Whitehead Graphs on Handlebodies

*John R. Stallings**

Abstract. A subset A of a free group F is called "separable" when there is a non-trivial free factorization $F = F_1 * F_2$ such that each element of A is conjugate to an element of F_1 or of F_2. A single element α is separable if and only if it belongs to a proper free factor. An algorithm is given to detect if a given finite set A is separable or not; this depends on cut vertices in the Whitehead graph of A relevant to a given free basis X of F. Disjoint simple closed curves A on the boundary of a handlebody H are said to be "geometrically separable" when there is a disk D properly and non-trivially embedded in H whose boundary does not intersect any element of A. It is shown that separable in the algebraic sense implies geometrically separable.

1991 Mathematics Subject Classification: 20E05, 20F32, 57N10.

0. Introduction

In [23], J. H. C. Whitehead invented a method to show whether certain subsets of a free group form part of a basis; this involved finding cut vertices in a certain graph and simplifying the situation by a certain kind of automorphism. In the current paper, a similar thing is done in Section 2, but with a different outcome in mind; we are particularly interested in discerning whether or not an element of a free group belongs to a proper free factor. Berge [1] has noted a result like this, and it can be said to be "obvious" to those who understood Whitehead's thoughts; in fact, much of this can be distilled from I.4 of Lyndon and Schupp [12]. The Whitehead graph and Whitehead automorphisms (in particular, the existence of cut vertices under certain situations) have been used by several people to say detailed results about the automorphism group of a free group; among these are Whitehead [24], Rapaport [17], Higgins and Lyndon [9], Hoare [10], Gersten [6], McCool [14], Goldstein and Turner [7]. Whitehead attributes some of his ideas to Singer [18]; cf. also Haken [8]. In the current paper, the theory is managed by using some constructions in 3-manifolds similar to those which Whitehead used. The basic result (Theorem 2.4) says this: If A is a finite subset of a free group $F(X)$ with given basis X, and if there is a non-trivial free factorization $F = F_1 * F_2$ in which the elements of A, suitably conjugated, lie in the factors, then the Whitehead

* This research was partly supported by NSF Grant DMS-9503034.

graph of A has a cut vertex; using this, an automorphism of F changing the basis can be found which reduces the total cyclic length of A.

A technique due to Volodin, Kuznetsov, and Fomenko [22], which is related to methods of Singer [18], uses a "wave" to change a Heegaard diagram. This idea applies to the picture of curves on the boundary of a handlebody. The result is (Theorem 3.2) that if A is a finite set of disjoint simple closed curves on the boundary ∂H of a handlebody, and if, on the group level, A can be separated into proper free factors of $F = \pi_1(H)$, then this separation can be detected geometrically by a disk D not intersecting A. An algorithm can be found to produce wave transformations of the description of H simplifying the Whitehead graph. Lyon [13] and Starr [20, 21] have their own versions of this. Here we use the Whitehead graph and cut vertices in order to produce algorithms to simplify the picture by Whitehead automorphisms or waves; much of the other work in the subject, such as [12], is concentrated on the idea of using all possible Whitehead automorphisms to find one which decreases the complexity. Starr [21] has geometric method using "pairs of pants" to find a wave in the handlebody situation.

The key idea for the handlebody picture here is that the Whitehead graph is embedded in a sphere and that a cut vertex produces an innermost complementary component around which a wave can be drawn. This yields information about "strongly irreducible" Heegaard diagrams; in particular, there is an algorithm using the Whitehead graph to see if certain situations involve incompressible surfaces in three-manifolds; see Przytycki[16], Domergue and Short [5], Jaco [11], Casson and Gordon [3], and Canary [2].

1. Basic definitions

Words, reduced, cyclically reduced. Let X be a set; let X^{-1} denote a set in one-to-one correspondence with X and disjoint from X, the element $x^{-1} \in X^{-1}$ corresponding to the element $x \in X$; let $(x^{-1})^{-1} = x$. A *word* is an ordered n-tuple $w = u_1 \cdots u_n$, where $u_i \in X \cup X^{-1}$; the number n is the *length* of w. A word w is *reduced* when, for all $i = 1, \ldots, n-1$, it is the case that $u_i^{-1} \neq u_{i+1}$. A word w is *cyclically reduced*, when it is reduced, and also $u_n^{-1} \neq u_1$. A *cyclic word* is to mean a word and all of its cyclic permutations.

Free group. The *free group* $F = F(X)$ with *basis* X can be defined in various ways. One characterization is that it consists of the reduced words in X, as defined above, with the group operation being the result of concatenation of words followed by a sequence of reductions resulting in a reduced word. Thus every element of F is represented by a unique reduced word, and every conjugacy class of F is represented by a unique cyclically reduced cyclic word.

Topological picture of a free group. One topological way to describe a free groups is that F is $\pi_1(,)$, the fundamental group of a graph , which is the wedge of circles. A disadvantage of this is that elements of the group F cannot be represented by embedded simple closed curves.

The 3-manifold picture, sphere structure. A technique used by Whitehead involves certain 3-manifolds; in a 3-manifold any closed curve is approximable by an embedding. The free group $F(X)$ (with $X = \{x_1, \ldots, x_k\}$ finite) is the fundamental group $\pi_1(M)$ of the 3-manifold obtained as the connected sum of k copies of $S^1 \times S^2$; in this topological picture everything is to be smooth and as transverse as possible; in such an M there is a collection of disjoint 2-spheres $\{\Sigma_1, \ldots, \Sigma_k\}$ related to a basis of the fundamental group: Each Σ_i has two sides, denoted by Σ_i^{+1} and Σ_i^{-1}. The element $x_i \in \pi_1(M)$ is represented by a closed path starting from the basepoint of M which belongs to none of the Σ_j, going to Σ_i^{-1}, piercing Σ_i, and returning to the basepoint from Σ_i^{+1}. Any directed closed path in M hitting the Σ_i transversely represents a word in X by the way it pierces each Σ_i and the order in which it does so, provided one chooses a starting point as the basepoint; without a basepoint chosen, such a closed path represents a conjugacy class, a cyclic word. The result of cutting M along all the Σ_i is denoted $\widehat{M}$; it is a 1-connected compact 3-manifold with boundary. The boundary $\partial(\widehat{M})$ consists of 2-spheres, $2k$ in number, which can be identified with the $\Sigma_i^{\pm 1}$. It is the case that $F = \pi_1(M)$ is free of rank k; a collection of k such 2-spheres in M, $\{\Sigma_1, \ldots, \Sigma_k\}$, whose complement in M is connected, determines a basis X of F; call any such system of 2-spheres a *sphere structure* in M.

Separable set in a free group. A set $A \subset F$ is said to be *separable*, if there exists a nontrivial free decomposition $F = F_1 * F_2$, such that for each $\alpha \in A$, there exists $w \in F$ such that $w\alpha w^{-1} \in F_1 \cup F_2$. In other words, the conjugacy classes of A can be separated into two sets, the first conjugate in F to elements of F_1 and the second to elements of F_2. In particular, a singleton subset $\{\alpha\}$ is separable if and only if α belongs to a proper free factor of F.

The Whitehead graph. Given a set X and a set of words A representing elements of the free group $F(X)$, define the *Whitehead graph* (or "coinitial" or "star" graph) , = , (A, X) as follows: Let V consist of the set $X \cup X^{-1}$ as in the definition of word in a free group; the vertices of , form the set V. Write each element $\alpha \in A$ as a word in terms of $X \cup X^{-1}$. If α has length n, then it creates n edges in , ; when, cyclically, α contains the word of length 2 of the form $v_1 v_2$, then there is an edge in , from v_1 to v_2^{-1}. Thus, if

$$\alpha = v_1 v_2 \cdots v_n, \qquad v_i \in V,$$

there are edges joining v_1 to v_2^{-1} and v_2 to v_3^{-1}, ..., v_{n-1} to v_n^{-1}, and v_n to v_1^{-1}. The valence (number of adjacent edges) of each vertex v is equal to the valence of the corresponding vertex v^{-1}. The shortest case, where $\alpha = v$ is of length 1, produces one edge which joins v to v^{-1}. If the letter $x \in X$ does not occur in any of the $\alpha \in A$, then the two vertices x, x^{-1} of , are isolated, not adjacent to any edge. An instance within α of $x_i x_i^{-1}$, yields a loop, an edge of , which starts and terminates at the vertex x_i; similarly, if α has initial letter and final letter which are inverses of each other, there is a loop.

The cyclically reduced situation. In case the set A consists of cyclically reduced words, the Whitehead graph , (A, X) depends only on the conjugacy classes of A; but it depends greatly on the basis X. In , there are no loops, but there may be several edges starting and ending at the same pair of vertices.

The topological picture of the Whitehead graph. Look at the 3-manifold picture as described above. A number of disjoint simple closed curves in M will represent the finite set A of words, so that the orders in which they pierce the Σ_i determine the words up to cyclic permutation. Cut M along the Σ_i to create $\widehat{M}$. Look at the pieces of the A curves inside of $\widehat{M}$. Interpret the Whitehead graph , (A, X) thus: The vertices correspond to the boundary spheres $\Sigma_i^{\pm 1}$ of $\widehat{M}$, and the edges correspond to the arcs of the A curves in $\widehat{M}$.

Whitehead automorphism. Divide the set $V = X \cup X^{-1}$ into two disjoint subsets; that is, $V = Y \cup Z$ with $Y \cap Z = \emptyset$; suppose that there is a vertex v which belongs to Y such that its inverse v^{-1} belongs to Z. Define the corresponding *Whitehead automorphism* $\phi_{(Y,Z;v)} : F(X) \to F(X)$ thus on the basis X: For $x \in X$:

If both x and x^{-1} belong to Y, then $\phi_{(Y,Z;v)}(x) = x$.
If both x and x^{-1} belong to Z, then $\phi_{(Y,Z;v)}(x) = vxv^{-1}$.
If $x = v$ or $x = v^{-1}$, then $\phi_{(Y,Z;v)}(x) = x$.
If x is neither v nor v^{-1}, then
if $x \in Y$ and $x^{-1} \in Z$, then $\phi_{(Y,Z;v)}(x) = vx$.
if $x \in Z$ and $x^{-1} \in Y$, then $\phi_{(Y,Z;v)}(x) = xv^{-1}$.

The interpretation in the topological picture of F as $\pi_1(M)$ as above is this: Find a 2-sphere S in $\widehat{M}$ which separates the boundary components of $\widehat{M}$ into the two parts Y and Z. Take the basepoint to be in the Y part of the picture. Identify the two boundary components of $\widehat{M}$ corresponding to v and v^{-1}, and cut along S. In other words, in the sphere structure on M replace one of the 2-spheres cut along (the v sphere) by a different one (the one labeled S). Describe the change of basis of the free group F that corresponds to this change; the result will be the automorphism $\phi_{(Y,Z;v)}$.

2. Basic facts about the Whitehead graph

The hypotheses in this section. There is a finite set X. This determines the free group $F = F(X)$. There is a finite set A of cyclically reduced words. Thus, A is a way of representing a finite set of conjugacy classes in F. Define the *complexity* of this situation to be the sum of the lengths of the elements of A. Let , $=$, (A, X) be the Whitehead graph.

Proposition 2.1. *If , contains an isolated vertex, then A is separable.*

Suppose, for instance, that this isolated vertex is labeled x_1. Then A has no instance of x_1 in any element. Thus, A belongs to the proper free factor with basis $\{x_2, \ldots, x_k\}$. □

Proposition 2.2. *Assume that , is not connected. Let , contain the non-empty connected component P; let Q denote the complement of P in , ; hence, $Q \neq \emptyset$. Consider the sets Y and Z of vertices of P and Q, respectively.*

(a) If Y and Z are closed under inverse (i.e., $v \in Y$ implies $v^{-1} \in Y$), then A is separable.

(b) If the other possibility occurs, so that there is $v \in Y$ and $v^{-1} \in Z$, then the Whitehead automorphism $\phi_{(Y,Z;v)}$ shows that A is separable. The result is that $\phi_{(Y,Z;v)}(A)$ does not involve the letter v, and thus comes under the situation of Proposition 2.1.

(Note that Proposition 2.1 is a special case of this Proposition.)

Case (a) is easy because the free group F splits into two factors corresponding to the splitting of the basis by Y and Z; the elements of A have no subwords joining an element of Y to an element of Z; thus each element of A involves only the generators in Y or only the generators in Z.

The second case involves a computation which an example may illustrate (so that the reader can generalize it to a proof). In $F(a, b, c, d)$ consider the element $\alpha = acd^{-1}b^{-1}dca^{-1}b$. In , there are edges connecting the following pairs: (a, c^{-1}), (c, d), (d^{-1}, b), (b^{-1}, d^{-1}), (d, c^{-1}), (c, a), (a^{-1}, b^{-1}), (b, a^{-1}). The first, second, fifth, and sixth of these form a square graph, which is disjoint from the square graph formed by the third, fourth, seventh, and eighth. The first part is P which has vertices $Y = \{a, c^{-1}, d, c\}$ and the second part is Q which has vertices $Z = \{b, d^{-1}, b^{-1}, a^{-1}\}$. Now, compute $\phi_{(Y,Z;a)}$, as follows:

$\phi_{(Y,Z;a)}(a) = a$.
$\phi_{(Y,Z;a)}(b) = aba^{-1}$.
$\phi_{(Y,Z;a)}(c) = c$.
$\phi_{(Y,Z;a)}(d) = ad$.

Then apply this to the element α:

$$\begin{aligned}
&\phi_{(Y,Z;a)}(a \cdot c \cdot d^{-1} \cdot b^{-1} \cdot d \cdot c \cdot a^{-1} \cdot b) \\
&\qquad = a \cdot c \cdot d^{-1}a^{-1} \cdot ab^{-1}a^{-1} \cdot ad \cdot c \cdot a^{-1} \cdot aba^{-1} \\
&\qquad = acd^{-1}b^{-1}dcba^{-1}.
\end{aligned}$$

This is conjugate to a cyclically reduced word in which a does not occur.

In the 3-manifold picture, what has happened is that the new sphere structure on M related to the change of basis given by the Whitehead automorphism $\phi_{(Y,Z;a)}$ includes one sphere which does not intersect a set of curves representing A (one might have to untangle the A curves somewhat, but the resulting new set still have the same intersections with the 2-spheres). □

Cut vertex. Let Γ be a connected graph. A *cut vertex* v of Γ is a vertex such that the graph decomposes into two non-trivial graphs Γ_1 and Γ_2 which intersect only in the vertex v. Thus, the edges incident to v in Γ decompose into two disjoint non-empty sets whose other vertices constitute sets of vertices such that any path in Γ connecting them must go through v.

Proposition 2.3. *Suppose that Γ is connected and that v is a cut vertex decomposing Γ into two non-trivial subgraphs Γ_1 and Γ_2, which intersect only in v. Suppose that Γ_2 contains the vertex v^{-1}. Let Y be the set of vertices of Γ_1, and Z the set of vertices of Γ_2 with the vertex v removed. Then the complexity of $\phi_{(Y,Z;v)}(A)$ is strictly less than the complexity of A.*

In fact what happens is that the complexity decreases by at least the number of edges of Γ_1 which are incident to v.

In the 3-manifold picture the sphere structure is changed by replacing the v sphere by a sphere which encloses the spheres labeled by Z; the valences of the v and v^{-1} vertices, which are each equal to the number of intersections of the v sphere with the A curves, have been replaced in the computation of the complexity by the number of intersections with the new sphere S in the picture, which corresponds to the part of Γ_2 which is incident with v. The resulting picture may involve non cyclically reduced words, but the process of reducing them further decreases the complexity.

An example to work out and thus generalize: The set $X = \{a, b, c, d\}$ and the set A consists of one element $\alpha = abad^{-1}cac^{-1}aab^{-1}d^{-1}$. The graph Γ has two cut vertices, a and a^{-1}; with respect to the vertex a, the sets Y and Z are: $Y = \{a, b^{-1}, d\}$ and $Z = \{a^{-1}, b, c, c^{-1}, d^{-1}\}$. The Whitehead automorphism $\phi_{(Y,Z;a)}$ takes a to a, b to ba^{-1}, c to aca^{-1}, and d to ad. This takes the element α to $abd^{-1}cac^{-1}aab^{-1}d^{-1}a^{-1}$, and this conjugates to an element of length two less. The complexity has been reduced from 11 to 9; the difference is the number of edges of Γ_1 incident to a. □

Theorem 2.4. *If Γ is connected and if A is separable in F, then there is a cut vertex in Γ.*

Proof. First, an outline of the proof, then more rigorous details.

Outline of proof: In the 3-manifold picture, represent $F = \pi_1(M)$, where M contains 2-spheres $\{\Sigma_1, \ldots, \Sigma_k\}$ corresponding to the basis X; represent A by a finite collection of disjoint simple closed curves in M. The assumption that A is separable in F implies that there exists a surface S in M which does not intersect A. The surface S corresponds to a non-trivial free factorization of F in such a way that a cover of S in the universal cover $\widetilde{M}$ is not null-homologous. Consider such a surface S which is minimal; it will be connected and will intersect the spheres Σ_i in a minimal total number of components. By lifting S to the universal cover $\widetilde{M}$, and projecting to a tree, the existence of at least one component T of $S \cap \widehat{M}$, with boundary ∂T contained in only one $\Sigma = \Sigma_i^{\pm 1}$ can be proved. If it happened that ∂T did not divide up that sphere Σ into two open sets both of which contain points of intersection with the A curves, then a contradiction would occur: Either T is homologically trivial in $(\widehat{M}, \partial \widehat{M})$, in which case an improvement of S could be done reducing the complexity of it; or else, T is not homologically trivial, in which case , is not connected. Thus T in fact divides the graph , into two parts intersecting only in $v = \Sigma$, and v is then a cut vertex of ,

More details:

(A) *The 3-manifold picture*. Look at the 3-manifold picture as described in Section 1. This involves M with $\pi_1(M) = F$, the sphere structure $\{\Sigma_1, \ldots, \Sigma_k\}\}$ related to the basis X, and a set of A curves $\{\alpha_1, \ldots, \alpha_n\}$ which intersect the sphere structure according to the expressions of the elements of A as cyclic words in X. The Whitehead graph , is evident in $\widehat{M}$ with the $2k$ vertices $\Sigma_j^{\pm 1}$ and the edges which are arcs in the A curves.

(B) *The universal cover*. Let $\widetilde{M}$ denote the universal covering space of M. There is an infinite tree in which every vertex has valence $2k$; each edge is directed and labeled with one of the k symbols Σ_j, so that each vertex is incident to one edge labeled Σ_j that leaves that vertex and to one such that enters the vertex. Each edge in this tree corresponds to a lift of one of the Σ_j, and each vertex corresponds to a lift of $\widehat{M}$. Thus, $\widetilde{M}$ is made of infinitely many copies of $\widehat{M}$ joined along their boundary spheres; a copy of Σ_j^{+1} in one copy of $\widehat{M}$ is identified with a copy of Σ_j^{-1} in an adjacent copy of $\widehat{M}$.

(C) *The separant surface S*. The assumption that A is separable in F means that there exists an isomorphism $F \approx F_1 * F_2$, where both F_1 and F_2 are non-trivial, such that, appropriate conjugates being chosen, each element of A corresponds to an element of F_1 or F_2. Realize $F_1 * F_2 = \pi_1(P_1 \vee P_2)$, where P_i is a graph with fundamental group F_i, and where "$\vee$" denotes the disjoint union with an arc added connecting basepoints; let p denote the central point of this arc. There exists (using the fact that $P_1 \vee P_2$ is aspherical) a continuous map $f : M \to P_1 \vee P_2$ giving this isomorphism on π_1. The curve α_i maps by f in such a way as to be homotopic to a map into P_1 or into P_2; the homotopy extension theorem then changes f so that $f(\alpha_i)$ is contained in either P_1 or P_2; smooth out f so that it behaves transversely to p and nicely with respect to the Σ_j. Now consider $f^{-1}(p) = S$; this is a surface in M, possibly not connected; it has three fundamental properties:

(a) For all j, the intersection $S \cap \alpha_j$ is empty.

(b) The surface S can be lifted homeomorphically to $\widetilde{M}$, to obtain a compact surface S' which represents a non-trivial element of $H_2(\widetilde{M})$, where this denotes homology with coefficients in $\mathbf{Z}/2\mathbf{Z}$.

(c) The surface S is transverse to all the Σ_j, and thus intersects each one in a finite number of simple closed curves.

Call any surface in M satisfying these properties, a *separant surface*.

The reason for (b) is that the universal cover of M maps to the universal cover $(P_1 \vee P_2)^{\sim}$ because of the isomorphism of fundamental groups. A lift of S in $\widetilde{M}$ corresponds to the pullback of a lift p' of the point p in the universal cover of $P_1 \vee P_2$; denote such a lift of S by S'. There is (because of the isomorphism on π_1) an isomorphism of $H^1_f((P_1 \vee P_2)^{\sim})$ and $H^1_f(\widetilde{M})$, cohomology with finite cochains. The point p' is dual to a non-zero element of $H^1_f((P_1 \vee P_2)^{\sim})$; it decomposes $(P_1 \vee P_2)^{\sim}$ into two non-compact pieces since F_1 and F_2 are nontrivial. Thus, S' is Poincaré dual to a non-zero element of $H^1_f(\widetilde{M})$ and is hence non-zero in $H_2(\widetilde{M})$.

(D) *The complexity of a separant surface*. Any separant surface S is compact and therefore has a finite number of connected components. By property (c), the number of components of intersection with each Σ_j is finite. Let $\beta(S)$ be the sum over all j of the number of components of $S \cap \Sigma_j$; and let $\gamma(S)$ be the number of components of S. The pair $(\beta(S), \gamma(S))$, lexicographically ordered, is the "complexity" of S. The possible complexities form a well-ordered set, and thus there is a minimum complexity.

(E) *Existence of cut vertex*. We now suppose that S is a separant surface of minimum complexity, and show how this produces a cut vertex in , (A, X).

(E.1). The separant surface S is connected. Otherwise, it would be the union of a finite number of components; the sum of these components, lifted to $\widetilde{M}$, homologically would equal S', and thus one such component of S' would be non-zero homologically. That component, S_1, viewed in M, would have smaller complexity since $\beta(S_1) \leq \beta(S)$ and $\gamma(S_1) < \gamma(S)$. It would clearly satisfy the three conditions listed in (C).

(E.2). The separant surface S intersects some Σ_j non-trivially. Otherwise, S would be totally contained in the interior of $\widehat{M}$; it separates $\widehat{M}$ into two pieces Y and Z. The homological non-triviality of a lift S' in $\widetilde{M}$ then implies that each of these pieces contains at least one of the boundary spheres $\Sigma_j^{\pm 1}$. Since S does not intersect the A curves, it follows that the Whitehead graph , (A, X) has been decomposed into two non-empty pieces which are not connected to each other. This is contrary to the hypothesis of Theorem 2.4 that , is connected.

(E.3). Lift S to S' in $\widetilde{M}$, and examine the tree-like structure of $\widetilde{M}$, discussed in (B). The image of S' in this tree is connected and has at least one edge; thus it is a tree itself, a finite tree; and thus it contains an extremal vertex, a vertex with valence one. What this means is that there is a copy of $\widehat{M}$ in $\widetilde{M}$, call it $\widehat{M}_1$, so that the intersection $T' = \widehat{M}_1 \cap S'$ is non-empty, and such that T' has all its boundary curves on one and only one boundary sphere of $\widehat{M}_1$. Look back down in $\widehat{M}$ at this situation. Define $T = \{T_1, \ldots, T_m\}$ to consist of those components of $S \cap \widehat{M}$ which intersect $\partial\widehat{M}$ only in a subset of the single sphere Σ that corresponds to the sphere in $\widetilde{M}$ containing $\partial T'$. The argument in the universal cover shows that T is non-empty. Each T_i may have several boundary curves, but they are all contained in the one sphere Σ. Note that each component T_i of T separates $\widehat{M}$ into two pieces, since $\partial T_i = T_i \cap \partial \widehat{M}$, and $\widehat{M}$ is simply connected.

Consider one component T_1 of T. Then ∂T_1 divides the sphere Σ into two parts K_1 and L_1 with $\partial K_1 = \partial L_1 = \partial T_1$; it is not necessarily the case that K_1 or L_1 is connected. The claim is that both K_1 and L_1 contain points of intersection of the A curves.

(E.4). Suppose that K_1, for instance, did not contain any point of the A curves. Then look at $T_1 \cup K_1$ and push slightly into the interior of $\widehat{M}$; the result is a closed surface U in $\widehat{M}$ which does not intersect A and does not intersect any Σ_i. The complexity of U is less than that of S, since S must (by E.2) intersect some Σ_i. Thus, U is not separant, and thus it must be null-homologous when lifted to $\widetilde{M}$. This says that the 3-dimensional region bounded by U in the interior of $\widehat{M}$ must contain no boundary sphere of $\widehat{M}$. This implies that T_1 separates $\widehat{M}$ into two parts, one of which intersects only the boundary sphere Σ in K_1 and no other part of $\partial \widehat{M}$. Call that part R_1, so that $\partial R_1 = T_1 \cup K_1$, and $R_1 \cap \partial \widehat{M} = K_1 \subset \Sigma$.

Let T_i be a component of T which intersects R_1. Then it is contained in R_1; it divides $\widehat{M}$ into two pieces, one of which, R_i, does not intersect T_1. Now the claim is that $R_i \subset R_1$; otherwise there would be points of R_i (near T_i) in the component R_1 of $\widehat{M} \setminus T_1$, and another point in the other component of $\widehat{M} \setminus T_1$; this would show that T_1 would not disconnect $\widehat{M}$.

Thus, there is an innermost component T_2 of T. This has the property that it separates $\widehat{M}$ into two parts, one of which is $R_2 \subset R_1$, with the property that R_2 does not otherwise intersect T. Let $K_2 = R_2 \cap \Sigma$. This is contained in K_1, and thus there is no intersection of an A curve with K_2. Change S by removing T_2 and adding K_2; then push slightly to the other side of the sphere Σ, getting a surface J. Now, homologically up in $\widetilde{M}$ it is the case that J' is homologous to S' and thus is non-zero. The number of components of J may be greater than 1, but the number of curves of intersection of J with the Σ_j is reduced. In other words, $\beta(J) < \beta(S)$, and no relation can be deduced about $\gamma(J)$ versus $\gamma(S)$. In any case, this would contradict the choice of S as a separant surface of minimal complexity.

Thus K_1 and, by symmetry, L_1 both intersect the A curves nontrivially.

(E.5). Since K_1 and L_1 both intersect the A curves, then the vertex represented by Σ is a cut vertex of , (A, X). For, T_1 shows how to divide , into two non-trivial pieces that intersect only in this vertex. □

Algorithm 2.5 to detect separability. Let a finite basis X of the free group F be given, and let a finite number of elements A of F be given.

(*) Represent each element of A by a cyclically reduced cyclic word in X. The complexity is the sum of the lengths of these elements.

Construct the Whitehead graph , (A, X).

If , is not connected, then by Propositions 2.1 and 2.2, A is separable.

If , is connected, search for a cut vertex.

If there is no cut vertex, then by Theorem 2.4, the set A is not separable.

If there is a cut vertex, apply the construction in Proposition 2.3, find a Whitehead automorphism. This changes the basis X in such a way as to reduce the complexity. Return to (*). $\square$

Remarks. It follows from this that one can consider another type of algorithm. List all possible Whitehead automorphisms, apply each and compute whether the complexity of A has been reduced; if the complexity is reduced, continue. If no Whitehead automorphism reduces the complexity, then check to see whether or not the basis X can be divided into two non-empty parts, such that each element of A is a cyclic word in one of the two parts. This algorithm is probably more time-consuming than the algorithm checking for the cut vertex, etc. If there is no cut vertex, then we have found non-separability, although there might yet exist a Whitehead automorphism reducing the complexity.

Corollary 2.5. *Let $F = F(X)$ be the free group on a finite basis X, and let $\alpha \in F$ be written as a cyclically reduced word. If the Whitehead graph $\Gamma(\alpha, X)$ is connected and contains no cut vertex, then α does not belong to any free factor of F.*

This is the case of Theorem 2.4 in which A is a singleton set. This was proved by Berge [1], using techniques of Goldstein and Turner [7].

3. Curves on a handlebody

Handlebody, disk structure, essential disk. A *handlebody* of genus k is a compact orientable 3-manifold H with boundary ∂H which is a connected surface of genus k, such that there exist k disks $\Delta_1, \dots, \Delta_k$ properly embedded in H (that is, $\Delta_i \cap \partial H = \partial \Delta$) and such that the result $\widehat{H}$ of cutting H along all the Δ_i is a 3-cell; call the handlebody H together with the collection of these disks a *disk structure* on H. Each disk Δ_j is two-sided; denote the two sides by Δ_j^{+1} and Δ_j^{-1}. In $\partial(\widehat{H})$ there are $2k$ disjoint disks that can be identified with $\Delta_j^{\pm 1}$. In general, in a handlebody H an *essential disk* is a disk Δ properly embedded in H such that, when H is cut along Δ, the result is either connected or else consists of two parts neither of which is a 3-cell.

There are many possible collections of disks that will have the property that the result of cutting H along them is a 3-cell; such a collection can be recognized (assuming everything is smooth) easily: There are k such disks, all disjoint and properly embedded in H, and their complement in H is connected. A choice of a collection of such disks gives a choice of a free basis of the free group $F = \pi_1(H)$ as in the topological picture used in Section 2; the 3-manifold M in Section 2 is analogous to H, and the spheres Σ_j there are analogous to the disks Δ_j here.

Waves and change of disk structure. Suppose H is a handlebody with disks $\Delta_1, \ldots, \Delta_k$ cutting H into $\widehat{H}$ which is a 3-cell. Call an arc $E \subset \partial H$ a *wave*, provided that E intersects the curves $\partial \Delta_j$ in only two points, the endpoints of E, and that these two points are both on the same side of a single $\partial \Delta_j$. In that case, on $\partial(\widehat{H})$ it is possible to draw the arc E, starting and ending at a single disk; to be specific, suppose that disk is the disk Δ_1^{+1}; then $E \cup \Delta_1^{+1}$ separates the 2-sphere $\partial(\widehat{H})$ into two pieces, one of which will contain the complementary disk Δ_1^{-1}; the boundary of that piece consists of the arc E together with one of the two arcs B_1, B_2 on $\partial \Delta_1^{+1}$ with endpoints ∂E; suppose this piece has boundary $C = B_1 \cup E$. Then C bounds a disk D properly embedded in $\widehat{H}$, since the latter is a 3-cell with C on its boundary. This can be seen also in H. Now, the disk Δ_1 can be removed from the picture and replaced with the disk D. The result of cutting H along the disks $\{D, \Delta_2, \ldots, \Delta_k\}$ is connected; the result of cutting along them consists of $\widehat{H}$ cut along D and with the two disks Δ_1^{+1} and Δ_1^{-1} identified; the choice of $\partial D = C$ separating Δ_1^{+1} and Δ_1^{-1} on $\partial \widehat{H}$ makes this connected. In this manner a wave determines a change of disk structure. This results in terms of fundamental group in a Whitehead automorphism. However, the topological situation is more constrained here than in the case of Section 2, and so not all Whitehead automorphisms can be constructed in this way.

The handlebody picture. Consider a handlebody H of genus k with disk structure $\Delta_1, \ldots, \Delta_k$. On the boundary ∂H, consider a collection of finitely many disjoint simple closed curves $A = \{\alpha_1, \ldots, \alpha_n\}$. Let F be the free group with basis $\{x_1, \ldots, x_k\}$, realized as $\pi_1(H)$, in which x_i is realized as a loop starting at some basepoint disjoint from the Δ_j, going to Δ_i^{-1}, through Δ_i, and back to the basepoint from Δ_i^{+1}. Everything is supposed to be smooth and transverse so that the A curves determine, by their intersections with the $\partial \Delta_i$ on ∂H, cyclic words in F with this basis; it is assumed that no α_i represents the identity element of F. Let , = , (A, X) be the Whitehead graph of A with respect to X. The complexity is defined to be the total number of intersections of the A curves with all the $\partial \Delta_j$.

Proposition 3.1. *In the handlebody picture, the cyclic words represented by A can be made into cyclically reduced words by a finite sequence of changes of disk structure by waves modeled on the reductions in the words of A.*

As an example, suppose that α_1 contains, considered as a cyclic word, a subword $x_1 x_1^{-1}$. Then α_1, as a curve on ∂H, contains an arc E which is a wave joining two points on the side Δ_1^{+1} of Δ_1. Do the change of disk structure modeled on E. As described above, ∂E consists of two points on the circle $\partial \Delta_1$, dividing the latter into two arcs B_1 and B_2. The change replaces Δ_1 by D which has $\partial D = E \cup B_1$, assuming one of the two possibilities for the location of Δ_1^{-1} in $\partial(\widehat{H})$. Part of the intersection of ∂D with α_1 is the arc E; push this arc of ∂D slightly in the direction of B_1; such a "direction" exists since ∂H is orientable. In the resulting situation the intersections of α_j with the boundary circles of the new disks have not changed, except that all the intersections with B_2 have been removed and the two intersections ∂E have been eliminated. Thus the complexity of the picture has been decreased. A finite number of these terminates with the situation where A represent cyclically reduced words. $\square$

Theorem 3.2. *In the handlebody picture, if A is separable in F, then a sequence of wave changes in the disks in the handlebody can be found which ends by finding a disk D essential in H such that ∂D intersects the A curves in the empty set.*

The proof involves decreasing the complexity to a minimum by waves that can be found by looking at the graph , as it is embedded in $\partial(\widehat{H})$. The Whitehead graph , (A, X) exists in $\partial(\widehat{H})$ as the arcs made out of pieces of the curves A, together with the disks $\Delta_i^{\pm 1}$ (these disks are the "vertices" of ,); thus, it is a planar graph, embedded in a particular way in the 2-sphere $\partial(\widehat{H})$.

If , is not connected, then it has various components. They are embedded in the 2-sphere $\partial(\widehat{H})$, and thus, there is an innermost such component; draw a curve around this innermost component; the result is a simple closed curve on $\partial(\widehat{H})$ which does not hit any of the disks $\Delta_i^{\pm 1}$; that curve bounds a disk D in $\widehat{H}$, and its preimage in H is the desired disk. The reason this D is essential is that it separates non-empty collections of the $\Delta_i^{\pm 1}$ on $\partial(\widehat{H})$.

If , is connected, but the A curves are not cyclically reduced, the wave construction in Proposition 3.1 will diminish the complexity. Then it might be that , has become disconnected, and one goes to the preceding step. If, eventually , is connected and the A are cyclically reduced, then, according to Theorem 2.4, , has a cut vertex; on $\partial(\widehat{H})$, this cut vertex corresponds to one of the disks $\Delta_i^{\pm 1}$, say to Δ_1^{+1}. The pieces of , into which this vertex cuts it lie in the 2-cell $\partial(\widehat{H}) \setminus \Delta_1^{+1}$. The important observation now is that there is an innermost such piece. An arc E can be drawn around this innermost piece, so that one can use it as a wave to change the disk structure of H. If that change replaces Δ_1 by D, then the boundary of D would be, say, $E \cup B_1$, and the intersections of A with B_2 have disappeared from the picture. Thus the complexity is reduced.

Eventually, then, the picture yields a non-connected , which produces the required D. $\square$

Remarks. The word "separable" as used here is an algebraic term. Call a finite set A of disjoint simple closed curves on the boundary of a handlebody H "geometrically separable" if there is a disk D properly and non-trivially embedded in H whose boundary is disjoint from A. The term "disk-busting" has also been used for "not geometrically separable" in a more general situation by Canary [2].

The value of Theorem 3.2 is to find the essential D by an algorithm; starting from a disk structure on H, decrease the complexity to get the curves A cyclically reduced by doing wave restructuring, look for a cut vertex in the Whitehead graph , (A, X); if there is none, then the situation is not geometrically separable; if there is a cut vertex, then a wave can be found to decrease the complexity. In the end, if A is geometrically separable on H, an essential disk D is found disjoint from A.

Another aspect of Theorem 3.2 is that it shows that algebraic separability implies geometric separability. There is another proof of this fact along these lines: Map H into $P_1 \vee P_2$ as in the construction of the separant surface in the proof of Theorem 2.4, part (C). The inverse image of p will be a surface properly embedded in H which maps trivially on fundamental group and which does not intersect A. Dehn's Lemma and the Loop Theorem ([19], [15], [4]) will simplify the picture to obtain an essential disk.

That geometrically separable implies the existence of a wave is to be found in work of Starr [20, 21].

It is not hard to see (using Dehn's Lemma and the Loop Theorem) that a singleton set A consisting of only one simple closed curve α is not geometrically separable in H, if and only if for every basepoint $\pi_1(\partial H \setminus \alpha) \to \pi_1(H)$ is injective. Theorem 3.2 yields an algorithm to determine whether α is geometrically inseparable in H; this helps determine whether certain Heegaard splittings are strongly irreducible [3]. This is useful for finding incompressible surfaces in certain 3-manifolds.

References

[1] Berge, J., Full-rank closure. Preprint (1993).

[2] Canary, R. D., Ends of hyperbolic 3-manifolds. J. Amer. Math. Soc. 6 (1993), 1–35.

[3] Casson, A. J., Gordon, C. McA., Reducing Heegaard splittings. Topology Appl. 27 (1987), 275–283.

[4] Dehn, M., Über die Topologie des dreidimensionalen Raumes. Math. Ann. 69 (1910), 137–168.

[5] Domergue, M., Short, H., Surfaces incompressibles dans les variétés obtenues par chirurgie à partir d'un nœud de S^3. Comptes Rendus Acad. Sc. Paris 300 (1985), 669–672.

[6] Gersten, S. M., On fixed points of certain automorphisms of free groups. Proc. London Math. Soc. (3) 48 (1984), 72–90.

[7] Goldstein, R. Z., Turner, E. C., Automorphisms of free groups and their fixed points. Invent. math. 78 (1984), 1–12.

[8] Haken, W., Some results on surfaces in 3-manifolds. In: Studies in Modern Topology, Prentice-Hall (1968), 39–98.

[9] Higgins, P. J., Lyndon, R. C., Equivalence of elements under automorphisms of a free group. J. London Math. Soc. (2) 8 (1974), 254–258.

[10] Hoare, A. H. M., On automorphisms of free groups I. J. London Math. Soc. (2) 38 (1988), 277–285.

[11] Jaco, W., Adding a 2-handle to a 3-manifold: an application to property R. Proc. Amer. Math. Soc. 92 (1984), 288–292.

[12] Lyndon, R. C., Schupp, P. E., Combinatorial Group Theory (Ergebnisse der Math. 89). Springer-Verlag, New York-Berlin-Heidelberg (1977).

[13] Lyon, H. C., Incompressible surfaces in the boundary of a handlebody—an algorithm. Canadian J. of Math. 32 (1980), 590–595.

[14] McCool, J., A presentation for the automorphism group of a free group of finite rank. J. London Math. Soc. (2) 8 (1974), 259–266.

[15] Papakyriakopoulos, C. D., On Dehn's lemma and the asphericity of knots. Ann. of Math. 66 (1957), 1–26.

[16] Przytycki, J. H., Incompressibility of surfaces after Dehn surgery. Michigan Math. J. 30 (1983), 289–308.

[17] Rapaport, E. S., On free groups and their automorphisms. Acta. Math. 99 (1958), 139–163.

[18] Singer, J., Three dimensional manifolds and their Heegaard diagrams. Trans. Amer. Math. Soc. 35 (1933), 88–111.

[19] Stallings, J. R., On the loop theorem. Ann. of Math. 72 (1960), 12–19.

[20] Starr, E. N., Curves in handlebodies. Doctoral Dissertation U. C. Berkeley (1992).

[21] —, A generalization of a theorem of Casson and Gordon, preprint (1996).

[22] Volodin, I. A., Kuznetsov, V. E., Fomenko, A. T., The problem of discriminating algorithmically the standard three-dimensional sphere. Uspehi Mat. Nauk 29:5 (1974), 72–168. English translation: Russian Math. Surveys 29:5 (1974), 71–172.

[23] Whitehead, J. H. C., On certain sets of elements in a free group. Proc. London Math. Soc. 41 (1936), 48–56.

[24] —, On equivalent sets of elements in a free group, Ann. of Math. 37 (1936), 768–800.