_m, and disjoint from $A \cup B \cup W$. These may intersect the V'_m in extra geometric points, because every time there is a W_ℓ, V'_m intersection there is a cancelling pair of S_{W_ℓ}, V'_m intersection points.

We want to make the $\{G_m\}$ into geometric duals for the V'_m. Choose Whitney discs Δ pairing up the extra intersections, in the complement of $A \cup B$. This can be done because A and B have dual spheres $D^A_{B,V}$ and $D^B_{A,V}$, for example, so $A \cup B$ is π_1-negligible. However Δ could still intersect $V' \cup W$. Where Δ intersections V', push Δ down to A, and tube the new intersections with A into $D^A_{B,V'}$. Where Δ intersects W, push W off Δ by performing a finger move of W through V'. This creates extra W, V' intersections, of course. We can now do the immersed Whitney move using the discs Δ. We have now made G_m into geometric duals for the $\{V'_m\}$.

Now choose immersed Whitney discs $\{L_p\}$ pairing up the $\mathring{V}'_m \pitchfork \mathring{W}_\ell$ intersections. Use $D^A_{B,W}$ and $D^B_{A,W}$ to make the $\{\mathring{L}_p\}$ miss $A \cup B \cup W$. Use the $\{G_m\}$ to make the $\{\mathring{L}_p\}$ miss V'. So the interiors of the L_p miss $A \cup B \cup V \cup W$. The $\{L_p\}$ have algebraic duals coming from Clifford tori. By essentially the argument we just used to make the $\{\mathring{L}_p\}$ disjoint from $A \cup B \cup V \cup W$, we have that $M \setminus (A \cup B \cup V \cup W)$ is 1-connected. So by the disc embedding theorem we can replace the $\{L_p\}$ with disjointly embedded, framed discs that we can use as Whitney discs to guide the desired isotopy of V'. □

14. A Cork Theorem for diffeomorphisms

In the smooth category, we cannot entirely simplify a pseudo-isotopy. Nonetheless, it is interesting to analysing its structure. We know from Theorem 8.1 that if $f\colon X \to X$ is pseudo-isotopy to Id then it is also stably smoothly isotopic to Id. We say that f is *1-stably isotopic to* Id if f becomes isotopic to Id after connected summing with one copy of $S^2 \times S^2$, i.e. if $f \# \mathrm{Id}_{S^2 \times S^2}$ is smoothly isotopic to $\mathrm{Id}_{X \# S^2 \times S^2}$.

Theorem 14.1 (Krushkal, Mukherjee, P, Warren)**.** *Let $f\colon X \to X$ be a diffeomorphism of a 1-connected, compact smooth 4-manifold X, and suppose that f is 1-stably isotopic to* Id. *Then there exists a contracible submanifold $U_1 \subseteq X$ such that f is smoothly isotopic to f', where f' is supported on U.*

We call the submanifold U_1 a *diff-cork*. Using this theorem we can construct new nontrivial diffeomorphisms on contractible 4-manifolds.

Sketch of proof. Let me give a brief sketch of the proof. It is inspired by the original cork theorem [CFHS96, Mat96, Kir96]. The hypothesis that f is 1-stably isotopic to the identity implies, by a result of Gabai [Gab22], that there is a pseudo-isotopy of f to the identity with a Cerf graphic having one eye. Arrange further than Quinn's embedded arc condition holds. We consider

$$Q := \nu(A \cup B \cup V \cup W) \subseteq M.$$

We have that $Q \simeq S^2 \vee S^2 \vee \bigvee^k S^1$. The S^1 factors correspond to extra intersections between the interior of V and the interior of W. There is a handle decomposition of M, built on Q, with a collection of 2-handles Y such that each 2-handle is attached to one free generator of $\pi_1(Q)$. Consider Q' given by $Q \cup Y$, thickened in the 5-dimensional direction (recall that $M \subseteq (X \times I)_{1/2}$ is the middle level of $X \times I$ with respect to $(g_{1/2}, \xi_{1/2})$. Now we have a 3-handle going up, in $X \times I$, with respect to $\xi_{1/2}$, corresponding to the 3-handle with attaching sphere B. Turning the cobordism upside down, we have a 3-handle attached to A (the upside down 2-handle). Taking the union of Q' with these two 3-handles, attached to Q' along push-offs to the boundary of A and B, we obtain a contractible sub-cobordism U in $X \times I$. Moreover the critical points cancel algebraically, so U is a contractible h-cobordism.

We can use the fact that U contains all the critical points, and all the trajectories between them, for all possible time values t, to isotope F to a diffeomorphism that is supported in U. The intersection $U_1 := U \cap (X \times \{1\})$ is contractible, because U is contractible and is an h-cobordism. Hence f can be isotoped to a diffeomorphism supported in U_1. $\square$

Example 14.2. I want to briefly give an example of a 1-stably isotopic diffeomorphisms, to which the cork theorem in the previous section applies. The original such examples were due to Ruberman [Rub98]. I will present an example due to Baraglia-Konno [BK20]. Let $X := K_3 \# S^2 \times S^2$. Let $\mathcal{K}$ be a smooth, closed 4-manifold that is homeomorphic to K_3 but not diffeomorphic to it, for example arising from knot surgery. Suppose also that $X' := \mathcal{K} \# S^2 \times S^2 \cong X$. Let $\phi\colon X \to X'$ be such a diffeomorphism. I claim that a choice of $\mathcal{K}$ exists such that a diffeomorphism with this property exists, and its action respects the decomposition of H_2 and acts as the identity on $H_2(S^2 \times S^2)$. Let $r\colon S^2 \times S^2 \to S^2 \times S^2$ be the composition of symplectic Dehn twists in the spheres representing $(1,1)$ and $(1,-1)$ in $H_2(S^2 \times S^2) \cong \mathbb{Z} \oplus \mathbb{Z}$. We consider the diffeomorphism

$$(\mathrm{Id}_{K_3} \# r) \circ \phi^{-1} \circ (\mathrm{Id}_{\mathcal{K}} \# r) \circ \phi\colon X \to X.$$

This acts trivially on $H_2(X)$, and so is topologically isotopic to the identity. It turns out to be 1-stably isotopic to Id_X, but we will not have time to explain it here.

15. Mapping class groups of 1-connected 4-manifolds with boundary

The pseudo-isotopy theorem of Perron and Quinn works just as well for 1-connected compact 4-manifolds with boundary. However the pseudo-isotopy classification is more subtle. In joint work with Orson, I figured out the details of this. Our work builds on important work of Saeki in this direction, who studied the analogous stable smooth mapping class group for smooth 1-connected 4-manifolds with connected nonempty boundary. The following description follows that in [OP22].

When $\partial X = \emptyset$, we have seen that if two o.p. homeomorphisms of X induce the same isometry of the intersection form then they are isotopic. When X has nonempty boundary, we need to consider a refinement of $\mathrm{Aut}(H_2(X), \lambda_X)$ to capture the algebraic data of a homeomorphism. A map $f \in \mathrm{Homeo}^+(X, \partial X)$ determines a homomorphism $\Delta_f\colon H_2(X, \partial X) \to H_2(X)$ called a *variation*, defined by $[x] \mapsto [x - f(x)]$. Using that X has Poincaré-Lefschetz duality, Saeki [Sae06] showed that Δ_f satisfies an additional condition, making it what we call a *Poincaré variation*. There is a binary operation on the set of Poincaré variations, together with which they form a group $\mathcal{V}(H_2(X), \lambda_X)$. The map $f \mapsto f_*$ factors through this group via homomorphisms:

$$\pi_0 \,\mathrm{Homeo}^+(X, \partial X) \xrightarrow{f \mapsto \Delta_f} \mathcal{V}(H_2(X), \lambda_X) \xrightarrow{\Delta \mapsto \mathrm{Id} - \Delta \circ j} \mathrm{Aut}(H_2(X), \lambda_X),$$

where $j\colon H_2(X) \to H_2(X, \partial X)$ is the quotient map. In general Δ_f contains more information than f_*, although if ∂X is a $\mathbb{Q}HS^3$ or a $\mathbb{Q}H(S^1 \times S^2)$ then the second map is an isomorphism. Saeki [Sae06] used $\mathcal{V}(H_2(X), \lambda_X)$ to describe the smooth stable mapping class group for simply connected 4-manifolds with nonempty, connected boundary.

Example 15.1. Let X be a 1-connected 4-manifold with boundary $T^3 = S^1_1 \times S^1_2 \times S^1_3$, the 3-torus. For example, such a 4-manifold arises by adding 0-framed 2-handles to D^4 along the Borromean rings. Rotating the S^1_1 direction yields a loop of diffeomorphism in $\pi_1 \operatorname{Diff}^+(T^3)$. We can apply this in a collar neighbourhood of ∂X to obtain a *generalised Dehn twist* $f\colon X \xrightarrow{\cong} X$. Since f is supported in $\partial X \times I$, it acts trivially on $H_2(X)$. However, the curve S^1_2 bounds a nontrivial relative homology class $x_2 \in H_2(X, \partial X)$. The difference $x - f(x)$ is the image under the injection $H_2(T^3) \rightarrowtail H_2(X)$ of the class $[S^1_1 \times S^1_2] \in H_2(T^3)$. Hence the variation Δ_f is nontrivial, and thus f is not isotopic rel. boundary to the identity. In contract, note that if the boundary is permitted to move in an isotopy, then f is isotopic to the identity.

When ∂X has more than one connected component and X admits a spin structure, there is a further invariant that does not appear in the closed case nor when the boundary is connected. For $f \in \operatorname{Homeo}^+(X, \partial X)$ we may compare a topological spin structure $\mathfrak{s}$ on X with the induced spin structure $f^*\mathfrak{s}$. The two agree on ∂X because f fixes the boundary pointwise. There is a free, transitive action of $H^1(X, \partial X; \mathbb{Z}/2)$ on the set of isomorphism classes of spin structures on X that agree on ∂X, and we denote by $\Theta(f) \in H^1(X, \partial X; \mathbb{Z}/2)$ the class representing the difference between $\mathfrak{s}$ and $f^*\mathfrak{s}$.

Example 15.2. Let $X := S^3 \times I$, and let $f\colon X \to X$ be the Dehn twist that we introduced earlier in the context of the connected sum sphere in $K3 \# K3$. This diffeomorphism necessarily acts trivially on $H_2(X) = 0$, has trivial Poincaré variation for the same reason. However, f is not (pseudo-) isotopic rel. boundary to Id_X, because it acts nontrivially on the relative spin structures of X (of which there are two).

In joint work with Orson, we showed that these invariants describe the entire topological mapping class group.

Theorem 15.3. *Let $(X, \partial X)$ be a compact, simply connected, oriented, topological 4-manifold.*

(i) When X is spin, the map $f \mapsto (\Theta(f), \Delta_f)$ induces a group isomorphism

$$\pi_0 \operatorname{Homeo}^+(X, \partial X) \xrightarrow{\cong} H^1(X, \partial X; \mathbb{Z}/2) \times \mathcal{V}(H_2(X), \lambda_X).$$

(ii) When X is not spin, the map $f \mapsto \Delta_f$ induces a group isomorphism

$$\pi_0 \operatorname{Homeo}^+(X, \partial X) \xrightarrow{\cong} \mathcal{V}(H_2(X), \lambda_X).$$

If the boundary is nonempty, then the $+$ is superfluous: every homeomorphism that acts as the identity on the boundary is o.p. In order to state the result for the case of empty and nonempty boundary simultaneously, we leave the $+$ in the notation.

In Example 15.1 we gave an example of a nontrivial diffeomorphism that acts trivially on $H_2(X)$, i.e. that lies in the Torelli subgroup of the mapping class group of X. The question then arose whether all the elements of the Torelli group can be smoothly realised.

Theorem 15.4 (Galvin-Ladu [GL23])**.** *There exists a 1-connected, smooth, compact 4-manifold X (which they construct explicitly) together with a homeomorphism $f\colon X \xrightarrow{\cong} X$ in the Torelli subgroup of $\pi_0(\operatorname{Homeo}^+(X, \partial X))$ that is not topologically isotopic to any diffeomorphism of X.*

This is in contrast to the closed case, when every element of the Torelli subgroup is isotopic to the identity, hence is certainly isotopic to a diffeomorphism.

16. Simple spines for knot traces

Here is another application. Let X be a compact 4-manifold with $X \simeq S^2$. Let F_0, F_1 be two locally flat embedded, oriented 2-spheres in X that represent the same generator of $H_2(X) \cong \mathbb{Z}$. These 2-spheres are called *spines* of X.

Definition 16.1. A spine F_i of X is called *simple* if $\pi_1(X \smallsetminus F_i)$ is abelian.

Theorem 16.2 (Orson-P. [OP24])**.** *If F_0 and F_1 are simple spines in X, then F_0 and F_1 are ambiently isotopic.*

To prove this theorem, first we used modified surgery to show that there is a homeomorphism $G \colon X \to X$ of X to itself, restricting to $\mathrm{Id}_{\partial X}$, sending F_0 to F_1. Then we applied our work on mapping class groups of manifolds with boundary, given above, to obtain an isotopy of G to the identity, which has the result of isotoping F_1 to F_0. In this case we use the following corollary of the result for mapping class groups of 1-connected 4-manifolds with boundary from the previous section.

Corollary 16.3. *Let X be a compact, simply connected 4-manifold such that ∂X has the rational homology of either S^3 or $S^1 \times S^2$. Let $G \colon X \to X$ be a homeomorphism that restricts to the identity on ∂X and is such that $G_* = \mathrm{Id} \colon H_2(X) \to H_2(X)$. Then G is topologically isotopic rel. boundary to the identity map of X.*

Simple spines have boundaries with the required homology, and the homeomorphism we constructed, by virtue of sending one spine to another, has to induce the identity on $H_2(X)$. So the corollary applies to give us the desired isotopy.

References

[Bau77] Hans J. Baues, *Obstruction theory on homotopy classification of maps*, Lecture Notes in Mathematics, vol. Vol. 628, Springer-Verlag, Berlin-New York, 1977.

[BG19] Ryan Budney and David Gabai, *Knotted 3-balls in S^4*, 2019. Preprint, available at arXiv:1912.09029.

[BK20] David Baraglia and Hokuto Konno, *A gluing formula for families Seiberg-Witten invariants*, Geom. Topol. **24** (2020), no. 3, 1381–1456.

[BKK$^+$21] Stefan Behrens, Boldizsár Kalmár, Min Hoon Kim, Mark Powell, and Arunima Ray (eds.), *The disc embedding theorem*, Oxford University Press, 2021.

[Cer70] 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** (1970), 5–173.

[CFHS96] C. L. Curtis, M. H. Freedman, W. C. Hsiang, and R. Stong, *A decomposition theorem for h-cobordant smooth simply-connected compact 4-manifolds*, Invent. Math. **123** (1996), no. 2, 343–348.

[CH90] Tim D. Cochran and Nathan Habegger, *On the homotopy theory of simply connected four manifolds*, Topology **29** (1990), no. 4, 419–440.

[COP20] Jae Choon Cha, Kent E. Orr, and Mark Powell, *Whitney towers and abelian invariants of knots*, Math. Z. **294** (2020), no. 1-2, 519–553.

[FQ90] Michael Freedman and Frank Quinn, *Topology of 4-manifolds*, Princeton Mathematical Series, vol. 39, Princeton University Press, 1990.

[Fre82] Michael Freedman, *The topology of four-dimensional manifolds*, J. Differential Geom. **17** (1982), no. 3, 357–453.

[Gab22] David Gabai, *3-spheres in the 4-sphere and pseudo-isotopies of* $S^1 \times S^3$, 2022. Preprint, available at arXiv:2212.02004.

[GGH+23] David Gabai, David T. Gay, Daniel Hartman, Vyacheslav Krushkal, and Mark Powell, *Pseudo-isotopies of simply connected 4-manifolds*, 2023.

[GL23] Daniel Galvin and Roberto Ladu, *Non-smoothable homeomorphisms of 4-manifolds with boundary*, 2023.

[HW73] Allen Hatcher and John Wagoner, *Pseudo-isotopies of compact manifolds*, Astérisque, No. 6, Société Mathématique de France, Paris, 1973.

[Igu21] Kyoshi Igusa, *Second obstruction to pseudoisotopy I*, 2021. Preprint, available at arXiv:2110.09659.

[Igu84] Kiyoshi Igusa, *What happens to Hatcher and Wagoner's formulas for* $\pi_0 C(M)$ *when the first Postnikov invariant of* M *is nontrivial?*, Algebraic K-theory, number theory, geometry and analysis (Bielefeld, 1982), 1984, pp. 104–172.

[Kir96] Rob Kirby, *Akbulut's corks and h-cobordisms of smooth, simply connected 4-manifolds*, Turk. J. Math. **20** (1996), no. 1, 85–93.

[Kis64] James M. Kister, *Microbundles are fibre bundles*, Ann. of Math. (2) **80** (1964), 190–199.

[KK24] Manuel Krannich and Alexander Kupers, *A note on homotopy and pseudoisotopy of diffeomorphisms of 4-manifolds*, 2024.

[KM20] Peter B. Kronheimer and Tomasz S. Mrowka, *The Dehn twist on a sum of two K3 surfaces*, Math. Res. Lett. **27** (2020), no. 6, 1767–1783.

[Kre79] Matthias Kreck, *Isotopy classes of diffeomorphisms of* $(k-1)$*-connected almost-parallelizable* $2k$*-manifolds*, Algebraic topology, Aarhus 1978 (Proc. Sympos., Univ. Aarhus), 1979, pp. 643–663.

[Kre99] Matthias Kreck, *Surgery and duality*, Ann. of Math. (2) **149** (1999), no. 3, 707–754.

[KS77] Robion C. Kirby and Laurence C. Siebenmann, *Foundational essays on topological manifolds, smoothings, and triangulations*, Princeton University Press, Princeton, N.J., 1977. With notes by John Milnor and Michael Atiyah, Annals of Mathematics Studies, No. 88.

[Mat96] R. Matveyev, *A decomposition of smooth simply-connected h-cobordant 4-manifolds*, J. Differ. Geom. **44** (1996), no. 3, 571–582.

[Mil61] John Milnor, *Microbundles and differentiable structures*, Princeton (1961). Mimeographed.

[Mil64] John Milnor, *Microbundles*: *Part I*, Topology **3** (1964), no. 1, 53–80.

[Mil65] John W. Milnor, *Lectures on the h-cobordism theorem*, Princeton University Press, Princeton, N.J., 1965. Notes by L. Siebenmann and J. Sondow.

[OP22] Patrick Orson and Mark Powell, *Mapping class groups of simply connected 4-manifolds with boundary*, 2022. Preprint, available at arXiv:2207.05986.

[OP24] Patrick Orson and Mark Powell, *Simple spines of homotopy 2-spheres are unique*, Proceedings of the London Mathematical Society **128** (2024), no. 2, e12583, available at `https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/plms.12583`.

[Per86] Bernard Perron, *Pseudo-isotopies et isotopies en dimension quatre dans la catégorie topologique*, Topology **25** (1986), no. 4, 381–397.

[Qui82] Frank Quinn, *Ends of maps III: Dimensions 4 and 5*, J. Differential Geom. **17** (1982), no. 3, 503–521.

[Qui86] Frank Quinn, *Isotopy of 4-manifolds*, J. Differential Geom. **24** (1986), no. 3, 343–372.

[Rob72] C. A. Robinson, *Moore-postnikov systems for non-simple fibrations*, Illinois Journal of Mathematics **16** (1972), no. 2, 234 –242.

[Rub98] Daniel Ruberman, *An obstruction to smooth isotopy in dimension 4*, Math. Res. Lett. **5** (1998), no. 6, 743–758.

[Sae06] Osamu Saeki, *Stable mapping class groups of 4-manifolds with boundary*, Trans. Amer. Math. Soc. **358** (2006), no. 5, 2091–2104.

[Sin21] Oliver Singh, *Pseudo-isotopies and diffeomorphisms of 4-manifolds* (2021). Preprint available at arXiv:2111.15658.

[Sma62] Stephen Smale, *On the structure of manifolds*, Amer. J. Math. (1962), 387–399.

[ST19] Rob Schneiderman and Peter Teichner, *The group of disjoint 2-spheres in 4-space*, Ann. of Math. (2) **190** (2019), no. 3, 669–750.

[SW00] Richard Stong and Zhenghan Wang, *Self-homeomorphisms of 4-manifolds with fundamental group* $\mathbb{Z}$, Topology Appl. **106** (2000), no. 1, 49–56.

[Tei92] Peter Teichner, *Topological 4-manifolds with finite fundamental group*, Ph.D. Thesis, 1992. Shaker Verlag, ISBN 3-86111-182-9.

[Tei97] Peter Teichner, *On the star-construction for topological* 4*-manifolds*, Geometric topology (Athens, GA, 1993), 1997, pp. 300–312.

[Wal63] C. T. C. Wall, *On the orthogonal groups of unimodular quadratic forms. II*, J. Reine Angew. Math. **213** (1963), 122–136.

School of Mathematics and Statistics, University of Glasgow, United Kingdom
Email address: `mark.powell@glasgow.ac.uk`