Embedded surfaces with infinite cyclic knot group

ANTHONY CONWAY

MARK POWELL

# Embedded surfaces with infinite cyclic knot group

Anthony Conway

Mark Powell

We study locally flat, compact, oriented surfaces in 4–manifolds whose exteriors have infinite cyclic fundamental group. We give algebraic topological criteria for two such surfaces, with the same genus $g$, to be related by an ambient homeomorphism, and further criteria that imply they are ambiently isotopic. Along the way, we provide a classification of a subset of the topological 4–manifolds with infinite cyclic fundamental group, and we apply our results to rim surgery.

## 1 Introduction

We study locally flat embeddings of compact, orientable surfaces in compact, oriented, simply connected topological 4–manifolds, where the complement of the surface has infinite cyclic fundamental group. Extending the terminology for knotted spheres, we call this group the *knot group*, so we shall study knotted surfaces with knot group $\mathbb{Z}$, or $\mathbb{Z}$*–surfaces*.

We will present algebraic criteria for pairs of $\mathbb{Z}$–surfaces to be ambiently isotopic. As part of the proof we obtain an algebraic classification of a certain subset of the 4–manifolds with boundary and fundamental group $\mathbb{Z}$, simultaneously generalising work of Freedman and Quinn [15] on the closed case, and of Boyer [5] on simply connected 4–manifolds with nonempty boundary; see Section 1.7. We apply our results to show that in simply connected 4–manifolds, 1–twisted rim surgery on a surface with knot group $\mathbb{Z}$ yields a topologically ambiently isotopic surface, extending results of Kim and Ruberman [31] and Juhász, Miller and Zemke [24]; see Section 1.5.

## 1.1 Surfaces in $S^4$ and $D^4$

We start with our main results on surfaces in $S^4$ and $D^4$ as the most important special cases, before going on to explain more general statements for surfaces in any simply connected 4–manifold.

**Theorem 1.1** *Any two locally flat, embedded, closed, orientable and genus $g \neq 1, 2$ surfaces $\Sigma_0, \Sigma_1 \subseteq S^4$ with $\pi_1(S^4 \setminus \Sigma_i) \cong \mathbb{Z}$ for $i = 0, 1$ are topologically ambiently isotopic.*

**Theorem 1.2** *Let $K$ be an Alexander polynomial one knot in $S^3$. Any two locally flat, embedded, compact, orientable genus $g \neq 1, 2$ surfaces $\Sigma_0, \Sigma_1 \subseteq D^4$ with boundary $K$ and $\pi_1(D^4 \setminus \Sigma_i) \cong \mathbb{Z}$ for $i = 0, 1$ are topologically ambiently isotopic rel boundary.*

In a previous paper, we proved Theorem 1.2 in the genus-zero case [10, Theorem 1.2]. With regards to Theorem 1.1, a genus $g$ surface $\Sigma \subseteq S^4$ is *unknotted* if it bounds a locally flat, embedded handlebody in $S^4$. The *unknotting conjecture* for locally flat surfaces posits that a locally flat, embedded, closed, oriented surface $\Sigma \subseteq S^4$ is unknotted if and only if $\pi_1(S^4 \setminus \nu\Sigma) = \mathbb{Z}$. The forwards direction holds because any two embeddings of a handlebody in $S^4$ are ambiently isotopic, so an unknotted surface is ambiently isotopic to a standard embedding. For the reverse direction, Freedman and Quinn [15, Theorem 11.7A] proved the $g = 0$ case of 2–knots, when $\Sigma \cong S^2$, while Hillman and Kawauchi [22] claimed it for $g \geq 1$. Theorem 1.1 offers a new proof for surfaces of genus $g \geq 3$. We also give new proofs for the $g = 0$ cases, but it should be noted that this specialisation produces somewhat similar proofs to those of Freedman and Quinn and our previous work [10].

Our proof of Theorem 1.1 differs from that of Hillman and Kawauchi in [22], who do not have any genus restrictions. In particular, in a key step in the proof, one considers a closed 4–manifold $M$ built from the two surface exteriors, and shows that $M$ is homeomorphic to $S^1 \times S^3 \# \#_{i=1}^{2g} S^2 \times S^2$. To see this, we control the $\mathbb{Z}[\mathbb{Z}]$–valued intersection form of $M$, whereas [22, Proof of Lemma 3.1] just calculates the $\mathbb{Z}$–valued intersection form, and appeals to separate work of Kawauchi [27], in which it was claimed that every closed, spin 4–manifold with fundamental group $\mathbb{Z}$ splits as $S^1 \times S^3 \# X$, where $X$ is a closed, simply connected 4–manifold. It would follow from this claim that computing the $\mathbb{Z}$–valued form suffices. A mistake in [27] was found by Hambleton and Teichner [21]. Although Kawauchi later updated his

theorem [28] to include the hypothesis that the $\mathbb{Z}$–intersection form be indefinite, which is the case for the 4–manifolds arising in [22], the consensus in the community seems to be that an independent account is also desirable.

## 1.2 Surfaces in simply connected 4–manifolds with boundary $S^3$

To fix our terminology, throughout the article, by a 4–*manifold* we shall always mean a compact, connected, oriented, topological 4–manifold with connected boundary. Surfaces will always be compact, connected, and orientable. We will consider locally flat, embedded, closed, oriented surfaces $\Sigma \subseteq X$ with $X$ a closed 4–manifold, and locally flat properly embedded surfaces $\Sigma \subseteq N$, where $N$ is a 4–manifold with boundary $S^3$ and $\partial\Sigma = K \subseteq S^3$ is a fixed knot. The exteriors of such surfaces will be denoted by $X_\Sigma$ and $N_\Sigma$ respectively.

Next we formulate our most general statement, for the nonempty boundary case, pairs of surfaces $\Sigma_i \subseteq N$, where $\partial\Sigma_i \subseteq \partial N = S^3$ is a fixed knot. Theorem 1.2 is an immediate consequence of Theorem 1.3 below, which we state after introducing the prerequisites. One quickly deduces consequences for closed surfaces in closed 4–manifolds $X$, in particular Theorem 1.1, by removing an unknotted $(D^4, D^2)$ pair from $(X, \Sigma)$, as we will explain in Section 1.3.

Given a compact, oriented 4–manifold $M$ with $\pi_1(M) \cong \mathbb{Z}$, we write $\pi_1(M) = \mathbb{Z}$ to mean that we have chosen an isomorphism between $\pi_1(M)$ and $\mathbb{Z}$. In the case of the exterior $X_\Sigma$ or $N_\Sigma$ of an oriented surface $\Sigma$, such an identification is determined by the orientations on $\Sigma$ and on $X$ or $N$ respectively. Set $\Lambda := \mathbb{Z}[t^{\pm 1}]$ and write $H_*(M; \Lambda)$ for the homology of the infinite cyclic cover $\widetilde{M}$ of $M$, considered as a $\Lambda$–module. Taking signed intersections in $\widetilde{M}$ endows the homology $\Lambda$–module $H_2(M; \Lambda)$ with a sesquilinear Hermitian intersection form

$$\lambda_M \colon H_2(M; \Lambda) \times H_2(M; \Lambda) \to \Lambda.$$

The adjoint of $\lambda_M$ is the $\Lambda$–module homomorphism

$$\widehat{\lambda}_M \colon H_2(M; \Lambda) \to \mathrm{Hom}_\Lambda(H_2(M; \Lambda), \Lambda) =: H_2(M; \Lambda)^*.$$

If two 4–manifolds $M_0, M_1$ with infinite cyclic fundamental group are orientation-preserving homeomorphic, then their $\Lambda$–intersection forms are isometric. That is, there exists an isomorphism $F \colon H_2(M_0; \Lambda) \xrightarrow{\cong} H_2(M_1; \Lambda)$ such that $\widehat{\lambda}_{M_0} = F^* \widehat{\lambda}_{M_1} F$. We write $F \colon \lambda_{M_0} \cong \lambda_{M_1}$ and call $F$ an *isometry*.

Let $Y$ be a compact oriented 3–manifold, and let $\varphi \colon \pi_1(Y) \twoheadrightarrow \mathbb{Z}$ be an epimorphism. Associated with this data, there is a $\Lambda$–module $H_1(Y; \Lambda)$, called the *Alexander module*, which is the first homology group of the infinite cyclic cover associated to $\ker(\varphi)$. If this module is torsion over $\Lambda = \mathbb{Z}[t^{\pm 1}]$, then it is endowed with a sesquilinear Hermitian *Blanchfield form*

$$\mathrm{Bl}_Y \colon H_1(Y; \Lambda) \times H_1(Y; \Lambda) \to \mathbb{Q}(t)/\Lambda.$$

More details on this pairing appear in Section 3.1, but we note that $\mathrm{Bl}_Y$ should be thought of as the analogue of the linking pairing of a $\mathbb{Q}$–homology sphere on the level of infinite cyclic covers.

As we recall in Sections 2 and 3, if $M_0, M_1$ are 4–manifolds with $\pi_1(M_i) = \mathbb{Z}$ whose boundaries $\partial M_i$ have torsion Alexander modules, and if $\pi_1(\partial M_i) \to \pi_1(M_i)$ is surjective for $i = 0, 1$, then an isometry $F \colon \lambda_{M_0} \cong \lambda_{M_1}$ of the $\Lambda$–intersection forms induces an isometry of the Blanchfield forms of the boundary,

$$\partial F \colon (H_1(\partial M_0; \Lambda), \mathrm{Bl}_{\partial M_0}) \xrightarrow{\cong} (H_1(\partial M_1; \Lambda), \mathrm{Bl}_{\partial M_1}).$$

Here is the construction: via universal coefficients, Poincaré–Lefschetz duality, and the long exact sequence of a pair, the composition

$$H_2(M_i; \Lambda)^* \cong H^2(M_i; \Lambda) \cong H_2(M_i, \partial M_i; \Lambda) \xrightarrow{\delta} H_1(\partial M_i; \Lambda)$$

induces an identification $\mathrm{coker}(\widehat{\lambda}_{M_i}) \cong H_1(\partial M_i; \Lambda)$, and the map

$$F^{-*} := (F^*)^{-1} \colon H_2(M_0; \Lambda)^* \to H_2(M_1; \Lambda)^*$$

induces $\partial F \colon \mathrm{coker}(\widehat{\lambda}_{M_0}) \to \mathrm{coker}(\widehat{\lambda}_{M_1})$.

Now we focus on the case where $M_i = N_{\Sigma_i}$ are surface exteriors, with $N$ a compact, simply connected 4–manifold with boundary $S^3$. The boundary $\partial N_{\Sigma_i}$ is homeomorphic to

$$M_{K,g} := E_K \cup_\partial (\Sigma_{g,1} \times S^1),$$

where $E_K = S^3 \setminus \nu(K)$ is the knot exterior and $\Sigma_{g,1}$ denotes the (abstract) genus $g$ surface with one boundary component. In Proposition 5.7, we show that every automorphism $h$ of the Blanchfield pairing $\mathrm{Bl}_{M_{K,g}}$ decomposes as $h_K \oplus h_\Sigma$, where $h_K$ is an automorphism of $\mathrm{Bl}_K := \mathrm{Bl}_{E_K}$ and $h_\Sigma$ is an automorphism of $\mathrm{Bl}_{\Sigma_{g,1} \times S^1}$.

Let $f_K \colon E_K \to E_K$ be an orientation-preserving homeomorphism that is the identity on $\partial E_K$. Extend $f_K$ via the identity on $\nu K$ to an orientation-preserving self-homeomorphism of $S^3$. The mapping class group of $S^3$ is trivial, so there is an

isotopy $\Theta(f_K)\colon S^3 \times [0, 1] \to S^3$ between the extension and the identity, such that $\Theta(f_K)|_{S^3 \times \{0\}} = \mathrm{Id}$ and $\Theta(f_K)|_{E_K \times \{1\}} = f_K$.

Here is the central theorem of the paper on embedded surfaces with knot group $\mathbb{Z}$.

**Theorem 1.3** *Let $X$ be a closed, simply connected and oriented 4–manifold, let $N = X \setminus \mathring{D}^4$ be a punctured $X$, and let $K \subseteq S^3 = \partial N$ be a knot. Let $\Sigma_0, \Sigma_1 \subseteq N$ be two locally flat, embedded, compact, oriented genus $g$ surfaces with the same oriented boundary $K$ and $\pi_1(N_{\Sigma_i}) = \mathbb{Z}$ for $i = 0, 1$. Suppose there is an isometry $F\colon \lambda_{N_{\Sigma_0}} \cong \lambda_{N_{\Sigma_1}}$ and write $\partial F = h_K \oplus h_\Sigma$.*

- *If $h_K$ is induced by an orientation-preserving homeomorphism $f_K\colon E_K \to E_K$ that is the identity on $\partial E_K$, then $f_K$ extends to an orientation-preserving homeomorphism of pairs*

$$(N, \Sigma_0) \xrightarrow{\;\cong\;} (N, \Sigma_1)$$

  *inducing the given isometry $F\colon H_2(N_{\Sigma_0}; \Lambda) \cong H_2(N_{\Sigma_1}; \Lambda)$.*

- *If in addition $N = D^4$, then for any choice of isotopy $\Theta(f_K)$, the surfaces $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic via an ambient isotopy of $D^4$ extending $\Theta(f_K)$.*

In particular note that if $h_K = \mathrm{Id}$, then we can take $f_K = \mathrm{Id}$ and $\Theta_{f_K}$ the constant isotopy, so that the homeomorphism of pairs in the first item can be assumed to fix the boundary pointwise, and the ambient isotopy in the second item can be assumed to be rel boundary.

In general, $\lambda_{N_{\Sigma_0}}$ and $\lambda_{N_{\Sigma_1}}$ need not be isometric, even for $N = D^4$, as shown by examples due to Oba [43].

The deduction of the last item uses Alexander's coning trick, which shows that every homeomorphism of $D^4$ that restricts to the identity on $\partial D^4$ is topologically isotopic to the identity. So in $D^4$, a homeomorphism of pairs can be upgraded to a topological ambient isotopy. See Section 5.3 for details.

As explained in Section 1.5 below, we will apply Theorem 1.3 to 1–twisted rim surgery, a method which has been effective at producing exotic embeddings of surfaces. In Section 1.6, we also apply Theorem 1.3 to study Seifert surfaces that are pushed in to $D^4$. Further applications can be obtained by finding classes of knots $K$ for which every automorphism of the Blanchfield pairing is realised by a symmetry of the knot exterior $E_K$.

## 1.3 Ambient isotopy in closed 4–manifolds

For closed surfaces embedded in closed, simply connected 4–manifolds, by applying the classification of self-homeomorphisms due to Kreck [36], Perron [44] and Quinn [45], we can potentially upgrade a homeomorphism of pairs obtained from Theorem 1.3 to an ambient isotopy. The theorem (see [45] or [15, Theorem 10.1]) is that two self-homeomorphisms of a closed, simply connected 4–manifold are isotopic if and only if they induce the same self-isomorphism on second homology. An analogous classification of homeomorphisms for simply connected 4–manifolds with boundary has not yet been proven.

**Theorem 1.4** *Let $X$ be a closed, simply connected, oriented 4–manifold. Let $\Sigma_0, \Sigma_1 \subseteq X$ be two locally flat, embedded, closed, oriented genus $g$ surfaces with $\pi_1(X_{\Sigma_i}) = \mathbb{Z}$ for $i = 0, 1$.*

(i) *If the intersection forms $\lambda_{X_{\Sigma_0}}$ and $\lambda_{X_{\Sigma_1}}$ are isometric via an isometry $F$, then there is an orientation-preserving homeomorphism of pairs*

$$\Phi\colon (X, \Sigma_0) \xrightarrow{\cong} (X, \Sigma_1)$$

*inducing the given isometry $\Phi_* = F\colon H_2(X_{\Sigma_0}; \Lambda) \cong H_2(X_{\Sigma_1}; \Lambda)$.*

(ii) *The isometry $F$ induces an isometry $F_{\mathbb{Z}}\colon H_2(X) \to H_2(X)$ of the standard intersection form $Q_X$ of $X$. The surfaces $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic if and only if $F_{\mathbb{Z}} = \mathrm{Id}$.*

With regards to (ii), Lemma 5.10 furnishes details of how the isometry $F\colon \lambda_{\Sigma_0} \cong \lambda_{\Sigma_1}$ induces an isometry $F_{\mathbb{Z}}$ of the standard intersection form $Q_X$. Theorem 1.4 also has applications to rim surgery, as we explain in Section 1.5.

Theorem 1.4(i) follows from Theorem 1.3. Here is a short outline; see Section 5.4 for details. After an isotopy, we assume $\Sigma_0$ and $\Sigma_1$ coincide on a disc $D^2$. Remove an open neighbourhood of this common disc $(\mathring{D}^4, \mathring{D}^2)$ from $(X, \Sigma_i)$ to obtain a pair $(N, \widetilde{\Sigma}_i)$, with $\partial N \cong S^3$ and $\partial \widetilde{\Sigma}_i$ an unknot $K$. The exterior of $\Sigma_i$ in $X$ equals the exterior of $\widetilde{\Sigma}_i$ in $N$, so the assumptions of Theorem 1.3 hold. Here, since the unknot has trivial Alexander module, $h_K = \mathrm{Id}$ and so we take $f_K = \mathrm{Id}$. Then Theorem 1.3 produces a homeomorphism of pairs $(N, \widetilde{\Sigma}_0) \cong (N, \widetilde{\Sigma}_1)$ rel boundary, which we complete with the identity on the missing 4–ball to prove (i). Deducing (ii) uses the classification of homeomorphisms from [45] and [15, Theorem 10.1] mentioned above.

**Remark 1.5**  It is worth contrasting Theorem 1.4 with a result of Sunukjian. Indeed, [49, Theorem 7.2] shows that closed surfaces $\Sigma_0, \Sigma_1 \subseteq X$ of the same genus with knot group $\mathbb{Z}$ are topologically isotopic provided $b_2(X) \geq |\sigma(X)| + 2$. Thus, when $X$ is "big enough", restrictions on the equivariant intersection form are not needed to establish isotopy.

## 1.4  Deducing Theorems 1.1 and 1.2 from Theorems 1.3 and 1.4

In the closed case, with $X = S^4$, if we have an isometry $F$ of the intersection form of $X_{\Sigma_0} = S^4_{\Sigma_0}$ and $X_{\Sigma_1} = S^4_{\Sigma_1}$ then the map $F_{\mathbb{Z}} \colon H_2(S^4) \to H_2(S^4)$ is automatically the identity automorphism.

For the case with nonempty boundary, with $N = D^4$, Alexander polynomial one implies that $H_1(E_K; \Lambda) = 0$, so $h_K = \text{Id}$ automatically, and we may take $f_K \colon E_K \to E_K$ also to be the identity.

Therefore, in both cases, it suffices to find an isometry between the intersection forms of the surface exteriors. It is an open question whether this is true in general. For genus zero, ie for discs and spheres, this is automatically the case since the $\Lambda$–coefficient second homology of the surface exterior vanishes. Therefore the genus-zero cases of Theorems 1.1 and 1.2, due to [15, Theorem 11.7A] and [10, Theorem 1.2], respectively, follow from Theorems 1.3 and 1.4.

For genus at least three, our strategy to show that the intersection forms are isometric is as follows. As explained above, the exterior of a closed surface in $S^4$ can be considered as the exterior of a properly embedded surface in $D^4$, so we discuss only the latter case.

By [2, Theorem 5] (see also Theorem A.1), $\Sigma_0$ and $\Sigma_1$ are stably equivalent, meaning that they become isotopic after adding some number of trivial tubes. Therefore the intersection forms of $\lambda_{D^4_{\Sigma_0}}$ and $\lambda_{D^4_{\Sigma_1}}$ satisfy $\lambda_{D^4_{\Sigma_0}} \oplus \mathcal{H}_2^{\oplus n} \cong \lambda_{D^4_{\Sigma_1}} \oplus \mathcal{H}_2^{\oplus n}$ for some $n \geq 0$, where

$$\mathcal{H}_2 := \begin{pmatrix} 0 & t-1 \\ t^{-1}-1 & 0 \end{pmatrix}.$$

Then provided $g$ is at least three we are able to leverage algebraic cancellation results for hyperbolic forms from Bass [1] (see also Crowley and Sixt [12], Hambleton and Teichner [21], Khan [29] and Magurn, van der Kallen and Vaserstein [41]) to improve such a stable isometry to an isometry. While we refer to Section 7.1 for details, we record one of the aforementioned intermediate results as it might be of independent interest.

**Proposition 1.6** *Let $X$ be a closed, simply connected, oriented 4–manifold. Let $\Sigma_0, \Sigma_1 \subseteq N = X \setminus \overset{\circ}{D}{}^4$ be two locally flat, properly embedded, compact, oriented genus $g$ surfaces with boundary the same oriented knot $K \subseteq S^3$, and $\pi_1(N_{\Sigma_i}) = \mathbb{Z}$ for $i = 0, 1$. There exists an integer $n \geq 0$ and an isometry*

$$\lambda_{N_{\Sigma_0}} \oplus \mathcal{H}_2^{\oplus n} \cong \lambda_{N_{\Sigma_1}} \oplus \mathcal{H}_2^{\oplus n}.$$

Whereas Theorems 1.1 and 1.2 hold for $g \neq 1, 2$, Proposition 1.6 leads to some results for arbitrary $g$. Indeed, Corollary 6.6 shows that for any genus $g$ surface $\Sigma \subseteq N = X \setminus \overset{\circ}{D}{}^4$ with knot group $\mathbb{Z}$ and $\partial \Sigma$ an Alexander polynomial one knot, $\lambda_{N_\Sigma} \oplus \mathcal{H}_2^{\oplus n} \cong Q_X \oplus \mathcal{H}_2^{\oplus(g+n)}$ for some $n \geq 0$, where $Q_X$ denotes the standard intersection form of $X$. The same result holds for closed surfaces $\Sigma \subseteq X$ with knot group $\mathbb{Z}$.

## 1.5 Application to rim surgery

Rim surgery is an effective way to produce smoothly inequivalent embeddings of surfaces; see Fintushel and Stern [14], Juhász, Miller and Zemke [24], Kim [30], Kim and Ruberman [31] and Mark [42]. Given a locally flat embedded closed oriented surface $\Sigma$ in a closed 4–manifold $X$, a knot $J \subseteq S^3$ and simple closed curve $\alpha \subseteq \Sigma$, *n–roll m–twist rim surgery* outputs another locally flat embedded closed oriented surface $\Sigma_n^m(\alpha, J) \subseteq X$. The case where $m = 0, n = 0$ was first introduced by Fintushel and Stern [14], while $m$–twisted rim surgery was introduced by Kim [30], and $n$–roll $m$–twist rim surgery first appeared in Kim and Ruberman [32].

Kim and Ruberman [32, Theorem 1.3 and Proposition 3.1] showed that if $X$ is a simply connected 4–manifold, $\pi_1(X_\Sigma) = \mathbb{Z}_d$ is a finite cyclic group and $(m, d) = 1$, then $\Sigma_n^m(\alpha, J)$ and $\Sigma$ are topologically isotopic. In order to extend this to infinite cyclic fundamental groups, we consider $n$–roll 1–twist rim surgery, and set $\Sigma_n(\alpha, J) := \Sigma_n^1(\alpha, J)$. In this case if $\pi_1(X_\Sigma) = \mathbb{Z}$, then also $\pi_1(X_{\Sigma_n(\alpha, J)}) = \mathbb{Z}$. Our first result on rim surgery then reads as follows.

**Theorem 1.7** *Let $X$ be a closed, simply connected 4–manifold, let $\Sigma \subseteq X$ be a locally flat, embedded, orientable surface with knot group $\mathbb{Z}$, let $\alpha \subseteq \Sigma$ be a simple closed curve, let $J$ be a knot, and let $n \in \mathbb{Z}$. Then the surfaces $\Sigma$ and $\Sigma_n(\alpha, J)$ are topologically ambiently isotopic.*

For good fundamental groups $\pi_1(X \setminus \Sigma)$ and $J$ a slice knot, Kim and Ruberman showed in [31, Theorem 4.5] that there is a homeomorphism of pairs $(X, \Sigma) \cong (X, \Sigma_0(\alpha, J))$, so for $\pi_1(X \setminus \Sigma) \cong \mathbb{Z}$ our result is an extension.

Rim surgery is also defined for a properly embedded surface in a 4–manifold $N$ with boundary. Juhász, Miller and Zemke [24, Corollary 2.7] showed that if $\Sigma \subseteq N$ is a properly embedded locally flat oriented surface, and if $\alpha$ bounds a locally flat disc in $N_\Sigma$, then $\Sigma_0(\alpha, J)$ is topologically isotopic to $\Sigma$. In the case where $\Sigma$ has knot group $\mathbb{Z}$, we generalise this result as follows.

**Theorem 1.8** *Let $X$ be a closed, simply connected 4–manifold, and let $N := X \setminus \overset{\circ}{D}{}^4$ be a punctured $X$. Let $\Sigma \subseteq N$ be a locally flat, properly embedded, orientable surface with knot group $\mathbb{Z}$, let $\alpha \subseteq \Sigma$ be a simple closed curve, let $J$ be a knot, and let $n \in \mathbb{Z}$ be an integer. There is a rel boundary orientation-preserving homeomorphism of pairs*

$$(N, \Sigma) \overset{\cong}{\longrightarrow} (N, \Sigma_n(\alpha, J)).$$

*If $N = D^4$, then the surfaces $\Sigma$ and $\Sigma_n(\alpha, J)$ are topologically ambiently isotopic rel boundary.*

Note that in both of the previous two results, we have no restrictions on the genera of the surfaces. This is because we show that the $\Lambda$–intersection forms of $N_{\Sigma_n(\alpha, J)}$ and $N_\Sigma$ coincide, allowing us to apply the first item of Theorem 1.3, for which there is no genus restriction.

For the case $N = D^4$, Theorem 1.8 means that the construction of exotic surfaces in [24] applies more generally than realised in that article, because the condition that $\alpha$ bounds a locally flat embedded disc in $N_\Sigma$ is not needed.

## 1.6 Application to pushed-in Seifert surfaces

Pushing a Seifert surface for a knot $K$ into $D^4$ yields a surface with knot group $\mathbb{Z}$. It is then intriguing to wonder whether any two Seifert surfaces of the same genus for a knot $K$ become isotopic once pushed into $D^4$; see for instance Problem 1.20(C) of Kirby [33] and Section 6 of Livingston [40]. In Theorem 7.11, relying on Theorem 1.3, we show that this is the case for Alexander polynomial one knots.

**Theorem 1.9** *If $F_0, F_1 \subseteq D^4$ are genus $g$ pushed-in Seifert surfaces for an Alexander polynomial one knot $K$, then they are topologically ambiently isotopic rel boundary.*

This result provides another setting where we are able to obtain results for $g = 1, 2$. In particular, it rules out the most naive potential counter examples to the conjecture that Theorem 1.2 holds for $\mathbb{Z}$–surfaces of any genus.

## 1.7  A classification result for topological 4–manifolds

Several of the steps from the proof of Theorem 1.3 fit in a classification scheme that applies to a wider class of 4–manifolds than surface exteriors. We state the resulting theorem, which is analogous to Boyer's classification [5] for the simply connected case.

We say that a compact, oriented, topological 4–manifold $M$ with nonempty connected boundary has *ribbon boundary* if the inclusion induced map $\iota : \pi_1(\partial M) \to \pi_1(M)$ is surjective. For example, in simply connected 4–manifolds, (connected) surface exteriors with infinite cyclic fundamental group have ribbon boundaries. Set $\Lambda := \mathbb{Z}[t^{\pm 1}]$. If we have an identification $\pi_1(M) = \mathbb{Z}$, then we have infinite cyclic covers $\widetilde{M} \to M$ and $\partial \widetilde{M} \to \partial M$ and $\Lambda$–modules $H_*(M; \Lambda) = H_*(\widetilde{M})$ and $H_*(\partial M; \Lambda) = H_*(\partial \widetilde{M})$.

Let $M_0$ and $M_1$ be two 4–manifolds with infinite cyclic fundamental group, whose boundaries are ribbon and have $\Lambda$–torsion Alexander modules. Fix an orientation-preserving homeomorphism $f : \partial M_0 \to \partial M_1$ that intertwines the inclusion induced maps $\varphi_i : \pi_1(\partial M_i) \twoheadrightarrow \mathbb{Z}$, and an isometry $F : (H_2(M_0; \Lambda), \lambda_{M_0}) \to (H_2(M_1; \Lambda), \lambda_{M_1})$ of the $\Lambda$–valued intersection forms $\lambda_{M_0}$ and $\lambda_{M_1}$. As described in Section 3.4 below, $f$ and $F$ induce isometries of the boundary Blanchfield forms:

$$f_*, \partial F : (H_1(\partial M_0; \Lambda), \mathrm{Bl}_{\partial M_0}) \to (H_1(\partial M_1; \Lambda), \mathrm{Bl}_{\partial M_1}).$$

We call $(f, F)$ a *compatible pair* if $f$ and $F$ induce the same isometry. For conciseness, we write

$$\mathrm{Homeo}_\varphi^+(\partial M_0, \partial M_1) := \{ f \in \mathrm{Homeo}^+(\partial M_0, \partial M_1) \mid \varphi_1 \circ f_* = \varphi_0 \}$$

for the set of orientation-preserving homeomorphisms of the boundaries that intertwine the maps to $\mathbb{Z}$.

Recall that the Kirby–Siebenmann invariant $\mathrm{ks}(M_i) \in \mathbb{Z}/2$ is the unique obstruction for the stable tangent bundle $M_i \to \mathrm{BTOP}$ to lift to $M_i \to \mathrm{BPL}$, or equivalently for $M_i$ to be smoothable after adding copies of $S^2 \times S^2$; see Freedman and Quinn [15, Theorem 8.6] and Friedl, Nagel, Orson and Powell [17, Section 8]. This is relevant in the next theorem for the nonspin case; note that for 4–manifolds with fundamental group $\mathbb{Z}$, whether or not they are spin is determined by the intersection pairing.

**Theorem 1.10** *Suppose $M_0$ and $M_1$ are two compact, oriented 4–manifolds with $\pi_1(M_i) = \mathbb{Z}$ for $i = 0, 1$, whose boundaries are ribbon and have $\Lambda$–torsion Alexander modules. Let $f \in \mathrm{Homeo}_\varphi^+(\partial M_0, \partial M_1)$ be a homeomorphism and let $F \colon \lambda_{M_0} \cong \lambda_{M_1}$ be an isometry. If $M_0$ and $M_1$ are not spin, assume that the Kirby–Siebenmann invariants satisfy $\mathrm{ks}(M_0) = \mathrm{ks}(M_1) \in \mathbb{Z}/2$. Then the following assertions are equivalent:*

(i) *The pair $(f, F)$ is compatible.*

(ii) *The homeomorphism $f$ extends to an orientation-preserving homeomorphism*

$$\Phi \colon M_0 \xrightarrow{\;\cong\;} M_1$$

*inducing the given isometry $F \colon H_2(M_0; \Lambda) \cong H_2(M_1; \Lambda)$.*

**Remark 1.11** Theorem 1.10 should be compared with both Boyer's classification [5] of simply connected 4–manifolds with boundary and with Freedman and Quinn's classification [15] of closed 4–manifolds with infinite cyclic fundamental group. Boyer uses a notion, similar to our compatible pairs, which he calls *morphisms*. In a nutshell, given simply connected 4–manifolds $M_0, M_1$ with rational homology spheres as their boundaries, Boyer [5, Theorem 0.7 and Proposition 0.8] shows that a morphism $(f, F)$ can be extended to a homeomorphism $M_0 \to M_1$. Note that Boyer's methods differ from ours, as he does not use the union of forms. On the other hand, Freedman and Quinn [15, Theorem 10.7A] show that closed 4–manifolds with infinite cyclic fundamental groups are classified by their $\Lambda$–intersection form and their Kirby–Siebenmann invariant; see also Stong and Wang [48] and Hambleton, Kreck and Teichner [20].

A major step in the proof of Theorem 1.10 is the following intermediate result, which might be of independent interest; a more detailed statement and a proof can be found in Theorem 3.12.

**Theorem 1.12** *Suppose $M_0$ and $M_1$ are two compact, oriented 4–manifolds with $\pi_1(M_i) = \mathbb{Z}$ for $i = 0, 1$, whose boundaries are ribbon and have $\Lambda$–torsion Alexander modules. Let $(f, F)$ be a compatible pair. If $M_0$ and $M_1$ are not spin, assume that the Kirby–Siebenmann invariants satisfy $\mathrm{ks}(M_0) = \mathrm{ks}(M_1) \in \mathbb{Z}/2$. Then there is an orientation-preserving homeomorphism*

$$M_0 \cup_f -M_1 \cong S^1 \times S^3 \mathbin{\#} \overset{a}{\underset{i=1}{\#}} S^2 \times S^2 \mathbin{\#} \overset{b}{\underset{j=1}{\#}} S^2 \widetilde{\times} S^2$$

*for some $a, b$ with $a + b = b_2(M_0)$. If $M_0$ and $M_1$ are spin, then $b = 0$.*

**Remark 1.13**  We compare Theorem 1.12 with a particular case of a result due to Hambleton and Teichner [21]. Hambleton and Teichner show that if $M$ is a closed, oriented, topological 4–manifold with infinite cyclic fundamental group and with $b_2(M) - |\sigma(M)| \geq 6$, then $M$ is homeomorphic to the connected sum of $S^1 \times S^3$ with a unique closed, simply connected 4–manifold [21, Corollary 3]. In most cases, this result is stronger than Theorem 1.12. However, there are some situations (such as the union of two genus one surface exteriors $X_{\Sigma_0}$, $X_{\Sigma_1} \subseteq S^4$ with $\pi_1(X_{\Sigma_i}) = \mathbb{Z}$ for $i = 0, 1$) where Theorem 1.12 applies while [21, Corollary 3] does not. Finally, note that Sunukjian [49, Theorem 7.2] has used these results to study closed surfaces $\Sigma \subseteq X$ with $\pi_1(X \setminus \nu\Sigma) = \mathbb{Z}$ in closed 4–manifolds that satisfy $b_2(X) \geq |\sigma(X)| + 6$.

## 1.8  Isometries of the Blanchfield form

Theorem 1.10 provides a criterion for extending a homeomorphism $f \colon \partial M_0 \to \partial M_1$ to a homeomorphism $M_0 \to M_1$: one must fit $f$ into a compatible pair. However, in practice, finding compatible pairs is difficult and so we provide some sufficient conditions for their existence.

Recall that if $V$ is a 4–manifold with $\pi_1(V) = \mathbb{Z}$ whose boundary $Y = \partial V$ is ribbon and has torsion Alexander module, then an isometry $F$ of the $\Lambda$–intersection form $\lambda_V$ induces an isometry

$$\partial F \colon (H_1(Y; \Lambda), \mathrm{Bl}_Y) \xrightarrow{\cong} (H_1(Y; \Lambda), \mathrm{Bl}_Y).$$

We write $\mathrm{Aut}(\lambda_V)$ and $\mathrm{Aut}(\mathrm{Bl}_Y)$ for the groups of isometries of $\lambda_V$ and $\mathrm{Bl}_Y$, and note that there is a left action of $\mathrm{Aut}(\lambda_V)$ on $\mathrm{Aut}(\mathrm{Bl}_Y)$ given by $F \cdot h = h \circ \partial F^{-1}$. We will be interested in a quotient of the orbit set

$$\mathrm{Aut}(\mathrm{Bl}_Y) / \mathrm{Aut}(\lambda_V).$$

In order to find compatible pairs, we need to know when an isometry of the Blanchfield form is induced by a homeomorphism. As we have already alluded to, we note in Proposition 3.7 that any homeomorphism $f \in \mathrm{Homeo}_\varphi^+(Y)$ induces a $\Lambda$–isometry $f_*$ of $\mathrm{Bl}_Y$ by lifting $f$ to the infinite cyclic covers. Here, $\mathrm{Homeo}_\varphi^+(Y) := \mathrm{Homeo}_\varphi^+(Y, Y)$ denotes the set of orientation-preserving self-homeomorphisms of $Y$ that intertwine the map $\pi_1(Y) \to \mathbb{Z}$. The assignment $(f, F) \cdot h = f_* \circ h \circ \partial F^{-1}$ then gives rise to a left action

$$\mathrm{Homeo}_\varphi^+(Y) \times \mathrm{Aut}(\lambda_V) \curvearrowright \mathrm{Aut}(\mathrm{Bl}_Y).$$

We return to our 4–manifolds $M_0$ and $M_1$, whose boundaries are ribbon and have torsion Alexander module. If $M_0$ and $M_1$ are orientation-preserving homeomorphic, then a compatible pair exists; recall Theorem 1.10. Assuming that we are given a homeomorphism $f' \in \mathrm{Homeo}_\varphi^+(\partial M_0, \partial M_1)$ and an isometry $F' \colon \lambda_0 \cong \lambda_1$, Proposition 3.14 notes that $(f', F')$ gives rise to a compatible pair $(f, F)$ if the composition $f'_* \circ \partial F'^{-1}$ is trivial in $\mathrm{Aut}(\mathrm{Bl}_{\partial M_1})/\mathrm{Homeo}_\varphi^+(\partial M_1) \times \mathrm{Aut}(\lambda_{M_1})$. Using this result, we obtain the following consequence of Theorem 1.10.

**Theorem 1.14** *Suppose $M_0$ and $M_1$ are two compact, oriented 4–manifolds with $\pi_1(M_i) = \mathbb{Z}$ for $i = 0, 1$, whose boundaries are ribbon and have $\Lambda$–torsion Alexander modules. Assume $M_0$ and $M_1$ have orientation-preserving homeomorphic boundaries, via a homeomorphism that intertwines the maps to $\mathbb{Z}$, and isometric intersection forms. If the orbit set*

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M_1})/\mathrm{Homeo}_\varphi^+(\partial M_1) \times \mathrm{Aut}(\lambda_{M_1})$$

*is trivial, then there is an orientation-preserving homeomorphism $M_0 \cong M_1$.*

Theorem 1.14 is helpful to summarise a programme to prove that two 4–manifolds $M_0$ and $M_1$ (with infinite cyclic fundamental group and identical ribbon boundaries with torsion Alexander modules) are homeomorphic:

  (i)  Decide whether the $\Lambda$–intersection forms $\lambda_{M_0}$ and $\lambda_{M_1}$ are isometric.
  (ii) Show that the orbit set $\mathrm{Aut}(\mathrm{Bl}_{\partial M_1})/\mathrm{Homeo}_\varphi^+(\partial M_1) \times \mathrm{Aut}(\lambda_{M_1})$ is trivial.

To illustrate this paradigm, we explain how Theorem 1.3 can be applied to classify $\mathbb{Z}$–surfaces for some knots purely in terms of their intersection pairings. For any knot $K \subseteq S^3$, multiplication by a monomial $\pm t^k$ gives rise to an automorphism of its Blanchfield pairing $\mathrm{Bl}_{E_K}$ on the exterior of the knot. For some knots $K$, these are the only automorphisms, and in this case we obtain a simpler classification of $\mathbb{Z}$–surfaces, ie locally flat, oriented surfaces $\Sigma$ in $N$ with $\pi_1(N_\Sigma) \cong \mathbb{Z}$, with boundary $K$.

**Corollary 1.15** *Let $X$ be a closed, simply connected, oriented 4–manifold, and let $N := X \setminus \mathring{D}^4$ be a punctured $X$. Let $\Sigma_0$ and $\Sigma_1$ be two $\mathbb{Z}$–surfaces in $N$ with the same genus and oriented boundary, so $\partial \Sigma_0 = \partial \Sigma_1 = K \subseteq S^3$. Suppose that $\mathrm{Aut}(\mathrm{Bl}_{E_K}) \subseteq \{\pm t^k \mid k \in \mathbb{Z}\}$.*

  (i)  *There is a rel boundary homeomorphism of pairs $(N, \Sigma_0) \cong (N, \Sigma_1)$ if and only if there is an isometry $\lambda_{N_{\Sigma_0}} \cong \lambda_{N_{\Sigma_1}}$.*
  (ii) *If $N = D^4$, then $\Sigma_0$ and $\Sigma_1$ are ambiently isotopic rel boundary if and only if there is an isometry $\lambda_{D_{\Sigma_0}^4} \cong \lambda_{D_{\Sigma_1}^4}$.*

**Proof** Let $F\colon \lambda_{N_{\Sigma_0}} \xrightarrow{\cong} \lambda_{N_{\Sigma_1}}$ be an isometry. Consider the isomorphism $g := \partial F|_{H_1(E_K;\Lambda)}$ induced on the Alexander module of $K$. It is an automorphism of the Blanchfield form and therefore by the hypothesis on $\text{Aut}(\text{Bl}_{E_K})$, $g$ is multiplication by $\pm t^k$ for some $k \in \mathbb{Z}$. Precompose $F$ with multiplication by $\pm t^{-k}$, to obtain a new isometry inducing Id on $H_1(E_K;\Lambda)$. Then apply Theorem 1.3 with $f_K = \text{Id}$. □

**Remark 1.16** If the Alexander module $H_1(E_K;\Lambda)$ is cyclic with order $\Delta_K$, then the automorphisms of the Blanchfield pairing can be computed directly as follows. Suppose that $\text{Bl}_K(1,1) = b/\Delta_K \in \mathbb{Q}(t)/\Lambda$ for some $b \in \Lambda$. Then

$$\text{Aut}(\text{Bl}_{E_K}) = \{[p] \in \Lambda/\Delta\Lambda \mid [p \cdot \overline{p} \cdot b] = [b] \in \Lambda/\Delta\Lambda\}.$$

In particular, for $K = \pm T_{2,3}$ a trefoil, it is not too hard to compute that $\text{Aut}(\text{Bl}_{E_K}) = \{\pm t^k \mid k = -1, 0, 1\}$, so that Corollary 1.15 applies to classify all $\mathbb{Z}$–surfaces with boundary a trefoil in terms of the isometry type of the intersection pairing.

## Organisation

In Section 2, we review some notions on linking forms and define the union of Hermitian forms over $\Lambda$. In Section 3, we show how this algebraic union can be used to express the intersection form of a union of two 4–manifolds; we also prove Theorem 1.12. In Section 4, we prove Theorem 1.10, our partial classification result for 4–manifolds with infinite cyclic fundamental group. In Section 5, we prove Theorems 1.3 and 1.4, our main results on properly embedded surfaces in simply connected 4–manifolds. In Section 6, we discuss the equivariant intersection forms for surfaces with knot group $\mathbb{Z}$, before focusing on surfaces in $S^4$ and $D^4$ in Section 7, where we prove Theorems 1.1 and 1.2. In Section 8, we prove Theorems 1.7 and 1.8 on rim surgery. In the appendix, we adapt the work of Baykur and Sunukjian to properly embedded surfaces in 4–manifolds with boundary.

## Conventions

(i)  From now on, all manifolds are assumed to be compact, connected, based and oriented; if a manifold has a nonempty boundary, then the basepoint is assumed to be in the boundary. We work in the topological category with locally flat embeddings unless otherwise stated.

(ii)  $X$ will always denote a closed, simply connected, oriented 4–manifold and we will write $N := X \setminus \mathring{D}^4$ for the complement of an open ball $\mathring{D}^4 \subseteq X$.

(iii)  A $\mathbb{Z}$–*surface* $\Sigma$ will always refer to a compact, connected, oriented, locally flat, embedded surface in a 4–manifold, whose knot group is isomorphic to $\mathbb{Z}$; a surface $\Sigma \subseteq X$ is understood to be closed, while a surface $\Sigma \subseteq N$ is understood to be properly embedded.

(iv)  If $P$ is manifold and $Q \subseteq P$ is a submanifold with tubular neighbourhood $\nu Q \subseteq P$, then $P_Q := P \setminus \nu Q$ will always denote the exterior of $Q$ in $P$. The only exception to this is that the exterior of a knot $K$ in $S^3$ will be denoted by $E_K$ instead of $S^3_K$.

(v)  Throughout the article, we set $\Lambda := \mathbb{Z}[t^{\pm 1}]$ for the ring of Laurent polynomials and $Q := \mathbb{Q}(t)$ for its field of fractions.

(vi)  For a manifold $X$, we write $\pi_1(X) = \mathbb{Z}$ to mean that there is an isomorphism $\pi_1(X) \cong \mathbb{Z}$, and that we have fixed a choice of such an isomorphism. For the exterior of a $\mathbb{Z}$–surface, the orientations on $X$ and $\Sigma$ determine an identification of the fundamental group with $\mathbb{Z}$. The choice of map $\pi_1(X) = \mathbb{Z}$ determines an isomorphism $\mathbb{Z}[\pi_1(X)] \to \Lambda$, which we use to define homology and cohomology with $\Lambda$ coefficients.

(vii)  We write $p \mapsto \bar{p}$ for the involution on $\Lambda$ induced by sending $t \mapsto t^{-1}$ and extending linearly. Given a $\Lambda$–module $H$, we write $\bar{H}$ for the $\Lambda$–module whose underlying abelian group is $H$ but with module structure given by $p \cdot h = \bar{p}h$ for $h \in H$ and $p \in \Lambda$.

(viii)  We write $H^* := \overline{\mathrm{Hom}_\Lambda(H, \Lambda)}$.

(ix)  For any ring $R$, elements of $R^n$ are considered as column vectors.

(x)  Let $f : X \to Y$ be a morphism in a category $\mathcal{C}$, and let $F : \mathcal{C} \to \mathcal{D}$ be a contravariant functor. Assume that $F(f)$ is invertible. We shall often write the induced morphism in $\mathcal{D}$ as $f^* = F(f)$, and denote its inverse by $f^{-*} := (f^*)^{-1} = F(f)^{-1}$.

# 2 The union of forms along isometries of their boundary linking forms

We develop the theory required for computing the intersection form of a closed 4–manifold obtained as the union of two compact 4–manifolds along their common boundary. Here is a summary of this section. In Section 2.1, we review the boundary linking form of a Hermitian form. In Section 2.2, we discuss isometries of Hermitian forms and linking forms. In Section 2.3, we study the union of two Hermitian forms over $\Lambda := \mathbb{Z}[t^{\pm 1}]$. In Section 2.4, we provide a condition for this union to admit a metaboliser.

## 2.1 The boundary of Hermitian forms and their isometries

We fix our terminology on Hermitian forms, linking forms and their boundaries. References include [47, Section 3.4] and [12, Section 6].

A *Hermitian form* over $\Lambda$ is a pair $(H, \lambda)$, where $H$ is a free $\Lambda$–module and

$$\lambda : H \times H \to \Lambda$$

is a sesquilinear Hermitian pairing. Here by sesquilinear, we mean that $\lambda(px, qy) = p\lambda(x, y)\overline{q}$ for all $x, y \in H$ and all $p, q \in \Lambda$. By Hermitian, we mean that $\lambda(y, x) = \overline{\lambda(x, y)}$ for all $x, y \in H$. The *adjoint* of $\lambda$ is the $\Lambda$–linear map

$$\widehat{\lambda} : H \to \overline{\mathrm{Hom}_\Lambda(H, \Lambda)} =: H^*$$

such that $\widehat{\lambda}(y)(x) = \lambda(x, y)$. A Hermitian form is *nondegenerate* if its adjoint is injective and *nonsingular* if its adjoint is an isomorphism. The *standard hyperbolic form* is

$$H^+(\Lambda) := \left( \Lambda \oplus \Lambda, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right).$$

**Remark 2.1** If $(H, \lambda)$ is a nondegenerate Hermitian form, then $\mathrm{coker}(\widehat{\lambda})$ is a torsion $\Lambda$–module. This can be seen by tensoring the following exact sequence with the field of fractions $Q := \mathbb{Q}(t)$ of $\Lambda$:

$$0 \to H \overset{\widehat{\lambda}}{\hookrightarrow} H^* \to \mathrm{coker}(\widehat{\lambda}) \to 0.$$

A *linking form* $(T, \beta)$ over $\Lambda$ consists of a torsion $\Lambda$–module $T$ together with a sesquilinear Hermitian form $\beta : T \times T \to Q/\Lambda$. The *adjoint* of $\beta$ is the $\Lambda$–linear map $\widehat{\beta} : T \to \overline{\mathrm{Hom}_\Lambda(T, Q/\Lambda)}$ such that $\widehat{\beta}(y)(x) = \beta(x, y)$. A linking form is *nondegenerate* if its adjoint is injective and *nonsingular* if its adjoint is an isomorphism.

**Definition 2.2** The *boundary linking form* of a nondegenerate Hermitian form $(H, \lambda)$ over $\Lambda$ is the linking form $(\mathrm{coker}(\widehat{\lambda}), \partial\lambda)$, where $\partial\lambda$ is defined as

$$\partial\lambda \colon \mathrm{coker}(\widehat{\lambda}) \times \mathrm{coker}(\widehat{\lambda}) \to Q/\Lambda, \quad ([x], [y]) \mapsto \frac{1}{s}(y(z)),$$

where, since $\mathrm{coker}(\widehat{\lambda})$ is $\Lambda$–torsion, there exists an $s \in \Lambda$ and a $z \in H$ such that $sx = \widehat{\lambda}(z)$.

It is not difficult to show that $\partial\lambda$ is independent of the choices involved, and is sesquilinear and Hermitian. We conclude with two remarks that we will use throughout this section.

**Remark 2.3** Let $(H, \lambda)$ be a Hermitian form, and set $H_Q := H \otimes_\Lambda Q$. Since $H$ is free, we can identify $\overline{\mathrm{Hom}_\Lambda(H, \Lambda)} \otimes_\Lambda Q$ with $H_Q^* := \overline{\mathrm{Hom}_Q(H_Q, Q)}$. As stated in Remark 2.1, if $(H, \lambda)$ is nondegenerate, then $\mathrm{coker}(\widehat{\lambda})$ is a torsion $\Lambda$–module and therefore $\widehat{\lambda}_Q := \widehat{\lambda} \otimes_\Lambda \mathrm{Id}_Q \colon H_Q \to H_Q^*$ is a nonsingular Hermitian form over $Q$.

Using this remark, we describe an equivalent definition of the boundary linking form.

**Remark 2.4** The boundary linking form of a nondegenerate Hermitian form $(H, \lambda)$ can be described as $\partial\lambda([x], [y]) = y(\widehat{\lambda}_Q^{-1}(x))$ for $[x], [y] \in \mathrm{coker}(\widehat{\lambda})$. Choose a basis $\boldsymbol{b} = (e_i)_{i=1}^n$ for $H$ and endow $\overline{\mathrm{Hom}_\Lambda(H, \Lambda)}$ with the dual basis $\boldsymbol{b}^*$. If $A_{ij} = \lambda(e_i, e_j)$ is a Hermitian matrix representing $\lambda$, then $\overline{A}$ is a matrix for $\widehat{\lambda}$ and $\partial\lambda([x], [y]) = y(\widehat{\lambda}_Q^{-1}(x)) = (\overline{A}^{-1}x)^T \overline{y} = x^T A^{-1} \overline{y}$. For future reference, note that since $A$ is Hermitian, we have $\overline{A} = A^T$.

## 2.2 Isometries of forms

We discuss isometries of Hermitian forms and linking forms. References include [47, Section 3.4] and [12, Section 6].

Let $(H_0, \lambda_0)$ and $(H_1, \lambda_1)$ be Hermitian forms over $\Lambda$. A $\Lambda$–linear isomorphism $F \colon H_0 \to H_1$ is an *isometry* if $\lambda_1(F(x), F(y)) = \lambda_0(x, y)$ for all $x, y \in H_0$. Let $\mathrm{Iso}(\lambda_0, \lambda_1)$ denote the set of isometries between the Hermitian forms $(H_0, \lambda_0)$ and $(H_1, \lambda_1)$, and let $\mathrm{Aut}(\lambda) := \mathrm{Iso}(\lambda, \lambda)$ denote the group of self-isometries of a Hermitian form $(H, \lambda)$.

We make the analogous definitions for linking forms. Let $(T_0, \beta_0)$ and $(T_1, \beta_1)$ be two linking forms over $\Lambda$. A $\Lambda$–linear isomorphism $h \colon T_0 \to T_1$ is an *isometry*

if $\beta_1(h(x), h(y)) = \beta_0(x, y)$ for all $x, y \in T_0$. We let $\mathrm{Iso}(\beta_0, \beta_1)$ denote the set of isometries between the linking forms $(T_0, \beta_0)$ and $(T_1, \beta_1)$, and let $\mathrm{Aut}(\beta) := \mathrm{Iso}(\beta, \beta)$ denote the group of self-isometries of a linking form $(T, \beta)$.

An isomorphism $F: H_0 \to H_1$ induces an isomorphism $F^{-*} := (F^*)^{-1}: H_0^* \to H_1^*$. If additionally, the isomorphism $F \in \mathrm{Iso}(\lambda_0, \lambda_1)$ is an isometry, then $F^{-*}$ descends to an isomorphism

$$\partial F := F^{-*}: \mathrm{coker}(\widehat{\lambda}_0) \to \mathrm{coker}(\widehat{\lambda}_1).$$

For later use, we note that $F: (H_0, \lambda_0) \to (H_1, \lambda_1)$ is an isometry if and only if $\widehat{\lambda}_0 = F^* \widehat{\lambda}_1 F$. Next, we verify that $\partial F$ is an isometry of the boundary linking forms.

**Lemma 2.5** *If $F: (H_0, \lambda_0) \to (H_1, \lambda_1)$ is an isometry of nondegenerate Hermitian forms, then $\partial F$ is an isometry of linking forms*:

$$\partial F: (\mathrm{coker}(\widehat{\lambda}_0), \partial\lambda_0) \to (\mathrm{coker}(\widehat{\lambda}_1), \partial\lambda_1).$$

*This construction provides a map*

$$\partial: \mathrm{Iso}(\lambda_0, \lambda_1) \to \mathrm{Iso}(\partial\lambda_0, \partial\lambda_1),$$

*which is a homomorphism on automorphism groups.*

**Proof**   Given $[x], [y] \in \mathrm{coker}(\widehat{\lambda}_0)$, there exist $s \in \Lambda$ and $z \in H_0$ such that $sx = \widehat{\lambda}_0(z)$ . Since $F$ is an isometry, it follows that $sF^{-*}(x) = F^{-*}\widehat{\lambda}_0(z) = \widehat{\lambda}_1 F(z)$. We can now conclude that $\partial F$ is an isometry because

$$\partial\lambda_1(\partial F([x]), \partial F([y])) = \partial\lambda_1(F^{-*}([x]), F^{-*}([y]))$$
$$= \frac{1}{s} F^{-*}(y) F(z)$$
$$= \frac{1}{s} y(z)$$
$$= \partial\lambda_0([x], [y]).$$

The last assertion follows from the equality $(G \circ F)^{-*} = G^{-*} \circ F^{-*}$.   $\square$

Using Lemma 2.5, we are led to the following definition.

**Definition 2.6**   The *boundary of an isometry* $F: (H_0, \lambda_0) \to (H_1, \lambda_1)$ of nondegenerate Hermitian forms is the isometry of linking forms

$$\partial F: \partial(H_0, \lambda_0) \to \partial(H_1, \lambda_1).$$

## 2.3 The union of nondegenerate Hermitian forms

We define the union of two Hermitian forms $(H_0, \lambda_0)$ and $(H_1, \lambda_1)$ over $\Lambda$ along an isometry of their boundary linking forms. The definition is inspired by [11, Chapter 3], which was concerned with forms over the integers.

We describe the main construction of this subsection. In what follows, for $i = 0, 1$, we use $\pi_i \colon H_i^* \to \operatorname{coker}(\widehat{\lambda}_i)$ to denote the canonical projections.

**Construction 2.7** Let $(H_0, \lambda_0)$ and $(H_1, \lambda_1)$ be two nondegenerate Hermitian forms over $\Lambda$, and let $h \colon (\operatorname{coker}(\widehat{\lambda}_0), \partial \lambda_0) \to (\operatorname{coker}(\widehat{\lambda}_1), \partial \lambda_1)$ be an isometry of their boundary linking forms. Consider the pair $(H, \lambda)$ with

$$H := \ker(h\pi_0 - \pi_1 \colon H_0^* \oplus H_1^* \to \operatorname{coker}(\widehat{\lambda}_1)),$$

$$\lambda\left(\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}\right) = \frac{1}{s_0} y_0(z_0) - \frac{1}{s_1} y_1(z_1) \in Q,$$

where since $\operatorname{coker}(\widehat{\lambda}_i)$ is torsion, there exist $s_i \in \Lambda$ and $z_i \in H_i$ such that $s_i x_i = \widehat{\lambda}_i(z_i)$. Since the Hermitian forms $\lambda_0$ and $\lambda_1$ are nondegenerate, it is not difficult to prove that the pairing $\lambda$ does not depend on the choice of $s_0, s_1, z_0, z_1$.

The next proposition establishes some facts about the pairing $(H, \lambda)$.

**Proposition 2.8** *The pair $(H, \lambda)$ from Construction 2.7 has the following properties.*

(i) *For $(x_0, x_1), (y_0, y_1) \in H$, the pairing $\lambda$ can equivalently be defined as*

$$\lambda\left(\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}\right) = y_0(\widehat{\lambda}_{0,Q}^{-1}(x_0)) - y_1(\widehat{\lambda}_{1,Q}^{-1}(x_1)).$$

(ii) *If we choose bases for $H_0$, $H_1$ and dual bases for $H_0^*$, $H_1^*$ and let $A_0$, $A_1$ be Hermitian matrices representing $\lambda_0, \lambda_1$, then*

$$\lambda\left(\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}\right) = x_0^T A_0^{-1} \overline{y}_0 - x_1^T A_1^{-1} \overline{y}_1.$$

(iii) *The pairing $\lambda$ is sesquilinear, Hermitian and takes values in $\Lambda$.*

(iv) *The following two maps, which we abridge by $\widehat{\lambda}_0, \widehat{\lambda}_1$, are injective:*

$$(H_0, \lambda_0) \xrightarrow{\begin{pmatrix} \widehat{\lambda}_0 \\ 0 \end{pmatrix}} (H, \lambda) \xleftarrow{\begin{pmatrix} 0 \\ \widehat{\lambda}_1 \end{pmatrix}} (H_1, -\lambda_1).$$

*Furthermore, these maps satisfy $\widehat{\lambda}_0(H_0)^{\perp} = \widehat{\lambda}_1(H_1)$.*

**Proof** To prove the first assertion, notice that since $\lambda_{i,Q}$ is a nonsingular pairing for $i = 0, 1$, we can write $\widehat{\lambda}_{i,Q}^{-1}(x_i) = z_i/s_i$ for some $z_i \in H$ and some $s_i \in \Lambda$. Consequently, $\widehat{\lambda}_i(z_i) = s_i x_i$ and we have $(1/s_i)y_i(z_i) = y_i(\widehat{\lambda}_{i,Q}^{-1}(x_i))$, as desired. This proves the first assertion. The second assertion now follows as in [Remark 2.4].

We prove the third assertion. The fact that $\lambda$ is sesquilinear and Hermitian follows from the second assertion and the fact that $A_0^{-1}$ and $A_1^{-1}$ are Hermitian. Next, using the definition of the boundary linking forms $\partial\lambda_0$ and $\partial\lambda_1$ from [Definition 2.2], we have

$$\lambda\left(\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}\right) = \partial\lambda_0([x_0], [y_0]) - \partial\lambda_1([x_1], [y_1]) \mod \Lambda.$$

By definition of $H$, we have $h[x_0] = x_1$ and $h[y_0] = y_1$. Then since $h$ is an isometry, this latter expression vanishes:

$$\begin{aligned} \partial\lambda_0([x_0], [y_0]) - \partial\lambda_1([x_1], [y_1]) &\equiv \partial\lambda_0([x_0], [y_0]) - \partial\lambda_1(h[x_0], h[y_0]) \\ &\equiv \partial\lambda_0([x_0], [y_0]) - \partial\lambda_0([x_0], [y_0]) \equiv 0 \in Q/\Lambda. \end{aligned}$$

It follows that $\lambda$ takes values in $\Lambda$, concluding the proof of the third assertion. We now prove the fourth assertion. First, we check that $\widehat{\lambda}_0$ is an isometric embedding; the proof for $\widehat{\lambda}_1$ is identical. Since the pairings are nondegenerate, the maps are injective and given $z_0, z_0' \in H_0$, we have

$$\lambda\left(\begin{pmatrix} \widehat{\lambda}_0(z_0) \\ 0 \end{pmatrix}, \begin{pmatrix} \widehat{\lambda}_0(z_0') \\ 0 \end{pmatrix}\right) = \widehat{\lambda}(z_0')(z_0) = \lambda_0(z_0, z_0').$$

It remains to check that $\widehat{\lambda}_0(H_0)^\perp = \widehat{\lambda}_1(H_1)$. The inclusion $\widehat{\lambda}_0(H_0)^\perp \supseteq \widehat{\lambda}_1(H_1)$ is clear and so we prove the inclusion $\widehat{\lambda}_0(H_0)^\perp \subseteq \widehat{\lambda}_1(H_1)$. Assume that $(x_0, x_1) \in H$ (with $sx_0 = \widehat{\lambda}_0(z_0)$) satisfies, for all $a_0 \in H_0$,

$$0 = \lambda\left(\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} \widehat{\lambda}_0(a_0) \\ 0 \end{pmatrix}\right) = \frac{1}{s}\widehat{\lambda}_0(a_0)(z_0) = \frac{1}{s}\lambda_0(z_0, a_0) \in \Lambda.$$

As $\lambda_0$ is nondegenerate, this implies that $z_0 = 0$ and therefore $x_0 = 0$. But since $(x_0, x_1) \in H$, we also have $[x_1] = h[x_0] = 0$ and so $x_1 = \widehat{\lambda}_1(z_1)$ for some $z_1 \in H_1$. We therefore conclude that $\widehat{\lambda}_0(H_0)^\perp = \widehat{\lambda}_1(H_1)$, establishing the proposition. $\square$

Note that [Proposition 2.8] does not contain a statement about the $\Lambda$–module $H$ being free. Since we defined Hermitian forms over free $\Lambda$–modules, a *Hermitian module* $(H, \lambda)$ will refer to a pair consisting of a finitely generated $\Lambda$–module $H$ and a sesquilinear Hermitian pairing $\lambda \colon H \times H \to \Lambda$; we drop the requirement that $H$ be free.

**Definition 2.9**  Let $(H_0, \lambda_0)$ and $(H_1, \lambda_1)$ be two nondegenerate Hermitian forms over $\Lambda$, and let $h: (\mathrm{coker}(\widehat{\lambda}_0), \partial\lambda_0) \to (\mathrm{coker}(\widehat{\lambda}_1), \partial\lambda_1)$ be an isometry of their boundary linking forms. The *union* of $(H_0, \lambda_0)$ and $(H_1, \lambda_1)$ along $h$ is the Hermitian module $(H_0 \cup_h H_1, \lambda_0 \cup_h -\lambda_1)$ described in Construction 2.7 and Proposition 2.8:

$$H_0 \cup_h H_1 := \ker(h\pi_0 - \pi_1 \colon H_0^* \oplus H_1^* \to \mathrm{coker}(\widehat{\lambda}_1)),$$

$$\lambda_0 \cup_h -\lambda_1 \left( \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \right) = \frac{1}{s_0} y_0(z_0) - \frac{1}{s_1} y_1(z_1) \in \Lambda,$$

where, since $\mathrm{coker}(\widehat{\lambda}_i)$ is torsion, there exist $s_i \in \Lambda$ and $z_i \in H_i$ such that $s_i x_i = \widehat{\lambda}_i(z_i)$.

**Remark 2.10**  We briefly discuss the nonsingularity of $(H_0 \cup_h H_1, \lambda_0 \cup_h -\lambda_1)$. Proposition 2.8 contains no statement about $(H, \lambda) := (H_0 \cup_h H_1, \lambda_0 \cup_h -\lambda_1)$ being nonsingular. While this holds over the integers [11, Lemma 3.6] and in the topological setting of Proposition 3.9, we will not prove it in the algebraic generality of this section. If $\widehat{\lambda} \colon H \xrightarrow{\cong} \overline{\mathrm{Hom}_\Lambda(H, \Lambda)} =: H^*$ is nonsingular, then $H$ must be free, since for any finitely generated $\Lambda$–module $H$, the dual $H^*$ is free [3, Lemma 2.1].

Given a $\Lambda$–module $H$, we set $H^* := \overline{\mathrm{Hom}_\Lambda(H, \Lambda)}$. If $H$ is free, then there is a $\Lambda$–linear evaluation isomorphism $\mathrm{ev} \colon H \to H^{**}$ given by $\mathrm{ev}_x(\varphi) = \overline{\varphi(x)}$, where $x \in H$ and $\varphi \in H^*$. The next remark uses $\mathrm{ev}$ to describe the adjoint of $\lambda_0 \cup_h -\lambda_1$.

**Remark 2.11**  Let $(H_0, \lambda_0)$ and $(H_1, \lambda_1)$ be two nondegenerate Hermitian forms over $\Lambda$, and let $h \in \mathrm{Iso}(\partial\lambda_0, \partial\lambda_1)$ be an isometry. The adjoint of $\lambda := \lambda_0 \cup_h -\lambda_1$ is given by

$$\widehat{\lambda} \colon H_0 \cup_h H_1 \to (H_0 \cup_h H_1)^*, \quad \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \mapsto (\mathrm{ev}\ \mathrm{ev}) \begin{pmatrix} \widehat{\lambda}_{0,Q}^{-1} & 0 \\ 0 & -\widehat{\lambda}_{1,Q}^{-1} \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}.$$

This follows by combining the definition of $\mathrm{ev}$ with the second and third items of Proposition 2.8. Indeed, for $(x_0, y_0) \in H_0 \cup_h H_1$, we have

$$\lambda \left( \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \right) = \overline{\lambda \left( \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \right)} = \overline{x_0(\widehat{\lambda}_{0,Q}^{-1}(y_0))} - \overline{x_1(\widehat{\lambda}_{1,Q}^{-1}(y_1))}$$

$$= (\mathrm{ev}_{\widehat{\lambda}_{0,Q}^{-1}(y_0)}\ \mathrm{ev}_{-\widehat{\lambda}_{1,Q}^{-1}(y_1)}) \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \widehat{\lambda} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}.$$

We conclude this section by describing how the union interacts with boundary isometries. The proof follows immediately from the definitions and is left to the reader.

**Proposition 2.12** *Let $(H_0, \lambda_0)$, $(H_1, \lambda_1)$ be Hermitian forms, and $h \in \mathrm{Iso}(\partial\lambda_0, \partial\lambda_1)$ be an isometry. If $F \in \mathrm{Iso}(\lambda_0, \lambda_1)$ is an isometry, then $F^{-*} \oplus \mathrm{Id}$ induces an isometry*

$$(H_0, \lambda_0) \cup_h (H_1, -\lambda_1) \to (H_1, \lambda_1) \cup_{h \circ \partial F^{-1}} (H_1, -\lambda_1).$$

## 2.4 Lagrangian complements

This subsection provides a criterion for a union of Hermitian forms to be metabolic. It motivates the notion of a compatible pair, which will be introduced in the next subsection. The idea for this criterion stems from work of Kreck [37, Proposition 8] and Crowley and Sixt [12, Theorem 5.11], and their work on the monoid $\ell_{2q+1}(\mathbb{Z}[\pi])$.

A *Lagrangian* for a nonsingular Hermitian form $(H, \lambda)$ is a direct summand $L \subseteq H$ such that $L^\perp = L$. A nonsingular Hermitian form that admits a Lagrangian is called *metabolic*. A *Lagrangian complement* for a half-rank direct summand $V \subseteq (H, \lambda)$ is a Lagrangian of $(H, \lambda)$ such that $L \oplus V = H$.

In the next proposition, for an isomorphism $F \colon H_0 \to H_1$ we consider the graph

$$\Gamma_{F^{-*}} = \{(x, F^{-*}(x)) \mid x \in H_0^*\} \subseteq H_0^* \oplus H_1^*.$$

Since, by definition of $\partial F$, we have $\partial F(x) - F^{-*}(x) = F^{-*}(x) - F^{-*}(x) = 0$, we deduce the inclusions $\Gamma_{F^{-*}} \subseteq \ker(\partial F \circ \pi_0 - \pi_1) = H_0 \cup_{\partial F} H_1 \subseteq H_0^* \oplus H_1^*$.

**Proposition 2.13** *Let $F \colon (H_0, \lambda_0) \to (H_1, \lambda_1)$ be a $\Lambda$–isometry such that $\lambda_0 \cup_{\partial F} -\lambda_1$ is nonsingular. The graph $\Gamma_{F^{-*}} \subseteq H_0 \cup_{\partial F} H_1$ is a Lagrangian complement for $\widehat{\lambda}_0(H_0)$:*

$$H_0 \cup_{\partial F} H_1 = \widehat{\lambda}_0(H_0) \oplus \Gamma_{F^{-*}}.$$

**Proof** Set $\lambda := \lambda_0 \cup_{\partial F} -\lambda_1$. We show that $\Gamma_{F^{-*}}$ is a Lagrangian. To see that $\Gamma_{F^{-*}} \subseteq \Gamma_{F^{-*}}^\perp$, we must show that for all $x_0, y_0 \in H_0^*$,

$$\lambda\left(\begin{pmatrix} x_0 \\ F^{-*}(x_0) \end{pmatrix}, \begin{pmatrix} y_0 \\ F^{-*}(y_0) \end{pmatrix}\right) = 0.$$

Pick $s_0 \in \Lambda$ and $z_0 \in H_0$ so that $s_0 x_0 = \widehat{\lambda}_0(z_0)$. We apply $F^{-*}$ to both sides of this equation. Since $F$ is an isometry, we have $F^* \widehat{\lambda}_1 F = \widehat{\lambda}_0$, and we therefore obtain $s_0 F^{-*}(x_0) = \widehat{\lambda}_1(F(z_0))$. Using the definition of $\lambda$, we now obtain the desired conclusion:

$$\lambda\left(\begin{pmatrix} x_0 \\ F^{-*}(x_0) \end{pmatrix}, \begin{pmatrix} y_0 \\ F^{-*}(y_0) \end{pmatrix}\right) = \frac{1}{s_0} y_0(z_0) - \frac{1}{s_0} F^{-*}(y_0)(F(z_0)) = 0.$$

Next, we show that $\Gamma^{\perp}_{F-*} \subseteq \Gamma_{F-*}$. We therefore assume that $(x_0, x_1) \in H_0 \cup_{\partial F} H_1$ satisfies, for all $y_0 \in H_0^*$, the equation

$$\lambda\left(\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ F^{-*}(y_0) \end{pmatrix}\right) = 0,$$

and we must show that $x_1 = F^{-*}(x_0)$. Pick $s_0, s_1 \in \Lambda$, $z_0 \in H_0$ and $z_1 \in H_1$ so that $s_0 x_0 = \hat{\lambda}_0(z_0)$ and $s_1 x_1 = \hat{\lambda}_1(z_1)$. We have $s_0 F^{-*}(x_0) = \hat{\lambda}_1(F(z_0))$. Use consecutively the definition of $\lambda$ and the definition of pullback to obtain

$$0 = \lambda\left(\begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \begin{pmatrix} y_0 \\ F^{-*}(y_0) \end{pmatrix}\right) = y_0\left(\frac{z_0}{s_0}\right) - F^{-*}(y_0)\left(\frac{z_1}{s_1}\right) = y_0\left(\frac{z_0}{s_0} - F^{-1}\left(\frac{z_1}{s_1}\right)\right).$$

Since this equation holds for all $y_0$, we deduce that

$$\frac{z_0}{s_0} = F^{-1}\left(\frac{z_1}{s_1}\right).$$

Using that $F$ is an isometry, as well as the definitions of $s_0$, $s_1$, $z_0$ and $z_1$, we obtain the desired equation:

$$x_0 = \hat{\lambda}_0\left(\frac{z_0}{s_0}\right) = \hat{\lambda}_0\left(F^{-1}\left(\frac{z_1}{s_1}\right)\right) = F^*\left(\hat{\lambda}_1\left(\frac{z_1}{s_1}\right)\right) = F^*(x_1).$$

To conclude the proof that $\Gamma_{F-*}$ is a Lagrangian, it remains to show that it is a direct summand of $H := \ker(\partial F \circ \pi_0 - \pi_1)$. In fact, we will show that

$$\hat{\lambda}_0(H_0) \oplus \Gamma_{F-*} = H.$$

First we establish the inclusion $\hat{\lambda}_0(H_0) + \Gamma_{F-*} \subseteq H$. To do so, we must prove that the sum of arbitrary elements $(\hat{\lambda}_0(z_0), 0) \in \hat{\lambda}_0(H_0)$ and $(y, F^{-*}(y)) \in \Gamma_{F-*}$ belongs to $H$. In other words, we must show that $\partial F[\hat{\lambda}_0(z_0) + y] = [F^{-*}(y)]$. This uses the fact that $F^{-*}$ induces $\partial F$ (by [Definition 2.6](#)) and the fact that, since $F$ is an isometry, $F^{-*} \circ \hat{\lambda}_0(z_0) = \hat{\lambda}_1(F(z_0))$ belongs to $\mathrm{im}(\hat{\lambda}_1)$ and so the class $[F^{-*} \circ \hat{\lambda}_0(z_0)]$ must vanish in the quotient $\mathrm{coker}(\hat{\lambda}_1)$:

$$\partial F[\hat{\lambda}_0(z_0) + y] = [F^{-*}\hat{\lambda}_0(z_0)] + [F^{-*}(y)] = [F^{-*}(y)].$$

Next, we prove the reverse inclusion, namely that $H \subseteq \hat{\lambda}_0(H_0) + \Gamma_{F-*}$. To show this, we must write $(x, y) \in H$ as a sum of an element in $\hat{\lambda}_0(H_0)$ with an element in $\Gamma_{F-*}$. Consider the decomposition

$$(x, y) = (x - F^*(y), 0) + (F^*(y), y).$$

The second term certainly lies in $\Gamma_{F-*}$, so we need to argue that the first term lies in $\hat{\lambda}_0(H_0)$. Since $(x, y) \in H$, we know that $\partial F[x] = [y]$ in $\mathrm{coker}(\hat{\lambda}_1)$. It follows that

$F^{-*}(x) - y \in H_1^*$ belongs to $\mathrm{im}(\widehat{\lambda}_1)$. Since $F$ is an isometry, this is equivalent to saying that $x - F^*(y) \in \mathrm{im}(\widehat{\lambda}_0)$, as desired.

Having proved that $\widehat{\lambda}_0(H_0) + \Gamma_{F^{-*}} = H$, it remains to check that $\widehat{\lambda}_0(H_0) \cap \Gamma_{F^{-*}} = 0$ (so that the sum is direct): if $(\widehat{\lambda}_0(x_0), 0) = (x, F^{-*}(x))$ belongs to this intersection, then $F^{-*}(x) = 0$ and therefore $0 = x = \widehat{\lambda}_0(x_0)$. This proves Proposition 2.13.  □

# 3  The intersection form of the union of two 4–manifolds

We start our study of 4–manifolds with infinite cyclic fundamental group and of their $\Lambda$–intersection forms. First, recall the following definition from the introduction.

**Definition 3.1**  A 4–manifold $M$ has *ribbon boundary* if $\partial M$ is nonempty and path-connected, and the map $\pi_1(\partial M) \to \pi_1(M)$ induced by the inclusion is surjective.

Here is a summary of Section 3. In Section 3.1, we show that if $M$ is a 4–manifold with infinite cyclic fundamental group whose boundary is ribbon and has torsion Alexander module, then the boundary linking form $\partial\lambda_M$ of $\lambda_M$ is isometric to minus the Blanchfield form of $\partial M$. In Section 3.2, we describe when homeomorphisms of 3–manifolds induce isometries of the Blanchfield pairing. In Section 3.3 we show that the $\Lambda$–intersection form of a union of two manifolds can be expressed using the algebraic union from Section 2.3. In Section 3.4, we introduce the notion of a compatible pair and prove Theorem 1.12 from the introduction.

## 3.1  The boundary of the intersection form

If a 4–manifold $M$ with $\pi_1(M) = \mathbb{Z}$ has ribbon boundary and $\Lambda$–torsion Alexander module, then $\partial M$ is endowed with a sesquilinear Hermitian nonsingular linking form over $\Lambda$, namely the Blanchfield form

$$\mathrm{Bl}_{\partial M} \colon H_1(\partial M; \Lambda) \times H_1(\partial M; \Lambda) \to Q/\Lambda.$$

This subsection establishes that the boundary linking form of the $\Lambda$–intersection form of $M$ is isometric to minus the Blanchfield form of $\partial M$. Since results of this form are known (see eg [4, Theorem 2.6]), we only outline the argument so as to fix some notation for later use.

We collect some homological facts about 4–manifolds with infinite cyclic fundamental group.

**Lemma 3.2** *Let $M$ be a 4–manifold with $\pi_1(M) = \mathbb{Z}$. When $\partial M$ is nonempty, we assume it is ribbon and has torsion Alexander module $H_1(\partial M; \Lambda)$. The following assertions hold:*

(i) *$H_0(M; \Lambda) \cong \mathbb{Z}$ and, if $\partial M \neq \varnothing$, then $H_2(\partial M; \Lambda) \cong \mathbb{Z}$.*

(ii) *The $\Lambda$–modules $H_1(M; \Lambda)$, $H_1(M, \partial M; \Lambda)$ and $H_3(M; \Lambda)$ all vanish.*

(iii) *The $\Lambda$–modules $H_2(M; \Lambda)$ and $H_2(M, \partial M; \Lambda)$ are free.*

(iv) *When $\partial M = \varnothing$, the $\Lambda$–intersection form $\lambda_M$ is nonsingular, whereas for $\partial M \neq \varnothing$, $\lambda_M$ is nondegenerate, and any matrix representing it presents the Alexander module of $\partial M$.*

**Proof** Throughout this proof we use that since the manifold has fundamental group $\mathbb{Z}$, homology with $\Lambda$–coefficients can be computed as the homology of the universal cover. Since universal covers are 1–connected, we immediately deduce that $H_0(M; \Lambda) \cong \mathbb{Z}$ and $H_1(M; \Lambda) = 0$. If the boundary is nonempty, the long exact sequence of the pair shows that $H_1(M, \partial M; \Lambda) = 0$; here we used that $\partial M$ is connected, as assumed in Definition 3.1. For later use, we also note that $H_0(M, \partial M; \Lambda) = 0$, also owing to the fact that $\partial M$ is connected. As $H_1(\partial M; \Lambda)$ is torsion, we have that $\mathrm{Hom}_\Lambda(H_1(\partial M; \Lambda), \Lambda) = 0$, and thus $H_2(\partial M; \Lambda) \cong \mathrm{Ext}^1_\Lambda(H_0(\partial M; \Lambda), \Lambda) \cong \mathbb{Z}$ by duality and the universal coefficient spectral sequence [39, Theorem 2.3]. Since $H_i(M, \partial M; \Lambda) = 0$ for $i = 0, 1$, duality and the universal coefficient spectral sequence (UCSS for short) also show that $H_3(M; \Lambda) = H^1(M, \partial M; \Lambda) = 0$. Thus we have proved the first two assertions; we now prove the third. Poincaré duality and the UCSS imply that

$$H_2(M; \Lambda) \cong H^2(M, \partial M; \Lambda) \cong \overline{\mathrm{Hom}_\Lambda(H_2(M, \partial M; \Lambda), \Lambda)},$$
$$H_2(M, \partial M; \Lambda) \cong H^2(M; \Lambda) \cong \overline{\mathrm{Hom}_\Lambda(H_2(M; \Lambda), \Lambda)}.$$

Since the dual of a finitely generated $\Lambda$–module is free [3, Lemma 2.1], we deduce that these second homology modules are free over $\Lambda$. This proves the third assertion and also establishes the fourth, namely that in the closed case, the $\Lambda$–intersection form is nonsingular. Finally, when the boundary is nonempty, the map

$$\mathbb{Z} \cong H_2(\partial M; \Lambda) \to H_2(M; \Lambda)$$

is the zero map (since $H_2(M; \Lambda)$ is free) and therefore the intersection form is nondegenerate; it presents $H_1(\partial M; \Lambda)$ because we established that $H_1(M; \Lambda) = 0$. $\square$

In the case with nonempty connected boundary, we fix bases for the free $\Lambda$–modules $H_2(M; \Lambda)$ and $H_2(M, \partial M; \Lambda)$.

**Remark 3.3** Let $M$ be a 4–manifold with $\pi_1(M) = \mathbb{Z}$ whose boundary is ribbon and has torsion Alexander module. Fix a basis $\boldsymbol{b}$ for $H := H_2(M; \Lambda)$, endow the dual $H^* := \overline{\mathrm{Hom}_\Lambda(H, \Lambda)}$ with the dual basis $\boldsymbol{b}^*$, and equip $H_2(M, \partial M; \Lambda)$ with the basis $\mathrm{PD} \circ \mathrm{ev}^{-1}(\boldsymbol{b}^*)$ coming from the isomorphisms $\mathrm{ev} \colon H^2(M; \Lambda) \to H^*$ and $\mathrm{PD} \colon H^2(M; \Lambda) \to H_2(M, \partial M; \Lambda)$. A short computation shows that if the matrix $A$ is defined as $A_{ij} := \lambda_M(b_i, b_j)$, then $A^T = \overline{A}$ is a matrix for $\widehat{\lambda}_M \colon H \to H^*$; recall [Remark 2.1](). The same conclusion holds for the map $H \to H_2(M, \partial M; \Lambda)$ induced by the inclusion [9, Section 5.2]. As mentioned in the fourth item of [Lemma 3.2](), the connecting homomorphism $\delta$ in the long exact sequence of the pair $(M, \partial M)$, together with Poincaré duality and the evaluation map, determines a map

$$D_M \colon \mathrm{coker}(\widehat{\lambda}_M) = \mathrm{coker}(A^T) \xrightarrow{\cong} H_1(\partial M; \Lambda), \quad [x] \mapsto \delta \circ \mathrm{PD} \circ \mathrm{ev}^{-1}(x),$$

where $x \in H_2(M; \Lambda)^*$. The map $D_M$ is well-defined and an isomorphism. That it is well-defined follows from a diagram chase in the next diagram, in which all homology has $\Lambda$ coefficients. The left two vertical maps are isomorphisms, and therefore by the five lemma so is $D_M$. To help the reader to parse the basis choices above, we also indicate the matrices representing the maps on the left of the diagram:

(1)
$$
\begin{array}{ccccccccc}
0 & \longrightarrow & H_2(M) & \xrightarrow[A^T]{\widehat{\lambda}_M} & H_2(M)^* & \longrightarrow & \mathrm{coker}(\widehat{\lambda}_M) & \longrightarrow & 0 \\
 & & \Big\downarrow{\scriptstyle I}{\,\mathrm{Id}} & & \Big\downarrow{\scriptstyle I}{\,\mathrm{PD} \circ \mathrm{ev}^{-1}} & & \Big\downarrow{D_M} & & \\
0 & \longrightarrow & H_2(M) & \xrightarrow[A^T]{j} & H_2(M, \partial M) & \xrightarrow{\delta} & H_1(\partial M) & \longrightarrow & 0
\end{array}
$$

Next, we briefly recall the definition of the Blanchfield pairing, referring to [9; 18] as references in which the conventions are identical to ours.

**Definition 3.4** Let $N$ be a closed 3–manifold with an epimorphism $\varphi \colon \pi_1(N) \twoheadrightarrow \mathbb{Z}$ such that the resulting Alexander module is $\Lambda$–torsion. The adjoint of the Blanchfield pairing is defined by the composition

$$H_1(N; \Lambda) \xrightarrow[\mathrm{PD}^{-1}]{\cong} H^2(N; \Lambda) \xrightarrow[\mathrm{BS}^{-1}]{\cong} H^1(N; Q/\Lambda) \xrightarrow[\mathrm{ev}]{\cong} \overline{\mathrm{Hom}_\Lambda(H_1(N; \Lambda), Q/\Lambda)}$$

of the inverse of Poincaré duality, the inverse of a Bockstein homomorphism, and the evaluation map. The Blanchfield pairing is a nonsingular linking form, so in particular it is sesquilinear and Hermitian.

Now we can prove the main result of this subsection.

**Proposition 3.5** *Let $M$ be a 4–manifold with $\pi_1(M) = \mathbb{Z}$ whose boundary is ribbon and has torsion Alexander module. The map $D_M$ defined in Remark 3.3 induces an isometry*

$$D_M : \partial(H_2(M; \Lambda), \lambda_M) = (\operatorname{coker}(\widehat{\lambda}_M), \partial\lambda_M) \xrightarrow{\cong} (H_1(\partial M; \Lambda), -\operatorname{Bl}_{\partial M}).$$

**Proof** Choose bases for $H_2(M; \Lambda)$ and $H_2(M, \partial M; \Lambda)$ as in Remark 3.3, and use the notation of that remark. The same argument as in [9, Section 5] shows that the following diagram (with exact columns) commutes:

$$
\begin{array}{ccc}
0 & & 0 \\
\downarrow & & \downarrow \\
H_2(M; \Lambda) \times H_2(M; \Lambda) & \xrightarrow{(x,y)\mapsto -x^T A \bar{y}} & \Lambda \\
\downarrow{\scriptstyle A^T \times A^T} & & \downarrow \\
H_2(M, \partial M; \Lambda) \times H_2(M, \partial M; \Lambda) & \xrightarrow{(x,y)\mapsto -x^T A^{-1} \bar{y}} & Q \\
\downarrow{\scriptstyle \delta \times \delta} & & \downarrow \\
H_1(\partial M; \Lambda) \times H_1(\partial M; \Lambda) & \xrightarrow{\operatorname{Bl}_{\partial M}} & Q/\Lambda \\
\downarrow & & \downarrow \\
0 & & 0
\end{array}
$$

In particular, the isomorphism

$$D_M : \operatorname{coker}(\widehat{\lambda}_M) = \operatorname{coker}(A^T) \xrightarrow{\cong} H_1(\partial M; \Lambda)$$

induced by $\delta$ and discussed in Remark 3.3 is an isometry between the linking forms $(\operatorname{coker}(\widehat{\lambda}_M), \partial\lambda_M)$ and $(H_1(\partial M; \Lambda), -\operatorname{Bl}_{\partial M})$. Here we are using the characterisation of $\partial\lambda_M$ from Remark 2.4. $\qquad\square$

The next remark records that the presence of a minus sign in Proposition 3.5 is immaterial once one passes to isometry groups, and defines the bijection $D^{\#}$ which identifies the two isometry groups.

**Remark 3.6** Given linking forms $(H_0, \beta_0)$ and $(H_1, \beta_1)$, observe that canonically we have $\operatorname{Iso}(\beta_0, \beta_1) = \operatorname{Iso}(-\beta_0, -\beta_1)$; the presence of the minus sign has no effect. In particular if $M_0$ and $M_1$ are 4–manifolds with $\pi_1(M_i) = \mathbb{Z}$ for $i = 0, 1$, whose boundaries are ribbon and have torsion Alexander modules, then we can use the maps $D_i := D_{M_i}$ from Proposition 3.5 to obtain an identification

$$D^{\#} : \operatorname{Iso}(\partial\lambda_{M_0}, \partial\lambda_{M_1}) \xrightarrow{\cong} \operatorname{Iso}(\operatorname{Bl}_{\partial M_0}, \operatorname{Bl}_{\partial M_1}), \quad \widetilde{f} \mapsto D_1 \circ \widetilde{f} \circ D_0^{-1}.$$

## 3.2 Homeomorphisms and isometries

In this short subsection, we provide a condition for a homeomorphism between 3–manifolds to induce an isometry of Blanchfield forms.

**Proposition 3.7** *Suppose that $Y_0^3$, $Y_1^3$ are 3–manifolds equipped with epimorphisms $\varphi_i \colon \pi_1(Y_i) \twoheadrightarrow \mathbb{Z}$ and assume that the resulting Alexander modules $H_1(Y_i; \Lambda)$ are $\Lambda$–torsion for $i = 0, 1$. If an orientation-preserving homotopy equivalence $f \colon Y_0^3 \to Y_1^3$ satisfies $\varphi_1 \circ f_* = \varphi_0$ on $\pi_1(Y_0)$, then it induces an isometry between the Blanchfield forms:*

$$f_* \colon (H_1(Y_0; \Lambda), \mathrm{Bl}_{Y_0}) \to (H_1(Y_1; \Lambda), \mathrm{Bl}_{Y_1}).$$

**Proof** Since we assumed that $\varphi_1 \circ f_* = \varphi_0$, the homotopy equivalence $f$ induces $\Lambda$–isomorphisms $f_* \colon H_*(Y_0; \Lambda) \to H_*(Y_1; \Lambda)$ (and similarly on cohomology). The fact that $f$ has degree one, together with the naturality of the UCSS and the Bockstein homomorphism, ensures that $f_*$ intertwines the Blanchfield pairings, concluding the proof of the proposition. $\qquad\square$

For pairs $(Y_0, \varphi_0)$ and $(Y_1, \varphi_1)$ as in the statement of Proposition 3.7, we consider those orientation-preserving homeomorphisms that intertwine the $\varphi_i$:

$$\mathrm{Homeo}_\varphi^+(Y_0, Y_1) = \{ f \in \mathrm{Homeo}^+(Y_0, Y_1) \mid \varphi_1 \circ f_* = \varphi_0 \}.$$

In the case where $(Y_0, \varphi_0) = (Y_1, \varphi_1)$, we simply write $\mathrm{Homeo}_\varphi^+(Y_0)$ for the resulting group.

## 3.3 Algebraic unions and topological unions

In this subsection, we show that under favourable conditions, the $\Lambda$–intersection form of a union of two 4–manifolds can be expressed using the algebraic union of Hermitian forms from Section 2.3. For that however, given two manifolds $M_0$, $M_1$ with $\pi_1(M_i) = \mathbb{Z}$, we need to verify that $\pi_1(M_0 \cup_f -M_1) = \mathbb{Z}$. Here and from now on, when $\pi_i(M_i) = \mathbb{Z}$, we take $\varphi_i \colon \pi_1(\partial M_i) \to \pi_1(M_i) = \mathbb{Z}$ to be the map induced by the inclusion. Thus, when we write $\mathrm{Homeo}_\varphi^+(\partial M_0, \partial M_1)$, it is with respect to these inclusion induced maps.

**Lemma 3.8** *If $f \in \mathrm{Homeo}_\varphi^+(\partial M_0, \partial M_1)$ is a homeomorphism where $M_0$, $M_1$ are 4–manifolds with ribbon boundaries and $\pi_1(M_i) = \mathbb{Z}$, then $\pi_1(M_0 \cup_f -M_1) = \mathbb{Z}$.*

**Proof** In order to apply van Kampen's theorem, for $i = 0, 1$ we write $\pi_1(M_i) = \mathbb{Z}\langle t_i \rangle$, so that $\pi_1(M_0 \cup_f -M_1)$ is generated by $t_0, t_1$. Consider the commutative diagram

$$
\begin{array}{ccc}
\pi_1(\partial M_0) & \xrightarrow{\quad f_*, \cong \quad} & \pi_1(\partial M_1) \\
\downarrow{\iota_0} \quad {\scriptstyle \varphi_0} & \qquad {\scriptstyle \varphi_1} & \quad \downarrow{\iota_1} \\
\pi_1(M_0) = \mathbb{Z}\langle t_0 \rangle \xrightarrow[\psi_0, \cong]{} \mathbb{Z}\langle t \rangle \xleftarrow[\psi_1, \cong]{} \mathbb{Z}\langle t_1 \rangle = \pi_1(M_1)
\end{array}
$$

Here, for $j = 0, 1$, $\iota_j : \pi_1(\partial M_j) \to \pi_1(M_j)$ is the inclusion induced map and $\psi_j : t_j \mapsto t$. In the group $\pi_1(M_0 \cup_f -M_1)$, the relations identify $\iota_1 f_*(g)$ with $\iota_0(g)$ for every $g \in \pi_1(\partial M_0)$. Since $\iota_0$ is surjective, there is a $\mu \in \pi_1(\partial M_0)$ with $\iota_0(\mu) = t_0$. We deduce that in $\pi_1(M_0 \cup_f -M_1)$, the relations identify $t_0 = \varphi_0(\mu)$ with $\iota_1(f_*(\mu)) = t_1^n$ for some $n$. Under the identification given by the $\psi_i$ and using that $\varphi_0(\mu) = \varphi_1(f_*(\mu))$, we deduce that $n = 1$. Thus $\pi_1(M_0 \cup_f -M_1)$ is cyclic. For every $g \in \pi_1(\partial M_0)$, we have $\iota_0(g) = t_0^m$ for some $m$, and the same reasoning as above shows that $\iota_0(g) = t_0^m$ is identified with $\iota_1(f_*(g)) = t_1^m$ in $\pi_1(M_0 \cup_f -M_1)$. It follows that $\pi_1(M_0 \cup_f -M_1)$ is isomorphic to $\langle t_0, t_1 \mid t_0 = t_1 \rangle \cong \mathbb{Z}$. $\qquad\square$

Next, we prove the main result of this section, after setting up some notation. Continuing with the setting of Lemma 3.8, we write $M := M_0 \cup_f -M_1$. Note that $\hat{\lambda}_M : H_2(M; \Lambda) \to H_2(M; \Lambda)^*$ is an isomorphism because $\partial M = \varnothing$; recall the fourth item of Lemma 3.2. This endows $H_2(M; \Lambda)^*$ with a Hermitian form, which we shall denote by $\lambda_M^{-1}$.

For $i = 0, 1$, we set $\hat{\lambda}_i := \hat{\lambda}_{M_i}$, and as above we write $D_i := D_{M_i}$. Combining Propositions 3.5 and 3.7, we have an isometry

$$
\tilde{f}_* : \operatorname{coker}(\hat{\lambda}_0) \xrightarrow{D_0} H_1(\partial M_0; \Lambda) \xrightarrow{f_*} H_1(\partial M_1; \Lambda) \xrightarrow{D_1^{-1}} \operatorname{coker}(\hat{\lambda}_1),
$$

which satisfies $D^{\#}(\tilde{f}_*) = f_*$.

**Proposition 3.9** *Let $M_0, M_1$ be two 4–manifolds with $\pi_1(M_i) = \mathbb{Z}$ for $i = 0, 1$, whose boundaries are ribbon and have torsion Alexander modules. Given a homeomorphism $f \in \operatorname{Homeo}_\varphi^+(\partial M_0, \partial M_1)$, set $M := M_0 \cup_f -M_1$, and let $\iota_j : M_i \to M$ be the inclusion. The map*

$$
\begin{pmatrix} \iota_0^* \\ \iota_1^* \end{pmatrix} : H_2(M; \Lambda)^* \to H_2(M_0; \Lambda)^* \oplus H_2(M_1; \Lambda)^*
$$

*induces an isometry*

$$
(2) \qquad (H_2(M; \Lambda)^*, \lambda_M^{-1}) \xrightarrow{\cong} (H_2(M_0; \Lambda) \cup_{\tilde{f}_*} H_2(M_1; \Lambda), \lambda_{M_0} \cup_{\tilde{f}_*} -\lambda_{M_1}).
$$

Furthermore, the map $\widehat{\lambda}_M$ induces an isometry $(H_2(M;\Lambda),\lambda_M) \to (H_2(M;\Lambda)^*,\lambda_M^{-1})$, which, using the identification from (2), makes the following diagram commute:

$$
(3) \quad
\begin{array}{ccc}
(H_2(M_0;\Lambda),\lambda_{M_0}) \overset{\iota_0}{\hookrightarrow} (H_2(M;\Lambda),\lambda_M) \overset{\iota_1}{\hookleftarrow} (H_2(M_1;\Lambda),-\lambda_{M_1}) \\
\searrow^{\widehat{\lambda}_{M_0}} \qquad \cong\downarrow^{\widehat{\lambda}_M} \qquad \swarrow^{\widehat{\lambda}_{M_1}} \\
(H_2(M_0;\Lambda)\cup_{\widetilde{f}_*} H_2(M_1;\Lambda),\lambda_{M_0}\cup_{\widetilde{f}_*}-\lambda_{M_1})
\end{array}
$$

In particular, $H_2(M_0;\Lambda)\cup_{\widetilde{f}_*} H_2(M_1;\Lambda)$ is a free $\Lambda$–module and $\lambda_{M_0}\cup_{\widetilde{f}_*}-\lambda_{M_1}$ is nonsingular.

**Proof** Lemma 3.8 implies that $\pi_1(M)=\mathbb{Z}$. Consider the following commutative diagram in which the first and third rows are exact and $\Lambda$ coefficients are understood:

$$
(4)
$$

$$
\begin{array}{c}
0 \longrightarrow H_2(M_0)\oplus H_2(M_1) \overset{(\iota_0\ \iota_1)}{\longrightarrow} H_2(M) \longrightarrow H_1(\partial M_1) \longrightarrow 0 \\
\downarrow^{j_0\oplus j_1} \qquad\qquad\qquad \downarrow \\
H_2(M_0,\partial M_0)\oplus H_2(M_1,\partial M_1) \qquad\qquad \downarrow^{\mathrm{PD}_M^{-1}} \\
\downarrow^{\mathrm{PD}_{M_0}^{-1}\oplus-\mathrm{PD}_{M_1}^{-1}} \qquad\qquad \downarrow^{\binom{\iota_0^*}{\iota_1^*}} \\
H^2(\partial M_1) \overset{f^{-*}\mathrm{incl}_0^*-\mathrm{incl}_1^*}{\longleftarrow} H^2(M_0)\oplus H^2(M_1) \longleftarrow H^2(M) \longleftarrow 0 \\
\cong\downarrow^{\mathrm{ev}\oplus\mathrm{ev}} \qquad\qquad \cong\downarrow^{\mathrm{ev}} \\
H_2(M_0)^*\oplus H_2(M_1)^* \overset{\binom{\iota_0^*}{\iota_1^*}}{\longleftarrow} H_2(M)^*
\end{array}
$$

We justify the three zeros that appear. In the first row, the rightmost zero comes from the fact that $H_1(M_i;\Lambda)=0$ for $i=0,1$ by the second item of Lemma 3.2. For the leftmost zero of the first row, the first and third items of Lemma 3.2 respectively imply that $H_2(\partial M_i;\Lambda)=\mathbb{Z}$ is $\Lambda$–torsion and that $H_2(M_i;\Lambda)$ is free. It follows that the inclusion induced homomorphism $H_2(\partial M_i;\Lambda)\to H_2(M_i;\Lambda)$ is zero. The same argument explains the appearance of the zero in the third row: $H^2(M;\Lambda)$ is free and $H^1(\partial M_1;\Lambda)$ is $\Lambda$–torsion. The commutativity of the middle square follows by applying [7, Lemma 8.2] with $K=M_0$ and $L=M_1$, as well as applying excision to the pairs $(M,M\setminus M_i)$.

Our next aim is to simplify the diagram in (4). We write $\pi_i\colon H_2(M_i;\Lambda)^*\to\mathrm{coker}(\widehat{\lambda}_i)$ for the canonical projection.

**Claim 1** *The isometry* $\widetilde{f}_* \colon \operatorname{coker}(\widehat{\lambda}_0) \overset{\cong}{\Longrightarrow} \operatorname{coker}(\widehat{\lambda}_1)$ *fits in the diagram*

(5)

$$
\begin{array}{ccccccc}
0 & \longrightarrow & \begin{array}{c} H_2(M_0;\Lambda) \\ \oplus \\ H_2(M_1;\Lambda) \end{array} & \xrightarrow{(\iota_0\ \iota_1)} & H_2(M;\Lambda) & \longrightarrow & H_1(\partial M_1;\Lambda) \\
& & \Big\downarrow{\widehat{\lambda}_0\oplus-\widehat{\lambda}_1} & & \Big\downarrow\cong{\widehat{\lambda}_M} & & \\
\operatorname{coker}(\widehat{\lambda}_1) & \xleftarrow{\widetilde{f}_*\pi_0-\pi_1} & \begin{array}{c} H_2(M_0;\Lambda)^* \\ \oplus \\ H_2(M_1;\Lambda)^* \end{array} & \xleftarrow{\left(\begin{smallmatrix}\iota_0^*\\ \iota_1^*\end{smallmatrix}\right)} & H_2(M;\Lambda)^* & \longleftarrow & 0
\end{array}
$$

*where the rows are exact and the square is commutative.*

**Proof** The two vertical compositions in (4) are by definition $\widehat{\lambda}_0 \oplus -\widehat{\lambda}_1$ and $\widehat{\lambda}_M$, respectively. Thus the exactness of the top row and the commutativity of the square follow from (4). It therefore remains to establish the exactness of the bottom row in (5). A second look at (4) shows a similar exact sequence with cohomology instead of duals and $H^2(\partial M_1;\Lambda)$ instead of $\operatorname{coker}(\widehat{\lambda}_1)$. We shall produce the commuting square

(6)

$$
\begin{array}{ccc}
H^2(\partial M_1;\Lambda) & \xleftarrow{f^{-*}\operatorname{incl}_0^*-\operatorname{incl}_1^*} & H^2(M_0;\Lambda)\oplus H^2(M_1;\Lambda) \\
\cong\Big\downarrow{G_1} & & \cong\Big\downarrow{\mathrm{ev}\oplus\mathrm{ev}} \\
\operatorname{coker}(\widehat{\lambda}_1) & \xleftarrow{\widetilde{f}_*\pi_0-\pi_1} & H_2(M_0;\Lambda)^*\oplus H_2(M_1;\Lambda)^*
\end{array}
$$

that can be appended to the bottom left of (4). The claim will then follow immediately.

We define isomorphisms $G_i$, for $i = 0, 1$, as the composition

$$
G_i := D_i^{-1} \circ \mathrm{PD} \colon H^2(\partial M_i;\Lambda) \overset{\cong}{\Longrightarrow} H_1(\partial M_i;\Lambda) \overset{\cong}{\Longrightarrow} \operatorname{coker}(\widehat{\lambda}_i).
$$

The map $G_i$ agrees with the rightmost vertical downwards composition (inverting $D_i$) in the next diagram. Again $\Lambda$ coefficients are understood, and the rows are exact:

(7)

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & H^2(M_i,\partial M_i) & \xrightarrow{j_i^*} & H^2(M_i) & \xrightarrow{\operatorname{incl}_i^*} & H^2(\partial M_i) & \longrightarrow & 0 \\
& & \mathrm{PD}\Big\downarrow\cong & & \mathrm{PD}\Big\downarrow\cong & & \mathrm{PD}\Big\downarrow\cong & & \\
0 & \longrightarrow & H_2(M_i) & \xrightarrow{j_i} & H_2(M_i,\partial M_i) & \xrightarrow{\delta_i} & H_1(\partial M_i) & \longrightarrow & 0 \\
& & \Big\Vert\,{\mathrm{Id}} & & \cong\Big\uparrow{\mathrm{PD}\circ\mathrm{ev}^{-1}} & & \cong\Big\uparrow{D_i} & & \\
0 & \longrightarrow & H_2(M_i) & \xrightarrow{\widehat{\lambda}_i} & H_2(M_i)^* & \xrightarrow{\pi_i} & \operatorname{coker}(\widehat{\lambda}_i) & \longrightarrow & 0
\end{array}
$$

The diagram commutes by the definition of $\widehat{\lambda}_i$ as $\text{ev} \circ \text{PD}^{-1} \circ j_i$, and using standard naturality of Poincaré duality; for the bottom right square, one uses the definition of $D_i$ from Remark 3.3. Via $G_i$ and ev, the concatenation of the two squares on the right identifies the maps $\pi_i$ and $\text{incl}_i^*$. That is, for $i = 0, 1$,

$$(8) \qquad G_i \circ \text{incl}_i^* = \pi_i \circ \text{ev} \colon H^2(M_i; \Lambda) \to \text{coker}(\widehat{\lambda}_i).$$

This shows that the bottom left square in the following diagram commutes. This diagram again has $\Lambda$ coefficients and vertical maps isomorphisms:

$$(9)$$

$$
\begin{array}{ccccc}
& & H_1(\partial M_0; \Lambda) & \xrightarrow{\ f_* \ }_{\cong} & H_1(\partial M_1; \Lambda) \\
& & \cong \downarrow {\scriptstyle \text{PD}^{-1}} & & \cong \downarrow {\scriptstyle \text{PD}^{-1}} \\
H^2(M_0) & \xrightarrow{\ \text{incl}_0^* \ } & H^2(\partial M_0) & \xrightarrow[\cong]{\ f^{-*} \ } & H^2(\partial M_1) \\
{\scriptstyle \text{ev}} \downarrow {\scriptstyle \cong} & & {\scriptstyle G_0} \downarrow {\scriptstyle \cong} & & {\scriptstyle G_1} \downarrow {\scriptstyle \cong} \\
H_2(M_0)^* & \xrightarrow{\ \pi_0 \ } & \text{coker}(\widehat{\lambda}_0) & \xrightarrow[\cong]{\ \widetilde{f}_* \ } & \text{coker}(\widehat{\lambda}_1)
\end{array}
$$

Commutativity of the top right square follows from naturality of Poincaré duality and the fact that all four maps are isomorphisms. Using this and that $D_i = \text{PD} \circ G_i^{-1}$ for $i = 0, 1$, starting with the definition of $\widetilde{f}_* \colon \text{coker}(\widehat{\lambda}_0) \to \text{coker}(\widehat{\lambda}_1)$ we obtain

$$\widetilde{f}_* = D_1^{-1} \circ f_* \circ D_0 = G_1 \circ \text{PD}^{-1} \circ f_* \circ \text{PD} \circ G_0^{-1} = G_1 \circ f^{-*} \circ G_0^{-1}.$$

In other words, the bottom right square of (9) commutes.

Therefore the bottom of diagram (9) yields $G_1 \circ f^{-*} \circ \text{incl}_0^* = \widetilde{f}_* \circ \pi_0 \circ \text{ev}$. Subtracting (8) with $i = 1$, we obtain the desired commutative square (6). The diagram in (5) then follows from the combination of (4) and (6), completing the proof of the claim. $\qquad \square$

Now we write $T := H_1(\partial M_1; \Lambda)$, as well as $H_i := H_2(M_i; \Lambda)$ for $i = 0, 1$. We also set $V := H_2(M; \Lambda)^*$ and identify $H_2(M; \Lambda)$ with the double dual $H_2(M; \Lambda)^{**} = V^*$ via the evaluation isomorphism ev. With this notation, the diagram from Claim 1 leads to the following commutative diagram in which both rows are exact:

$$(10)$$

$$
\begin{array}{ccccccc}
0 & \longrightarrow & H_0 \oplus H_1 & \xrightarrow{\ (\text{ev ev}) \ } & V^* & \longrightarrow & T \\
& & \downarrow {\scriptstyle \widehat{\lambda}_0 \oplus -\widehat{\lambda}_1} & & \cong \downarrow {\scriptstyle \widehat{\lambda}_M} & & \\
\text{coker}(\widehat{\lambda}_1) & \xleftarrow{\ \widetilde{f}_* \pi_0 - \pi_1 \ } & H_0^* \oplus H_1^* & \longleftarrow & V & \longleftarrow & 0
\end{array}
$$

Identify $V = H_2(M; \Lambda)^*$ with $\ker(\widetilde{f}_* \pi_0 - \pi_1)$ and view it as a subspace of $H_0^* \oplus H_1^*$. The diagram in (10) therefore yields the equation

(11)
$$\begin{pmatrix} \widehat{\lambda}_0 & 0 \\ 0 & -\widehat{\lambda}_1 \end{pmatrix} = \widehat{\lambda}_M \circ (\text{ev} \ \ \text{ev}).$$

Since Lemma 3.2 implies that $H_2(M; \Lambda)$ is a free $\Lambda$–module and $H_1(M; \Lambda) = 0$, we may identify $H^2(M; Q)$ with $H^2(M; \Lambda) \otimes_\Lambda Q$, and similarly for

$$H^2(M_i, \partial M_i; Q) \cong H^2(M_i, \partial M_i; \Lambda) \otimes_\Lambda Q \quad \text{and} \quad H^2(M_i; Q) \cong H^2(M_i; \Lambda) \otimes_\Lambda Q$$

for $i = 0, 1$. As a consequence, we may identify the tensored up intersection pairings $\lambda_M \otimes_\Lambda \text{Id}_Q$ and $\lambda_i \otimes_\Lambda \text{Id}_Q$ with the nonsingular $Q$–valued intersections pairings on $H_2(M; Q)$ and $H_2(M_i; Q)$. Since $\widehat{\lambda}_M$ is invertible by Lemma 3.2, we deduce from (11) that

$$\widehat{\lambda}_M^{-1} = (\text{ev} \ \ \text{ev}) \circ \begin{pmatrix} \widehat{\lambda}_{0,Q}^{-1} & 0 \\ 0 & -\widehat{\lambda}_{1,Q}^{-1} \end{pmatrix}.$$

As noted in Remark 2.11, this is precisely the adjoint of $\lambda_0 \cup_{\widetilde{f}_*} -\lambda_1$.

We have thus proved that the inclusions $\iota_0$ and $\iota_1$ induce an isometry

$$(H_2(M; \Lambda)^*, \lambda_M^{-1}) \cong (H_0 \cup_{\widetilde{f}_*} H_1, \lambda_0 \cup_{\widetilde{f}_*} -\lambda_1).$$

Now, since $\lambda_M^* = \lambda_M$, a quick verification shows that the isomorphism

$$\widehat{\lambda}_M \colon H_2(M; \Lambda) \to H_2(M; \Lambda)^*$$

induces an isometry between the forms $(H_2(M; \Lambda), \lambda_M)$ and $(H_2(M; \Lambda)^*, \lambda_M^{-1})$. Indeed, writing $H := H_2(M; \Lambda)$, this follows from the following commutative diagram:

$$\begin{array}{ccc} H & \xrightarrow[\cong]{\widehat{\lambda}_M} & H^* \\ \widehat{\lambda}_M \downarrow & & \downarrow \widehat{\lambda}_M^{-1} \\ H^* & \xleftarrow[\widehat{\lambda}_M^* = \widehat{\lambda}_M]{\cong} & H \end{array}$$

The commutativity of the diagram in (3) follows from the commutativity of (10): we have shown that $(H_0 \cup_{\widetilde{f}_*} H_1, \lambda_0 \cup_{\widetilde{f}_*} -\lambda_1) \cong (H_2(M; \Lambda), \lambda_M)$ and, recalling that $(\text{ev}, \text{ev}) = (\iota_0, \iota_1)$, the diagram in (10) gives $\widehat{\lambda}_M \circ \iota_0(x_0, 0) = \widehat{\lambda}_0(x_0, 0)$, leading to the commutativity of the left-hand triangle of (3); the reasoning for the right-hand triangle is identical. $\square$

## 3.4 Compatible pairs

Throughout this section, $M_0$ and $M_1$ are 4–manifolds with $\pi_1(M_i) = \mathbb{Z}$ for $i = 0, 1$, whose boundaries are ribbon and have torsion Alexander modules. This section introduces a notion of compatibility for an isometry $F: H_2(M_0; \Lambda) \to H_2(M_1; \Lambda)$ and a homeomorphism $f: \partial M_0 \to \partial M_1$. We then prove Theorem 1.12 from the introduction.

Since the boundaries are ribbon, and we have identifications $\pi_1(M_i) = \mathbb{Z}$, the inclusions $\partial M_i \hookrightarrow M_i$ induce surjections $\varphi_i: \pi_1(M_i) \twoheadrightarrow \mathbb{Z}$ for $i = 0, 1$. Recall from Section 3.2 that we set

$$\text{Homeo}_\varphi^+(\partial M_0, \partial M_1) = \{f \in \text{Homeo}^+(\partial M_0, \partial M_1) \mid \varphi_1 \circ f_* = \varphi_0\}.$$

Proposition 3.7 ensures that a homeomorphism $f \in \text{Homeo}_\varphi^+(\partial M_0, \partial M_1)$ induces an isometry $f_* \in \text{Iso}(\text{Bl}_{\partial M_0}, \text{Bl}_{\partial M_1})$ of the Blanchfield forms. On the other hand, recall from Lemma 2.5 that an isometry $F \in \text{Iso}(\lambda_{M_0}, \lambda_{M_1})$ of the $\Lambda$–intersection forms of $M_0$ and $M_1$ induces an isometry $\partial F = F^{-*}$ of the boundary linking forms. Using Proposition 3.5, we identify these boundary linking forms with minus the corresponding Blanchfield forms, so that using Remark 3.6 we may consider $D^\#(\partial F) \in \text{Iso}(\text{Bl}_{\partial M_0}, \text{Bl}_{\partial M_1})$.

**Definition 3.10** An orientation-preserving homeomorphism $f \in \text{Homeo}_\varphi^+(\partial M_0, \partial M_1)$ is *compatible* with an isometry $F \in \text{Iso}(\lambda_{M_0}, \lambda_{M_1})$ if, using the identification $D^\#$ from Remark 3.6, we have

$$D^\#(\partial F) = f_*: (H_1(\partial M_0; \Lambda), \text{Bl}_{\partial M_0}) \to (H_1(\partial M_1; \Lambda), \text{Bl}_{\partial M_1}).$$

In this case, the pair $(f, F)$ is called a *compatible pair*.

The next proposition shows that the existence of a compatible pair is a necessary condition for a homeomorphism $f \in \text{Homeo}_\varphi^+(\partial M_0, \partial M_1)$ to extend to a homeomorphism $M_0 \to M_1$.

**Proposition 3.11** Let $f \in \text{Homeo}_\varphi^+(\partial M_0, \partial M_1)$ be a homeomorphism, and let $F \in \text{Iso}(\lambda_{M_0}, \lambda_{M_1})$ be an isometry. If $f$ extends to an orientation-preserving homeomorphism $\Phi: M_0 \xrightarrow{\cong} M_1$ that induces $F$, then $(f, F)$ is a compatible pair.

**Proof** Write $\lambda_i := \lambda_{M_i}$ and $D_i := D_{M_i}$ for $i = 0, 1$. The homeomorphism $\Phi$ induces an isomorphism $\Phi_*: \pi_1(M_0) \to \pi_1(M_1)$. By the assumption on $f$, the isomorphisms $\pi_1(M_i) = \mathbb{Z}$ are such that $\Phi_*$ intertwines them. This follows because in the next

diagram both triangles, the top square, and the large outside square commute:

$$
\begin{array}{ccc}
\pi_1(\partial M_0) & \xrightarrow{\ f_* \ } & \pi_1(\partial M_1) \\
\downarrow & & \downarrow \\
\pi_1(M_0) & \xrightarrow{\ \Phi_* \ } & \pi_1(M_1) \\
\parallel & & \parallel \\
\mathbb{Z} & ==== & \mathbb{Z}
\end{array}
$$

with $\varphi_0$ and $\varphi_1$.

It follows that the bottom square commutes too. By assumption $\Phi$ induces the maps

$$F\colon H_2(M_0;\Lambda) \xrightarrow{\ \cong\ } H_2(M_1;\Lambda) \quad \text{and} \quad F^{-*}\colon H_2(M_0;\Lambda)^* \xrightarrow{\ \cong\ } H_2(M_1;\Lambda)^*.$$

By Lemma 2.5 the map $F^{-*}$ induces an isometry

$$\partial F\colon (\operatorname{coker}(\widehat{\lambda}_0), \partial\lambda_0) \xrightarrow{\ \cong\ } (\operatorname{coker}(\widehat{\lambda}_1), \partial\lambda_1).$$

We assert that $\partial F$ fits into a commuting diagram of isometries

(12)
$$
\begin{array}{ccc}
\operatorname{coker}(\widehat{\lambda}_0) & \xrightarrow[\cong]{\ \partial F\ } & \operatorname{coker}(\widehat{\lambda}_1) \\
\cong \downarrow{\scriptstyle D_0} & & \cong \downarrow{\scriptstyle D_1} \\
H_1(\partial M_0;\Lambda) & \xrightarrow[\cong]{\ f_*\ } & H_1(\partial M_1;\Lambda)
\end{array}
$$

The maps $D_i$ are isometries by Proposition 3.5, and the map $f_*$ in the bottom row is an isometry by Proposition 3.7. By definition of $D^\#$, commutativity of the diagram (12) is equivalent to $D^\#(\partial F) = f_*$, ie to the fact that $(f, F)$ is a compatible pair. To show that (12) commutes, we use the next diagram, in which all homology groups have $\Lambda$ coefficients and the rightmost horizontal maps are all surjections:



The map $\partial F\colon \operatorname{coker}(\widehat{\lambda}_0) \to \operatorname{coker}(\widehat{\lambda}_1)$ is by definition the map induced by the upper part of the diagram. The squares in the leftmost cube commute either trivially, by

the definition of $\hat{\lambda}_i$, or by naturality of Poincaré duality. The potentially contentious point is the latter justification: naturality in fact says that $F_* \circ \mathrm{PD}_{M_0} \circ F^* = \mathrm{PD}_{M_1}$. But since $F^*$ is an isomorphism, with inverse $F^{-*}$, it follows that the middle vertical square commutes as shown. A straightforward diagram chase now shows, as desired, that the rightmost vertical square, which is equivalent to (12), also commutes. This completes the proof that $(f, F)$ is a compatible pair. $\qquad\square$

The next result shows that the existence of a compatible pair $(f, F)$ imposes strong restrictions on the topology of $M := M_0 \cup_f -M_1$. This is Theorem 1.12 from the introduction.

**Theorem 3.12** *Let $(f, F)$ be a compatible pair. If $M_0$ and $M_1$ are not spin, assume that the Kirby–Siebenmann invariants satisfy $\mathrm{ks}(M_0) = \mathrm{ks}(M_1) \in \mathbb{Z}/2$. Then there is an orientation-preserving homeomorphism*

$$M := M_0 \cup_f -M_1 \cong S^1 \times S^3 \mathbin{\#} \overset{a}{\underset{i=1}{\#}} S^2 \times S^2 \mathbin{\#} \overset{b}{\underset{j=1}{\#}} S^2 \widetilde{\times} S^2$$

*for some $a, b$ with $a + b = b_2(M_0)$. If $M_0$ and $M_1$ are spin, then $b = 0$.*

*Moreover, it can be assumed that the $\Lambda$–isometry induced by this homeomorphism takes the Lagrangian $\hat{\lambda}_M^{-1}(\Gamma_{F^{-*}}) \subseteq H_2(M; \Lambda)$ to the Lagrangian that is generated by $\{[\{\mathrm{pt}\} \times S^2]\}_{i=1}^{b_2(M_0)}$.*

**Proof** Lemma 3.8 establishes that $\pi_1(M) = \mathbb{Z}$. Proposition 3.9 proves that $\lambda_M$ is isometric to $\lambda_{M_0} \cup_{\widetilde{f}_*} -\lambda_{M_1}$. We have $\lambda_{M_0} \cup_{\widetilde{f}_*} -\lambda_{M_1} = \lambda_{M_0} \cup_{\partial F} -\lambda_{M_1}$ because the pair $(f, F)$ is compatible. It follows from Proposition 2.13 that $\lambda_{M_0} \cup_{\partial F} -\lambda_{M_1}$ is metabolic with $\Gamma_{F^{-*}}$ as a Lagrangian. We deduce that $M$ is a closed 4–manifold with $\pi_1(M) = \mathbb{Z}$ and metabolic $\Lambda$–intersection form. The (nonsingular) intersection form of $M$ (which can be obtained from $\lambda_M$ by setting $t = 1$) is therefore isometric to

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{\oplus a} \oplus \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{\oplus b}$$

for some $a, b$. As $\mathrm{ks}(M_0) = \mathrm{ks}(M_1)$, by additivity of the Kirby–Siebenmann invariant [17, Theorem 8.2] under gluing it follows that $\mathrm{ks}(M) = 0$. Apart from the statement that $b = 0$ for $M_i$ spin, the theorem now follows from the classification of closed 4–manifolds with infinite cyclic fundamental group due to [15, Theorem 10.7A(2)] and [48]: every isometry between the $\Lambda$–intersection forms of two closed, oriented 4–manifolds with fundamental group $\mathbb{Z}$ can be realised by a homeomorphism.

Now assume that $M_0$ and $M_1$ are spin. To show that $b = 0$, it is sufficient to show that $\lambda_M$ is hyperbolic, as the result will then once again follow from [15, Theorem 10.7A(2)] and [48]. We already argued that $\lambda_M$ is isometric to $\lambda_{M_0} \cup_{\partial F} -\lambda_{M_1}$ because the pair $(f, F)$ is compatible. Applying Proposition 2.12 to the isometry $h = \partial F$, this latter form is isometric to $\lambda_{M_1} \cup_{\mathrm{Id}} -\lambda_{M_1}$. The identity certainly belongs to $\mathrm{Homeo}^+_\varphi(\partial M_1, \partial M_1)$ and so a second application of Proposition 3.9 ensures that $\lambda_{M_1} \cup_{\mathrm{Id}} -\lambda_{M_1}$ is isometric to $\lambda_{M_1 \cup_{\mathrm{Id}} -M_1}$. Summarising, we have the following sequence of isometries:

$$(13) \qquad \lambda_M \cong \lambda_{M_0 \cup_{\widetilde{f_*}} -M_1} \cong \lambda_{M_0} \cup_{\partial F} -\lambda_{M_1} \cong \lambda_{M_1} \cup_{\mathrm{Id}} -\lambda_{M_1} \cong \lambda_{M_1 \cup_{\mathrm{Id}} -M_1}.$$

Since $M_1$ is spin, the double $DM_1 := M_1 \cup_{\mathrm{Id}} -M_1$ is also spin. As $\pi_2(DM_1) = H_2(DM_1; \Lambda)$ is free (by the third item of Lemma 3.2 and Lemma 3.8) and $DM_1$ is spin, it follows from [26, Remark 1.7] that $\lambda_{DM_1}$ is even, meaning that $\lambda_{DM_1} = q + \overline{q}$ for some sesquilinear form $q$. Since $\lambda_{DM_1}$ and $\lambda_M$ are isometric by (13), it follows that $\lambda_M$ is even. Now since $\lambda_M$ is an even metabolic nonsingular Hermitian form on a finitely generated free $\Lambda$–module, it is hyperbolic by [34, Corollary 3.7.3] with $\Gamma_{F-*}$ as a Lagrangian. $\square$

We make the Lagrangian $\widehat{\lambda}_M^{-1}(\Gamma_{F-*}) \subseteq H_2(M; \Lambda)$ somewhat more concrete. In the setting of Theorem 3.12, combine the inclusion induced maps as

$$\iota = (\iota_0 \; \iota_1): H_2(M_0; \Lambda) \oplus H_2(M_1; \Lambda) \to H_2(M; \Lambda).$$

We record the following fact; compare with [5, Proposition 4.2] in the simply connected case.

**Lemma 3.13** *The Lagrangian $\widehat{\lambda}_M^{-1}(\Gamma_{F-*}) \subseteq H_2(M; \Lambda)$ contains the graph $\Gamma_{-F}$ of the isometry $-F: H_2(M_0; \Lambda) \to H_2(M_1; \Lambda)$.*

**Proof** Given $x_0 \in H_2(M_0; \Lambda)$, we have

$$\iota\begin{pmatrix} x_0 \\ -F(x_0) \end{pmatrix} = \widehat{\lambda}_M^{-1}\begin{pmatrix} \widehat{\lambda}_{M_0}(x_0) \\ \widehat{\lambda}_{M_1}(F(x_0)) \end{pmatrix}$$

$$= \widehat{\lambda}_M^{-1}\begin{pmatrix} \widehat{\lambda}_0(x_0) \\ F^{-*}(\widehat{\lambda}_{M_0}(x_0)) \end{pmatrix}.$$

Here, the first equality holds by the commutativity of the diagram in (10), while the second holds because $F$ is an isometry. It follows that $\Gamma_{-F} \subseteq \widehat{\lambda}_M^{-1}(\Gamma_{F-*})$, as desired. $\square$

We introduce some notation needed to provide a sufficient criterion to produce compatible pairs. Under the identification $\mathrm{Aut}(\partial\lambda_{M_1}) = \mathrm{Aut}(\mathrm{Bl}_{\partial M_1})$ as in Remark 3.6, there is a left action

$$\big(\mathrm{Homeo}_\varphi^+(\partial M_1) \times \mathrm{Aut}(\lambda_{M_1})\big) \times \mathrm{Aut}(\mathrm{Bl}_{\partial M_1}) \to \mathrm{Aut}(\mathrm{Bl}_{\partial M_1}),$$
$$(f, F) \cdot h := f_* \circ h \circ \partial F^{-1}.$$

The next proposition gives a criterion to find a compatible pair.

**Proposition 3.14** *For $M_0$ and $M_1$ as above, if there exist $f \in \mathrm{Homeo}_\varphi^+(\partial M_0, \partial M_1)$ and $F \in \mathrm{Iso}(\lambda_{M_0}, \lambda_{M_1})$, and if the orbit set*

$$\mathrm{Aut}(\mathrm{Bl}_{\partial M_1})/\mathrm{Homeo}_\varphi^+(\partial M_1) \times \mathrm{Aut}(\lambda_{M_1})$$

*is trivial, then a compatible pair $(f', F')$ exists.*

**Proof** The composition $f_* \circ \partial F^{-1}$ is an isometry of $\partial\lambda_{M_1} \cong \mathrm{Bl}_{\partial M_1}$. By the assumption, there exist a homeomorphism $g \in \mathrm{Homeo}_\varphi^+(\partial M_1)$ and an isometry $G \in \mathrm{Aut}(\lambda_{M_1})$ such that we have $f_* \circ \partial F^{-1} = g_*^{-1} \circ \partial G$. Equivalently, we can write $\widetilde{(g \circ f)}_* = \partial(G \circ F)$ and consequently the pair $(f', F') := (g \circ f, G \circ F)$ is compatible. $\qquad\square$

# 4   Partial classification of compact 4–manifolds with fundamental group $\mathbb{Z}$

In this section, we prove Theorem 1.10 from the introduction. The proof of this result was inspired by [5] and [37].

**Theorem 4.1** *Let $M_0$ and $M_1$ be two 4–manifolds with $\pi_1(M_i) = \mathbb{Z}$ for $i = 0, 1$, whose boundaries are ribbon and have torsion Alexander modules. Suppose that $f \in \mathrm{Homeo}_\varphi^+(\partial M_0, \partial M_1)$ is an orientation-preserving homeomorphism, and suppose that $F \in \mathrm{Iso}(\lambda_{M_0}, \lambda_{M_1})$ is an isometry. If $M_0$ and $M_1$ are not spin, assume that the Kirby–Siebenmann invariants satisfy $\mathrm{ks}(M_0) = \mathrm{ks}(M_1) \in \mathbb{Z}/2$. Then the following assertions are equivalent:*

  (i)   *The pair $(f, F)$ is compatible.*

 (ii)  *The homeomorphism $f$ extends to an orientation-preserving homeomorphism*

$$\Phi \colon M_0 \xrightarrow{\cong} M_1$$

     *inducing the given isometry $F \colon H_2(M_0; \Lambda) \cong H_2(M_1; \Lambda)$.*

**Proof** Proposition 3.11 shows that (ii) implies (i). To prove the converse, given a compatible pair $(f, F)$, we must therefore establish the existence of a homeomorphism $\Phi$. Under this assumption, and in the nonspin case assuming that the Kirby–Siebenmann invariants coincide, we saw in Theorem 3.12 that for some $a, b$ such that $a+b = b_2(M_0)$, there is a homeomorphism

$$M_0 \cup_f -M_1 \cong S^1 \times S^3 \# \overset{a}{\underset{i=1}{\#}} S^2 \times S^2 \# \overset{b}{\underset{i=1}{\#}} S^2 \tilde\times S^2.$$

Define $M := M_0 \cup_f -M_1$. With this notation, Theorem 3.12 additionally states that this homeomorphism takes the Lagrangian $\hat\lambda_M^{-1}(\Gamma_{F-*}) \subseteq H_2(M; \Lambda)$ to the Lagrangian generated by the spheres $\{[\{pt\} \times S^2]\}_{i=1}^{b_2(M_0)}$. Consider the 5–dimensional null-bordism for $M$ given by

$$W := S^1 \times D^4 \natural \overset{a}{\underset{i=1}{\natural}} S^2 \times D^3 \natural \overset{b}{\underset{i=1}{\natural}} S^2 \tilde\times D^3,$$

where $S^2 \tilde\times D^3$ denotes the twisted linear $D^3$ bundle over $S^2$. We prove that $W$ is an $h$–cobordism; it will then automatically be an $s$–cobordism as $\mathrm{Wh}(\mathbb{Z}) = 0$. As $H_i(W, M_0; \Lambda) = 0$ for $i \neq 2, 3$, we must show that $H_2(W, M_0; \Lambda) = 0$ for $i = 2, 3$.

Use $D_{1/2}^3 \subseteq D^3$ to denote a 3–ball of smaller radius inside $D^3$. Choosing the obvious set of generators for $H_2(W; \Lambda)$ and representing them by embedded spheres, we obtain

$$U := \overset{a}{\underset{i=1}{\natural}} S^2 \times D_{1/2}^3 \natural \overset{b}{\underset{i=1}{\natural}} S^2 \tilde\times D_{1/2}^3 \subseteq W.$$

Note that $H_*(U; \Lambda) = H_*(W; \Lambda)$ and $H_*(\partial U; \Lambda) = H_*(M; \Lambda)$. In particular, under this identification, we have

$$(14) \quad \ker\big(H_2(M; \Lambda) \overset{j}{\to} H_2(W; \Lambda)\big) = \mathrm{im}(H_3(U, \partial U; \Lambda) \to H_2(\partial U; \Lambda))$$
$$= \hat\lambda_M^{-1}(\Gamma_{F-*}).$$

In what follows, we identify $H_3(U, \partial U; \Lambda)$ with its image in $H_2(\partial U; \Lambda)$ allowing ourselves for instance to write $H_3(U, \partial U; \Lambda) \subseteq H_2(\partial U; \Lambda)$. We also think of the adjoint $\hat\lambda_M$ of the intersection form on $M$ as having $H_2(\partial U; \Lambda)$ as its domain.

**Claim 2** *The connecting homomorphism*

$$\partial_0 \colon H_3(W \setminus \mathring U, M_0 \sqcup \partial U; \Lambda) \to H_2(M_0; \Lambda),$$

*which arises from the long exact sequence of the triple* $(W \setminus \mathring U, M_0 \sqcup \partial U, \partial U)$, *is an isomorphism.*

**Proof of Claim 2**  It suffices to note that by excision and by our choice of $U$ and $W$, we have $H_i(W \setminus \mathring{U}, \partial U; \Lambda) \cong H_i(W, U; \Lambda) = 0$ for $i = 2, 3$.  □

**Claim 3**  *The following diagram commutes*:

$$
\begin{array}{ccccc}
H_2(M_0; \Lambda) & \xrightarrow{\ \widehat{\lambda}_0\ } & (H_2(M_0; \Lambda) \cup_{\partial F} H_2(M_1; \Lambda), \lambda_0 \cup_{\partial F} -\lambda_1) & \longleftrightarrow & \Gamma_{F-*} \\
{\scriptstyle\cong} \big\uparrow {\scriptstyle\partial_0} & & {\scriptstyle\cong} \big\uparrow {\scriptstyle\widehat{\lambda}_M} & & {\scriptstyle\cong} \big\uparrow {\scriptstyle\widehat{\lambda}_M} \\
H_3(W \setminus \mathring{U}, M_0 \sqcup \partial U; \Lambda) & \hookrightarrow & (H_2(\partial U; \Lambda), \lambda_{\partial U}) & \longleftrightarrow & H_3(U, \partial U; \Lambda)
\end{array}
$$

**Proof of Claim 3**  The commutativity of the right square follows from (14), and so we focus on the left square. Consider the long exact sequence of the triples

$$(W \setminus \mathring{U}, M_0 \sqcup \partial U, \partial U) \quad \text{and} \quad (W \setminus \mathring{U}, M_0 \sqcup \partial U, M_0).$$

The portions of interest can be seen in the two horizontal rows of the following diagram, where $\Lambda$ coefficients are understood:

(15)

$$
\begin{array}{ccccc}
H_3(W\setminus\mathring{U}, \partial U)=0 \longrightarrow H_3(W\setminus\mathring{U}, M_0\sqcup\partial U) \xrightarrow{\ \partial_0\ } H_2(M_0) \longrightarrow 0 = H_2(W\setminus\mathring{U}, \partial U)
\end{array}
$$

$$
\begin{array}{c}
\Big\| \qquad\qquad \begin{array}{c}\iota_0 \\ \Big\downarrow\end{array} \qquad \begin{array}{c}\widehat{\lambda}_0\\ \searrow\end{array} \\
\qquad\qquad\qquad\qquad H_2(M) \xrightarrow[\widehat{\lambda}_M]{\ \cong\ } \begin{array}{c} H_2(M_0) \\ \cup_{\partial F} \\ H_2(M_1)\end{array} \\
\qquad\qquad\qquad\qquad \Big\| \qquad \begin{array}{c}\cong \\ \nearrow \\ \widehat{\lambda}_M\end{array}
\end{array}
$$

$$
\begin{array}{ccccc}
H_3(W\setminus\mathring{U}, M_0)=0 \longrightarrow H_3(W\setminus\mathring{U}, M_0\sqcup\partial U) \longrightarrow H_2(\partial U) \longrightarrow H_2(W\setminus\mathring{U}, M_0)
\end{array}
$$

The claim will follow from the central portion of the diagram, once we explain all its features and its commutativity. The zeros in the first row are a consequence of Claim 2. The bottom leftmost zero is stated in Kreck's work [37, page 734], but we outline a proof in this setting. The exact sequence of the triple $(W \setminus \mathring{U}, M_1 \sqcup \partial U, M_1)$, together with a Mayer–Vietoris argument, give rise to an isomorphism

$$H_1(W \setminus \mathring{U}, M_1 \sqcup \partial U; \Lambda) \cong H_0(\partial U; \Lambda) = \Lambda.$$

Similarly, using the long exact sequence of the triple $(W \setminus \mathring{U}, M_1 \sqcup \partial U, \partial U)$, we deduce that $H_2(W \setminus \mathring{U}, M_1 \sqcup \partial U; \Lambda) = 0$. Since we also have $H_0(W \setminus \mathring{U}, M_1 \sqcup \partial U; \Lambda) = 0$, Poincaré duality and the UCSS imply that

$$H_3(W \setminus \mathring{U}, M_0; \Lambda) = H^2(W \setminus \mathring{U}, M_1 \sqcup \partial U; \Lambda) = 0.$$

To establish the claim, it only remains to show that the diagram in (15) is commutative. The middle square clearly commutes. Proposition 3.9 establishes the commutativity of the triangles but with $H_2(M_0) \cup_{\tilde{f}_*} H_2(M_1)$ in place of $H_2(M_0) \cup_{\partial F} H_2(M_1)$. However, since $(f, F)$ is a compatible pair, these two modules are isomorphic. This concludes our explanation of the diagram in (15) and concludes the proof of Claim 3. $\square$

Write $A$ and $B$ for the images of $H_3(U, \partial U; \Lambda)$ and $H_3(W \setminus \mathring{U}; M_0 \sqcup \partial U; \Lambda)$ in $H_2(\partial U; \Lambda)$. As Proposition 2.13 implies that

$$H_2(M_0; \Lambda) \cup_{\partial F} H_2(M_1; \Lambda) = \Gamma_{F^{-*}} \oplus H_2(M_0; \Lambda),$$

we deduce from Claim 3 that

$$H_2(\partial U; \Lambda) = A \oplus B.$$

Assemble the long exact sequence of the triple $(W \setminus \mathring{U}, M_0 \sqcup \partial U, M_0)$ and the long exact sequence of the triple $(W, W \setminus \mathring{U}, M_0)$ into the following diagram, based on that from [37, page 738], in which $\Lambda$–coefficients are understood:

$$
\begin{array}{ccc}
& 0 & \\
& \downarrow & \\
& B \cong H_3(W \setminus \mathring{U}, M_0 \sqcup \partial U) & \\
& \downarrow & \\
A \cong H_3(U, \partial U) \longrightarrow & H_2(\partial U) & \\
\downarrow{\scriptstyle\cong} & \downarrow & \\
0 \to H_3(W, M_0) \to H_3(W, W \setminus \mathring{U}) \to H_2(W \setminus \mathring{U}, M_0) \to H_2(W, M_0) \to 0 \\
& \downarrow & \\
& 0 &
\end{array}
$$

Here, the left vertical isomorphism comes from excision. Since $H_2(\partial U; \Lambda) = A \oplus B$, the right–down composition $A \to H_2(W \setminus \mathring{U}, M_0; \Lambda)$ is an isomorphism. It follows that the central map in the long row is an isomorphism, and therefore $H_i(W, M_0; \Lambda) = 0$ for $i = 2, 3$. Thus $W$ is a relative $h$–cobordism, and therefore an $s$–cobordism, as desired.

Since $\mathbb{Z}$ is a good group, by the topological $s$–cobordism theorem [15, Theorem 7.1A] $M_0$ and $M_1$ are homeomorphic via a homeomorphism $\Phi$ that extends $f$. It remains to show that $\Phi$ induces the isometry $F$ on $H_2(-; \Lambda)$. The inclusions of $M_0, M_1$ into $M = M_0 \cup_f -M_1 \subseteq W$ give rise to the homomorphism

$$H_2(M_0) \oplus H_2(M_1) \xrightarrow{\iota = (\iota_0 \ \iota_1)} H_2(M) \xrightarrow{j} H_2(W).$$

Set $j_i = j \circ \iota_i$; these are isomorphisms because $W$ is an $s$–cobordism. By definition $j_1^{-1} j_0$ is the isometry induced by the homeomorphism $\Phi$. We noted that $H_2(U; \Lambda) = H_2(W; \Lambda)$ as well as $H_2(\partial U; \Lambda) = H_2(M; \Lambda)$. We also noted in (14) that $\ker(j) = \widehat{\lambda}_M^{-1}(\Gamma_{F-*})$. Using Lemma 3.13, we deduce that $\iota(\Gamma_{-F}) \subseteq \ker(j)$. This inclusion implies that for all $x_0 \in H_2(M_0; \Lambda)$,

$$0 = j\iota \begin{pmatrix} x_0 \\ -F(x_0) \end{pmatrix} = j\iota_0(x_0) - j\iota_1(F(x_0)) = j_0(x_0) - j_1(F(x_0)).$$

Since the $j_i$ are isomorphisms, we obtain the desired conclusion: $j_1^{-1} j_0(x_0) = F(x_0)$. This concludes the proof of Theorem 4.1.      □

# 5   Knotted surfaces in simply connected 4–manifolds

The goal of this section is to use Theorem 4.1 to prove Theorems 1.3 and 1.4 from the introduction. Recall from our conventions that $X$ refers to a closed, simply connected, oriented 4–manifold, while $N = X \setminus \mathring{D}^4$, and $\Sigma$ always denotes a $\mathbb{Z}$–surface of genus $g$ either embedded in $X$ or properly embedded in $N$ (recall that a $\mathbb{Z}$–*surface* refers to an oriented surface whose knot group is $\mathbb{Z}$).

This section is organised as follows. In Section 5.1, we collect some initial facts concerning $\mathbb{Z}$–surfaces. Section 5.2 then shows that any isometry of $\mathrm{Bl}_{\Sigma_{g,1} \times S^1}$ can be realised by a homeomorphism of $\Sigma_{g,1} \times S^1$ (allowing us to construct a compatible pair between $\mathbb{Z}$–surfaces exteriors $N_{\Sigma_0}$ and $N_{\Sigma_1}$). In Sections 5.3 and 5.4, respectively, we prove Theorems 1.3 and 1.4 from the introduction.

## 5.1   Facts about $\mathbb{Z}$–surfaces

**Lemma 5.1**   *If $\Sigma \subseteq N$ is a $\mathbb{Z}$–surface, then it is null-homologous.*

**Proof**   We must show that $[\Sigma, \partial\Sigma] = 0 \in H_2(N, \partial N)$. The intersection form $Q_N$ pairs $H_2(N, \partial N)$ with $H_2(N)$ nonsingularly because $N$ is simply connected. Thus, it is equivalent to show that $Q_N([\Sigma, \partial\Sigma], x) = 0$ for every class $x \in H_2(N)$. Represent such an $x \in H_2(N)$ by a closed surface $S \subseteq N$ that intersects $\Sigma$ transversely in points $p_1, \ldots, p_n$, so that $Q_N([\Sigma, \partial\Sigma], x) = \sum_{i=1}^n \varepsilon(p_i)$, where $\varepsilon(p_i) = \pm 1$. Now the intersection $S \cap N_\Sigma$ is a properly embedded surface in $N_\Sigma$ with oriented boundary (homologous to) $\sum_{i=1}^n \varepsilon(p_i)\mu_\Sigma$. This implies that $\sum_{i=1}^n \varepsilon(p_i)\mu_\Sigma = 0 \in H_1(N_\Sigma)$. But now, since the homology group $H_1(N_\Sigma) = \pi_1(N_\Sigma) = \mathbb{Z}\mu_\Sigma$ is torsion-free, we therefore deduce that $Q_N([\Sigma, \partial\Sigma], x) = \sum_{i=1}^n \varepsilon(p_i) = 0$, establishing that $\Sigma$ is null-homologous.      □

It follows easily that a closed $\mathbb{Z}$–surface $\Sigma \subseteq X$ is also null-homologous in $H_2(X)$. For the next lemma, recall that orientable 2–plane bundles over compact, orientable surfaces are classified up to isomorphism by their (relative) Euler number.

**Lemma 5.2** *Every $\mathbb{Z}$–surface has trivial normal bundle. In the case of nonempty boundary this holds*, *moreover*, *relative to the Seifert framing on the boundary*.

**Proof** In the case that the boundary is nonempty, the surface is oriented and therefore the normal bundle is trivial. However, we need to show that this holds relative to the Seifert framing, and we need to show that the normal bundle is likewise trivial in the closed case. Continuing with the nonempty boundary case, cap off the $\mathbb{Z}$–surface $\Sigma$ with a Seifert surface $F$ in $S^3$. Form a push-off by pushing $F$ off itself in $S^3$ and extending this to a normal push-off in $N$ along a generic section that agrees with the push-off of $F$ along $\partial F = \partial \Sigma$.

Now the argument is the same in the closed case and in the nonempty boundary case (in the closed case take a push-off of $\Sigma$ using a generic section of the normal bundle). Now in both cases, by the proof of Lemma 5.1, the push-off intersects $\Sigma$ algebraically trivially. Therefore the normal bundle of $\Sigma$ has vanishing Euler number, relative to the Seifert framing in the case of nonempty boundary. Thus the normal bundle of $\Sigma$ is trivial, again relative to the Seifert framing in the case of nonempty boundary. $\square$

The boundary of the exterior of $N_\Sigma$ is homeomorphic to $M_{K,g} := E_K \cup_\partial (\Sigma_{g,1} \times S^1)$, where $\Sigma_{g,1}$ is the orientable genus $g$ surface with one boundary component and $E_K := S^3 \setminus \nu(K)$. We give more details on this identification as we will then make it implicitly throughout the remainder of the paper.

**Remark 5.3** If $\Sigma \subseteq X$ is a closed oriented surface with $\pi_1(X_\Sigma) = \mathbb{Z}$, then the homotopy class of the meridian of $\Sigma \subseteq X$ is the unique nontrivial primitive class in $\pi_1(\partial X_\Sigma)$ that bounds a disc in $\Sigma \times D^2$ and maps to $1 \in \mathbb{Z}$. Fix a framing of the normal bundle of $\Sigma$, ie an identification $\nu\Sigma \cong \Sigma \times \mathbb{R}^2$ compatible with the orientation, with the property that for each simple closed curve $\gamma_k \subseteq \Sigma$, we have that $\gamma_k \times \{e_1\} \subseteq X \setminus \Sigma$ is null-homologous in $H_1(X \setminus \nu\Sigma) \cong \mathbb{Z}$. Use a choice of an identification $\Sigma \cong \Sigma_g$ and this condition to fix an identification of the boundary of the exterior with $\Sigma_g \times S^1$. Any two choices now differ by an element of the mapping class group of $\Sigma_g$. Similarly, if $\Sigma \subseteq N$ is a properly embedded, oriented surface, then we can identify $\partial N_\Sigma$ with $E_K \cup_\partial (\Sigma_{g,1} \times S^1)$.

**Lemma 5.4**  *The exterior of a $\mathbb{Z}$–surface $\Sigma$ has ribbon boundary.*

**Proof**  Since $N$ is simply connected, we have

$$\{1\} = \pi_1(N_\Sigma) *_{\pi_1(\Sigma \times S^1)} \pi_1(\Sigma \times D^2),$$

which implies that $\pi_1(\Sigma \times S^1) \to \pi_1(N_\Sigma)$ is onto. In fact, by the parametrisation of $\Sigma \times S^1$ described in Remark 5.3, we have that $\pi_1(\Sigma)$ maps trivially to $\pi_1(N_\Sigma)$, so the fundamental group of $N_\Sigma$ is generated by a meridian of $\Sigma$. The closed case follows, since the exterior of a closed surface in $X$ can be thought of as the exterior of a surface in $N = X \setminus \mathring{D}^4$.                                                                                 □

**Lemma 5.5**  *The boundary of the exterior of a $\mathbb{Z}$–surface $\Sigma$ has $\Lambda$–torsion Alexander module. More precisely, the following hold.*

(1)  *If $\Sigma \subseteq X$ is closed, then*

$$H_1(\partial X_\Sigma; \Lambda) = H_1(\Sigma_g \times S^1; \Lambda) = (\Lambda/(t-1))^{\oplus 2g}.$$

(2)  *If $\Sigma \subseteq N$ has boundary a knot $K$, then the inclusion induces an isomorphism*

$$H_1(E_K; \Lambda) \oplus H_1(\Sigma_{g,1} \times S^1; \Lambda) \cong H_1(\partial N_\Sigma; \Lambda).$$

*In particular, if $K$ has Alexander polynomial one, then*

$$H_1(\partial N_\Sigma; \Lambda) = (\Lambda/(t-1))^{\oplus 2g}.$$

**Proof**  Using infinite cyclic covers, the assertion in the closed case is immediate:

$$H_1(\partial X_\Sigma; \Lambda) = H_1(\Sigma_g \times S^1; \Lambda) = H_1(\widetilde{\Sigma_g \times S^1}; \mathbb{Z}) = H_1(\Sigma_g \times \mathbb{R})$$
$$= \mathbb{Z}^{2g} = (\Lambda/(t-1))^{\oplus 2g}.$$

We prove the second assertion. Consider the Mayer–Vietoris sequence for the decomposition $M_{K,g} = E_K \cup_\partial (\Sigma_{g,1} \times S^1)$ with $\Lambda$ coefficients,

$$\cdots \to H_1(S^1 \times S^1; \Lambda) \to H_1(E_K; \Lambda) \oplus H_1(\Sigma_{g,1} \times S^1; \Lambda) \to H_1(M_{K,g}; \Lambda) \to 0.$$

We have
$$H_1(S^1 \times S^1; \Lambda) = H_1(S^1 \times \mathbb{R}) = \mathbb{Z},$$
$$H_1(\Sigma_{g,1} \times S^1; \Lambda) = H_1(\Sigma_{g,1} \times \mathbb{R}) = \mathbb{Z}^{2g}.$$

The map $H_1(S^1 \times S^1; \Lambda) \to H_1(E_K; \Lambda)$ is the zero map: it sends the generator to the lift of the longitude of $E_K$; this is null-homologous in the infinite cyclic cover $E_K^\infty$ of $E_K$. Next, the homomorphism $H_1(S^1 \times S^1; \Lambda) \to H_1(\Sigma_{g,1} \times S^1; \Lambda)$ is also the zero map: it coincides with the map $H_1(S^1) \to H_1(\Sigma_{g,1})$ sending the generator to $[\partial \Sigma_{g,1}]$, which vanishes in $H_1(\Sigma_{g,1})$. This establishes the second assertion.    □

## 5.2 Isometries of the Blanchfield pairing

Given an orientable genus $g$ surface $\Sigma_{g,1}$ with one boundary component, we determine the isometries of the Blanchfield pairing on $\Sigma_{g,1} \times S^1$ and show that they can be realised by orientation-preserving homeomorphisms. We then describe the Blanchfield pairing of

$$M_{K,g} := E_K \cup (\Sigma_{g,1} \times S^1),$$

where $E_K := S^3 \setminus \nu(K)$ is the knot exterior.

To describe the Blanchfield pairing $\mathrm{Bl}_{\Sigma_{g,1} \times S^1}$, we fix some notation on the symplectic group:

$$J_g := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{\oplus g}, \quad \mathrm{Sp}_{2g}(\mathbb{Z}) := \{A \in M_{2g \times 2g}(\mathbb{Z}) \mid A^T J_g A = J_g\}.$$

Given a genus $g$ surface $\Sigma_{g,1}$ with one boundary component, we call a basis of $H_1(\Sigma_{g,1})$ *symplectic* if the intersection form on $\Sigma_{g,1}$ with respect to this basis is represented by $J_g$. We also consider the map $\varphi \colon \pi_1(\Sigma_{g,1} \times S^1) \twoheadrightarrow \mathbb{Z}$ induced by the projection on the second coordinate and recall that

$$\mathrm{Homeo}_{\varphi}^+(\Sigma_{g,1} \times S^1) = \{f \in \mathrm{Homeo}^+(\Sigma_{g,1} \times S^1) \mid f_* \circ \varphi = \varphi\}.$$

The next proposition identifies the isometries of $\mathrm{Bl}_{\Sigma_{g,1} \times S^1}$ with $\mathrm{Sp}_{2g}(\mathbb{Z})$ and uses this fact to show that the orbit set $\mathrm{Aut}(\mathrm{Bl}_{\Sigma_{g,1} \times S^1})/\mathrm{Homeo}_{\varphi}^+(\Sigma_{g,1} \times S^1)$ is trivial.

**Proposition 5.6**  *A choice of a symplectic basis for $H_1(\Sigma_{g,1})$ yields an identification*

$$\mathrm{Aut}(\mathrm{Bl}_{\Sigma_{g,1} \times S^1}) = \mathrm{Sp}_{2g}(\mathbb{Z}).$$

*Every element of the group $\mathrm{Aut}(\mathrm{Bl}_{\Sigma_{g,1} \times S^1})$ can be realised by a homeomorphism in $\mathrm{Homeo}_{\varphi}^+(\Sigma_{g,1} \times S^1)$:*

$$\mathrm{Aut}(\mathrm{Bl}_{\Sigma_{g,1} \times S^1})/\mathrm{Homeo}_{\varphi}^+(\Sigma_{g,1} \times S^1) = \{\mathrm{Id}\}.$$

*Any such homeomorphism can be assumed to be of the form $j \times \mathrm{Id}_{S^1}$ for some $j \in \mathrm{Homeo}^+(\Sigma_{g,1})$ that fixes the boundary of $\Sigma_{g,1}$ pointwise.*

**Proof**  Let $h$ be an isometry of the Blanchfield pairing $\mathrm{Bl}_{\Sigma_{g,1} \times S^1}$. Since we saw in Lemma 5.5 that $H_1(\Sigma_{g,1} \times S^1; \Lambda) = H_1(\Sigma_{g,1})$, the isometry $h$ can be thought of as a $\mathbb{Z}$–linear map. We assert that it preserves the intersection form on the surface $\Sigma_{g,1}$. Pick a symplectic basis for $H_1(\Sigma_{g,1})$ so that its intersection form is represented by $J_g$.

By [18, Corollary 1.2], the Blanchfield form on the fibred manifold $\Sigma_{g,1} \times S^1$ is given by

$$(v, w) \mapsto v^T (t^{-1} - 1)^{-1} J_g w \in \Lambda_S / \Lambda.$$

Here, $\Lambda_S = \mathbb{Z}[t^{\pm 1}, (t-1)^{-1}]$ denotes the ring obtained from $\Lambda = \mathbb{Z}[t^{\pm 1}]$ by inverting $(t-1)$. Since $h$ can be thought of as a $\mathbb{Z}$–linear map, we represent it by a matrix $H$ with coefficients in $\mathbb{Z}$. As $h$ is an isometry, we have the following equalities of matrices with coefficients in $\Lambda_S / \Lambda$:

$$(t^{-1} - 1)^{-1} H^T J_g H = J_g (t^{-1} - 1)^{-1}.$$

Since $H$ and $J_g$ take values in $\mathbb{Z}$, the coefficients of the matrix $H^T J_g H - J_g$ take values in $(t-1)\Lambda$. On the other hand, $H^T J_g H - J_g$ has coefficients in $\mathbb{Z}$, so $H^T J_g H - J_g = 0$, and we deduce that $H^T J_g H = J_g$. This shows that $H$ is a symplectomorphism, proving the first assertion.

We prove the second assertion. We may realise any $h \in \mathrm{Sp}_{2g}(\mathbb{Z})$ as the map on $H_1(\Sigma_{g,1})$ induced by an orientation-preserving homeomorphism $j \colon \Sigma_{g,1} \to \Sigma_{g,1}$ that fixes the boundary pointwise [13, Section 2.1 and the discussion following Theorem 6.4]. Cross $j$ with the identity on $S^1$ to obtain an orientation-preserving homeomorphism $j \times \mathrm{Id} \colon \Sigma_{g,1} \times S^1 \to \Sigma_{g,1} \times S^1$ that belongs to $\mathrm{Homeo}^+_\varphi(\Sigma_{g,1} \times S^1)$. $\qquad\square$

The next proposition describes the Blanchfield pairing of $M_{K,g} = E_K \cup (\Sigma_{g,1} \times S^1)$ and its automorphism group; here recall that $E_K = S^3 \setminus \nu(K)$ is the knot exterior.

**Proposition 5.7** *For a knot $K \subseteq S^3$ there is an isometry $\mathrm{Bl}_{M_{K,g}} \cong \mathrm{Bl}_{E_K} \oplus \mathrm{Bl}_{\Sigma_{g,1} \times S^1}$, and*

$$\mathrm{Aut}(\mathrm{Bl}_{M_{K,g}}) \cong \mathrm{Aut}(\mathrm{Bl}_{E_K}) \oplus \mathrm{Aut}(\mathrm{Bl}_{\Sigma_{g,1} \times S^1}).$$

**Proof** We proved in Lemma 5.5 that the inclusion induces a $\Lambda$–isomorphism

$$H_1(E_K; \Lambda) \oplus H_1(\Sigma_{g,1} \times S^1; \Lambda) \cong H_1(M_{K,g}; \Lambda).$$

The isometry $\mathrm{Bl}_{M_{K,g}} \cong \mathrm{Bl}_{E_K} \oplus \mathrm{Bl}_{\Sigma_{g,1} \times S^1}$ now follows from [16, Theorem 1.1].

We must now show that $\mathrm{Aut}(\mathrm{Bl}_{E_K} \oplus \mathrm{Bl}_{\Sigma_{g,1} \times S^1}) = \mathrm{Aut}(\mathrm{Bl}_{E_K}) \oplus \mathrm{Aut}(\mathrm{Bl}_{\Sigma_{g,1} \times S^1})$. The order of $H_1(M_K; \Lambda)$ is the Alexander polynomial $\Delta_K$ (which is not divisible by $(t-1)$) and the order of $H_1(\Sigma_{g,1} \times S^1; \Lambda)$ is $(t-1)^{2g}$. An automorphism $h$ in $\mathrm{Aut}(\mathrm{Bl}_{E_K}) \oplus \mathrm{Aut}(\mathrm{Bl}_{\Sigma_{g,1} \times S^1})$ can be written as

$$h = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}.$$

Since the orders of $H_1(M_K; \Lambda)$ and $H_1(\Sigma_{g,1} \times S^1; \Lambda)$ are coprime, we deduce that $h_{12} = h_{21} = 0$. $\qquad\square$

## 5.3 Surfaces in manifolds with boundary

The aim of this subsection is to prove Theorem 1.3 from the introduction. Given a knot $K \subseteq S^3$, thanks to Proposition 5.7, we can write automorphisms of $\mathrm{Bl}_{M_{K,g}}$ as $h = h_K \oplus h_\Sigma$ with $h_K \in \mathrm{Aut}(\mathrm{Bl}_{E_K})$ and $h_\Sigma \in \mathrm{Aut}(\mathrm{Bl}_{\Sigma_{g,1} \times S^1})$.

Let us recall the construction of an isotopy, as introduced in Section 1.3. Suppose that $f_K \colon E_K \to E_K$ is an orientation-preserving homeomorphism that is the identity on $\partial E_K$. Extend $f_K$ via the identity on $\nu K$ to an orientation-preserving self-homeomorphism $\widetilde{f}_K$ of $S^3$. The mapping class group of $S^3$ is trivial, so there is an isotopy $\Theta(f_K)_t \colon S^3 \to S^3$ between the extension and the identity, such that $\Theta(f_K)_0 = \mathrm{Id}$ and $\Theta(f_K)_1 = \widetilde{f}_K$. We can now prove Theorem 1.3.

**Theorem 5.8** *Let $\Sigma_0, \Sigma_1 \subseteq N$ be two $\mathbb{Z}$–surfaces of genus $g$ with boundary $K$. Suppose there is an isometry $F \colon \lambda_{N_{\Sigma_0}} \cong \lambda_{N_{\Sigma_1}}$, and write $\partial F = h_K \oplus h_\Sigma$.*

- *If $h_K$ is induced by an orientation-preserving homeomorphism $f_K \colon E_K \to E_K$ that is the identity on $\partial E_K$, then $f_K$ extends to an orientation-preserving homeomorphism of pairs*

$$(N, \Sigma_0) \xrightarrow{\ \cong\ } (N, \Sigma_1)$$

  *inducing the given isometry $F \colon H_2(N_{\Sigma_0}; \Lambda) \cong H_2(N_{\Sigma_1}; \Lambda)$.*

- *If, in addition, $N = D^4$, then for any choice of isotopy $\Theta(f_K)$, the surfaces $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic via an ambient isotopy of $D^4$ extending $\Theta(f_K)$.*

**Proof** By assumption, the boundary of the isometry $F$ is $\partial F = h_K \oplus h_\Sigma$ for some $h_K \in \mathrm{Aut}(\mathrm{Bl}_{E_K})$ and some $h_\Sigma \in \mathrm{Aut}(\mathrm{Bl}_{\Sigma_{g,1} \times S^1})$. By Proposition 5.6, we can realise $h_\Sigma$ by an orientation-preserving homeomorphism of the form $j \times \mathrm{Id}_{S^1}$, with $j \colon \Sigma_{g,1} \to \Sigma_{g,1}$ an orientation-preserving homeomorphism that fixes the boundary pointwise. Moreover, $h_K$ is realised by an orientation-preserving homeomorphism $f_K \colon E_K \to E_K$ that fixes the boundary $\partial E_K$. Gluing these two homeomorphisms together, we obtain a homeomorphism

$$f := f_K \cup (j \times \mathrm{Id}_{S^1}) \colon M_{K,g} \to M_{K,g}.$$

Since $f$ is the identity on $\partial E_K \subseteq M_{K,g}$, we deduce that $f \in \mathrm{Homeo}^+_\varphi(M_{K,g})$. It follows that $(f, F)$ is a compatible pair. The $N_{\Sigma_i}$ are 4–manifolds whose boundaries are ribbon (recall Lemma 5.4) and have torsion Alexander modules for $i = 0, 1$ (by Lemma 5.5). Since $\Sigma_i \times D^2$ is a smooth manifold, $\mathrm{ks}(\Sigma_i \times D^2) = 0$. Thus for $i = 0, 1$,

by additivity of ks under gluing we have $\text{ks}(N_{\Sigma_i}) = \text{ks}(N_{\Sigma_i}) + \text{ks}(\Sigma_i \times D^2) = \text{ks}(N)$. In particular, $\text{ks}(N_{\Sigma_0}) = \text{ks}(N_{\Sigma_1})$. We may therefore apply Theorem 4.1 to extend $f$ to an orientation-preserving homeomorphism $\Phi \colon N_{\Sigma_0} \to N_{\Sigma_1}$ that induces $F$. As a consequence, we obtain the required homeomorphism of pairs:

$$\Phi' := \Phi \cup (j \times \text{Id}_{D^2}) \colon (N, \Sigma_0) \to (N, \Sigma_1).$$

This establishes the first assertion.

For the second assertion, let $N = D^4$ and let $\Theta(f_K)_t \colon S^3 \to S^3$ be an isotopy of self-homeomorphisms of $S^3$, with $\Theta(f_K)_0 = \text{Id}_{S^3}$ and $\Theta(f_K)_1$ equal to the extension $\widetilde{f}_K$ of $f_K$ by the identity to all of $S^3$. Cone $\Theta(f_K)$ to obtain an isotopy $C(\Theta(f_K)_t) \colon D^4 \to D^4$ with $C(\Theta(f_K)_0) = \text{Id}_{D^4}$ and $C(\Theta(f_K)_1)$ equal to the cone $C(\widetilde{f}_K)$ of $\widetilde{f}_K$.

Next, note that $C(\widetilde{f}_K)$ and $\Phi' \colon D^4 \to D^4$ are two homeomorphisms of $D^4$ that restrict to the same homeomorphism of $S^3$ on the boundary. Therefore $C(\widetilde{f}_K)^{-1} \circ \Phi' \colon D^4 \to D^4$ is a homeomorphism restricting to $\text{Id}_{S^3}$ on the boundary. By the Alexander trick, $C(\widetilde{f}_K)^{-1} \circ \Phi'$ is isotopic rel boundary to the identity, via an isotopy $G_t \colon D^4 \to D^4$ with $G_0 = \text{Id}_{D^4}$ and $G_1 = C(\widetilde{f}_K)^{-1} \circ \Phi'$. Note that $G_t|_{S^3} = \text{Id}_{S^3}$ for all $t \in [0, 1]$. Then let

$$H_t := C(\Theta(f_K)_t) \circ G_t \colon D^4 \to D^4.$$

This is an isotopy with

$$H_0 = \text{Id}_{D^4} \circ \text{Id}_{D^4} = \text{Id}_{D^4} \quad \text{and} \quad H_1 = C(\widetilde{f}_K) \circ C(\widetilde{f}_K)^{-1} \circ \Phi' = \Phi'.$$

In addition, for every $t \in [0, 1]$ we have that $H_t|_{S^3} = \Theta(f_K)_t \circ G_t|_{S^3} = \Theta(f_K)_t$. Since $\text{Id}_{D^4}(\Sigma_0) = \Sigma_0$ and $\Phi'(\Sigma_0) = \Sigma_1$, the isotopy $H_t$ is a topological ambient isotopy between $\Sigma_0$ and $\Sigma_1$, and it extends $\Theta(f_K)$, as required.                                □

We also note the following corollary, in which we give a slightly relaxed criterion, but without precise control on the isotopy of $K$. It is sometimes easier to construct homeomorphisms of $E_K$ that do not fix the boundary $\partial E_K$, but rather only fix the basepoint, so this could be a useful variation.

**Corollary 5.9**  *Let $K \subseteq S^3$ be a knot and fix a basepoint in $\partial E_K$. Suppose that $f \colon S^3 \to S^3$ is an orientation-preserving, basepoint-preserving homeomorphism with $f(K) = K$ as oriented knots, so also $f(E_K) = E_K$. Then $f$ induces an isomorphism $h \colon H_1(E_K; \Lambda) \to H_1(E_K; \Lambda)$. Let $\Sigma_0, \Sigma_1$ be two $\mathbb{Z}$–surfaces in $N = X \setminus \mathring{D}^4$ with boundary $K \subseteq S^3 = \partial N$, and let $F \colon \lambda_{N_{\Sigma_0}} \cong \lambda_{N_{\Sigma_1}}$ be an isometry with $\partial F = h$. Then $\Sigma_0$ and $\Sigma_1$ are related by a homeomorphism of pairs, and if $N = D^4$, then they are ambiently isotopic.*

**Proof** Since $f$ is orientation-preserving as a map from $S^3$ to $S^3$, and as a map $K \to f(K)$, it restricts to an orientation-preserving map on $\partial E_K$ that preserves the homotopy class of the zero-framed longitude and therefore preserves the orientation of the meridian. Thus $f$ commutes with the map $\pi_1(E_K) \to \mathbb{Z}$ inducing the $\Lambda$– coefficients, and therefore induces a map $h \colon H_1(E_K; \Lambda) \to H_1(E_K; \Lambda)$ as claimed; recall Proposition 3.7. Let $\theta_t \colon S^3 \to S^3$ be an isotopy from the identity of $S^3$ to $f$. Let $S^3 \times [0, 1] \hookrightarrow N$ be a collar on the boundary of $N$ with $S^3 \times \{0\}$ mapping to $\partial N$. Let $\Theta_t \colon N \to N$ be the isotopy that is the identity outside $S^3 \times [0, 1]$ and defined by $\Theta_t(x, s) = \theta_{t(1-s)}(x)$. This performs $\theta_t$ on the boundary, and tapers it in the collar. Let $\Sigma'_0 := \Theta_1(\Sigma_0)$, and note that $\Sigma'_0$ is ambiently isotopic (not necessarily rel boundary) to $\Sigma_0$. In addition, $\partial(\Theta_1)_* = \partial F = h$ as isomorphisms of $H_1(E_K; \Lambda)$. Therefore $G := F \circ (\Theta_1)_*^{-1} \colon \lambda_{N_{\Sigma'_0}} \xRightarrow{\cong} \lambda_{N_{\Sigma_1}}$ is an isometry with $\partial G = \mathrm{Id}$. By Theorem 1.3, $\Sigma'_0$ and $\Sigma_1$ are related by a homeomorphism of pairs restricting to the identity on the boundary, and if $N = D^4$ then they are ambiently isotopic rel boundary. The corollary now follows immediately. $\square$

## 5.4 Surfaces in closed manifolds

Now we recover the statements in the closed setting, proving Theorem 1.4 from the introduction. As discussed in the introduction, for a closed, simply connected 4– manifold $X$ there is a classification of self-homeomorphisms of $X$ in terms of their action on $H_2(X)$. The hypothesis of Theorem 1.4 is in terms of an isometry between the $\Lambda$–intersection forms of the surface exteriors $X_{\Sigma_0}$ and $X_{\Sigma_1}$. In order to carefully state Theorem 1.4, first we need to understand the relationship between such an isometry and isomorphisms of the second homology $H_2(X)$ of the ambient 4–manifold.

**Lemma 5.10** *Let $\Sigma_0, \Sigma_1 \subseteq X$ be closed $\mathbb{Z}$–surfaces of the same genus.*

(i) *An isometry $F \colon \lambda_{X_{\Sigma_0}} \xRightarrow{\cong} \lambda_{X_{\Sigma_1}}$ induces an isomorphism $F_{\mathbb{Z}} \colon H_2(X) \xRightarrow{\cong} H_2(X)$.*

(ii) *If the isometry $F$ is induced by a homeomorphism $\Phi' \colon (X, \Sigma_0) \to (X, \Sigma_1)$ of pairs, then $F_{\mathbb{Z}} = \Phi'_*$; in particular, $F_{\mathbb{Z}}$ is an isometry of the standard intersection form $Q_X$.*

**Proof** We claim that $H_2(X_{\Sigma_i}; \Lambda) \otimes_\Lambda \mathbb{Z} \cong H_2(X_{\Sigma_i})$. To prove this, we use the universal coefficient spectral sequence with second page $E^2_{p,q} = \mathrm{Tor}^\Lambda_p(H_q(X_{\Sigma_i}; \Lambda), \mathbb{Z})$ and which converges to $H_*(X_{\Sigma_i})$. The claim now follows promptly from the fact that $H_1(X_{\Sigma_i}; \Lambda) = 0$ and

$$\mathrm{Tor}^\Lambda_2(H_0(X_{\Sigma_i}; \Lambda), \mathbb{Z}) = \mathrm{Tor}^{\mathbb{Z}[\mathbb{Z}]}_2(\mathbb{Z}, \mathbb{Z}) = H_2(\mathbb{Z}; \mathbb{Z}) = H_2(S^1) = 0.$$

The claim implies that the intersection form on $H_2(X_{\Sigma_i})$ is obtained from the intersection form on $H_2(X_{\Sigma_i}; \Lambda)$ by tensoring down. Therefore $F$ induces an isometry $F \otimes_\Lambda \mathrm{Id}_{\mathbb{Z}} \colon H_2(X_{\Sigma_0}) \to H_2(X_{\Sigma_1})$. A Mayer–Vietoris argument yields the commutative diagram

(16)
$$
\begin{array}{ccccccccc}
0 & \longrightarrow & \mathbb{Z}^{2g} & \longrightarrow & H_2(X_{\Sigma_0}) & \xrightarrow{p_0} & H_2(X) & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle (F \otimes_\Lambda \mathrm{Id}_{\mathbb{Z}})|} & & \downarrow{\scriptstyle F \otimes_\Lambda \mathrm{Id}_{\mathbb{Z}}} & & \downarrow{\scriptstyle F_{\mathbb{Z}}} & & \\
0 & \longrightarrow & \mathbb{Z}^{2g} & \longrightarrow & H_2(X_{\Sigma_1}) & \xrightarrow{p_1} & H_2(X) & \longrightarrow & 0
\end{array}
$$

where the dotted map labelled $F_{\mathbb{Z}}$ is defined as the unique homomorphism that makes the right-hand square commute (both existence and uniqueness follow from a short diagram chase).

The second assertion is now immediate: the homeomorphism $\Phi'$ induces an isometry of $(H_2(X), Q_X)$ that satisfies $\Phi'_* \circ p_0 = p_1 \circ (F \otimes_\Lambda \mathrm{Id}_{\mathbb{Z}})$, so $F_{\mathbb{Z}} = \Phi'_*$ is also an isometry. $\qquad\square$

We move on to the proof of Theorem 1.4, whose statement we recall here for the benefit of the reader.

**Theorem 5.11** *Let $\Sigma_0, \Sigma_1 \subseteq X$ be two closed $\mathbb{Z}$–surfaces of the same genus.*

(i) *If the intersection forms $\lambda_{X_{\Sigma_0}}$ and $\lambda_{X_{\Sigma_1}}$ are isometric via an isometry $F$, then there is an orientation-preserving homeomorphism of pairs*
$$
\Phi \colon (X, \Sigma_0) \xrightarrow{\cong} (X, \Sigma_1)
$$
*inducing the given isometry $\Phi_* = F \colon H_2(X_{\Sigma_0}; \Lambda) \xrightarrow{\cong} H_2(X_{\Sigma_1}; \Lambda)$.*

(ii) *The isometry $F$ induces an isometry $F_{\mathbb{Z}} \colon H_2(X) \to H_2(X)$ of the standard intersection form $Q_X$ of $X$ by Lemma 5.10. The surfaces $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic if and only if $F_{\mathbb{Z}} = \mathrm{Id}$.*

**Proof** After an ambient isotopy, assume that $\Sigma_0$ and $\Sigma_1$ coincide on a disc $D^2 \subseteq \Sigma_0 \cap \Sigma_1$. Assume that the normal bundles also coincide over this $D^2$. Consider the preimage $\mathring{D}^2 \times \mathbb{R}^2 \subseteq \nu\Sigma_i$. This is homeomorphic to an open 4–ball $\mathring{D}^4$. Remove this $(\mathring{D}^4, \mathring{D}^2)$ from $(X, \Sigma_i)$ to obtain $(N, \widetilde{\Sigma}_i)$, with $\partial\widetilde{\Sigma}_i = \Sigma_i \cap \partial N$ the unknot $K$ in $S^3$. Then the exterior of $\Sigma_i$ in $X$ equals the exterior of $\widetilde{\Sigma}_i$ in $N$.

Since $X_{\Sigma_i} = N_{\widetilde{\Sigma}_i}$, the $\Lambda$–intersection forms are unchanged and the isometry $F$ also induces an isometry $F \colon \lambda_{N_{\widetilde{\Sigma}_0}} \xrightarrow{\cong} \lambda_{N_{\widetilde{\Sigma}_1}}$. Write $\partial F = h_K \oplus h_\Sigma$ as in Theorem 5.8.

The unknot $K \subseteq S^3$ has trivial Alexander module, so $h_K = \text{Id}$ and so $h_K$ is realised by the homeomorphism $f_K = \text{Id}: E_K \to E_K$. Theorem 5.8 provides a rel boundary homeomorphism of pairs $\Phi': (N, \tilde{\Sigma}_0) \to (N, \tilde{\Sigma}_1)$ that induces the isometry $F$ on the $\Lambda$–homology of the surface exteriors. We recover the required homeomorphism of pairs $\Phi: (X, \Sigma_0) \to (X, \Sigma_1)$ by gluing $\Phi'$ with the identity homeomorphism $(D^4, D^2) \to (D^4, D^2)$.

Lemma 5.10 implies that the isometry $F$ induces an isometry $F_\mathbb{Z}$ of the standard intersection form $Q_X$. In particular, Lemma 5.10(ii) specifies that $F_\mathbb{Z}$ is induced by the homeomorphism $\Phi$. If $F_\mathbb{Z} = \text{Id}$, then $\Phi$ is a self-homeomorphism of $X$ inducing the identity on $H_2(X)$, so by [45, Theorem 1.1] and [15, Theorem 10.1], $\Phi$ is isotopic to the identity. It follows that $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic.

On the other hand, if $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic, then the induced homeomorphism between their exteriors extends to a homeomorphism from $X$ to itself that is isotopic to the identity, and so certainly induces the identity map on $H_2(X)$. □

# 6 Equivariant intersection forms of surface exteriors

The goal of this section is to collect some results about the intersection forms of $\mathbb{Z}$–surface exteriors. The main result of Section 6.1 shows that up to direct summands with the Hermitian form

$$\mathcal{H}_2 := \left( \Lambda^2, \begin{pmatrix} 0 & t-1 \\ t^{-1}-1 & 0 \end{pmatrix} \right),$$

the equivariant intersection forms of any two $\mathbb{Z}$–surface exteriors are isometric; this will be Proposition 6.3. Section 6.2 will then focus on the case of $\mathbb{Z}$–surfaces $\Sigma \subseteq N = X \setminus \mathring{D}^4$ with boundary an Alexander polynomial one knot $K \subseteq S^3$. The main result is Corollary 6.6 which shows that $\lambda_{N_\Sigma}$ becomes isometric to $Q_X \oplus \mathcal{H}_2^{\oplus g}$ after adding sufficiently many copies of $\mathcal{H}_2$, where $Q_X$ is the $\mathbb{Z}$–valued intersection form of $X$ (note that $Q_X = Q_N$).

## 6.1 Equivariant intersection forms of $\mathbb{Z}$–surface exteriors

The goal of this section is to study the $\Lambda$–intersection form of $\mathbb{Z}$–surface exteriors up to stabilisations by $\mathcal{H}_2$.

We start by describing the $\Lambda$–intersection form of unknotted surfaces in $S^4$.
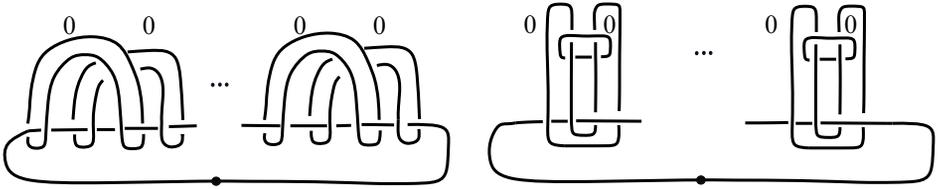
Figure 1: Two handle diagrams for the exterior of a standardly embedded genus $g$ surface $F \subseteq D^4$ with boundary the unknot in $S^3$.

**Lemma 6.1** *If $\Sigma \subseteq S^4$ is an unknotted oriented closed genus $g$ surface, then the $\Lambda$–intersection form of $S^4_\Sigma$ is isometric to $\mathcal{H}_2^{\oplus g}$.*

**Proof** We carry out the computation for the standardly embedded genus $g$ surface $\Sigma \subseteq S^4$. Use $U \subseteq S^3$ to denote the unknot. Slice $(S^4, \Sigma)$ along an equatorial $(S^3, U)$ to obtain a decomposition $(S^4, \Sigma) = (D^4, \overset{\circ}{\Sigma}) \cup (D^4, D)$, with $D \subseteq D^4$ a disc bounding $U$ and $\overset{\circ}{\Sigma} \subseteq D^4$ a punctured unknotted surface in $D^4$.

Note that $S^4_\Sigma$ and $D^4_{\Sigma^\circ} := D^4 \setminus \nu \overset{\circ}{\Sigma}$ are homeomorphic, since we can assume that the removed $\overset{\circ}{D^4}$ lies in the regular neighbourhood $\nu \Sigma$ removed from $S^4$ to form $S^4_\Sigma$. It follows that $H_2(S^4_\Sigma; \Lambda) = H_2(D^4_{\Sigma^\circ}; \Lambda)$ and that the $\Lambda$–intersection forms agree.

Thus it remains to compute the $\Lambda$–intersection form of the exterior $D^4_F$ of a properly embedded unknotted surface $F \subseteq D^4$. A handle diagram with a single one-handle and $2g$ two-handles for $D^4_F$ appears in the left-hand side of Figure 1, produced using the formalism of [19, Section 6.2]. It can then be isotoped as in the right-hand side of Figure 1, leading to a handle diagram for the infinite cyclic cover of $D^4_F$ depicted in Figure 2. From this diagram, by taking the union of the cores of the 2–handles with null-homotopies of their attaching curves in the 4–ball, we obtain generators of $\pi_2(D^4_F) = H_2(D^4_F; \Lambda) = \Lambda^{2g}$. The $\Lambda$–intersection form can be computed via (equivariant) linking numbers, yielding the required result. $\qquad\square$
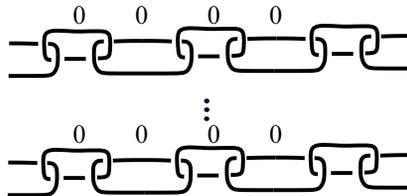


Figure 2: A handle diagram for the infinite cyclic cover of $D^4_F$, where $F \subseteq D^4$ is an unknotted punctured surface.

We recall the concept of a 1–handle stabilisation for a surface in a 4–manifold. The following definition was motivated by [25]. Let $\Sigma \subseteq V$ be a locally flat (connected) surface embedded in a 4–manifold $V$. Let $B$ be an embedding of $D^4$ into $V$ such that $\partial B$ intersects $\Sigma$ transversely in a 2–component unlink $L$, and $B$ intersects $\Sigma$ transversely in two discs $D_0$ and $D_1$, which can be simultaneously isotoped within $B$ to lie in $\partial B$. Suppose that, for $i = 0, 1$, a 3–dimensional 1–handle $D^2 \times [0, 1]$ is embedded in the interior of $B$ such that $D^2 \times \{i\} = D_i$. The resulting 1–*handle stabilisation of* $\Sigma$ is defined as

$$\Sigma' = (\Sigma \cap (S^4 \setminus B)) \cup_L (S^1 \times [0, 1]).$$

If there exists a path $\gamma$ in $\Sigma$ between $(1, 0)$ and $(1, 1)$ in $S^1 \times [0, 1]$ such that $\gamma \cup (\{1\} \times [0, 1])$ is a null-homotopic loop in $V$, via a null-homotopy $h \colon D^2 \to V$ with $h(\mathring{D}^2) \subseteq V \setminus \Sigma$, then we call the stabilisation *trivial*.

A detailed discussion of this construction in the locally flat setting can be found in [17, Proposition 9.1]. The next result describes the effect of trivial 1–handle stabilisation on the $\Lambda$–intersection form of surface exteriors.

**Lemma 6.2** *If* $\Sigma \subseteq N$ *is a* $\mathbb{Z}$–*surface and* $\Sigma' \subseteq N$ *is obtained from* $\Sigma$ *by a trivial* 1–*handle stabilisation, then*

$$\lambda_{N_{\Sigma'}} \cong \lambda_{N_\Sigma} \oplus \mathcal{H}_2.$$

**Proof** Since the 1–handle stabilisation is trivial, one can write

$$(N, \Sigma') = (N, \Sigma) \# (S^4, T^2),$$

where $T^2 \subseteq S^4$ denotes a standardly embedded torus, and $\#$ denotes the interior connect sum. It follows that $N_{\Sigma'} = N_\Sigma \cup S^4_{T^2}$, where the identification takes place along thickened meridians: $\mu_\Sigma \times D^2$ is identified with $\mu_{T^2} \times D^2$. One thus deduces the isomorphism $\pi_1(N_{\Sigma'}) \cong \pi_1(N_\Sigma) = \mathbb{Z}$. As the coefficient system maps these meridians to $1 \in \mathbb{Z}$, a straightforward Mayer–Vietoris argument shows that

$$H_2(N_{\Sigma'}; \Lambda) = H_2(N_\Sigma; \Lambda) \oplus H_2(S^4_{T^2}; \Lambda),$$

noting that $H_1(\mu_\Sigma; \Lambda) = 0$. It then follows that $\lambda_{N_{\Sigma'}} = \lambda_{N_\Sigma} \oplus \lambda_{S^4_{T^2}}$. The result is now a consequence of Lemma 6.1. $\qquad\square$

The next proposition shows that the $\Lambda$–intersection forms of any two $\mathbb{Z}$–surfaces exteriors become isometric after adding sufficiently many $\mathcal{H}_2$ summands.

**Proposition 6.3** *For $\mathbb{Z}$–surfaces $\Sigma_0, \Sigma_1 \subseteq N$ of the same genus with boundary $K$, there exists an integer $n \geq 0$ and an isometry*

$$\lambda_{N_{\Sigma_0}} \oplus \mathcal{H}_2^{\oplus n} \cong \lambda_{N_{\Sigma_1}} \oplus \mathcal{H}_2^{\oplus n}.$$

**Proof** Any two null-homologous surfaces in a 4–manifold can be made isotopic by enough 1–handles stabilisations to each surface [2, Theorem 5]. While Baykur and Sunukjian prove this result in the smooth category for surfaces in closed manifolds, it also applies in the topological category for properly embedded surfaces in 4–manifolds with boundary; see Theorem A.1 for a detailed proof in the case at hand.

We apply this to $\Sigma_0, \Sigma_1$. This is possible because $\mathbb{Z}$–surfaces are null-homologous by Lemma 5.1. Since $\pi_1(\partial(\nu\Sigma_i)) \to \pi_1(N_{\Sigma_i}) = \mathbb{Z}$ is surjective, all 1–handle stabilisations can be taken to be trivial [2, Lemma 3]. We deduce that after sufficiently many (say $n$) trivial 1–handle stabilisations, $\Sigma_0$ becomes isotopic to $\Sigma_1$, also stabilised $n$ times. Using Lemma 6.2, each such stabilisation adds an $\mathcal{H}_2$–summand to the $\Lambda$–intersection form of the corresponding surface exterior, from which the result follows. $\square$

## 6.2 $\mathbb{Z}$–surfaces for Alexander polynomial one knots

We now restrict our attention to $\mathbb{Z}$–surfaces with boundary Alexander polynomial one knots. An Alexander polynomial one knot $K \subseteq S^3$ bounds a disc $D \subseteq D^4$ with $\pi_1(D^4 \setminus D) = \mathbb{Z}$; see [15]. We call this disc *a $\mathbb{Z}$–slice disc of $K$*. When $S^3 = \partial N$ with $N = X \setminus \mathring{D}^4$, we arrange that this disc $D$ belong to a collar neighbourhood $S^3 \times [0, 1] \subseteq N$ of $S^3 = \partial N$. We use this disc to build a genus $g$ surface with boundary $K$.

**Definition 6.4** A *genus $g$ target surface* $\Sigma_g^t$ for an Alexander polynomial one knot $K$ is an embedded surface obtained from a $\mathbb{Z}$–disc $D$ of $K$ by $g$ trivial 1–handle stabilisations.

Although we do not need this fact in the sequel, an Alexander polynomial one knot $K \subseteq S^3$ bounds a *unique* $\mathbb{Z}$–disc: this follows either from [10, Theorem 1.2] or from Theorem 5.8 above, because $H_2(D_D^4; \Lambda) = 0$. This is why we do not keep track of the $\mathbb{Z}$–disc $D$ in the notation for target surfaces.

The next lemma describes the $\Lambda$–intersection form of a target surface exterior.

**Lemma 6.5** *If $D \subseteq S^3 \times [0,1] \subseteq N$ is a $\mathbb{Z}$–disc, then $\lambda_{N_D} = Q_X$. If $\Sigma_g^t$ is a target surface, then $\lambda_{N_{\Sigma_g^t}} \cong Q_X \oplus \mathcal{H}_2^{\oplus g}$.*

**Proof** Since $D$ is properly embedded in a collar neighbourhood of $\partial N$, $N_D$ is homeomorphic to the interior connected sum $N_D \cong D_D^4 \# X$. The exterior $D_D^4$ of a $\mathbb{Z}$–slice disc is aspherical, and therefore we have $H_i(D_D^4; \Lambda) = 0$ for $i > 0$ [10, Lemma 2.1]. The first assertion now follows from a straightforward Mayer–Vietoris argument. Since a target surface is obtained from a $\mathbb{Z}$–slice disc by $g$ handle additions, the second assertion follows from the first and Lemma 6.2. □

Finally, we describe the $\Lambda$–intersection form of a $\mathbb{Z}$–surface exterior for an Alexander polynomial one knot $K$.

**Corollary 6.6** *If $\Sigma \subseteq N = X \setminus \mathring{D}^4$ is a genus $g$ $\mathbb{Z}$–surface for an Alexander polynomial one knot $K$, then $\lambda_{N_\Sigma} \oplus \mathcal{H}_2^{\oplus n} \cong Q_X \oplus \mathcal{H}_2^{\oplus(n+g)}$ for some integer $n \geq 0$.*

**Proof** Since $K$ has Alexander polynomial one, it bounds a genus $g$ target surface $\Sigma_g^t$. Proposition 6.3 ensures that $\lambda_{N_\Sigma} \oplus \mathcal{H}_2^{\oplus n}$ is isometric to $\lambda_{N_{\Sigma_g^t}} \oplus \mathcal{H}_2^{\oplus n}$ for some $n \geq 0$. The result now follows from Lemma 6.5, thanks to which $\lambda_{N_{\Sigma_g^t}} = Q_X \oplus \mathcal{H}_2^{\oplus g}$. □

# 7 Knotted surfaces in $S^4$ and $D^4$

As we outlined in Section 1.4, our results about surfaces in $D^4$ and $S^4$ can be deduced from Theorem 5.8 once we prove that the surface exteriors have isometric equivariant intersection forms. In Section 7.1, we show that if $\Sigma \subseteq D^4$ is a genus $g \geq 3$ $\mathbb{Z}$–surface, with boundary an Alexander polynomial one knot $K$, then $\lambda_{D_\Sigma^4} \cong \mathcal{H}_2^{\oplus g}$. In Section 7.2 we then combine this result with Theorem 5.8 to deduce our results about surfaces in $D^4$ and $S^4$, in particular proving Theorems 1.1 and 1.2. In Section 7.3, we apply these results to study Seifert surfaces pushed into $D^4$.

## 7.1 Intersection forms in $D^4$

The aim of this subsection is to show that if $\Sigma \subseteq D^4$ is a $\mathbb{Z}$–surface of genus $g \geq 3$ with $\Delta_{\partial \Sigma} \doteq 1$, then $\lambda_{D_\Sigma^4} \cong \mathcal{H}_2^{\oplus g}$.

Before we continue, we need to recall some general theory on $\varepsilon$–Hermitian forms, and on their $\varepsilon$–quadratic counterparts; further details can be found in [47, Section 1.1]

and [12, Section 2], where the terminology "$\varepsilon$–symmetric" is used, as is customary in $L$–theory. Let $R$ be a ring with involution and let $\varepsilon \in R$ be a central unit with $\varepsilon\bar{\varepsilon} = 1$. Let $M$ be a finitely generated free left $R$–module. Let $\mathrm{Sesq}(M)$ denote the abelian group of $R$–sesquilinear forms on $M$, meaning that $b(rx, sy) = r\,b(x, y)\bar{s}$ for $b \in \mathrm{Sesq}(M)$, $x, y \in M$, and $r, s \in R$. Define an involution

$$T_\varepsilon \colon \mathrm{Sesq}(M) \to \mathrm{Sesq}(M), \quad (T_\varepsilon b)(y, x) = \varepsilon\overline{b(x, y)}.$$

A quick check shows that $(T_\varepsilon^2 b)(x, y) = \varepsilon b(x, y)\bar{\varepsilon} = \varepsilon\bar{\varepsilon}b(x, y) = b(x, y)$, so the conditions that $\varepsilon$ is central and that $\varepsilon\bar{\varepsilon} = 1$ are crucial for $T_\varepsilon$ to be an involution. This enables us to define the *symmetric Q–group* $Q^\varepsilon(M)$ of $M$ and the *quadratic Q–group* $Q_\varepsilon(M)$ of $M$ via the exact sequence

$$0 \to Q^\varepsilon(M) \to \mathrm{Sesq}(M) \xrightarrow{\;1-T_\varepsilon\;} \mathrm{Sesq}(M) \to Q_\varepsilon(M) \to 0.$$

In other words,

$$Q^\varepsilon(M) := \ker(1 - T_\varepsilon) \quad \text{and} \quad Q_\varepsilon(M) := \mathrm{coker}(1 - T_\varepsilon).$$

A sesquilinear form $b \colon M \times M \to R$ is called $\varepsilon$–*Hermitian* if it belongs to $Q^\varepsilon(M)$, that is if $b(y, x) = \varepsilon\overline{b(x, y)}$ for every $x, y \in M$. A $(+1)$–Hermitian form is a Hermitian form in the usual sense, while a $(-1)$–Hermitian form is a skew-Hermitian form in the usual sense. A $\varepsilon$–Hermitian form is called *hyperbolic* if it is isometric to

$$H^\varepsilon(R)^{\oplus g} = \left( R^{2g}, \begin{pmatrix} 0 & 1 \\ \varepsilon & 0 \end{pmatrix}^{\oplus g} \right) \quad \text{for some } g \geq 0.$$

The reason for introducing this terminology is that we will shortly be concerned with $(-t)$–Hermitian forms over $\Lambda$, and their quadratic analogues.

An element $\psi \in Q_\varepsilon(M)$ is called an $\varepsilon$–*quadratic form*. To an $\varepsilon$–quadratic form in $Q_\varepsilon(M)$ is associated its *symmetrisation* $\varphi := (1 + T_\varepsilon)(\psi) \in Q^\varepsilon(M)$. Given $\varphi \in Q^\varepsilon(M)$, a quadratic form $\psi$ with $(1 + T_\varepsilon)(\psi) = \varphi$ is called a *quadratic refinement of* $\varphi$. The symmetrisation is well-defined on equivalence classes in $Q_\varepsilon(M)$ because $(1 + T_\varepsilon)(1 - T_\varepsilon) = 1 - T_\varepsilon^2 = 0$. Note that quadratic forms are considered up to addition of forms in the image of $1 - T_\varepsilon$. A quadratic form over $R$ is *hyperbolic* if it is isometric to $H_\varepsilon(R)^{\oplus g} = \left( R^{2g}, \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)\right)^{\oplus g}$ for some $g \geq 0$. Let us emphasise that subscripts denote quadratic forms, while superscripts denote $\varepsilon$–Hermitian forms.

Here is the key relationship between $(-t)$–Hermitian and $(-t)$–quadratic forms that we shall exploit in the proof of Theorem 7.4 below.

**Lemma 7.1** *Let $M$ be a finitely generated free $\Lambda$–module. Then the map*

$$(1 + T_{-t}): Q_{-t}(M) \to Q^{-t}(M)$$

*is injective.*

**Proof** To prove the lemma we will show the following facts:

(i) For $\varphi \in Q^{-t}(M)$, the assignment $\varphi \mapsto (t^{-1} - 1)\varphi$ determines a map

$$\Omega: Q^{-t}(M) \to Q^{+1}(M).$$

(ii) For $\psi \in \mathrm{Sesq}(M)$, the assignment $\psi \mapsto (t^{-1} - 1)\psi$ determines an injective map

$$\Xi: Q_{-t}(M) \to Q_{+1}(M).$$

(iii) The map $(1 + T_{+1}): Q_{+1}(M) \to Q^{+1}(M)$ is injective.

(iv) The diagram

$$
\begin{array}{ccc}
Q_{-t}(M) & \xrightarrow{\ \Xi\ } & Q_{+1}(M) \\
{\scriptstyle 1+T_{-t}}\big\downarrow & & \big\downarrow{\scriptstyle 1+T_{+1}} \\
Q^{-t}(M) & \xrightarrow{\ \Omega\ } & Q^{+1}(M)
\end{array}
$$

commutes.

Then, since the right–down composition is injective, it follows that the left vertical map is injective, as desired. So it suffices to prove these four assertions.

For (i), suppose that $\varphi \in Q^{-t}(M) = \ker(1 - T_{-t})$, that is $\varphi = -t\bar{\varphi}^T$. Then

$$\overline{(t^{-1} - 1)\varphi}^T = (t - 1)\bar{\varphi}^T = (t^{-1} - 1)(-t)\bar{\varphi}^T = (t^{-1} - 1)\varphi.$$

Therefore $\Omega(\varphi) = (t^{-1} - 1)\varphi \in \ker(1 - T_{+1}) = Q^{+1}(M)$, as desired.

To prove (ii), first we show that $\Xi: Q_{-t}(M) \to Q_{+1}(M)$ is well-defined. We need to see that for every sesquilinear form $\theta \in \mathrm{Sesq}(M)$, the element $(1 - T_{-t})(\theta) = \theta - (-t)\bar{\theta}^T = \theta + t\bar{\theta}^T$ maps to the trivial element in $Q_{+1}(M)$. This is indeed the case since

$$(t^{-1} - 1)(\theta + t\bar{\theta}^T) = (t^{-1} - 1)\theta - (t - 1)\bar{\theta}^T = (1 - T_{+1})((t^{-1} - 1)\theta)$$

is trivial in $\mathrm{coker}(1 - T_{+1}) = Q_{+1}(M)$. Therefore $\Xi$ is well-defined.

Now we prove that $\Xi: Q_{-t}(M) \to Q_{+1}(M)$ is injective. Let $B$ and $C$ be two $(-t)$–quadratic forms in $Q_{-t}(M)$, and suppose that $Z := (t^{-1} - 1)(B - C) = 0 \in Q_{+1}(M)$. That is, $Z \in \mathrm{im}(1 - T_{+1})$, so, choosing a basis for $M \cong \Lambda^n$, there exists a matrix $X$ over $\Lambda$ such that

$$(t^{-1} - 1)(B - C) = Z = X - \bar{X}^T.$$

We assert that there exists a matrix $Y$ over $\Lambda$ such that

$$Z = (t^{-1} - 1)Y - (t - 1)\overline{Y}^T.$$

Write the entries of the matrix $Z = X - \overline{X}^T$ as $z_{ij}$. Since $X - \overline{X}^T$ is skew-Hermitian, we have $\overline{z}_{ij} = -z_{ji}$. Also, $z_{ij}$ is divisible by $t^{-1} - 1$ for every $i, j$. For $i < j$, define the entries $y_{ij}$ of $Y$ by the formula $y_{ij} := z_{ij}/(t^{-1} - 1)$. For $i > j$ define $y_{ij} := 0$. This way $(t^{-1} - 1)y_{ij} - (t - 1)\overline{y}_{ji} = z_{ij}$ for $i \neq j$. It remains to define the diagonal entries of $Y$. For each $i$, since the Laurent polynomial $z_{ii}$ satisfies $z_{ii} = -\overline{z}_{ii}$, we have $z_{ii}(-1) = -z_{ii}(-1)$, and so $z_{ii}(-1) = 0$. Therefore $z_{ii}$ is divisible by $t + 1$. Since $t^{-1} - 1$ and $(t + 1)$ are coprime, we deduce that $z_{ii} = (t + 1)(t^{-1} - 1)q$ for some polynomial $q$ that satisfies $q = \overline{q}$. Now set $y_{ii} := q$. We compute

$$(t^{-1} - 1)y_{ii} - (t - 1)\overline{y}_{ii} = (t^{-1} - 1 - (t - 1))q = (t^{-1} - t)q = (t + 1)(t^{-1} - 1)q = z_{ii}.$$

This concludes the proof of the assertion that $Z = (t^{-1} - 1)Y - (t - 1)\overline{Y}^T$ for some matrix $Y$. Using this assertion, as well as the definitions of $X$ and $Y$, we obtain

$$Z = (t^{-1} - 1)(B - C) = X - \overline{X}^T = (t^{-1} - 1)Y - (t - 1)\overline{Y}^T = (t^{-1} - 1)(Y + t\overline{Y}^T).$$

This implies that $B - C = Y + t\overline{Y}^T = Y - (-t)\overline{Y}^T$, which is zero in $Q_{-t}(M)$. Therefore the map $\Xi \colon Q_{-t}(M) \to Q_{+1}(M)$ is injective, as desired. This completes the proof of (ii).

Next we prove (iii), that $1 + T_{+1} \colon Q_{+1}(M) \to Q^{+1}(M)$ is injective. For an arbitrary finitely generated free module $M$, the map $1 + T_\varepsilon$ fits into the exact sequence

$$0 \to \hat{Q}^{-\varepsilon}(M) \to Q_\varepsilon(M) \xrightarrow{1 + T_\varepsilon} Q^\varepsilon(M).$$

The group $\hat{Q}^{-\varepsilon}(M) := \ker(1 + T_\varepsilon)$ is a *hyperquadratic $Q$–group*; its purpose is to measure the difference between Hermitian and quadratic forms [47]. We have

$$(17) \quad \hat{Q}^{-\varepsilon}(M) = \ker\left(1 + T_\varepsilon \colon Q_\varepsilon(M) \to Q^\varepsilon(M)\right) = \frac{\ker(1 + T_\varepsilon)}{\operatorname{im}(1 - T_\varepsilon)} = \frac{\ker(1 - T_{-\varepsilon})}{\operatorname{im}(1 + T_{-\varepsilon})}.$$

We shall use the penultimate description, but include the last equation to show why the $-\varepsilon$ appears in the notation for the group $\hat{Q}^{-\varepsilon}(M)$. Writing $M = \Lambda^n$, there is an isomorphism $\hat{Q}^{-\varepsilon}(M) = \hat{Q}^{-\varepsilon}(\Lambda^n) \cong \hat{Q}^{-\varepsilon}(\Lambda)^n$ [12, Remark 3.4]. By (17) with $\varepsilon = +1$ and $M = \Lambda$, we have $\hat{Q}^{-1}(\Lambda) = \ker(1 + T_{+1})/\operatorname{im}(1 - T_{+1})$. For $p(t) \in \Lambda$, if $p(t) + p(t^{-1}) = 0$, then $p(t)$ must be of the form $p(t) = r(t) - r(t^{-1})$ for some $r(t) \in \Lambda$. It follows that $\hat{Q}^{-1}(\Lambda) = 0$. Thus $\ker(1 + T_{+1} \colon Q_{+1}(M) \to Q^{+1}(M)) = 0$ and so $1 + T_{+1}$ is indeed injective. This concludes the proof of the third item of the lemma.

For (iv), let $\psi \in Q_{-t}(M)$. Then

$$(1 + T_{+1}) \circ \Xi(\psi) = (1 + T_{+1})(t^{-1} - 1)\psi = (t^{-1} - 1)\psi + (t - 1)\overline{\psi}^T$$
$$= (t^{-1} - 1)(\psi - t\overline{\psi}^T) = \Omega \circ (1 + T_{-t})(\psi),$$

so the diagram commutes, which proves (iv) and completes the proof that the map $(1 + T_{-t}): Q_{-t}(M) \to Q^{-t}(M)$ is injective. $\qquad\square$

Next we state the algebraic cancellation result that we will use. The idea behind the proof of Theorem 7.4 below will be to engineer a situation in which we can apply cancellation.

**Definition 7.2** The *Witt index* $\mathrm{ind}(H, \theta)$ of an $\varepsilon$–quadratic form is the largest integer $k$ such that a subform of $(H, \theta)$ is isometric to $H_\varepsilon(R)^{\oplus k}$.

**Proposition 7.3** *Let $\varepsilon \in \Lambda$ be a central unit with $\varepsilon\overline{\varepsilon} = 1$, and let $(H, \theta)$ and $(H', \theta')$ be $\varepsilon$–quadratic forms over $\Lambda$. Assume that for some $n \geq 0$ there is an isometry*

$$(18) \qquad\qquad (H, \theta) \oplus H_\varepsilon(\Lambda)^{\oplus n} \cong (H', \theta') \oplus H_\varepsilon(\Lambda)^{\oplus n}.$$

*If $\mathrm{ind}(H', \theta') \geq 3$, then there is an isometry*

$$(19) \qquad\qquad (H, \theta) \cong (H', \theta').$$

**Proof** We apply a result due to Bass [1, Corollary IV.3.6]. Given a ring $R$ with involution, we write $R_0 = \{\sum_i x_i \overline{x}_i \mid x_i \in R\}$ for the norm subring of $R$, as well as $\mathrm{maxspec}(R_0)$ for the set of all maximal ideals of $R_0$ under the Zariski topology, and $d_R := \dim \mathrm{maxspec}(R_0)$. Detailed definitions of these notions are irrelevant: we need only know that $d_\Lambda = 2$; see eg [29, Proposition 2.2] or [21, page 439]. Let $(H, \theta)$ and $(H', \theta')$ be quadratic forms over $R$. If there is an isometry

$$(H, \theta) \oplus H_\varepsilon(R)^{\oplus n} \cong (H', \theta') \oplus H_\varepsilon(R)^{\oplus n},$$

and if $\mathrm{ind}((H', \theta') \oplus H_\varepsilon(R)) \geq d_R + 2$, then there is an isometry $(H, \theta) \cong (H', \theta')$ [1, Corollary IV.3.6]. In particular, this cancellation result holds if $\mathrm{ind}(H', \theta') \geq d_R + 1$. As mentioned above, for $R = \Lambda$, we have $d_\Lambda = 2$. Since

$$\mathrm{ind}((H', \theta') \oplus H_\varepsilon(\Lambda)) \geq 3 + 1 = d_\Lambda + 2,$$

the result of Bass says that (18) implies (19). $\qquad\square$

The next theorem is the main result of this subsection.

**Theorem 7.4** If $\Sigma \subseteq D^4$ is a $\mathbb{Z}$–surface of genus $g \geq 3$ whose boundary is an Alexander polynomial one knot $K \subseteq S^3$, then

$$\lambda_{D_\Sigma^4} \cong \mathcal{H}_2^{\oplus g}.$$

**Proof** We outline the strategy of the proof. Since the knot $K$ has Alexander polynomial one and $Q_{D^4} = 0$, by Corollary 6.6, there exists an integer $n \geq 0$ and an isometry

$$(20) \qquad\qquad \lambda_{D_\Sigma^4} \oplus \mathcal{H}_2^{\oplus n} \cong \mathcal{H}_2^{\oplus g+n}.$$

We cannot apply the cancellation result of Proposition 7.3 directly, since $\mathcal{H}_2^{\oplus n}$ is not hyperbolic over $\Lambda$. However, by switching the basis elements $\mathcal{H}_2$ is isometric to a multiple of the $(-t)$–hyperbolic form:

$$\begin{pmatrix} 0 & t^{-1} - 1 \\ t - 1 & 0 \end{pmatrix} = (t^{-1} - 1) \cdot \begin{pmatrix} 0 & 1 \\ -t & 0 \end{pmatrix} = (t^{-1} - 1)\mathcal{H}^{-t}(\Lambda).$$

If we could remove the $(t^{-1} - 1)$ factor, we could apply Proposition 7.3. The idea for this is to use the relative intersection pairing

$$\lambda_{D_\Sigma^4}^\partial : H_2(D_\Sigma^4; \Lambda) \times H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda) \to \Lambda$$

instead of the "absolute" intersection pairing $\lambda_{D_\Sigma^4}$. We will see, using (20) and appropriate bases, that the pairing $\lambda_{D_\Sigma^4}^\partial$ is represented by a $(-t)$–Hermitian matrix over $\Lambda$. We will then apply the cancellation result of Proposition 7.3 with $\varepsilon = -t$ to these forms, before deducing the desired conclusion on the original absolute pairings. Note that the adjoints of the relative and absolute pairings fit into the commutative diagram

$$(21)$$

$$H_2(D_\Sigma^4; \Lambda) \xrightarrow{\quad\widehat{\lambda}\quad} \overline{\mathrm{Hom}_\Lambda(H_2(D_\Sigma^4; \Lambda), \Lambda)}$$

$$\searrow q \qquad\qquad \nearrow \widehat{\lambda}^\partial$$

$$H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda)$$

For the first step of the proof we will choose bases for $H_2(D_\Sigma^4; \Lambda)$ and $H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda)$ and describe the map $q$ with respect to these bases. First, a short rank computation:

**Claim 4** $$H_2(D_\Sigma^4; \Lambda) \cong \Lambda^{2g}.$$

**Proof** Since $H_2(D_\Sigma^4; \Lambda)$ is free (by Lemma 3.2), we need only show that its rank is $b_1(\Sigma) = 2g$. Using $H_1(D_\Sigma^4) = \mathbb{Z}$, a Mayer–Vietoris argument shows that $b_2(D_\Sigma^4) = b_1(\Sigma)$ as well as $b_3(D_\Sigma^4) = 0$. Since we also have $b_4(D_\Sigma^4) = 0$ and $b_0(D_\Sigma^4) = 1$,

we get $\chi(D_\Sigma^4) = b_1(\Sigma)$. The Euler characteristic can also be computed with $Q = \mathbb{Q}(t)$ coefficients. By Lemma 3.2, we have $b_i^Q(D_\Sigma^4) = 0$ for $i \neq 2$, and therefore

$$\operatorname{rk}_\Lambda H_2(D_\Sigma^4; \Lambda) = \chi^Q(D_\Sigma^4) = b_1(\Sigma),$$

as asserted. □

**Claim 5** *There is a choice of bases with respect to which the homomorphism*

$$q: H_2(D_\Sigma^4; \Lambda) \to H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda)$$

*is represented by the matrix* $(t-1)\operatorname{Id}$.

**Proof** We first base $H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda)$ using the fact that the Alexander polynomial of $K$ is one. We have an exact sequence as part of the long exact sequence of the pair $(D_\Sigma^4, \partial D_\Sigma^4)$:

$$H_2(D_\Sigma^4; \Lambda) \xrightarrow{q} H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda) \xrightarrow{\delta} H_1(\partial D_\Sigma^4; \Lambda) \to 0.$$

Here $H_1(D_\Sigma^4; \Lambda) = 0$ because $\pi_1(D_\Sigma^4) \cong \mathbb{Z}$. In particular, the map $q$ presents the $\Lambda$–module $H_1(\partial D_\Sigma^4; \Lambda)$. Use Lemmas 3.2 and 5.5, Claim 4 and the fact that $\Delta_K \doteq 1$ to deduce that

$$H_2(D_\Sigma^4; \Lambda) \cong \Lambda^{2g} \cong H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda) \quad \text{and} \quad H_1(\partial D_\Sigma^4; \Lambda) \cong \bigoplus_{i=1}^{2g} \Lambda/(t-1).$$

Choose a set of generators $\{\gamma_k\} \subseteq H_1(\partial D_\Sigma^4; \Lambda)$, represented in the form

$$g_k \times \{\operatorname{pt}\} \subseteq \Sigma \times \{\operatorname{pt}\} \subseteq \Sigma \times S^1 \subseteq \partial D_\Sigma^4,$$

where $\{g_k\}$ is a symplectic basis of curves on $\Sigma$. Then choose a basis $\{x_i\}$ for $H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda)$ such that $\delta(x_i) = \gamma_i$ for $1 \leq i \leq 2g$. The homology classes $x_k$ can be represented by embedded surfaces $S_i \subseteq D_\Sigma^4$ with $\partial S_i = \gamma_i$ for $1 \leq i \leq 2g$ [17, Section 10.3].

Next, we base $H_2(D_\Sigma^4; \Lambda)$. For each $i = 1, \ldots, 2g$, define a closed surface representing a class $y_i \in H_2(D_\Sigma^4; \Lambda)$, as follows. Consider the torus $g_i \times S^1 = \partial(\overline{\nu}\Sigma|_{g_i})$, and surger it along $\gamma_i$ using the surface $S_i$. That is, remove an annular neighbourhood $g_i \times (p, p') \subseteq g_i \times S^1$ of $\gamma_i = g_i \times \{\operatorname{pt}\}$, leaving $B := g_i \times (S^1 \setminus (p, p'))$. Then glue $-S_i$ to $g_i \times \{p\}$, and a push-off $S_i'$ of $S_i$ to $g_i \times \{p'\}$. Call the resulting surface $T_i$, and set $y_i := [T_i] \in H_2(D_\Sigma^4; \Lambda)$ for $1 \leq i \leq 2g$. Then, for $i = 1, \ldots, 2g$, in $H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda)$ we have

$$(22) \qquad q(y_i) = q([T_i]) = [-S_i] + [B \cup S_i'] = -x_i + tx_i = (t-1)x_i.$$
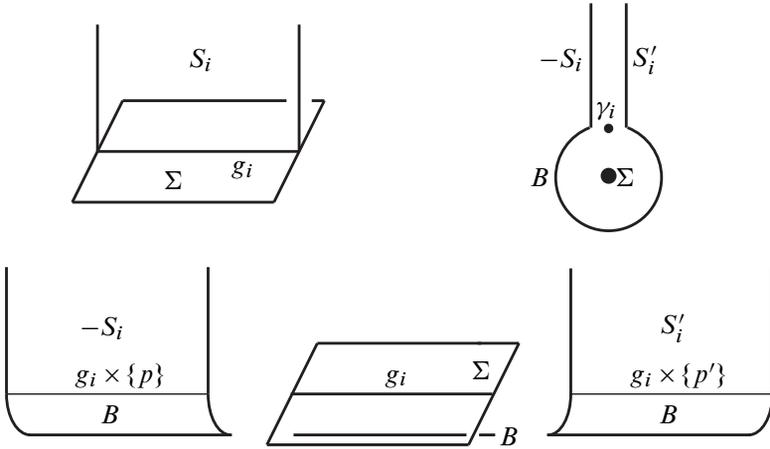
Figure 3: Representing $y_i \in H_2(D_\Sigma^4; \Lambda)$ by an immersed surface $T_i$ in $D_\Sigma^4$ obtained by surgering the torus $g_i \times S^1$, using a surface $S_i$ representing the relative homology class $x_i$. The top left picture shows a neighbourhood of a portion of $g_i$, intersected with a carefully chosen 3–dimensional subspace that contains the intersections of $\Sigma$ and $S_i$ with this neighbourhood. Bottom, a 4–dimensional picture of $T_i$, consisting of three 3–dimensional slices at certain special values of the fourth coordinate (time), in which $\Sigma$, $-S_i$, $S_i'$ and most of $B$ appear. The annulus $B$ joins $-S_i$ and $S_i'$ together; in intermediate time values only a line of $B$ appears, including in the middle slice. The top right picture shows a cross section of $T_i$, looking along $g_i$.

Here the $t$ arises because the annulus $B$ wraps around a meridian of $\Sigma$. This process is illustrated in Figure 3.

Consider the diagram of short exact sequences

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & \Lambda^{\oplus 2g} & \xrightarrow{(t-1)\,\mathrm{Id}} & \Lambda^{\oplus 2g} & \xrightarrow{\mathrm{proj}} & (\Lambda/(t-1))^{\oplus 2g} & \longrightarrow & 0 \\
& & \downarrow{y} & & \downarrow{x} & & \downarrow{\gamma} & & \\
0 & \longrightarrow & H_2(D_\Sigma^4; \Lambda) & \xrightarrow{q} & H_2(D_\Sigma^4, \partial D_\Sigma^4; \Lambda) & \longrightarrow & H_1(\partial D_\Sigma^4; \Lambda) & \longrightarrow & 0
\end{array}
$$

where the maps $y$ and $x$ send the canonical basis $\{e_i\}_{i=1}^{2g}$ of $\Lambda^{2g}$ to $\{y_i\}_{i=1}^{2g}$ and $\{x_i\}_{i=1}^{2g}$, respectively, and $\gamma$ maps the $[e_i]$ to the $\gamma_i$. This diagram commutes thanks to (22). By construction $x$ and $\gamma$ are isomorphisms, and therefore so is $y$ by the five lemma. Thus $\{y_i\}_{i=1}^{2g}$ forms a basis of $H_2(D_\Sigma^4; \Lambda)$ and $q$ is represented by $(t-1)\,\mathrm{Id}$. This completes the proof of Claim 5.                                                                              □

Recall from (20) that there is an isometry $\alpha\colon \lambda_{D_\Sigma^4} \oplus \mathcal{H}_2^{\oplus n} \cong \mathcal{H}_2^{\oplus(g+n)}$. Denote the matrices of $\lambda_{D_\Sigma^4}$ and $\lambda^\partial_{D_\Sigma^4}$ with respect to the previously described bases by $L$ and $L^\partial$, respectively. Also represent the isometry $\alpha$ by a matrix $A$ with respect to these bases. Therefore we have a congruence

$$(23) \qquad A^T\left(L \oplus \begin{pmatrix} 0 & t-1 \\ t^{-1}-1 & 0 \end{pmatrix}^{\oplus n}\right)\overline{A} = \begin{pmatrix} 0 & t-1 \\ t^{-1}-1 & 0 \end{pmatrix}^{\oplus(g+n)}.$$

Our aim is now to factor out a $(t^{-1}-1)$ from both sides of this equation. Recall from (21) that the adjoints of the pairings $\lambda_{D_\Sigma^4}$ and $\lambda^\partial_{D_\Sigma^4}$ are related by $\hat{\lambda}^\partial_{D_\Sigma^4} \circ q = \hat{\lambda}_{D_\Sigma^4}$. With respect to our bases (and recalling (22)), for every $x, y \in H_2(D_\Sigma^4; \Lambda)$ we therefore obtain

$$x^T L^\partial (t^{-1}-1)\overline{y} = x^T L^\partial \overline{(t-1)y} = \lambda^\partial_{D_\Sigma^4}(x, q(y)) = \hat{\lambda}^\partial(q(y))(x)$$
$$= \hat{\lambda}(y)(x) = \lambda(x, y)$$
$$= x^T L \overline{y}.$$

It follows that $L^\partial(t^{-1}-1) = L$. Combine this with

$$\mathcal{H}_2 = \begin{pmatrix} 0 & t-1 \\ t^{-1}-1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}(t^{-1}-1)$$

to rewrite (23) as

$$A^T\left(L^\partial(t^{-1}-1) \oplus \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}(t^{-1}-1)\right)\overline{A} = \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus g+n}(t^{-1}-1).$$

As $(t^{-1}-1)\cdot \mathrm{Id}$ is diagonal with constant diagonal coefficients, it is central and thus we obtain

$$(t^{-1}-1)\cdot A^T\left(L^\partial \oplus \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus n}\right)\overline{A} = (t^{-1}-1)\begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus(g+n)}.$$

As $(t^{-1}-1)\cdot \mathrm{Id}$ is nondegenerate over $\Lambda$, it follows that

$$(24) \qquad A^T\left(L^\partial \oplus \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus n}\right)\overline{A} = \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus(g+n)}.$$

In particular $L^\partial$ stabilises, via $(-t)$–Hermitian matrices, to a $(-t)$–Hermitian form, so it follows that $L^\partial$ is itself $(-t)$–Hermitian.

In order to apply cancellation, we refine (24) to a statement about quadratic forms. Set

$$\Psi := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Then $(1 + T_{-t})\Psi = \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}$, so $\Psi$ is a quadratic refinement for $\begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}$. Therefore $\Psi^{\oplus(g+n)}$ is a quadratic refinement for $\begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus(g+n)}$, which is the right-hand side of (24). Thus $\Psi^{\oplus(g+n)}$ is a quadratic refinement for

$$A^T \left( L^\partial \oplus \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus n} \right) \bar{A}.$$

Rearranging, we deduce that $A^{-T}\Psi^{\oplus(g+n)}\bar{A}^{-1}$ is a quadratic refinement for the $(-t)$–Hermitian matrix $L^\partial \oplus \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus n}$. Now, corresponding to this decomposition, write

$$A^{-T}\Psi^{\oplus(g+n)}\bar{A}^{-1} = \begin{pmatrix} \Psi^\partial & U \\ V & \Psi_{\mathrm{ind}} \end{pmatrix},$$

where this defines the block entries on the right-hand side. Since this is a quadratic refinement of the block diagonal $(-t)$–Hermitian matrix

$$L^\partial \oplus \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus n},$$

we deduce that $V - t\bar{U}^T = 0$ and therefore

$$(25) \quad A^{-T}\Psi^{\oplus(g+n)}\bar{A}^{-1} = \begin{pmatrix} \Psi^\partial & 0 \\ 0 & \Psi_{\mathrm{ind}} \end{pmatrix} + \begin{pmatrix} 0 & U \\ t\bar{U}^T & 0 \end{pmatrix} = \Psi^\partial \oplus \Psi_{\mathrm{ind}} \in Q_{-t}(\Lambda^{\oplus 2(g+n)}).$$

This procedure therefore endows both the $(-t)$–Hermitian matrices $L^\partial$ and $\begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus n}$ with quadratic refinements, which we denoted by $\Psi^\partial$ and $\Psi_{\mathrm{ind}}$ respectively. Rearranging (25) we obtain the following isometry of quadratic forms:

$$(26) \qquad\qquad A^T(\Psi^\partial \oplus \Psi_{\mathrm{ind}})\bar{A} = \Psi^{\oplus(g+n)} \in Q_{-t}(\Lambda^{\oplus 2(g+n)}).$$

In order to apply cancellation, we need $\Psi_{\mathrm{ind}}$ to agree with the quadratic refinement $\Psi^{\oplus n}$ of $\begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus n}$ in the $Q$–group $Q_{-t}(\Lambda^{2n})$. The fact that both are quadratic refinements of the form $\begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus n}$ in $Q^{-t}(\Lambda^{2n})$ can be restated as

$$(1 + T_{-t})(\Psi_{\mathrm{ind}}) = (1 + T_{-t})(\Psi^{\oplus n}) \in Q^{-t}(\Lambda^{2n}).$$

By Lemma 7.1, $(1 + T_{-t})$ is injective, and so $\Psi_{\mathrm{ind}} = \Psi^{\oplus n}$ in $Q_{-t}(\Lambda^{2n})$, as desired.

We deduce from (26) that

$$(27) \qquad\qquad A^T(\Psi^\partial \oplus \Psi^{\oplus n})\bar{A} = \Psi^{\oplus(g+n)} \in Q_{-t}(\Lambda^{2(g+n)}).$$

Since

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

we see that $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ is congruent to the hyperbolic $(-t)$–quadratic form $H_{-t}(\Lambda)$. We have established the following isometry of $(-t)$–quadratic forms:

$$(28) \qquad \Psi^{\partial} \oplus H_{-t}(\Lambda)^{\oplus n} \cong H_{-t}(\Lambda)^{\oplus g} \oplus H_{-t}(\Lambda)^{\oplus n} \in Q_{-t}(\Lambda^{2(g+n)}).$$

Since $g \geq 3$, we now apply the cancellation result of Proposition 7.3 to (28) and symmetrise the result to deduce that $L^{\partial}$ is congruent to $H^{-t}(\Lambda)^{\oplus g}$, and therefore to $\begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus g}$. Multiplying both sides of this congruence by $t^{-1} - 1$ shows that $L$ is congruent to $\mathcal{H}_2^{\oplus g}$. This concludes the proof of Theorem 7.4. $\qquad \square$

## 7.2 Surfaces in $S^4$ and $D^4$

We collect and prove our results concerning (rel boundary) ambient isotopy of surfaces in $D^4$ and $S^4$. Since the proofs of these results were already outlined in the introduction or elsewhere, we proceed swiftly.

**Theorem 7.5** Let $\Sigma_0, \Sigma_1 \subseteq D^4$ be $\mathbb{Z}$–surfaces of genus $g$ with boundary $K$.

(i) *Suppose there is an isometry $F \in \mathrm{Iso}(\lambda_{D_{\Sigma_0}^4}, \lambda_{D_{\Sigma_1}^4})$ and write $\partial F = h_K \oplus h_{\Sigma}$. If $h_K$ is represented by an orientation-preserving homeomorphism that is the identity on $\partial E_K$, then $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic rel boundary.*

(ii) *If $\Delta_K = 1$, $g = 1, 2$ and $\lambda_{D_{\Sigma_0}^4} \cong \lambda_{D_{\Sigma_1}^4}$, then $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic rel boundary.*

(iii) *If $\Delta_K = 1$ and $g \neq 1, 2$, then $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic rel boundary.*

**Proof** The first assertion is the second item of Theorem 5.8. When $\Delta_K = 1$, we have $H_1(E_K; \Lambda) = 0$, so $h_K = \mathrm{Id}$ is automatic, and the second assertion follows from the first. For the third assertion, additionally use that $\lambda_{D_{\Sigma_0}^4} \cong \lambda_{D_{\Sigma_1}^4}$ when $g \neq 1, 2$ by Theorem 7.4. $\qquad \square$

**Theorem 7.6** Let $\Sigma_0, \Sigma_1 \subseteq S^4$ be closed $\mathbb{Z}$–surfaces of genus $g$.

(i) *For $g = 1, 2$, if $\lambda_{\Sigma_0} \cong \lambda_{\Sigma_1}$, then $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic.*

(ii) *For $g \neq 1, 2$, the surfaces $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic.*

**Proof** Both assertions follow from Theorem 7.5 by removing a $(D^4, D^2)$–pair from $(S^4, \Sigma_i)$. $\qquad \square$

### 7.3 Pushed-in Seifert surfaces for Alexander polynomial one knots

We prove Theorem 1.9 from the introduction, which states that any two pushed-in Seifert surfaces of the same genus for an Alexander polynomial one knot are topologically ambiently isotopic in $D^4$. Recall that a *sublagrangian* for a nonsingular $\varepsilon$–Hermitian form $(H, \lambda)$ is a direct summand $M \subseteq H$ such that $\lambda(M \times M) = 0$, ie $M \subseteq M^\perp$. A sublagrangian is a *Lagrangian* if $M = M^\perp$. For a nonsingular $\varepsilon$–quadratic form $(H, \lambda, \mu)$, the definitions are identical, with the additional requirement that $\mu(M) = 0$. The following lemma is known to those familiar with L–theory. For instance, given a sublagrangian $i \colon M \hookrightarrow H$, the result follows quite promptly from [46, Proposition 2.2], provided one assumes that $i^* \widehat{\lambda} \colon H \to M^*$ is surjective. Since our argument is elementary (Ranicki's proof is more involved because his statement is more general) and since most of our work goes into establishing that $i^* \widehat{\lambda}$ is surjective, we include a proof for the readers' convenience.

**Lemma 7.7** *Let* $(H, \lambda)$ *be a nonsingular $\varepsilon$–Hermitian form on a free $\Lambda$–module $H$. If* $M \subseteq H$ *is a half-rank sublagrangian, then $M$ is a Lagrangian.*

**Proof** We know that $M \subseteq M^\perp$ and must show that $M = M^\perp$. Equivalently, we must show that $M/M^\perp = 0$. We will show that $M/M^\perp$ is both torsion and torsion-free.

**Claim 6** *The inclusion $M^\perp \hookrightarrow H$ is split and $M^\perp \subseteq H$ is a half-rank summand.*

**Proof** Consider the exact sequence $0 \to M^\perp \to H \xrightarrow{\iota^* \widehat{\lambda}} M^*$. It now suffices to show that $\iota^* \widehat{\lambda}$ is surjective: since $M^*$ is free and $\mathrm{rk}_\Lambda(M^*) = \mathrm{rk}_\Lambda(M) = \frac{1}{2} \mathrm{rk}_\Lambda(H)$, the result would then follow. Since $\lambda$ is nonsingular, $\widehat{\lambda}$ is surjective and so it suffices to prove that $\iota^*$ is injective. This occurs if and only if $\mathrm{Ext}^1_\Lambda(H/M, \Lambda) = 0$. But since $M$ is a summand, it follows that $H \cong M \oplus H/M$. Since $H$ is free, $H/M$ is projective over $\Lambda$ and therefore free over $\Lambda$ [38, Chapter V, Corollary 4.12]. Thus $\mathrm{Ext}^1_\Lambda(H/M, \Lambda) = 0$, so $\iota^* \widehat{\lambda}$ is surjective and thus $M^\perp \hookrightarrow H$ is a split injection, as claimed. ☐

The fact that $M/M^\perp$ is torsion follows from the claim because $M^\perp \subseteq M$ and $\mathrm{rk}_\Lambda(M^\perp) = \mathrm{rk}_\Lambda(M)$. We now show that $M/M^\perp$ is torsion-free. Since submodules of torsion-free modules are torsion-free, it suffices to show that $H/M^\perp \supseteq M/M^\perp$ is torsion-free. But now by the claim, we know that $H/M^\perp$ is (isomorphic to) a submodule of the free module $H \cong M^\perp \oplus H/M^\perp$ and so it is indeed torsion-free. This concludes the proof of the lemma. ☐

Given a Seifert matrix $A$ for a genus $g$ Seifert surface $S$, we write $(H, A) := (\mathbb{Z}^{2g}, A)$. Use $F \subseteq D^4$ to denote the result of pushing $S$ into $D^4$. After some choice of bases, the equivariant intersection form $(H_2(D_F^4; \Lambda), \lambda_{D_F^4})$ is isometric to

$$(H \otimes_{\mathbb{Z}} \Lambda, (1-t)A^T + (1-t^{-1})A).$$

See [35, Section 3], also [8, proof of Lemma 5.4].

**Remark 7.8**  $(H \otimes_{\mathbb{Z}} \Lambda, A - tA^T, A)$ defines a $(-t)$–quadratic form. Furthermore, if $A$ is a Seifert matrix for an Alexander polynomial one knot $K$, then this quadratic form is nonsingular since $\det(A - tA^T) \doteq \Delta_K(t) \doteq 1$.

Recall that a *metaboliser* for a Seifert matrix $(\mathbb{Z}^{2g}, A)$ is a half-rank direct summand $L \subseteq \mathbb{Z}^{2g}$ such that $x^T A y = 0$ for all $x, y \in L$. Since Alexander polynomial one knots are (in particular) algebraically slice, their Seifert matrices admit metabolisers.

**Proposition 7.9**  *Let $(H, A) := (\mathbb{Z}^{2g}, A)$ be a Seifert matrix for an Alexander polynomial one knot $K$. If $L \subseteq H$ is a metaboliser for $(H, A)$, then $L \otimes_{\mathbb{Z}} \Lambda$ is a Lagrangian for $(H \otimes_{\mathbb{Z}} \Lambda, A - tA^T, A)$, and so this latter $(-t)$–quadratic form is hyperbolic.*

**Proof**  We verify the three points in the definition of a Lagrangian. Since $A$ vanishes on $L$, the extended form (also denoted by $A$) vanishes on $M := L \otimes_{\mathbb{Z}} \Lambda$. It follows that the $(-t)$–Hermitian form $A - tA^T$ vanishes on $M \times M$. Since $L \subseteq H$ is a summand, so is $M \subseteq H \otimes_{\mathbb{Z}} \Lambda$. Thus $M \subseteq (H \otimes_{\mathbb{Z}} \Lambda, A - tA^T)$ is a sublagrangian of a $(-t)$–Hermitian form, with $H \otimes_{\mathbb{Z}} \Lambda$ a free module. We showed in Lemma 7.7 that since $M$ is half-rank, this forces $M^\perp = M$.

The last sentence of the proposition follows because it is known that if a nonsingular $\varepsilon$–quadratic form $(H, \lambda, \mu)$ admits a Lagrangian, then it is isometric to the standard hyperbolic form [46, Proposition 2.2]. □

**Proposition 7.10**  *If $A$ is a Seifert matrix for an Alexander polynomial one knot $K$, then*

$$(H \otimes_{\mathbb{Z}} \Lambda, (1-t)A^T + (1-t^{-1})A) \cong \mathcal{H}_2^{\oplus g}.$$

**Proof**  Using Proposition 7.9, we deduce the following sequence of isometries:

$$(H \otimes_{\mathbb{Z}} \Lambda, A - tA^T, A) \cong H_{-t}(\Lambda)^{\oplus g} = \left( \Lambda^{2g}, \begin{pmatrix} 0 & 1 \\ -t & 0 \end{pmatrix}^{\oplus g}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^{\oplus g} \right)$$

$$\cong \left( \Lambda^{2g}, \begin{pmatrix} 0 & -t \\ 1 & 0 \end{pmatrix}^{\oplus g}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}^{\oplus g} \right).$$

Multiplying both sides by $(1-t^{-1})$ then gives the assertion. □

We are now ready to prove Theorem 1.9 from the introduction.

**Theorem 7.11**  *If $F_0$, $F_1 \subseteq D^4$ are genus $g$ pushed-in Seifert surfaces for an Alexander polynomial one knot $K$, then they are topologically ambiently isotopic rel boundary.*

**Proof**  Let $A_0$ and $A_1$ be Seifert matrices for $F_0$ and $F_1$. As mentioned above, the equivariant intersection form $(H_2(D_{F_i}^4; \Lambda), \lambda_{D_{F_i}^4})$ is isometric to

$$(H \otimes_{\mathbb{Z}} \Lambda, (1-t)A_i^T + (1-t^{-1})A_i).$$

By Proposition 7.10, both forms are isometric to $\mathcal{H}_2^{\oplus g}$. Since $K$ is an Alexander polynomial one knot, the result now follows from Theorem 7.5. □

# 8 Rim surgery on surfaces with knot group $\mathbb{Z}$

We prove Theorems 1.7 and 1.8, which concern rim surgery on surfaces in 4–manifolds whose knot group is $\mathbb{Z}$. In Section 8.1, we review the definition of rim surgery; in Sections 8.2 and 8.3 we prove the main results.

## 8.1 Knot surgery and twist rim surgery

We review some facts about knot surgery and rim surgery. For references see for instance [14; 30; 32; 2].

Let $Z$ be a compact, oriented 4–manifold containing a locally flat embedded torus $T$ with trivial normal bundle, and let $J \subseteq S^3$ be a knot. Use $\mu_T$ to denote the meridian of $T \subseteq Z$, and $\mu_J$ and $\lambda_J$ respectively for the meridian and 0–framed longitude of $J$. Let $\varphi \colon \partial \nu(T) \to S^1 \times \partial E_J$ be any diffeomorphism such that $\varphi_*(\mu_T) = \lambda_J$. The 4–manifold obtained by *knot surgery* along $J$ and $\varphi$ is defined as

$$Z_J(\varphi) = (Z \setminus \nu(T)) \cup_\varphi (S^1 \times E_J).$$

Given a compact 4–manifold $W$ and a locally flat embedded orientable surface $\Sigma \subseteq W$, we assume either that $W$ and $\Sigma$ are closed or that $\Sigma \subseteq W$ is a properly embedded. As we now describe, rim surgery arises from a particular type of knot surgery on the surface exterior $W_\Sigma := W \setminus \nu(\Sigma)$. Choose a simple closed curve $\alpha \subseteq \Sigma$ and a trivialisation of the normal bundles over $\alpha$ and $\Sigma$ so that $(\nu(\Sigma), \nu(\alpha)) = (\Sigma \times D^2, \alpha \times I \times D^2)$; in other words, the normal bundle of $\alpha$ inside $\Sigma$ is $\alpha \times I$. This way, it is understood that $\mu_\Sigma = \{\text{pt}\} \times \partial D^2$, and the *rim torus* $T$ is $\partial(\nu(\Sigma)|_\alpha) = \alpha \times \mu_\Sigma$. The rim torus is framed: if we write $\nu(\Sigma)|_{\alpha \times I} = \alpha \times I \times D^2$, then a framing of $\nu T$ is given by the $I$–direction and the radial direction in the polar coordinates of $D^2$.

In order to perform a knot surgery on $W_\Sigma$ along $J$ using the rim torus $T$, for integers $m$ and $n$ we consider the homeomorphism $\varphi_n^m \colon \partial\overline{\nu}(T) \to S^1 \times \partial E_J$ determined by

$$\varphi_n^m(\alpha) = m\mu_J + n\lambda_J + S^1, \quad \varphi_n^m(\mu_\Sigma) = \mu_J, \quad \varphi_n^m(\mu_T) = \lambda_J.$$

Gluing $\Sigma \times D^2$ back into the result of this knot surgery produces a manifold $W_J(\varphi_n^m)$ that is homeomorphic to $W$, and in fact if $W$ is smooth then $W_J(\varphi_n^m)$ is diffeomorphic to $W$; see [30, Lemma 2.4], in which a specific diffeomorphism is constructed. Thus, we obtain a new embedding $\Sigma_n^m(\alpha, J) \subseteq W$.

**Definition 8.1** Let $\alpha \subseteq \Sigma$ be a simple closed curve, let $J$ be a knot, and let $m, n \in \mathbb{Z}$ be integers. The *n–roll m–twist* rim surgery of a locally flat, properly embedded, orientable surface $\Sigma \subseteq W$ is the image $\Sigma_n^m(\alpha, J)$ of $\Sigma$ under the homeomorphism $W \cong W_J(\varphi_n^m)$ mentioned above. If $W$ is smooth and $\Sigma$ is smoothly embedded, then since $W \cong W_J(\varphi_n^m)$ is a diffeomorphism, $\Sigma_n^m(\alpha, J)$ is smoothly embedded.

The exterior $W_{\Sigma_n^m(\alpha, J)}$ of the rim surgery $\Sigma_n^m(\alpha, J)$ is the knot surgery on $W_\Sigma$ along the homeomorphism $\varphi_n^m$. The proof of the following lemma can be found in [30, Proposition 3.3].

**Lemma 8.2** *For a compact 4–manifold $W$ and a locally flat, embedded, orientable surface $\Sigma \subseteq W$, we assume either that $W$ and $\Sigma$ are closed or that $\Sigma \subseteq W$ is properly embedded. Let $\alpha \subseteq \Sigma$ be a simple closed curve, let $J$ be a knot, and let $m, n \in \mathbb{Z}$. If $\pi_1(W_\Sigma) = \mathbb{Z}_d$ with $d = \pm 1 \mod m$, then*

$$\pi_1(W_\Sigma) = \pi_1(W_{\Sigma_n^m(\alpha, J)}) = \mathbb{Z}_d.$$

In our case, the knot group is infinite cyclic, so we take $d = 0$ and restrict to $n$–roll 1–twist surgeries. We write $\Sigma_n(\alpha, J) := \Sigma_n^1(\alpha, J)$ and $\varphi_n := \varphi_n^1$.

**Remark 8.3** If $J = U$ is the unknot then $\Sigma_n(\alpha, U) = \Sigma$, by [32, Lemma 2.2].

## 8.2 The $\Lambda$–intersection form

Now we focus on the case where the ambient manifold $N$ has a boundary, and show that the $\Lambda$–intersection form of $N_{\Sigma_n(\alpha, J)}$ agrees with that of $N_\Sigma$.

We start with an observation concerning the rim torus.

**Lemma 8.4**  *Let $\Sigma \subseteq N = X \setminus \overset{\circ}{D}{}^4$ be a $\mathbb{Z}$–surface. The exterior of a rim torus $T$ satisfies*

$$H_1(N_\Sigma \setminus \nu(T)) = \mathbb{Z}\langle \mu_\Sigma \rangle \oplus \langle \mathbb{Z}\mu_T \rangle.$$

**Proof**  Consider the Mayer–Vietoris sequence of $N_\Sigma = (N_\Sigma \setminus \nu(T)) \cup \overline{\nu}(T)$:

$$(29) \quad \cdots \to H_2(N_\Sigma) \xrightarrow{\partial} H_1(T \times S^1) \xrightarrow{\theta} H_1(\overline{\nu}(T)) \oplus H_1(N_\Sigma \setminus \nu(T)) \to H_1(N_\Sigma) \to 0.$$

We show that $\partial = 0$. Note that

$$H_1(T \times S^1) \cong H_1(T) \oplus H_1(S^1) \xrightarrow{\theta} H_1(\overline{\nu}(T)) \cong H_1(T)$$

is projection onto the first factor, so for $x \neq 0$ in $H_1(T)$, $\theta(x, n) \neq 0$. Therefore in the image of $\partial$, the first coordinate vanishes. Now for a closed surface $S \subseteq N$, the definition of the connecting homomorphism implies that

$$\partial([S]) = (0, Q_N([T], [S])) \in H_1(T \times S^1) \cong H_1(T) \oplus H_1(S^1).$$

Since $T$ is isotopic to a torus in $\Sigma \times S^1 \subseteq \partial N_\Sigma$, we deduce that $Q_N([T], [S]) = 0$. It follows that $\partial = 0$ as asserted. The Mayer–Vietoris sequence in (29) then reduces to

$$0 \to \mathbb{Z}\langle \mu_T \rangle \to H_1(N_\Sigma \setminus \nu(T)) \to H_1(N_\Sigma) \to 0.$$

Since $H_1(N_\Sigma) = \mathbb{Z}\langle \mu_\Sigma \rangle$ is free, this sequence splits, establishing the lemma.  $\square$

Next, we extend the coefficient system on $\pi_1(N_\Sigma)$ over $\pi_1(N_{\Sigma_n(\alpha, J)})$.

**Remark 8.5**  The map $\pi_1(N_\Sigma) \to \mathbb{Z}$ restricts to a map $\pi_1(N_\Sigma \setminus \nu(T)) \to \mathbb{Z}$ that sends $\mu_\Sigma$ to 1 and $\mu_T$ to 0; recall from Lemma 8.4 that $H_1(N_\Sigma \setminus \nu(T)) \cong \mathbb{Z}\langle \mu_\Sigma \rangle \oplus \mathbb{Z}\langle \mu_T \rangle$. Consider the map $\pi_1(S^1 \times E_J) \to \mathbb{Z}$ that sends the class of the $S^1$–factor to $-1$ and that is the abelianisation on $\pi_1(E_J)$. We verify that these maps extend to a map $\pi_1(N_{\Sigma_n(\alpha, J)}) \to \mathbb{Z}$. We are identifying the 3–torus $\partial \overline{\nu}(T) \subseteq \partial N_\Sigma \setminus \nu(T)$ with $S^1 \times \partial E_J$ via the homeomorphism $\varphi_n$. We have the identifications $[\alpha] \sim [S^1][\mu_J][\lambda_J]^n$, $[\mu_\Sigma] \sim [\mu_J]$ and $[\mu_T] \sim [\lambda_J]$. The loop $\alpha \subseteq \Sigma$ is mapped to zero under $\pi_1(N_\Sigma) \to \mathbb{Z}$; the same is true for $[S^1][\mu_J][\lambda_J]^n$ under the map $\pi_1(S^1 \times E_J) \to \mathbb{Z}$. Similarly, both $\mu_\Sigma$ and $\mu_J$ (resp $\mu_T$ and $\lambda_J$) are sent to 1 (resp. 0) under the respective maps to $\mathbb{Z}$.

Arguing as in [32, Proof of Lemma 2.3], we construct a degree one map

$$\Psi : N_{\Sigma_n(\alpha, J)} \to N_\Sigma.$$

Using obstruction theory, one can construct a degree one map $E_J \to E_U$, where $U \subseteq S^3$ is the unknot, that takes meridian to meridian and longitude to longitude. Cross it with the identity to obtain a degree one map $S^1 \times E_J \to S^1 \times E_U$. This can be glued to the identity map $\mathrm{Id}_{N_\Sigma \setminus \nu(T)}$ to obtain a degree one map $N_{\Sigma_n(\alpha, J)} \to N_{\Sigma_n(\alpha, U)}$. By [32, Lemma 2.2], there is a homeomorphism $N_{\Sigma_n(\alpha, U)} \cong N_\Sigma$ and we have therefore obtained a degree one map

$$\Psi \colon N_{\Sigma_n(\alpha, J)} \to N_{\Sigma_n(\alpha, U)} \cong N_\Sigma.$$

By construction, $\Psi$ is the identity on $\partial N_\Sigma$. Furthermore, since the fundamental groups of both $N_{\Sigma_n(\alpha, J)}$ and $N_\Sigma$ are generated by the meridians of the surface, $\Psi$ induces an isomorphism on fundamental groups: indeed by construction $\Psi$ takes meridian to meridian. In particular $\Psi$ induces homomorphisms on the $\Lambda$–homology groups.

In order to prove that $\Psi$ in fact induces an isometry of the equivariant intersection forms, we compute the $\Lambda$–homology of $S^1 \times E_J$.

**Lemma 8.6** Let $J \subseteq S^3$ be a knot. With respect to the coefficient system on $S^1 \times E_J$ introduced in Remark 8.5, we have $H_2(S^1 \times E_J; \Lambda) = 0$ and $H_1(S^1 \times E_J; \Lambda) = \mathbb{Z}$. Additionally, the degree one map $\Psi$ induces an isometry

$$\Psi_* \colon (H_2(N_{\Sigma_n(\alpha, J)}; \Lambda), \lambda_{N_{\Sigma_n(\alpha, J)}}) \xrightarrow{\cong} (H_2(N_\Sigma; \Lambda), \lambda_{N_\Sigma}).$$

**Proof** We use the Künneth spectral sequence with

$$E^2_{p,q} = \bigoplus_{q_1+q_2=q} \mathrm{Tor}_p(H_{q_1}(S^1; \Lambda), H_{q_2}(E_J; \Lambda)),$$

which converges to $H_*(S^1 \times E_J; \Lambda \otimes_\Lambda \Lambda) \cong H_*(S^1 \times E_J; \Lambda)$, where $\Lambda \otimes_\Lambda \Lambda$ is a right $\mathbb{Z}[\pi_1(E_K) \times \pi_1(S^1)]$–module via the diagonal action induced by $p \otimes q \cdot (e, s) = pe \otimes qs$ for $p, q \in \Lambda$ and $(e, s) \in \pi_1(E_K) \times \pi_1(S^1)$. Write $\varepsilon \colon \Lambda \to \mathbb{Z}$, $p(t) \mapsto p(1)$ for the augmentation map and $\mathbb{Z}_\varepsilon$ for the resulting $\Lambda$–module structure on $\mathbb{Z}$. By definition of the coefficient system in Remark 8.5, we have $H_1(S^1; \Lambda) = 0$ and $H_0(S^1; \Lambda) = \Lambda/(t-1) = \mathbb{Z}_\varepsilon$. We also have $H_0(E_J; \Lambda) = \mathbb{Z}$ and $H_i(E_J; \Lambda) = 0$ for $i \geq 2$. This implies that $E^2_{0,2} = 0$. In fact, we also deduce that $E^2_{1,0} = \mathbb{Z}$ and $E^2_{2,0} = 0$ because

$$E^2_{i,0} = \mathrm{Tor}_i^{\mathbb{Z}[\mathbb{Z}]}(\mathbb{Z}, \mathbb{Z}) = H_i(\mathbb{Z}; \mathbb{Z}) = H_i(S^1).$$

We claim that $\mathrm{Tor}_i^\Lambda(\mathbb{Z}_\varepsilon, H_1(E_J; \Lambda)) = 0$ for $i = 0, 1$. These Tor groups are computed as the homology of the complex obtained tensoring the resolution $0 \to \Lambda \xrightarrow{t-1} \Lambda \to \mathbb{Z}_\varepsilon \to 0$ with the Alexander module $H_1(E_J; \Lambda)$. Since multiplication by $t-1$ induces an

isomorphism on the Alexander module [39, Proposition 1.2], these Tor groups vanish and the claim is proved.

Using the claim, we deduce that $E_{0,1}^2 = 0$ and $E_{1,1}^2 = 0$. It follows that $d_{2,0} = 0$ and now the first assertion is a consequence of a standard spectral sequence computation.

We now assert that $\Psi$ induces an isomorphism on $H_2(-; \Lambda)$. Recall the decompositions

$$N_{\Sigma_n(\alpha, J)} = (N_\Sigma \setminus \nu(T)) \cup_{\varphi_n} (S^1 \times E_J),$$
$$N_\Sigma \cong N_{\Sigma_n(\alpha, U)} = (N_\Sigma \setminus \nu(T)) \cup_{\varphi_n} (S^1 \times E_U).$$

By construction, $\Psi$ restricts to the identity on $N_\Sigma \setminus \nu(T)$ and $\partial \overline{\nu}(T)$ and therefore induces the identity on the $\Lambda$–homology of these spaces. Next, $\Psi_* \colon H_1(S^1 \times E_J; \Lambda) \to H_1(S^1 \times E_U; \Lambda)$ is also an isomorphism: indeed, the first assertion shows that both modules are isomorphic to $\mathbb{Z} = \Lambda/(t-1)$, and this term comes from $H_0(E_J; \Lambda) = \mathbb{Z}_\varepsilon$, on which $\Psi$ does indeed induce an isomorphism. We showed that $H_2(S^1 \times E_J; \Lambda) = 0$, so the assertion now follows from the five lemma applied to the following commutative diagram, where $\Lambda$–coefficients are understood:

$$\begin{array}{ccccccccc}
H_2(\partial\overline{\nu}(T)) & \to & H_2(N_\Sigma\setminus\nu(T)) & \to & H_2(N_{\Sigma_n(\alpha,J)}) & \to & H_1(\partial\overline{\nu}(T)) & \to & \begin{matrix}H_1(N_\Sigma\setminus\nu(T))\\ \oplus \\ \mathbb{Z}\end{matrix} \to 0 \\
\Psi_* \| & & \Psi_* \| & & \Psi_* \downarrow & & \Psi_* \| & & \Psi_* \downarrow \cong \\
H_2(\partial\overline{\nu}(T)) & \to & H_2(N_\Sigma\setminus\nu(T)) & \to & H_2(N_\Sigma) & \to & H_1(\partial\overline{\nu}(T)) & \to & \begin{matrix}H_1(N_\Sigma\setminus\nu(T))\\ \oplus \\ \mathbb{Z}\end{matrix} \to 0
\end{array}$$

Since $\Psi$ induces a $\Lambda$–isomorphism on $H_2(-; \Lambda)$, it induces one on $H_2(-; \Lambda)^*$, and therefore on $H^2(-; \Lambda)$ (recall the UCSS argument from Lemma 3.2) and thus on the second relative $\Lambda$–homology groups (because $\deg(\Psi) = 1$). We conclude that $\Psi_*$ in fact induces an isometry of the $\Lambda$–intersection forms. □

## 8.3  Topological triviality of rim surgery on surfaces with knot group $\mathbb{Z}$

Next we prove Theorem 1.8 from the introduction.

**Theorem 8.7**  *Let $\Sigma \subseteq N = X \setminus \mathring{D}^4$ be a $\mathbb{Z}$–surface, let $\alpha \subseteq \Sigma$ be a simple closed curve, let $J$ be a knot, and let $n \in \mathbb{Z}$ be an integer. There is a rel boundary orientation-preserving homeomorphism of pairs*

$$\Phi \colon (N, \Sigma_n(\alpha, J)) \xrightarrow{\cong} (N, \Sigma)$$

that induces the same isometry as the degree one map $\Psi$ on the equivariant intersection forms of the surface exteriors. If $N = D^4$, then the surfaces are topologically isotopic rel boundary.

**Proof** By Lemma 8.2 both surfaces have knot group $\mathbb{Z}$. Lemma 8.6 states that the degree one map $\Psi \colon N_{\Sigma_n(\alpha, J)} \to N_\Sigma$ induces an isometry $\Psi_* \colon \lambda_{N_{\Sigma_n(\alpha, J)}} \cong \lambda_{N_\Sigma}$. As $\Psi$ is the identity on $\partial N_\Sigma$, we see that $\partial \Psi$ induces the identity isometry between the boundary Blanchfield forms. The result now follows from Theorem 5.8. $\qquad\square$

The remaining result we owe a proof of is Theorem 1.7 from the introduction. To prove this result, which involves obtaining an isotopy between closed surfaces, we need to control the map $F_\mathbb{Z}$ discussed in Lemma 5.10. This will rely on the following result.

**Lemma 8.8** *Let $\Sigma \subseteq N$ be a $\mathbb{Z}$–surface, let $\alpha \subseteq \Sigma$ be a simple closed curve with associated rim torus $T$, and let $J$ be a knot. The following sequences of inclusion induced maps are exact:*

$$0 \to \mathbb{Z}\langle \alpha \times \mu_T \rangle \to H_2(N_\Sigma \setminus \nu(T); \Lambda) \to H_2(N_{\Sigma_n(\alpha, J)}; \Lambda) \to 0,$$
$$0 \to \mathbb{Z}\langle \alpha \times \mu_T \rangle \to H_2(N_\Sigma \setminus \nu(T); \Lambda) \to H_2(N_\Sigma; \Lambda) \to 0.$$

**Proof** As we noted in Lemma 8.2 that $\pi_1(N_{\Sigma_n(\alpha, J)}) = \mathbb{Z}$ and since $N_{\Sigma_n(\alpha, J)}$ has ribbon boundary, we know that $H_1(N_{\Sigma_n(\alpha, J)}; \Lambda) = 0$ as well as $H_3(N_{\Sigma_n(\alpha, J)}; \Lambda) = 0$, again by Lemma 3.2. Consequently, the Mayer–Vietoris sequence for $N_{\Sigma_n(\alpha, J)} = N_\Sigma \setminus \nu(T) \cup_{\varphi_n} (S^1 \times E_J)$ with $\Lambda$ coefficients gives

$$0 \to H_2(\partial \overline{\nu}(T); \Lambda) \to H_2(N_\Sigma \setminus \nu(T); \Lambda) \oplus H_2(S^1 \times E_J; \Lambda) \to H_2(N_{\Sigma_n(\alpha, J)}; \Lambda)$$
$$\to H_1(\partial \overline{\nu}(T); \Lambda) \to H_1(N_\Sigma \setminus \nu(T); \Lambda) \oplus H_1(S^1 \times E_J; \Lambda) \to 0.$$

Since there is a homotopy equivalence $\overline{\nu}(T) \simeq \alpha \times \mu_\Sigma$, the discussion of coefficient systems from Remark 8.5 implies that

$$H_i(\partial \overline{\nu}(T); \Lambda) = H_i(\alpha \times \mu_\Sigma \times \mu_T; \Lambda) = H_i(\alpha \times \mu_T) \quad \text{for } i = 1, 2.$$

We saw in Lemma 8.6 that $H_2(E_J \times S^1; \Lambda) = 0$ and $H_1(E_J \times S^1; \Lambda) = \mathbb{Z}$. The previous sequence therefore reduces to

$$(30) \quad 0 \to \mathbb{Z}\langle \alpha \times \mu_T \rangle \to H_2(N_\Sigma \setminus \nu(T); \Lambda) \to H_2(N_{\Sigma_n(\alpha, J)}; \Lambda) \to \mathbb{Z}\langle \alpha \rangle \oplus \mathbb{Z}\langle \mu_T \rangle$$
$$\to H_1(N_\Sigma \setminus \nu(T); \Lambda) \oplus \mathbb{Z} \to 0.$$

When $J = U$ is the unknot, $H_1(S^1 \times E_J; \Lambda) = H_1(\alpha) = \mathbb{Z}\langle \alpha \rangle$, and therefore $\mathbb{Z}\langle \mu_T \rangle$ surjects onto $H_1(N_\Sigma \setminus \nu(T); \Lambda)$. In particular, as an abelian group, $H_1(N_\Sigma \setminus \nu(T); \Lambda)$ is cyclic.

**Claim 7** *The map $\mathbb{Z}\langle\mu_T\rangle \xrightarrow{\iota} H_1(N_\Sigma \setminus \nu(T); \Lambda)$ is an isomorphism.*

**Proof** We already noted that, as an abelian group, $H_1(N_\Sigma \setminus \nu(T); \Lambda)$ is cyclic, so it suffices to show that $H_1(N_\Sigma \setminus \nu(T); \Lambda) \otimes_{\mathbb{Z}} \mathbb{Q} \neq 0$. Equivalently, we must establish that

$$H := H_1(N_\Sigma \setminus \nu(T); \mathbb{Q}[t^{\pm 1}]) \neq 0.$$

By way of contradiction, assume that $H = 0$. Let $\varepsilon\colon \Lambda \to \mathbb{Z}$ again denote the augmentation map and write $\mathbb{Z}_\varepsilon$ and $\mathbb{Q}_\varepsilon$ for the resulting $\Lambda$–module structures. Since $\text{Tor}_2^\Lambda(\mathbb{Z}_\varepsilon, \mathbb{Q}_\varepsilon) = 0$, the universal coefficient spectral sequence applied to

$$H_1(N_\Sigma \setminus \nu(T); \mathbb{Q}) = H_1(N_\Sigma \setminus \nu(T); \mathbb{Q}[t^{\pm 1}] \otimes_\Lambda \mathbb{Q}_\varepsilon)$$

produces the short exact sequence

$$0 \to H \otimes_\Lambda \mathbb{Q}_\varepsilon \to H_1(N_\Sigma \setminus \nu(T); \mathbb{Q}) \to \text{Tor}_1^\Lambda(\mathbb{Z}_\varepsilon, \mathbb{Q}_\varepsilon) \to 0.$$

We assumed that $H = 0$, so $H \otimes_\Lambda \mathbb{Q}_\varepsilon = 0$, and thus $H_1(N_\Sigma \setminus \nu(T); \mathbb{Q}) \cong \text{Tor}_1^\Lambda(\mathbb{Z}, \mathbb{Q})$. Using Lemma 8.4, $H_1(N_\Sigma \setminus \nu(T); \mathbb{Q}) = \mathbb{Q}^2$. On the other hand, $\text{Tor}_1^\Lambda(\mathbb{Z}_\varepsilon, \mathbb{Q}_\varepsilon) = H_1(S^1; \mathbb{Q}) = \mathbb{Q}$. This is a contradiction and the claim is established. □

Using Claim 7, the penultimate map in (30) is a surjection $\mathbb{Z}^2 \to \mathbb{Z}^2$ and therefore an isomorphism. The first short exact sequence in the statement of the lemma now follows from the one displayed in (30). The exactness of the second sequence follows from exactness of the first by taking $J = U$ and recalling that $\Sigma_n(\alpha, U) = \Sigma$ by Remark 8.3. □

**Theorem 8.9** *Let $\Sigma \subseteq X$ be a closed $\mathbb{Z}$–surface, let $\alpha \subseteq \Sigma$ be a simple closed curve, let $J$ be a knot, and let $n \in \mathbb{Z}$. The surfaces $\Sigma$ and $\Sigma_n(\alpha, J)$ are topologically ambiently isotopic.*

**Proof** After an ambient isotopy, we may assume that the surfaces $\Sigma_0 := \Sigma_n(\alpha, J)$ and $\Sigma_1 := \Sigma$ coincide on a disc $D^2 \subseteq \Sigma \cap \Sigma_n(\alpha, J)$ that does not intersect $\alpha \subseteq \Sigma$. Assume that the normal bundles also coincide over this $D^2$. Consider the preimage $\mathring{D}^2 \times \mathbb{R}^2 \subseteq \nu\Sigma_i$. This is homeomorphic to an open 4–ball $\mathring{D}^4$. Remove this $(\mathring{D}^4, \mathring{D}^2)$ from $(X, \Sigma)$ and $(X, \Sigma_n(\alpha, J))$ to obtain properly embedded surfaces $\widetilde{\Sigma} \subseteq N$ and $\widetilde{\Sigma}_n(\alpha, J) \subseteq N$ with the unknot as a common boundary. By construction, $\widetilde{\Sigma}_n(\alpha, J)$ is obtained by rim surgery on $\widetilde{\Sigma}$. Apply Theorem 8.7 to obtain a rel boundary homeomorphism of pairs $\Phi\colon (N, \widetilde{\Sigma}_n(\alpha, J)) \to (N, \widetilde{\Sigma})$. Recall that on $H_2(-; \Lambda)$, we have $\Phi_* = \Psi_*$, where $\Psi$ is the degree one map described above Lemma 8.6. Construct

a homeomorphism of pairs $(X, \Sigma_n(\alpha, J)) \to (X, \Sigma)$ by gluing $\Phi$ with the identity homeomorphism $(D^4, D^2) \to (D^4, D^2)$.

We refine this argument to obtain the required ambient isotopy. Use $F$ to denote the isometry induced by these homeomorphisms on $H_2(-; \Lambda)$ of the surface exteriors. In Lemma 8.6, we argued that this isometry fits into the following commutative diagram with exact rows (here we also used Lemma 8.8 to simplify the Mayer–Vietoris sequences):

$$(31) \qquad \begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{Z}\langle \alpha \times \mu_T \rangle & \longrightarrow & H_2(N_{\widetilde{\Sigma}} \setminus \nu(T); \Lambda) & \longrightarrow & H_2(N_{\widetilde{\Sigma}_n(\alpha,J)}; \Lambda) & \longrightarrow & 0 \\ & & \Big\| \text{Id} & & \Big\| \text{Id} & & \Big\downarrow F=\Psi_* & & \\ 0 & \longrightarrow & \mathbb{Z}\langle \alpha \times \mu_T \rangle & \longrightarrow & H_2(N_{\widetilde{\Sigma}} \setminus \nu(T); \Lambda) & \longrightarrow & H_2(N_{\widetilde{\Sigma}}; \Lambda) & \longrightarrow & 0 \end{array}$$

Recall that $\Sigma_0 := \Sigma_n(\alpha, J)$ and $\Sigma_1 := \Sigma$, so $X_{\Sigma_0} \cong N_{\widetilde{\Sigma}_n(\alpha,J)}$ and $X_{\Sigma_1} \cong N_{\widetilde{\Sigma}}$. In Lemma 5.10, $F_{\mathbb{Z}}$ was defined by the following commutative diagram with exact rows:

$$(32) \qquad \begin{array}{ccccccccc} 0 & \longrightarrow & \mathbb{Z}^{2g} & \longrightarrow & H_2(X_{\Sigma_0}) & \xrightarrow{p_0} & H_2(X) & \longrightarrow & 0 \\ & & \Big\downarrow (F \otimes_\Lambda \text{Id}_{\mathbb{Z}})| & & \Big\downarrow F \otimes_\Lambda \text{Id}_{\mathbb{Z}} & & \Big\downarrow F_{\mathbb{Z}} & & \\ 0 & \longrightarrow & \mathbb{Z}^{2g} & \longrightarrow & H_2(X_{\Sigma_1}) & \xrightarrow{p_1} & H_2(X) & \longrightarrow & 0 \end{array}$$

By (32), every element $x$ of $H_2(X)$ can be represented by a class in $H_2(X_{\Sigma_0})$, and by (31) every class here can be represented by a surface $S$ in $N_{\widetilde{\Sigma}_n(\alpha,J)} \setminus \nu(T) \subseteq N_{\widetilde{\Sigma}_n(\alpha,J)} \cong X_{\Sigma_0}$. Since $F$ is induced by the degree one map $\Psi$, mapping our surface to the same surface $S$ in $N_{\widetilde{\Sigma}} \setminus \nu(T) \subseteq N_{\widetilde{\Sigma}} \cong X_{\Sigma_1}$ yields $(F \otimes_\Lambda \text{Id}_{\mathbb{Z}})([S]) \in H_2(X_{\Sigma_1})$. The inclusion induced map $p_1 \colon H_2(X_{\Sigma_1}) \to H_2(X)$ then sends $(F \otimes_\Lambda \text{Id}_{\mathbb{Z}})([S])$ to $x$. Therefore $F_{\mathbb{Z}} \colon H_2(X) \to H_2(X)$ is indeed the identity, so the second item of Theorem 5.11 implies that $\Sigma_0$ and $\Sigma_1$ are topologically ambiently isotopic. $\qquad \square$

# Appendix   Stable isotopy for surfaces in a topological 4–manifold with boundary

The next result is an extension of [2, Theorem 5] to the topological case, and allowing nonempty boundary, but restricting the ambient 4–manifolds somewhat. Since we need the given extension, we provide details of the proof. The main ideas are due to [2]. We fill in some details in their argument for constructing a map to $S^1$ in the course of the proof. The case with nonempty boundary was also stated as [25, Proposition 2.13] in the smooth category.

**Theorem A.1**  *Let $N$ be an oriented, connected, simply connected, compact topological 4–manifold with boundary $S^3$. Let $\Sigma$ be a compact, connected, orientable surface with one boundary component. Let $\Sigma_0, \Sigma_1 \subseteq N$ be two properly embedded, locally flat, oriented surfaces homeomorphic to $\Sigma$, and suppose that $\partial \Sigma_0 = \partial \Sigma_1$. Assume that for each $i$, $\pi_1(N \setminus \Sigma_i) \cong \mathbb{Z}$ is infinite cyclic generated by an oriented meridian to $\Sigma_i$. Then some finite number of trivial one-handle stabilisations results in ambiently isotopic surfaces $\Sigma_0'$ and $\Sigma_1'$, with $\partial \Sigma_i = \partial \Sigma_i'$ for $i = 0, 1$.*

We begin with a lemma on framings, some of which is a recollection from Section 5. Note that any locally flat embedded surface in a 4–manifold admits a normal bundle with linear structure group [15, Theorem 9.3].

**Lemma A.2**  *For $i = 0, 1$ the normal bundle of $\Sigma_i$ is trivial and for any choice of framing there is a well-defined induced (homotopy class of) framing on the normal bundle of $\partial \Sigma_i$. The surface $\Sigma_i$ is null-homologous in $H_2(N, \partial N)$, and the induced framing on $\partial \Sigma_i$ equals the Seifert framing.*

**Proof**  Since the surface $\Sigma_i$ has nonempty boundary and is oriented, its normal bundle is trivial. Two choices of framings on $\Sigma_i$ differ by a map $\Sigma_i \to SO(2)$. The difference between the two framings on the boundary $\partial \Sigma_i$ is governed by the homotopy class of the composite

$$S^1 \xrightarrow{\cong} \partial \Sigma_i \to \Sigma_i \to SO(2) \cong S^1.$$

Since the boundary determines a commutator in $\pi_1(\Sigma_i)$, the displayed map is null-homotopic. It follows that the two framings induced on $\partial \Sigma_i$ agree up to homotopy. This proves that the induced framing is well-defined. The second sentence was already shown in Lemmas 5.1 and 5.2.                                                                                     $\square$

**Proof of Theorem A.1**  Push the boundary of $\Sigma_1$ off the boundary of $\Sigma_0$, using the Seifert framing, to arrange that $\partial \Sigma_0$ and $\partial \Sigma_1$ in $\partial N \cong S^3$ are parallel circles. We may assume that there is a collar $\partial N \times I$ with $\partial N \times \{0\} = \partial N$, and an embedding $g \colon S^1 \times I \times I \hookrightarrow \partial N \times I$ corresponding to the Seifert framing push-off, with

$$g(S^1 \times I \times \{i\}) = \Sigma_i \cap (\partial N \times I) \quad \text{for } i = 0, 1,$$
$$g(S^1 \times \{t\} \times I) \subseteq \partial N \times \{t\} \qquad \text{for all } t \in I.$$

Let $A \subseteq \partial N \cong S^3$ be the annulus

$$A := g(S^1 \times \{0\} \times I)$$
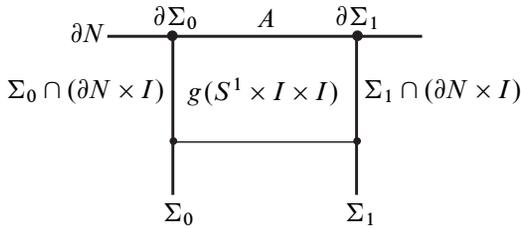
Figure 4: A schematic diagram indicating the relation of $g(S^1 \times I \times I) \subseteq$ $\partial N \times I$ to $\partial \Sigma_0$, $\partial \Sigma_1$, and the annulus $A \subseteq S^3$ that joins them.

connecting $\partial \Sigma_0$ and $\partial \Sigma_1$, arising from the trace of the push. Note that $A$ induces the Seifert framing on $\partial \Sigma_i$ for $i = 0, 1$, and that by Lemma A.2 this equals the framing induced by some choice of framing of the normal bundle of $\Sigma_i$.

By an isotopy, arrange further that $\Sigma_0$ and $\Sigma_1$ intersect transversely in their interiors [15, Theorem 9.5], and since $\Sigma$ is compact we may also assume there are finitely many intersection points.

Cap off each $\partial \Sigma_i$ with a Seifert surface, and by a small isotopy of the capped-off $\Sigma_1$, arrange that the capped-off surfaces are disjoint in $\partial N \times I$. By Lemma A.2 (and since $H_2(\partial N) = H_2(S^3) = 0$), the capped-off surfaces are null-homologous in $H_2(N)$, and so intersect algebraically zero times. By tubing $\Sigma_0$ to itself, that is by 1–handle stabilisations, arrange that $\Sigma_0$ and $\Sigma_1$ are disjoint. To achieve this, pair up points with opposite signs in $\Sigma_0 \pitchfork \Sigma_1$ and for each pair $\{p, q\}$ choose a path $\gamma$ in $\Sigma_1$ connecting the two intersection points and away from the other intersection points. We can choose these paths to be mutually disjoint, but this is not obligatory. For each pair of points $\{p, q\}$, remove two open discs from $\Sigma_0$, one for each pair of points, and add a tube, coming from the normal circle bundle of $\Sigma_1$ restricted to $\gamma$, as shown in Figure 5. In case different paths on $\Sigma_1$ intersect, vary the radii of the tubes to keep them disjoint. This stabilises $\Sigma_0$ to a surface disjoint from $\Sigma_1$.

Perform the same number of trivial 1–handle stabilisations to $\Sigma_1$, so that the surfaces are still abstractly homeomorphic. We will abuse notation and still refer to the resulting surfaces as $\Sigma_0$ and $\Sigma_1$. Let

$$N_{\Sigma_0, \Sigma_1} := N \setminus (\nu \Sigma_0 \cup \nu \Sigma_1).$$

We will construct an oriented 3–dimensional relative Seifert manifold $Y \subseteq N_{\Sigma_0, \Sigma_1}$, that is a rel boundary cobordism between $\Sigma_0$ and $\Sigma_1$, embedded in $N$ and with $\partial Y = S$, where

$$S := \Sigma_0 \cup A \cup -\Sigma_1.$$
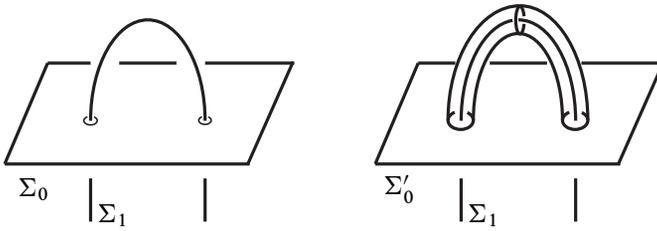
Figure 5: Stabilising $\Sigma_0$ to $\Sigma_0'$ by adding a 1–handle in a neighbourhood of an arc $\gamma$ in $\Sigma_1$.

**Claim** *There is a locally flat, embedded, compact, orientable 3–manifold $Y$ with $\partial Y = S$ and corners at $\Sigma_i \cap A$, as described in the preceding paragraph.*

**Proof** The strategy to construct $Y$ is as follows: define a suitable map $\partial N_{\Sigma_0,\Sigma_1} \to S^1$, extend it to a map $N_{\Sigma_0,\Sigma_1} \to S^1$ while controlling the restriction to $\partial N_{\Sigma_0,\Sigma_1}$, and then take $Y$ to be the inverse image of a transverse regular point in $S^1$.

The first step is to construct a map $\alpha\colon N_{\Sigma_0,\Sigma_1} \to S^1$. Recall that $N_{\Sigma_i} := N \setminus \nu\Sigma_i$. We can express $N = N_{\Sigma_0} \cup N_{\Sigma_1}$, with $N_{\Sigma_0,\Sigma_1} = N_{\Sigma_0} \cap N_{\Sigma_1}$. A portion of the resulting Mayer–Vietoris sequence (with $\mathbb{Z}$ coefficients) is

$$H_2(N) \xrightarrow{\delta} H_1(N_{\Sigma_0,\Sigma_1}) \to H_1(N_{\Sigma_0}) \oplus H_1(N_{\Sigma_1}) \to H_1(N).$$

By hypothesis $H_1(N) = 0$ and $H_1(N_{\Sigma_i}) \cong \mathbb{Z}$, so that we obtain a short exact sequence

$$0 \to \operatorname{im}\delta \to H_1(N_{\Sigma_0,\Sigma_1}) \to \mathbb{Z} \oplus \mathbb{Z} \to 0.$$

Since $\mathbb{Z} \oplus \mathbb{Z}$ is free abelian, this splits and we have that

$$H_1(N_{\Sigma_0,\Sigma_1}) \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \operatorname{im}\delta.$$

By the hypotheses of Theorem A.1, the first two summands are generated by meridians $\mu_{\Sigma_0}$ and $\mu_{\Sigma_1}$ to the surfaces $\Sigma_0$ and $\Sigma_1$, respectively. Consider the dual element

$$\alpha := \mu_{\Sigma_0}^* - \mu_{\Sigma_1}^* \in \operatorname{Hom}(H_1(N_{\Sigma_0,\Sigma_1}), \mathbb{Z}) \cong H^1(N_{\Sigma_0,\Sigma_1}) \cong [N_{\Sigma_0,\Sigma_1}, S^1],$$

sending the $\operatorname{im}\delta$ summand identically to 0. We will also write $\alpha\colon N_{\Sigma_0,\Sigma_1} \to S^1$ for a corresponding representing map. We use that since $N_{\Sigma_0,\Sigma_1}$ is a topological 4–manifold, it is homotopy equivalent to a CW–complex [17, Theorem 4.5], and therefore we may identify $H^1(N_{\Sigma_0,\Sigma_1}) \cong [N_{\Sigma_0,\Sigma_1}, S^1]$.

Now we consider the restriction to $\partial N_{\Sigma_0,\Sigma_1}$ under the inclusion-induced map

$$\iota^*\colon H^1(N_{\Sigma_0,\Sigma_1}) \to H^1(\partial N_{\Sigma_0,\Sigma_1}) \cong [\partial N_{\Sigma_0,\Sigma_1}, S^1],$$

where $\iota\colon \partial N_{\Sigma_0,\Sigma_1} \to N_{\Sigma_0,\Sigma_1}$ is the inclusion map. We have a decomposition

$$\partial N_{\Sigma_0,\Sigma_1} = \Sigma_0 \times S^1 \cup \partial_0 N_{\Sigma_0,\Sigma_1} \cup \Sigma_1 \times S^1.$$

An elementary Mayer–Vietoris argument yields that

$$H^1(\partial N_{\Sigma_0,\Sigma_1}) \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}^{2g} \oplus \mathbb{Z}^{2g}.$$

The first two summands are generated by duals to the meridians, $\mu_{\Sigma_0}^*$ and $\mu_{\Sigma_1}^*$. By our work with framings above, for $i = 0, 1$ there is a framing of the normal bundle of $\Sigma_i$, determining an identification of the tubular neighbourhood $\overline{\nu}\Sigma_i$ with $\Sigma_i \times D^2$, that agrees on $\partial\Sigma_i$ with a corresponding identification induced from the Seifert framing on $\partial\Sigma_i$. This determines an identification

$$\partial\overline{\nu}\Sigma_i \setminus \nu\partial\Sigma_i \cong \Sigma_i \times S^1.$$

The two $\mathbb{Z}^{2g}$ summands of $H^1(\partial N_{\Sigma_0,\Sigma_1})$ are generated by dual classes to curves of the form $\gamma_k \times \{-1\}$, where $\gamma_k$ is a simple closed curve forming part of a symplectic basis for $H_1(\Sigma_i)$, for some $i$. Here we use the chosen framing of the normal bundle of $\Sigma_i$ to fix representatives for the $\mathbb{Z}^{2g}$ summands.

The restriction $\iota^*\alpha \in H^1(\partial N_{\Sigma_0,\Sigma_1})$ is $(1, -1, x, y)$, for some $x, y \in \mathbb{Z}^{2g}$. But by changing the choice of framing of $\nu\Sigma_i$ along the curves $\gamma_k$, we can arrange that $x = y = 0$. By changing the framing of the normal bundle of $\Sigma_i$ by some number of full twists along some basis curve $\gamma_k$ in $H_1(\Sigma_i)$, we change the meaning of $\gamma_k \times \{-1\}$ in the previous paragraph.

Let us explain the operation of "changing the framing" in more detail. Such a change is occasioned by the action of $\left[\bigcup_k \gamma_k, \mathrm{SO}(2)\right]$ on the set of framings of the normal bundle, to alter the given framing on the union of the curves $\{\gamma_k\}$ by some number of full twists for each curve. Any such alteration automatically extends over the 2–skeleton since the attaching map of the 2–cell of $\Sigma_i$ is a commutator in the $\gamma_k$ times $\partial\Sigma_i$. Any two choices of extension over the 2–cells are homotopic, since $\pi_2(\mathrm{SO}(2)) = 0$. Therefore we have a well-defined notion of altering the framing along the curves $\gamma_k$. This changes the entry of $(x, y)$ corresponding to $\gamma_k$, since the map $\alpha\colon N_{\Sigma_0,\Sigma_1} \to S^1$ now sends $\gamma_k \times \{-1\}$ to a curve representing a different element of $H_1(S^1)$. By Lemma A.2, changing the choice of framing on a basis element for $H_1(\Sigma_i)$ does not change the induced framing on the boundary.

Now we define a map $f\colon \partial N_{\Sigma_0,\Sigma_1} \to S^1$. On $\Sigma_i \times S^1 \subseteq \partial N_{\Sigma_0,\Sigma_1}$, define the map to $S^1$ by the projection $f|_{\Sigma_i \times S^1}\colon \Sigma_i \times S^1 \to S^1$ onto the second factor. On the remainder

of $\partial N_{\Sigma_0,\Sigma_1}$, namely

$$\partial_0 N_{\Sigma_0,\Sigma_1} := \partial N \setminus (\nu\Sigma_0 \cup \nu\Sigma_1),$$

or in other words the link exterior $S^3 \setminus (\nu\partial\Sigma_0 \cup \nu\partial\Sigma_1)$, define a Pontryagin–Thom-style collapse map $f|_{\partial_0 N_{\Sigma_0,\Sigma_1}} : \partial_0 N_{\Sigma_0,\Sigma_1} \to S^1$ by choosing a tubular neighbourhood $A \times [-1,1] \subseteq \partial_0 N_{\Sigma_0,\Sigma_1}$ and sending $(a,x) \mapsto e^{\pi i x} \in S^1$ for $a \in A$ and $x \in [-1,1]$, then sending $\partial_0 N_{\Sigma_0,\Sigma_1} \setminus (A \times [-1,1])$ to $-1 \in S^1$. Then $f|_{\partial_0 N_{\Sigma_0,\Sigma_1}}^{-1}(\{1\}) = A \cap \partial_0 N_{\Sigma_0,\Sigma_1}$. Note that we may assume that our maps to $S^1$ agree on the torus overlaps $\partial\Sigma_i \times S^1$, since the framings on $\partial\Sigma_i$ all agree up to homotopy. This completes the construction of a map $f : \partial N_{\Sigma_0,\Sigma_1} \to S^1$. Note that $f : \partial N_{\Sigma_0,\Sigma_1} \to S^1$ corresponds to the element

$$(1,-1,0,0) \in \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}^{2g} \oplus \mathbb{Z}^{2g}.$$

Therefore the cohomology classes $\iota^*\alpha$ and $f$ agree in $H^1(\partial N_{\Sigma_0,\Sigma_1}) \cong [\partial N_{\Sigma_0,\Sigma_1}, S^1]$, and so by a homotopy of $\alpha$ supported in a collar of $\partial N_{\Sigma_0,\Sigma_1}$ we obtain a map $F : N_{\Sigma_0,\Sigma_1} \to S^1$ with $F|_{\partial N_{\Sigma_0,\Sigma_1}} = f : \partial N_{\Sigma_0,\Sigma_1} \to S^1$.

The inverse image under $F$ of a transverse regular point in $S^1$ yields a 3–dimensional relative Seifert manifold, locally flatly embedded in $N_{\Sigma_0,\Sigma_1} = N \setminus (\nu\Sigma_0 \cup \nu\Sigma_1)$. See [17, Section 10.2] for information on map transversality in the topological category. Add collars $\Sigma_0 \times I$ and $\Sigma_1 \times I$ in $\nu\Sigma_0$ and $\nu\Sigma_1$, respectively, to obtain the 3–manifold $Y$ that we seek. This completes the proof of the claim.                                   □

Morse theory on $Y$ gives rise to a Heegaard decomposition relative to the annulus $A$. The Heegaard surface can be obtained from both $\Sigma_0$ and $\Sigma_1$ by 1–handle stabilisations and ambient isotopy.

Since by our assumptions $\pi_1(N_{\Sigma_i})$ is generated by a meridian to $\Sigma_i$, Boyle's [6] proof shows that every 1–handle stabilisation is a trivial stabilisation. He applied [23, Theorem 4] of Hudson for the statement that $D^1$ cores of handle additions that are homotopic rel endpoints are in fact smoothly ambiently isotopic fixing the endpoints. To apply Boyle's work to topologically embedded surfaces in a compact 4–manifold $N$, remove a point from $N$ and smooth $N \setminus \{pt\}$ in such a way that $\Sigma$ is smoothly embedded. Then Boyle's application of Hudson's result yields a smooth ambient isotopy, which gives rise to a topological ambient isotopy once the point is added back to $N$. We note that Boyle works in $S^4$, but his proof applies to any oriented ambient 4–manifold.

We therefore have that after finitely many trivial stabilisations, $\Sigma_0$ and $\Sigma_1$ are ambiently isotopic in $N$ relative to the constant isotopy on the boundary, as desired.              □

# References

[1]  **H Bass**, *Unitary algebraic K–theory*, from "Algebraic $K$–theory, III: Hermitian $K$–theory and geometric applications" (H Bass, editor), Lecture Notes in Math. 343, Springer (1973) 57–265  MR  Zbl

[2]  **R İ Baykur**, **N Sunukjian**, *Knotted surfaces in 4–manifolds and stabilizations*, J. Topol. 9 (2016) 215–231  MR  Zbl

[3]  **M Borodzik**, **S Friedl**, *On the algebraic unknotting number*, Trans. London Math. Soc. 1 (2014) 57–84  MR  Zbl

[4]  **M Borodzik**, **S Friedl**, *The unknotting number and classical invariants, I*, Algebr. Geom. Topol. 15 (2015) 85–135  MR  Zbl

[5]  **S Boyer**, *Simply-connected 4–manifolds with a given boundary*, Trans. Amer. Math. Soc. 298 (1986) 331–357  MR  Zbl

[6]  **J Boyle**, *Classifying 1–handles attached to knotted surfaces*, Trans. Amer. Math. Soc. 306 (1988) 475–487  MR  Zbl

[7]  **G E Bredon**, *Topology and geometry*, Graduate Texts in Math. 139, Springer (1993)  MR  Zbl

[8]  **T D Cochran**, **K E Orr**, **P Teichner**, *Structure in the classical knot concordance group*, Comment. Math. Helv. 79 (2004) 105–123  MR  Zbl

[9]  **A Conway**, **S Friedl**, **E Toffoli**, *The Blanchfield pairing of colored links*, Indiana Univ. Math. J. 67 (2018) 2151–2180  MR  Zbl

[10]  **A Conway**, **M Powell**, *Characterisation of homotopy ribbon discs*, Adv. Math. 391 (2021) art. id. 107960  MR  Zbl

[11]  **D J Crowley**, *The classification of highly connected manifolds in dimensions 7 and 15*, PhD thesis, Indiana University (2002)  MR  Available at  https://www.proquest.com/docview/305508775

[12]  **D Crowley**, **J Sixt**, *Stably diffeomorphic manifolds and $l_{2q+1}(\mathbb{Z}[\pi])$*, Forum Math. 23 (2011) 483–538  MR  Zbl

[13]  **B Farb**, **D Margalit**, *A primer on mapping class groups*, Princeton Mathematical Series 49, Princeton Univ. Press (2012)  MR  Zbl

[14]  **R Fintushel**, **R J Stern**, *Surfaces in 4–manifolds*, Math. Res. Lett. 4 (1997) 907–914  MR  Zbl

[15]  **M H Freedman**, **F Quinn**, *Topology of 4–manifolds*, Princeton Mathematical Series 39, Princeton Univ. Press (1990)  MR  Zbl

[16]  **S Friedl**, **C Leidy**, **M Nagel**, **M Powell**, *Twisted Blanchfield pairings and decompositions of 3-manifolds*, Homology Homotopy Appl. 19 (2017) 275–287  MR  Zbl

[17]  **S Friedl**, **M Nagel**, **P Orson**, **M Powell**, *A survey of the foundations of four-manifold theory in the topological category*, preprint (2019)  arXiv 1910.07372

[18]  **S Friedl**, **M Powell**, *A calculation of Blanchfield pairings of* 3*–manifolds and knots*, Mosc. Math. J. 17 (2017) 59–77   MR   Zbl

[19]  **R E Gompf**, **A I Stipsicz**, 4*–Manifolds and Kirby calculus*, Graduate Studies in Math. 20, Amer. Math. Soc. (1999)   MR   Zbl

[20]  **I Hambleton**, **M Kreck**, **P Teichner**, *Topological* 4*–manifolds with geometrically two-dimensional fundamental groups*, J. Topol. Anal. 1 (2009) 123–151   MR   Zbl

[21]  **I Hambleton**, **P Teichner**, *A non-extended Hermitian form over* $\mathbb{Z}[\mathbb{Z}]$, Manuscripta Math. 93 (1997) 435–442   MR   Zbl

[22]  **J A Hillman**, **A Kawauchi**, *Unknotting orientable surfaces in the* 4*–sphere*, J. Knot Theory Ramifications 4 (1995) 213–224   MR   Zbl

[23]  **J F P Hudson**, *Embeddings of bounded manifolds*, Proc. Cambridge Philos. Soc. 72 (1972) 11–20   MR   Zbl

[24]  **A Juhász**, **M Miller**, **I Zemke**, *Transverse invariants and exotic surfaces in the* 4*–ball*, Geom. Topol. 25 (2021) 2963–3012   MR   Zbl

[25]  **A Juház**, **I Zemke**, *Stabilization distance bounds from link Floer homology*, preprint (2018)   arXiv 1810.09158

[26]  **D Kasprowski**, **M Powell**, **P Teichner**, *Algebraic criteria for stable diffeomorphism of spin* 4*–manifolds*, preprint (2020)   arXiv 2006.06127

[27]  **A Kawauchi**, *Splitting a* 4*–manifold with infinite cyclic fundamental group*, Osaka J. Math. 31 (1994) 489–495   MR   Zbl

[28]  **A Kawauchi**, *Splitting a* 4*–manifold with infinite cyclic fundamental group, revised*, J. Knot Theory Ramifications 22 (2013) art. id. 1350081   MR   Zbl

[29]  **Q Khan**, *Cancellation for* 4*–manifolds with virtually abelian fundamental group*, Topology Appl. 220 (2017) 14–30   MR   Zbl

[30]  **H J Kim**, *Modifying surfaces in* 4*–manifolds by twist spinning*, Geom. Topol. 10 (2006) 27–56   MR   Zbl

[31]  **H J Kim**, **D Ruberman**, *Smooth surfaces with non-simply-connected complements*, Algebr. Geom. Topol. 8 (2008) 2263–2287   MR   Zbl

[32]  **H J Kim**, **D Ruberman**, *Topological triviality of smoothly knotted surfaces in* 4*–manifolds*, Trans. Amer. Math. Soc. 360 (2008) 5869–5881   MR   Zbl

[33]  **R Kirby**, *Problems in low dimensional manifold theory*, from "Algebraic and geometric topology, II" (R J Milgram, editor), Proc. Sympos. Pure Math. 32, Amer. Math. Soc., Providence, RI (1978) 273–312   MR   Zbl

[34]  **M-A Knus**, *Quadratic and Hermitian forms over rings*, Grundl. Math. Wissen. 294, Springer (1991)   MR   Zbl

[35]  **K H Ko**, *A Seifert-matrix interpretation of Cappell and Shaneson's approach to link cobordisms*, Math. Proc. Cambridge Philos. Soc. 106 (1989) 531–545   MR   Zbl

[36] **M Kreck**, *Isotopy classes of diffeomorphisms of* $(k-1)$*–connected almost-parallel-izable* $2k$*–manifolds*, from "Algebraic topology" (J L Dupont, I H Madsen, editors), Lecture Notes in Math. 763, Springer (1979) 643–663  MR Zbl

[37] **M Kreck**, *Surgery and duality*, Ann. of Math. 149 (1999) 707–754  MR Zbl

[38] **T Y Lam**, *Serre's problem on projective modules*, Springer (2006)  MR Zbl

[39] **J Levine**, *Knot modules, I*, Trans. Amer. Math. Soc. 229 (1977) 1–50  MR Zbl

[40] **C Livingston**, *Surfaces bounding the unlink*, Michigan Math. J. 29 (1982) 289–298  MR Zbl

[41] **B A Magurn**, **W van der Kallen**, **L N Vaserstein**, *Absolute stable rank and Witt cancellation for noncommutative rings*, Invent. Math. 91 (1988) 525–542  MR Zbl

[42] **T E Mark**, *Knotted surfaces in* 4*–manifolds*, Forum Math. 25 (2013) 597–637  MR Zbl

[43] **T Oba**, *Surfaces in the* 4*–disk with the same boundary and fundamental group*, Math. Res. Lett. 27 (2020) 265–279  MR Zbl

[44] **B Perron**, *Pseudo-isotopies et isotopies en dimension quatre dans la catégorie topologique*, Topology 25 (1986) 381–397  MR Zbl

[45] **F Quinn**, *Isotopy of* 4*–manifolds*, J. Differential Geom. 24 (1986) 343–372  MR Zbl

[46] **A Ranicki**, *The algebraic theory of surgery, I: Foundations*, Proc. London Math. Soc. 40 (1980) 87–192  MR Zbl

[47] **A Ranicki**, *Exact sequences in the algebraic theory of surgery*, Mathematical Notes 26, Princeton Univ. Press (1981)  MR Zbl

[48] **R Stong**, **Z Wang**, *Self-homeomorphisms of* 4*–manifolds with fundamental group* $\mathbb{Z}$, Topology Appl. 106 (2000) 49–56  MR Zbl

[49] **N S Sunukjian**, *Surfaces in* 4*–manifolds: concordance, isotopy, and surgery*, Int. Math. Res. Not. 2015 (2015) 7950–7978  MR Zbl

Department of Mathematics, Massachussetts Institute of Technology
Cambridge, MA, United States

School of Mathematics and Statistics, University of Glasgow
Glasgow, United Kingdom

anthonyyconway@gmail.com,   mark.powell@glasgow.ac.uk

# GEOMETRY & TOPOLOGY

msp.org/gt

See inside back cover or msp.org/gt for submission instructions.

© 2023 Mathematical Sciences Publishers

# Geometry & Topology