# ISOTOPY CLASSES OF INCOMPRESSIBLE SURFACES IN IRREDUCIBLE 3-MANIFOLDS 

Jeffrey L. TOLLEFSON

(Received April 25, 1994)

## 1. Introduction

Let $M$ be a compact, orientable, irreducible, $\partial$-irreducible 3-manifold with a fixed triangulation $\mathscr{T}$. We are interested in the isotopy and projective isotopy classes of compact, orientable, incompressible, $\partial$-incompressible surfaces in $M$ and how they are represented in the projective solution space $\mathscr{P}_{\mathscr{F}}$ from normal surface theory. Two surfaces belong to the same projective isotopy class if there exist multiples of each which are isotopic. We show that $\mathscr{P}_{\mathscr{g}}$ has maximal faces, called complete lw-faces, which have the following properties: (i) if a complete lw-face carries one surface in a projective isotopy class then it carries every least weight normal surface in that projective isotopy class and (ii) every surface carried by an lw-face is least weight in its isotopy class. Each complete lw-face is partitioned by compact linear cells in such a way that the set of surfaces carried by such a linear cell is precisely the set of all least weight surfaces in the corresponding projective isotopy class.

In normal surface theory there is associated to the triangulation $\mathscr{T}$ a system of matching equations whose admissible integral $n$-tuple solutions are in a one-to-one correspondence with the normal surfaces in $M$. The projections of such solutions to the unit sphere are called the projective normal classes of the corresponding normal surfaces and are contained in $\mathscr{P}_{\mathscr{F}}$, the compact, convex, linear cell of solutions to the normalized matching equations for $\mathscr{T}$. The weight of a normal surface $F$, denoted by $\mathrm{wt}(F)$, is the number of points in which $F$ interesects the 1 -skeleton of $\mathscr{T}$ and we say that $F$ is a least weight surface if it is least wieght relative to its isotopy class. It has been shown in [2], [3], and [4] that least weight surfaces exhibit some strong and useful properties.

A face $C$ of the compact, convex linear cell $\mathscr{P}_{\mathscr{F}}$ is said to be an $l w$-face if every normal surface carried by $C$ is incompressible, $\partial$-incompressible and least weight. A complete $l w$-face $C$ is an $l w$-face with the additional property that if a normal surface $F$ is carried by $C$ then every least weight normal surface isotopic to $F$ is also carried by $C$. We show in Theorem 4.5 that there is a finite collection of complete lw-faces such that every compact, orientable, incompressible, $\partial$ -
incompressible, least weight, normal surface is carried by one in the collection. A important ingredient in the proof is Theorem 4.2 which states that if $F$ is such a surface then the minimal face $C_{F}$ of $\mathscr{P}_{\mathscr{F}}$ carrying $F$ has the property that every normal surface carried by $C_{F}$ is also a least weight surface. This last theorem was announced by Oertel in [6].

Let $G$ and $H$ be normal surfaces intersecting transversely such that each component of $G \cap H$ is a regular curve. There is the normal sum $G+H$ defined by canonical cut-and-paste operations (called regular exchanges) along the intersection curves. Suppose that the normal sum $G+H$ is a compact, orientable, incompressible, $\partial$-incompressible, least weight surface. The existence of complete lw-faces is used to prove (Corollary 4.3) that given any least weight normal surfaces $G^{\prime}$ isotopic to $G$ and $H^{\prime}$ isotopic to $H$, the normal sum $G^{\prime}+H^{\prime}$ is defined and isotopic to $G+H$. Applications to simplicical group actions are given in Section 6. For example, suppose $G$ is a finite group of simplicial homeomorphisms of $M$ and $F$ is a compact, orientable, incompressible, $\partial$-incompressible, least weight, normal surface such that $g(F)$ is homotopic to a surface disjoint from $F$ for each $g \in G$. Then the normal sum $S=\Sigma_{g} g(F)$ is a $G$-equivariant collection of pairwise disjoint normal surfaces, one of which is isotopic to $F$. Moreover, if $K$ is any normal surface carried by the minimal face of $\mathscr{P}_{\mathscr{F}}$ carrying $S$, the normal sum $\Sigma_{g} g(K)$ is defined and is itself a $G$-invariant, injective, $\partial$-injective, least weight, normal surface.

Consider a set $\left\{F_{1}, \cdots, F_{k}\right\}$ of compact, orientable, incompressible, $\partial-$ incompressible, least weight, normal surfaces all in the same projective isotopy class. Thus, a multiple of $F_{i}$ is isotopic to a multiple of $F_{j}$ for each pair $i, j$. Let $V$ denote the subspace of $\boldsymbol{R}^{n}$ spanned by the normal classes $\left\{\vec{F}_{1}, \cdots, \vec{F}_{k}\right\}$. We will show that every orientable normal surface carried by the linear cell $A=V \cap \mathscr{P}_{\mathscr{F}}$ is a least weight normal surface in the same projective isotopy class as the $F_{i}$ 's. If $\left\{F_{1}, \cdots, F_{k}\right\}$ is chosen to be a maximal linearly independent set then every least weight normal surface in this projective isotopy class will be carried by $A$. We obtain a similar result for branched surfaces as a byproduct of the proof of Lemma 5.2.

In Theorem 5.5 we show that each 1 w -face $C$ of $\mathscr{P}_{\mathscr{F}}$ is partitioned by linear cells associated with projective isotopy classes. A proper face $D$ of $C$ is said to be $C$-independent if no normal surface carried by $D$ belongs to the same projective isotopy class as one carried by $\mathscr{C}$. Let $\mathscr{D}$ denote the union of all $C$-independent faces of $C$. We construct a family of $k_{C}$-dimensional subspaces $V_{\alpha}$ such that (i) each $C$-independent face $D$ is a union of a subcollection of the linear cells $V_{\alpha} \cap C$, (ii) the family $\left\{A_{\alpha}\right\}$ of cells $V_{\alpha} \cap C$ meeting $\mathscr{C}$ forms a partition of $C-\mathscr{D}$ by linear $k_{c}$-dimensional cells with the property that the collection of normal surfaces carried by each $A_{\alpha}$ is a complete projective isotopy class of surfaces carried by $C$. Next consider a maximal cell $D$ in $\mathscr{D}$ and let $\mathscr{E}$ denote the union of all $D$-independent faces. Just as for $C$, there exists a family of $k_{D}$-dimensional subspaces $V_{\beta}$ satisfying (i) each $D$-independent face is a union of some of the linear cells $D \cap V_{\beta}$, (ii) the family $\left\{B_{\alpha}\right\}$ of $V_{\beta} \cap D$ meeting $D$ forms a partition of $D-\mathscr{E}$ by


Fig. 1. A model of elementary disks
linear $k_{D}$-dimensional cells with the property that the collection of normal surfaces carried by each $B_{\beta}$ is a complete projective isotopy class of surfaces carried by $D$ and also (iii) each cell $V_{\beta} \cap D$ is contained in some $C \cap V_{\alpha}$. Repetition of the process gives rise to a partition $\left\{X_{\alpha}\right\}$ of $C$, called a PIC-partition, where $X_{\alpha}$ is a linear cell in $C$ such that the set of normal surfaces carried by each $X_{\alpha}$ consists of all the surfaces carried by $C$ and belonging to a given projective isotopy class.

In this paper all manifolds and maps are assumed to be PL and $M$ is always a compact, orientable, irreducible, $\partial$-irreducible 3-manifold with a fixed triangulation $\mathscr{T}$.

## 2. Normal surfaces

A normal surface $F$ in $M$ is a properly embedded surface in general position with the 1 -skeleton $\mathscr{T}^{(1)}$ and intersects each tetrahedron $\Delta$ of $\mathscr{T}$ in properly embedded elementary disks which intersect each edge in at most one point and each face in at most one straight line. If $E$ is an elementary disk then $\partial E$ is uniquely determined by the points $E \cap \mathscr{T}^{(1)}$ and it is convenient that $E$ itself be uniquely determined by these points. Thus, we require that each elementary disk $E$ in $\Delta$ be a planar disk if $E \cap \mathscr{T}^{(1)}$ is a planar set and otherwise require that $E$ equal the cone $b^{*} \partial E$ where $b$ is the centroid of the 3 -simplex in $\Delta$ spanned by $E \cap \mathscr{T}^{(1)}$. With this convention, a normal surface $F$ is uniquely determined by the set of points $F \cap \mathscr{T}^{(1)}$. A normal isotopy of $M$ is an isotopy which leaves the simplices of $\mathscr{T}$ invariant. An elementary disk is determined, up to normal isotopy, by the manner in which it separates the vertices of $\Delta$ and we refer to the normal isotopy class of an elementary disk as its disk type. There are seven possible disk types in each tetrahedron corresponding to the seven possible separations of its four vertices;
four consisting of triangles and three consisting of quadrilaterals. The normal isotopy class of an arc in which an elementary disk meets a 2 -face of $\Delta$ is called an arc type.

After fixing an ordering of $\left(d_{1}, \cdots, d_{7 t}\right)$ the disk types in $\mathscr{T}$, we assign to a normal surface $F$ a $7 t$-tuple $\vec{F}=\left(x_{1}, \cdots, x_{7 t}\right)$, called the normal coordinates of $F$, by letting $x_{i}$ denote the number of elementary disks in $F$ of type $d_{i}$. A normal surface is uniquely determined, up to normal isotopy, by its normal coordinates. Among $7 t$-tuples of non-negative integers $\vec{x}=\left(x_{1}, \cdots, x_{7 t}\right)$, those corresponding to normal surfaces are characterized by two constraints. The first constraint is that it must be possible to realize the required 4 -sided disk types $d_{i}$ corresponding to nonzero $x_{i}$ 's by disjoint elementary disks. This is equivalent to allowing no more than one 4 -sided disk type to be represented in each tetrahedron. The second constraint concerns the matching of the edges of elementary disks along incident 2 -faces of tetrahedra. Consider two tetrahedra meeting along a common 2 -face and fix an arc type in this 2-face. There are exactly two disk types from each of the tetrahedra whose elementary disks meet this 2 -face in arcs of the given arc type. If the $7 t$-tuple is to correspond to a normal surface then there must be the same number of elementary disks on both sides of the incident 2 -face meeting it in arcs of the given type. This constraint is given as a system of matching equations, one equation for each arc type in the 2 -simplexes of $\mathscr{T}$ interior to $M$.

$$
\begin{gather*}
\text { Matching Equations } \\
x_{i}+x_{j}=x_{k}+x_{l}  \tag{1}\\
0 \leq x_{i}, \quad 1 \leq i \leq 7 t
\end{gather*}
$$

The non-negative solutions to the matching equations (1) form an infinite linear cone $\mathscr{S}_{\mathscr{T}} \subset \boldsymbol{R}^{7 t}$. The normalizing equation $\Sigma_{i=1}^{7 t} x_{i}=1$ is added to form the system of normal equations for $\mathscr{T}$. The solution space $\mathscr{P}_{\mathscr{F}} \subset \mathscr{S}_{\mathscr{F}}$ becomes a compact, convex, linear cell and is referred to as the projective solution space for $\mathscr{T}$. The projective normal class $\vec{F}^{*}$ of a normal surface $F$ is the image of the normal coordinates of $F$ under the projection $\mathscr{S}_{\mathscr{I}} \rightarrow \mathscr{P}_{\mathscr{F}}$. A rational point $\vec{z} \in \mathscr{P}_{\mathscr{F}}$ is said to be an admissible solution if corresponding to each tetrahedron there is at most one of the quadrilateral variables which is nonzero. Every admissible solution is the projective normal class of an embedded normal surface.

The carrier of a normal surface $F$ is the unique minimal face $C_{F}$ of $\mathscr{P}_{\mathscr{F}}$ that contains the projective class of $F$ and a normal surface is said to be carried by $C_{F}$ if its projective normal class belongs to $C_{F}$. The normal surfaces $G$ which are carried by $C_{F}$ are characterized by the property that there exists a normal surface $H$ such that $m F=G+H$ for some positive integer $m$. Each rational point in $C_{F}$ is an admissible solution since its coordinates are zero in any variable corresponding to a disk type not already represented in $F$.

If two elementary disks $E_{1}, E_{2}$ in a tetrahedron $\Delta$ intersect transversely then
$E_{1} \cap E_{2}$ is an arc $\alpha$ properly emebedded in $\Delta$ and $\alpha$ spans the interior of distinct 2 -faces of $\Delta$. We say that $\alpha$ is a regular arc of intersection if there exists a pair of disjoint elementary disks having the same disk types as $E_{1}$ and $E_{2}$. This is equivalent to the property that the union of the vertices of $E_{1}$ and $E_{2}$ span a disjoint pair of elementary disks and this is alwasy the case except when $E_{1}$ and $E_{2}$ are quadrilateral disks of different disk types. Two normal surfaces $G$ and $H$ are said to intersect transversely if each pair of elementary disks from $G$ and $H$, respectively, intersect transversely. Let $G$ and $H$ be normal surfaces intersecting transversely such that each intersection curve of $G \cap H$ is regular in that it is a union of regular arcs. The points $(G \cup H) \cap \mathscr{T}^{(1)}$ determine a unique (embedded) normal surface $G+H$ called the normal sum of $G$ and $H$. The normal coordinates of $G+H$ coincide with the vector sum of the normal coordinates of $H$ and $G$, that is $\overrightarrow{G+H}=\vec{G}+\vec{H}$. It follows that the normal sum is an associative operation.

A regular intersection curve $\alpha$ is always orientation preserving [3]. The unique cut-and-paste operation along $\alpha$ which preserves the normal isotopy classes of the elementary disks is called a regular exchange. The normal sum $G+H$, as defined above, is the surface which results from performing a regular exchange along each (regular) intersection curve of $G \cap H$ and then straightening by a normal isotopy. The trace curves associated to the sum $G+H$ are the curves in the surface $G+H$ that result from the identification of the curves in $G$ and $H$ along which the cuts are made during the cut-and-paste operation.

It is sometimes useful to notice that regular exhanges are completely determined by the corresponding regular exchanges between the arcs in which the surfaces meet the 2 -simplices of $\mathscr{T}$. Since a regular exchange does not introduce new disk types, irregular intersection curves are never introduced during the process.

Consider an intersection curve $\alpha$ of two normal surfaces $G$ and $H$ and suppose one performs a cut-and-paste operation along $\alpha$ that is not a regular exchange. In this case there exists a tetrahedron $\Delta$ containing elementary disks $E^{\prime} \subset G, E^{\prime \prime} \subset H$ which intersect in an arc $E^{\prime} \cap E^{\prime \prime} \subset \alpha$ such that one of the disks produced by the cut-and-paste operation along $E^{\prime} \cap E^{\prime \prime}$ meets a 2 -face of $\Delta$ in an arc $\beta$ where $\partial \beta$


Fig. 2. Fold created by cut-and-paste along a irregular curve


Fig. 3. Branched surface constructed from a normal surface
is contained in a 1 -simplex. We call such an arc $\beta$ a fold. More generally, a surface $K$ that intersects $\mathscr{T}^{(2)}$ transversely is said to contain a fold if there exists a 2 -simplex $\sigma$ such that some component of $\sigma \cap K$ is an arc with both endpoints in a 1 -simplex of $\sigma$. Whenever a surface contains a fold there exists an isotopy removing the fold and decreasing the weight of the surface.

Given a pair of normal surfaces $G$ and $H$ which intersect transversely, a component $\alpha$ of $G \cap H$ is called an inessential intersection curve if $\alpha$ is the frontier of a disk in either $G$ or $H$. In certain cases one can eleminate all inessential intersection curves by regular exchanges which exchange pairs of innermost disks. A disk $D$ in $G$ with $\operatorname{fr}(D) \subset G \cap H$ is called an innermost disk if $D \cap H=\operatorname{fr}(D)$ and $\operatorname{fr}(D)$ is connected. A least weight innermost disk is an innermost disk which has minimal weight relative to all other innermost disks in $G$ and $H$.

## 3. Branched surfaces constructed from normal surfaces

Let $S^{\prime}$ be a compact, orientable, incompressible, $\partial$-incompressible, least weight normal surface carried by the interior of a face $C$ of $\mathscr{P}_{\mathscr{F}}$. We modify slightly the construction in [5] to from a branched surface from the normal surface $S^{\prime}$. Let $S=2 S^{\prime}$ be the boundary of a regular neighborhood of the two-sided surface $S^{\prime}$. The branched surface $\tilde{\mathscr{B}}_{S}$ is obtained by taking one disk of each type found in $S$, identifying all the edges of the same arc type in each 2 -simple of $\mathscr{T}$, and flattening appropriately. This flattening to achieve a generic branch locus may create some very thin sectors which lie within a small ragular neighborhood of the 2-skeleton. The surfaces carried by $\widetilde{\mathscr{B}}_{s}$ are precisely the normal surfaces carried by the face $C$. If $G$ is a two-sided normal surface carried by the interior of $C$ then the branched surface $\tilde{\mathscr{B}}_{G}$ constructed from $G$, using this procedure, is identical to $\tilde{\mathscr{B}}$, up to normal isotopy. There is a fibered neighborhood $\tilde{N}$ for $\tilde{\mathscr{B}}_{s}$ which intersects both the 1 -skeleton and 2 -skeleton in fibers and such that $\partial_{h}(\tilde{\mathcal{N}}) \subset S$.

The construction of the incompressible branched surface $\mathscr{B}=\mathscr{B}_{S}$ from $\tilde{\mathscr{B}}$ is accomplished by removing disks of contact from $\tilde{\mathscr{B}}$ as follows. Form the $I$-bundle
$\tilde{L}$ by cutting $\tilde{N}$ on the closure in $S$ of $S \cap \operatorname{int}(\tilde{N})$ and taking the fibers to agree with the fibers of $\tilde{N}$. A component $J$ of $\tilde{L}$ is trivial if every properly embedded curve in $\partial_{h} J$ bounds a disk or half-disk in $S$. Every trivial $J$ has the form $J=C \times I \subset \tilde{L}$, where $C$ is a punctured disk and there exist disks $E_{0}, E_{1}$ in $S$ such that $C \times 0 \subset E_{0}, C \times 1 \subset E_{1}, \partial E_{0}=\alpha \times 0 \subset \partial C \times 0$, and $\partial E_{1}=\alpha \times 1 \subset \partial C \times 1$. Define the fibered neighborhood $N$ of the branched surface $\mathscr{B}$ by $N=\tilde{N}$-interiors of fibers of trivial components of $\tilde{L}$. $\mathscr{B}$ carries $S$ with positive weights and, since $S$ is a least weight, incompressible, $\partial$-incompressible surface, it follows as in Lemma 4.6 [5] that $\mathscr{B}$ is a RIB (Reebless incompressible branched surface).

In our contex we do not allow $N$ to vary beyond a normal isotopy. Thus, if a surface $G$ is carried by $\mathscr{B}$ then $G$ is a normal surface carried by the face C. Although all surfaces carried by $C$ are carried by $\mathscr{\mathscr { B }}$, not all need be carried by $\mathscr{B}$ since it may not be possible to eliminate intersections with trivial components of $\tilde{L}$ by a normal isotopy.

Lemma 3.1. If $S$ is a compact, orientable, incompressible, $\partial$-incompressible least weight, normal surface then every normal surface carried by the carrier $C_{S}$ of $S$ is isotopic to one carried by the branched surface $\mathscr{B}_{S}$ obtained from $S$.

Proof. Suppose $G$ is a normal surface carried by $C_{S}$ and hence by $\mathscr{B}$. Assume that $G$ intersects $S$ transversely in such a way that the number of components of $S \cap G$ is minimal relative to normal isotopy. If $G$ is not carried by $\mathscr{B}$ then $G$ intersects a trivial component of $\tilde{L}$ and hence $S \cap G$ has an inessential component. It follows from [9] that there exists a pair of innermost disks $D \subset S$ and $E \subset G$ which are switched by a regular exchange along the curve $D \cap E$, which is the common frontier of $D$ and $E$. Replace $G$ by the isotopic surface $(G-E) \cup D$ and pull it off $S$ along $\mathrm{fr}(E)$. Repeating this construction, we eventually obtain a surface $G^{\prime}$ carried by $\mathscr{B}$ and isotopic to $G$.

Let $\mathscr{B}$ denote a branched surface constructed as above from a normal surface and assume that $\mathscr{B}$ has $s$ sectors. If $F$ is carried by $\mathscr{B}$ then we let $\overrightarrow{f=}\left(f_{1}, \cdots, f_{s}\right) \in \boldsymbol{R}^{s}$ denote the integral invariant measure associated to $F$ and write $F=\mathscr{B}(\vec{f})$. An invariant measure $\vec{w}=\left(w_{1}, \cdots, w_{s}\right)$ on $\mathscr{B}$ is a nonnegative solution to the system of branching equations $w_{i}+w_{j}=w_{k}$. The set $\mathscr{C}(\mathscr{B})$ of all invariant measures on $\mathscr{B}$ is called the cone of invariant measures and is the cone over $\mathscr{M}(\mathscr{B})$, the subset of $\mathscr{C}(\mathscr{B})$ satisfying the additional equation $\Sigma\left|w_{i}\right|=1$. The convex compact polyhedron $\mathscr{M}(\mathscr{B})$ is called the cell of invariant projective measures. We let $\vec{w}^{*}$ denote the projection of the invariant measere $\vec{w}$ into $\mathscr{M}(\mathscr{B})$.

The surface $F=\mathscr{B}(\vec{f})$ can be viewed as a normal surface and its normal coordinates $\vec{F}=\left(x_{1}, \cdots, x_{n}\right)$ are related to the invariant measure $\vec{f}$ by a linear function $\rho: \mathscr{C}(\mathscr{B}) \rightarrow \boldsymbol{R}^{n}$. Associated to each sector of $\mathscr{B}$ is a string of normal coordinates corresponding to the elementary disks comprising the sector. If $\vec{w} \in \mathscr{C}(\mathscr{B})$ define
$\rho(\vec{w})=\Sigma \rho\left(w_{i}\right)$ where $\rho\left(w_{i}\right)=\left(w_{i} m_{1}, \cdots, w_{i} m_{t}\right)$ and $m_{j}$ denotes the number of elementary disks of type $j$ in the $i$-th sector of $\mathscr{B}$. If the $i$-th sector is one of the thin sectors created by flattening then $\rho\left(w_{i}\right)=(0, \cdots, 0)$. Thus, if $\vec{f}=\left(f_{1}, \cdots, f_{s}\right)$ is integral then $\rho(\vec{f})=\vec{F}$ gives the normal surface coordinates of the normal surface $F=\mathscr{B}(\vec{f})$. There is also the affine map $\rho^{*}: \mathscr{M}(\mathscr{B}) \rightarrow \mathscr{P}_{\mathscr{F}}$ induced by $\rho$.

There are simple relationships among the three commonly used measurements on normal surfaces carried by the branched surface $\mathscr{B}$. If we let $\vec{f}=\left(f_{1}, \cdots, f_{s}\right)$ denote the invariant measure for $F=\mathscr{B}(\vec{f})$, we have the $\mathscr{B}$-complexity of $F$ defined by $\gamma_{g}(F)=\Sigma f_{i}$. Viewed as a normal surface with normal coordinates $\vec{F}=\left(x_{1}, \cdots, x_{n}\right)$, the complexity of $F$ is defined by $\mathrm{cp}(F)=\Sigma x_{i}$, which is just the total number of elementary disks contained in $F$. The weight of $F$ is the number of points in the intersection of $F$ with the 1 -skeleton of $\mathscr{T}$.

Lemma 3.2. (1) Let $p$ denote the minimum number of disk types occurring in any one sector of $\mathscr{B}$ and let $P$ denote the maximum number. Then $\operatorname{cp}(F) / P \leq \gamma_{\mathscr{B}}(F)$ $\leq \mathrm{cp}(F) / p$.
(2) Let $d$ denote the maximum number of 2 -simplices in $\mathscr{T}$ containing any single 1 -simplex. Then $\mathrm{cp}(F) \leq \frac{d}{3} \mathrm{wt}(F)$. If $F$ is a surface with $\chi(F) \leq 0$ and $\partial F=\emptyset$ we also heve $\operatorname{wt}(F) \leq \mathrm{cp}(F)$.

Proof.
(1) $\operatorname{cp}(F)=\Sigma \rho\left(f_{i}\right)=\Sigma p_{i} f_{i}$, where $p_{i}$ denotes the number of elementary disk types that occur in the $i$-th sector of $\mathscr{B}$. Since $p \leq p_{i} \leq P$ we have $p \gamma_{\mathscr{B}}(F)=p \Sigma f_{i} \leq \operatorname{cp}(F)$ $\leq P \Sigma f_{i}=P \gamma_{\overparen{B}}(F)$.
(2) Consider the cell decomposition of $F$ defined by the elementary disks and let $\delta_{3}, \delta_{4}$ denote the number of 3 -sided, 4 -sided, respectively, elementary disks in $F$. Then $\operatorname{cp}(F)=\delta_{3}+\delta_{4}$ and $\mathrm{wt}(F) \geq\left(3 \delta_{3}+4 \delta_{4}\right) / d \geq 3 \mathrm{cp}(F) / d$. For the second inequality, compute the Euler characteristic $\chi(F)$ from the induced cell decomposition to obtain the equation $\mathrm{wt}(F)=\operatorname{cp}(F)+\chi(F)-\delta_{3} / 2+b / 2$, where $b$ denotes the number of edges of the elementary disks in $F$ which lie in $\partial M$. Thus $\operatorname{wt}(F) \leq \mathrm{cp}(F)$ when $b=0$ and $\chi(F) \leq 0$.

## 4. Least weight incompressible surfaces

Suppose $F$ is a compact, orientable, incompressible, $\partial$-incompressible, least weight, normal surface in $M$ and $C_{F}$ is the carrier of $F$ in $\mathscr{P}_{\mathscr{F}}$. It is known that all normal surfaces carried by $C_{F}$ are incompressible [2] and $\partial$-incompressible [4]. In this section we apply [6] to show that such surfaces are also least weight. This leads to the existence of a complete lw-face $C$ carrying $F$ which is maximal in the sense that if a normal surface $G$ is carried by $C$ then every least weight normal surface isotopic to $G$ is also carried by $C$.

Lemma 4.1. Suppose $G$ and $H$ are compact, orientable, incompressible, д-incompressible normal surfaces in $M$ such that the normal sum $F=G+H$ is a compact, orientable, incompressible, $\partial$-incompressible, least weight surface. If $K$ and $L$ are normal surfaces isotopic to and having no greater weight than $G$ and $H$, respectively, then every intersection curve of $K \cap L$ is regular and the normal sum $K+L$ is isotopic to $F$.

Proof. Assume that $G$ and $H$ intersect transversely. The intersection $G \cap H$ is said to be simplified if
(1) $G \cap H$ contains no inessential intersection curves,
(2) there are no products in $M$ between surfaces in $G$ and $H$,
(3) there are no products in $\partial M$ between arcs of $\partial G$ and $\partial H$.

Step A. We construct surfaces $G_{4}$ and $H_{4}$ isotopic to $G$ and $H$, respectively, such that the intersection $G_{4} \cap H_{4}$ is simplified and a double-curve sum $G_{4} \triangleright \square H_{4}$ (not necessarily a normal sum) is isotopic to $F$. (In fact, $G_{4} \triangleright \triangleleft H_{4}$ will coincide with the normal sum $G+H$ outside a collar neighborhood of $\partial M$, but this is not used.)
(1) If there exists a contractible simple closed curve in $G \cap H$ then (from the proof of Lemma 2.1 in [2]) there exists a pair of innermost disks of equal weight which are switched by a regular exchange along their common boundary. We may assume that all such exchanges have been made and that we have $F=G_{1}+H_{1}$, where $G_{1}, H_{1}$ are isotopic to and have the same weight as $G$, $H$, respectively, and there are no contractible simple closed curves in $G_{1} \cap H_{1}$. Similarly, if there exists an arc component of $G_{1} \cap H_{1}$ that is inessential then (from the proof of Lemma 6.6 in [4]) there exists a pair of innermost disks of equal weight switched by a regular exchange along the arc forming their common frontier. Therefore we can write $F=G_{2}+H_{2}$, where $G_{2}, H_{2}$ are isotopic to have the same weight as $G, H$, respectively, and there are no inessential intersection curves in $G_{2} \cap H_{2}$.
(2) Suppose there exists a product in $M$ between subsurfaces of $G_{2}$ and $H_{2}$. Then there exists a compact connected surface $W$ and a compact 1 -submanifold $\gamma$ of $\partial W$ such that if we let $P=(W \times I) / \sim$ denote the product with intervals $x \times 1$ collapsed to points for each $x \in \gamma$, then $(P,(\partial W-\stackrel{\circ}{\gamma}) \times I)$ is embedded in $(M, \partial M)$ with $W \times 0 \subset G_{2}, W \times 1 \subset H_{2}$ and $P \cap\left(G_{2} \cup H_{2}\right)=W \times 0 \cup W \times 1$. Consider the set of cut-and-paste operations along the intersection curves $\gamma / \sim$ which result in the switching of $W \times 0$ and $W \times 1$. We will show that these particular cut-and-paste operations are regular exchanges along each component of $\gamma / \sim$ and thus coincide with what occurs along $\gamma / \sim$ when forming the normal $\operatorname{sum} F=G_{2}+H_{2}$. It will then follow that after performing regular exchanges along all such products we can write $F=G_{3}+H_{3}$ where $G_{3}, H_{3}$ are isotopic (by isotopies across the products) to $G, H$, respectively, there are no inessential intersection curves in $G_{3} \cap H_{3}$, and no products between subsurfaces of $G_{3}$ and $H_{3}$.

Suppose there exists a component $\alpha$ of $\gamma / \sim$ along which the above cut-and-paste
operation switching $W \times 0$ and $W \times 1$ is an irregular exchange. We will show that this assumption leads to a contradiction.

First consider the case there exists an essential arc $\lambda$ in $W$ with endpoints in $\alpha$, where $\lambda$ is essential in the sense that it does not cobound a disk in $W$ with a subarc of $\alpha$. Since switching $W \times 0$ and $W \times 1$ is associated with an irregular exchange along $\alpha$ and in the construction of $F$ as the normal sum $G_{2}+H_{2}$ a regular exchange along $\alpha$ is used, it follows that $D=(\lambda \times I) / \sim$ is a compression disk for $F$. Since $F$ is incompressible, $\partial D$ bounds a disk $D^{\prime} \subset \stackrel{\circ}{F}$. Now $D^{\prime} \cap(W \times 0)$ is a disk-with-holes and must contain a component of $G_{2} \cap H_{2}-\alpha$ since $\lambda$ is essential in $W$. This is impossible since $G_{2} \cap H_{2}$ has no contractible simple closed curve components. Therefore there does not exist an essential an essential arc $\lambda$ in $W$ with endpoints in a component of $\gamma / \sim$ along which the cut-and-paste operation switching $W \times 0$ and $W \times 1$ is an irregular exchange.

It follows that either $W$ is a disk or $W$ is an annulus where each component of $\gamma / \sim$ along which the cut-and-paste operation switching $W \times 0$ and $W \times 1$ is an irregular exchange must be a component of $\partial W$. Since we have already eliminated inessential intersection curves, it follows that when $W$ is a disk, each components of $\gamma / \sim$ is an arc and there must be at least two components.

Suppore that $W$ is a disk and $\gamma / \sim$ has three or more components. Let $\beta$, $\delta$ denote the two components of $\gamma / \sim$ which are adjacent to $\alpha$ along $\partial W \times 0$. Since $F$ cannot have any contractible boundary components, the cut-and-paste operation along $\gamma / \sim$ switching $W \times 0$ and $W \times 1$ is a regular exchange along each of $\beta$ and $\delta$. Choose an arc $\lambda$ in $W$ with one endpoint in $\alpha$ and the other in $W \cap \partial M$ such that $\beta$ and $\delta$ lie in different components of $W-\lambda$. The disk $D=(\lambda \times I) / \sim$ is a $\partial$-compression disk for $F$. Since $F$ is $\partial$-incompressible there exists a disk $D^{\prime} \subset F$ with $\operatorname{fr}\left(D^{\prime}\right)=D \cap F$. Then $D^{\prime}$ contains a trace curve of either $\beta$ or $\delta$, say $\beta^{\prime}$ of $\beta$. There exists a disk $E^{\prime} \subset D^{\prime}$ such that $\partial E^{\prime}=\beta \cup\left(E^{\prime} \cap \partial M\right)$. Now $E^{\prime}$ cannot contain a simple closed curve trace curve of $G_{2}+H_{2}$. Thus there exists an outermost trace curve $\gamma^{\prime}$ in $E^{\prime}$ corresponding to an inessential intersection curve $\gamma$ in $G_{2} \cap H_{2}$, which is impossible.

Suppose that $W$ is a disk and $\gamma / \sim$ has just two components, $\alpha$ and a second arc $\beta$. As before, the cut-and-paste operation along $\gamma / \sim$ switching $W \times 0$ and $W \times 1$ is a regular exchanges along $\beta$. There exists a 0 -weight disk $B$ containing $\beta$ such that $B \cap F$ consists of two disjoint trace curves $\beta^{\prime}, \beta^{\prime \prime}$ in $\partial B$ parallel to $\beta$ and $B \cap \partial M=\overline{\partial B}-\left(\beta^{\prime} \cup \beta^{\prime \prime}\right)$. Splitting $F$ along $B \cap F$ and capping with two copies of $B$ produces a compressible annulus component $A$ and a surface $F^{\prime}$ isotopic to $F$. Described another way, the disjoint union of $A$ and $F^{\prime}$ is be obtained from $G_{2} \cup H_{2}$ by performing regular exchanges along all intersection curves except $\beta$ and an irregular exchange along $\beta$. Since the annulus component $A$ has nonzero weight, it follows that the surface $F^{\prime}$, which is isotopic to $F$, has less weight than $F$. This contradicts the assumption that $F$ is least weight.

Suppose that $W$ is an annulus and $\partial W=\gamma / \sim$, where $\partial W$ consists of the two
simple closed curves $\alpha$ and $\beta$. The cut-and-paste operation switching $W \times 0$ and $W \times 1$ must be an regular exchange along $\beta$ for otherwise $F$ would have a compressible torus component bounding a solid torus approximated by $W \times I / \sim$. Thus, if we perform regular exchanges along all intersection curves except $\beta$ and an irregular exchange along $\beta$, we obtain from $G_{2} \cup \mathrm{H}_{2}$ an compressible torus $A^{\prime}$ with positive weight and a surface $F^{\prime}$ isotopic to $F$. But then $F^{\prime}$ has less weight than $F$, a contradiction.

Finally suppose that $W$ is an annulus and $\partial W \neq \gamma / \sim$. Then there exists an arc $\lambda$ in $W$ with one endpoint in $\alpha$ and the other endpoint in $W \cap \partial M$. The disk $D=(\lambda \times I) / \sim$ is an essential $\partial$-compression disk for $F$ since $\lambda$ does not separate $W \subset F$. Again this is impossible.

Thus the above cut-and-paste operation switching $W \times 0$ and $W \times 1$ corresponds to regular exchanges along each component of $\gamma / \sim$. It follows that after performing regular exchanges along all such products we can write $F=G_{3}+H_{3}$ where $G_{3}, H_{3}$ are isotopic to $G, H$, respectively, all intersection curves in $G_{3} \cap H_{3}$ are essential, and there does not exist any products between subsurfaces of $G_{3}$ and $H_{3}$.
(3) Suppose there exists a product $R=(\mu+I) / \sim$ in $\partial M$ such that the arc $\mu \times 0 \subset G_{3}$, the arc $\mu \times 1 \subset H_{3}$ and $R \cap\left(G_{3} \cup H_{3}\right)=\partial R$. Let $\alpha$ and $\beta$ denote the intersection curves of $G_{3} \cap H_{3}$ which have an endpoint in $\partial \mu \times 0$. We will first show that a regular exchange along $\alpha \cup \beta$ interchanges $\mu \times 0$ and $\mu \times 1$. Suppose that $\alpha=\beta$. A regular exchange along $\alpha$ must either switch $\mu \times 0$ and $\mu \times 1$ or form a contractible boundary component of $F$ from $\mu \times 0 \cup \mu \times 1$. Since $F$ is incompressible, the arcs $\mu \times 0$ and $\mu \times 1$ are switched.

Suppose that $\alpha \neq \beta$. There are three possibilities to consider: (i) irregular exchanges along both $\alpha$ and $\beta$ exchange $\mu \times 0$ and $\mu \times 1$, (ii) an irregular exchange along one, say $\alpha$, and a regular exchange along $\beta$ exchanges $\mu \times 0$ and $\mu \times 1$, (iii) regular exchange along both $\alpha$ and $\beta$ exchange $\mu \times 0$ and $\mu \times 1$. As before, (i) can be ruled out since otherwise a contractible boundary component of $F$ would result.

Suppose we have (ii). Let $r$ be the arc in $\partial F$ arising from the cut-and-paste operation on $\mu \times 0$ and $\mu \times 1$ produced by regular exchanges along $\alpha$ and $\beta$. There exists a 0 -weight disk $B$ containing $\beta$ such that $\partial B=a \cup b \cup \beta^{\prime} \cup \beta^{\prime \prime}$, where $\beta^{\prime}, \beta^{\prime \prime}$ are the two trace curves $B \cap F$ and $B \cap \partial M=a \cup b$ with notation chosen such that $\partial a=\partial r$. Then $a \cup r$ bounds a disk $R^{\prime} \subset R$. Fix a collar neighborhood $\partial M \times[0,1]$ of $\partial M=\partial M \times 0$ such that $B \cap \partial M \times[0,1]=(a \cup b) \times[0,1]$ and $F \cap(\partial M \times[0,1])$ $=\partial F \times[0,1]$. The disk $D=R^{\prime} \times 1 \cup(B-(a \times[0,1))$ is a $\partial$-compression disk for $F$. Thus there exists a disk $D^{\prime} \subset F$ with $\partial D^{\prime}=x \cup y$, where $x=D \cap F$ and $y=D^{\prime} \cap \partial M$.

If $r \times[0,1] \subset D^{\prime}$ then let $E$ be the closure of the component of $D^{\prime}-r \times[0,1]$ for which $\beta^{\prime} \subset E$. We then have $\partial E=\beta^{\prime} \cup(E \cap \partial M)$. If the disk $E \subset F$ does not contain any trace curves other than $\beta^{\prime}$ then $\beta$ would be an inessential intersection curve. If there are other trace curves in $F$ (relative to the sum $F=G_{3}+H_{3}$ ) then an outtermost one relative to $\beta^{\prime}$ corresponds to an inessential intersection
curve. Since $G_{3} \cap H_{3}$ contains no inessential curves, this cannot occur. Thus $r \times[0,1] \not \subset D^{\prime} . \quad$ In this case $E=D^{\prime} \cup(r \times[0,1])$ is a disk such that $E \cap B=\beta^{\prime} \cup \beta^{\prime \prime}$. We then have the compressible annulus $B \cup E$ and the surface $F^{\prime}=(F-E) \cup B$ is isotopic to $F$. But this is impossible since $\mathrm{wt}\left(F^{\prime}\right)<\mathrm{wt}(F)$. We are left only with (iii) as the only possibility.

We have shown that regular exchanges along both $\alpha$ and $\beta$ exchange $\mu \times 0$ and $\mu \times 1$. Use an isotopy with support in a neighborhood of $R$ to move $G_{3}$ to a surface $G_{3}^{\prime}$ by pulling $\mu \times 0$ across $R$ to the other side of $\mu \times 1$. That is, we want $G_{3}^{\prime} \cup H_{3}$ to coincide with $G_{3} \cup H_{3}$ outside $R \times[0,1]$ and, in effect, the endpoints of $\alpha, \beta$ in $R$ are joined and pushed into $\stackrel{\circ}{M}$. The sum $G_{3}^{\prime} \triangleright \triangleleft H_{3}$ obtained by cut-and-paste operations agreeing with those of the normal sum outside a neighborhood of $R$ is clearly isotopic to $F$ and no inessential intersection curves have been introduced in $G_{3}^{\prime} \cap H_{3}$ since products between subsurfaces have already been eliminated. Repeating this construction, we eventually obtain surfaces $G_{4}$, $H_{4}$ isotopic to $G, H$, respectively, such that a double-curve sum $G_{4} \triangleright \triangleleft H_{4}$ is isotopic to $F$ and the intersection $G_{4} \cap H_{4}$ is simplified and $G_{4} \triangleright \triangleleft H_{4}$ coincides with the normal sum $G_{3}+H_{3}$ off $\partial M \times[0,1]$.

Step B. The next step is to find surfaces $K^{\prime}, L^{\prime}$, isotopic to and having no greater weight than $K, L$, respectively, such that the intersection $K^{\prime} \cap L^{\prime}$ is almost simplified. The products in $\partial M$ cannot be eliminated without the possibility of increasing weight.
(1) If there exists an inessential intersection curve in $K \cap L$ then we can find an innermost disk $D$ in either $K$ or $L$ such that $\operatorname{fr}(D)$ is a component of $K \cap L$ and $D$ is least weight among all such innermost disks in $K$ and $L$. We suppose $D \subset K$ as the argument is the same for $D \subset L$. There exists a disk $E \subset L$ such that $D \cap L=\mathrm{fr}(E)$. The disks $D$ and $E$ are parallel so we can isotope $L$ to $(L-E) \cup D$ and then pull it off from $K$ along $\operatorname{fr}(D)$. Since $\operatorname{wt}(D) \leq \mathrm{wt}(E)$ we do not increase weight by this construction and we have reduced the number of intersection curves. Repeating this construction we eventually obtain surfaces $K_{1}, L_{1}$ isotopic to and having no greater weight than $K, L$, respectively, and such that all intersection curves in $K_{1} \cap L_{1}$ are essential.
(2) Suppose there exists a product $(P,(\partial W-\gamma) \times I)$ embedded in $(M, \partial M)$ with $W \times 0 \subset K_{1}$ and $W \times 1 \subset L_{1}$. Choose notation such that $\mathrm{wt}(W \times 0) \leq \mathrm{wt}(W \times 1)$. Consider the cut-and-paste operation along the intersection curves $\gamma / \sim$ on $L_{1}$ which replaces $W \times 1$ by $W \times 0$. After an isotopy with support in a neighborhood of $W \times 0$, one obtains isotopic surfaces of no greater weight and the number of intersection curves is decreased. Repeating such constructions we eventually obtain the desired surfaces $K^{\prime}$ and $L^{\prime}$.

Step C. We now associate with $K^{\prime}, L^{\prime}$ isotopic surfaces $K^{\prime \prime}, L^{\prime \prime}$ which have simplified intersections. The surfaces $K^{\prime \prime}, L^{\prime \prime}$ are constructed from $K^{\prime}, L^{\prime}$ by isotopies which remove product regions in $\partial M$ in a controlled manner. Assume that we have already removed $i$ product regions and in the process we have constructed
the surfaces $K_{i}^{\prime}, L_{i}^{\prime}$. Suppose there remains a product $R_{i}=\left(\mu_{i} \times I\right) / \sim$ in $\partial M$ with $\mu_{i} \times 0 \subset K_{i}^{\prime}, \mu_{i} \times 1 \subset L_{i}^{\prime}$ and $R_{i} \cap\left(K_{i}^{\prime} \cup L_{i}^{\prime}\right)=\partial R_{i} . \quad$ Let $\alpha_{i}$ and $\beta_{i}$ denote the intersection curves of $K_{i}^{\prime} \cap L_{i}^{\prime}$ which have an endpoint in $\partial \mu_{i} \times 0$. Form $K_{i+1}^{\prime}, L_{i+1}^{\prime}$ by using an isotopy with support in a small neighborhood of $R_{i}$ that slides each of $\mu_{i} \times 0$ and $\mu_{i} \times 1$ across to the opposite side. Let $\gamma_{i}$ denote the new intersection curve resulting from the ends of $\alpha_{i+1}$ and $\beta_{i+1}$ being joined after this isotopy. Continuing in this way, we eventually obtain surfaces $K^{\prime \prime}, L^{\prime \prime}$, which are isotopic to $G, H$, have simplified intersection $K^{\prime \prime} \cap L^{\prime \prime}$, and coincide with $K^{\prime}, L^{\prime}$ off $\partial M \times[0,1]$.

Step D. At this stage we have the two pairs of surfaces $G_{4}, H_{4}$ and $K^{\prime \prime}, L^{\prime \prime}$ such that (a) the intersections $G_{4} \cap H_{4}$ and $K^{\prime \prime} \cap L^{\prime \prime}$ are simplified, (b) $K^{\prime \prime}$ and $G_{4}$ are isotopic to $G$, (c) $L^{\prime \prime}$ and $H_{4}$ are isotopic to $H$, (d) a double-curve sum $G_{4} \triangleright \square H_{4}$ is isotopic to $F$. By Theorem 3 [6] there exists a double-curve sum $F^{\prime \prime}=K^{\prime \prime} \triangleright \triangleleft L^{\prime \prime}$ which is isotopic to $F$. We will show that there is a corresponding double-curve sum $K^{\prime} \triangleright L^{\prime}$ which is isotopic to $F$. It is sufficient to show that for each of the intersection curves $\gamma_{i}$ formed in Step C, the cut-and-paste operation involved in $K^{\prime \prime} \triangleright \triangleleft L^{\prime \prime}$ along $\gamma_{i}$ agrees outside of $\partial M \times[0,1]$ with the cut-and-paste operation of $K^{\prime}$ and $L^{\prime}$ along $\alpha_{i}, \beta_{i}$ which interchanges $\mu_{i} \times 0$ and $\mu_{i} \times 1$.

Consider a product region $R_{i}=\mu_{i} \times I / \sim$ and suppose that we have already established that the two cut-and-paste operations agree near each product region interior to $R_{i}$. Thus, performing the cut-and-paste operations of the sum $K^{\prime \prime} \triangleright \triangleleft L^{\prime \prime}$ in the interior of $R_{i}$ has the same result as an isotopy pulling the curves $K^{\prime \prime} \cap \AA_{i}$ apart from the curves $L^{\prime \prime} \cap \stackrel{\circ}{R}_{i}$. There exists a disk $D$ such that $\partial D=x \cup y \cup z$ where $x, y, z$ are arcs with disjoint interiors, $D \cap \partial M=D \cap R_{i}=x, D \cap K^{\prime \prime}=y$, $D \cap L^{\prime \prime}=z$, and $y \cap z \in \gamma_{i}$. In performing the sum $K^{\prime \prime} \triangleright \triangle L^{\prime \prime}$, there are just two possibilities for the cut-and-paste operation along $\gamma_{i}$. One agrees outside of $\partial M \times[0,1]$ with the cut-and-paste operation on $K^{\prime}$ and $L^{\prime}$ along the intersection curves $\alpha_{i}, \beta_{i}$ in $K^{\prime} \cap L^{\prime}$ which interchanges $\mu_{i} \times 0$ and $\mu_{i} \times 1$. The other results in $D$ becoming a $\partial$-compression disk $D^{\prime}$ for $F^{\prime \prime}=K^{\prime \prime} \triangleright \triangleleft L^{\prime \prime}$. In the latter case, since $F^{\prime \prime}$ is $\partial$-incompressible, there exists a disk $D^{\prime \prime} \subset F^{\prime \prime}$ with $\operatorname{fr}\left(D^{\prime \prime}\right)=D^{\prime} \cap F^{\prime \prime}$. As in Step (A), one can show that the disk $D^{\prime \prime} \subset F^{\prime \prime}$ leads to the existence of inessential intersection curves in $K^{\prime \prime} \cap L^{\prime \prime}$. Since this is impossible, the only alternative is that the cut-and-paste operation involved in $K^{\prime \prime} \triangleright \triangleleft L^{\prime \prime}$ along $\gamma_{i}$ agrees outside of $\partial M \times[0,1]$ with the cut-and-paste operations along $\alpha_{i}, \beta_{i}$ which interchanges $\mu_{i} \times 0$ and $\mu_{i} \times 1$.

Step E. From Step D it follows that there is a double-curve sum $F^{\prime}=K^{\prime} \triangleright \triangleleft L^{\prime}$ isotopic to $F$ such that $\mathrm{wt}\left(K^{\prime}\right) \leq \mathrm{wt}(K)$ and $\mathrm{wt}\left(L^{\prime}\right) \leq \mathrm{wt}(L)$. Hence $\mathrm{wt}\left(F^{\prime}\right)$ $=\mathrm{wt}\left(K^{\prime}\right)+\mathrm{wt}\left(L^{\prime}\right) \leq \mathrm{wt}(K)+\mathrm{wt}(L) \leq \mathrm{wt}(G)+\mathrm{wt}(H)=\mathrm{wt}(F)$. But $F$ is least weight and hence $\mathrm{wt}(F)=\mathrm{wt}\left(F^{\prime}\right)$. It follows that $\mathrm{wt}(G)=\mathrm{wt}(K)$ and $\mathrm{wt}(H)=\mathrm{wt}(L)$. Because $K$ and $L$ are arbitrary surfaces isotopic to and having no greater weight than $G$ and $H$, respectively, we can conclude that both $G$ and $H$ are least weight surfaces. At this point in the proof we are also able to cunclude that $K$ and $L$ are least weight
surfaces.
Step F. It remains to show that every intersection curve of $K \cap L$ is regular and the normal sum $K+L$ is isotopic to $F$. We return to Step B and start over using the new information (from Step E) that $K$ and $L$ are both least weight.
(1) If there exists inessential intersection curves in $K \cap L$ then it follows from Proposition 1 of [9] or Proposition 3.4 in [3] that regular exchanges along the inessential intersection curves yieled surfaces $K^{\prime}, L^{\prime}$ isotopic to and of the same weight as $K, L$, respectively, and such that all intersection curves of $K^{\prime} \cap L^{\prime}$ are essential.
(2) Suppose there exists a product $(P,(\partial W-\dot{\gamma}) \times I)$ embedded in $(M, \partial M)$ with $W \times 0 \subset K^{\prime}$ and $W \times 1 \subset L^{\prime}$. Since both $K^{\prime}, L^{\prime}$ are least weight it follows that $\mathrm{wt}(W \times 0)=\mathrm{wt}(W \times 1)$. The cut-and-paste operation along all the intersection curves $\gamma / \sim$ on $K$ and $L$ which interchanges $W \times 0$ and $W \times 1$ does not change the weight and decreases the number of intersection curves. By repeating this construction we eventually produce surfaces $K^{\prime \prime}$ and $L^{\prime \prime}$ isotopic to having the same weight as $K$ and $L$, respectively, such that all intersection curves of $K^{\prime \prime} \cap L^{\prime \prime}$ are essential and there do not exist any products between surfaces of $K^{\prime \prime}$ and $L^{\prime \prime}$.
(3) By Steps (C) and $(D)$ there is a double-curve sum $F^{\prime \prime}=K^{\prime \prime} \triangleright \triangleleft L^{\prime \prime}$ isotopic to $F$. If any of the cut-and-paste operations used are not regular exchanges then a fold exists in the sum $K^{\prime \prime} \triangleright L^{\prime \prime}$ (see Lemma 1 [8]) and the weight can be reduced. But this is impossible since $F$ is already least weight. Thus every intersection curve in $K^{\prime \prime} \triangleright \triangleleft L^{\prime \prime}$ is regular and the double-curve sum $K^{\prime \prime} \triangleright \triangleleft L^{\prime \prime}$ is a normal sum. Since $K^{\prime \prime} \cup L^{\prime \prime}$ was obtained from $K \cup L$ by a sequence of regular exchanges, it follows that $K+L=K^{\prime \prime}+L^{\prime \prime}$. Therefore, the normal sum $K+L$ is defined and is isotopic to $F$.

For a surface $F$ in $M$ let $N(F)$ denote a regular neighborhood of $F$. We say that $F$ is injective if $\partial N(F)$ is incompressible and $\partial$-injective if $\partial N(F)$ is $\partial$-incompressible. Thus an orientable surface is injective ( $\partial$-injective) if and only if it is incompressible ( $\partial$-incompressible).

Theorem 4.2. Let $F$ be a compact, orientable, least weight, incompressible, $\partial$-incompressible, normal surface in the compact, irreducible, $\partial$-irreducible, orientable 3-manifold $M$. Then every normal surface carried by $C_{F}$ (the carrier of $F$ ) is injective, $\partial$-injective, and least weight.

Proof. Let $G$ be any normal surface carried by $C_{F}$. There exists a surface $H$ also carried by $C_{F}$ such that $m F=G+H$ for some positive integer $m$. If either $G$ or $H$ is nonorientable we can consider $2 m F=2 G+2 H$, where both $2 G$ and $2 H$ are orientable boundaries of regular neighborhoods of $G$ and $H$, respectively. Thus, we may assume that both $G$ and $H$ are orientable. It follows that $G$ and $H$ are incompressible by Theorem 2.2 in [2] and $\partial$-incompressible by

Theorem 6.5 in [4].
Consider a normal surface $K$ isotopic to $G$ and suppose that $\mathrm{wt}(K)<\mathrm{wt}(G)$. By Lemma 4.1 it follows that $K+H$ is defined and isotopic to $m F$. But $\mathrm{wt}(K+H)=\mathrm{wt}(K)+\mathrm{wt}(H)<\mathrm{wt}(G)+\mathrm{wt}(H)=\mathrm{wt}(m F)$, contradicting that $m F$ is least weight. Therefore $G$ is also a least weight surface.

Corollary 4.3. Let $G, H$ be orientable normal surfaces such that the normal sum $G+H$ is defined and is a compact, orientable, incompressible, $\partial$-incompressible, least weight surface. If $G^{\prime}$ and $H^{\prime}$ are least weight normal surfaces such that $G^{\prime}$ is isotopic to $G$ and $H^{\prime}$ is isotopic to $H$ then the normal sum $G^{\prime}+H^{\prime}$ is defined and is isotopic to $G+H$.

Proof. By Theorem 4.2, both $G$ and $H$ are compact, orientable, incompressible, $\partial$-incompressible, least weight surfaces. Hence the conclusion follows directly from Lemma 4.1.

Corollary 4.4. Let $F$ be a compact, orientable, incompressible, $\partial$-incompressible, surface. Suppose $G_{1}, G_{2}$ are least weight normal surfaces isotopic to $n_{1} F, n_{2} F$, respectively, where $n_{1}, n_{2}$ are positive integers. Then every intersection curve in $G_{1} \cap G_{2}$ is regular and $G_{1}+G_{2}$ is a least weight surface isotopic to $\left(n_{1}+n_{2}\right) F$.

Recall that a face $C$ of $\mathscr{P}_{\mathscr{T}}$ is an $l w$-face if every normal surface carried by $C$ is incompressible, $\partial$-incompressible and least weight. A complete $l w$-face of $\mathscr{P}_{\mathscr{F}}$ is an lw-face $C$ with the additional property that if $G$ is any normal surface carried by $C$ then every least weight surface isotopic to $G$ is also carried by $C$.

Theorem 4.5. Let $M$ be a compact, orientable, irreducible, $\partial$-irreducible, 3-manifold. There exists a finite collection of complete lw-faces of $\mathscr{P}_{\mathscr{F}}$ carrying compact, orientable, incompressible, $\partial$-incompressible, least weight, normal surface in M.

Proof. Suppose $F$ is a compact, orientable, incompressible, $\partial$-incompressible, least weight surface. Let $F_{1}, \cdots, F_{k}$ be a set of normal surfaces representing all normal isotopy classes of least weight normal surfaces isotopic to $F$. By Corollary 4.4, the normal sum $H=F_{1}+\cdots+F_{k}$ is defined and is a least weight normal surface isotopic to $k F$. It follows from Lemma 4.2 that the carrier of $H$, which we denote by $C_{1}$, is an lw-face. Let $G_{1}$ be an orientable normal surface carried by $\dot{C}_{1}$ and suppose that $K_{1}$ is isotopic to $G_{1}$ and least weight. Then $H_{1}=G_{1}+K_{1}$ is least weight and isotopic to $2 G_{1}$. Moreover, if $C_{2}$ is the carrier of $H_{1}$ then $C_{2}$ is an lw-face and $C_{1} \subset C_{2}$. Suppose we have constructed the sequence $C_{1} \subset C_{2} \subset \cdots$ $\subset C_{i}$ of lw -faces. If there exists a compact, orientable, incompressible, $\partial$ incompressible, least weight surface $K_{i}$ isotopic to a surface $G_{i}$ carried by $\mathscr{C}_{i}$ then


Fig. 4. Branched surface formed from $S \cup A \cup B$ in $M=S \times S^{1}$
set $H_{i}=G_{i}+K_{i}, \quad H_{i}$ is a least weight surface isotopic to $2 K_{i}$. The carrier $C_{i+1}$ of $H_{i}$ contains $C_{i}$ and is an lw-face. Since $\mathscr{P}_{\mathscr{F}}$ has only a finite number of faces, we eventually obtain an lw-face $C$ with the property that if $G$ is carried by $C$ then every least weight normal surface isotopic to $G$ is also carried by $C$.

To finish, consider a normal surface $G$ carried by the boundary of $C . G$ is a least weight surface and we can find least weight orientable normal surfaces $H$ and $X$ carried by $\mathcal{C}$ such that $H=2 G+X$. Suppose $K$ is a least weight normal surface isotopic to $2 G$. By Lemma 4.1 the sum $K+X$ is defined and isotopic to $H$. Since $\mathrm{wt}(K+X)=\mathrm{wt}(H)$ it follows that $K+X$ is least weight and hence carried by $C$. Therefore $K$ is also carried by $C$, proving that $C$ is a complete $l w$-face.

Corollary 4.6. Let $C_{1}, C_{2}$ be complete $l w$-faces of $\mathscr{P}_{\mathscr{g}}$. Suppose $G$ is an orientable normal surface carried by $C_{1}$ and $G$ is isotopic to a surface carried by $C_{2}$. Then $G$ is carried by $C_{1} \cap C_{2}$.

## 5. Projective isotopy classes

The following example illustrates the PIC-partitioning in a simple case. This example is essentially Example 1.5 in [5] presented from our point of view.

Example. Let $M=S \times S^{1}$ where $S$ is a closed orientable surface of genus 3 and $a, b$ are simple closed curves in $S$ as shown in Figure 4. $A=a \times S^{1}$ and $B=b \times S^{1}$ are non-isotopic vertical tori in $M$. We let $S=S \times *$, where $*$ is a point in $S^{1}$. The branched surface $\mathscr{B}$ is formed from $S \cup A \cup B$ with six sectors labeled $\mathscr{E}_{1}, \cdots, \mathscr{E}_{6}$ as shown schematically in the figure. Thus $A=\mathscr{B}(\vec{a})$ with $\vec{a}=(0,0,1,0,1,0)$ and $B=\mathscr{B}(\vec{b})$ with $\vec{b}=(0,0,0,1,0,1)$. For each positive integer $n$ we have the pair
of isotopic surfaces $F_{n}=\mathscr{B}(1,1, n-1,0, n, 1)$ and $G_{n}=\mathscr{B}(1,1,0, n-1,1, n)$. The isotopic between $F_{n}$ and $G_{n}$ is obtained by repeating $n-1$ times the simple isotopy associated to $\vec{p}=(1,0,0,1,0,1)-(1,0,1,0,1,0)=(0,0,-1,1,-1,1)$. (Simple isotopies are described in the proof of Lemma 5.2.) Observe that we can write $F_{n}=C+(n-1) A$ and $G_{n}=C+(n-1) B$ where $C=\mathscr{B}(1,1,0,0,1,1)$. Looking at the projective classes in $\mathscr{M}(\mathscr{B})$ we see that as $n \rightarrow \infty$ we have $\vec{F}_{n}{ }^{*}=(1,1, n-1,0, n, 1)^{*} \rightarrow \vec{a}^{*}$ and $\vec{G}_{n}{ }^{*}=(1,1,0, n-1,1, n)^{*}$ $\rightarrow b^{*}$. Although $F_{n}$ is isotopic to $G_{n}$, the torus $\mathscr{B}(\vec{a})$ carried by $\vec{a}^{*}$ is not in the same projective isotopy class as the torus $\mathscr{B}(\vec{b})$ carried by $\vec{b}^{*}$.

From the branching equations $x_{1}+x_{3}=x_{5}, x_{1}+x_{4}=x_{6}, x_{2}+x_{3}=x_{5}$, and $x_{2}+x_{4}=x_{6}$ one computes that the cell of invariant projective measures $\mathscr{M}(\mathscr{B})$ is a 2 -simplex with vertices $\vec{a}^{*}, \vec{b}^{*}$ and $\vec{c}^{*}$. Let $\mu: \mathscr{M}(\mathscr{B}) \rightarrow[0,1]$ be the linear map defined by $\mu\left(\vec{a}^{*}\right)=\mu\left(\vec{b}^{*}\right)=0$ and $\left.\mu \vec{c}^{*}\right)=1$. For each $t>0$, every surface carried by the fiber $\mu^{-1}(t)$ belongs to the some projective isotopy class. Surfaces carried by different fibers do not belong to the same projective isotopy classes. However, in the fiber $\mu^{-1}(0)$, which is a proper face of $\mathscr{M}(\mathscr{B})$, we find that distinct rational points correspond to distinct projective isotopy classes.

To view this example in the context of normal surfaces, triangulate $M$ such that $H=F_{2}+G_{2}=\mathscr{B}(2,2,1,1,3,3)$ is a least weight normal surface $\mathscr{B}_{H}$ and the branched surface $\mathscr{B}_{H}$ as constructed in section 3 is the branched surface $\mathscr{B}$ described above with $\mathscr{E}_{5}$ and $\mathscr{E}_{6}$ the thin sectors arising from flattening. It is easy to see that the carrier $C_{H}$ of $H$ is also a 2-simplex in which the vertices are the projective normal classes $\vec{A}^{*}, \vec{B}^{*}$, and $\vec{C}^{*} . \quad C_{H}$ is an lw-face and the linear cells of the PIC-partition for $C_{H}$ coincide with the fibers of the linear map $\mu: C_{H} \rightarrow[0,1]$ defined by $\mu\left(\vec{A}^{*}\right)=\mu\left(\vec{B}^{*}\right)=0$ and $\mu\left(\vec{C}^{*}\right)=1$, except for $\mu^{-1}(0)$ where the cells of the partition are points. The face $\mu^{-1}(0)$ is the only maximal $C_{H}$-independent face.

The remainder of this section is devoted to considering the representation of projective isotopy classes in $\mathscr{P}_{\mathscr{F}}$ and proving that the behavior exhibited in the above example is typical.

Lemma 5.1. Suppose that $G_{1}, G_{2}$ are compact, orientable, incompressible, $\partial$-incompressible, least weight, normal surfaces in the same projective isotopy class. Let $L$ denote the line segment in $\mathscr{P}_{\mathscr{F}}$ with endpoints $\vec{G}_{1}^{*}$ and $\vec{G}_{2}^{*}$, that is the intersection of $\mathscr{P}_{\mathscr{F}}$ with the straight line passing through $\vec{G}_{1}^{*}$ and $\vec{G}_{2}^{*}$. Every orientable normal surface carried by $\operatorname{int}(L)$ is in the same projective isotopy class as $G_{1}$ and $G_{2}$.

Proof. Let $H$ be an orientable normal surface carried by int $(L)$. If $\vec{H}^{*}$ lies on $L$ between $\vec{G}_{1}^{*}$ and $\vec{G}_{2}^{*}$ then there exist positive integers $m, n_{1}, n_{2}$ such that $m H=n_{1} G_{1}+n_{2} G_{2}$ and the conclusion follows from Corollary 4.4. If $\vec{H}^{*}$ does not lie between $\vec{G}_{1}^{*}$ and $\vec{G}_{2}^{*}$ then choose notation such that $\vec{H}^{*}$ is between $\vec{G}_{2}^{*}$ and the endpoint $\vec{G}_{3}^{*}$ of $L$. Since $\vec{G}_{3}^{*}=t \vec{G}_{1}^{*}+(1-t) \vec{G}_{2}^{*}$ and $\vec{G}_{3}^{*}$ has at least one zero coordinate that is not zero in $\vec{G}_{1}^{*}$ it follows that $t$, and hence $\vec{G}_{3}{ }^{*}$, is rational. We
will construct a sequence of surfaces $S_{k}$ such that $\vec{S}_{k}^{*}$ is a sequence in $L$ converging to $\vec{G}_{3}{ }^{*}$ and $S_{k}$ belongs to the projective isotopy class of $G_{1}$. The conclusion will then follow from Corollary 4.4 as before.

There exist positive integers $a, b, c$ such $a \vec{G}_{2}=b \vec{G}_{1}+c \vec{G}_{3}$, where $\vec{G}_{3}$ is an integral $n$-tuple which is a multiple of $\vec{G}_{3}{ }^{*}$. To simplify notation we let $F=b G_{1}, S=a G_{2}$, $G=c G_{3}$. Then $F$ and $S$ are compact, orientable, incompressible, $\partial$-incompressible, least weight, normal surfaces where $S=F+G$ and $m F$ is isotopic to $n S$ for some positive integers $m, n$. Set $S_{0}=S$ and recursively define a sequence of normal surfaces using the formula $S_{k+1}=m^{k+1} G+n S_{k}$.

The following induction argument shows that $S_{k}$ is isotopic to $m^{k} S$. Assuming $S_{k}$ is isotopic to $m^{k} S$, appply Corollary 4.3 once, replacing $n S$ by $m F$, to see that $S_{k+1}$ is isotopic to $m^{k+1} G+n m^{k} S$ and again to see that $S_{k+1}$ is isotopic to $m^{k+1} G+m^{k+1} F=m^{k+1}(G+F)=m^{k+1} S$.

A computation based on the recursion formula shows that $S^{k}=\left(m^{k}+m^{k-1} n\right.$ $\left.+\cdots+m n^{k-1}+n^{k}\right) G+n^{k} F$. Hence $\vec{S}_{k}^{*}=\left(1-\lambda_{k}\right) \vec{G}^{*}+\lambda_{k} \vec{F}^{*}$, where $\lambda_{k}=n^{k} \operatorname{cp}(F) /$ $\left[\left(m^{k}+m^{k-1} n+\cdots m n^{k-1}+n^{k}\right) \operatorname{cp}(G)+n^{k} \operatorname{cp}(F)\right]$. Rewrite $\lambda_{k}$ as

$$
\lambda_{k}=\frac{1}{\left[(m / n)^{k}+(m / n)^{k-1}+\cdots+(m / n)+1\right] \operatorname{cp}(G) / \operatorname{cp}(F)+1} .
$$

If $m / n>1$ then it is clear that $\lambda_{k}$ converges to 0 and $\vec{S}_{k}^{*}$ converges to the endpoint $\vec{G}^{*}$ of $L$. To show that $m / n>1$, first observe that since $m F$ is isotopic to $n S$ it follows from Corollary 4.3 that $m S=m F+m G$ is isotopic to $n S+m G$. Since all surfaces involved are least weight we have $\mathrm{wt}(m S)=\mathrm{wt}(n S)+\mathrm{wt}(m G)$. Hence $m[\operatorname{wt}(S)]>n[\operatorname{wt}(S)]$ and thus $m>n$ as required.

Next we consider normal surfaces carried by an endpoint of the line segment $L$. Using the construction described in section 3, construct a branched surface $\mathscr{B}_{S}$ from a normal surface $S$ carried by $\operatorname{int}(L)$. Recall that $\mathscr{M}\left(\mathscr{B}_{S}\right)$ denotes the cell of invariant projective measures on $\mathscr{B}_{S}$ and is obtained by projectivizing the cone of invariant measures $\mathscr{C}\left(\mathscr{B}_{S}\right)$.

Lemma 5.2. Let $\mathscr{B}=\mathscr{B}_{S}$ be a RIB constructed from a compact, orientable, incompressible, $\partial$-incompressible, least weight, normal surface $S$ and assume that $\partial_{h} N(\mathscr{B}) \subset S$. Suppose that $F=\mathscr{B}(\vec{f})$ is an orientable surface in the projective isotopy class of $S$ and $S=\mathscr{B}(\vec{S})$. Let $L$ denote the line segment which is the intersection of $\mathscr{M}(\mathscr{B})$ with the line through $\vec{s}^{*}$ and $\vec{f}^{*}$. Every orientable surface $G=\mathscr{B}(\vec{g})$ carried by $L$ belongs to the projective isotopy class of $S$.

Proof. The image of $L$ under $\rho^{*}: \mathscr{M}\left(\mathscr{B}_{S}\right) \rightarrow C_{S}$ is the line segment through $\vec{S}^{*}$ and $\vec{F}^{*}$. Thus if $\vec{g}^{*} \in \operatorname{int}(L)$ then it follows from Lemma 5.1 that $\mathscr{B}(\vec{g})$ is in the projective isotopy class of $S$.

Assume that $\vec{g}$ is an endpoint of $L$ and choose notation such that $\vec{f} *$ lies between $\vec{s}^{*}$ and $\vec{g}^{*}$ on $L$. There exist positive integers $a, b, c$ such that $a \vec{f}=b \vec{s}+c \vec{g}$. Then $a F=b S+c G$ and there exist positive integers $m, n$ such that $m F$ is isotopic to $n S$. Since $S$ has positive weights in $\mathscr{B}$, it follows from Theorem 1 in [5] that there exists a sequence of simple isotopies and fiber-preserving isotopies in $N(\mathscr{B})$ (which are normal isotopies) carrying $n S$ to $m F$. A simple isotopy (relative to $N(\mathscr{B})$ ) is associated with a component $P$ of $\overline{M-N(\mathscr{B})}$ endowed with a product structure $P=W \times I$ which extends the I-bundle structure of $L$ (the I-bundle obtained from $N(\mathscr{B})$ by splitting it along $S$ ). Furthermore, $\partial W \times I=\partial_{v} P \subset \partial_{v} N(\mathscr{B})$ and $W \times \partial I=\partial_{h} P \subset \partial_{h} N(\mathscr{B})$. Let $\vec{a}=\left(a_{1}, \cdots, a_{s}\right)$ and $\vec{b}=\left(b_{1}, \cdots, b_{s}\right)$ denote the weights in $N(\mathscr{B})$ correponding to $W \times 0$ and $W \times 1$, respectively. If $\mathscr{B}(\vec{w})$ is a surface carried by $\mathscr{B}$ such that $\vec{w}-\vec{a}$ has no negative entries, then there is an obvious isotopy across $P$, called a simple isotopy, moving $\mathscr{B}(\vec{w})$ to $\mathscr{B}(\vec{w}+\vec{b}-\vec{a})$. Observe that the weights $a_{i}, b_{i}$ can take on only the values 0,1 or 2 and $a_{i}+b_{i} \leq 2$.

The sequence of simple isotopies carrying $n S$ to $m F$ corresponds to a sequence of weights $\vec{p}_{1}=\vec{b}_{1}-\vec{a}_{1}, \cdots, \vec{p}_{q}=\vec{b}_{q}-\vec{a}_{q}$ associated with products $P_{i}=W_{i} \times I$ which support the simple isotopies. Thus $n \vec{s}+\vec{p}_{1}+\cdots+\vec{p}_{q}=m \vec{f}$ and $n \vec{s}+\vec{p}_{1}+\cdots+\vec{p}_{i}-\vec{a}_{i+1}$ $\geq 0$ for each $i$.

Let $H_{0}=\mathscr{B}\left(\overrightarrow{h_{0}}\right)$ denote a multiple of $S$ such that $\operatorname{wt}\left(H_{0}\right) \geq N_{0}$, where $N_{0}$ is a fixed positive integer to be chosen later. Assume that for $0 \leq j \leq r$ we can construct $H_{j}=\mathscr{B}\left(\vec{h}_{j}\right)$ from $H_{j-1}=\mathscr{B}\left(\vec{h}_{j-1}\right)$ by the above sequence of simple isotopies associated to the weights $\vec{p}_{1}, \cdots, \vec{p}_{q}$. Then $\vec{h}_{j}=\vec{h}_{j-1}+\vec{p}_{1}+\cdots+\vec{p}_{q}=\vec{h}_{0}+j\left(\vec{p}_{1}+\cdots+\vec{p}_{q}\right)$ and $\vec{h}_{j-1}+\vec{p}_{1}+\cdots+\vec{p}_{i}-\vec{a}_{i+1} \geq 0$ for each $i$. Observe that $\vec{h}_{j}^{*}$ is on the line $L$ between $\vec{h}_{j-1}^{*}$ and $\vec{g}^{*}$. Because all surfaces carried by $\mathscr{B}$ (and hence by the carrier $C_{S}$ of $S$ ) are least weight, it follows that $\mathrm{wt}\left(H_{j}\right)=\mathrm{wt}\left(H_{0}\right)$. Thus, from Lemma 3.2 we have $\gamma_{\Re 8}\left(H_{j}\right) \leq(d / 3 p) \mathrm{wt}\left(H_{0}\right)$. It follows that there exists a $\delta>0$, independent of $j$, such that the distance on $L$ between and $\vec{h}_{j}^{*}$ and $\vec{h}_{j-1}^{*}$ is greater than $\delta$. Consequently, we eventually reach an $H_{r}$ in the construction such that the sequence of simple isotopies cannot be carried out on $H_{r}$ to construct the next surface $H_{r+1}$. Let $H^{\prime}$ be the last surface that can be constructed from $H_{r}$ in the attempt to form $H_{r+1}$. We shall show that by making a suitable choice for the number $N_{0}$ we can arrange matters such that (i) there exists a surface $S^{\prime}$ carried by $\partial \mathscr{M}(\mathscr{B})$, where either $S^{\prime}=H^{\prime}$ or $S^{\prime}$ is obtained from $2 H^{\prime}$ by applying the next simple isotopy in the sequence, and (ii) $G$ is carried by the minimal face of $\mathscr{C}^{*}(B)$ carrying $S^{\prime}$. The construction can then be repeated in the proper sub-branched surface $\mathscr{M}(\mathscr{B})$.

Choose an $\varepsilon$-neighborhood $U$ in $\boldsymbol{R}^{s}$ about the straight line passing through $\vec{s}^{*}=\vec{h}_{0}^{*}$ and $\vec{g}^{*}$ and set $U^{*}=U \cap \mathscr{M}(\mathscr{B})$. We may assume that is chosen such that the component $\partial_{+} U^{*}$ of $U^{*} \cap \partial \mathscr{M}(\mathscr{B})$ containing $g^{*}$ meets only faces of $\partial \mathscr{M}(\mathscr{B})$ which also contain $\vec{g}^{*}$. This is equivalent to requiring that if the $i$-th coordinate of $\vec{g}^{*}$ is positive then every point in $\partial U_{+}^{*}$ has a positive $i$-th coordinate.

We choose the integer $N_{0}$ sufficiently large so that for any orientable surface
$K=\mathscr{B}(\vec{k})$ with $\vec{k}^{*}$ on $L$ between $\vec{s}^{*}$ and $\vec{g}^{*}$, the vectors $\left(\vec{k}+\vec{p}_{1}+\cdots+\vec{p}_{j}\right)^{*}$ and $\left(\vec{k}+\vec{p}_{1}+\cdots+\vec{p}_{j}-\vec{a}_{j+1}\right)^{*}$ (for $\left.j=1, \cdots, q\right)$ all lie in $U$, and none meet $U^{*}$ $\cap \partial \mathscr{M}(\mathscr{B})-\partial_{+} U^{*}$. With this choice of $N_{0}$ we return to $H_{r}$ and determine the surface $S^{\prime}$. If $\vec{h}_{r}^{*}$ has a zero coordinate then we let $S^{\prime}=H_{r}$. Otherwise, set $K_{0}=H_{r}$ and let $K_{i}=\mathscr{B}\left(\vec{k}_{i}\right)$ denote the surface constructed from $K_{0}$ by using the first $i$ simple isotopies of the sequence. The next surface $K_{i+1}=\mathscr{B}\left(\vec{k}_{i+1}\right)$, with $\vec{k}_{i+1}=\vec{k}_{i}+\vec{p}_{i+1}$, can be formed provided that $\vec{k}_{i}-\vec{a}_{i+1} \geq 0$. Let $t$ be the first index such that $\vec{k}_{t}-\vec{a}_{t+1}$ has a negative coordinate, say the $j$-th coordinate. If the $j$-th coordinate of $\vec{k}_{t}$ is $\cdot 0$ then let $S^{\prime}=K_{t}$. If the $j$-th coordinate of $\vec{k}_{t}$ is greater than 0 then the $j$-th coordinates of $\vec{a}_{t+1}$ and $\vec{b}_{t+1}$ are 2 and 0 , respectively. We let $S^{\prime}=\mathscr{B}\left(\overrightarrow{s^{\prime}}\right)$ where $\vec{s}^{\prime}=2 \vec{k}_{t}+\vec{b}_{t+1}-\vec{a}_{t+1}$. In each case, the surface $S^{\prime}$ has some multiple $x S^{\prime}$ isotopic to some multiple $y S$ and $\vec{s}^{*} \in \partial \mathscr{M}(\mathscr{B})$. Because of our choice of $N_{0}, G$ is carried by the minimal face of $\mathscr{M}(\mathscr{B})$ carrying $S^{\prime}$.

Form the surface $F^{\prime}=x S^{\prime}+G$. It follows from Corollary 4.3 that $F^{\prime}=x S^{\prime}+G$ is isotopic to $y S+G$ and from Lemma 5.1 that a multiple of $y S+G$ is isotopic to a multiple of $S$. Thus a multiple of $F^{\prime}$ is isotopic to a multiple of $S$. Let $\tilde{\mathscr{B}}^{\prime}$ be the sub-branched surface of $\mathscr{B}$ carrying $F^{\prime}$ with positive weights (obtained by deleting the sectors of $\mathscr{B}$ in which $F^{\prime}$ has weight 0 ). It follows from Theorem 2.3 in [7] that $\tilde{\mathscr{B}}^{\prime}$ satisfies the properties to be a RIB except it may have disks of contact. Let $\mathscr{B}^{\prime}$ be the RIB obtained from $\mathscr{\mathscr { B }}^{\prime}$ by removing disks of contact as described in section 3 in such a way that $F^{\prime}$ is carried by $\mathscr{B}^{\prime}$ with positive weights. Although $G$ is carried by $\tilde{\mathscr{B}}^{\prime}$ it may not be carried by $\mathscr{B}^{\prime}$ in which case we isotope $G$ to another least weight surface $G^{\prime}$ that is carried by $\mathscr{B}{ }^{\prime}$, say $G^{\prime}=\mathscr{B}^{\prime}\left(\vec{g}^{\prime}\right)$.

Again we have $F^{\prime}+G^{\prime}$ and $F^{\prime}$ isotopic to a multiple of $S$. If $\vec{g}^{\prime *}$ is not in $\partial \mathscr{M}\left(\mathscr{B}^{\prime}\right)$ then we are done by Lemma 5.1. Otherwise we repeat the above argument using $F^{\prime}, G^{\prime}$ and $S^{\prime}$ in place of $F, G$ and $S$. After each repetition, the number of sectors in the branched surface strictly decreases. It is certainly the case that the sub-branched surface $\mathscr{\mathscr { B }}^{\prime}$ has fewer sectors than $\mathscr{B}$ and removing disks of contact to form $\mathscr{B}^{\prime}$ does not increase the number of sectors but rather enlarges existing ones. Thus, we eventually obtain a surface $S^{\prime \prime}$ in the projective isotopy class of $S$, and a branched surface $\mathscr{B}^{\prime \prime}$ carrying $S^{\prime \prime}$ with positive weights such that either $g^{\prime *}$ is not on $\partial \mathscr{M}\left(\mathscr{B}^{\prime \prime}\right)$ or a multiple of $S^{\prime \prime}$ is equal to a multiple of $G$. In either case we are finished.

Theorem 5.3. Suppose $F_{1}, \cdots, F_{k}$ are least weight normal surfaces in the projective isotopy class of a compact, orientable, incompressible, $\partial$-incompressible surface $F$. Let $V$ denote the subspace spanned by the normal classes $\vec{F}_{1}, \cdots, \vec{F}_{k}$ in $\boldsymbol{R}^{n}$. Then every normal surface carried by the linear cell $V \cap \mathscr{P}_{\mathscr{F}}$ is least weight and in the projective isotopy class of $F$.

Proof. Let $C$ denote the complete lw-face carrying $F$. It is sufficient to prove that if $S_{1}, S_{2}$ are least weight, orientable, normal surfaces each with a multiple
isotopic to a multiple of $F$ then every orientable surface $G$ carried by the line $L$ in $C$ passing through $\vec{S}_{1}^{*}$ and $\vec{S}_{2}^{*}$ has a multiple isotopic to a multiple of $F$. If $G$ is carried by the interior of $L$ then this follows from Lemma 5.1. Suppose then that $\vec{G}^{*}$ is an endpoint of $L$. Let $S$ be an orientable normal surface carried by the interior of $L$ between $\vec{S}_{1}{ }^{*}$ and $\vec{G}^{*}$. Form the branched surface $\mathscr{B}=\mathscr{B}_{s}$. There exist integers $a, b$ and $c$ such that $a S=b S_{1}+c G$. If $G$ is carried by $\mathscr{B}$ then, by Lemma 5.2, a multiple of $G$ is isotopic to a multiple of $F$. If $G$ is not carried by $\mathscr{B}$ then, by Lemma 3.1, there is a surface $G^{\prime}$ isotopic to $G$ and carried by $\mathscr{B}$. Since $S+G$ is in the projective isotopy class of $F$ it follows from Corollary 4.3 that $F^{\prime}=S+G^{\prime}$ also belongs to the projective isotopy class of $F$. But now $F^{\prime}, S$ and $G$ are carried by $\mathscr{B}$ and we apply Lemma 5.2 as before to conclude that $G^{\prime}$ is in the projective isotopy class of $F$.

Corollary 5.4. Suppose $C^{\prime}$ and $C^{\prime \prime}$ are lw-faces of $\mathscr{P}_{\mathscr{F}}$ such that $C^{\prime} \cap C^{\prime \prime}$ is a proper face of each. If $C^{\prime} \cap C^{\prime \prime}$ is not $C^{\prime}$-independent then there exists an lw-face $C$ such that $C^{\prime}$ and $C^{\prime \prime}$ are both faces of $C$ and $C^{\prime \prime}$ is not $C$-independent.

Proof. Since $C^{\prime} \cap C^{\prime \prime}$ is not $C^{\prime}$-independent there exist rational points $\vec{x} \in C^{\prime} \cap C^{\prime \prime}$ and $\vec{y} \in C^{\prime}$ associated to the same projective isotopy class. Let $\vec{z}$ be any rational point in $\dot{C}^{\prime \prime}$. Then, by Theorem $5.3,(\vec{x}+\vec{z}) / 2$ and $(\vec{y}+\vec{z}) / 2$ represent the same projective isotopy class. It follows from Theorem 4.5 that there exists a complete lw-face $\hat{C}$ carrying both $(\vec{x}+\vec{z}) / 2$ and $(\vec{y}+\vec{z}) / 2$. Let $C$ be the face of $\hat{C}$ containing $(\vec{y}+\vec{z}) / 2$ in its interior. Then $C^{\prime}, C^{\prime \prime}$ are both faces of $C$ since $C$ carries an interior point of each, namely $\vec{y}$ and $\vec{z}$. Moreover, $C^{\prime \prime}$ is not $C$-independent since $(\vec{x}+\vec{z}) / 2 \in C$ is in the same projective isotopy class as $(\vec{y}+\vec{z}) / 2 \in \mathscr{C}$.

The notion of a PIC-partition was defined in the introduction and the next theorem gives the existence of such a partition for each lw-face of $\mathscr{P}_{\mathscr{F}}$.

Theorem 5.5. Each $l w$-face $C$ in has a PIC-partition into compact linear cells $X_{\alpha}$ which have the property that the set of normal surfaces carried by an $X_{\alpha}$ is precisely the set of surfaces carried by $C$ and in the corresponding projective isotopy class.

Proof. Suppose $C$ is an lw-face and let $\mathscr{D}$ denote the union of all $C$-independent faces of $C$. Thus each face in $\mathscr{D}$ does not carry a surface in the same projective isotopy class as a surface carried by $\mathscr{C}^{C}$. The first step is to construct a family of $k_{C^{-}}$-dimensional subspaces $V_{\alpha}$ such that $C-\mathscr{D}$ is partitioned by the $k_{C^{-}}$-dimensional cells $V_{\alpha} \cap C$ meeting $\mathscr{C}$ in such a way that the projective isotopy classes correspond to the rational cells in the partition. We then repeat this construction for each maximal $C$-independent face $D$ of $C$ to extend the partition to $D-\mathscr{E}$, where $\mathscr{E}$ denotes the union of all $D$-independent faces of $D$. This process is continued
until we have constructed a partition of the entire face $C$.
Start with an orientable normal surface $F_{1}$ carried by $\mathscr{C}$ and choose a maximal collection of normal surfaces $\left\{F_{1}, \cdots, F_{k}\right\}$ carried by $C$ and in the same projective isotopy class as $F_{1}$ such that the set of vectors $\left\{\vec{F}_{1}, \cdots, \vec{F}_{k}\right\}$ is linear independent. By taking appropriate integral multiples, we may assume that the surfaces $F_{i}$ are mutually isotopic. Let $V$ denote the $k$-dimensional subspace spanned by the vectors $\left\{\vec{F}_{1}, \cdots, \vec{F}_{k}\right\}$ and consider the compact linear cell $A=V \cap C$. By Theorem 5.3, every normal surface carried by $A$ belongs to the projective isotopy class of $F_{1}$. The maximality of the set $\left\{F_{1}, \cdots, F_{k}\right\}$ ensures that every normal surface carried by $C$ and in the projective isotopy class of $F_{1}$ is carried by $A$.

To construct the first family of linear cells in the PIC-partition, choose an arbitrary orientable normal surface $H$ carried by $C-A$. Let $t$ be a fixed rational number with $0 \leq t<1$ and choose a positive integer $m$ such that $m t$ is an integer. Form the normal surface $G_{i}=m t H+m(1-t) F_{i}$ for each $i$. We may assume that the $G_{i}$ are orientable by replacing $m$ with $2 m$ if necessary. It follows from Corollary 4.3 that $G_{i}$ is isotopic to $G_{1}$ for each $i$. The vectors $\left\{\vec{G}_{1}, \cdots, \vec{G}_{k}\right\}$ span a $k$-dimensional subspace $V^{\prime}$. As before, every normal surface with projective normal coordinates in the linear cell $A^{\prime}=V^{\prime} \cap C$ belongs to the same projective isotopy class as $G_{1}$. A dimension argument in which the roles of $V$ and $V^{\prime}$ are reversed shows that every normal surface $S$ in the projective isotopy class of $G_{1}$ and carried by $C$ is carried by $A^{\prime}$. If $A^{\prime \prime}=V^{\prime \prime} \cap C$ is any other linear cell construced in this way then $A^{\prime \prime}$ meets $C^{C}$ and it follows from the previous comment that either $A^{\prime}=$ $A^{\prime \prime}$ or $A^{\prime} \cap A^{\prime \prime}=\emptyset$. Consider the family $\left\{V_{\alpha}\right\}$ of $k$-dimensional vector subspaces which intersect $C$ and are spanned by vectors $\left\{\vec{a}_{1}, \cdots, \vec{a}_{k}\right\}$ of the form $\overrightarrow{a_{i}}=t \vec{F}_{i}+(1-t) \vec{H}$, $t>0$, where $\vec{H}$ is any orientable normal surface carried by $C$. If we let $\left\{A_{\alpha}\right\}$ denote the linear cells of the form $A_{\alpha}=V_{\alpha} \cap C$ which meet $\mathscr{C}$, then we obtain the PIC-partition $\left\{A_{\alpha}\right\}$ of $C-\mathscr{D}$ into linear cells of dimension $k_{C}=k$.

If the given $C$ is a face of an lw-face $\hat{C}$ then each cell $A_{\alpha}$ is contained in one of the $k_{\hat{C}}$-dimensional subspaces used to form the PIC-partition of $\hat{C}$. To see this, let $\hat{H}$ be any orientable normal surface carried by the interior of $\hat{C}$. Then the normal surfaces $\hat{F}_{1}=\hat{H}+F_{1}, \cdots, \hat{F}_{n}=\hat{H}+F_{n}$ are isotopic, carried by the interior of $\hat{C}$, and the collection can be enlarged so as to span a linear cell in the PIC-partition of $\hat{C}$.

The PIC-partition for the remainder of $C$ is obtained by repeating this construction for the maximal $C$-independent faces of $C$, and then continuing the process until a partition is obtained that covers all of $C$.

## 6. Applications to finite group actions

Let $G$ be a finite group of simplicial homemorphisms of the compact, orientable, irredudible, $\partial$-irreducible 3-manifold $M$ with a fixed triangulation $\mathscr{T}$. Assume that $\operatorname{Fix}(G)=\{x \mid g(x)=x$ for some $g \in G\}$ is a subcomplex of $\mathscr{T}$. A surface $F$ is
$G$-equvariant if for each component $K$ of $F$ and each $g \in G$ either $g(K)=K$ or $g(K) \cap K=\emptyset$. A $G$-invariant face of $\mathscr{P}_{\mathscr{F}}$ is one invariant under the linear action of $G$ on $\mathscr{P}_{\mathscr{T}} \cap Q^{n}$ defined by $g\left(\vec{F}^{*}\right)=\overrightarrow{g(F)}^{*}$. If a compact, orientable, incompressible, $\partial$-incompressible, least weight, normal surface $F$ is $G$-equivariant then the carrier of $\cup g(F)$ is a $G$-invariant lw-face of $\mathscr{P}_{\mathscr{F}}$. On the other hand, if $F$ is a normal surface carried by a $G$-invariant 1 w -face $C$ then it is always the case that the normal sum $\Sigma_{g \in G} g(F)$ is defiend.

Theorem 6.1. Let $F$ be a compact, orientable, incompressible, $\partial$-incompressible, least weight, normal surface carried by an lw-face $C$ such that $g\left(\vec{F}^{*}\right) \in C$ for all $g \in G$. Then the normal sum $S=\Sigma_{g \in G} g(F)$ is an injective, $\partial$-injective, least weight surface which is invariant under $G$ and $\vec{S}^{*}$ is fixed by $G$. Moreover, the isotoppy class of $S$ depends only on the isotopy classes of the surfaces $g(F)$.

Proof. Since each $g(F)$ is carried by $C$ the normal sum $S=\Sigma_{g \in G} g(F)$ is defined. It follows from Theorem 4.2 that $S$ (or else $2 S$ ) is an orientable, incompressible, $\partial$-incompressible, least weight, normal surface. By Corollary 4.3, the isotopy class of $S$ ( or $2 S$ ) depends only on the isotopy classes of the surfaces $g(F)$.

It remains to show that the normal sum $\Sigma_{g \in G} g(F)$ is $G$-invariant. First consider the case in which $G$ is acting freely. The normal surface $\Sigma_{g \in G} g(F)$ is obtained by forming the geometric sum of all the images of $F$ under $G$. But this is the unique normal surface determined by the $G$-invariant collection of points $\left[\bigcup_{g \in G} g(F)\right] \cap \mathscr{T}^{(1)}$ and is automatically $G$-equivariant.

If $\operatorname{Fix}(G) \neq \emptyset$ then the normal surfaces $F$ and $g(F)$ do not intersect transversely (in our sense) along Fix $(g) \cap F$. However, there is a unique (up to normal isotopy) geometric sum $F+g(F)$ obtained by first moving the surfaces by normal isotopies to make them intersect transversely and then forming the geometric sum. This


Fig. 5. Equivariant regular exchange along $\operatorname{Fix}(G)$
can be done in such a way that the resulting sum will be $G$-equivariant. Choose a small neighborhood $I_{v}$ in $\mathscr{T}^{(1)}$ about each point $v \in \operatorname{Fix}(G) \cap F \cap \mathscr{T}^{(1)}$ such that the family of neighborhoods $\left\{I_{v}\right\}$ is $G$-equivariant. Now the argument is the same as in the free case if, when forming the geometric sum, we relace each $v$ by the two end points $v^{\prime}, v^{\prime \prime}$ of $I_{v}$ as shown in Figure 5.

Corollary 6.2. Suppose that $F$ is an orientable normal surface. The normal sum $S=\Sigma_{g} g(F)$ is defined and $2 S$ is a G-equivariant compact, orientable, incompressible, $\partial$-incompressible, least weight surface if and only if there exists a $G$-invariant lw-face $C$ of $\mathscr{P}_{\mathscr{F}}$ carrying $F$.

Proof. If $2 S$ is $G$-equivariant then the carrier $C_{S}$ of $S$ is a $G$-invariant lw-face of $\mathscr{P}_{\mathscr{F}}$ carrying $F$.

Corollary 6.3. Suppose $G$ is homotopically trivial. Then every lw-complete face in $\mathscr{P}_{\mathscr{T}}$ is $G$-invariant. If $F$ is any compact, orientable, incompressible, $\partial$-incompressibe, least weight, normal surface then $S=\Sigma_{g} g(F)$ is a $G$-equivariant collection of pairwise disjoint normal surfaces isotopic to $F$.

Proof. By [10], each $g \in G$ is isotopic to the identity.
Corollary 6.4. Suppose $F$ is a compact, orientable, incompressible, $\partial$ incompressible, least weight, normal surface such that for each $g \in G$, the normal surface $g(F)$ is isotopic to a surface disjoint from $F$. Then $F$ is carried by some $G$-invariant $l w$-face $C$ and the normal sum $S=\Sigma_{g} g(K)$ is defined and is a $G$-equivariant collection of pairwise disjoint surfaces. Moreover, if $K$ is any compact, orientable, incompressible, $\partial$-incompressible, least weight, normal surface carried by $C$ then the normal sum $\Sigma_{g} g(F)$ is a G-equivariant, injective, $\partial$-injective, least weight surface.

Proof. It follows from Proposition 3.7 in [3] that the normal sum $S=\Sigma_{g} g(K)$ is defined and is least weight. Therefore the carrier of $S$ is the desired $G$-invariant lw-face.

## References

[1] W. Haken: Theorie der Normal Flächen, Acta Math. 105 (1961), 245-375.
[2] W. Jaco and U. Oertel: An Algorithm to Decide if a 3-manifold is a Haken Minifold, Topology 23 (1984), 195-209.
[3] W. Jaco and J. Rubinstein: PL equivariant surgery and invariant decompositions of 3-manifolds, Advances in Mathematics 73 (1989), 149-191.
[4] W. Jaco and J. Tollefson: Algorithms for the complete decomposition of a closed 3-manifold, Illinois Journal of Mathematics (to appear).
[5] U. Oertel: Incompressible branched surfaces, Invent. Math. 76 (1984), 385-410.
[6] U. Oertel: Sums of Incompressible Surfaces, Proc. Amer. Math. Soc. 102 (1988), 711-719.
[7] U. Oertel: Measured Laminations in 3-manifolds, Trans. Amer. Math. Soc. 305 (1988), 531-573.
[8] J. Tollefson: Normal Surfaces Minimizing Weight in a Homology Class, Topology and Its Applications 50 (1993), 63-71.
[9] J. Tollefson: Innermost Disk Pairs in Least Weight Normal Surfaces, Topology and Its Applications (to appear).
[10] F. Waldhausen: On irreducible 3-manifolds which are sufficiently large, Ann. of Math. (2) 87 (1968), 56-88.

Department of Mathematics University of Connecticut
Room 111, U-9196 Auditorium Road
Storrs, CT 06269-3009 U.S.A.

