

DUAL DECOMPOSITIONS OF 4-MANIFOLDS

FRANK QUINN

ABSTRACT. This paper concerns decompositions of smooth 4-manifolds as the union of two handlebodies, each with handles of index ≤ 2 . In dimensions ≥ 5 results of Smale (trivial π_1) and Wall (general π_1) describe analogous decompositions up to diffeomorphism in terms of homotopy type of skeleta or chain complexes. In dimension 4 we show the same data determines decompositions up to 2-deformation of their spines. In higher dimensions spine 2-deformation implies diffeomorphism, but in dimension 4 the fundamental group of the boundary is not determined. Sample results: (1.5) Two 2-complexes are (up to 2-deformation) spines of a dual decomposition of the 4-sphere if and only if they satisfy the conclusions of Alexander-Lefschetz duality ($H_1K \simeq H^2L$ and $H_2K \simeq H^1L$). (3.3) If $(N, \partial N)$ is 1-connected then there is a “pseudo” handle decomposition without 1-handles, in the sense that there is a pseudo collar $(M, \partial N)$ (a relative 2-handlebody with spine that 2-deforms to ∂N) and N is obtained from this by attaching handles of index ≥ 2 .

1. RESULTS

A *dual decomposition* of a 4-manifold is a description as the union of two 2-handlebodies along their boundary. The two pieces are “dual” in the sense that each is the complement of the other. There is a temptation to call these “4-dimensional Heegard decompositions.” However their properties and uses are very different from the 3-dimensional situation, and the use of the same term seems to be misleading.

This paper concerns existence of dual decompositions as measured by chain complexes and 2-complex spines. In these terms they are fairly flexible: Corollary 1.5, for instance, generalizes the theorem of Huck [H] that any two acyclic 2-complexes can be (up to 2-deformation) dually embedded in S^4 . Generally this continues the theme developed by C. T. C. Wall [W1, W2, W3, W4] that CW and handlebody structures are faithfully described by cellular chain complexes. This development was motivated by potential applications to knot theory and invariants of smooth 4-manifolds, see §2.3.

The paper is organized as follows: Results and definitions are given in this section. Section 2 lists facts and questions about dual decompositions. More detailed relative versions of the main results are given in §3. Section 4 begins the proof by showing how to align 1-skeleta of chain and CW complexes. The proof of the chain realization theorem is given in §5. Section 6 gives a characterization of chain equivalence of 2-complexes in terms of geometric moves. Finally the spine realization theorem is proved in §7.

Received by the editors October 2, 2000 and, in revised form, August 4, 2001.
2000 *Mathematics Subject Classification*. Primary 57R65, 57M20.

In the following statements N is a closed connected smooth 4-manifold, and chain complexes are finitely generated free based complexes over $Z\pi_1N$. If $K \rightarrow N$ is a CW complex then C_*K denotes the cellular chains of the cover of K induced from the universal cover of N . Other definitions are given after the statements, and relative versions are given in §3. We try to have hypotheses as weak as possible (chain complexes), and conclusions as strong as possible (2-deformation, which is probably stronger than simple homotopy).

1.1 Theorem (Chain realization). *Suppose D_* is a chain complex with a chain map $D_* \rightarrow C_*(N)$. Then D_* is simple chain equivalent to the cellular chains of one side of a dual decomposition of N if and only if it is homologically 2-dimensional and $H_0D \rightarrow H_0C_*N (= Z)$ is an isomorphism.*

In the next statement we hold the chains fixed and vary the spine. “Ambient deformations” of M do not change the spine of M or the chains of the complement, but can change the spine of the complement arbitrarily within these constraints.

1.2 Theorem (Spine realization). *Suppose $N = M \cup W$ is a dual decomposition and $K \rightarrow N$ is a 2-complex. Then there is an ambient deformation of M to a decomposition $M' \cup W'$ with a 2-deformation of K to the spine of W' if and only if there is a simple chain equivalence $C_*K \rightarrow C_*W$ that chain-homotopy commutes with the inclusions.*

The remainder of this section contains corollaries and definitions.

1.3 Corollary (Spine characterization). *A finite CW 2-complex $K \rightarrow N$ 2-deforms to the spine of half a dual decomposition if and only if K is connected and $\pi_1K \rightarrow \pi_1N$ is onto.*

Proof. If $N = M \cup W$ then W is connected and onto π_1N because its 1-skeleton is a 1-skeleton of N . For the converse note that connected and onto π_1N implies that the cover of K induced from the universal cover of N is connected. This means $H_0C_*K \rightarrow H_0C_*N \simeq Z$ is an isomorphism. According to the chain realization theorem there is a dual decomposition $N = M \cup W$ with $C_*K \simeq C_*W$ a simple equivalence. But then the spine realization theorem shows we can deform M to realize K (up to 2-deformation) as the spine of the complement. \square

In the next corollary we specify the spines of both sides simultaneously. This requires a hypothesis that encodes algebraic duality in N . Suppose $N = M \cup W$ is a dual decomposition. Then there is a diagram of chain complexes,

$$\begin{array}{ccccc}
 C_*M & \xrightarrow{\text{dual}} & C^{4-*}(M, \partial M) & \xrightarrow{\text{excision}} & C^{4-*}(N, W) \\
 \downarrow & & & & \downarrow \\
 C_*N & \xrightarrow{\text{dual}} & C^{4-*}N & \xrightarrow{=} & C^{4-*}N \\
 & & & & \downarrow \\
 & & & & C^{4-*}W
 \end{array}$$

in which the rows are chain equivalences and the right column is homotopy-exact. The corresponding diagram of homology may be more familiar; there chain equivalence of the rows corresponds to isomorphism on homology, and homotopy-exactness on the right gives the long exact sequence in cohomology.

Consider the composition from the upper left to the lower right. Homotopy exactness and the fact that it factors through the upper right complex gives a chain nullhomotopy of this composition. Conversely a chain nullhomotopy of the composition determines a lift of $C_*M \rightarrow C^{4-*}N$ into $C^{4-*}(N, W)$. The natural nullhomotopy specifies the lift coming from duality and excision, which is a simple equivalence.

We abstract this by saying a *simple algebraic duality* of K, L is a chain nullhomotopy of the corresponding composition, so that the induced lift is a simple equivalence. More explicitly the nullhomotopy is for the composition

$$C_*K \rightarrow C_*N \xrightarrow{\text{dual}} C^{4-*}N \rightarrow C^{4-*}L.$$

The nullhomotopy induces a lift $C_*K \rightarrow C^{4-*}(N, L)$, and we require this to be a simple equivalence. The next corollary shows this characterizes spines of dual decompositions.

1.4 Corollary (Characterization of dual spines). *2-dimensional CW complexes $K, L \rightarrow N$ 2-deform to spines of a dual decomposition if and only if there is a chain nullhomotopy giving simple algebraic duality on the chain level.*

The 2-deformations in the conclusion actually preserve the data in the sense that $K \rightarrow M \subset N$ and $L \rightarrow W \subset N$ are homotopic to the original maps, and these homotopies together with the canonical chain homotopy for $M \cup W$ give (up to chain homotopy) the chain homotopy in the hypothesis.

Proof of 1.4. The ‘only if’ part of the statement is explained above. For the converse suppose a chain homotopy is given. Use the first corollary to realize K as the spine of M in a decomposition $M \cup W$. The canonical duality for the decomposition gives a simple equivalence $C_*K \rightarrow C^{4-*}(N, W)$. The hypothesized duality for K, L gives a simple equivalence $C_*K \rightarrow C^{4-*}(N, L)$. Composing one with the inverse of the other and taking duals gives a simple equivalence $C_*(N, L) \rightarrow C_*(N, W)$. This chain-commutes with the maps from C_*N , so via quotients induces a simple equivalence $C_*L \rightarrow C_*W$. We can now apply the deformation theorem to ambiently 2-deform M to M' so that L 2-deforms to the skeleton of the new complement. This gives the desired dual decomposition. □

The special case of $N = S^4$ can be described more explicitly because its homology vanishes, and chain complexes over the trivial group are determined by homology. The conclusion is that Alexander-Lefschetz duality tells the whole story.

1.5 Corollary (Dual spines in S^4). *Two 2-complexes K, L occur (up to 2-deformation) as spines of a dual decomposition of S^4 if and only if they are connected and there are isomorphisms $H_1(K) \simeq H^2(L)$ and $H_2(K) \simeq H^1(L)$.*

Huck [H] used a direct explicit construction to realize arbitrary acyclic ($H_1K = H_2K = 0$) 2-complexes as spines. His proof could probably be elaborated to prove 1.5. Given a 2-complex K we note there is a particularly simple choice for dual spine in S^4 . Suppose $H_1K \simeq (\bigoplus_i Z/n_i) \oplus Z^{\beta_1}$ and $H_2(K) \simeq Z^{\beta_2}$. Then the 1-point union

$$L = \left(\bigvee_i S^1 \cup_{n_i} D^2 \right) \vee \left(\bigvee^{\beta_2} S^1 \right) \vee \left(\bigvee^{\beta_1} S^2 \right)$$

occurs (after 2-deformation) as a dual spine for K .

Proof of 1.5. Choose (necessarily nullhomotopic) maps $K, L \rightarrow S^4$ satisfying the hypotheses. A standard exact sequence computation shows there are isomorphisms $H_*K \simeq H^{4-*}(S^4, L)$. Chain complexes over the integers are determined up to chain homotopy equivalence by homology, so there is a chain equivalence $C_*K \simeq C^{4-*}(S^4, L)$. This is simple because the Whitehead group of $Z[1]$ is trivial. The composition $C_*K \rightarrow C^{4-*}L$ used in 1.4 is nullhomotopic because it factors through $C_*(S^4)$. As explained before 1.4, a nullhomotopy determines a lift $C_*K \rightarrow C^{4-*}(S^4, L)$. A random homotopy can be changed to one whose lift is the chain equivalence found above. This is a simple algebraic duality in the sense of 1.4 so implies the existence of a dual decomposition with the desired properties. \square

We now provide definitions for terms used in the statements of 1.1 and 1.2.

1.6 Definition (homologically 2-dimensional). A finitely generated free chain complex D_* over a ring R is *homologically n -dimensional* if $H_j D = 0$ for all $j \geq n+1$ and $H^{j+1} D = 0$.

Wall [W3], [W4] shows that a finitely generated free based complex is simple equivalent to an n -dimensional one if and only if it is homologically n -dimensional. He shows further that a finite CW complex is simple equivalent to an n -complex, for $n \neq 2$, if and only if the chain complex of the universal cover is homologically n -dimensional over the fundamental group. The 2-dimensional case is still mysterious. Note that 1.1 gives realizations of chain complexes by 2 CW complexes (the spine of half the dual decomposition), but only by homology equivalence over the fundamental group of N . These are almost never homotopy equivalences because the fundamental groups are usually different.

1.7 Definition (homotopy 2-deformation of complexes). Suppose K, L are CW complexes. A *2-deformation* $K \rightarrow L$ is a sequence of moves starting with K and ending with L . Each move is either an elementary expansion or collapse of dimension ≤ 2 , or change of attaching map of a 2-cell by homotopy. Explicitly, these expansions are attaching of a 1- or 2-cell by an endpoint or arc in the boundary, respectively, and the collapses remove cells attached in this special way. In some of the literature these moves are called “3-deformations” because homotopy of 2-cells is viewed as obtained by expansion and then slightly different collapse of a 3-cell. Our terminology follows Wall [W2]. Note that a 2-deformation determines a simple homotopy equivalence $K \rightarrow L$. The (generalized) Andrews-Curtis conjecture is that any simple equivalence of 2-complexes comes from a 2-deformation. No one expects this to be true, however.

In §6 we define a homological analog where attaching maps of 2-cells are allowed to change by homology in a cover. Unlike homotopy 2-deformation there is a global characterization of the resulting equivalence relation in terms of simple equivalence of chain complexes.

1.8 Definition (ambient 2-deformation of a sub-handlebody). These are defined for sub-handlebodies in a handlebody decomposition of a 4-manifold. The number and indexes of handles in the sub-handlebody and the complement do not change, so deforming a dual decomposition gives another dual decomposition. As the terminology suggests, the spine of the sub-handlebody changes by homotopy 2-deformation.

An *ambient deformation* of M in N is a sequence of moves of two types. The first type are standard handle moves in M : introduction or omission of cancelling

pairs of 0- and 1-handles, or 1- and 2-handles, and isotopy of the attaching maps of 2-handles. These moves do not change the underlying decomposition of N . The second type of moves are *homotopically trivial handle moves over handles in the complement*. To describe these we suppose A is a 2-handle in M and B is a 2-handle in the complement attached to ∂M disjointly from A . Let f, g be embedded paths from A to B in ∂M disjoint from $A \cup B$ except for their endpoints. Suppose also that they are homotopic rel ends by a homotopy in $M - A$. Obtain a 2-handle A' by pushing A over B twice: once along f , and once along g with opposite sign. Define M' to be $(M - A) \cup A'$.

If A' is obtained from A by a homotopically trivial handle move, then the attaching maps are homotopic in $M - A$. Therefore there is a 2-deformation between the spines of M and M' . The handles of the complement have been rearranged, possibly by more than 2-deformation, but no new ones have been added.

The moves actually used in the proof have quite a bit more structure. We have not tried to capture this in the definition since it plays no role here, but it will be important in refined versions of the theory.

In higher dimensions the homotopically trivial move would be equivalent to a handle move in M . In higher dimensions homotopy in $M - A$ would imply homotopy in $\partial M - A$, and homotopy of paths implies isotopy. But algebraically cancelling handle moves using *isotopic* curves actually cancels geometrically, and changes the attaching map of A by isotopy. In dimension 4 the fundamental groups of the boundary and inside may be different, and the paths are in the 3-manifold ∂M so may be knotted.

2. FACTS AND QUESTIONS

In this section we provide background and pose further questions

2.1 Facts.

- *Existence.* Compact smooth 4-manifolds have dual decompositions. To see this begin with a handlebody decomposition that is *indexed* in the sense that all k -handles are attached before any $(k + 1)$ -handles are attached. Divide the 2-handles into two sets, and define M to be the union of the 0- and 1-handles, and the 2-handles in one of these sets. Define W to be the rest of N , namely the 4- and 3-handles and the remaining 2-handles. The dual handles give a description of W as a 2-handlebody, and $N = M \cup W$.

Note the complete freedom in dividing up the 2-handles between the two sides. The 3-dimensional situation is more determined: since we want each side to be a 1-handlebody, one side must consist of all of the 0- and 1-handles, while the other must be all of the 2- and 3-handles. The extra degrees of freedom in dimension 4 should be thought of as partial compensation for the greater complexity of 2-handlebodies. Rather than studying a single dual decomposition, as in dimension 3, we should expect to have to modify it to improve the pieces.

- *Smoothness.* Donaldson, cf. [DK], has shown that not all 4-manifolds have smooth structures, and when they do the set of such structures is quite mysterious. However (equivalence classes of) handlebody structures correspond exactly to (diffeomorphism classes of) smooth structures. This is because handlebodies are determined by the attaching maps in one lower dimension, and attaching maps in 3-manifolds can be uniquely smoothed. Thus a dual decomposition encodes a smooth structure, and may give access to invariants of this structure.

- *Pictures.* Dual decompositions give explicit pictures of smooth 4-manifolds. The Kirby calculus, extended to include 1-handles, gives explicit pictorial descriptions of 4-dimensional 2-handlebodies in terms of links in the 3-sphere, see [GS]. The main theorem of the calculus is that the boundaries of two such handlebodies are diffeomorphic if and only if one link diagram can be deformed to the other using “Kirby moves.” The proof gives a sharper conclusion, namely that a particular sequence of moves determines a particular diffeomorphism (up to isotopy). A dual decomposition, including the glueing map on the boundary, can therefore be described as two link diagrams together with a sequence of Kirby moves relating them.

Previous descriptions of 4-manifolds have used link diagrams to represent the whole 2-skeleton. This is ok when there are no 3-handles, and in principle could be useful even when there are 3-handles because the attaching maps are fairly rigidly determined. In practice these attaching maps can be very complicated and hard to see. Whether a dual Kirby-move description is more useful remains to be seen.

2.2 Cautions.

- *2-handlebody existence.* There are no topological criteria for a 4-manifold to be a 2-handlebody. Thus we cannot decompose a manifold and deduce after the fact that it is a dual decomposition. There are easy homological criteria for n -handlebodies when $n \neq 2$ [W3], [W4], but we expect the analog to be false when $n = 2$. In particular there should be smooth 4-manifolds that are simple equivalent to 2-complexes but are not 2-handlebodies. Note that the simple equivalence hypothesis does not eliminate algebraic topology from the picture since the boundary 3-manifold may have a useful fundamental group. Casson has used representations of fundamental groups in compact connected Lie groups to detect counterexamples to a relative version.

- *2-handlebody uniqueness.* It seems highly unlikely that 2-handlebody structures are unique in the sense that one can be deformed to any other through 2-handlebodies, even though the corresponding result is true for k -handlebodies, $k \neq 2$ [W2]. In other words we expect smooth 4-manifolds with several different 2-handlebody “structures”, so the definition of “dual decomposition” must include the handlebody structures, not just the underlying manifolds.

- *2-deformation.* The results give embeddings of 2-complexes as spines, but only after 2-deformation. The original 2-complexes may not be embeddable at all, see [FKT]. Or they may be embeddable but with strong restrictions on the possible complementary spines [KT]. In either case preliminary deformations are necessary to destroy structures that lead to such restrictions.

- *Homology equivalence.* A homology equivalence of 2-complexes induces a chain homotopy equivalence of cellular chains, but the converse is not true. In fact a map inducing a homology equivalence must preserve quite a bit of other structure, such as quotients of fundamental groups by various commutator subgroups. Therefore “simple homology equivalence” cannot be substituted for “simple chain equivalence” in Theorem 1.2.

2.3 Questions.

- *Reconstruction.* In 2.1 we observed that a dual decomposition can be pictured as a pair of link diagrams together with a sequence of Kirby moves going between them. We need an explicit equivalence relation (in terms of “moves”) on such data

so that equivalence classes of data correspond to diffeomorphism classes of manifolds. Two sorts of moves are needed: one would show how to transfer a 2-handle from one side of the decomposition to the other. The other sort should explain how to relate two sequences of Kirby moves that yield isotopic diffeomorphisms of the boundary 3-manifold. The second part may be more accessible than one might expect. The *existence* of a sequence of Kirby moves corresponds to the existence of a diffeomorphism between 3-manifolds, and this seems to be algorithmically unsolvable in terms of the data given. However *equivalence* of two sequences would correspond to automorphisms of 3-manifolds, modulo some sort of bordism. These may be more detectable.

- *Geometric structures.* Eliashberg has observed that a Stein surface has a 2-handlebody structure and Gompf has characterized the handlebodies obtained this way, see [GS]. Akbulut and Matveyev [AM] have shown that all smooth 4-manifolds have dual decompositions with Stein pieces. Is there a deformation theory for these in the spirit of 1.2? The boundary of a Stein surface has a contact structure. Are there decompositions with Stein pieces in which the glueing map has some nice relationship with the contact structures? The goal is to find a comprehensible link between a combinatorial structure and Donaldson or Seiberg-Witten invariants. The combinatorial structure would be a dual decomposition with Stein pieces, together with a sequence of restricted Kirby moves giving a “contact-nice” glueing map. The link would go through gauge theory on Stein surfaces.

- *Uniqueness.* Suppose N has two dual decompositions, $M \cup W$ and $M' \cup W'$, and suppose there is a homotopy 2-deformation from the spine of M to that of M' . Is there an ambient 2-deformation of M to M' ? This seems unlikely but there are no immediately obvious obstructions.

A version of “concordance” gives another way to compare decompositions. A 2-deformation $K \rightarrow K'$ has a “track” which is a 3-complex that collapses to K and K' . Define a *concordance* between 2-handlebodies $M, M' \subset N$ to be an embedding of the track of a 2-deformation between the spines, in $N \times I$. There are obstructions to concordance, seen as follows. If M and M' are concordant then the complements are related: $N \times \{0\} - M \rightarrow N \times I - \text{track} \leftarrow N \times \{1\} - M'$ are homology equivalences with $Z[\pi_1 N]$ coefficients. But generally such homology equivalences do not exist. More explicitly, the spine realization theorem shows 2-complexes can be realized as spines of the complement provided only that they have the right homology. Homology equivalence through maps induces, by Stallings’ theorem, isomorphism on fundamental groups modulo a lower central series construction. But there are 2-complexes with the same homology and different π_1 modulo lower central series. This situation is familiar in knot theory. Note the concordance question is “high-dimensional” since it concerns embedding in a 5-manifold.

3. RELATIVE VERSIONS

First recall that a relative 2-handlebody structure on $(M, \partial_0 M)$ is a description as a collar $\partial_0 M \times I$ with handles of index ≤ 2 attached on $\partial_0 M \times \{1\}$. The result has $\partial_0 M \times \{0\}$ embedded in its boundary, and this is identified with $\partial_0 M \subset \partial M$.

A dual decomposition of a compact 4-manifold *with boundary* is a decomposition $N = M \cup W$ into codimension 0 submanifolds, together with relative 2-handlebody decompositions on $(M, M \cap \partial N)$ and $(W, W \cap \partial N)$.

In this description $\partial_0 M = M \cap \partial N$, so the handlebodies are joined along their “top boundaries.” The collars on $\partial_0 M$, $\partial_0 W$ fit together to give a collar of ∂N . Alternatively we could start in the middle and work outward. Consider the interface $X = M \cap W$, and thicken it to a collar $X \times I \subset N$. Duals of the handles used above describe M as obtained from X by attaching 2-, 3-, and 4-handles on $X \times \{1\}$, and similarly N is obtained by attaching handles to $X \times \{0\}$.

In the following N is a compact smooth 4-manifold and “chain complex” means finitely generated free based complex over $Z[\pi_1 N]$. The statements focus on geometric conclusions from algebraic hypotheses. The algebraic hypotheses can be sharpened to be necessary as well as sufficient, but trying to include this in the statements makes them unnecessarily complicated.

3.1 Theorem (Relative chain realization). *Suppose a decomposition $\partial N = A \cup V$ is given, D_* is a homologically 2-dimensional chain complex, and $D_* \rightarrow C_*(N, A)$ induces isomorphism on H_0 , epimorphism on H_1 . Then there exists a dual decomposition $N = M \cup W$ with $M \cap \partial N = A$ and a simple chain equivalence $D_* \rightarrow C_*(M, A)$*

Refinement. *The composition $D_* \rightarrow C_*(M, A) \rightarrow C_*(N, A)$ is chain homotopic to the original chain map.*

The point of the refinement is that the geometric decomposition realizes the data: we only have to change the chain map by homotopy to geometrically realize it. It sometimes happens in low dimensions that existence of data implies geometric conclusions, but corresponding to different data. The discrepancy between given and realized data then gives rise to higher-order obstructions. Fortunately this does not happen here.

Setting $\partial N = \emptyset$ in 3.1 gives 1.1, except for the H_1 hypothesis. However the homology is with $Z\pi_1 N$ coefficients, or in other words homology of the universal cover of N . This is simply connected so H_1 vanishes, and in this case the hypothesis holds trivially.

To prepare for the deformation statement we note that ambient deformation does not change the boundary decomposition $\partial N = \partial_0 M \cup \partial_0 W$. It also does not change the chains of the complement, in the sense that if $(M, \partial_0 M)$ ambiently 2-deforms to $(M', \partial_0 M')$, and the complements are $(W, \partial_0 W)$ and $(W', \partial_0 W')$ respectively, then there is a canonical (up to chain homotopy) simple chain equivalence $C_*(W, \partial_0 W) \rightarrow C_*(W', \partial_0 W')$, together with a chain homotopy making the diagram induced by inclusions

$$\begin{array}{ccc} C_*(W, \partial_0 W) & \longrightarrow & C_*(W', \partial_0 W') \\ \downarrow & & \downarrow \\ C_*(N, \partial_0 W) & \xrightarrow{=} & C_*(N, \partial_0 W) \end{array}$$

commute. This can be seen either directly from the definition of ambient 2-deformation or by using duality and the corresponding fact for the 2-deformation $M \rightarrow M'$.

3.2 Theorem (Relative spine realization). *Suppose*

$$(N, \partial N) = (M, \partial_0 M) \cup (W, \partial_0 W)$$

is a dual decomposition, $(K, \partial_0 W) \rightarrow (N, \partial_0 W)$ is a map of a relative 2-complex, and there is a simple chain equivalence $C_(K, \partial_0 W) \rightarrow C_*(W, \partial_0 W)$ together with*

a chain homotopy making the diagram induced by inclusions commute. Then there is an ambient 2-deformation of M to give $(M', \partial_0 M) \cup (W', \partial_0 W)$, and a relative 2-deformation from $(K, \partial_0 W)$ to the spine of $(W', \partial_0 W)$.

Refinement. *There are*

1. a homotopy from the composition of the 2-deformation with the inclusion, $(K, \partial_0 W) \rightarrow (W', \partial_0 W) \rightarrow (N, \partial_0 W)$, to the original map;
2. a chain homotopy from given chain equivalence $C_*(K, \partial_0 W) \rightarrow C_*(W, \partial_0 W)$ to the the composition $C_*(K, \partial_0 W) \rightarrow C_*(W', \partial_0 W) \rightarrow C_*(W, \partial_0 W)$, where the first map is induced by the 2-deformation to the skeleton and the second is dual to the ambient 2-deformation of M ; and
3. note that combining the homotopies of (1) and (2) give a chain homotopy making the diagram induced by inclusions commute. Then there is a chain homotopy between this homotopy and the one given in the data.

Again the refinement asserts that the output realizes the input data, so no higher obstructions arise from discrepancies between the two. From a technical point of view the extra precision becomes important in inductions or other arguments applying the theorem several times.

To illustrate use of the relative versions we give a version of “geometrical connectivity” for 4-manifolds. A special case of Wall’s results [W1] is that in higher dimensions a 1-connected manifold pair is obtained from a collar of the submanifold of the boundary by adding handles of index ≥ 2 . Casson has shown that this is not generally true for 4-manifolds. However there is a version where the handles are added to something which is not quite a collar on the boundary.

3.2 Corollary. *Suppose N is a compact smooth 4-manifold, $\partial_0 N \subset \partial N$ is a codimension 0 submanifold, and $(N, \partial_0 N)$ is relatively 1-connected. Then there is a “fake collar”, ie a codimension 0 submanifold $M \subset N$ with $M \cap \partial N = \partial_0 N$ and $(M, \partial_0 N)$ a relative 2-handlebody whose skeleton 2-deforms rel $\partial_0 N$ to $\partial_0 N$, and so that N is obtained from M by adding handles of index ≥ 2 .*

Proof. Let $\partial_1 N$ denote the closure of the complement of $\partial_0 N$ in ∂N . The 1-connected hypothesis is equivalent to $C_*(N, \partial_1 N)$ being homologically 2-dimensional. Applying the realization theorem gives a relative dual decomposition $N = M \cup W$ with the inclusion $C_*(W, \partial_1 N) \rightarrow C_*(N, \partial_1 N)$ a simple chain equivalence over $Z\pi_1 N$. The inclusion of the trivial relative 2-complex $(\partial_0 N, \partial_0 N) \rightarrow (M, \partial_0 N)$ gives a simple equivalence of chains, so by the deformation theorem W can be ambiently 2-deformed to give a new M whose skeleton relatively 2-deforms to $\partial_0 N$. The duals of the handles in $(W, \partial_1 N)$ describe N as obtained from M by adding handles of index ≥ 2 , so this decomposition satisfies the conclusions of Corollary 3.2. □

Note that $(W, \partial_1 N) \rightarrow (N, \partial_1 N)$ is a homology equivalence over $Z\pi_1 N$. This is a homotopy equivalence if and only if $\pi_1 W \rightarrow \pi_1 N$ is an isomorphism. In this case (N, W) has the homotopy type of a 2-complex, as well as being homologically 2-dimensional. Conversely if (N, V) were known to have the homotopy type of a relative 2-complex then we could realize that complex (up to 2-deformation) as the spine of W , and get $(W, \partial_1 N) \rightarrow (N, \partial_1 N)$ a homotopy equivalence.

We close the section with a few observations about the fake collars appearing in the corollary. First, 2-deformations can be realized by handle moves in a 5-manifold

[AC], so $(M \times I, \partial_0 N \times I) \simeq (\partial_0 N \times I^2, \partial_0 N \times I \times \{0\})$. The second observation is that $(M, \partial_0 N)$ is a simple homology H -cobordism with coefficients $\pi_1 M = \pi_1 \partial_0 N$, from $\partial_0 N$ to the closure of its complement in ∂M . This is usually not an s -cobordism because $\partial M - \partial_0 N$ usually has a different fundamental group. This must happen if $\pi_1 W \rightarrow \pi_1 N$ is not an isomorphism, and in particular if $(N, \partial_1 N)$ is not equivalent to a 2-complex.

4. 1-SKELETON ALIGNMENT

The objective here is to modify chain complexes to standardize their 1-skeletons. If the complex is cellular chains of a CW complex we want to achieve this standardization through 2-deformations of the CW complex. The result is similar to the low-dimensional cases of Wall [W2], see particularly Lemma 3B.

Fix a group π . “Chain complex” will mean a finitely generated free based $Z[\pi]$ complex which vanishes in negative degrees. CW complexes will be assumed to come with a homomorphism $\pi_1 K \rightarrow \pi$, and maps, deformations, etc. of CW complexes are understood to commute with these homomorphisms. The “cellular chains” of a CW complex are the cellular chains of the induced cover with covering group π , with a free basis obtained by lifting cells.

Lemma (1-skeleton alignment). *Suppose $f : C_* \rightarrow D_*$ is a chain map of free based $Z[\pi]$ complexes which is an isomorphism on H_0 and an epimorphism on H_1 . Then there is a chain homotopy commutative diagram*

$$\begin{array}{ccc} C_* & \xrightarrow{f} & D_* \\ \downarrow & & \downarrow \\ C'_* & \xrightarrow{f'} & D'_* \end{array}$$

so the vertical maps are simple equivalences and isomorphisms in degrees ≥ 3 , and f' is a basis-preserving isomorphism in degrees 0, 1.

If D_* is the cellular chains of a CW complex or pair (resp. 4-d handlebody) with connected π cover then $D_* \rightarrow D'_*$ can be arranged to be the map on cellular chains induced by a 2-deformation (resp. handlebody moves).

If both C_* and D_* are cellular chains of CW complexes or pairs (resp. 4-d handlebodies) with connected π covers and the isomorphism on H_0 is the identity, then both $C_* \rightarrow C'_*$ and $D_* \rightarrow D'_*$ can be arranged to be induced by 2-deformation (resp. handle moves).

In the last case recall that H_0 of a connected CW complex is canonically identified with Z . A chain map of cellular chains that induces -1 on H_0 with respect to these identifications cannot be made basis-preserving in degree 0 because there are no orientation-reversing endomorphisms of a point. In the relative case $H_0 = 0$ so the hypothesis is automatic.

Proof. We prove this by collecting data and then explicitly giving chain complexes, homotopies, etc. The first step is to align the 0-skeletons, i.e. arrange $C_0 \rightarrow D_0$ to be an isomorphism. Assuming this we then show how to align the 1-skeleton.

There are several procedures for aligning skeleta, depending on dimension and whether we want to realize the changes by 2-deformation. The simplest procedure fixes a dimension and changes both complexes one dimension higher. The bound on the changes means it can be used for 1-skeleta. This does not work for 0-skeleta

when one of the complexes is geometric because the changes cannot be realized geometrically. In that case we give a more elaborate argument that leaves one complex (the geometric one) unchanged and modifies the other. This argument changes the complex two dimensions above the target dimension, so cannot be used for 1-skeleta.

Semi-algebraic 0-skeleton alignment. Here we are not trying to realize changes in the domain by 2-deformation. The range complex is unchanged, so can be geometric. The data is a chain map of complexes

$$\begin{array}{ccc}
 C_2 & \xrightarrow{f_2} & D_2 \\
 \downarrow \partial & & \downarrow \partial \\
 C_1 & \xrightarrow{f_1} & D_1 \\
 \downarrow \partial & & \downarrow \partial \\
 C_0 & \xrightarrow{f_0} & D_0
 \end{array}$$

Since this is onto H_0 the map $(f_0 \ \partial) : D_0 \oplus C_1 \rightarrow D_0$ is onto. Since D_0 is free there is a splitting $(\begin{smallmatrix} g_0 \\ s \end{smallmatrix})$. In other words

$$(1) \quad f_0 g_0 + \partial s = \text{id}.$$

A similar argument using the fact that f_* induces an isomorphism on H_0 shows there is $t : C_0 \rightarrow C_1$ satisfying

$$(2) \quad g_0 f_0 + \partial t = \text{id}.$$

In the next step we improve this t . Consider

$$\begin{array}{ccc}
 & C_1 \oplus D & \\
 & \downarrow (f_1 \ \partial) & \\
 C_0 & \xrightarrow{sf_0 - f_1 t} & D_1
 \end{array}$$

We claim there is a lifting of the horizontal map to $(\begin{smallmatrix} \alpha \\ u \end{smallmatrix}) : C_0 \rightarrow (\ker \partial) \oplus D_2$. It suffices to show the image of the horizontal map is in the image of this subspace. Since f_* is onto on H_1 , the image of $(\ker \partial) \oplus D_2 \rightarrow D_1$ is the kernel of $\partial : D_1 \rightarrow D_0$. Therefore to verify the image of the horizontal map is in this image it suffices to verify the composition with ∂ vanishes. But $\partial(sf_0 - f_1 t) = (\partial s)f_0 - f_0(\partial t)$, which vanishes by (1) and (2).

Replace t by $t + \alpha$. Since α has image in $\ker \partial$, property (2) still holds. The factorization property becomes

$$(3) \quad \partial u + f_1 t = s f_0.$$

Now form the diagram

$$\begin{array}{ccc}
 C_2 \oplus C_0 & \xrightarrow{(f_2 \ u)} & D_2 \\
 \downarrow \begin{pmatrix} \partial & -t \\ 0 & f_0 \end{pmatrix} & & \downarrow \partial \\
 C_1 \oplus D_0 & \xrightarrow{(f_1 \ s)} & D_1 \\
 \downarrow (f_0 \partial \ \partial s) & & \downarrow \partial \\
 D_0 & \xrightarrow{\text{id}} & D_0
 \end{array}$$

The vertical sequences continue up with higher-degree terms of C_* , D_* , if there are any. Identity (3) shows the left vertical composition is trivial, so gives a chain complex, and also that the upper square commutes so the horizontal maps give a chain map.

Define the left-hand complex to be C'_* and the map to be f'_* . Define $D'_* = D_*$; since it is unchanged the change is realized by a 2-deformation when D_* is geometric. f'_* is a based isomorphism in degree 0. To finish the construction we need a simple equivalence $C_* \rightarrow C'_*$.

There is a natural chain map of C_* into C'_* ,

$$\begin{array}{ccc}
 C_2 & \xrightarrow{\begin{pmatrix} \text{id} \\ 0 \end{pmatrix}} & C_2 \oplus C_0 \\
 \downarrow \partial & & \downarrow \begin{pmatrix} \partial & -t \\ 0 & f_0 \end{pmatrix} \\
 C_2 & \xrightarrow{\begin{pmatrix} \text{id} \\ 0 \end{pmatrix}} & C_1 \oplus D_0 \\
 \downarrow \partial & & \downarrow (f_0 \partial \ \partial s) \\
 C_0 & \xrightarrow{f_0} & D_0
 \end{array}$$

It is easy to see that composition of this with f'_* gives f_* , so it suffices to show this is a simple equivalence. This can be done by adding the trivial complex $\text{id} : C_0 \rightarrow C_0$ in degrees 0 and 1, then doing elementary moves to transform the result to C_* plus the trivial complex

$$C_0 \xrightarrow{\begin{pmatrix} \text{id} \\ 0 \end{pmatrix}} C_0 \oplus D_0 \xrightarrow{(0 \ \text{id})} D_0.$$

We omit the details of this.

Geometric 0-skeleton alignment. Here both C_* and D_* are geometric, and we want the modifications realized by geometric 2-deformations. We begin with the non-relative case. Since both complexes are connected there are 2-deformations to CW complexes with single 0-cells. In the manifold case there are handle moves to handlebody structures with single 0-handles. Let C'_* , D'_* be the chains of these complexes.

Since the π covers of the spaces are connected the single basepoints give a single copy of $Z\pi$ in degree 0. Thus the chain map has the form

$$\begin{array}{ccc} C_1 & \xrightarrow{f_1} & D_1 \\ \downarrow \partial & & \downarrow \partial \\ Z\pi & \xrightarrow{f_0} & Z\pi \end{array}$$

By hypothesis f_* induces the identity on H_0 . So does the identity map on $Z\pi$, so the image of $(\text{id} - f_0) : Z\pi \rightarrow Z\pi$ lies in the image of the boundary in D_* . Let $s : Z\pi \rightarrow C_1$ be a lift. Then s defines a chain homotopy of f_* to the chain map

$$\begin{array}{ccc} C_1 & \xrightarrow{f_1+s\partial} & D_1 \\ \downarrow \partial & & \downarrow \partial \\ Z\pi & \xrightarrow{\text{id}} & Z\pi \end{array}$$

This is a basis-preserving isomorphism in degree 0 so serves as f'_* .

We now consider the relative case, where C and D are the relative chains of pairs. Since the spaces are connected there are 2-deformations to relative 2-complexes that have no 0-cells. Since the chain groups in dimension 0 are now trivial the chain map trivially induces an isomorphism. However there is a little more to do for pairs (K, A) when A is not connected. The condition actually used in the 1-skeleton argument is that all 1-cells are attached at a single point. If A is not connected choose a linear order of the components of A and deform the CW structure so there is a 1-cell joining each component to the next one. All other 1-cells can be assumed to be attached at a point. The alignment needed is to deform the other CW complex to have a similar structure, so that the chain map preserves the classes of the arcs connecting pieces of A . This begins by finding maps of arcs rel endpoints into the complex rel A that are liftings on the chain level. The mapping cylinder of these maps has a CW structure with new 1- and 2-cells. This 2-deforms to the complex, but now has 1-cells in the right place algebraically. Next we manipulate the CW structure to get all other 1-cells to be attached at a point. This can be done to end up with the desired alignment of the special 1-cells. We omit details since the argument is basically the same as others in this section.

Algebraic 1-skeleton alignment. The hypotheses are as in the lemma, and additionally f_0 is a basis-preserving isomorphism. To reflect this we denote the bottom term of D_* also by C_0 , and $f_0 = \text{id}$. Here we do algebraic moves to get a basis-preserving isomorphism in degree 1 as well. First we collect some data.

Since f_0 and $H_0(f_*)$ are isomorphisms, f_0 induces an isomorphism on the images $\partial C_1 \rightarrow \partial C_2$. This together with the hypothesis that $H_1(f_*)$ is onto shows that $f_1 : C_1 \rightarrow D_1/\partial D_2$ is onto. This means there is a lift $g : D_1 \rightarrow C_1$ so $f_1g = \text{id}$ modulo ∂D_2 . This in turn implies a lift $s : D_1 \rightarrow D_2$ so

$$(1) \quad \partial s = \text{id} - f_1g.$$

Now consider the diagram

$$\begin{array}{ccccccc}
 C_2 \oplus D_1 & \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} & C_2 \oplus D_1 & \xrightarrow{\begin{pmatrix} \partial & -g \\ f & s \end{pmatrix}} & C_1 \oplus D_2 & \xleftarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} & C_1 \oplus D_2 \\
 \downarrow \begin{pmatrix} \partial & 0 \\ 0 & 1 \end{pmatrix} & & \downarrow \begin{pmatrix} \partial & -g \\ 0 & 1 \end{pmatrix} & & \downarrow \begin{pmatrix} \partial & 0 \\ -f & 1 \end{pmatrix} & & \downarrow \begin{pmatrix} \partial & 0 \\ 0 & 1 \end{pmatrix} \\
 C_1 \oplus D_1 & \xrightarrow{\begin{pmatrix} 1 & -g \\ 0 & 1 \end{pmatrix}} & C_1 \oplus D_1 & \xrightarrow{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}} & C_1 \oplus D_1 & \xleftarrow{\begin{pmatrix} 1 & 0 \\ -f & 1 \end{pmatrix}} & C_1 \oplus D_1 \\
 \downarrow \begin{pmatrix} \partial & 0 \end{pmatrix} & & \downarrow \begin{pmatrix} \partial & \partial \end{pmatrix} & & \downarrow \begin{pmatrix} \partial & \partial \end{pmatrix} & & \downarrow \begin{pmatrix} \partial & 0 \end{pmatrix} \\
 C_0 & \longrightarrow & C_0 & \longrightarrow & C_0 & \longleftarrow & C_0
 \end{array}$$

The identity (1) shows that the columns are all chain complexes and the horizontal maps give chain maps. Let the second and third columns be C'_* and D'_* respectively, and the map between them f'_* . Note f'_* is a basis-preserving isomorphism in degrees 0 and 1. The first column is a stabilization of C_* and the map to the second column is a simple equivalence. Thus this gives the simple equivalence $C_* \rightarrow C'_*$. Similarly the right column is a stabilization of D_* and the map to the third column is simple. This gives $D_* \rightarrow D'_*$. It only remains to check that the square obtained by adding the original $f_* : C_* \rightarrow D_*$ chain-homotopy commutes. This too is a simple consequence of (1). These data therefore give the conclusion of the lemma.

Geometric 1-skeleton alignment. By construction the chain maps $C_* \rightarrow C'_*$ and $D_* \rightarrow D'_*$ consist of stabilizations adding cancelling 1- and 2-dimensional generators, and endomorphisms using elementary (upper or lower triangular) matrices on the 1-dimensional generators. Stabilization can be realized by elementary expansions, or introduction of cancelling 1-, 2-handle pairs in the handlebody case. 1-cells or 1-handles all attached to a single 0-cell or 0-handle can be moved over each other by homotopy (resp. isotopy) to realize elementary-matrix endomorphisms of the chain complex. Therefore if one or both complexes starts out as cellular chains then the modified complexes are realized by 2-deformation (resp. handle moves).

This completes the proof of the 1-skeleton alignment lemma.

5. PROOF OF THE REALIZATION THEOREM

Suppose, as in the statement of 3.1, that D_* is a homologically 2-dimensional $Z\pi_1 N$ -complex and $D_* \rightarrow C_*(N, A)$ induces isomorphism on H_0 , epimorphism on H_1 . According to Wall [W3], [W4], it is simple equivalent to a 2-dimensional chain complex. Use the 1-skeleton alignment lemma to change D by simple equivalence to another 2-dimensional complex, and the handle decomposition on (N, A) by handle moves, so $D_* \rightarrow C_*(N, A)$ is a basis-preserving isomorphism in degrees 0 and 1.

We claim the handlebody structure can be further changed so $D_2 \rightarrow C_2(N, A)$ is a basis-preserving isomorphism on a summand. Assuming this we complete the proof of the theorem. Simply take M to be the union of the 0-handles, the 1-handles, and the 2-handles whose corresponding basis elements in $C_2(N, A)$ come from D_2 . Then $C_*(M, A)$ is the image of D_* in $C_*(N, A)$, and the map $D_* \rightarrow C_*(M, A)$ is a basis-preserving isomorphism and therefore a simple equivalence. Since M contains all the 0- and 1-handles the complement is a 2-handlebody, and gives a dual decomposition.

We now prove the claim. The situation is a chain map which is a based isomorphism in degrees 0, 1.

$$\begin{array}{ccccccc}
 D_2 & \longrightarrow & D_1 & \longrightarrow & D_0 & & \\
 \downarrow f & & \downarrow = & & \downarrow = & & \\
 C_3(N, A) & \longrightarrow & C_2(N, A) & \longrightarrow & C_1(N, A) & \longrightarrow & C_0(N, A)
 \end{array}$$

Introduce in N cancelling 2, 3-handle pairs indexed by the basis of D_2 . Do elementary moves among the 2-handles to change the cellular chains by the chain map

$$\begin{array}{ccccccc}
 C_3 \oplus D_2 & \xrightarrow{\begin{pmatrix} \partial & 0 \\ 0 & 1 \end{pmatrix}} & C_2 \oplus D_2 & \xrightarrow{\begin{pmatrix} \partial & 0 \end{pmatrix}} & C_1 & \longrightarrow & \dots \\
 \downarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & & \downarrow \begin{pmatrix} 1 & -f \\ 0 & 1 \end{pmatrix} & & \downarrow = & & \\
 C_3 \oplus D_2 & \xrightarrow{\begin{pmatrix} \partial & -f \\ 0 & 1 \end{pmatrix}} & C_2 \oplus D_2 & \xrightarrow{\begin{pmatrix} \partial & \partial \end{pmatrix}} & C_1 & \longrightarrow & \dots
 \end{array}$$

The composition of the original map from D into this gives

$$\begin{array}{ccccccc}
 D_2 & \longrightarrow & D_1 & \longrightarrow & \dots & & \\
 \downarrow \begin{pmatrix} f \\ 0 \end{pmatrix} & & \downarrow = & & & & \\
 C_3 \oplus D_2 & \xrightarrow{\begin{pmatrix} \partial & -f \\ 0 & 1 \end{pmatrix}} & C_2 \oplus D_2 & \xrightarrow{\begin{pmatrix} \partial & \partial \end{pmatrix}} & C_1 & \longrightarrow & \dots
 \end{array}$$

But now $\begin{pmatrix} 0 \\ 1 \end{pmatrix} : D_2 \rightarrow C_3 \oplus D_2$ is a chain homotopy that changes this chain map to be the inclusion of the second summand in dimension 2. This is a basis-preserving injection to a based summand of the 2-chains, proving the claim and the theorem.

6. HOMOLOGICAL 2-DEFORMATION

This is a characterization of simple chain equivalence of 2-complexes in terms of geometric moves. To some extent it is a substitute for the unknown and probably false characterization of simple homotopy by 2-deformation. This provides a new derivation and wider context for the “s-moves” shown to give simple homotopy in [Q]. Homological 2-deformations are used in the proof of the deformation theorem 3.2.

Definition (homological 2-deformation). Suppose π is a group and K is a 2-complex with a homomorphism $\pi_1 K \rightarrow \pi$. A homological 2-deformation (with $Z\pi$ coefficients) is a sequence of moves, each of which is either a 2-deformation or change of attaching map of a 2-cell by $Z\pi$ homology in the 1-skeleton.

We describe this last move more explicitly. Suppose $\alpha, \beta : S^1 \rightarrow K^1$ are maps with the same basepoint. Here K^1 denotes the 1-skeleton of K . These maps differ by $Z\pi$ homology in the 1-skeleton if the composition $\alpha^{-1}\beta$ extends to a map of an orientable surface with one boundary component $S \rightarrow K^1$, and this map lifts to the π cover of K^1 . The lifting property is equivalent to triviality of the composition on $\pi_1, \pi_1 S \rightarrow \pi_1 K^1 \rightarrow \pi$.

Proposition (Homological deformation is chain equivalence). *Suppose π is a group and K, L are 2-complexes with homomorphisms $\pi_1 \rightarrow \pi$. A homological 2-deformation $K \rightarrow L$ over π induces a simple chain equivalence of $Z\pi$ cellular*

chain complexes $C_*K \rightarrow C_*L$. Conversely if K and L are connected and the homomorphisms to π are onto, then any simple chain equivalence is realized (up to chain homotopy) by a homological 2-deformation.

A straightforward relative version also holds.

Proof. For the first statement recall that 2-deformations induce simple chain equivalences over π_1 and therefore over any other coefficients. Changing attaching map of a 2-cell by homology does not change the cellular chain complex at all. The only issue is the boundary homomorphism $C_2 \rightarrow C_1$, which is defined using homology classes in the 1-skeleton. Therefore a homological 2-deformation also induces a simple chain equivalence.

For the other direction first apply the 1-skeleton alignment lemma. This gives 2-deformations to K' and L' together with a chain map of cellular chain complexes (over π) that is a basis-preserving isomorphism on C_0 and C_1 . Since it is a chain homotopy equivalence and the complexes are 2-dimensional it must be an isomorphism on C_2 as well. Since the equivalence is simple this isomorphism is also simple. This means that (possibly after stabilization) it is a product of elementary matrices and diagonal matrices with group elements on the diagonal. Such a matrix can be realized by homotopies of attaching maps and changing choice of liftings used as bases. Doing these moves to K gives K, L and a basis-preserving isomorphism of cellular chains.

There is a unique (up to homotopy) way to identify the 1-skeletons of K, L so that the identification induces the chain isomorphism. We now regard K, L as 2-complexes with the same 1-skeleton. The chain isomorphism gives a bijection between 2-cells so corresponding cells have the same boundaries in C_1K^1 . But C_1 is defined to be $H_1(K^1, K^0; Z\pi)$, so the attaching maps are homologous in the π cover of K^1 . Equivalently the difference is nullhomologous. But a map of S^1 into a space is nullhomologous if and only if it extends to a map of an oriented surface. Therefore the attaching maps in the two complexes differ by oriented surfaces as specified in the definition, so they are obtained from each other by homological 2-deformation.

Note the final homological moves realize the chain map (which is by then the identity). This chain map is homotopic to the original, so the original is realized up to homotopy. \square

6.2 Simple homotopy equivalence. A simple *homotopy* equivalence of 2-complexes $K \rightarrow L$ is a map which is an isomorphism on π_1 and a simple equivalence on cellular chain complexes over π_1L . This is in particular a homology equivalence so comes from a homological 2-deformation. However the intermediate stages in a homological 2-deformation will usually not be homotopy equivalences (the fundamental groups change) so further data is needed to encode the fact that the composition is. The proof shows that the homological moves can be done together, so there are maps

$$K \rightarrow K' \rightarrow L' \rightarrow L$$

whose composition is the given map, the first and last are 2-deformations, and $K' \rightarrow L'$ changes (all) 2-cell attaching maps by $Z\pi_1$ homology in the 1-skeleton.

$Z\pi_1$ homology corresponds to surfaces whose π_1 maps trivially into both K' and L' . Therefore curves on the surface extend to maps of 2-disks into both K'

and L' . Conversely if there are enough 2-disks on the surface to surger away the extra π_1 then this shows the attaching maps are homotopic and the complexes are homotopy equivalent. Thus simple *homotopy* equivalence can be characterized as 2-deformations, and simultaneous homologies of attaching maps together with appropriate nullhomotopies of curves on the homology surfaces. An explicit description of this data is given in [Q], where it is called an “s-move.”

7. PROOF OF THE DEFORMATION THEOREM

The data is a decomposition $(N, \partial N) = (M, A) \cup (W, V)$, a map of a relative 2-complex $(K, A) \rightarrow (N, A)$, and a simple equivalence of chain complexes over $Z\pi_1 N$, $C_*(K, A) \rightarrow C_*(M, A)$.

The first step is to use the 1-skeleton alignment lemma: change (K, A) by relative 2-deformation and (M, A) by handle moves so the chain map becomes a basis-preserving isomorphism in degrees 0 and 1. Then since the complexes are 2-dimensional and the map is an equivalence, it must be an isomorphism in degree 1 too. Since the chain map is simple and the other degrees are basis-preserving, this isomorphism is simple. This means that after stabilization it is a product of a diagonal matrix with group elements on the diagonal, and elementary matrices. The diagonal matrix can be realized by changing choice of lift to the cover used as basis element. Elementary matrices can be realized by handle slides in M , or homotopy of attaching maps in K . After doing this the chain map becomes a basis-preserving isomorphism in degree 2 as well.

Choose a homotopy equivalence of (relative) 1-skeletons $(K^1, A) \rightarrow (M^1, A)$ realizing the bijection of basis elements in the cellular chain complexes. Use this to replace the 1-skeleton of K , and regard K as obtained from M^1 by attaching 2-cells. The bijection of bases in C_* gives a correspondence between 2-cells of (K, A) and 2-handles of (M, A) . Since the boundary homomorphisms agree the attaching maps are homologous in the 1-skeleton.

Choose surfaces mapping into the 1-skeleton representing the homologies between attaching maps of K -2-cells and M -2-handles. Choose symplectic basis curves for the fundamental groups of these surfaces. These are collections of simple closed curves in the surfaces so that each curve intersects exactly one other curve, in a single point. Since the homology takes place in the cover corresponding to $\pi_1 N$ each of these curves is nullhomotopic in N . Choose nullhomotopies. These extend the surfaces to maps of capped surfaces into N . See [FQ] for details about capped surfaces.

The objective is to do ambient 2-deformations of W to change the 2-handles of M so their attaching maps become homotopic to the 2-cells of K . The capped surfaces are used to see how to do this.

The surfaces map into the 1-skeleton. After homotopy the caps can be arranged to have subdisks that map to parallels of the cores of the 2-handles, and the complement of these subdisks maps into M^1 . A cap can be split along an arc by joining the base surface to itself by a tube along the arc. The fragments give caps for the new surface, with dual caps given by parallel copies of the dual to the original. Repeated splitting gives a capped surface so that each cap has a single subdisk going over a 2-handle.

The next step is a subdivision trick in W that reduces to the case where all caps go over (duals of) 2-handles in W , and each such 2-handle intersects a single

cap. This is in preparation for homotopically trivial handle moves in W which will change the decomposition. In principle a lot could be done without changing the decomposition, using handle moves in M . Suppose there is a cap that goes over a handle h_i in M , and this cap is on the homology surface between an attaching map in K and a handle h_j in M , with $i \neq j$. The cap describes how to do two handle moves of h_j over h_i to move past the cap and reduce the genus of the homology surface by one. The remainder of the surface shows this move does not change the chain complex data. However moving h_j disturbs all the caps passing over it, adding new subdisks passing over h_i . Care would be necessary to be sure we are making progress. Since we have to do the more disruptive moves anyway, we use them for everything and dodge this point.

We describe the subdivision trick. Suppose h is a 2-handle of the decomposition of (N, A) that does not lie in M . In other words it is dual to a 2-handle in the structure on (W, V) , and is given by an embedding $h : (D^2 \times D^2, S^1 \times D^2) \rightarrow (W, \partial W - V)$. Suppose n caps pass over h . This means they intersect the image in disks $D^2 \times \{p_i\}$ for p_i points in the other D^2 factor. Think of D^2 as a handlebody with n 0-handles and $n - 1$ 1-handles, for instance with spine an interval subdivided to have n vertices. Identify (by isotopy) the p_i with the center points of the 0-handles. Taking the product of (D^2, S^1) with this structure gives a subdivision of h into n 2-handles and $n - 1$ 3-handles. This separates the caps so each passes over a distinct 2-handle. Recall that h is the dual of a handle in the structure on (W, V) . The dual of this operation is to split h^* into 1- and 2-handles, so the 2-handlebody condition is preserved.

The point here is that “passing over” is not symmetric under duality. If a disk passes over a 2-handle then it is parallel to the core. In the dual the core becomes the transverse disk which does not “go over” the dual. Deforming it down to the attaching region gives a picture of it “going under” the dual in the sense that it cuts through the attaching region but does not enter the interior of the handle. Getting confused about these dual pictures is a mistake that has been discovered by all the foremost 4-dimensional topologists.

There is a variation on the subdivision trick that applies when h is a handle in M . Again suppose n caps pass over it. Subdivide it into $n + 1$ 2-handles and n 3-handles so that each cap gets its own 2-handle and there is one left over. Shrinking h down to the left-over handle changes M by ambient isotopy and puts the rest of the handles into W . There they are dual to n each 1- and 2-handles in (W, V) . W has not been changed up to diffeomorphism, and in fact these new handles geometrically cancel.

We now have caps going over distinct 2-handles lying in W . We further improve the surfaces after introducing some notation. Let S_j denote the surface with boundary the composition of loops $(\partial h_j)(\partial d_j)^{-1}$, where d is a 2-cell of K and h is a 2-handle of M . Let $D_{j,*}$ denote the caps attached to S_j . We want:

1. each surface is divided into pieces $S_j = S'_j \cup R_j$ that intersect in an arc;
2. R_j is a disk containing ∂d_j ;
3. S'_j contains ∂h_j and the boundaries of all the cap disks;
4. the sets $S'_j \cup D_{j,*}$ are disjointly embedded; and
5. S'_j is contained in $\partial M - A$.

This is not hard because the S_j have 1-dimensional spines and it suffices to embed the spines. In detail first choose a tree in each S_j that joins ∂h_j to the intersection

point in each dual pair of cap boundaries. Let S'_j be a small neighborhood in S_j of the union of the tree, ∂h_j , and the cap boundaries. Denote by R_j the closure of the complement; by construction this is a disk containing ∂d_j . Denote by T_j the tree union with an arc in each cap which connects the intersection point in the cap boundaries to the subdisk mapping to the handle core. These start off in M^1 but since they are 1-dimensional we can push them, rel endpoints on ∂h_j and handle cores, to lie in $\partial M - A$. Since this is 3-dimensional we can approximate the map to be disjoint embeddings on the T_j . We can extend the embedding $T_j \subset \partial M - A$ to an embedding of a neighborhood of T_j in S_j , and thus an embedding $S'_j \subset \partial M - A$.

The next step is to do handle moves. Choose one in each dual pair of caps, and use it to do surgery on S'_j . Surgery consists of cutting out a neighborhood of the boundary of the cap and replacing it by two parallel copies. We choose the parallel copies to be on the boundary of the handle whose core is the cap. The net effect is to convert each S_j into a disk embedded in the boundary of $M^1 \cup (2\text{-handles})$. Push the h_j across these disks by isotopy. On the handlebody level this induces handle slides of the h_j over cap 2-handles. Afterwards the attaching maps of the handles are homotopic in the 1-skeleton to the K 2-cell attaching maps, via the disks R_j . Therefore the move changes M to have 2-skeleton equal to K (that is, equal to the complex currently denoted K , and obtained from the original K by 2-deformation).

The final step is to describe these handle moves from the dual point of view, and recognize them as ambient 2-deformations of W . First recall that surgery on S'_j using a cap yields two parallel copies of the cap in the resulting disk. Changing the attaching map of h_j by isotopy across the disk does two handle additions of h_j over the cap handle. The dual cap gives a homotopy between arcs recording the homotopy classes of these additions. This homotopy lies in N so shows the two additions cancel in the chain complex with $Z\pi_1 N$ coefficients. If the homotopy were in M it would show the additions cancel homotopically, and give an ambient 2-deformation of M in N . However these homotopies go over handles in W and are not in M . Some can be deformed into M , but this is unusual.

The dual of a handle move of h over g is a move of g^* over h^* . The dual of the procedure above thus involves moves of 2-handles of (W, V) over duals of handles in (M, A) . These moves still occur in algebraically cancelling pairs, with a homotopy between the addition arcs provided by the dual cap. Since the cap is in W the homotopy also lies in W . Thus viewed from this side the moves are homotopically cancelling handle additions and therefore ambient 2-deformations of W .

This completes the geometric part of the theorem. The ‘‘Refinement’’ asserts that when the geometric output is converted back into algebra it matches up with the algebraic input. This is supposed to be clear since at every step the geometric moves were modeled on the algebra. We omit details since they are routine but long.

REFERENCES

- [AM] S. Akbulut and R. Matveyev, *A convex decomposition theorem for 4-manifolds*, Internat. Math. Res. Notices **7** (1998), 371–381. MR **99f**:57030
- [AC] J. J. Andrews and M. L. Curtis, *Free groups and handlebodies*, Proc. Amer. Math. Soc. **16** (1965), 192–195. MR **30**:3454

- [DK] S. K. Donaldson and P. B. Kronheimer, *The geometry of four-manifolds*, Oxford Science Publications. Oxford University Press, New York, 1990. MR **92a**:57036
- [FKT] M. H. Freedman, V. S. Krushkal and P. Teichner, *Van Kampen's embedding obstruction is incomplete for 2-complexes in R^4* , Math. Res. Lett. **1** (1994), 167–176. MR **95c**:57003
- [FQ] M. H. Freedman and F. Quinn, *Topology of 4-manifolds*, Princeton University Press, 1990.
- [GS] R. E. Gompf and A. I. Stipsicz, *4-manifolds and Kirby calculus.*, Graduate Studies in Mathematics, vol. 20, American Mathematical Society, Providence, RI, 1999. MR **2000h**:57038
- [H] G. Huck, *Embeddings of acyclic 2-complexes in S^4 with contractible complement*, Springer Lecture Notes in Math. **1440** (1990), 122–129. MR **92a**:57002
- [K1] V. S. Krushkal, *Embedding obstructions and 4-dimensional thickenings of 2-complexes*, Proc. Amer. Math Soc. **128** (2000), 3683–3691. MR **2001b**:57006
- [KT] V. S. Krushkal and P. Teichner, *Alexander duality, gropes and link homotopy*, Geometry & Topology **1** (1997), 51–69. MR **98i**:57013
- [Q] F. Quinn, *Handlebodies and 2-complexes*, Springer Lecture Notes in Math. **1167** (1985), 245–259. MR **88a**:57008
- [W1] C. T. C. Wall, *Geometrical Connectivity*, J. London Math. Soc **3** (1971), 597–608. MR **44**:7569a
- [W2] ———, *Formal deformations*, Proc. London Math. Soc **16** (1966), 342–352. MR **33**:1851
- [W3] ———, *Finiteness conditions for CW-complexes*, Ann. Math. **81** (1965), 56–69. MR **30**:1515
- [W4] ———, *Finiteness conditions for CW-complexes II*, Proc. Roy. Soc. **295** (1966), 129–139. MR **35**:2283

DEPARTMENT OF MATHEMATICS, VIRGINIA TECH, BLACKSBURG, VIRGINIA 24061-0123
E-mail address: quinn@math.vt.edu