Heegaard splittings of 3-manifolds

Martin Scharlemann
Department of Mathematics, University of California
Santa Barbara, CA 93106 USA

1. Introduction

In 1898 P. Heegaard [He] gave a simple way to build a complicated 3-manifold. Begin with the 3-ball B^3 and in its boundary pick out two disjoint 2-disks D_0 and D_1. Using those disks, attach to B^3 a handle, that is a copy of $D^2 \times I$, by identifying $D^2 \times \{i\}$ with D_i, $i = 0, 1$. Depending on the orientation with which the ends of the handle are attached, the result is either $D^2 \times S^1$ or a solid Klein bottle. One can continue to attach more handles to B^3 in a similar way. The result of attaching g handles to B^3 is called a handlebody of genus g. Topologically there are exactly two handlebodies of genus g, one of them orientable and the other not orientable. Now suppose that H_1 and H_2 are handlebodies of the same genus and orientability. Then ∂H_1 and ∂H_2 are homeomorphic. One can construct a complicated 3-manifold by attaching H_1 to H_2 by a possibly complicated homeomorphism from ∂H_1 to ∂H_2. The resulting closed 3-manifold M can be written $M = H_1 \cup_S H_2$, where S is the surface ∂H_i in M. This structure on M is called a Heegaard splitting of M and S is called a Heegaard splitting surface. Two Heegaard splittings of a closed 3-manifold M are equivalent if their Heegaard splitting surfaces are isotopic in M. They are homeomorphic if there is a homeomorphism of M carries one to the other.

Now natural questions arise: How universal is this construction? Is there a natural extension to 3-manifolds with boundary? This is considered in section 2. In section 3 we shall study four important structures of Heegaard splittings:

1) Stabilization
2) Reducibility
3) ∂-reducibility
4) Weak reducibility

As an application, in section 4 we shall give a proof to a conjecture on tunnel numbers of composite knots.

These notes were taken by Ruifeng Qiu, Jilin University, Changchun, China

2. Some notions and results on 3-manifolds

In this section, we shall introduce some notions and results on 3-manifolds.

2.1. The Heegaard splittings of 3-manifolds.

Definition 2.1. A 3-manifold M is a separable metric space locally homeomorphic to $R^3_+ = \{(x_1, x_2, x_3)|x_3 \geq 0\}$. That is, for any $x \in M$, there is an open neighborhood of x, say U, and an open set of R^3_+, say V, so that U is homeomorphic to V. If the homeomorphism carries x to $R^2 = \{(x_1, x_2, x_3)|x_3 = 0\}$, then $x \in \partial M$. Thus ∂M is well defined. ∂M is called the boundary of M.

A foundational theorem of Moise[Mo] (see also [Bi]) says that all 3-manifolds can be triangulated. That is the following theorem:

Theorem 2.2[Mo]. 1) Any compact 3-manifold is homeomorphic to a finite simplicial complex.

2) If M is homeomorphic to two simplicial complexes K and L, then the homeomorphism from K to L is isotopic to a homeomorphism which is piecewise linear.

Exercise 2.3. Suppose that K is a simplicial complex.

1) When is K a compact 3-manifold?

2) When is K a compact m-manifold?

3) When is K a manifold with $\partial M \neq \phi$?

Suppose that $(M, \partial M)$ and $(N, \partial N)$ are two compact manifolds. An inclusion $(N, \partial N) \subset (M, \partial M)$ is proper if $\partial M \cap N = \partial N$.

PL topology implies the following:

1) Any point in the interior of a 3-manifold M has a neighborhood homeomorphic to a 3-ball.

2) Any properly embedded arc α in $(M, \partial M)$ has a neighborhood homeomorphic to $\alpha \times D^2$, where D^2 is a disk.

3) If M is orientable and c is a circle in the interior of a 3-manifold, then c has a neighborhood homeomorphic to $c \times D^2$.

4) If M is orientable and S is a properly embedded orientable surface in M, then S has a neighborhood homeomorphic to $S \times I$.

Attaching a 1-handle to a 3-manifold M means that $M \cup_h (I \times D^2)$ where D^2 is a disk and $h : (\partial I) \times D^2 \longrightarrow \partial M$ is an embedding map. Attaching a 2-handle to a 3-manifold M means that $M \cup_h (D^2 \times I)$ where $h : (\partial D^2) \times I \longrightarrow \partial M$ is an embedding map. If h is orientation preserving, then the result is orientable. A handlebody of genus g is the 3-manifold obtained from a 3-ball B^3 by attaching g 1-handles.

Examples. 1) If Γ is a connected finite graph properly embedded in a 3-manifold M with v vertices and e edges, then $\eta(\Gamma)$ is a handlebody of genus $(e - v + 1)$, where $\eta(\Gamma)$ is a neighborhood of Γ.

2) If M is a closed 3-manifold and K is a triangulation of M, then $\eta(K^1)$ and $M - int\eta(K^1)$ are handlebodies, where K^1 is the 1-skeleton of K.

Proof. 1) Let Γ' be a maximal tree of Γ. Since Γ contains v vertices, Γ' contains $v-1$ edges. Since Γ contains e edges, $\eta(\Gamma)$ is obtained by attaching $e-v+1$ 1-handles to $\eta(\Gamma')$. Since $\eta(\Gamma')$ is a 3-ball, $\eta(\Gamma)$ is a handlebody of genus $e - v + 1$.

2) By 1), $\eta(K^1)$ is a handlebody.

Now let Γ be the dual 1-skeleton of K which is defined as follows. The vertices of Γ are the barycenters of the 2- and 3-simplices of K and edges connect the barycenter of a 3-simplex to the barycenter of each of its faces. By 1), $\eta(\Gamma)$ is a handlebody. It is easy to see that $M - int\eta(K^1)$ is homeomorphic to $\eta(\Gamma)$. Thus $M - int\eta(K^1)$ is also a handlebody.

□

Corollary 2.4. Let M be a closed 3-manifold. Then $M = H_1 \cup_h H_2$, where H_1 and H_2 are handlebodies and $h : \partial H_1 \longrightarrow \partial H_2$ is a homeomorphism.

Definition 2.5. A Heegaard splitting of a closed 3-manifold M, denoted by $M = H_1 \cup_S H_2$, is a surface S in M which divides M into two handlebodies H_1 and H_2.

Two Heegaard splittings of a closed 3-manifold are equivalent if their Heegaard surfaces are isotopic. By Corollary 2.4, any closed 3-manifold has a Heegaard splitting.

Exercise 2.6. Construct Heegaard splittings for the following 3-manifolds:
1) $M = T^3 = S^1 \times S^1 \times S^1$.
2) $M = S^1 \times S^2$.
3) $M = S^1 \times S_g$ where S_g is a closed surface.
4) $M = RP^3$.

Exercise 2.7. Classify all the closed 3-manifolds with genus one Heegaard splittings.

Definition 2.8. Let F be a closed surface not homeomorphic to a 2-sphere. A compression body H is the manifold obtained from $F \times I$ by attaching 1-handles to $F \times \{1\}$. We denote by $\partial_- H = F \times \{0\}$, $\partial_+ H = \partial H - F \times \{0\}$. If no 1-handles are attached, $F \times I$ is called a trivial compression body.

Remarks

1) Let S be a closed surface, and let M be the manifold obtained from $S \times I$ by attaching 2-handles to $S \times \{1\}$ and capping off 2-spheres with 3-balls. Then M is either a handlebody or a compression body. In the following argument, a handlebody is viewed as a compression body with $\partial_- H = \phi$.

2) The cores of the handles are called meridian disks and a collection of meridian disks is said to be complete if each of its complementary components is either a 3-ball or $\partial_- H \times I$.

The construction of Heegaard splittings for closed 3-manifolds in the previous examples suggests several possible ways of extending the definition of Heegaard splitting to cover the case in which the 3-manifold has boundary. The most useful is the following: Write ∂M as the disjoint union of two sets of components, $\partial_1 M$ and $\partial_2 M$. Choose a triangulation K of M so that no simplex of K is incident to more than one boundary component. Let K' be its barycentric subdivision. Delete the interior of all simplices of K' incident to $\partial_2 M$. The resulting 3-manifold M' is homeomorphic to M, since only a collar of $\partial_2 M$ has been removed; let $\partial_2' M$ denote $\partial_2 M$ in this new triangulation. Then $\partial_2' M$ contains the subcomplex of the dual subcomplex Γ determined by simplices incident to $\partial_2 M$. Let $\Gamma_1 \subset M'$ be the union of $\partial_1 M$ and all vertices and edges not incident to $\partial_2 M$. Let Γ_2 be the union of $\partial_2' M$ and all vertices and edges of the dual 1-complex $\Gamma \cap M'$ not incident to Γ_1. Again it is easy to check that M is the union of regular neighborhoods of the complexes Γ_1 and Γ_2 along their homeomorphic boundary, which is still a closed connected surface. It is easy to see that a regular neighborhood of Γ_i in M is a compression body, $i = 1, 2$.

This construction suggests the following way of defining a Heegaard splitting on a 3-manifold with boundary.

Definition 2.9. Suppose that M is a 3-manifold with $\partial M = \partial_1 M \cup \partial_2 M$. A Heegaard splitting of $(M, \partial_1 M, \partial_2 M)$ is a surface S dividing each component of M into two compression bodies H_1 and H_2 with $\partial_- H_1 = \partial_1 M$ and $\partial_- H_2 = \partial_2 M$.

Remarks.

1) By the above argument, any compact 3-manifold with any division $\partial M = \partial_1 M \cup \partial_2 M$ has a Heegaard splitting.

2) A Heegaard splitting $H_1 \cup_S H_2$ is said to be trivial if one of H_1 or H_2 is a trivial compression body, i.e. is homeomorphic to $S \times I$

3) Suppose $H_1 \cup_S H_2$ is a Heegaard splitting of a 3-manifold $(M, \partial_1 M, \partial_2 M)$. Then H_1 is obtained from $\partial_1 M \times I$ by attaching 1-handles and H_2 is obtained from $S = \partial_+ H_1$ by attaching 2- and 3-handles. From this point of view a Heegaard splitting is just a standard handle decomposition of M viewed as a cobordism between $\partial_1 M$ and $\partial_2 M$.

2.2. Surfaces in 3-manifolds. Henceforth we will, for simplicity, consider only orientable 3-manifolds and surfaces. Suppose that M is a 3-manifold, P is a p-manifold properly embedded in M and Q is a q-manifold properly embedded in M. In the following argument, we shall assume that $P \cap Q$ is a $(p+q-3)$-manifold properly embedded in M. General position means that

1) two 1-manifolds in a 3-manifold are disjoint,

2) the intersection of a 1-manifold and a 2-manifold in a 3-manifold is a finite set of points.

3) the intersection of two surfaces in a 3-manifold is a 1-manifold.

Definition 2.10. Let F be a surface.

1) A simple closed curve in F is said to be inessential if it divides F into two components, one of which is a disk; otherwise it is said to be essential.

2) An arc properly embedded in F is said to be inessential if it, together with some arc on ∂F, bounds a disk; otherwise it is said to be essential.

The following observations are very useful in studying 3-manifolds.

Let Γ be a 1-manifold properly embedded in a surface F.

1) If a component of Γ is an inessential simple closed curve, then some component α of Γ bounds a disk in F which is disjoint from Γ. α is called an innermost circle of Γ.

2) If a component of Γ is an inessential arc, then there is either an innermost circle or an arc β which, together with some arc on ∂F, bounds a disk in F which is disjoint from Γ. β is called an outermost arc of Γ.

Definition 2.11. Let T be a surface properly embedded in a 3-manifold M.

1) T is said to be compressible if either T bounds a 3-ball in M, or there is an essential simple closed curve in T which bounds a disk D in M such that int$D$ is disjoint from T; otherwise it is said to be incompressible.

2) T is said to be ∂-compressible if there is an essential arc in T which, together with some arc in ∂M, bounds a disk D in M such that $intD$ is disjoint from T; otherwise it is said to be ∂-incompressible.

3) T is said to be essential if T is incompressible and ∂-incompressible.

Theorem 2.12 (Loop theorem). Suppose that M is a 3-manifold, T is a surface properly embedded in M and i is the inclusion map. If $i_* : \pi_1(T) \longrightarrow \pi_1(M)$ is not injective, then T is compressible.

Remark. In fact a surface T not homeomorphic to a 2-sphere is incompressible in a 3-manifold M if and only if $i_* : \pi_1(T) \longrightarrow \pi_1(M)$ is injective. (The situation would be a bit more complicated without our ongoing assumption that T and M are orientable.)

Corollary 2.13. Any properly embedded surface T in S^3 is compressible.

Proof If T is not homeomorphic to a 2-sphere S^2, then $i_* : \pi_1(T) \longrightarrow \pi_1(S^3)$ is not injective. By Theorem 2.12, T is compressible. If T is homeomorphic to S^2, then T bounds a 3-ball in M. Thus T is compressible.

□

Exercise 2.14. 1) Suppose H is a compression body. Is $\partial_- H$ incompressible in H?

2) Let T be a surface properly embedded in a handlebody. Show that either
a) T is a meridian disk or
b) T is compressible or
c) T is ∂-compressible.

3) Suppose T is a surface properly embedded in a compression body H. Show that either
a) T is a meridian disk or
b) T is compressible or
c) T is ∂-compressible or
d) T is a spanning annulus or
e) T is parallel to some components of $\partial_- H$.

Let T be a surface properly embedded in a 3-manifold M. If there is a disk D in M so that $D \cap T = \partial D$. Then there is a 3-ball $B^3 = D \times I$ such that $B^3 \cap T = \partial D \times I$. Now let $T' = (T - \partial D \times I) \cup D \times \{0, 1\}$. The process to obtain T' from T is called doing a 2-surgery on T along D. It is easy to see that $\chi(T') = \chi(T) + 2$.

Definition 2.15. Let M be a 3-manifold. M is said to be reducible if M contains an incompressible 2-sphere; otherwise it is said to be irreducible.

Let M be a reducible 3-manifold.Suppose P in M is a reducing 2-sphere and there is a disk D in M such that $D \cap P = \partial D$. Now let P' be the surface obtained by doing a 2-surgery on P along D. Then P' contains two components and one of them is also a reducing 2-sphere of M.

Suppose that M_1 and M_2 are two 3-manifolds, and S_i is a 2-sphere bounding a 3-ball B_i^3 in M_i, $i = 1, 2$. The manifold $(M_1 - intB_1^3) \cup (M_2 - intB_2^3)$, denoted by $M_1 \# M_2$, is called a connected sum of M_1 and M_2. In fact, any 3-manifold is a connected sum of irreducible 3-manifolds.

Theorem 2.16([Kn], [Mil]). Let M be a compact orientable 3-manifold. Then $M = M_1 \# \dots M_n \# (S^1 \times S^2) \# \dots \# (S^1 \times S^2)$, where each M_i is irreducible. Furthermore, this decomposition is unique.

Proposition 2.17. Suppose that M is irreducible, S and T are incompressible surfaces in M. Then S and T can be isotoped so that each of circle components of $S \cap T$ is essential in both surfaces.

Proof. Suppose that one circle component of $S \cap T$ is inessential in S. Then there is a component of $S \cap T$, say α, which bounds a disk D in S such that int$D$ is disjoint from T. If α is essential in T, then T is compressible, a contradiction. So α bounds a disk D' in T. Thus $D \cup D'$ is a 2-sphere. Since M is irreducible, $D \cup D'$ bounds a 3-ball in M. Now let $T' = (T - D') \cup D$. Then T' is isotopic to T But $|S \cap T'| < |S \cap T|$ (by pushing T' slightly). By induction on $|S \cap T|$, the result follows.

□

Definition 2.18. Let M be a 3-manifold. M is said to be ∂-reducible if there is an essential simple closed curve in ∂M which bounds a disk in M; otherwise it is said to be ∂-irreducible.

Theorem 2.19. Let M be an irreducible, ∂-irreducible 3-manifold. If S and T are incompressible, ∂-incompressible surfaces in M, then S and T can be isotoped so that each component of $S \cap T$ is essential in both surfaces.

Proof. By Proposition 2.17, we may assume that each circle of $S \cap T$ is essential in both S and T. Now suppose that one arc α of $S \cap T$ is inessential in one of S and T, say S. Then α, together with some arc β_1 in ∂S, bounds a disk D in S. D contains no circle of $S \cap T$, because such a circle is inessential in S. Without loss of generality, we assume that α is outermost in S. Then int$D$ is disjoint from T. Since T is ∂-incompressible, α, together with some arc β_2 in ∂T, bounds a disk D' in T. Thus $\beta_1 \cup \beta_2$ bounds a properly embedded disk $D \cup D'$ in M. Since M is ∂-irreducible, $\beta_1 \cup \beta_2$ bounds a disk D'' in ∂M. Thus $D \cup D' \cup D''$ is a 2-sphere which bounds a 3-ball. Let $T' = (T - D') \cup D$. Then T' is isotopic to T, but $|S \cap T'| < |S \cap T|$ (by pushing T' slightly). By induction $|S \cap T|$, the result follows.

□

3. Structures of Heegaard splittings

In this section, we shall study structures of Heegaard splittings.

3.1. Stabilization, reducibility and ∂-reducibility.

Definition 3.1. A Heegaard splitting $H_1 \cup_S H_2$ of a 3-manifold M is said to be stabilized if there are properly embedded disks $D_i \subset H_i$ so that $|\partial D_1 \cap \partial D_2| = 1$.

Definition 3.2. A Heegaard splitting $H_1 \cup_S H_2$ of a 3-manifold M is said to be reducible if there is a 2-sphere P such that $|P \cap S| = 1$ and $P \cap S$ is essential in S.

Remark.

There is an equivalent definition of a reducible Heegaard splitting $H_1 \cup_S H_2$. A Heegaard splitting $H_1 \cup_S H_2$ is reducible if and only if there are two properly embedded disks $D_i \subset H_i$ so that $\partial D_1 = \partial D_2$, and ∂D_i is essential in S.

There is a connection between stabilization and reducibility, given by the following proposition:

Proposition 3.3. Suppose a Heegaard splitting $H_1 \cup_S H_2$ is stabilized. Then either it is reducible or it is the standard genus one splitting of S^3.

Proof. Let $D_i \subset H_i$ be properly embedded disks so that $\partial D_1 \cap \partial D_2$ is a single point. Let B be the union of a bicollar of D_1 in H_1 and a bicollar of D_2 in H_2 along the square in which they intersect. B is a 3-ball whose boundary sphere P can be moved slightly so that P intersects each H_i in a single hemisphere and so that the curve $c = P \cap S$ cuts off from S a punctured torus. Unless this curve is inessential in S the boundary of the 3-ball is a reducing sphere. If the curve is inessential, then S is a torus dividing M into two solid tori, whose meridians intersect in a single point. This is the genus one Heegaard splitting of S^3.

□

In the following argument, a new notion is necessary.

Definition 3.4. Let H be a compression body. A spine Σ of H is a finite graph in H so that $\sum \cap \partial H = \sum \cap \partial_- H$ consists only of valence one vertices and H deformation retracts to $\sum \cup \partial_- H$.

By Definition 3.4, The spine of a compression body is not uniquely defined and a spine of a handlebody H is a finite graph in H to which H deformation retracts.

Suppose that M has a Heegaard splitting $H_1 \cup_S H_2$ and M' has a Heegaard splitting $H_1' \cup_{S'} H_2'$. Then the 3-manifold $M \# M'$ has a natural Heegaard splitting as follows: Let D be a disk in S and D' be a disk in S'. Let $H_1'' = H_1 \cup_{D=D'} H_1'$ and $H_2'' = H_2 \cup_{D=D'} H_2'$. It is easy to see that H_1'' and H_2'' are compression bodies, and $H_1'' \cup_{\partial_+ H_1'' = \partial_+ H_2''} H_2''$ is a reducible Heegaard splitting of the manifold $M \# M'$. Now natural question arises: Is any Heegaard splitting of a reducible 3-manifold reducible? One of the first major theorems on Heegaard splittings, due to Haken, is that any Heegaard splitting of a reducible manifold is reducible. The theorem is important not just for what it says, but for the type of argument which is used.

Theorem 3.5([Ha]). Suppose M is a reducible manifold with a Heegaard splitting $H_1 \cup_S H_2$. Then there is a reducing sphere P for M so that $P \cap S$ is a single circle

Proof. There are two ways to prove Haken's theorem. Similar to Haken's original proof is that given in [Ja1]. One can assume that P intersects one of the compression bodies only in disks. The idea will be to minimize the number of circles of intersection, under the assumption that P intersects one of the compression bodies only in disks. If P intersects H_2 only in disks, consider the planar surface $P_1 = P \cap H_1$. Compress and ∂-compress P_1 as much as possible. Compressions of P_1 will convert P into two spheres, at least one of which, say P_2, is a reducing sphere. We shall restrict attention to P_2. At the end of this process P_1 will be converted to a surface P' which is disjoint from a complete collection of meridian disks for H_1 (otherwise curves of intersection can be used to compress or ∂-compress) and, for any essential curve α in $\partial_- H_1$, disjoint from a spanning annulus $\alpha \times I$. It follows that $P' \cap H_1$ is a collection of disks. What is not obvious, but can be explicitly calculated, is that the number of disks in $P' \cap H_1$ is lower that the original $P \cap H_2$. The process is continued, switching the roles of H_1 and H_2 until there is only one intersection curve.

Another approach is given in [ST2]. The following is a sketch of the proof.

Let $\sum$ be a spine of H_2. Put it transverse to P. Let Δ be a complete collection of compressing disks for H_1 viewed as a ∂-singular collection of disks in the complement of $\sum$. Put Δ transverse to P. Circles of intersection can be removed, just as in the previous argument, so that $(\sum \cup \Delta) \cap P$ becomes a graph $\Gamma \subset P$ with vertices $\sum \cap P$ and edges $\Delta \cap P$. Trivial loops of Γ can be eliminated at the cost of merely changing Δ, and a vertex incident to some edges but no loops can be used to slide edges of $\sum$ in a way that lowers $\sum \cap P$. (This is the hard part to see.) The upshot is that, eventually, there is guaranteed to be an isolated vertex. This picks out a meridian μ of H_1 which is disjoint from a complete collection of meridian disks for H_2. If H_2 is a handlebody this implies that $\partial\mu$ bounds a meridian in H_2 and so $H_1 \cup_S H_2$ is reducible. If H_2 is merely a compression body, we can only conclude that there is a ∂-reducing disk for M which intersects S in a single curve. But we can surger M along this disk to get a new reducible 3-manifold and continue the process until an appropriate sphere is found.

□

Corollary 3.6. [Fr] Suppose that $M = H_1 \cup_S H_2$ is a Heegaard splitting $M \neq S^3$, and a spine $\sum$ for H_1 has property that some circuit in $\sum$ lies in a 3-ball. Then S is reducible.

Proof. Let σ be a circuit of $\sum$ which is contained in a 3-ball B^3 and $\eta(\sigma)$ is a regular neighborhood in B^3. Then $(H_1 - int\eta(\sigma)) \cup_S H_2$ is a Heegaard splitting of $M' = M - int\eta(\sigma)$. Since M' is reducible, by Theorem 3.5, $(H_1 - int\eta(\sigma)) \cup_S H_2$ is a reducible splitting. Hence $H_1 \cup_S H_2$ is also a reducible splitting.

□

Now we give a new notion:

Definition 3.7. A Heegaard splitting $M = H_1 \cup_S H_2$ is ∂-reducible if there is a ∂-reducing disk for M which intersects S in a single curve.

It suggests the following analogue to Theorem 3.5.

Proposition 3.8. Any Heegaard splitting of a ∂-reducible 3-manifold is ∂-reducible.

Proof. The proof of this theorem is similar to the one of Theorem 3.5.

□

3.2. Weakly reducibility. In 1987 Casson and Gordon discovered a new structure on Heegaard splittings which is perhaps less natural than those described above but which has turned out to be quite useful.

Definition 3.9. A Heegaard splitting $H_1 \cup_S H_2$ is said to be weakly reducible if there are essential disks $D_i \subset H_i$ so that ∂D_1 and ∂D_2 are disjoint in S; otherwise it is said to be strongly irreducible.

Proposition 3.10. 1) A reducible Heegaard splitting $H_1 \cup_S H_2$ is weakly reducible.

2) A nontrivial ∂-reducible Heegaard splitting $H_1 \cup_S H_2$ is weakly reducible.

Proof. 1) Let P be a sphere which intersects H_i in a disk D_i, $i = 1, 2$. Let $\partial D_i \times I$ be a regular neighborhood of ∂D_i in S. Then $\partial D_i \times \{0\}$ bounds a disk in H_1 and $\partial D_i \times \{1\}$ bounds a disk in H_2. Since $\partial D_i \times \{0\} \cap \partial D_i \times \{1\} = \phi$, $H_1 \cup_S H_2$ is weakly reducible.

2) Let D be a ∂-reducing disk for $M = H_1 \cup_S H_2$ which intersects S in a single curve α. Without loss of generality, we suppose that α bounds a disk in H_1. Since $H_1 \cup_S H_2$ is nontrivial, there is a meridian disk in H_2. Such a meridian disk D' can be found so that D' is disjoint from the spanning annulus $D \cap H_2$. In particular then $\partial D' \cap \alpha = \phi$. Thus $H_1 \cup_S H_2$ is weakly reducible.

□

Here are some applications of this structure.

Theorem 3.11([CG]). If $M = H_1 \cup_S H_2$ is a weakly reducible Heegaard splitting then either $H_1 \cup_S H_2$ is reducible or M contains an incompressible surface.

Proof. Since the splitting is weakly reducible, S can be compressed simultaneously in both directions, that is, both into H_1 and simultaneously into H_2. Let $\Delta_1 \subset H_1$ and $\Delta_2 \subset H_2$ be collections of disjoint meridians in the respective compression bodies so that $\partial\Delta_1$ and $\partial\Delta_2$ are disjoint in S and the families Δ_i are maximal with respect to this property. (Since $H_1 \cup_S H_2$ is weakly reducible, $\Delta_i \neq \phi$, $i = 1, 2$.) That is, if S_i represents the surface in H_i obtained by compressing S along Δ_i, then any further compressing disks of S_i into H_i will necessarily have boundaries intersecting the boundaries of the other disk family.

Let $\bar{S}$ be the surface obtained by compressing S_1 along Δ_2, or, dually, S_2 along Δ_1. Then S separate M into the manifolds W_1 and W_2 so that H_1, say, can be recovered from W_1 by removing some neighborhoods of arcs from W_1 (arcs dual to Δ_2) and attaching some 1-handles in W_2. A helpful and vivid picture is to imagine H_1 red and H_2 blue. The compressions of S to $\bar{S}$ along the Δ_i cover $\bar{S}$ with both red and blue spots, two red spots for each disk in Δ_1 and two blue spots for each disk in Δ_2. S is recovered from $\bar{S}$ by attaching red tubes in W_2 with ends on red spots and blue tubes in W_1 with ends on blue spots

The surface $\bar{S}$ is incompressible in M. To see this, suppose that $\bar{S}$ compresses into W_1, say. After pushing $\bar{S}$ slightly into W_2, we can view S_1 as a Heegaard splitting surface of W_1, that is $W_1 = (W_1 \cap H_1) \cup_{S_1} (W_1 \cap H_2)$ and each of these pieces is a compression body. The compression of $\bar{S}$ is a ∂-reduction of W_1. By Proposition 3.8, there is a ∂-reducing disk D that intersects S_1 in a single curve. We can take ∂D to be disjoint from Δ_1 (i. e. the red spots) and, after 2-handle slides among the Δ_2, we can make Δ_2 disjoint from the annulus $D - (W_1 \cap H_1)$. But then $D \cap H_1$ makes S_1, hence also S, compressible in H_1 via a disk disjoint from Δ_2, contradicting the maximality of Δ_1.

Unless $\bar{S}$ is a collection of spheres, we are through. Suppose that $\bar{S}$ is a collection of spheres. Note that at least one, $\bar{S}_0$, has both a red spot and a blue spot. For otherwise, when S is recovered from $\bar{S}$ by attaching red and blue tubes, S would consist of two components: one containing all red tubes and one containing all blue tubes. Choose in $\bar{S}_0$ a simple closed curve that separates in the sphere $\bar{S}_0$ the red spots from the blue spots. Push the interior of the disk in $\bar{S}_0$ that contains the red spots (resp. blue spots) completely into H_1 (resp. H_2). Then $\bar{S}_0$ is the union of a red disk and a blue disk along a curve, i. e. it is a reducing sphere for the original Heegaard splitting.

□

Note that at the end of the proof of Theorem 3.11 we have $\bar{S}$ dividing M into two 3-manifolds W_1 and W_2, each of which inherits a Heegaard splitting surface (a component of S_i) of lower genus than S. This splitting itself may be weakly reducible and we can continue the process. Ultimately an irreducible Heegaard splitting $M = H_1 \cup_S H_2$ can be broken into a series of strongly irreducible splittings (see [ST3]). That is, we can begin with the handle structure determined by $H_1 \cup_S H_2$ and rearrange the order of the 1- and 2-handles, so that

$$M = M_0 \cup_{\bar{S}_1} M_1 \cup_{\bar{S}_2} \cup \ldots \cup_{\bar{S}_m} M_m$$

The 1- and 2-handles which occur in M_i provide it with a strongly irreducible splitting $A_i \cup_{P_i} B_i$ with $\partial_- A_i = \bar{S}_i$, $\partial_- B_{i-1} = \bar{S}_i$ for $1 \leq i \leq m$, $\partial_- A_0 = \partial_- H_1$, $\partial_- B_m = \partial_- H_2$. Each component of each $\bar{S}_i$ is a closed incompressible surface of positive genus. None of the compression bodies $A_i, B_{i-1}, 1 \leq i \leq m$ is trivial, though components of each may be. If $\partial_- A_0$ or $\partial_- B_m$ is compressible in M then A_0 or B_m is trivial. Such a rearrangement of handles will be called an untelescoping of the Heegaard splitting.

In 1968 Waldhausen[Wa] showed that any positive genus Heegaard splitting of S^3 is stabilized. That implies that any positive genus Heegaard splitting of S^3 is obtained by stabilizing the unique zero splitting into 3-balls. So a Heegaard splitting of S^3 is completely determined by its genus. This is the first uniqueness result. Here is a sketch of the proof of Waldhausen's theorem given in [ST2].

Theorem 3.12([Wa]). Every positive genus Heegaard splitting of S^3 is stabilized.

Proof. Suppose that $S^3 = H_1 \cup_S H_2$ and $\sum$ is a spine of H_1. We may assume that $\sum$ is a trivalent graph in S^3 and we are allowed to do edge-slides. Choose a Morse function $h : S^3 \longrightarrow [-1, 1]$ which has a single minimum(at height -1) and a single maximum(at height 1) and which restricts to a Morse function on $\sum$. Put $\sum$ in thin position with respect to this height function. In outline, this means that you cannot push down a maximum so that it moves below a minimum without introducing new critical points.

It suffices to show that there is an unknotted cycle $\gamma \subset \sum$. For then S would also be a Heegaard splitting surface for the solid torus $S^3 - \eta(\gamma)$. By Proposition 3.8, this splitting would necessarily be boundary reducible. That means that the original splitting S is stabilized.

Consider a complete collection Δ of meridian disks of H_2, extended into H_1, so that its interior is embedded in $S^3 - \sum$ and its (singular) boundary lies in $\sum$. The first observation is that we may as well assume $\partial\Delta$ runs across every edge of $\sum$, for otherwise $H_1 \cup_S H_2$ would be reducible(see Theorem 3.11). If the splitting were reducible then a reducing 2-sphere splits S into two Heegaard splittings of S^3 each of smaller positive genus, and we would be done by induction.

Consider when a level sphere $S_t = h^{-1}(t)$ cuts off from Δ a subdisk sufficiently simple that it can be used to slide part of an edge of $\sum$ so that it lies on S_t. It is easy to see that this is true just below the highest point of $\sum$ and just above the lowest point. In the former case the disk can be used to lower the maximum slightly and in the latter to raise the minimum. Suppose we simultaneously have two subdisks of Δ, one of which lowers a maximum and the other of which raises a minimum. Then either this violates thin position or the two edges which we have pushed onto the level sphere have the same ends, i. e. they create an unknotted cycle and we are done.

We know then that a sufficiently high sphere cuts off a subdisk of Δ lowering a maximum, a sufficiently low sphere cuts off a subdisk raising a minimum and if subdisks of both types are cut off simultaneously, then we are done. So it suffices to eliminate the possibility that neither type occurs, that is, there is a height t_0 so that no subdisk cut off by S_{t_0} from Δ can be used either to raise a minimum or lower a maximum. But this situation cannot in fact occur, by an argument-reminiscent of the proof of Theorem 3.5, with S_{t_0} playing the role of reducing sphere.

□

Armed with Theorem 3.12 we can prove a sort of converse to Proposition 3.3.

Theorem 3.13. Suppose M is an irreducible 3-manifold and $H_1 \cup_S H_2$ is a reducible Heegaard splitting of M. Then $H_1 \cup_S H_2$ is stabilized.

Proof. Let P be a sphere which intersects S in a single essential circle. Since M is irreducible, P bounds a 3-ball in M, so the manifold obtained by reducing M along P is the disjoint union of S^3 and a homomorphism of M. The induced Heegaard splitting of the former is stabilized. Its stabilizing disks, when viewed back in $H_1 \cup_S H_2$ show that S was also stabilized.

□

In 1982 W. Jaco gave a famous theorem on handle additions which plays an important role in studying 3-manifolds.

Theorem 3.14([Ja2]). Suppose M is a compact, orientable, ∂-reducible 3-manifold and M' is obtained by attaching a 2-handle to M along a simple closed curve c in ∂M. If $\partial M - c$ is incompressible in M, then M' is ∂-irreducible.

Proof. We denote by T the component of ∂M on which c lies. So $\partial M - T$ is incompressible in M. Via the untelescoping argument following 3.11, M can be separated by incompressible surfaces into n 3-manifolds $M_1, \dots, M_n, n \geq 1$, such that

1) M_1 is a compression body with $\partial_+ M_1 = T$, and

2) M_i is ∂-irreducible, $2 \leq i \leq n$.

We denote by M_1' the manifold obtained by attaching a 2-handle to M_1 along c. Since $\partial M_i (2 \leq i \leq n)$ is incompressible in M, M' is ∂-reducible if and only if M_1' is ∂-reducible.

Note that there is a natural Heegaard splitting $A \cup_T B$ of M_1' such that

1) $A = M_1$ with $\partial_+ A = T$, and

2) B is obtained from $T \times I$ by attaching a 2-handle along $c \times \{0\}$ on $T \times \{0\}$ with $\partial_+ B = T \times \{1\}$.

Now suppose, otherwise, that M_1' is ∂-reducible. By Propositions 3.8 and 3.10, $A \cup_S B$ is weakly reducible. Thus there are two disks $D_1 \subset A$ and $D_2 \subset B$ such that ∂D_1 and ∂D_2 are disjoint in T. There are two possibilities on ∂D_2:

1) ∂D_2 is isotopic to c, and

2) ∂D_2 is coplanar to c, that is, ∂D_2 is separating in $T \times \{1\}$ and one component of $T \times \{1\} - c \cup \partial D_2$ is a planar surface.

For each case, D_1 can be isotoped so that ∂D_1 is disjoint from c. Thus $\partial M - c$ is compressible in M, a contradiction.

□

Proposition 3.15. Suppose $A \cup_S B$ is a strongly irreducible Heegaard splitting of a 3-manifold and D is a disk in M such that D is transverse to S with $\partial D \subset S$. Then D can be isotoped rel ∂D so that $intD \cap S = \phi$.

Proof. This proof is by induction on $|S \cap intD|$. If the interior of D is disjoint from S there is nothing to prove. If $S - D$ has any disk components D' then, by replacing the subdisk of D bounded by $\partial D'$ by a parallel copy of D' we can decrease $|S \cap intD|$. So we may assume that each simple closed curve in $S \cap D$ is essential in S.

A disk component of $D - S$ compresses S in one of the two compression bodies, say A. Then by strong irreducibility of S, all the disk components of $D - S$ lie in A. If any pair of curves of $D \cap S$ are nested in D then the outer curve of the innermost such pair cuts off a component P of $D - S$ so that all but one of the curves in ∂P are adjacent to disks in A and precisely one, denoted by α, is not. Compress S into A along 2-handles whose cores are the disks with boundaries on ∂P. Let M_- be the manifold obtained from B by attaching these 2-handles to B Then $\alpha \subset \partial M_-$ is inessential in M_- so, by strong irreducibility and Proposition 3.8, α is inessential in ∂M_-. Push the disk bounded by α in ∂M_- slightly into H_1 and obverse that this is then a disk D'' in H_1 whose boundary is parallel to α in the component of D adjacent to P across α. Replacing the subdisk of D bounded by α with D'' lowers $|S \cap intD|$.

□

Proposition 3.16. Let $A \cup_S B$ be a strongly irreducible Heegaard splitting of a 3-manifold M, and let T be a surface in M and P be a closed incompressible surface in M. Then

1) P can be isotoped so that no circle of $P \cap T$ is inessential in T.

2) S can be isotoped so that $S \cap T$ can not contain two circles which are inessential and nested in T.

Furthermore, if T is incompressible in M, then

3) P can be isotoped so that each circle of $P \cap T$ is essential in both P and T.

4) S can be isotoped so that each circle of $S \cap T$ is essential in both S and T.

Proof. 1) Suppose that P' is a surface isotopic to P such that $|P' \cap T|$ is minimal among all the surfaces isotopic to P. Suppose, otherwise, that one circle c of $P' \cap T$ is inessential in T. Without loss of generality, we assume that c bounds a disk in T such that $intD$ is disjoint from P'. Since P' is incompressible, c also bounds a disk D' in P'. Let $P'' = (P' - D') \cup D$. Since $A \cup_S B$ is a strongly irreducible Heegaard splitting of M, M is irreducible. Hence $D \cup D'$ bounds a 3-ball in M. That means that P'' is isotopic to P', but $|P'' \cap T| < |P' \cap T|$ (by pushing P'' slightly), a contradiction.

2) This proof is similar, using Lemma 3.15.

3) This is similar to the proof of 1).

4) Let $\sum_A$ be a spine of A and $\sum_B$ be a spine of B. Then $M - (\sum_A \cup \sum_B)$ is homeomorphic to $S \times (0,1)$. Without loss of generality, we assume that $\sum_A \cup \sum_B$ are in general position with T. If one of $\sum_A$ and $\sum_B$, say $\sum_A$, is disjoint from T, then we are done by taking $S = \eta(\Sigma_A)$. If neither is disjoint, denote by S_t the surface $S \times \{t\}$ and observe that for small ϵ, $T - S_\epsilon$ contains a meridian disk for A and $T - S_{1-\epsilon}$ contains a meridian disk for B. For no t can it be true that $T - S_t$ contains meridian disks for both A and B, since the splitting is strongly irreducible.

It follows (with a little bit of argument about non-generic points) that there is a generic value for t so that $T - S_t$ contains no meridian disks for either A or B. In particular, any component of $T \cap S_t$ that is inessential in T is also inessential in S_t and so can be eliminated by an isotopy of S_t. Since T is incompressible, any component of $T \cap S_t$ that is inessential in S_t would also have been inessential in T. So after these isotopies, T and S_t intersect only in curves that are essential in both surfaces.

□

4. An application to knot theory

In this section, we shall introduce an application of structures on Heegaard splittings. The following notions are necessary.

Definition 4.1. Let K be a knot in S^3. A system of arcs attached to K are tunnels on K. If their complement is a handlebody then it is called an unknotting system of tunnels. The number, denoted by $t(K)$, of tunnels contained in a minimal collection of unknotting tunnels is called the tunnel number of K.

Definition 4.2. Suppose that K_1 and K_2 are two knots which lie in the distinct sides of a 2-sphere S^2 in S^3. Let $b : I \times I \longrightarrow S^3$ be an embedding map such that $b(I \times I) \cap K_1 = b(I \times \{0\})$, $b(I \times I) \cap K_2 = b(I \times \{1\})$ and $b(I \times I) \cap S_2 = b(I \times \{1/2\})$. Then the knot $(K_1 - b(I \times \{0\}) \cup b(\partial I \times I) \cup (K_2 - b(I \times \{1\})$, denoted by $K_1 \# K_2$, is called a composite knot.

Remark.
Let $K = K_1 \# K_2$, and A_i be an annulus on $\partial(S^3 - int\eta(K_i)$ whose core bounds a disk in $\eta(K_i)$, $i = 1, 2$. Then $S^3 - int\eta(K)$ is homeomorphic to the manifold $(S^3 - int\eta(K_1)) \cup_{A_1 = A_2} (S^3 - int\eta(K_2))$, and A_i is essential in $S^3 - int\eta(K)$.

Theorem 4.3.([SS]) Suppose that $K_1, \dots, K_n$ are n prime knots in S^3, and $K = K_1 \# \dots \# K_n$. Then $t(K) \geq n$.

Proof: Let $\alpha_1, \dots, \alpha_t$ be a minimal system of unknotting tunnels. Then $S^3 - int\eta(K \cup (\cup\alpha_i))$ is a handlebody. Let $S = \partial\eta(K \cup (\cup\alpha_i))$. Then S is a Heegaard splitting of S^3. It is easy to see that S is also a Heegaard splitting of $M_K = S^3 - \eta(K)$. We denote by g the genus of S. Since $t = g - 1$, $-\chi(S) = 2t$. We denote by $J(H)$ the number of 1-handles in a compression body H.

Fact 1. $J(H) = (\chi(\partial_- H) - \chi(\partial_+ H))/2$.

Now suppose that $A \cup_S B$ is a Heegaard splitting of M_K, and

$$(A_1 \cup_{S_1} B_1) \cup_{F_2} \cup \dots \cup_{F_m} (A_m \cup_{S_m} B_m)$$

is an untelescoping of $A \cup_S B$. Then

$$J(A) = \sum_1^m J(A_i) = (\sum_1^m \chi(F_i) - \sum_1^m \chi(S_i))/2 = -\chi(S)/2 = t$$

where $F_1 = \partial_- A_1$.

Since $K = K_1 \# \dots \# K_n$,

$$S^3 - \eta(K) = W_1 \cup_{a_1} W_2 \cup \dots \cup_{a_{n-1}} W_n$$

where $W_i = S^3 - \eta(K_i)$ and a_i is an annulus.

Since an incompressible surface in a compression body is ∂-compressible or a spanning annulus, one can prove the following fact.

Fact 2. If a is a system of incompressible annuli in a compression body H and X is a component of $H - a$, then

1) X is a compression body, and

2) $\chi(\partial_+ H \cap X) \leq \chi(\partial_- H \cap X)$.

Fact 3. Let L be a prime knot in S^3 and a be an meridian annulus in $S^3 - \eta(L)$. Then a is isotopic to $\partial(S^3, \eta(L))$ rel ∂a.

By Proposition 3.15, F_i and S_i can be isotoped so that they intersect a_j only in essential curves. Since F_i and S_i are separating in M, $|F_i \cap a_j|$ and $|S_i \cap a_j|$ are even. Hence $\chi(F_i \cap W_j)$ and $\chi(S_i \cap W_j)$ is even.

Let $x_{ij} = -\chi(F_i \cap W_j)/2$, $y_{ij} = -\chi(S_i \cap W_j)/2$. Then $\sum_j y_{ij} = -\chi(S_i)$, and $\sum_j x_{ij} = -\chi(F_i)$. Since F_i and S_i intersect a_j in essential curves, the following fact follows:

Fact 4. $0 \leq x_{ij} \leq y_{ij}, 0 \leq x_{i-1,j} \leq y_{ij}$.

If $x_{ij} = y_{ij} = 0$ for $1 \leq i \leq m$ then, for each i, F_i and S_i intersect W_j only in annuli. Thus one component of the complement is homeomorphic to W_j. That means that W_j is contained in a compression body. Hence K_j is trivial, a contradiction. Thus the following fact follows:

Fact 5. For each j, there exists i such that $y_{ij} \neq 0$.

Now we claim that, for each j, $\sum_i y_{ij} > \sum_i x_{ij}$.

If, for each i, $x_{ij} = 0$, then, by Fact 5, $\sum y_{ij} > \sum x_{ij}$. If there exists i such that $x_{ij} \neq 0$, then, using Fact 2 and Fact 3, one can prove that $\sum y_{ij} > \sum x_{ij}$.

Now $\sum_i y_{ij} \geq \sum_i x_{ij} + 1$ and $\sum_{ij}(y_{ij} - x_{ij}) \geq n$. Hence $t \geq n$.

□

References

[Bi] R. H. Bing, An alternative proof that 3-manifolds can be triangulated, Ann. Math. 69 (1959), 37-65.

[CG] A. Casson and C. McA. Gordon, Reducing Heegaard splittings, Topology and its applications, 27 (1987), 275-283.

[Fr] Froham, C., The topological uniqueness of triply periodic minimal surfaces in R^3, J. Diff. Geometry, 31 (1990), 277-283.

[Ha] W. Haken, Some results on surfaces in 3-manifolds, Studies in Modern Topology, Math. Assoc. Am., Prentice Hall, 1968, 34-98.

[He] P. Heegaard, Forstudier til en topologisk teori for de algebraiske fladers sammenhaeng, Ph. D. thesis, Copenhagen, 1898.

[Hem] J. Hempel, 3-manifolds, Annals of Math Studies 86 (1976), Princeton University press.

[Ja1] W. Jaco, Lectures on Three-manifold Topology, Regional Conference Series in Mathematics 43 (1981), Amer. Math. Soc..

[Ja2] W. Jaco, Adding a 2-handle to a 3-manifold: An application to Property R, Proc. Amer. Math. Soc. 92 (1984), 288-292.

[Kn] H. Kneser, Geschlossene Flächen in dreidimensionalen Mannigfaltigkeiten, Jahresbericht der Deut Math. Verein., 38 (1929), 248-260.

[Mi1] J. Milnor, A unique factorization theorem for 3-manifolds, Amer. J. Math. 79 (1957), 623-630.

[Mi2] J. Milnor, Morse Theory, Annals of Math. Studies 51 1963, Princeton University press.

[Mo] E. Moise, Affine structures in 3-manifolds V: the triangulation theorem and Hauptvermutung, Ann. of Math. 55 (1952), 96-114.

[RS] C. Rourke and B. Sanderson, Piece linear topology, Springer, 1982.

[SS] M. Scharlemann and J. Schultens, Tunnel number of the connect sum on n knots is at least n, Topology 38 (1999), 265-270.

[ST1] M. Scharlemann and A. Thompson, Heegaard splittings of (surface)$\times I$ are standard, Math. Ann., 295 (1993), 549-564.

[ST2] M. Scharlemann and A. Thompson, Thin position and Heegaard splittings of the 3-sphere, Jour. Diff. Geom., 39 (1994), 343-357.

[ST3] M. Scharlemann and A. Thompson, Thin position for 3-manifolds, AMS Contemporary Math. 164 (1994), 231-238.

[Wa] F. Waldhausen, Heegaard-Zerlegungen der 3-sphäre, Topology, 7 (1968), 195-203.