Three-dimensional Topology

<u>Introduction</u> We summarize the basic definitions. For three-dimensional topology, intuitive understanding is much more important than technical details.

An <u>n-manifold</u> is a Hausdorff space M such that every point of M has a neighborhood homeomorphic to $\mathbb{R}^n$. We generalize this to allow <u>n-manifolds with boundary</u>, in which every point has a neighborhood homeomorphic to $\mathbb{R}^n_+ = \{(x_1 \dots x_n) : x_n \geq 0\}$ or $\mathbb{R}^n$. The <u>boundary</u> ∂M consists of all points of M which don't have neighborhoods homeomorphic to $\mathbb{R}^n$. ∂M is a manifold, $\partial(\partial M) = \emptyset$ and $\dim(\partial M) = \dim M - 1$. M is a <u>closed manifold</u> if M is compact and $\partial M = \emptyset$.

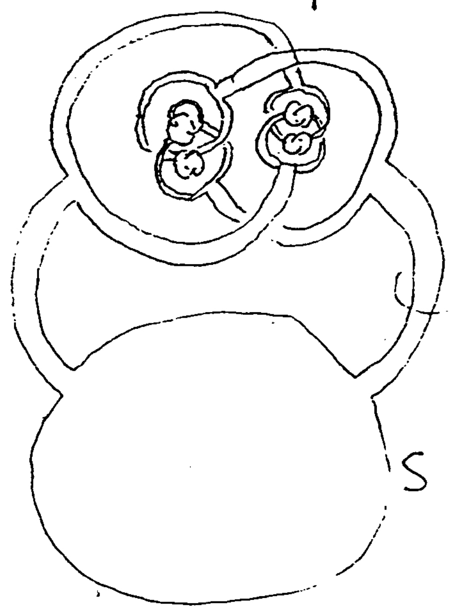
We would like to generalize the well-known classification of closed surfaces

2-sphere S^2 2-torus T^2 $T^2 \# T^2$ (connected sum)   . . .

projective plane P^2 Klein bottle $P^2 \# P^2$ $P^2 \# P^2 \# P^2$ . . .

Thurston proposed his "Geometrization Conjecture" as a classification of 3-manifolds. Many cases have been proved, but many remain unsolved (including the Poincaré Conjecture, that every simply connected closed 3-manifold is homeomorphic to $S^3 = \mathbb{R}^3 \cup \{\infty\}$)

We shall often study surfaces in 3-manifolds. To avoid 'pathological' examples like the Alexander horned sphere.



we shall impose differentiable (C^∞) or piecewise linear (PL) structures on our manifolds.

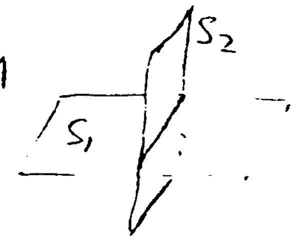
Let E be an open set in $\mathbb{R}^n$. A map $h: E \to \mathbb{R}^n$ is C^∞ if it has continuous partial derivatives of all orders. $h: E \to \mathbb{R}^n$ is PL if there is a locally finite covering of E by simplexes such that h restricts to a linear map on each simplex of the covering. A C^∞ or PL structure on a manifold M is defined by a covering of M by open sets U_i together with chart maps $\phi_i : U_i \to \mathbb{R}^n$ (which are homeomorphism onto open sets in $\mathbb{R}^n$) such that the overlap maps $\phi_j \phi_i^{-1} \mid \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$ are C^∞ or PL. It is not hard to prove that a compact PL manifold is homeomorphic to a finite simplicial complex.

Every manifold of dimension ≤ 3 has C^∞ and PL structures, which are unique up to C^∞ diffeomorphism and PL homeomorphism respectively. This was proved in the 1950's by Moise and Bing. For manifolds of dimension ≥ 4, the topological, PL and C^∞ classifications are different.

S is a submanifold of M if every $x \in S$ has a neighborhood U such that $(U, U \cap S)$ is (PL) (C^∞) homeomorphic to the standard pair $(\mathbb{R}^m, \mathbb{R}^s)$. $f : S \to M$ is an embedding if f defines a homeomorphism from S to a submanifold of M.

The following ideas are explained from the C^∞ point of view, for surfaces in 3-manifolds. Again, I emphasise that an intuitive geometric understanding is much more important than the technical details.

Surfaces S_1, S_2 in 3-manifold M are <u>transverse</u> ($S_1 \pitchfork S_2$) if they intersect as shown.



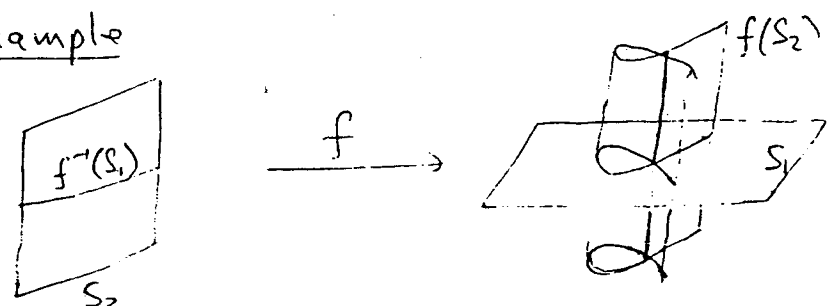
Technically, transversality means that for all $x \in S_1 \cap S_2$, $T_x(S_1) + T_x(S_2) = T_x(M)$

An equivalent definition, which can also be used for PL manifolds, is that $S_1 \pitchfork S_2$ iff every $x \in S_1 \cap S_2$ has a neighborhood U such that $(U, S_1 \cap U, S_2 \cap U) \cong (\mathbb{R}^3, \mathbb{R}^2 \times 0, 0 \times \mathbb{R}^2)$

A surface S and a curve C are transverse iff they intersect like the (x, y) plane and the z axis in $\mathbb{R}^3$.

A map $f: S_2 \to M$ is transverse to a submanifold $S_1 \subset M$ if, for all $x \in f^{-1}(S_1)$, $T_{f(x)}(S_1) + df(T_x(S_2)) = T_{f(x)}(M)$.

<u>Example</u>



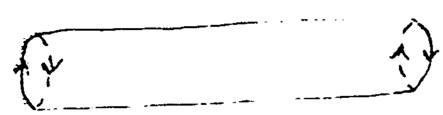
This map f is an <u>immersion</u>: locally it is an embedding, but two lines in S_2 are identified to a single line in $f(S_2)$, where $f(S_2)$ intersects itself transversely

Theorem (A) If S_1 is a submanifold of M then any map $f: S_2 \to M$ can be approximated by a map $g: S_2 \to M$ which is transverse to S_1.

(B) If $f: S_2 \to M$ is transverse to S_1 then $f^{-1}(S_1)$ is a submanifold of S_2;
$\dim(S_2) - \dim(f^{-1}(S_1)) = \dim(M) - \dim(S_1)$
If S_1, S_2 are surfaces and M is a 3-manifold then $f^{-1}(S_1)$ is a 1-dimensional manifold.

You can find proofs in introductory books on differential topology. For us, the proofs are not important, just the ideas.

Submanifolds of 3-manifolds have neighborhoods of simple types. If S is a circle in an orientable 3-manifold M, then S has a neighborhood homeomorphic to a solid torus $S^1 \times \mathbb{R}^2$. In general, an orientation-preserving circle has a solid torus neighborhood, an orientation-reversing circle has a solid Klein bottle neighborhood, obtained from a cylinder by identifying the ends with a 'twist'.



An orientable surface S in an orientable 3-manifold M has a neighborhood homeomorphic to $M \times \mathbb{R}$ (a "product neighborhood"). [If S or M is non-orientable, the neighborhood may be a twisted $\mathbb{R}$-bundle over S.] These facts follow from the tubular neighborhood theorem in differential topology: again, details are not important.

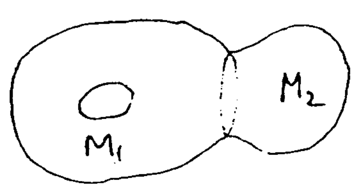
The Schönflies Theorem

We use the notation
B^n (n-ball) = D^n (n-disk) = $\{x \in \mathbb{R}^n : \|x\| \leq 1\}$
and $S^n = \partial B^{n+1}$: S^n is homeomorphic to $\mathbb{R}^n \cup \{\infty\}$

2.1 Theorem Any (C^∞ or PL) 2-sphere in $\mathbb{R}^3$ bounds a 3-ball in $\mathbb{R}^3$.

Remarks The original Schönflies theorem states that every Jordan curve in $\mathbb{R}^2$ bounds a disk: this is easy to prove for PL (polygonal) curves. The theorem is known for PL n-spheres in $\mathbb{R}^{n+1}$ except for $n = 3$.
We present a C^∞ proof, which can be made into a PL proof. We need the following fact, which can be deduced from the (PL or C^∞) Schönflies theorem in the plane.

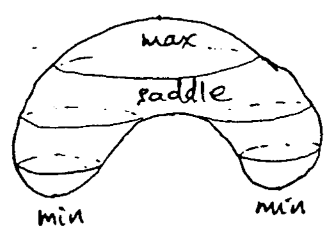
2.2 Proposition Let $M = M_1 \cup M_2$, where M_1 is a 3-manifold, $M_2 \cong B^3$ and $M_1 \cap M_2 = \partial M_1 \cap \partial M_2 \cong D^2$. Then $M \cong M_1$

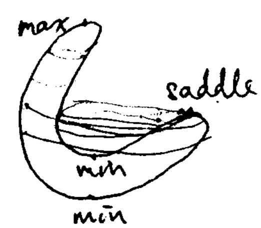


Proof of Schönflies Th. Let $S \subset \mathbb{R}^3$ be a C^∞ sphere. S may be perturbed so that the "height function" $h : \mathbb{R}^3 \to \mathbb{R}$ has a finite number of maxima, minima and saddle points and no other critical points. We prove that S bounds a ball by induction on the number n of saddle points.

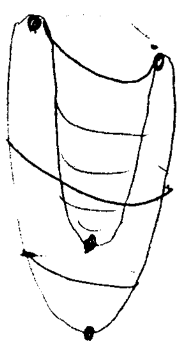
If $n=0$, S has one minimum and one maximum so bounds a ball.

If $n=1$, there are 2 minima and one maximum (or 2 maxima and one minimum, but this is exactly similar). There are 2 possibilities:



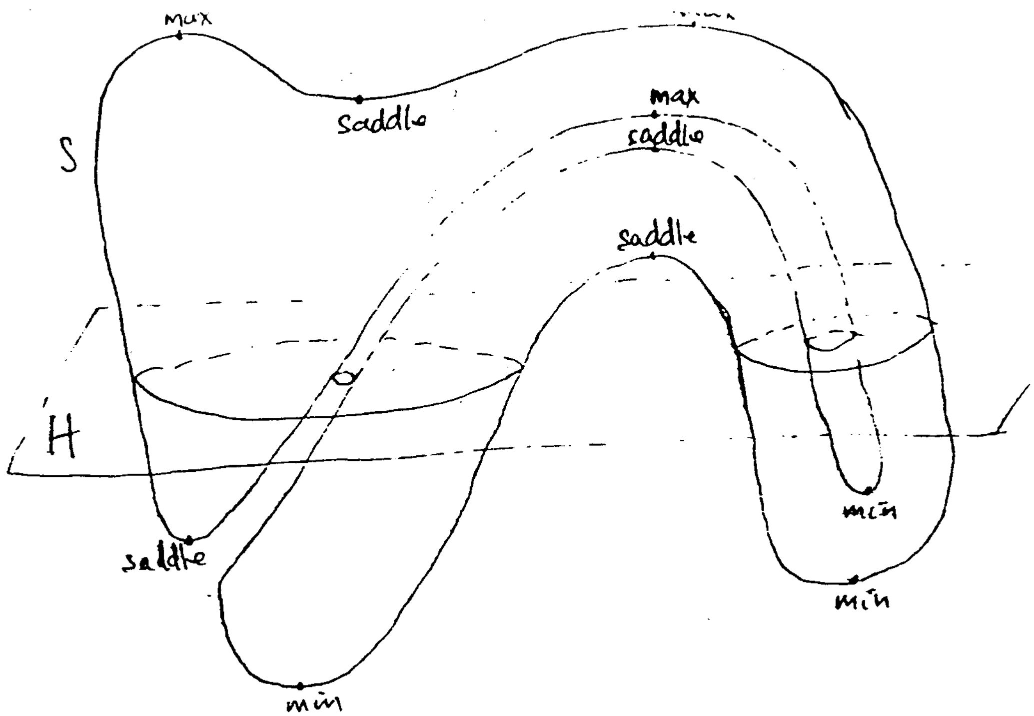


('tilted bowl')



In each case S bounds a ball in $\mathbb{R}^3$.

Suppose $n \geq 2$, and suppose every 2-sphere in $\mathbb{R}^3$ with $< n$ saddle points bounds a ball. $\exists$ horizontal plane H disjoint from critical points and with ≥ 1 saddle point on each side of H. $H \cap S$ is a compact 1-manifold, no boundary, so $H \cap S$ is a finite union of disjoint circles in H.



Let C be an <u>innermost</u> <u>component</u> of $H \cap S$, so $\exists$ disk $D \subset H$ with $D \cap S = \partial D = C$.
C separates S into disks D_1, D_2.
$S_1 := D \cup D_1$, $S_2 = D \cup D_2$ are embedded 2-spheres in $\mathbb{R}^3$. There are 2 cases.

<u>Case 1</u> If S_1, S_2 both have saddles, both have $< n$ saddles, so S_1, S_2 bound balls in $\mathbb{R}^3$. By the proposition, S bounds a ball.

ould have
$S_1 \subset \mathrm{int}(S_2)$??

<u>Case 2</u> If S_1 has no saddle, S_1 bounds a ball, so S bounds a ball iff S_2 bounds a ball.
Push S_2 to reduce number of components of $S_2 \cap H$ (eliminating C), and repeat the process with a new innermost circle.

What if $S' \subseteq S'$?

Since H was chosen to separate the saddles of S we eventually reach case 1), completing the proof

The same argument also gives the following theorem of Alexander.

Q6 (could be) Why not both

2.3 **Theorem** If T is a (C^∞ or PL) torus in $S^3 = \mathbb{R}^3 \cup \{\infty\}$, then one of the components of $S^3 \setminus T$ has closure homeomorphic to a solid torus $S^1 \times D^2$.

A 3-manifold M is *irreducible* if every 2-sphere in M bounds a 3-ball in M. We have shown that S^3, $\mathbb{R}^3$, B^3 are all irreducible.

Suppose S is a 2-sphere in a 3-manifold M and D is a 2-disk in M with $D \cap S = \partial D$. ∂D separates S into disks D', D'': let $S' = D \cup D'$, $S'' = D \cup D''$.

2.4 **Lemma** If two of S, S', S'' bound balls in M then the third also bounds a ball in M.

Proof Suppose S', S'' bound balls $B', B'' \subset M$ (other cases are exactly similar). If $S \subset B'$ or $S \subset B''$ the result follows from the Schönflies theorem. Otherwise $B' \cap B'' = D$, so $B' \cup B''$ is a ball with boundary S.

Suppose a 3-manifold M contains a 2-sphere S that separates M into N_1 and N_2. Let $M_1 = N_1 \cup B^3$ and $M_2 = N_2 \cup B^3$. We say that M is the connected sum $M_1 \# M_2$ of M_1 and M_2.

It can be shown that M is uniquely determined as an oriented manifold by M_1, M_2 together with their orientations. (If we ignore orientation, M_1 and M_2 may have <u>two</u> non-homeomorphic connected sums.) Note that $M \# S^3 \cong M$.

M is <u>prime</u> if $M = M_1 \# M_2 \Rightarrow M_1$ or $M_2 \cong S^3$.

2.5 <u>Lemma</u> Every irreducible 3-manifold is prime. A closed orientable prime 3-manifold is <u>either</u> irreducible <u>or</u> homeomorphic to $S^1 \times S^2$.

<u>Proof</u> Clearly irreducible manifolds are prime. Suppose M prime and $S^2 \subset M$. If S^2 separates M into N_1, N_2 then N_1 or N_2 is a ball (since M is prime: use Schönflies theorem).
Suppose S^2 does not separate M. Since S^2, M are orientable, S^2 has a neighborhood homeomorphic to $S^2 \times I$.
Let $N = \overline{M \setminus S^2 \times I}$ and let $a \in S^2$. Since S^2 does not separate M, there is an embedding $\alpha : I \to N$ with $\alpha(0) = a \times 0$ and $\alpha(1) = a \times 1$. Let C be the circle $\alpha(I) \cup (a \times I) \subset M$. Since M is orientable, C has a neighborhood homeomorphic to $D^2 \times S^1$. By choosing the neighborhoods small, we may arrange that $S^2 \times I \cap D^2 \times S^1 = D^2 \times I$, where D^2 is identified with a small disk neighborhood

of a in S^2 and I is identified with a short interval in S^1.

So $M \supset W = S^2 \times I \cup D^2 \times S^1$

$$= S^2 \times S^1 \setminus ((S^2 \setminus D^2) \times (S^1 \setminus I))$$
$$\cong S^2 \times S^1 \setminus \text{3-ball}.$$

But $\partial W \cong S^2$ separates M, so ∂W bounds a 3-ball in M, so $M \cong S^2 \times S^1$.

Let $S = S_1 \cup \dots \cup S_k$ be a disjoint union of 2-spheres in a 3-manifold M. S is independent if no component of $M \setminus S$ is homeomorphic to a punctured sphere (ie to $S^3 \setminus$ (finite disjoint union of balls)).

If $M = M_0 \# M_1 \# \cdots \# M_k$ with M_i not homeomorphic to S^3 then M contains an independent set of k 2-spheres. We shall prove that, for each M, there is an upper bound to the number of independent 2-spheres in M.

Lemma 2.6 Suppose $S = S_1 \cup \dots \cup S_k$ is an independent set of 2-spheres in M. Suppose D is a 2-disk in M with $D \cap S = \partial D$, and suppose ∂D separates S_i into disks D', D''. Let $S_i' = D \cup D'$, $S_i'' = D \cup D''$, and let S', S'' be obtained from S by replacing S_i with S_i', S_i'' respectively. Then one of S', S'' is independent.

Proof Follows from lemma 2.4.

Exercises

1) Define $f : \mathbb{C} \to \mathbb{C} \times \mathbb{R}$ by
$$f(z) = (z^2, \text{imaginary part of } z).$$
Is f an immersion? Is the restriction of f to $\{z : 1 \leq |z| \leq 2\}$ an immersion?

2) Draw a picture of a closed orientable surface S of genus 2 (ie $T^2 \# T^2$) embedded in $\mathbb{R}^3$ so that neither component of $\mathbb{R}^3 \cup \{\infty\} \setminus S$ is standard.

3) Let $\tilde{M}$ be a covering space of a 3-manifold M. Prove that if $\tilde{M}$ is irreducible then M is irreducible. Deduce that $T^3 = S^1 \times S^1 \times S^1$ is irreducible.

4) Show how to deduce proposition 2.2 from the Schönflies theorem for $S^1 \subset S^2$.

5) Prove Theorem 2.3 (use a similar argument to Th 2.1)

§3 Prime Factorization, Normal Surfaces

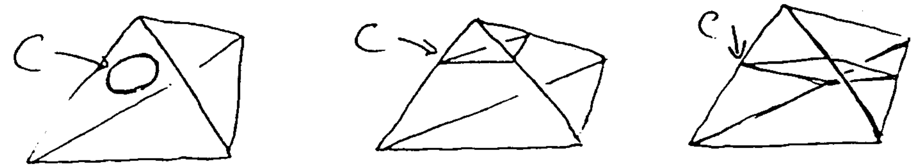
A homeomorphism $H: M \times I \to M \times I$ such that $H(x,t) = (H_t(x), t)$ is called an isotopy from $H_0: M \to M$ to $H_1: M \to M$. Surfaces S_0, S_1 are isotopic if $\exists$ isotopy H such that $H_0 = 1$ and $H_1(S_0) = S_1$. Isotopy is an equivalence relation.

Let M be a closed PL 3-manifold. M is homeomorphic to a simplicial complex K; in fact we identify M with K. Let K^i be the i-skeleton of K, the union of all the simplices of K of dimension $\le i$.

Let S be a surface embedded in M (not necessarily connected). S can be perturbed by a small isotopy so that S meets all simplices of K transversely. So $S \cap K^0 = \emptyset$, $S \cap K^1$ is a finite number of points, and for each 2-simplex Δ of K, $S \cap \Delta$ is a finite disjoint union of circles and arcs embedded in Δ with end-points in $\partial\Delta$. Define the weight of S to be $w(S) = |S \cap K^1|$, the number of points in $S \cap K^1$. Choose S to have least weight among all surfaces isotopic to S.

Lemma 3.1 If S has least weight in its isotopy class, then for each 3-simplex Δ

of K, each component C of $S \cap \partial\Delta$ has one of the following types.

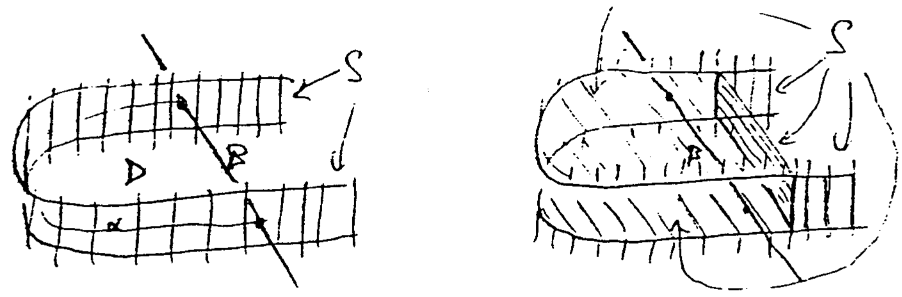


<u>Proof</u> C is a simple closed curve in $\partial\Delta$, disjoint from the vertices and transverse to the edges of Δ. If C is not one of the types shown, it contains one of the following configurations.

In either case there is an arc $\alpha' \subset C$ and an arc β in an edge of Δ such that $\alpha' \cup \beta$ is a simple closed curve in $\partial\Delta$, bounding a disk $D' \subset \partial\Delta$.

Let F be the component of $S \cap \Delta$ containing C. Let α be an arc in F, close to α', with $\partial\alpha = \partial\alpha' = \alpha \cap \partial F$. Let D be a disk in Δ, close to D', with $\partial D = \alpha \cup \beta$. D meets S transversely along α. (There may be other intersections of D with S in the interior of D or on β; intersections on β could have been avoided by choosing β innermost.)

Now there is an isotopy supported in a neighborhood of D (ie the identity <u>outside</u> a neighborhood of D) carrying S to a surface with fewer intersections with K^1 (the intersections at the end-points of β are eliminated, and no new intersections are created).

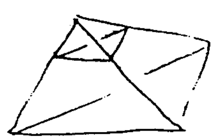


before isotopy after isotopy

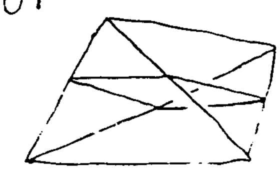
This contradicts the assumption that S has least weight, completing the proof.

Define S to be a <u>normal</u> <u>surface</u> with respect to K if

1) S is transverse to all simplices of K,
2) for each 3-simplex Δ of K, each component C of $S \cap \partial\Delta$ is of type

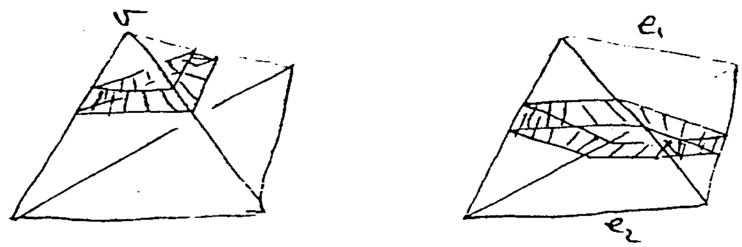


or



3) each component of $S \cap \Delta$ is a disk (for each 3-simplex Δ of K).

If X is the closure of a good component of $M \setminus S$, we shall construct an equivalence relation $\sim$ on X such that each equivalence class is homeomorphic to a closed interval. Let Δ be a 3-simplex of K and let C be a component of $X \cap \Delta$.



If C is a (triangle) $\times I$, meeting the 3 edges of Δ that pass through the vertex v, the equivalence classes are the intersections of C with the lines through v. If C is a (square) $\times I$ not meeting the edges e_1, e_2 of Δ, the equivalence classes are intersections of C with lines that meet both e_1 and e_2. (Note that every point of Δ is on a unique line joining a point of e_1 and a point of e_2.) These definitions fit together to give a well-defined equivalence relation on X.

Let $Y = X/\sim$ be the set of equivalence classes, with the quotient topology. Y is a connected closed surface. The projection map $p: X \to Y$ restricts to a two-to-one map $p|\partial X \to Y$ which is locally a homeomorphism. Each component of ∂X is mapped onto Y (because it has open and closed image), so either

1) ∂X has 2 components, each mapped homeomorphically to Y, or
2) ∂X has 1 component, which is a double covering of Y.

In case 1), it is easy to see that $X \cong Y \times I$, since the intervals $p^{-1}(y)$ $(y \in Y)$ can be oriented coherently.

In case 2), the intervals $p^{-1}(y)$ cannot be oriented coherently and X is a 'twisted I-bundle over Y'. In this case, the image of $\pi_1(\partial X)$ in $\pi_1(X)$ (under the map induced by the inclusion $i: \partial X \to X$) is a subgroup of index 2. It follows from Van Kampen's theorem, and the fact that $H_1(M: \mathbb{Z}_2)$ is obtained from $\pi_1(M)$ by Abelianizing and reducing modulo 2 that each twisted I-bundle component of $M \setminus S$ contributes 1 to $\dim H_1(M: \mathbb{Z}_2)$. Alternatively, you could use the Mayer-Vietoris sequence.

<u>Lemma 3.2</u> If M contains an independent set of k 2-spheres, then M contains an independent set $S = S_1 \cup \dots \cup S_k$ which is a normal surface with respect to K.

<u>Proof</u> Among all independent sets of k 2-spheres which have minimal weight, choose $S = S_1 \cup \dots \cup S_k$ to minimize

$$m(S) = \sum_{2\text{-simplices } \Delta} (\text{number of components of } S \cap \Delta).$$

Suppose there is a 2-simplex Δ such that $S \cap \Delta$ has circle components. Let C be an innermost circle of $S \cap \Delta$, bounding disk $D \subset \Delta$ with $D \cap S = C$. C separates some S_i into disks D', D''. Let S', S'' be obtained from S by replacing S_i with $D \cup D'$, $D \cup D''$ respectively, and pushing slightly to eliminate the intersection component C. One of S', S'' is independent by lemma 2.6, contradicting the minimality of $m(S)$.

Similarly, ~~the~~ suppose that for some 3-simplex Δ of K a component F of $S \cap \Delta$ is not a disk. Let C be a component of ∂F, bounding a disk $D' \subset \partial\Delta$. ~~the~~ Choose F, C and D' so that D' is as small as possible; in particular D' does not contain in its interior ~~the~~ a boundary component of any non-disk component of $S \cap \Delta$. Then $\exists$ disk $D \subset \Delta$ with $D \cap S = \partial D$. As before, construct S', S'' with $w(S') = w(S'') = w(S)$, $m(S') < m(S) > m(S'')$. Since one of S' is independent by lemma 2.6, this is a contradiction, so S must be normal.

<u>Theorem 3.3</u> (H. Kneser) Suppose M has a triangulation with t 3-simplices. If M contains an independent set of k 2-spheres, then $k < 6t + 2 \dim H_2(M : \mathbb{Z}_2)$.

<u>Proof</u> Suppose M contains k independent 2-spheres with $k \geq 6t + 2 \dim H_2(M : \mathbb{Z}_2)$. By lemma 3.2, M contains an independent set $S = S_1 \cup \dots \cup S_k$ which is a normal surface.
Let Δ be a 3-simplex of K. A component of $(\partial\Delta) \setminus S$ is called <u>good</u> if it is an annulus and contains no vertex of Δ. At most 6 components of $(\partial\Delta) \setminus S$ are bad.
A component X of $M \setminus S$ is <u>good</u> if every component of $X \cap \partial\Delta$ is good for every 3-simplex Δ of K. At most $6t$ components of $M \setminus S$ are bad.
If $k \geq 6t + 2 \dim H_2(M : \mathbb{Z}_2)$ then $M \setminus S$ has at least $6t + 1 + \dim H_2(M : \mathbb{Z}_2)$ components.
So there are at least $1 + \dim H_2(M : \mathbb{Z}_2)$

good components.

A good component X is made up of regions homeomorphic to (triangle) $\times I$ and (square) $\times I$, meeting their neighbors along (edge) $\times I$. So X is an I-bundle over a surface. <u>Every non-trivial I-bundle contributes one $\mathbb{Z}_2$ direct summand to $H_2(M;\mathbb{Z}_2)$</u>, so there is at least one trivial I-bundle X. Since ∂X consists of 2-spheres, $X \cong S^2 \times I$, contradicting the independence of S. (see pp 13, 16)

<u>Corollary</u> Every closed 3-manifold can be expressed as a connected sum of finitely many prime 3-manifolds.

<u>Theorem 3.4</u> Suppose M is a closed oriented 3-manifold. If $M \cong M_1 \# \cdots \# M_k$ and $M \cong N_1 \# \cdots \# N_k$ with M_i, N_j prime (and not S^3), then $k = l$ and, after re-numbering, $M_i \cong N_i$ by an orientation-preserving homeomorphism

<u>Proof</u> First suppose that every 2-sphere in M separates M; this implies that M_i and N_j are irreducible, by lemma 2.5. Let S be a disjoint union of 2-spheres in M such that the closures of the components of $M \setminus S$ are $M_1^*, \ldots M_k^*$, where M_i^* is

a punctured M_i (homeomorphic to the closure of $M_i \setminus$ finitely many disjoint balls).

Let Σ be a 2-sphere in M separating N_1^* from $(N_2 \# \cdots \# N_k)^*$. Make S transverse to Σ, so $S \cap \Sigma$ consists of disjoint simple closed curves on Σ. Choose S to minimize the number of components of $S \cap \Sigma$.

If $S \cap \Sigma \neq \emptyset$, let C be an innermost component of $S \cap \Sigma$, bounding a disk $D \subset \Sigma$ with $D \cap S = C = \partial D$. Suppose $D \subset M_i^*$. Since M_i is irreducible, one of the components of $M_i^* \setminus D$ has closure a punctured ball P. Let $C \subset S_j \subseteq \partial M_i^*$, and let $D' = \overline{S_j \setminus P}$. Replace S_j by $D \cup D'$, pushed slightly to eliminate the intersection component C. This contradicts the choice of S to minimize the number of components of $S \cap \Sigma$.

So $S \cap \Sigma = \emptyset$. Since M_i and N_1 are irreducible, S and Σ are isotopic (across an $S^2 \times I$ region) so $N_1 \cong$ some M_i.

Now suppose M contains a non-separating 2-sphere S. Let Σ be a disjoint set of separating 2-spheres in M such that the components of $M \setminus \Sigma$ are $M_1^*, \ldots M_k^*$. Make S, Σ transverse with minimum number of components of $S \cap \Sigma$. If $S \cap \Sigma \neq \phi$, let C be an innermost component of $S \cap \Sigma$, bounding a disk $D \subset \Sigma$. Replace S by S' or S'' as usual, whichever does not separate M, and push to eliminate the intersection component C. This contradicts minimality, so $S \cap \Sigma = \phi$. Suppose $S \subset M_i^*$; by lemma 2.5, $M_i \cong S^1 \times S^2$. Similarly one of the N_j is $S^1 \times S^2$. Repeat as often as needed to reach the first case.

A closed embedded surface $S \subseteq M$ is <u>compressible</u> if <u>either</u>

1) some component of S is contained in a ball

<u>or</u> 2) there is a disk $D \subset M$ such that $D \cap S = \partial D$ and ∂D <u>does</u> <u>not</u> bound a disk in S.

S is <u>incompressible</u> if it is not compressible. Incompressible surfaces are very important. The idea is that a compressible surface can be simplified. In case 1) we leave out any component that is contained in a ball. In case 2), D has a neighborhood in M homeomorphic to $D \times I$, meeting S in $\partial D \times I$. We replace S by $S' = \overline{S \setminus (\partial D \times I)} \cup (D \times \partial I)$; this is a 'simpler' surface than S.

cut along the compressing disc

S is incompressible if it cannot be simplified in this way.

<u>Lemma 3.5</u> An incompressible surface in an irreducible 3-manifold is isotopic to a normal surface.

<u>Proof</u> If D is a 2-disk in M with $D \cap S = \partial D$ then (since S is incompressible) ∂D bounds a disk $D' \subset S$. The sphere $D \cup D'$ bounds a ball B in M (since M is irreducible) which is disjoint from $S \setminus D'$ (since S is incompressible). Let $S'' = \overline{S \setminus D'} \cup D$; then S is isotopic to S'' (push across the ball B). Now the proof of lemma 3.2 applies to show that S is isotopic to a normal surface.

Two connected surfaces S_1, S_2 in M are <u>parallel</u> if they are disjoint and $M \setminus (S_1 \cup S_2)$ has a component X with $\partial \overline{X} = S_1 \cup S_2$ and $\overline{X} \cong S_1 \times I$ (so $S_1 \cong S_2$).

<u>Theorem 3.6</u> (W. Haken). Let M be a compact irreducible 3-manifold. There is a number $h(M)$ with the following property. If S is a closed incompressible surface in M and no two components of S are parallel, then S has $< h(M)$ components.

<u>Proof</u> Similar to Kneser's theorem 3.3; set $h(M) = 6t + 2 \dim H_2(M : \mathbb{Z}_2)$, where t is the number of 3-simplices in a triangulation K of M.

§4 Examples of 3-manifolds

1) If S is a closed surface then $M = S \times S^1$ is a closed 3-manifold. $S \times \{\text{point}\}$ is an incompressible surface in M.

2) More generally, let $h: S \to S$ be any homeomorphism (not necessarily homotopic to the identity. Let
$M = S \times I / ((x,0) \sim (h(x),1) \text{ for all } x \in S)$:
this is a closed 3-manifold, and again $S \times \frac{1}{2}$ is an incompressible surface in M.
M is called a fiber bundle over S^1 with fiber S; the projection $p: M \to S^1$ can be defined by $p(x,t) = e^{2\pi i t}$.

3) We can write $S^3 = \{(z,w) : z, w \in \mathbb{C}, |z|^2 + |w|^2 = 1\}$.
Let p, q be coprime integers, with $0 < q < p$.
Define $g: S^3 \to S^3$ by
$g(z,w) = (e^{2\pi i/p} z, e^{2\pi i q/p} w)$.
Then $G = \{1, g, g^2, \dots g^{p-1}\}$ is a cyclic group of homeomorphisms of S^3, and all elements of G other than the identity are fixed-point free. It follows that the orbit space $S^3/G = S^3/(x \sim \gamma x \text{ for all } x \in S^3, \gamma \in G)$ is a 3-manifold, called the lens space $L_{p,q}$.

4) Let W, W' be "handlebodies" of genus g, ie the solids in $\mathbb{R}^3$ bounded by the

standard genus g surfaces.

W $\qquad\qquad$ W'

Let $h: \partial W \to \partial W'$ be a homeomorphism and let $M = W \cup W' / (x \sim h(x)$ for all $x \in \partial W)$ be the space obtained by 'glueing' W and W' along their boundaries using the homeomorphism h. This is called a Heegaard splitting of M into handlebodies W, W'.

Theorem 4.1 Every closed orientable 3-manifold has a Heegaard splitting.

Proof Start with a triangulation K of a closed orientable 3-manifold M. Let K'' be the second barycentric derived subdivision of K, and let $W = N(K^1, K'')$, the union of the simplices of K'' that meet K^1. W is made from $N(K^0, K'')$, a finite disjoint union of balls, by attaching finitely many solid handles (neighborhoods of edges). W is orientable (since M is) and it follows that W is a handlebody. Similarly, $W' = \overline{M \setminus W}$ is a handlebody, and

$M = W \cup W'$, $W \cap W' = \partial W$, so M has a Heegaard splitting.

5) Let K be a knotted circle in S^3,

e.g.  (the figure eight knot).

Let N be a solid torus neighborhood of K and let $X = \overline{S^3 \setminus N}$. Then $\partial X = \partial N$ is a torus. Choose a homeomorphism $h : \partial X \to \partial(S^1 \times D^2)$ and let

$M = X \cup (S^1 \times D^2) / (y \sim h(y) \text{ for all } y \in \partial X)$.

M is a closed 3-manifold: this construction is known as <u>Dehn surgery</u> on the knot K.

<u>Every closed orientable 3-manifold can be made by Dehn surgery on some *link* of sufficiently many disjoint circles in S^3.</u>

Exercises

1) Let S be a connected surface in a 3-manifold M, let $i: S \to M$ be the inclusion map and let $a \in S$. Prove that if $i_*: \pi_1(S, a) \to \pi_1(M, a)$ is 1-1 and S is not homeomorphic to S^2 then S is incompressible in M.
(We shall prove a partial converse to this result later in the course.)

2) Prove that a disconnected surface S is incompressible in M if and only if each component of S is incompressible in M. (One way is trivial, the other requires proof.)

3) Prove that if S is incompressible in M and $M \setminus S$ is irreducible, then M is irreducible.

In problem 4 of the first set, you may assume that ∂M_i has a neighborhood in M_i homeomorphic to $(\partial M_i) \times I$ $(i = 1, 2)$.

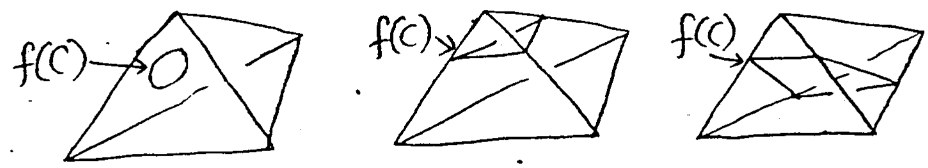
§5 PL Minimal Surfaces.

Minimal surfaces give the best method known for simplifying maps of surfaces into 3-manifolds. The method was first used by Meeks and Yau to give a new proof of the Sphere Theorem. This states that if M is an orientable 3-manifold and there is a map $f: S^2 \to M$ which is not homotopic to a point, then there is an embedding $e: S^2 \hookrightarrow M$ which is not homotopic to a point. Meeks and Yau also found stronger versions of the sphere theorem by their method.

Unfortunately the existence theorems for C^∞ surfaces of least area are very hard to prove. We base our proofs on the theory of "PL minimal surfaces", which is technically much simpler.

First we re-work the theory of normal surfaces for maps $f: S \to M$ which are not necessarily 1-1. Let K be a triangulation of a closed 3-manifold M and let S be a closed surface. Every map $f: S \to M$ can be changed by a small homotopy so that f is transverse to all the simplices of K. Define the <u>weight</u> of f to be $w(f) = |f^{-1}(K^1)|$.

Lemma 5.1 There is a map $f: S \to M$ such that
1) f has least weight among all maps in its homotopy class, and
2) for each 3-simplex Δ of K and each component C of $f^{-1}(\partial\Delta)$, $f(C)$ is a simple closed curve in $\partial\Delta$ of one of the following types.

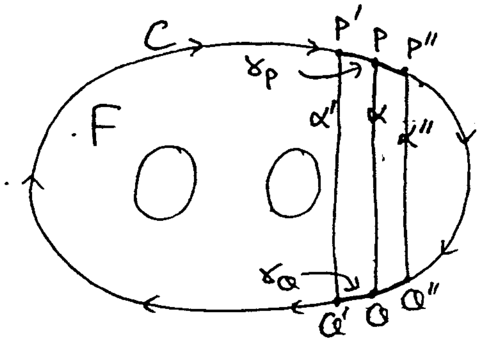


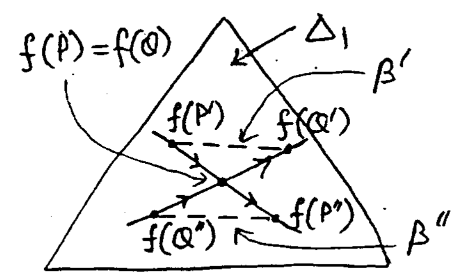
Note if C_1, C_2 are different components of $f^{-1}(\partial\Delta)$ then $f(C_1)$ and $f(C_2)$ may intersect.

Proof Choose $f: S \to M$ transverse to all simplices of K and minimizing $w(f)$. For each 2-simplex Δ^2 of K, $f^{-1}(\Delta^2)$ is a finite disjoint union of embedded arcs and circles in S, but $f|f^{-1}(\Delta^2)$ is not necessarily 1-1. We can perturb f so that, for all 2-simplices Δ^2, the self-intersections of $f(f^{-1}(\Delta))$ are transverse double points in the interior of Δ^2. Among the f of least weight, choose f to minimize

$$\sum_{\Delta^2} (\text{number of double points of } f(f^{-1}(\Delta^2))).$$

Now consider a 3-simplex Δ. Since f is transverse to $\partial\Delta$, $f^{-1}(\partial\Delta)$ is a disjoint union of simple closed curves in S. Suppose that for some component C of $f^{-1}(\partial\Delta)$, $f|C$ is not 1–1. Let F be the component of $f^{-1}(\Delta)$ that contains C as a boundary component. Suppose $P, Q \in C$ are mapped to the same point of $\partial\Delta$. Join P, Q by an embedded arc α in F. Let α', α'' be arcs in F parallel to α, close to α and on opposite sides of α. Let α' have end-points P', Q', with P' near P, Q' near Q. Similarly let α'' have end-points P'', Q''. Let γ_P, γ_Q be the short arcs of C joining $P'P''$, $Q'Q''$. Choose them so short that $f(\gamma_P)$, $f(\gamma_Q)$ are embedded arcs in a single 2-simplex $\Delta_1 \subset \Delta$, and intersecting only at $f(P) = f(Q)$.





Let A be the closed disk in F bounded by α', α'', γ_P, γ_Q. Choose embedded arcs $\beta', \beta'' \subset \Delta_1$ so that $\partial\beta' = \{f(P'), f(Q')\}$, $\partial\beta'' = \{f(P''), f(Q'')\}$ and $\beta' \cap f(C) = \partial\beta'$, $\beta'' \cap f(C) = \partial\beta''$.

We wish to replace f by a homotopic map $g : S \to M$ such that

$$g(g^{-1}(\Delta_1)) = (f(f^{-1}(\Delta_1)) \setminus (f(\gamma_P) \cup f(\gamma_Q))) \cup \beta' \cup \beta''$$

and for all other 2-simplexes Δ^2,

$$g(g^{-1}(\Delta^2)) = f(f^{-1}(\Delta_2)).$$

This will contradict the minimality of

$$\sum_{\Delta^2} |\text{double points of } f(f^{-1}(\Delta^2))|.$$

First replace f by a homotopic map $f_1 : S \to M$ such that $f_1 | \overline{S \setminus F} = f | \overline{S \setminus F}$, $f_1(\alpha') = \beta'$, $f_1(\alpha'') = \beta''$ and $f_1(A) \subset \Delta_1$. Now push $f_1(A \setminus (\alpha' \cup \alpha''))$ outside Δ, without moving $f_1 | \alpha'$ or $f_1 | \alpha''$. This homotopy can be small, and constant outside a neighborhood of $A \setminus (\alpha' \cup \alpha'')$ in S. The resulting map g is of the required type.

It follows that the map $f : S \to M$ chosen at the start of the proof maps each component C of $f^{-1}(\partial \Delta)$ to a simple closed curve in $\partial \Delta$. The rest of the proof is similar to, and simpler than, lemma 3.1.

A map $f: S \to M$ is <u>normal</u> with respect to the triangulation K if

1) f is transverse to all simplices of K,
2) for each 3-simplex Δ of K, and each component C of $f^{-1}(\partial\Delta)$, $f(C)$ has one of the following types.

3) for each 3-simplex Δ of K, each component D of $f^{-1}(\Delta)$ is a disk, and $f|D$ is 1-1.

Note that a normal map $f: S \to M$ is an immersion but not necessarily an embedding: if D_1, D_2 are two components of $f^{-1}(\Delta)$, $f(D_1)$ may intersect $f(D_2)$. $f(S)$ is often called an immersed normal surface.

Let $\mathcal{C}$ be a class of maps of connected closed surfaces into a 3-manifold M. We call $\mathcal{C}$ a <u>normal class</u> if

1) $\mathcal{C}$ is closed under homotopy,
2) $\mathcal{C}$ does not contain a null-homotopic map,

3) if $f: S \to M$ is in $\mathscr{C}$, and $f|C$ is null-homotopic for some simple closed curve $C \subset S$, then C separates S and either $f': S' \to M$ or $f'': S'' \to M$ is in $\mathscr{C}$, where f', f'' are constructed as follows. Let R', R'' be the closures of the components of $S \setminus C$, and let $S' = R' \cup D$, $S'' = R'' \cup D$, where D is a 2-disk with boundary C. The null-homotopy of $f|C$ gives an extension of $f|C$ over D, which we use to define $f': S' \to M$ and $f'': S'' \to M$.

For example,
$\mathscr{C} = \{ f: S^2 \to M : f \text{ is not null-homotopic} \}$ is a normal class. If S is a closed orientable surface of genus ≥ 1 and $\phi: \pi_1(S) \to \pi_1(M)$ is a fixed injective homomorphism, then $\mathscr{C} = \{ f: S \to M : f_* = \phi \}$ is a normal class. On the other hand, given $\phi: S^2 \to M$, $\mathscr{C} = \{ f: S^2 \to M : f \text{ is homotopic to } \phi \}$ is not usually normal.

<u>Lemma 5.2</u> Every non-empty normal class contains a normal map $f: S \to M$.

<u>Proof</u> Similar to, and simpler than, lemmas 3.2 and 3.5.

Digression: the hyperbolic plane.

In $\mathbb{R}^3$, let X be the part of the surface $z^2 - x^2 - y^2 = 1$ in the half-space $z > 0$. Give X a Riemannian metric by the formula $ds^2 = -dz^2 + dx^2 + dy^2$. One can verify the following.

1) ds^2 is a positive definite metric on X,
2) the geodesics are the intersections of X with planes through the origin in $\mathbb{R}^3$,
3) X is complete,
4) every linear transformation $T: \mathbb{R}^3 \to \mathbb{R}^3$ that leaves X invariant induces an isometry of X (we don't need this here).

X is called the hyperbolic plane, often written $\mathbb{H}^2$: it is clearly homeomorphic to $\mathbb{R}^2$. Hyperbolic n-space $\mathbb{H}^n$ can be defined similarly, starting with the hypersurface $x_{n+1}^2 - x_1^2 - \cdots - x_n^2 = 1$ in $\mathbb{R}^{n+1}$.

We only use $\mathbb{H}^2$ in the following very simple way. Let L be the cone in $\mathbb{R}^3$ defined by $z^2 - x^2 - y^2 > 0$, $z > 0$, and define projection $p: L \to X$ by $p(x,y,z) = (z^2 - x^2 - y^2)^{-\frac{1}{2}}(x,y,z)$. Let $A_1 = (1,0,1)$, $A_2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2}, 1)$, $A_3 = (-\frac{1}{2}, -\frac{\sqrt{3}}{2}, 1)$: these are in $\bar{L} \setminus L$.

If Δ is a 2-simplex with vertices v_1, v_2, v_3, there is a unique linear map $\ell : \Delta \to \mathbb{R}^3$ with $\ell(v_i) = A_i$ $(i = 1, 2, 3)$. Let $\Delta^* = \Delta \setminus \{v_1, v_2, v_3\}$: then $\ell(\Delta^*) \subset L$, so $h = p\ell : \Delta^* \longrightarrow X$ is a well-defined embedding.

The <u>hyperbolic metric</u> on Δ^* is the one corresponding to the metric ds^2 on X defined above. We need the following.

1) The shortest path between two points $a, b \in \Delta^*$ is the straight line.

2) If v is a vertex of Δ and $a \in \Delta^*$ then distance$(a, x) \to \infty$ as $x \to v$.

Corrections to proof of Th 3.4

lines 24-26 of proof: So $S \cap \Sigma = \phi$.
If $S \cap N_1^* \neq \phi$ then, since N_1 is irreducible, some component of S is parallel to Σ, and some $M_i^* \subset N_1^*$ with $\overline{N_1^* \setminus M_i^*} \cong S^2 \times I$.
If $S \cap N_1^* = \phi$ then $N_1^* \subset$ some M_i^*: since M_i is irreducible, $\overline{M_i^* \setminus N_1^*}$ is a punctured ball.
In either case $M_i \cong N_1$ and also
$M_1 \# \cdots \# M_{i-1} \# M_{i+1} \# \cdots \# M_k \cong N_2 \# \cdots \# N_\ell$.
The result follows by induction.

At the end of the proof we need to show that if $M' \# (S^1 \times S^2) \cong N' \# (S^1 \times S^2)$ then $M' \cong N'$.
This can be done (when M', N' are orientable), but it is more convenient to modify lines 27-41 of the proof as follows.
Let S be a maximal set of 2-spheres in M such that $M \setminus S$ is connected: suppose S has r components. Use the innermost circle argument to replace S by $S', S'', \ldots S^{(n)}$ so that $S^{(n)} \cap \Sigma = \phi$. Now it follows that, after re-numbering, $M_1 \cong M_2 \cong \cdots \cong M_r \cong S^1 \times S^2$.
If $S^{(n)} \times I$ is a (small) product neighborhood of $S^{(n)}$ then $\overline{M \setminus (S^{(n)} \times I)} \cong (M_{r+1} \# \cdots \# M_k)^*$.
The same applies to the factorization $M = N_1 \# \cdots \# N_\ell$. To complete the proof,

we show that $\overline{M \setminus (S^{(n)} \times I)} \cong \overline{M \setminus (S \times I)}$.
It is sufficient to show that if D is a disk in M with $D \cap S = \partial D$, and ∂D bounds a disk $D' \subset S$, and $S' = \overline{S \setminus D'} \cup D$ does not separate M, then $\overline{M \setminus (S \times I)} \cong \overline{M \setminus (S' \times I)}$.
Since S was maximal, $D \times I$ separates $\overline{M \setminus (S \times I)}$ into two components, with closures U and V (say). Now

$$\overline{M \setminus (S \times I)} \cong U \underset{D}{\cup} V \,, \quad \overline{M \setminus (S' \times I)} \cong U \underset{D'}{\cup} V \,;$$

these are homeomorphic, since both are obtained from $U \cup V$ by glueing disks in their boundaries.

§5 (continued) PL minimal surfaces.

We have defined the 'hyperbolic metric' on a 2-simplex minus its vertices. If K is a triangulation of a 3-manifold M, the metrics on the 2-simplices agree on the edges, so define a metric on $K^2 \setminus K^0$.

If $f: S \to M$ is a normal map, then $f^{-1}(K^2)$ is a graph in S, with edges $\alpha_1, \dots \alpha_n$ such that $f(\alpha_i)$ is an embedded arc in a 2-simplex of K. Define $\ell(f) = \sum_{i=1}^{n} \text{length}(f(\alpha_i))$ (using the hyperbolic metric). Roughly, $\ell(f)$ is the length of $f(S) \cap K^2$. The <u>PL area</u> $A(f)$ of f is the ordered pair $(w(f), \ell(f))$, where $w(f)$ is the weight of f. We shall say that $A(f_1) < A(f_2)$ iff either $w(f_1) < w(f_2)$ or $w(f_1) = w(f_2)$ and $\ell(f_1) < \ell(f_2)$. It is easy to see that any non-empty subset of $\{(w, \ell) : w \in \mathbb{Z}, \ell \in \mathbb{R}, w, \ell \geq 0\}$ has an infimum (greatest lower bound) with respect to this ordering.

<u>Theorem 5.3</u> If $\mathcal{C}$ is a non-empty normal class of maps of surfaces into M, there is a map $f \in \mathcal{C}$ such that $A(f) = \inf\{A(g) \mid g \in \mathcal{C}, g \text{ normal}\}$

Proof By lemma 5.2, $\mathcal{C}$ contains a normal map, so $A_0 = \inf\{A(g) \mid g \in \mathcal{C},\ g \text{ normal}\}$ is finite. Call $g : S \to M$ a *straight* normal map if, for each 2-simplex Δ^2 of K, and each component α of $g^{-1}(\Delta^2)$, $g(\alpha)$ is a geodesic arc in Δ^2. Clearly $A_0 = \inf\{A(g) \mid g \in \mathcal{C},\ g \text{ straight normal}\}$.

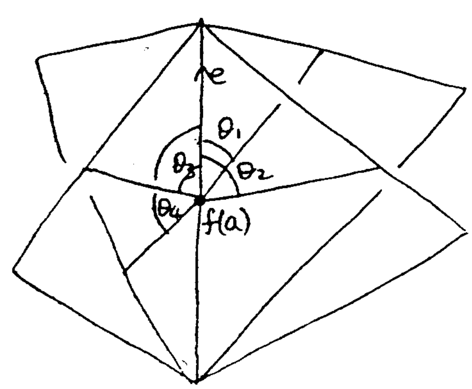
Let $g_n : S_n \to M$ be a *minimizing sequence* of straight normal maps, so $A(g_n) \to A_0$ as $n \to \infty$. Then $w(g_n)$, $\ell(g_n)$ are bounded, say $w(g_n) \le W$ and $\ell(g_n) \le L$ for all n. It follows that $g_n(S_n) \cap K^2$ is the union of at most $\binom{W}{2}$ geodesic arcs in 2-simplices (with end-points in K^1) and that the length of each arc is at most L.

Suppose that some vertex v of K is in the closure of $\bigcup_{n=1}^{\infty} g_n(S_n) \cap K^2$. Since $g_n(S_n) \cap K^2$ is connected, and is made up of $\le \binom{W}{2}$ arcs of length $\le L$, it follows that every neighborhood of v in K contains $g_n(S_n) \cap K^2$ for some n. (Recall that if $a \in K^2 \setminus K^0$ then $d(a, x) \to \infty$ as $x \to v$.) This would imply that some g_n is null-homotopic, a contradiction.

Therefore there is a compact set $C \subset K^2 \setminus K^0$ such that $g_n(S_n) \cap K^2 \subseteq C$ for all n.

Each $g_n(S_n) \cap K^2$ is made up of at most $\binom{W}{2}$ geodesic arcs in C, each embedded in a 2-simplex, with end-points on K^1. A simple compactness argument shows that there is a subsequence $g_{n_r} : S_{n_r} \to M$ and a straight normal map $f : S \to M$ such that the arcs of $g_{n_r}(S_r) \cap K^2$ converge to the arcs of $f(S) \cap K^2$. It follows that $A(f) = A_0$.

<u>Lemma 5.4</u> Suppose $f : S \to M$ has least PL area in its homotopy class. Let $a \in f^{-1}(K^1)$, let e be the edge containing $f(a)$ and let $\alpha_1, \ldots, \alpha_n$ be the arcs of $f^{-1}(K^2)$ with end-point a. Let θ_i be the angle between e (with a fixed orientation) and the geodesic $f(\alpha_i)$, measured using the hyperbolic metric on $K^2 \setminus K^0$. Then $\cos\theta_1 + \cdots + \cos\theta_n = 0$.



Proof If we move $f(a)$ a small distance h along the edge e, the length of $f(\alpha_1)$ increases by approximately $h \cos \theta_1$: the error is $O(h^2)$. The result follows.

If this condition is satisfied for all points of $f^{-1}(K^1)$, $f(S)$ is called a PL minimal surface. Minimal surfaces are "stationary points for the area function"; least area surfaces attain the absolute minimum area.

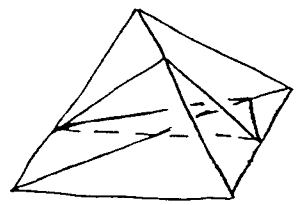
The requirement that $f: S \to M$ be PL minimal or PL least area imposes no conditions on $f(S) \setminus K^2$. Let $f_1 : S_1 \to M$, $f_2 : S_2 \to M$ be PL minimal surfaces (possibly with $f_1 = f_2$), let Δ be a 3-simplex of K and let D_i be a component of $f_i^{-1}(\Delta)$ $(i = 1, 2)$. We would like the following conditions to be satisfied.

1) $f_1(D_1)$ is transverse to $f_2(D_2)$ iff $f_1(\partial D_1)$ is transverse to $f_2(\partial D_2)$ in $\partial\Delta$.

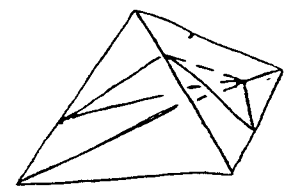
2) $f_1(D_1) = f_2(D_2)$ iff $f_1(\partial D_1) = f_2(\partial D_2)$.

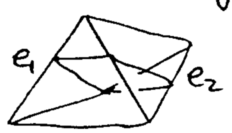
3) Every component of $f_1(D_1) \cap f_2(D_2)$ meets $\partial\Delta$.

Here is one way to achieve this. If D is a 3-sided component of $f^{-1}(\Delta)$, choose $f(D)$ to be a linear 2-simplex. If D is a 4-sided component of $f^{-1}(\Delta)$, choose $f(D)$ to be the union of two 2-simplices. There are two possible choices:



or



For each 3-simplex we make the same choice consistently, ie for each curve of type 

we use the diagonal joining edges e_1, e_2.

For any PL minimal surface $f: S \to M$ we will always make such a choice for $f(S) \setminus K^2$

<u>Corollary to Lemma 5.4</u> Suppose $f: S \to M$ and $g: T \to M$ are PL minimal surfaces. If $a \in S$, $b \in T$ and $f(a) = g(b) \in K^1$ then there are neighborhoods U of a in S, V of b in T such that <u>either</u> $f(U) = g(V)$ <u>or</u> $g(V)$ contains points on both sides of $f(U)$. (Of course, $f(U)$ doesn't separate M, but it does separate a small neighborhood of $f(a)$ in M.)

Proof Let $\alpha_1, \dots \alpha_n$ be the arcs of $f^{-1}(K^2)$ with end-point a, and let $\beta_1, \dots \beta_n$ be the arcs of $g^{-1}(K^2)$ with end-point b. Let e be the edge containing $f(a) = g(b)$ and let θ_i, ϕ_i be the (hyperbolic) angles between e and $f(\alpha_i), g(\beta_i)$. If the corollary were false we would have $\theta_i \neq \phi_i$ for some i and either $\theta_i \leq \phi_i$ for all i or $\theta_i \geq \phi_i$ for all i. This contradicts the equations

$$\sum_{i=1}^{n} \cos\theta_i = 0 = \sum_{i=1}^{n} \cos\phi_i .$$

Lemma 5.5 Suppose $f_1, f_2 : S^2 \to M$ have least PL area in the class of maps which are not null-homotopic. Suppose also that f_1, f_2 are 1-1. Then either $f_1(S^2) = f_2(S^2)$ or $f_1(S^2) \cap f_2(S^2) = \phi$.

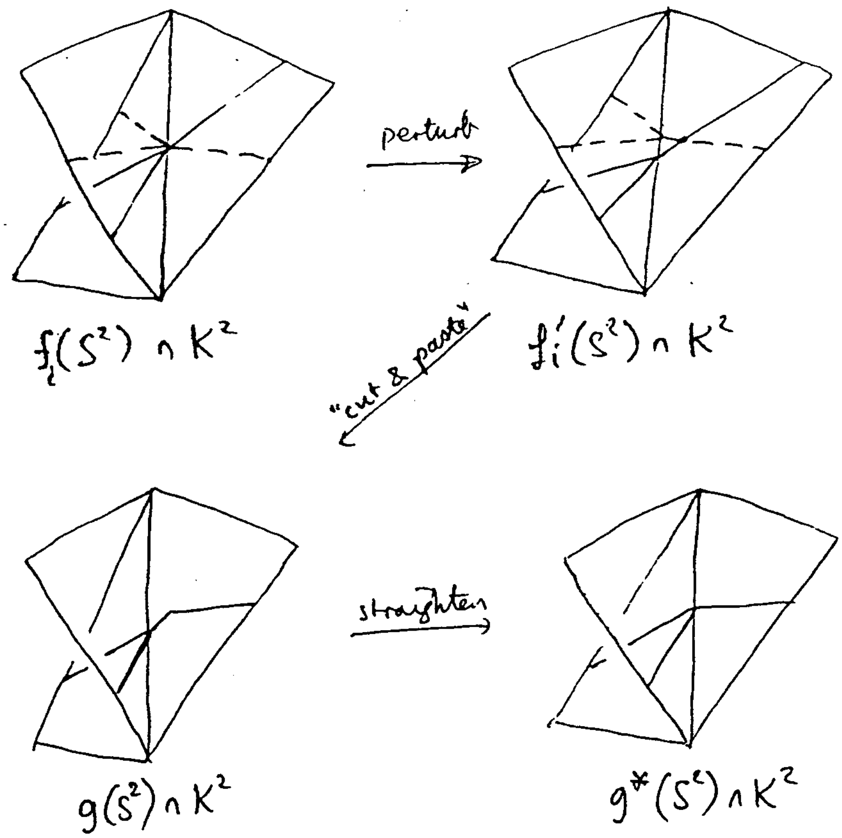
Remark This is a step in the proof of the Sphere Theorem, which implies that f_1, f_2 are automatically either 1-1 or 2-1 maps.

Proof First suppose $f_1(S^2), f_2(S^2)$ meet transversely and $f_1(S^2) \cap f_2(S^2) \cap K^1 = \phi$. Let C be a component of $f_1(S^2) \cap f_2(S^2)$, bounding disks D_i, D_i' in $f_i(S^2)$. Define $A(D_i), A(D_i')$ as for closed surfaces, and suppose $A(D_2) \leq A(D_1), A(D_1'), A(D_2')$.

Let $g, g' : S^2 \to M$ be immersions with images $D_1 \cup D_2$, $D_1' \cup D_2$. One of g, g' is not null-homotopic, say g. If g is not a normal map then there is a normal map $g^* : S \to M$ with $w(g^*) < w(g)$ and g^* not null-homotopic. This contradicts the minimality of $A(f_1)$, since $w(g) = w(D_1) + w(D_2) \le w(D_1) + w(D_1') = w(f_1)$. If g is a normal map, then $w(g) \le w(f_1)$ and (similarly) $\ell(g) \le \ell(f_1)$. But g is not a straight normal map, so it is homotopic to a straight normal map g^* with $\ell(g^*) < \ell(g)$. Again this contradicts the minimality of $A(f_1)$.

If $f_1(S^2)$ and $f_2(S^2)$ are not transverse (or if $f_1(S^2) \cap f_2(S^2)$ meets K^1) we use the "Meeks–Yau trick". Arbitrarily close to $f_1, f_2 : S^2 \to M$ we can find straight normal maps $f_1', f_2' : S^2 \to M$ such that $f_1'(S^2), f_2'(S^2)$ are transverse and $f_1'(S^2) \cap f_2'(S^2) \cap K^1 = \emptyset$. We apply the above argument to f_1', f_2'. If $f_1(S^2) \neq f_2(S^2)$ we show that there is a $\delta > 0$, independent of the (small) distance between f_i and f_i', such that $\ell(g) - \ell(g^*) > \delta$. If f_i' is sufficiently close to f_i, $\ell(f_i') < \ell(f_i) + \delta$, so we still have $\ell(g^*) < \ell(f_i)$.

A "worst case" is illustrated below.



So far, we may not have decreased $\ell(g)$ very much. However, the straight normal map g^* cannot be PL minimal (see lemma 5.4), so $\ell(g^*)$ can be further reduced by a homotopy through straight normal maps. This decreases $\ell(g^*)$ by an amount depending on the angles between the arcs of $f_i(S^2) \cap K^2$ and the edges of K, not on the size of the perturbation from f_i to f_i'.

Exercises (For these you should use the loop and sphere theorems)

1) Prove that if $S \not\cong S^2$ is an embedded surface in $\mathbb{R}^3$ then there is a disk $D \subset \mathbb{R}^3$ such that $D \cap S = \partial D$ and ∂D is an essential curve in S.

2) Prove that if M is a compact, connected, irreducible 3-manifold with $\partial M \cong$ torus, and the map $\pi_1(\partial M) \to \pi_1(M)$ induced by inclusion is onto, then $M \cong S^1 \times D^2$.

[Hint: the map $H_1(\partial M) \to H_1(M)$ induced by inclusion cannot be 1-1.

Proof Consider the exact sequence of $(M, \partial M)$:

$$H_2(M) \to H_2(M, \partial M) \to H_1(\partial M) \to H_1(M) \to 0.$$

By Poincaré-Lefschetz duality,

$H_2(M) \cong H^1(M, \partial M) = 0$ and

$H_2(M, \partial M) \cong H^1(M) \cong \mathrm{Hom}(H_1(M), \mathbb{Z})$.

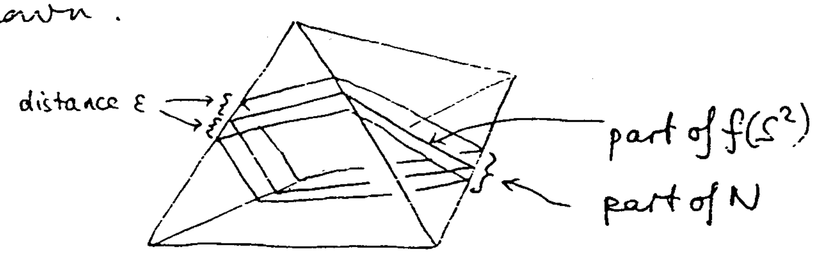
From this it follows that the rank (ie the number of $\mathbb{Z}$ summands) of $H_1(\partial M)$ is twice the rank of $H_1(M)$. If you are not familiar with Poincaré duality, just assume the hint above.]

3) Generalize 2) to the case $\partial M \cong$ closed orientable surface of genus $g > 1$.

§6 The Sphere Theorem

Theorem 6.1 If M is a closed 3-manifold with $\pi_2(M) \neq 0$, and $f: S^2 \to M$ has least PL area among maps $S^2 \to M$ which are not null-homotopic, then either f is an embedding or f double covers an embedded projective plane.

Proof By 5.3, such an f exists, and is a straight normal map. First assume that f is self-transverse, and the double points of f are disjoint from K^1. Let N be a thin regular neighborhood of $f(S^2)$, constructed as shown.



(We make the same convention about $N \setminus K^2$ as for $f(S^2) \setminus K^2$.) If the 'radius' ε of N is small enough, N is a compact 3-manifold and $f(S^2)$ is a deformation retract of N.

Now we use the famous 'tower construction' introduced by Papakyriakopoulos in 1957. Set $f_0 = f$ and $N_0 = N$. Suppose we have already constructed a compact 3-manifold N_{i-1} and a map $f_{i-1}: S^2 \to N_{i-1}$. Let $\tilde{N}_{i-1}$ be the universal covering of N_{i-1}. If this is a non-trivial covering (ie if $\pi_1(N_{i-1}) \neq \{1\}$), let $f_i : S^2 \longrightarrow \tilde{N}_{i-1}$ be a lift of f_{i-1}. Let N_i be a regular neighborhood of $f_i(S^2)$ in $\tilde{N}_{i-1}$. If N_{i-1} is simply connected, the tower terminates.

$$\begin{tikzcd}
& N_k \arrow[r, phantom, "\subset"] & \tilde{N}_{k-1} \arrow[dl] \\
& N_{k-1} \arrow[r, phantom, "\subset"] & \tilde{N}_{k-2} \\
& \vdots & \vdots \\
& N_2 \arrow[r, phantom, "\subset"] & \tilde{N}_1 \arrow[dl] \\
& N_1 \arrow[r, phantom, "\subset"] & \tilde{N}_0 \arrow[dl] \\
S^2 \arrow[r, "f_0 = f"] \arrow[ur, "f_1"] \arrow[uur, "f_2"] \arrow[uuuur, "f_{k-1}"] \arrow[uuuuur, "f_k"] & N_0 \arrow[r, phantom, "\subset"] & M
\end{tikzcd}$$

<u>Step 1</u> For some k, N_k is simply connected, so the tower terminates.

<u>Proof</u> Let $\Sigma_i = \{x \in S^2 : f_i(x) = f_i(y) \text{ for some } y \neq x\}$. Since f (therefore f_i) is self-transverse, Σ_i is a finite graph (ie 1-complex) in S^2. Clearly Σ_{i+1} is a subgraph of Σ_i. If $\Sigma_{i+1} = \Sigma_i$ then the covering $\tilde{N}_i \to N_i$ induces a homeomorphism $f_{i+1}(S^2) \to f_i(S^2)$. This implies that $\pi_1(\tilde{N}_i)$ maps onto $\pi_1(N_i)$, so $\tilde{N}_i$ is a trivial covering and N_i is in fact simply connected. Since the sequence

$\Sigma_0 \supset \Sigma_1 \supset \cdots$ must stabilize, say with $\Sigma_k = \Sigma_{k+1}$, N_k is simply connected for some k. We choose the least such k, so that $\pi_1(N_i) \neq \{1\}$ for $i < k$ and $\pi_1(N_k) = \{1\}$

<u>Step 2</u> If $f_k : S^2 \to N_k$ is 1-1, then $k = 0$, implying that f is an embedding.

<u>Proof</u> If $k \geq 1$, we can consider f_k as a map from S^2 to $\tilde{N}_{k-1}$. Let G be the group of covering translations of $\tilde{N}_{k-1}$. If $\gamma \in G \setminus \{1\}$ then $f_k(S^2)$, $\gamma f_k(S^2)$ are embedded 2-spheres in $\tilde{N}_{k-1}$ that project to PL least area maps in M. Since they are transverse, it follows as in lemma 5.5 that $f_k(S^2) \cap \gamma f_k(S^2) = \emptyset$. This holds for all $\gamma \in G \setminus \{1\}$, implying that $f_{k-1} : S^2 \to N_{k-1}$ is 1-1. But then $N_{k-1} \simeq S^2$ is simply connected, a contradiction.

<u>Step 3</u> $f_k : S^2 \to N_k$ is 1-1.

We need two propositions from algebraic topology.

<u>Proposition 1</u> If W is a compact 3-manifold with $\pi_1(W) = \{1\}$, then all components of ∂W are spheres.

<u>Proof</u> Since W is simply connected, it is orientable. By the Poincaré-Lefschetz duality theorem, $H_2(W, \partial W) \cong H^1(W) = 0$. Now the exact sequence of the pair $(W, \partial W)$ includes

$H_2(W, \partial W) \longrightarrow H_1(\partial W) \longrightarrow H_1(W)$.

Since $H_2(W, \partial W) = 0 = H_1(W)$, $H_1(\partial W) = 0$, implying that ∂W consists of spheres.

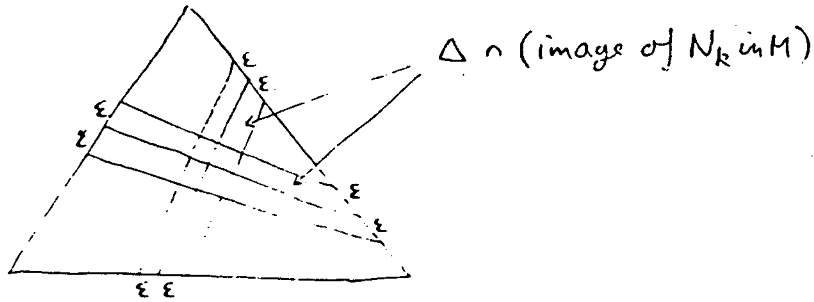
<u>Proposition 2</u> If W is a compact 3-manifold with $\pi_1(W) = \{1\}$ then $\pi_2(W)$ is generated by the homotopy classes of the components of ∂W.

<u>Proof</u> Since W is simply connected, $\pi_2(W) \cong H_2(W)$ by the Hurewicz theorem. The exact sequence $H_2(\partial W) \to H_2(W) \to H_2(W, \partial W)$, in which $H_2(W, \partial W) = 0$, shows that $H_2(\partial W)$ maps onto $H_2(W)$.

<u>Corollary</u> Some component of ∂N_k is not null-homotopic. In fact, since the sum of all the components of ∂N_k is zero in $\pi_2(N_k)$, at least two components of ∂N_k are essential in M.

To complete the proof (when f is self-transverse) we show that if f_k is not 1-1 then the total PL area of ∂N_k is strictly less than $2A(f)$. This would imply that at least one component of ∂N_k is an essential sphere in M with PL area less than $A(f)$, a contradiction.

If f_k is self-transverse and not 1-1, there is a 2-simplex Δ where two arcs of $f_k(S^2)$ (projected into M) meet transversely.



The image of N_k in M meets Δ in (at least) two intersecting strips. Since the part of the boundary of one strip that is in the interior of the other strip is not in ∂N_k, and since f_k is a PL minimal surface, we find that

$\ell(\partial N_k) = 2\ell(f) - (\text{constant}) \cdot \varepsilon + o(\varepsilon)$.

So if ε is small enough, $\ell(\partial N_k) < 2\ell(f)$

Since $w(\partial N_k) \leq 2w(f)$, $A(\partial N_k) < 2A(f)$.

Finally, if f is not self-transverse we apply the Meeks-Yau trick. This succeeds unless f is a covering of an embedded surface, necessarily a double covering of a projective plane.

A surface S in M is <u>2-sided</u> if it has a neighborhood homeomorphic to $S \times I$. We call M <u>P^2-irreducible</u> if M is irreducible and contains no 2-sided projective plane. For orientable manifolds, P^2-irreducible $\Leftrightarrow$ irreducible; but $P^2 \times S^1$ is irreducible but not P^2-irreducible.

<u>Corollary 1</u> If M is P^2-irreducible then $\pi_2(M) = 0$.

Proof If $\pi_2(M) \neq 0$ then M contains an essential embedded S^2 or P^2. An S^2 or 2-sided P^2 would contradict P^2-irreducibility. So M contains a 1-sided P^2, with a neighborhood N which is a twisted I-bundle over P^2, with $\partial N \cong S^2$. Since M is irreducible, ∂N bounds a ball in M, so $M \cong P^3$; but $\pi_2(P^3) = 0$.

Corollary 2 If M is orientable and $\pi_2(M) \neq 0$ then M contains an essential embedded S^2.

Proof Similar.

Corollary 3 If M is a closed P^2-irreducible 3-manifold then the universal cover $\tilde{M}$ satisfies:

a) if $\pi_1(M)$ is finite, $\tilde{M}$ is homotopy equivalent to S^3

b) if $\pi_1(M)$ is infinite, $\tilde{M}$ is contractible.

Proof $\pi_1(\tilde{M}) = \pi_2(\tilde{M}) = 0$ in both cases.
If $\pi_1(M)$ is infinite, $\tilde{M}$ is non-compact, so $H_3(\tilde{M}) = 0$, and clearly $H_i(\tilde{M}) = 0$ for $i > 3$. By the Hurewicz theorem, $\pi_i(\tilde{M}) = 0$ for $i \geq 3$, in fact for all i. By the Whitehead theorem, $\tilde{M}$ is contractible.
If $\pi_1(M)$ is finite, $\tilde{M}$ is a closed 3-manifold, orientable since $\pi_1(\tilde{M}) = \{1\}$. So $H_3(\tilde{M}) \cong \mathbb{Z}$, and by the Hurewicz theorem, $\pi_3(\tilde{M}) \cong \mathbb{Z}$. Let $f: S^3 \to \tilde{M}$ represent a generator of $\pi_3(\tilde{M})$. Both S^3, $\tilde{M}$ are simply connected, and f induces isomorphisms in homology groups of all dimensions. By the Hurewicz and

Whitehead theorems, f is a homotopy equivalence.

Corollary 3 gives a lot of information about the possible fundamental groups of closed P^2-irreducible 3-manifolds. One consequence is that if M is P^2-irreducible and $\pi_1(M)$ is infinite then $\pi_1(M)$ is torsion-free.

<u>Conjecture</u> If M is a closed P^2-irreducible 3-manifold with $\pi_1(M)$ infinite then $\tilde{M} \cong \mathbb{R}^3$.

This is known to be true if $\pi_1(M) \supset \pi_1(S)$, where S is a closed orientable surface of genus ≥ 1. The following theorem is a substantial contribution to this problem.

<u>Theorem 6.2</u> (Meeks, Simon, Yau) If M is a P^2-irreducible 3-manifold then the universal covering $\tilde{M}$ of M is irreducible.

<u>Proof</u> Let K be a triangulation of M and let $\tilde{K}$ be the 'lifted' triangulation of $\tilde{M}$ (in general, $\tilde{K}$ is an infinite complex). The group $G \cong \pi_1(M)$ of covering translations acts simplicially on $\tilde{K}$.

Let $\mathscr{C}$ be the class of 1-1 maps $g: S^2 \to \tilde{M}$ such that $g(S^2)$ does not bound a ball in $\tilde{M}$. If $\tilde{M}$ is not irreducible, $\mathscr{C}$ is non-empty.

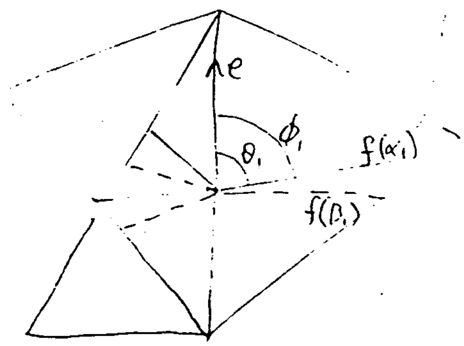
As in lemmas 3.1, 3.2, $\mathcal{C}$ contains a normal embedding $g : S^2 \longrightarrow \tilde{M}$. The proof of Theorem 5.3 shows that there is a straight normal map $f : S^2 \longrightarrow \tilde{M}$ that is in the closure of $\mathcal{C}$ (so f can be approximated arbitrarily closely by essential embeddings) with

$$A(f) = \inf \{ A(g) \mid g \in \mathcal{C}, g \text{ normal} \}.$$

(Although $\tilde{M}$ is non-compact, there is a compact subset $C \subset \tilde{M}$ such that $\tilde{M} = \bigcup_{\gamma \in G} \gamma C$: this is sufficient for the proof of 5.3.)

Suppose f is not 1-1. Then $\exists\ a, b \in S^2$ such that $f(a) = f(b) \in K^1$ (since f can be approximated by embeddings). Let $\alpha_1, \ldots \alpha_n$ be the arcs of $f^{-1}(K^2)$ with end-point a, let $\beta_1, \ldots \beta_n$ be the arcs of $f^{-1}(K^2)$ with end-point b. Let e be the edge containing $f(a) = f(b)$. If θ_i is angle between e and $f(\alpha_i)$, and ϕ_i is angle between e and $f(\beta_i)$, we have (without loss of generality) $\theta_i \leq \phi_i$ for all i and $\theta_i \neq \phi_i$ for some i (if a, b chosen suitably: otherwise f would be a covering of a projective plane, contradicting simple connectivity of $\tilde{M}$).

So $$\sum_{i=1}^{n} \cos\theta_i > \sum_{i=1}^{n} \cos\phi_i .$$



It follows that $l(f)$ can be reduced by shifting $f(a)$ up or $f(b)$ down on e. This removes the self-intersection $f(a) = f(b)$. By repeating this operation we can replace f by an embedding with smaller PL area, a contradiction. So $f: S^2 \to \tilde{M}$ is an embedding. If $f(S^2)$ bounds a ball in $\tilde{M}$, it is contained in the interior of a slightly larger ball $B \subset \tilde{M}$. Since f can be approximated arbitrarily closely by members of $\mathcal{C}$, $\exists\, g \in \mathcal{C}$ with $g(S^2) \subset B$. By the Schoenflies theorem, $g(S^2)$ bounds a ball, a contradiction. It follows that $f \in \mathcal{C}$, and has least PL area of all members of $\mathcal{C}$.

First suppose that, for all $\gamma \in G \setminus \{1\}$, $f(S^2)$ and $\gamma f(S^2)$ are transverse, and their intersection is disjoint from K^1. We shall show that $f(S^2) \cap \gamma f(S^2) = \emptyset$. If not, let D be a disk in $f(S^2)$ or $\gamma f(S^2)$ with $\partial D \subset f(S^2) \cap \gamma f(S^2)$, chosen so that $A(D)$

is as small as possible. Suppose $D \subset \gamma f(S^2)$. Then D is innermost, ie $D \cap f(S^2) = \partial D$. Let ∂D bound disks $D', D'' \subset f(S^2)$. Let $f', f'' : S^2 \to \tilde{M}$ be embeddings with images $D \cup D', D \cup D''$. At least one is essential, and as in lemma 5.5 we can find an embedding $f^* : S^2 \to \tilde{M}$ with $A(f^*) < A(f)$, a contradiction. So $f(S^2) \cap gf(S^2) = \phi$ for all $g \in G \setminus \{1\}$. This implies that f projects to an embedded sphere in M that does not bound a ball, contradicting irreducibility of M.

If the intersections are not transverse, we use the Meeks-Yau trick to conclude that for all $\gamma \in G$, $f(S^2)$ and $\gamma f(S^2)$ are either equal or disjoint. This implies that $f(S^2)$ projects to an essential S^2 or P^2 embedded in M, contradicting P^2-irreducibility.

<u>Theorem 6.3</u> (Loop theorem) Let M be a 3-manifold, let S be a connected surface contained in ∂M and suppose that the map $\pi_1(S) \to \pi_1(M)$ induced by inclusion is <u>not</u> 1-1. Then there is an embedded disk $D \subset M$ such that $\partial D \subset S$ and ∂D does not bound a disk in S.

<u>Proof</u> We have to adapt the theory of PL minimal surfaces to manifolds with boundary. Surprisingly little needs to be changed. We use the same definition of 'normal map' and 'PL area'. Triangulate M by a complex K such that S is a subcomplex of K. Let $\mathcal{C}$ be the class of maps $g : D \longrightarrow M$ such that $g(\partial D) \subset S$ and $g|\partial D$ is not null-homotopic in S. Then $\mathcal{C}$ is a normal class in the usual sense, and in addition, the following condition is satisfied. Suppose α is an arc in D with end-points in ∂D, and D', D'' are the closures of the components of $D \setminus \alpha$. If $g \in \mathcal{C}$ is such that $g|\alpha$ is homotopic (rel $\partial\alpha$) to an arc in S, define $g' : D \to M$ by $g'|D' = g|D'$, $g'|D'' =$ homotopy from $g|\alpha$ into S. Similarly define $g'' : D \to M$; then one of g', g'' is in $\mathcal{C}$.

This allows us to show that $\mathcal{C}$ contains normal maps. The existence proof for PL least area maps is just as in Theorem 5.3. Now we use a tower argument that is similar to, but simpler than, the one for the sphere theorem.

<u>Corollary 1</u> (Dehn's lemma) If C is a simple closed curve in ∂M that is null-homotopic in M, then C bounds an embedded disk in M.

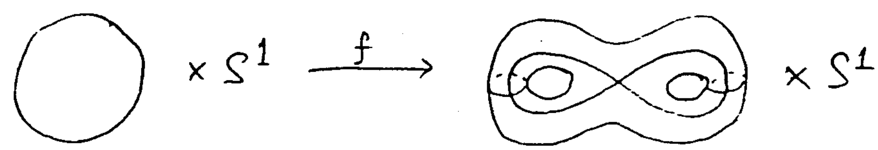
<u>Proof</u> Since C is null-homotopic in M, it is an orientation-preserving curve and has a neighborhood S in ∂M homeomorphic to $C \times I$. Now apply the loop theorem ~~with~~ to $S \subset \partial M$.

<u>Corollary 2</u> Let S be a 2-sided incompressible surface in M. Then the inclusion map $\pi_1(S) \to \pi_1(M)$ is 1-1.

<u>Proof</u> Let $f: S^1 \to S$ be a map which is null-homotopic in M, so f extends to a map $g: D^2 \to M$. If $g^{-1}(S) = \partial D^2$, notice that $g(D^2)$ intersects a small neighborhood of S only on one side of S. Embed $S \times I$ in M with $S \times 0 = S$ and $S \times 1$ on the other side of S from $g(D^2)$. Now the loop theorem, applied to $\overline{M \setminus S \times I}$, shows that f is null-homotopic. In general, make g transverse to S, minimizing the number of components of $g^{-1}(S)$. Let C be an innermost component, bounding disk $D' \subset D^2$. Then $g|C \to S$ is null-homotopic, so we can replace $g(D')$ by a disk close to S, reducing number of components of $g^{-1}(S)$. It follows that $f: S^1 \to S$ is null-homotopic.

§7 The Torus Theorem

Suppose M is a closed orientable irreducible 3-manifold and $\pi_1(M)$ contains a subgroup Γ isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$. Then there is a map $f: T \to M$, where T is a torus, such that $f_*: \pi_1(T) \to \pi_1(M)$ is 1-1 and has image Γ. It is not reasonable to expect f to be homotopic to an embedding, as the following example shows.



A good question is whether f is homotopic to a map without triple points.

The class $\mathcal{C}$ of maps $g: T \to M$ such that g_* maps $\pi_1(T)$ isomorphically onto Γ is normal, so by Theorem 5.3 it contains a map $f: T \to M$ of least PL area.

Theorem 7.1 Let M be a closed orientable irreducible 3-manifold, with universal covering $\tilde{M}$. If $f: T \to M$ has least PL area in $\mathcal{C} = \{ g: T \to M \mid g_*: \pi_1(T) \cong \Gamma \}$, then

a) every lift $\tilde{f} : \tilde{T} \to \tilde{M}$ of f is 1–1,

b) if $\tilde{f}_1, \tilde{f}_2$ are two lifts of f, then $\tilde{f}_1(\tilde{T}), \tilde{f}_2(\tilde{T})$ are either equal, disjoint or intersect transversely in a line (ie intersection is homeomorphic to $\mathbb{R}$).

This is a special case of a theorem of Freedman, Hass and Scott (Inventiones Mathematicae 71 (1983) pages 609–642). Their theorem applies to arbitrary surfaces, but the conclusion is more complicated to state.

<u>Proof</u> (outline)

<u>Step 1</u> Let $\hat{M}$ be the covering of M with $\pi_1(\hat{M}) = \Gamma$, and let $\hat{f} : T \to \hat{M}$ be a lift of f. Then $\hat{f}$ is 1–1 (this implies conclusion a).

<u>Proof</u> Tower argument, sketched below. For more details, see the paper of Freedman, Hass, Scott quoted above, in particular §2. Let $\tilde{N}_i$ be a double covering of N_i to which f_i lifts, if there is one (otherwise, tower stops!). Let f_{i+1} be a lift of f_i to $\tilde{N}_i$, let N_{i+1} be a regular neighborhood of $f_{i+1}(T)$ in $\tilde{N}_i$. Tower does stop, as in proof of sphere theorem.

$$\begin{tikzcd}
 & N_k \arrow[r, phantom, "\subset"] \arrow[d, phantom, "\vdots"] & \tilde{N}_{k-1} \arrow[d, phantom, "\vdots"] \\
 & {} & {} \\
 & N_2 \arrow[r, phantom, "\subset"] & \tilde{N}_1 \arrow[dl] \\
 & N_1 \arrow[r, phantom, "\subset"] & \tilde{N}_0 \arrow[dl] \\
T^2 \arrow[r, "f_0 = \hat{f}"'] \arrow[ur, "f_1"] \arrow[uur, "f_2"] \arrow[uuuur, dashed, "f_k"] & N_0 \arrow[r, phantom, "\subset"] & \hat{M}
\end{tikzcd}$$

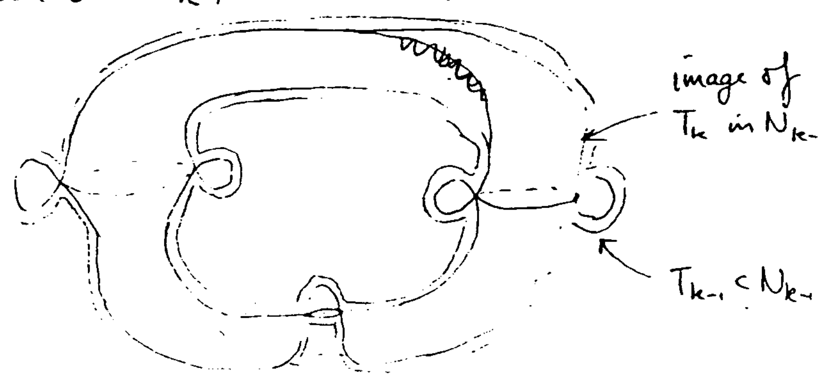
Since N_k has no double covering to which f_k lifts, $f_{k*} : H_1(T^2 : \mathbb{Z}_2) \to H_1(N_k : \mathbb{Z}_2)$ is onto. It is also 1-1 since it projects to $\hat{f}_* : H_1(T : \mathbb{Z}_2) \to H_1(\hat{M} : \mathbb{Z}_2)$, which is an isomorphism.
Poincaré-Lefschetz duality gives
$H_2(N_k, \partial N_k : \mathbb{Z}_2) \cong H^1(N_k : \mathbb{Z}_2) \cong H^1(T : \mathbb{Z}_2)$.
Exact sequence of pair $(N_k, \partial N_k)$ gives
$H_2(N_k, \partial N_k : \mathbb{Z}_2) \to H_1(\partial N_k : \mathbb{Z}_2) \to H_1(N_k : \mathbb{Z}_2)$
exact, so $\dim H_1(\partial N_k : \mathbb{Z}_2) \leq 4$. We wish to show that ∂N_k is a union of arbitrarily many spheres and exactly two tori.

If $x, y \in N_k \setminus f_k(T^2)$, define $i(x,y) \in \mathbb{Z}_2$ by joining x, y by an arc α transverse to $f_k(T^2)$ and setting $i(x,y) = |\alpha \cap f_k(T^2)| \pmod 2$. This is unchanged if α is varied by a homotopy (rel $\partial\alpha$). Using the fact that

$f_{k*}: H_1(T:\mathbb{Z}_2) \to H_1(N_k:\mathbb{Z}_2)$ is an isomorphism, one can ~~show that $i(x,y)$ is completely independe~~ of the choice of α.

Choose $a \in \partial N_k$ and let W_k be the closure of $\{x \in N_k \setminus f_k(T^2) : i(a,x) = 0\}$. Let $X_k = W_k \cap \partial N$. Thinking of W_k as a 3-chain (mod 2), $\partial W_k = X_k - f_k(T^2)$, showing that X_k represents the same element of $H_2(N_k:\mathbb{Z}_2)$ as $f_k(T^2)$. Similarly, $X_k' = \partial N_k \setminus X_k$ represents the same element. Since $\pi_2(\hat{M}) = 0$, and X_k, X_k' project to non-trivial homology classes in $H_2(\hat{M}:\mathbb{Z}_2)$, X_k, X_k' each contain one torus component, say T_k, T_k'. As in the proof of the sphere theorem, at least one, say T_k, has smaller PL area than $f(T)$.

Now consider the projection of T_k in N_{k-1}. Since $\tilde{N}_{k-1}$ is a double covering, T_k is immersed in N_{k-1} with disjoint double curves



Cut and paste along double curves to give embedded torus $T_{k-1} \subset N_{k-1}$, with smaller PL area than T_k and representing the same homology class in $H_2(\hat{M} : \mathbb{Z}_2)$. Continue down the tower until we get embedded torus $T_0 \subset \hat{M}$ with $A(T_0) < A(\hat{f})$ and representing non-trivial $\mathbb{Z}_2$ homology class. One shows that the map $\pi_1(T_0) \to \pi_1(\hat{M}) \cong \Gamma$ induced by inclusion is onto, so this contradicts the minimality of $A(f)$.

<u>Step 2</u> Each lift $\tilde{f} : \tilde{T} \to \tilde{M}$ is least area, in the following sense. If D is a compact surface <u>in $\tilde{T}$</u> and $f' : \tilde{T} \to \tilde{M}$ agrees with $\tilde{f}$ on $\overline{\tilde{T} \setminus D}$, then $A(f'(D)) \geq A(\tilde{f}(D))$.

<u>Proof</u> Let $p : \tilde{T} \to T$ be the universal covering. The conclusion of step 2 is easy to prove for surfaces $D \subset \tilde{T}$ such that $p|D$ is 1-1; we could use a map $f' : D \to \tilde{M}$ that agrees with $\tilde{f}$ on ∂D and has $A(f'(D)) < A(\tilde{f}(D))$ to reduce the area of $f : T \to M$.

In general, even if $p|D$ is not 1-1, there is a finite covering T_1 of T, with universal covering $p_1 : \tilde{T} \to T_1$, such that $p_1|D$ is 1-1. So the proof of Step 2 reduces to proving that, if $q : T_1 \to T$ is a finite covering, then $fq : T_1 \longrightarrow M$ is a least area map.

Let $\Gamma_1 = f_* q_* (\pi_1(T_1)) \subset f_*(\pi_1(T)) = \Gamma$, and let $\hat{M}_1$ be the covering of M with $\pi_1(\hat{M}_1) = \Gamma_1$. $fq : T_1 \to M$ is homotopic to a least area map $g : T_1 \to M$, which lifts to $\hat{g} : T_1 \to \hat{M}_1$. By Step 1, $\hat{g}$ is an embedding, and $i(x,y) \in Z_2$ can be defined for $x, y \in \hat{M}_1 \setminus \hat{g}(T_1)$, showing that $\hat{M}_1 \setminus \hat{g}(T_1)$ has 2 components.

The covering $r : \hat{M}_1 \longrightarrow \hat{M}$ has a finite group of covering translations, isomorphic to Γ/Γ_1. If γ is a covering translation, an argument similar to 5.5 shows that $\hat{g}(T_1)$, $\gamma\hat{g}(T_1)$ are either equal or disjoint. (For details, see lemma 1.3 of the paper by Freedman, Hass and Scott.) Therefore $r\hat{g} : T_1 \to \hat{M}$ is a covering of $g'(T)$ for some embedding $g' : T \to \hat{M}$. g' must induce a 1-1 map $\pi_1(T) \longrightarrow \pi_1(\hat{M}) = \Gamma$: in fact one can show that $g'_*(\pi_1(T)) = \Gamma$. Since f was least area, $A(g') \geq A(f)$, implying that $A(g) = |\Gamma/\Gamma_1| A(g') \geq |\Gamma/\Gamma_1| A(f) = A(fq)$. So $fq : T_1 \to M$ is least area, completing Step 2.

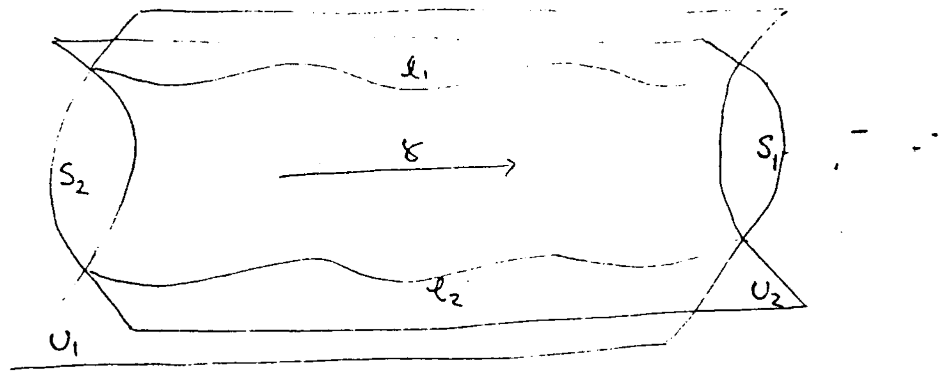
Step 3 If $\tilde{f}_1, \tilde{f}_2 : \tilde{T} \to \tilde{M}$ are lifts of f, then $\tilde{f}_1(\tilde{T}) \cap \tilde{f}_2(\tilde{T})$ has no compact component.

Proof Follows from Step 2 (using the Meeks–Yau trick).

Step 4 If $\tilde{f}_1, \tilde{f}_2 : \tilde{T} \to \tilde{M}$ are lifts of f, and $\tilde{f}_1(\tilde{T}), \tilde{f}_2(\tilde{T})$ are not equal or disjoint, then $\tilde{f}_1(\tilde{T}) \cap \tilde{f}_2(\tilde{T}) \cong \mathbb{R}$.

Proof Let $U_i = \tilde{f}_i(\tilde{T}) \subset \tilde{M}$, and first suppose U_1, U_2 are transverse. The covering translation group G of $\tilde{M}$ has subgroups $\Gamma_{U_1}, \Gamma_{U_2} \cong \mathbb{Z} \oplus \mathbb{Z}$ such that Γ_{U_i} leaves U_i invariant. Γ_{U_i} is the group of translations of the covering $U_i \xrightarrow{(\tilde{f}_i)^{-1}} \tilde{T} \longrightarrow T$.

Any (non-compact) component of $U_1 \cap U_2$ is a covering of an essential curve in T, therefore is invariant under an infinite cyclic subgroup of $\Gamma_{U_1} \cap \Gamma_{U_2}$. If $U_1 \cap U_2$ has more than one component, we can find a strip $S_1 \subset U_1$ bounded by two components l_1, l_2 of $U_1 \cap U_2$, with $S_1 \cap U_2 = \partial S_1$. Let S_2 be the strip in U_2 bounded by l_1, l_2. S_1, S_2, l_1, l_2 are invariant under some infinite cyclic subgroup $\langle \gamma \rangle \subset \Gamma_{U_1} \cap \Gamma_{U_2}$.



There is a finite covering $q : T_1 \to T$ such that $p_1 (\tilde{f}_i)^{-1} : U_i \to T_1$ restricts to a covering map on the strip S_i. As in steps 2, 3, we can now decrease the area of $fq : T_1 \to M$ by exchanging S_1 and S_2.

If U_1, U_2 are not transverse we use the Meeks-Yau trick. It turns out that the area can always be reduced in this case, and we can conclude that the least area map $f : T \to M$ is self-transverse. See §6 of the paper by Freedman, Hass and Scott for details.

Theorem 7.2 If M is a closed orientable irreducible 3-manifold and $\pi_1(M)$ contains a subgroup Γ isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$ then either M contains an embedded incompressible torus or $\pi_1(M)$ contains a normal subgroup isomorphic to $\mathbb{Z}$.

This form of the torus theorem is due to Peter Scott. The only known 3-manifolds M for which $\pi_1(M)$ contains an infinite cyclic normal subgroup are the Seifert fibered spaces. Let S be a closed surface and let $D_1, \ldots D_k$ be disjoint disks in S. Set $M = \overline{S \setminus (D_1 \cup \cdots \cup D_k)} \times S^1 \cup k$ solid tori, where the solid tori are attached by homeomorphisms of their boundaries (that do not make point $\times S^1$ bound disks in any solid tori). This is not quite the most general Seifert fibered space; we could start with a non-orientable S^1-bundle over $\overline{S \setminus D_1 \cup \cdots \cup D_k}$.

Conjecture Under the hypotheses of Theorem 7.2 either M contains an embedded incompressible torus or M is a Seifert fibered space.

<u>Proof of Theorem 7.2</u> We start with the result of 7.1, which we summarize as follows. Let $\tilde{M}$ be the universal covering of M, let G be the group of covering translations. Let $f: T \to M$ be a least area map with $f_*(\pi_1(T)) = \Gamma$. The image of a lift of f in $\tilde{M}$ will be called a <u>plane</u> in $\tilde{M}$.
Each plane $U \subset \tilde{M}$ is invariant under a subgroup $\Gamma_U \subset G$ isomorphic to $\mathbb{Z} \oplus \mathbb{Z}$, namely the translation group of the covering $U \to T$.
If U_1, U_2 are planes, either $U_1 = U_2$, $U_1 \cap U_2 = \phi$ or $U_1 \cap U_2 \cong \mathbb{R}$. In the third case, $U_1 \cap U_2$ is called a <u>line</u> in $\tilde{M}$, and is invariant under an infinite cyclic subgroup of $\Gamma_{U_1} \cap \Gamma_{U_2}$. Two lines l_1, l_2 in a plane U are <u>parallel</u> if $\exists\, \gamma \in \Gamma_U \setminus \{1\}$ such that $\gamma l_i = l_i$ $(i = 1, 2)$.

<u>Step 1</u> If U_1, U_2 meet U in non-parallel lines, then $\pi_1(M)$ contains $\mathbb{Z}^3$ as a subgroup of finite index.

<u>Proof</u> If $l_1 = U \cap U_1$ and $l_2 = U \cap U_2$ are not parallel, they meet in a finite non-zero number of points. So $U_1 \cap U_2$

is a line ℓ, and $\ell \cap U = \ell_1 \cap \ell_2$.
Let $\gamma \in \Gamma_{U_1} \cap \Gamma_{U_2}$ leave ℓ invariant, and let $\gamma_i \in \Gamma_U \cap \Gamma_{U_i}$ leave ℓ_i invariant.
Then $\gamma, \gamma_1, \gamma_2$ commute, since Γ_U, Γ_{U_i} are Abelian. No power of γ is in the subgroup generated by γ_1, γ_2, otherwise ℓ would meet U in infinitely many points. So $\gamma, \gamma_1, \gamma_2$ generate a subgroup H of $G \cong \pi_1(M)$ that is isomorphic to $\mathbb{Z}^3$.

Let $\hat{M}$ be the covering of M with $\pi_1(\hat{M}) = H$.
$\exists$ map $f: \hat{M} \to T^3$ inducing an isomorphism $f_*: \pi_1(\hat{M}) \to \pi_1(T^3)$, where $T^3 = S^1 \times S^1 \times S^1$.
Since $\pi_i(\hat{M}) \cong \pi_i(T^3) = 0$ for $i \geqslant 2$, f is a homotopy equivalence by the Whitehead theorem. Therefore $H_3(\hat{M}) \cong \mathbb{Z}$, so $\hat{M}$ is a closed manifold, a finite cover of M, so H has finite index in $\pi_1(M)$.

<u>Proposition</u> A torsion-free group G with a subgroup H of finite index isomorphic to $\mathbb{Z}^3$ contains a normal subgroup isomorphic to $\mathbb{Z}$.

See P. Scott "The Geometries of 3-manifolds" Bulletin of London Mathematical Society 15 (1983) 401-487, in particular p 444-5.

In view of this proposition, we may assume that whenever planes U_1, U_2 meet another plane U in lines ℓ_1, ℓ_2, then ℓ_1 and ℓ_2 are parallel.

An element $\gamma \in \Gamma \cong \mathbb{Z} \oplus \mathbb{Z}$ is primitive if $\gamma = \beta^m \Rightarrow m = \pm 1$.

Step 2 If planes U_1, U meet in a line ℓ, then $\exists$ primitive element $\gamma \in \Gamma_U$ such that $\gamma \ell = \ell$.

Proof Suppose $\gamma \in \Gamma_U$ satisfies $\gamma^m(\ell) = \ell$ with $m \geq 1$ and minimal. We must show that $m = 1$.

U separates $\tilde{M}$ into 2 components: let X be the closure of one of them. Since M, T are orientable, γ preserves orientations of $\tilde{M}$, U, so $\gamma X = X$. Let $Y = U_1 \cap X$, a closed half-plane. Operating on U_1 by powers of γ gives distinct planes $U_1, \gamma U_1, \ldots \gamma^{m-1} U_1$. Any two of these intersect in (at most) a line invariant under γ^m. It follows that there is a compact set $C < X$ such that $C^* = \bigcup_{n \in \mathbb{Z}} \gamma^n C$ contains ~~all~~

$\gamma^i Y \cap \gamma^j Y$ and $\gamma^i Y \cap \partial X$ for $0 \leq i \neq j < m$.

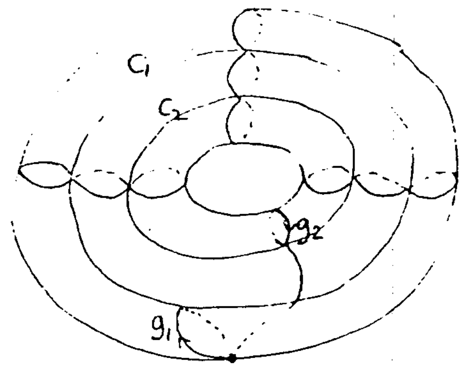
Let $a \in \partial X$ be far from C^*, so $\forall n \in \mathbb{Z}$ $\gamma^n a$ is in the same component of $\partial X \setminus C^*$ as a. Let A be the closure of the component of $X \setminus (C^* \cup \bigcup_{n\in\mathbb{Z}} \gamma^n Y)$ that contains a. Since each $\gamma^n Y$ separates X, there is just one $i \in \{0, \dots m-1\}$ such that $A \cap \gamma^i U_1 / \langle \gamma^m \rangle$ is non-compact. This is impossible if $m > 1$, since $\gamma A = A$ and $\langle \gamma \rangle$ permutes $U_1, \gamma U_1, \dots \gamma^{m-1} U_1$ transitively. Therefore $m = 1$.

<u>Corollary</u> $\Sigma = \{x \in T : \exists y \neq x \text{ such that } f(x) = f(y)\}$ consists of immersed curves on T which are all in the same (primitive) homotopy class.

<u>Step 3</u> $\pi_1(f(T))$ has a normal subgroup isomorphic to $\mathbb{Z}$.

<u>Proof</u> Let $K = f(T)$. We can write $K = \bar{f}(F)$, where F is a surface covered by T (a torus or Klein bottle) and $\bar{f} : F \to M$ is self-transverse. Let $C_1, \dots C_k$ be the double curves of $\bar{f}(F)$. Let $g_1, \dots g_R \in \pi_1(K$

be as shown.



Then $\pi_1(K)$ is generated by $\pi_1(F), g_1, \dots g_k$. [Otherwise, there is a covering $\hat{K}$ of K such that $\pi_1(\hat{K})$ contains $\pi_1(F), g_1, \dots g_k$. It follows that $\bar{f}: F \to K$ lifts to $\hat{K}$, and that the covering map $\hat{K} \to K$ is 1-1.]

The subgroup of $\pi_1(K)$ generated by $C_1, \dots C_k$ is infinite cyclic and normal.

Completion of proof of torus theorem

Let N be obtained from a regular neighborhood of $f(T)$ by attaching 3-balls to any 2-sphere boundary components. Since M is irreducible, $N \subset M$, and $\pi_1(N) \cong \pi_1(f(T))$ contains a normal $\mathbb{Z}$.

If ∂N were compressible in N, with compressing disk D, either

$N \setminus D$ has 2 components N_1, N_2,
or $N \setminus D = N_1$ is connected.

By Van Kampen's theorem, in the first case $\pi_1(N) \cong \pi_1(N_1) * \pi_1(N_2)$: in the second case, $\pi_1(N) \cong \pi_1(N_1) * \mathbb{Z}$. This contradicts the existence of a normal $\mathbb{Z}$ subgroup of $\pi_1(N)$ (except in the case $\pi_1(N_1) \cong \pi_1(N_2) \cong \mathbb{Z}_2$ but in this case N would be closed).

So ∂N is incompressible in N.

If ∂N had a component S of genus > 1, $\pi_1(S) \cap (\text{normal } \mathbb{Z}) = \{1\}$, implying that $\pi_1(S) \cdot (\text{normal } \mathbb{Z})$ is of finite index in $\pi_1(N)$, and N is closed.

So all components of ∂N are tori.

Either some component of ∂N is incompressible in M, or all components of ∂N bound solid tori in M. In the second case, $\pi_1(M)$ is a quotient of $\pi_1(N)$, so contains a normal $\mathbb{Z}$ subgroup.