TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 310, Number 1, November 1988

THE EXISTENCE OF LEAST AREA SURFACES IN 3-MANIFOLDS

JOEL HASS AND PETER SCOTT

ABSTRACT. This paper presents a new and unified approach to the existence theorems for least area surfaces in 3-manifolds.

Introduction. A surface F smoothly embedded or immersed in a Riemannian manifold M is *minimal* if it has mean curvature zero at all points. It is a *least area* surface in a class of surfaces if it has finite area which realizes the infimum of all possible areas for surfaces in this class. The connection between these two ideas is that a surface which is of least area in a reasonable class of surfaces must be minimal. The converse is false; minimal surfaces are in general only critical points for the area function. There are close analogies between these two concepts and the theory of geodesics in a Riemannian manifold. Minimal surfaces correspond to geodesics, and least area surfaces correspond to geodesic arcs or closed geodesics which have shortest length in some class of paths. Any geodesic λ in a Riemannian manifold M has the property that it is locally shortest, i.e., if P and Q are nearby points on λ, then the subarc of λ which joins P and Q is the shortest path in M from P to Q. It can also be proved that minimal surfaces are locally of least area, but the proof is difficult and involves substantial knowledge of the theory of partial differential equations.

There are now a large number of theorems asserting the existence of surfaces of least area in various classes. Surfaces of this type have become an important tool in 3-dimensional topology. In this paper we present a new approach to the proofs of these existence theorems. This yields a simplified and unified method for the proof of the existence of minimal surfaces in 3-dimensional Riemannian manifolds.

In 1930 Douglas [**Do**] and Rado [**Ra**] independently showed that a simple closed curve in R^n which bounds a disk of finite area bounds a disk of least possible area. This result was extended by Morrey [**MoI**] in 1948 to a general class of Riemannian manifolds, the homogeneously regular manifolds, a class which includes all closed manifolds. Work of Osserman [**O**] and Gulliver [**G**] later showed that least area disks in 3-manifolds were immersed in their interiors. More recently there has been a series of new existence results for closed surfaces of least area in manifolds of any dimension. It follows from work of Sacks and Uhlenbeck [**S-UI**] that if M is closed and $\pi_2(M)$ is nonzero then there is an essential map of the 2-sphere into M which has least area among all essential maps. Sacks and Uhlenbeck [**S-UII**] and Schoen and Yau [**S-Y**] independently showed that if $f\colon F \to M$ is a map of a closed orientable surface F, not the 2-sphere, into a closed Riemannian manifold M, such

Received by the editors June 23, 1986.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 53A10; Secondary 57M35.
The first author was supported by NSF Postdoctoral Fellowship DMS84-14097.

that $f_*: \pi_1(F) \to \pi_1(M)$ is injective, then there is a map g of F into M such that g_* equals f_* up to conjugacy and g has least possible area among all such maps.

Each of the Douglas-Rado, Schoen-Yau and Sacks-Uhlenbeck theorems are about the minimization of area in homotopy classes of maps. There is an alternative approach to finding minimal surfaces known as geometric measure theory. The techniques of this theory were used by Meeks, Simon and Yau [**M-S-Y**] to prove the existence of surfaces of least area in certain isotopy classes in a 3-manifold. In particular, if F is an incompressible embedded surface in a closed 3-manifold and I is the infimum of the areas of all surfaces isotopic to F then either there is a surface L isotopic to F with area I, or there is a one-sided embedded surface L' with area $I/2$ such that the boundary of a regular neighborhood of L' is isotopic to F.

In this paper we start with Morrey's result on the existence of least area disks and the results of Meeks-Yau and Gulliver and Osserman on the regularity properties of these disks. Meeks and Yau established that if the boundary of a 3-manifold has positive mean curvature and if a least area disk has boundary properly embedded in the boundary of the manifold, then the least area disk is embedded. This key result is special to dimension three, as are most of the techniques used in this paper. We use these results to establish the existence of surfaces which are of least area in their isotopy classes, thus obtaining results similar to those of Meeks-Simon-Yau, but without the use of geometric measure theory.

We begin by considering a minimizing sequence of embedded surfaces $\{S_i\}$, i.e. a sequence whose area approaches the infimum of all possible areas among surfaces isotopic to a given one. Instead of directly examining the convergence of the sequence, we construct a new minimizing sequence by intersecting the surfaces S_i with small balls and replacing the intersections with least area disks. The problem of convergence is reduced to that of convergence of a sequence of least area disks in a 3-ball, and this is handled in Lemma 3.3, our key convergence result. We observe in passing that this lemma implies the existence of curvature estimates for least area disks embedded in 3-manifolds. These are special cases of curvature bounds found by Schoen [**S**] by less geometric techniques. By the above methods we establish existence results in the embedded setting for both spheres and surfaces of higher genus.

Having established existence results for least area embedded surfaces in 3-manifolds, we turn to the existence theorems for least area maps which are not assumed to be embeddings. Via topological techniques involving covering spaces and cut and paste arguments, we show that when the ambient manifold is 3-dimensional, the results of Sacks-Uhlenbeck and Schoen-Yau follow from the existence of least area embedded surfaces. Thus one conclusion of our methods is that these results can be deduced from geometric measure theory and topological techniques, without resort to the analysis of Morrey, Sacks-Uhlenbeck and Schoen-Yau. However, since we can establish the existence results for least area embedded surfaces without geometric measure theory, another conclusion is that the proofs of all the existence theorems follow from Morrey's theorem, the Meeks-Yau embeddedness result for disks, and arguments involving 3-dimensional topology. Since the Meeks-Yau techniques are basically topological, the net result is to concentrate the analysis involved into Morrey's theorem. Note that the results of Sacks-Uhlenbeck and Schoen-Yau also relied essentially on Morrey's theorem, but contain much additional analysis.

In a sequel to this paper [**H**] additional existence results, many of them new, will be established using the same basic technique. These include the existence of surfaces minimizing area in a homology class, cases of surfaces with boundary, and the existence of certain noncompact minimal surfaces.

This paper is organized as follows. In §1 we summarize the definitions which we will use later. In §2 we survey the "well-known" material from minimal surface theory which we will need. We include this because many of the proofs of these results are either scattered through the literature or not written down, and because one aim of this paper is to make this material accessible to those not familiar with the techniques of minimal surface theory. In §3 we prove our key results on the convergence of least area 2-disks in 3-balls. In §4 we apply these results to demonstrate the existence of least area 2-spheres in 3-manifolds in various situations. This is the main section of the paper. In §5 we show how similar methods to those of §4 can be used to prove that certain 3-manifolds contain closed surfaces of positive genus which are of least area. Up to this point, we have considered closed 3-manifolds and surfaces only. In §6 we discuss the modifications needed to our arguments to handle the cases where the surfaces and 3-manifolds have boundary, and the case of noncompact 3-manifolds.

1. Preliminaries. We will work in the category of smooth manifolds and maps unless otherwise specified. We will denote the boundary of a manifold M by ∂M and its interior by $\operatorname{int}(M)$. A map $f\colon (F, \partial F) \to (M, \partial M)$ of manifolds with boundary is said to be *proper* if $f^{-1}(\partial M) = \partial F$ and the preimage of any compact subset of M is compact. A 3-manifold M is *irreducible* if every embedded 2-sphere in M bounds a 3-ball embedded in M. M is P^2-*irreducible* if it is irreducible and contains no embedded 2-sided projective planes. A 2-sphere embedded in a 3-manifold is *incompressible* if it does not bound a 3-ball. An embedded surface F in M other than a 2-sphere is *incompressible* if whenever D is a 2-disk embedded in M such that $D \cap F = \partial D$, then ∂D bounds a 2-disk in F. Two-sided embedded surfaces other than S^2 are incompressible if and only if their inclusion is injective on fundamental groups. One-sided embedded incompressible surfaces do not necessarily inject on fundamental groups.

In a Riemannian manifold M, we will use the notation $d_M(x, y)$ for the distance from x to y, and $B_M(x, \varepsilon)$ for the ball of radius ε centered at x. We drop the subscript M if there is no ambiguity. A Riemannian manifold M is said to have *convex* boundary if, given x in ∂M and $\varepsilon > 0$, any two points of $B(x, \varepsilon) \cap \partial M$ can be joined by a geodesic λ of M which lies in $B(x, \varepsilon)$. ∂M is *strictly convex* if $\lambda \cap \partial M = \partial \lambda$.

If F is a surface, then an immersion f of F in M induces a Riemannian metric on F, by pulling back the quadratic form on each tangent space. The *area* of f is defined to be the area of F in this induced metric. Clearly one can extend this definition to maps which are smooth immersions almost everywhere. If one alters f to f' by composing with a diffeomorphism of F, the induced metric on F will be altered, but the area of f' will be the same as that of f. In order to pick out a particularly nice map which can be obtained from f in this way, one considers the energy of the maps involved. To define energy, one needs a Riemannian metric on

F and on M. Then the *energy* of a map $f\colon F \to M$ is defined by

$$E(f) = \frac{1}{2}\int_F |df|^2$$

whenever this integral exists. For maps of surfaces to 3-manifolds, the energy integral can be locally expressed in a coordinate neighborhood U of F by

$$\frac{1}{2}\int_F \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial x} \right\rangle + \left\langle \frac{\partial f}{\partial y}, \frac{\partial f}{\partial y} \right\rangle \, dx\, dy,$$

where $\partial/\partial x$, $\partial/\partial y$ form an orthonormal basis for $TF|_U$ with dual basis dx, dy, and where $\langle\ ,\ \rangle$ denotes inner product given by the Riemannian metric on M. Note that energy is defined for a much larger class of maps than smooth immersions, namely maps in the function space L^2_1, essentially the set of functions on which one can integrate the squares of the functions and the squares of their first derivatives. Note also that the energy depends on the choice of metric for F. When f is a smooth immersion, the area of f is always less than or equal to its energy, whatever metric is chosen on F. For surfaces in R^3, this follows as the area of a map is locally expressed by the formula

$$\mathrm{Area}(f|_D) = \int_D \left\| \frac{\partial f}{\partial x} \times \frac{\partial f}{\partial y} \right\| \, dx\, dy,$$

where D is the standard unit disk and $f|_D$ is a diffeomorphism that gives a coordinate neighborhood of F. The area integrand satisfies

$$\begin{aligned} \|\partial f/\partial x \times \partial f/\partial y\| &= [\|\partial f/\partial x\|^2 \|\partial f/\partial y\|^2 - \langle \partial f/\partial x, \partial f/\partial y\rangle^2]^{1/2} \\ &\leq \|\partial f/\partial x\|\, \|\partial f/\partial y\| \leq \tfrac{1}{2}(\|\partial f/\partial x\|^2 + \|\partial f/\partial y\|^2), \end{aligned}$$

where the final expression is the energy integrand. Equality holds if and only if $\langle \partial f/\partial x, \partial f/\partial y\rangle = 0$ and $\|\partial f/\partial x\| = \|\partial f/\partial y\|$. If these conditions hold then f is said to be *almost conformal.* f is *conformal* if the additional condition $|\partial f/\partial x| \neq 0$ also holds. By selecting appropriate coordinates, one can always find a metric on F such that f is almost conformal and the energy and area of f are indeed equal. This metric is unique up to a conformal diffeomorphism of F, and such coordinates are called *isothermal coordinates.* For the existence of such coordinates see [**Ch**].

Finally, we give the definition of convergence which we will use in this paper. This appears to be a somewhat nonstandard concept. Let $\{F_i\}$ be a sequence of embedded surfaces in the Riemannian 3-manifold M. We will say that the sequence $\{F_i\}$ *converges* to the surface F if the following conditions hold:

1. F consists of all the limit points of the sequence, i.e. given a sequence of points x_i in F_i with limit x, then x lies in F, and each point of F is the limit of such a sequence.

2. Given x in F and x_i in F_i as above, there is a disk neighborhood N of x in F, disk neighborhoods N_i of x_i in F_i, and diffeomorphisms $f\colon D \to N$ and $f_i\colon D \to N_i$, where D is the unit disk in R^2, such that f_i converges to f in the C^∞-topology on maps of D to M.

If the sequence $\{F_i\}$ has no limit points, it will be convenient to say that $\{F_i\}$ converges to the empty surface.

2. Results in minimal surface theory. We present here some results which are used in this paper which are more or less well known in minimal surface theory but may not be familiar to those outside the field. We start with a result which states that the length of a curve gives an upper bound to the area of a least area disk spanning the curve. This is a simple example of a set of inequalities called isoperimetric inequalities.

LEMMA 2.1. *Let M be a closed Riemannian manifold. There exist constants $\alpha > 0$ and $\delta > 0$ such that for any closed curve Γ contained in a ball of radius α, there is a disk D spanning Γ with* $\operatorname{Area}(D) \leq \delta \operatorname{Length}(\Gamma)^2$.

PROOF. As M is closed, there is an $\alpha > 0$ such that balls of radius α in M are taken to the unit ball in R^n by a map φ with uniformly bounded dilation. In R^n the above inequality holds for the curve $\varphi(\Gamma)$, with $\delta = 1$, by a computation of the area of the disk D' formed by taking the cone over a point on $\varphi(\Gamma)$. The ratio of the areas of D' and $\varphi^{-1}(D') = D$ is bounded by a constant given by the dilation bound on $d\varphi^{-1}$, which is bounded since M is compact. The result follows.

We next state a result relating, for sufficiently small r, the growth of area of a surface in a ball of radius r to the length of the curve of intersection of the surface and the sphere of radius r. This is known as the co-area formula. It does not depend on the surface being minimal. Let M be a Riemannian manifold, let x be a point in M and let $B(r)$ be the ball of radius r about x in M with boundary $S(r)$. Let F be a surface meeting $B(r)$, let $A(r)$ be the area of $F \cap B(r)$ and when F meets $S(r)$ transversely let $L(r)$ be the length of $F \cap S(r)$.

LEMMA 2.2. *If r is less than the injectivity radius of M at x and less than the distance of x from ∂M, then $\partial(A(r))/\partial r \geq L(r)$.*

PROOF. For brevity we give the proof only for the case when M is 3-dimensional. The general case is similar. For almost all values of r we can assume by Sard's theorem that $F \cap S(r)$ consists of a finite number of smooth curves on $S(r)$. Let e_1, e_2, e_3, be an orthonormal basis for TM in some neighborhood U of $F \cap S(r)$, such that e_1 is tangent to the curve $F \cap S(r)$, e_2 is tangent to F and normal to e_1, and e_3 is normal to F. Let ω_1, ω_2, ω_3 be the dual basis of one-forms. Then the formulas for the area $A_U(r)$ of $F \cap U \cap B(r)$ and length $L_U(r)$ of $F \cap U \cap S(r)$ can be expressed by

$$A_U(r) = \int_{F \cap U \cap B(r)} \omega_1 \wedge \omega_2 \quad \text{and} \quad L_U(r) = \int_{F \cap S(r) \cap U} \omega_1.$$

Integrating $L_U(r)$ over r in U gives

$$\int_{F \cap U} L_U(r)\, dr = \int_U \left(\int_{F \cap S(r)} \omega_1 \right) dr = \int_{F \cap U} \omega_1 \wedge dr.$$

Since ω_1 and dr both have norm one, and since $\omega_1 \wedge \omega_2$ restricted to F has norm one, we have that $\omega_1 \wedge \omega_2 \geq \omega_1 \wedge dr$ and so $A_U(r) \geq \int_{F \cap U} L_U(r)\, dr$. The co-area formula follows.

The next result, called the monotonicity formula, deals with the growth of the area of a minimal surface. In applications it is often used to show that the area of a minimal surface cannot be too small.

LEMMA 2.3. *Let M be a closed Riemannian manifold. There exists $\alpha > 0$ and $\beta > 0$ such that if D is a least area disk in M, x a point on D and if $B(r)$ is a ball of radius $r < \alpha$ in M centered at x with $\partial D \cap B(r) = \emptyset$, then the area of $D \cap B(r)$ exceeds βr^2.*

PROOF. Choose α as in Lemma 2.1. Let D_r denote the intersection of D with the ball of radius r. D_r in general consists of a collection of planar domains which may not be disks. Each component of D_r however is least area rel boundary. The intersection of D_r with $\partial B(r)$ consists of a collection of curves $\{\gamma_i\}$. By Lemma 2.1 we can find a disk C_i in $B(r)$ for each i which satisfies $(A(C_i)) < \delta \operatorname{length}(\gamma_i)^2$. Since D was a least area disk we have that $A(D_r) \leq \sum_i A(C_i)$ so that

$$\frac{A(D_r)}{\delta} \leq \sum_i \operatorname{length}(\gamma_i)^2 \leq \left(\sum_i \operatorname{length}(\gamma_i)\right)^2 = \operatorname{length}(\partial D_r)^2.$$

Let $A(r)$ be the area of D_r and let $L(r)$ be the length of $\partial D_r = \sum_i \operatorname{length}(\gamma_i)$. Then $\sqrt{A(r)/\delta} \leq L(r)$ and by the co-area formula, $\partial/\partial r(A(r)) \geq L(r)$. Combining these two inequalities gives that $\partial A(r)/\partial r \geq \sqrt{A(r)/\delta}$. Since $A(r) > 0$ it follows that $A(r) > y(r)$, where $y(r) = r^2/4\delta$ is the solution to the O.D.E. $\partial y/\partial r = \sqrt{y/\delta}$, $y(0) = 0$. Thus we have $A(r) > r^2/4\delta$ and the result follows by taking $\beta = 1/4\delta$.

The next lemma relates the lengths of curves in a surface to the energy of the surface. One would like to be able to say that if the energy of a map of a surface is small then it sends short curves in the domain to short curves in the image. This need not be true in general, as indicated by Example 2.5, but the following important lemma says that at least some curves have short image. It is often referred to as the Courant-Lebesgue lemma. Note that the function $\varepsilon(\delta)$ in the lemma is independent of x and f. Thus the lemma says that given a bound K on the energy of f there is always some curve C_r, with r uniformly bounded from below, whose image has short length.

LEMMA 2.4. *Let K be a constant and let f be a smooth map of the 2-disk D^2 to a Riemannian manifold M with the energy of f satisfying $E(f) < K$. Let x be a point in D^2 and let C_r be the set of points in D^2 at Euclidean distance r from x. Thus C_r is a circle or part of a circle of radius r about x. Let δ be a constant, $0 < \delta < 1$. Then there is an r_0 with $\delta < r_0 < \sqrt{\delta}$ such that* $\operatorname{length}[f(C_{r_0})] < \varepsilon(\delta) = 2K/\log(1/\sqrt{\delta})$.

REMARK. Note that $2K/\log(1/\sqrt{\delta}) \to 0$ as $\delta \to 0$. An important special case is when x lies on ∂D. Note that it need not be true that the length of $f(C_r)$ tends to zero as $r \to 0$. Example 2.5, which follows the proof of the lemma, demonstrates this.

PROOF. The energy $E(f)$ is given in the region $A = \{\delta \leq r \leq \sqrt{\delta}\}$ by

$$\frac{1}{2}\iint_A \left(\left|\frac{\partial f}{\partial x}\right|^2 + \left|\frac{\partial f}{\partial y}\right|^2\right) dx\, dy \leq K$$

or, in polar coordinates, by

$$\frac{1}{2}\iint_A \left(\left|\frac{\partial f}{\partial r}\right|^2 + \frac{1}{r^2}\left|\frac{\partial f}{\partial \theta}\right|^2\right) r\, dr\, d\theta \leq K.$$

Discarding the first term gives

$$\frac{1}{2}\iint_A \frac{1}{r}\left|\frac{\partial f}{\partial\theta}\right|^2 dr\,d\theta \leq K.$$

Consider r_0 with $\delta \leq r_0 \leq \sqrt{\delta}$ such that $\int |\partial f(r_0,\theta)/\partial\theta|^2\,d\theta$ is minimized. Then

$$\frac{1}{2}\iint_A \frac{1}{r}\left|\frac{\partial f(r_0,\theta)}{\partial\theta}\right|^2 dr\,d\theta \leq K.$$

Integrating gives

$$\frac{1}{2}(\log\sqrt{\delta} - \log\delta)\int\left|\frac{\partial f(r_0\theta)}{\partial\theta}\right|^2 d\theta \leq K,$$

so that

$$\int\left|\frac{\partial f(r_0,\theta)}{\partial\theta}\right|^2 d\theta \leq \frac{2K}{\log(1/\sqrt{\delta})}.$$

The result follows.

EXAMPLE 2.5. We construct in this example a mapping f of the unit disk into R^3 of bounded energy with the property that there are circles about the origin C_r of arbitrarily small radius r whose image has length greater than one. This mapping is depicted in Figure 1. To construct f we consider the function g_k on $D(1/k)$, the disk of radius $1/k$ in R^2, defined by $g_k(r,\theta) = 1-(kr)^{1/k}$. Calculation shows that $\int_{D(1/k)}|dg_k|^2 = \pi/k$. g_k is smooth except at the origin, and is zero on the circle of radius $1/k$ about the origin. If we let h_k be a smooth function on the plane with support on the disk of radius $1/k$, defined by smoothing g_k near the origin and near the circle of radius $1/k$ in such a way that $h_k(0)$ is less than 1 but still greater than $1/2$, we can arrange that $\|dh_k\| < \|dg_k\|$ at each point. Letting k go to infinity we have that $E(h_k)\to 0$ and $\text{support}(h_k)\to 0$. Thus we can construct smooth "spike" functions f_n with support on the disk of radius $1/2^n$ whose value at the origin is between $1/2$ and 1 and such that

$$\int_{D(1/2^n)}\|df_n\|^2 < 1/2^n.$$

Spacing these spikes along the x-axis so that the function f_n is centered at the point $(1/n,0)$, $n=2,3,\dots$, we get a function $h\colon D^2\to R$ with support on the interior of the unit disk, with $E(h) < 1$ and with $h(1/n,0) > 1/2$ for $n = 2,3,4,\dots$. We now form the function $f\colon D^2\to R^3$ by letting $f(x,y) = (x,y,h(x,y))$. Then $E(f) < 2\pi+1$ but the image of the circle of radius $1/n$ under f has length greater than 1 for $n=2,3,\dots$.

The following lemma, sometimes known as the maximal principle for minimal surfaces, describes the intersection of two minimal surfaces near a point of tangency.

LEMMA 2.6. *Let M be a Riemannian 3-manifold and let F_1, F_2, be closed minimal surfaces immersed in M. Suppose that F_1 and F_2 are tangent at a point P. Then either F_1 and F_2 coincide or there is a C^1 coordinate chart (x^2,x^2,x^3) about P in which F_1 is given by $x^3=0$ and F_2 is given by $x^3 = \mathrm{Real}(x^1+ix^2)^n$ for some $n\geq 2$.*

PROOF. In local coordinates the minimality of F_1 and F_2 yields a partial differential equation for their difference, expressed as a graph over their common tangent

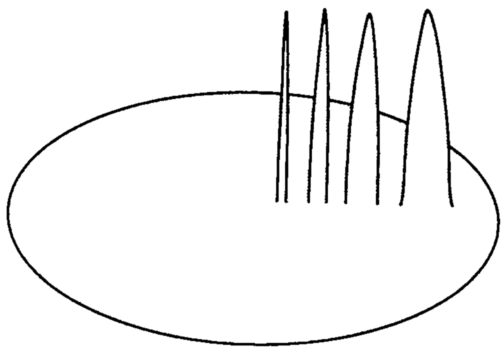

FIGURE 1

plane. This equation is of the quasi-linear type. The proof proceeds by applying a variation of the Hopf maximal principle of partial differential equations to the equation for the difference of two minimal surfaces near a point of tangency [**M-YI**, **F-H-S**].

We next summarize some of the basic existence and regularity theorems for minimal surfaces and give sketches of their proof. Let M be a closed Riemannian n-manifold. Let γ be a simple closed curve in M. Let F be the class of maps f of the two-disk D into M with $E(f)$ bounded and such that f takes the boundary of the two-disk homeomorphically onto γ.

THEOREM 2.7 (MORREY). *There exists a map f in F such that $E(f) = \inf\{E(g) | g \in F\}$ and any such map is smooth.*

REMARK. Morrey's result applies to a larger class of manifolds, namely the homogeneously regular manifolds. These are discussed in §6. Morrey's result also applies to the case where γ is only topologically embedded, so long as γ bounds a disk of finite area.

PROOF. We give a very short sketch of the proof of this theorem. The proof for Euclidean space, due to Douglas [**Do**] and Rado [**Ra**], can be found in [**L**].

One works in the Hilbert space of L^2-functions from F to M which have bounded energy, known as the L_1^2-functions. Take a sequence of functions in **F** whose energy converges to the infimum. Arguments similar to those of §4 establish that, with appropriate reparametrization, this sequence is equicontinuous and has a subsequence converging in L_1^2 to a continuous limit f in **F** whose energy realizes the infimum. One then proves the smoothness of f by estimates involving the homogeneous regularity and a bootstrapping process common in P.D.E. theory.

The least area disks whose existence is established in Theorem 2.7 have additional regularity properties.

THEOREM 2.8 (GULLIVER, OSSERMAN). *A least area map of a surface to a 3-manifold, when restricted to the interior of the surface, is the composite of a branched cover of the surface to another surface and an immersion of the second surface into the 3-manifold.*

REMARK. A least area torus in $S^2 \times S^1$ can be constructed by composing the double branched cover of a torus over the 2-sphere with the inclusion of the 2-sphere into $S^2 \times S^1$ as one factor. This gives an example of a torus which is least area in its homotopy class and factors through a branched cover.

PROOF. We give an outline of the proof of Osserman and Gulliver. Since being immersed is a local property it suffices to prove this theorem for least area disks, which is the case treated by the work of Osserman and Gulliver. Least area disks can be parametrized to be least energy. They then give conformal harmonic maps. Such maps have first derivatives which vanish at isolated points, and these are called branch points. The map is an immersion away from these points. If the least area surface has nonimmersed image, then there is a canonical representation for the local picture around a branch point. Topologically the picture is the cone over an immersed curve in the 2-sphere surrounding the branch point. There is a process, well known to topologists and dating back to Dehn, of pushing the branch point along a double curve, which reparametrizes the map in a neighborhood of the branch point, without changing its image or its area. The resulting map has the same area as the original map but is not immersed along an arc of what was previously a double curve. Since the nonimmersed points of least area disks are isolated, a small disk about this folding curve can be replaced by one of less area, contradicting the assumption of the original surface being least area.

The next result states that with some assumptions on the boundary curve γ, a least area disk is actually embedded.

THEOREM 2.9 (MEEKS-YAU). *Let M be a compact Riemannian 3-manifold whose boundary is strictly convex, and let γ be a simple closed curve in M which is null-homotopic in M. Then γ bounds a least area disk D in M. If γ is embedded in ∂M then any such least area disk D is properly embedded in M. Moreover if D' is another such disk with boundary $\gamma \subset \partial M$, then* $\mathrm{int}(D) \cap D'$ *is empty or $D = D'$. If $\overline{\gamma}$ is a simple closed curve disjoint from γ bounding a least area disk $\overline{D}$, then $D \cap \overline{D} = \varnothing$.*

REMARK. There is a similar result when the boundary of M satisfies a weaker condition than strict convexity. This is stated in Theorem 6.3. If one does not work in the smooth category, then one must make an additional assumption that γ bounds a disk of finite area.

PROOF. The disk D can be shown to exist in the case where M has nonempty boundary by extending the techniques of Morrey discussed in Theorem 2.7. Assume that the disk D is in general position and its boundary is an embedded curve γ on ∂M. One can then apply covering space techniques originally due to Papakyriakopoulos [**P**], but used here in a Riemannian setting, to reduce the problem to the case where there are no triple points of self-intersection in the least area disk. This is done by passing to a sequence of double covers of regular neighborhoods of the disk, called a tower, until one reaches a cover where the boundary of the regular neighborhood consists entirely of 2-spheres. γ bounds two embedded disks on the sphere component of the boundary on which it lies in this cover, and it is not hard to see that one of them is smaller than the lift of D to this cover unless this lift is embedded. If the lift is embedded, project down one stage to a space two-fold covered by the previous one. In this cover the self-intersections of the lift of D consist of pairs of identified simple closed curves on D. These can be exchanged to give a new map of D^2 into M with the same boundary as the old map and having the same area, but with the new map nonimmersed along the former double curve of self-intersection. This contradicts Theorem 2.8 if this is a least area disk. By a perturbation argument; cf. [**F-H-S**, §1] one can justify the assumption that the disk D is in general position, proving the first embeddedness result claimed in the theorem. The other statements follow similarly.

3. Convergence of least area disks. This section contains results on the properties of least area disks and the convergence of sequences of such disks. Most of these results do not seem to have been stated in this form before, and so we give more detailed proofs than in the previous section. The first lemma states that the least area disk spanning a simple closed curve lying on a small sphere lies in the ball bounded by the sphere.

LEMMA 3.1. *Let M be a closed Riemannian 3-manifold. Then there is an $\varepsilon > 0$ such that for any point x in M, the ball of radius ε about x, $B(x, \varepsilon)$, has the property that if $\Gamma \subset \partial B(x, \varepsilon)$ is a simple closed curve and if D is a least area disk in M which spans Γ, then D is properly embedded in $B(x, \varepsilon)$.*

PROOF. From Theorem 2.7 we know that given a curve Γ as above there exists a least area disk D in M with boundary Γ. We will show that if ε is chosen sufficiently small, D lies inside $B(x, \varepsilon)$.

We first pick an $R > 0$ such that the boundary of $B(x, r)$ is strictly convex for any x in M and any r with $0 < r < R$. The existence of such an r follows from the compactness of M. Consider now a curve Γ on $\partial B(x, R/K)$ for K some large integer. There is a constant C_K such that $\partial B(x, R/K)$ has area less than C_K for all $x \in M$. Moreover $C_K \to 0$ as $K \to \infty$. There is another constant $\delta_0 > 0$, given by Lemma 2.3, such that for any $x \in M$, a least area disk F_0 meeting x and having no boundary in $B(x, R/2)$ has area larger than δ_0. Let $\varepsilon = R/K_0$ where K_0 is sufficiently large so that $K_0 > 2$ and $C_{K_0} < \delta_0$. Now consider a curve Γ on $\partial B(x, \varepsilon)$ bounding a least area disk D.

If $D \subset B(x, R)$ and D does not lie in $B(x, \varepsilon)$ then there is an R_0 with $\varepsilon < R_0 < R$ such that $R_0 = \sup\{r | D \subset B(x, r)\}$. But $\partial B(x, R_0)$ is convex and D meets $\partial B(x, R_0)$ and lies on its convex side. This contradicts the maximal principle, Lemma 2.6.

Thus either $D \subset B(x, \varepsilon)$ or D intersects $M - B(x, R)$. In the latter case we can find a point $x_0 \in B(x, R)$ such that $B(x_0, R/2) \subset B(x, R) - B(x, R/2)$ and $x_0 \in D$. As ∂D cannot meet $B(x_0, R/2)$, Lemma 2.3 shows that $D \cap B(x_0, R/2)$ has area $\geq \delta_0$. But $\delta_0 > C_{K_0}$ for K_0 large and Γ bounds a disk in $B(x, \varepsilon)$ of area less than C_{K_0}, so that D cannot be a least area disk. Thus the least area disk lies in $B(x, \varepsilon)$. Now Theorem 2.9 shows that D is properly embedded in $B(x, \varepsilon)$, concluding the proof of the lemma.

We will need to know that least area disks do not wander too far even when their boundaries are not simple closed curves lying on spheres. The next lemma states that least area disks spanning short curves do not wander too far.

LEMMA 3.2. *Let M be a closed Riemannian 3-manifold. Then there exists $r > 0$ such that if $\varepsilon < r$ and Γ is a closed curve of length less then ε contained in $B(x, \varepsilon)$, then any least area disk D spanning Γ lies in the ball $B(x, \varepsilon)$.*

PROOF. The proof is similar to Lemma 3.1, but uses Lemma 2.1, the isoperimetric inequality, rather then the area of a 2-sphere on which the curve lies, to find a disk with small area bounded by Γ.

In §4, we will consider the convergence of a sequence $\{S_i\}$ of 2-spheres in a closed Riemannian 3-manifold M. We will consider a 3-ball X in M with strictly convex boundary and will arrange that $S_i \cap X$ consists of a collection of disjoint least area disks. We will then apply the following lemma, which contains the key convergence technique of the paper. Part of the proof is based on ideas in [**M-YII**].

LEMMA 3.3. *Let X be a compact Riemannian 3-manifold with strictly convex boundary. Let $\{D_i\}$ be a sequence of properly embedded least area disks in X which have uniformly bounded area. Then there is a subsequence $\{D_i\}$ which converges smoothly in* $\mathrm{int}(X)$ *to a properly embedded minimal surface T, which may be empty.*

Let X_ε denote X with an open ε-neighborhood of ∂X removed. Then there is a $\beta > 0$ such that the intersection of T with X_ε consists of a collection of disjoint embedded least area disks for each ε with $0 < \varepsilon < \beta$.

PROOF. Let $\varepsilon_0 > 0$ be any constant. We show first that $\{D_i\}$ has a subsequence converging on $\mathrm{int}(X_{\varepsilon_0})$. As ε_0 is arbitrary this implies convergence on $\mathrm{int}(X)$.

If the sequence $\{D_j\}$ has no limit point in $\mathrm{int}(X_{\varepsilon_0})$, then the result is trivial. So we suppose that P is a limit point which lies in $\mathrm{int}(X_{\varepsilon_0})$.

Recall that a least area map has least energy if it is parametrized conformally, and that such a parametrization can always be found. Pick such isothermal parametrizations of $\{D_i\}$ to get a sequence of maps $f_i\colon D \to X$ where f_i is energy minimizing for its boundary values. By composing with a conformal map of the disk and passing to a subsequence, we can assume that $f_i(0) \to P$. In isothermal coordinates the energy and area are equal, so that there is a constant K such that $E(f_i) < K$ for all i.

Let C be a compact subset of $\mathrm{int}(D)$. We next establish the equicontinuity of f_i on C. Let x, y be two points in C with distance $d(x, y) < \delta_1$, where $\sqrt{\delta_1} < d(C, \partial X)$. Then by Lemma 2.4 for any given i, there is a circle C_r of radius r about x with $\delta_1 < r < \sqrt{\delta_1}$ such that $f_i(C_r)$ has length less than $\varepsilon(\delta_1) = 2K/\log(1/\sqrt{\delta_1})$, and $\varepsilon(\delta_1) \to 0$ as $\delta_1 \to 0$. Now for $\varepsilon(\delta_1)$ sufficiently small, $f_i(C_r)$ lies in a small ball B in X of radius $\varepsilon(\delta_1)$. From Lemma 3.2 it follows that the entire least area subdisk bounded by C_r also lies in this ball. It follows that the distance in X between $f_i(x)$ and $f_i(y)$ is less than $2\varepsilon(\delta_1)$. Thus $\{f_i\}$ is equicontinuous on compact subsets of $\mathrm{int}(D)$ and the Arzela-Ascoli theorem implies that there is a continuous function $f\colon \mathrm{int}(D) \to X$ with a subsequence of $\{f_i\}$ converging pointwise to f on $\mathrm{int}(D)$.

We show next that $\{f_i\}$ converges smoothly to f on compact subsets of $\mathrm{int}(D)$. As above we can assume that a small disk of radius δ_0 about a point x, $D(x, \delta_0)$, has image lying in a coordinate ball B. We can identify this ball with the unit ball in 3-space by a diffeomorphism, and the minimal surface equation in the corresponding coordinates becomes:

$$\Delta f_i^j = \frac{\partial^2 f_i^j}{\partial x^2} + \frac{\partial^2 f_i^j}{\partial y^2} = -\sum_{k,l} \Gamma_{kl}^j \left\{ \frac{\partial f_i^k}{\partial x}\frac{\partial f_i^l}{\partial x} + \frac{\partial f_i^k}{\partial y}\frac{\partial f_i^l}{\partial y} \right\},$$

where f_i^j denotes the jth component of f_i and Γ_{kl}^j are the Christoffel symbols of the metric on B. We can use this equation together with Sobolov and L^p estimates to deduce that f_i is uniformly bounded in the $C^{k,\alpha}$ norm [**M-YII**]. Standard P.D.E. arguments then give that f is smooth and that the sequence $\{f_i\}$ has a subsequence converging smoothly in $\mathrm{int}(X)$ to f.

Next we show that given $\alpha > 0$ there is an r with $0 < r < 1$ such that the annulus A_r of points in $\mathrm{int}(D)$ lying within distance r of ∂D contains an essential curve σ_i whose image under f_i lies within $X - X_\alpha$. This implies, in particular, that f is not a constant map. For each point y in ∂D and each small δ we can find, using Lemma 2.4, a curve C_r with $\delta < r < \sqrt{\delta}$ whose image under f_i lies within $\varepsilon(\delta)$ of ∂X. By taking many such points y_k on ∂D and stringing the corresponding curves together, we can find annuli $A(r_l, r_m)$ with $0 < r_l < r_m < 1$ and $r_l, r_m \to 1$, such that each f_i maps some curve σ_i essential in $A(r_l, r_m)$ within distance $\varepsilon(\delta)$ of ∂X.

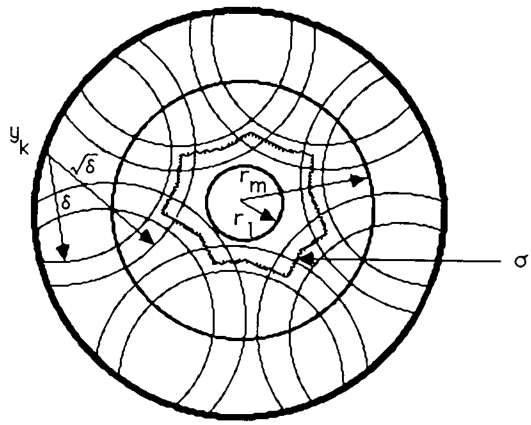

FIGURE 2

We next show that f is a least area map. If not there is a compact subdisk D' of D such that $f|D'$ is not least area. Thus there exists a disk D'' with smaller area and with the same boundary as D', and there exists a $\beta > 0$ with $\mathrm{Area}(f(D')) - \mathrm{Area}(f(D'')) > \beta$. Now since f is the smooth limit of f_i on $\mathrm{int}(D)$ there are disks D_i' such that $f_i(D_i')$ converges to $f(D_i')$ as $i \to \infty$. In particular, for large i,

$$|\,\mathrm{Area}(f_i(D_i')) - \mathrm{Area}(f(D_i'))| < \beta/2,$$

and $f_i(\partial D_i')$ is sufficiently close to $f(\partial D_i')$ so that $f_i(\partial D_i') \cup f(\partial D_i')$ cobound an annulus A in X with $\mathrm{Area}(A) < \beta/2$. Since $\mathrm{Area}(A \cup f(D_i'')) < \mathrm{Area}(f_i(D_i')$ this gives a contradiction to the assumption that f_i is least area, and we conclude that D'' cannot exist and that f is a least area map. Now f is minimal and a smooth limit of embedded disks and thus has embedded image; if not then it would have a point of transverse self-intersection which would also appear in a nearby approximating surface. Conceivably f could factor through a branched cover, as in the remark following Theorem 2.8. However if so then f could not be approximated by embeddings in a neighborhood of the branch point. So f has no branch points and is an embedding itself.

Since ∂X is strictly convex so is $\partial X_{\varepsilon_0}$ for sufficiently small ε_0. Thus the intersection of $f(\mathrm{int}(D))$ with X_{ε_0} contains only disks, and no planar domains. For a planar domain would imply the existence of a least area planar domain in $X - X_{\varepsilon_0}$ with boundary on $\partial X_{\varepsilon_0}$, which contradicts the maximal principle, as such a surface would meet a convex surface, namely ∂X_α for some small α, on its convex side without crossing it. Thus the intersection of $f(\mathrm{int}(D))$ with X_{ε_0} consists of a collection of embedded disjoint least area disks.

It may happen that $f(\mathrm{int}(D))$ is not the full limit set of $\{D_i\}$ in $\mathrm{int}(X_{\varepsilon_0})$ (see Example 3.4 below). If so there is a point P_1 in $\mathrm{int}(X_{\varepsilon_0})$ which is not in $f(\mathrm{int}(D))$ but is a limit point of the sequence $\{D_i\}$. We can then repeat the above construction using P_1 instead of P, obtaining a new limiting map $f^1\colon \mathrm{int}(D) \to X$ whose image contains P_1. The union of the images of f and f^1 gives a collection of embedded disjoint least area disks in X_{ε_0}, as above. We then repeat this process for a point P_2 if $f \cup f^1$ does not give the full limit set of $\{D_i\}$ in X_{ε_0}. Since the addition of a component containing P_1 added a finite amount of area estimable from below in terms of ε, via Lemma 2.3, and since the total area of f is finite and the area of the maps in the sequence $\{f_i\}$ is uniformly bounded, it follows that after a finite number of repetitions we recover the full limit set in X_{ε_0}. Letting $\varepsilon_0 \to 0$ and letting $T = f(D) \cup f^1(D) \cup f^2(D) \cup f^3(D) \cdots$ we obtain the desired surface.

REMARK. It follows from the above proof that the curvature of any least area disk in X is bounded above in X_{ε_0}. For if not, take a sequence of least area disks whose curvature blows up in X_{ε_0}. Such a sequence cannot converge to a nonbranched embedding, contradicting the above argument. This result is one of a collection of such curvature estimates for minimal surfaces which are due to Schoen, Simon, Yau and others (see [S] for the 3-manifold case).

EXAMPLE 3.4. The sequence of curves in the boundary of the unit ball in R^3, depicted in Figure 3, bounds a sequence of least area disks whose limit in the interior of the ball is two disks.

In this example the convergence is complicated by the convergence of two arcs to a single arc. In the following lemma we see that if this phenomenon is avoided, we get convergence up to the boundary. Let H^+, H^-, S^+ and S^- refer to the upper and lower hemispheres of the 2-sphere and 1-sphere respectively.

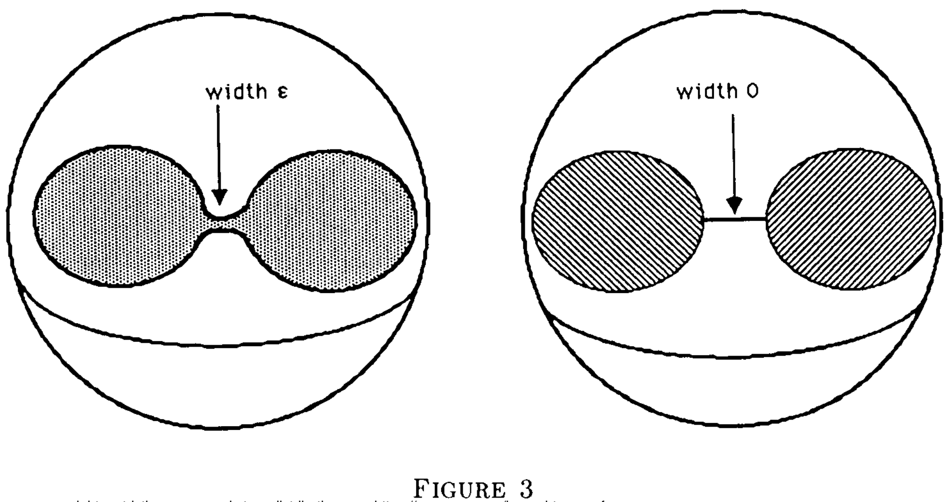

FIGURE 3

LEMMA 3.5. *Let X be a 3-ball with a Riemannian metric such that X has strictly convex boundary. Let $\{D_i\}$ be a sequence of properly embedded least area disks in X which converge to T as in Lemma* 3.3. *Suppose there is a smooth simple arc Γ_0 properly embedded in H^+ and suppose there is a sequence of simple arcs $\Gamma_i \subset \partial D_i$ such that $D_i \cap H^+ = \Gamma_i$ and $\{\Gamma_i\}$ converges to Γ_0. Then a component of T extends smoothly to Γ_0. That is, there is a smooth map $f_0\colon \mathrm{int}(D) \cup S^+ \to \mathrm{int}(M) \cup H^+$ and $\{f_i\}$ converges on $\mathrm{int}(D) \cup S^+$ to f_0.*

PROOF. Parametrize the sequence $\{D_i\}$ to obtain energy minimizing maps $f_i\colon D^2 \to X$. These energy minimizing maps are not unique, as conformal maps of the disk leave the energy unchanged when composed with f. We now use a standard technique to normalize $\{f_i\}$ so that the images of three points are fixed. Let a, b, c be points in Γ_0 such that $\partial\Gamma_0 = a \cup b$ and $c \in \mathrm{int}(\Gamma_0)$, and reparametrize f_i conformally so that as $i \to \infty$, $f_i(-1) \to a$, $f_i(+1) \to b$ and $f_i(\sqrt{-1}) \to c$. As in the proof of Lemma 3.3, the sequence $\{f_i\}$ is equicontinuous on compact subsets of $\mathrm{int}(D)$. We will show that $\{f_i\}$ is equicontinuous on compact subsets of $\mathrm{int}(D)\cup S^+$. Let x be any point on S^+, and fix a constant $\varepsilon > 0$. For δ sufficiently small Lemma 2.4 implies that the curve C_r defined in Lemma 2.4, for some r, $\delta < r < \sqrt{\delta}$, has image $f_i(C_r)$ of length less than $\varepsilon/2$. Thus the endpoints e_1, e_2, of the arc on ∂D containing x and running between the endpoints of C_r have images lying within $\varepsilon/2$ of one another. Since the maps $\{f_i|S^+\}$ have images converging to Γ_0, it follows that if i is chosen sufficiently large, then the length of the image of the arc on ∂D between $f_i(e_1)$ and $f_i(e_2)$ and containing x is no more than ε. Thus for each f_i there is a value r_i with $\delta < r_i < \sqrt{\delta}$ such that the boundary curve of the disk $D(x, r_i)$ about x bounded by C_r and ∂D has image under f_i of length less than 2ε. By Lemma 3.2 we can conclude that the entire disk $D(x, \delta)$ has image $f_i(D(x, \delta))$ that lies within 2ε of $f_i(x)$. Thus by starting with i large, we can assume that the entire sequence $\{f_i\}$ maps $D(x, \delta)$ to within 2ε of $f(x)$, and we conclude that $\{f_i\}$ is equicontinuous on compact subsets of $\mathrm{int}(D) \cup S^+$. The limit map f maps S^+ to Γ_0. As in Lemma 3.3 the image of f may not consist of the entire set of points to which $\{D_i\}$ accumulate, but f is nonconstant and its image intersects X_ε in a disjoint union of embedded disks for ε sufficiently small. Finally, f is smooth on $D \cup S^+$ by results of Hildebrandt [**Hi**].

LEMMA 3.6. *Let X be a compact Riemannian* 3*-manifold with strictly convex boundary. Let T_i be a countable collection of properly embedded least area disks in X with the area of T_i uniformly bounded. Then there is a subsequence $\{T_j\}$ of $\{T_i\}$ which converges in* $\mathrm{int}(X)$ *to a countable collection T* (*possibly empty*) *of least area open disks properly embedded in* $\mathrm{int}(X)$. *If $\{T_i\}$ has a limit point in* $\mathrm{int}(X)$ *then T is not empty.*

Note. This convergence may be with multiplicity. A large number of disks in $\{T_j\}$ may converge to a single disk in T. Note also that we do not assume that the disks in T_i are disjoint from one another.

PROOF OF LEMMA 3.6. First note that since ∂X is strictly convex, so is ∂X_ε for ε sufficiently small. Thus we can pick ε small enough so that the intersection of a least area disk with X_ε consists of least area disks, as in Lemma 3.3. By the monotonicity lemma, Lemma 2.3, only a finite number of components of $\{T_i\}$ can intersect X_ε. Picking one such component, if it exists, from each T_i, we obtain a sequence of least area disks $\{D_i^1\}$. A subsequence $\{D_j^1\}$ converges in X_ε by Lemma 3.3. In the corresponding subsequence $\{T_j\}$ of $\{T_i\}$ we pick a second component intersecting X_ε, if it exists, and obtain a new sequence of least area disks $\{D_i^1\}$. Again Lemma 3.3 implies convergence of a subsequence on X_ε, possibly to the same limiting disk as before. Since there is a uniform bound on the number of components of $\{T_i\}$ which meet X_ε, we repeat this process a finite number of times to obtain a subsequence of $\{T_i\}$ which converges in X_ε. We call this sequence $\{T_{i,\varepsilon}\}$.

We next consider the intersection of $\{T_{i,\varepsilon}\}$ with $X_{\varepsilon/2}$. As above we get a convergent subsequence in $X_{\varepsilon/2}$, which we call $\{T_{i,2}\}$. Proceeding in this way we get sequences $\{T_{i,n}\}$ converging in $X_{\varepsilon/n}$, with $\{T_{i,n}\}$ a subsequence of $\{T_{i,n-1}\}$. Taking the diagonal sequence $\{T_{i,i}\}$ we get a sequence converging in $\operatorname{int}(X)$. The limit is a countable, possibly empty, collection of open disks, intersecting $X_{\varepsilon/n}$ in a finite collection of least area disks for any $n \geq 1$. It is nonempty by the construction if there is a limit point of $\{T_i\}$ in $\operatorname{int}(X)$.

4. Least area spheres. In this section we present some existence results for 2-spheres of least area in a closed 3-manifold. In Theorem 4.1, we use the results of §3 to establish the existence of a least area 2-sphere among the family of all embedded 2-spheres which do not bound a 3-ball. This argument is the heart of our paper, and similar arguments will be applied in §5 to show that least area embedded surfaces of higher genus also exist. In Theorem 4.2, we use Theorem 4.1 and some standard topological arguments to prove the existence of a least area map among all homotopically nontrivial maps of the 2-sphere into a 3-manifold. Theorem 4.3 extends this to find a finite collection of disjoint minimal 2-spheres which generate $\pi_2(M)$ as a $\pi_1(M)$-module.

THEOREM 4.1. *Let M be a closed Riemannian 3-manifold which contains an embedded 2-sphere that does not bound a 3-ball. Let F denote the set of all piecewise smooth embedded 2-spheres in M which do not bound a 3-ball, and let $I = \inf\{\text{Area}(S)\colon S \text{ in } F\}$. Then either there is an embedded sphere Σ in F with area equal to I or there is an embedded one-sided projective plane P in M with area equal to $I/2$ and with the boundary of a regular neighborhood of P an element of F.*

PROOF. Let $\{S_i\}$ be a sequence of embedded spheres in M, each in F, whose area tends to I as $i \to \infty$. Cover M with the interiors of balls $B_1, B_2, \dots, B_n$, whose boundary in M is strictly convex, each having radius sufficiently small so that the condition in Lemma 3.1 holds. We will form a series of new minimizing sequences out of $\{S_i\}$, so that the final one will converge. We first consider the ball B_1. We can assume that S_i is transverse to ∂B_j for each i, j by decreasing the radius of B_j slightly if necessary. Let Γ_i be a simple closed curve in $S_i \cap \partial B_1$. Suppose first that the intersection of S_i with B_1 consists of a single disk D_i whose boundary is Γ_i. Let D_i' be a disk of least area in M which spans Γ_i. By our choice of the B_i's, D_i' lies in B_1. We can form a new embedded 2-sphere S_i' by setting $S_i' = S_i - D_i + D_i'$. Clearly S_i' has area no larger than that of S_i and does not bound a 3-ball, so that S_i' is an element of F.

In general, the intersection of S_i with B_1 consists of a collection of planar surfaces. In this case, we alter S_i by replacing $S_i \cap B_1$ with a collection of least area disks spanning the curves of $S_i \cap \partial B_1$. These disks lie inside B_1 and are embedded, by Lemma 3.1 and Theorem 2.9, and they are disjoint as their boundaries are disjoint. Thus we obtain a collection of embedded 2-spheres $\Sigma_1, \Sigma_2, \ldots, \Sigma_k$. Clearly each Σ_j satisfies $\text{Area}(\Sigma_j) \leq \text{Area}(S_i)$, and at least one Σ_j does not bound a 3-ball. We denote this Σ_j by $\overline{S_i}$. The new sequence $(\overline{S_i})$ thus obtained still has the property that $\lim_{i\to\infty}(\text{Area}(\overline{S_i})) = I$. Each member of this new sequence intersects B_1 in a collection of least area disks. Lemma 3.6 now gives a subsequence $\{S_{i,1}\}$ which converges on $\text{int}(B_1)$ to a union of least area open disks. Note that this limit surface may be empty.

We next consider $S_{i,1} \cap B_2$. As before, form $\{\overline{S}_{i,1}\}$ by replacing $S_{i,1} \cap B_2$ by least area disks and picking a component $\overline{S}_{i,1}$ that does not bound a 3-ball. Lemma 3.6 gives a subsequence $\{S_{i,2}\}$ which converges in $\text{int}(B_2)$ to a surface T_2 which is a union of open least area disks. Now $S_{i,2} \cap (B_1 - B_2)$ consists of some subcollection of the components of $S_{i,1} \cap (B_1 - B_2)$, as $S_{i,2}$ was formed from $S_{i,1}$ by discarding some components. Thus the convergence of $\{S_{i,1}\}$ on $\text{int}(B_1)$ implies the convergence of $\{S_{i,2}\}$ on $\text{int}(B_1) - \text{int}(B_2)$ to a surface which we denote by T_1. We will show that $T_1 \cup T_2$ is a smooth subsurface of $\text{int}(B_1) \cup \text{int}(B_2)$. This is clear if $B_1 \cap B_2$ is empty, so we assume that they intersect. As the convergence of $\{S_{i,1}\}$ was smooth on $\text{int}(B_1)$, T_1 meets $\partial B_2 \cap \text{int}(B_1)$ in smooth open arcs. If this collection of arcs is empty, it is again clear that $T_1 \cup T_2$ is smooth, so we consider a component E of T_1 such that E intersects $\partial B_2 \cap \text{int}(B_1)$ and we consider an arc Γ of this intersection. Then Γ is the limit of a sequence of open arcs $\{\Gamma_i\}$ in $\partial B_2 \cap \text{int}(B_1)$, where Γ_i is a component of $S_{i,1} \cap \partial B_2 \cap \text{int}(B_1)$. By construction, Γ_i forms part of the boundary of a least area disk F_i in B_2 which is a component of $S_{i,2} \cap B_2$. We let A_i denote the closed arc $\partial F_i - \Gamma_i$. Let x be any point of Γ. We will show that $T_1 \cup T_2$ is a smooth surface in a neighborhood of x. The first step is to show that, near x, T_2 is a smooth surface with boundary on Γ. Thus, in a neighborhood of x, $T_1 \cup T_2$ is a continuous surface which is smooth except possibly for a bend along Γ. Then we will show that no bending can occur.

Case 1. There is a constant $\varepsilon > 0$ such that $A_i \cap B(x, \varepsilon)$ is empty for all i.

In this case, Lemma 3.5 implies that $\{F_i\}$ converges in B_2 to a surface which extends continuously up to the boundary in a neighborhood of x in B_2.

Case 2. The sequence A_i intersects each neighborhood of x.

Let C_i be a component of $A_i \cap \text{int}(B_1)$ which contains a point y_i such that $y_i \to x$ as $i \to \infty$. Since C_i and Γ_i are both components of $S_{i,2} \cap \partial B_2 \cap \text{int}(B_1)$, and $S_{i,2}$ is embedded, we must have that both arcs converge to Γ. Thus $S_{i,2}$ is converging with multiplicity at least two to E in $\text{int}(B_1) - \text{int}(B_2)$. In this case, we have to consider how the disk F_i converges in $\text{int}(B_2)$. Note that, for small values of ε, the intersection of F_i with $B(x, \varepsilon) \cap B_2$ consists of disjoint 2-disks. This is because F_i is a least area disk in B_2.

Subcase 2(i). There is an $\varepsilon > 0$ such that the disk of $F_i \cap B(x, \varepsilon) \cap B_2$ which contains Γ_i in its boundary intersects $\partial B_2 \cap \text{int}(B_1)$ only in Γ, for all i.

In this case, the argument of Lemma 3.5 again implies that F_i converges to a smooth surface in B_2 which extends continuously up to the boundary in a neighborhood of x in B_2.

Subcase 2(ii). There is a sequence of points x_i on Γ_i which converges to x, and a sequence of points y_i on an arc C_i of A_i also converging to x and a sequence of paths d_i in F_i joining x_i to y_i such that the length of $d_i \to 0$ as $i \to \infty$.

We will show that this case is impossible, by using a cut and paste argument. Choose a small half ball B in $B_1 - \text{int}(B_2)$ centered on x such that $E \cap B$ is a disk E_1. Then, for large values of i, $S_{i,2} \cap B$ has components $D_{i,1}$ and $D_{i,2}$ which are 2-disks such that $D_{i,1} \cap \partial B_2$ is contained in Γ_i and $D_{i,2} \cap \partial B_2$ is contained in C_i. We let G_i denote the disk formed from the union of $D_{i,1}$ and $D_{i,2}$ with a thin strip about d_i. We also construct a 2-disk G_i' by starting with the annulus in ∂B between $\partial D_{i,1}$ and $\partial D_{i,2}$, removing a small segment running between the two boundary components of the annulus near x, and adding two small disks D', D'' near x which fill in the circles formed by the boundary of the segment and the boundary of the strip about d_i (see Figure 4). That disks of small area actually do exist can be seen by taking d_i to be an arc in the intersection of the boundary of a small regular neighborhood N of E_1 with the surface $S_{i,2} \cap B_2$. ∂N intersects B_2 in a disk, of small area compared to E_1, which is cut into two disks by d_i. Note that d_i need not be chosen short, as long as the areas of D' and D'' are small.

The area of G_i tends to at least twice the area of E_1 as $i \to \infty$. The area of G_i' becomes smaller than the area of E_1 as $i \to \infty$, because $\partial D_{i,1}$ and $\partial D_{i,2}$ converge

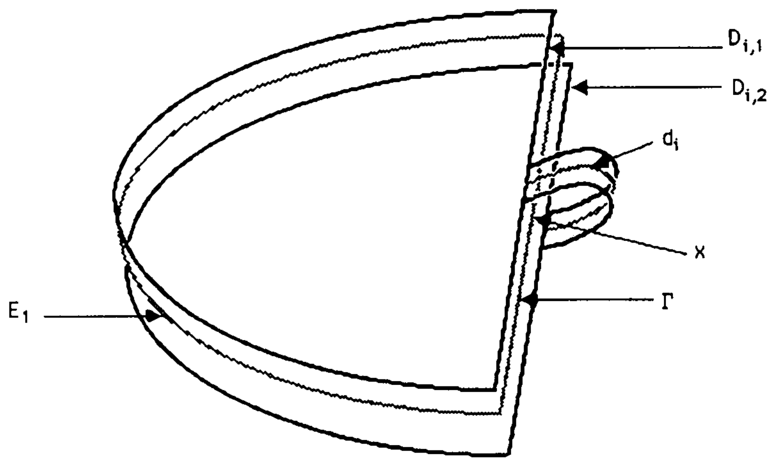

FIGURE 4a

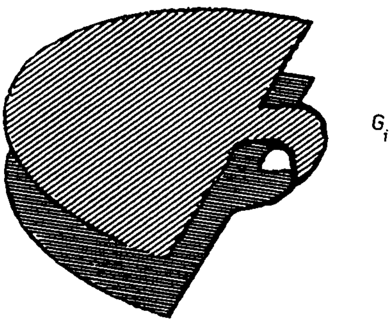

FIGURE 4b

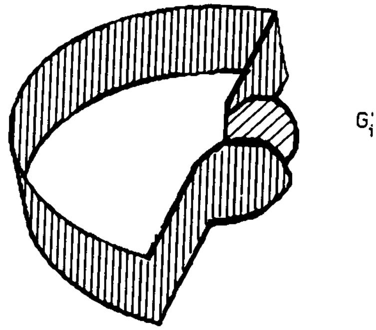

FIGURE 4c

to the same curve ∂E_1 and ∂B, and the area of the disks D', D'' can be taken close to zero. Thus there is a positive constant δ such that $\text{Area}(G_i') < \text{Area}(G_i) - \delta$, for all large enough values of i.

Now consider the sphere $S_{i,2}'$ formed from $S_{i,2}$ by replacing G_i by G_i'. If this new surface is embedded for each i, we obtain a contradiction as follows. Because $\{S_{i,2}\}$ is a minimizing sequence, $\text{Area}(S_{i,2}) < I + \delta$ for large enough values of i. This inequality together with the previous inequality relating the areas of G_i and G_i' shows that $S_{i,2}'$ has area less than I and so must bound a 3-ball. But the union of G_i and G_i' is an embedded 2-sphere which bounds a 3-ball, by construction, so it follows that $S_{i,2}$ itself must bound a 3-ball for large enough values of i, which is the required contradiction.

If $S_{i,2}'$ is singular, we argue in a similar way as follows. It is still true that the area of $S_{i,2}'$ is less than I for large enough values of i. Let X denote the 2-disk obtained from $S_{i,2}$ by removing G_i. Then $S_{i,2}' = G_i' \cup X$ and clearly the double curves of $S_{i,2}'$ are disjoint and simple and are the intersection curves of the interiors of G_i' and of X. Each double curve in $S_{i,2}'$ bounds two 2-disks in $S_{i,2}'$. Out of all these disks, let D be one of least possible area. Clearly D contains no double curves in its interior and D is a subdisk of G_i' or of X. Suppose that D lies in X. Let D' denote the subdisk of G_i' with the same boundary as D. Then $D \cup D'$ is an embedded 2-sphere of area less than that of $S_{i,2}'$, so it must bound a 3-ball. By isotoping D' across this 3-ball, we obtain an isotopy of G_i' which removes at least one double curve from $S_{i,2}'$ and simultaneously reduces its area. By repeating this procedure we find boundary fixing isotopies of G_i' and of X to new disks G_i'' and X'' which have less area and which have disjoint interiors. Thus $G_i'' \cup X''$ is an embedded 2-sphere of area less than I and so bounds a 3-ball. Now these isotopies are all made in the complement of the 2-disk G_i, so it follows that the embedded 2-sphere $G_i \cup X''$, which is isotopic to $G_i'' \cup X''$, bounds a 3-ball. Hence $S_{i,2}$ itself bounds a 3-ball for large enough values of i, because $S_{i,2} = G_i \cup X$ is isotopic to $G_i \cup X''$. This contradiction completes the proof that subcase 2(ii) cannot occur.

As the above discussion applies to any point x of $T_1 \cap \partial B_2$, we conclude that $T_1 \cup T_2$ is a continuous surface which is smooth except possibly for bending along the arcs of $T_1 \cap \partial B_2$. Suppose that $T_1 \cup T_2$ fails to be smooth along an arc Γ of $T_1 \cap \partial B_2$ at a point x of Γ. Choose a small ball B centered at x so that the component of $T_1 \cup T_2$ which contains Γ meets B in a 2-disk H. As H has a bend along G near x, it cannot be a least area disk. We let H' denote a least area disk in B with the same boundary as H, and we choose a number α such that $0 < \alpha < \text{Area}(H) - \text{Area}(H')$. Now for large enough values of i, the intersection of $S_{i,2}$ with B must include a 2-disk H_i such that the H_i's converge to H. Also the curves ∂H_i converge to ∂H. Thus the annulus L_i on ∂B bounded by ∂H and ∂H_i has area which tends to zero as $i \to \infty$. Hence, for large enough values of i, we have $\text{Area}(L_i) < \alpha/4$, and $|\text{Area}(H_i) - \text{Area}(H)| < \alpha/4$. Now consider the surface $S_{i,2}' = S_{i,2} - H_i + L_i + H'$. For large values of i, we have

$$\begin{aligned}\text{Area}(S_{i,2}') &= \text{Area}(S_{i,2}) - \text{Area}(H_i) + \text{Area}(L_i) + \text{Area}(H') \\ &< \text{Area}(S_{i,2}) - (\text{Area}(H) - \alpha/4) + \alpha/4 + (\text{Area}(H) - \alpha) \\ &= \text{Area}(S_{i,2}) - \alpha/2.\end{aligned}$$

Now Area$(S_{i,2}) < I + \alpha/2$ for large values of i, so that $S'_{i,2}$ has area less than I. If $S'_{i,2}$ is embedded, it must bound a 3-ball and hence so does $S_{i,2}$, which is a contradiction. If $S'_{i,2}$ is singular, we argue as before to obtain a contradiction. This completes the proof of our claim that $T_1 \cup T_2$ is a smooth subsurface of $\operatorname{int}(B_1) \cup \operatorname{int}(B_2)$.

A similar argument can now be applied to $B_3, B_4, \ldots, B_n$ to obtain a sequence $(S_{i,n})$, $i \geq 1$, of embedded 2-spheres which converges piecewise smoothly in M to a smooth surface S. Note that S cannot be empty as any sequence of surfaces in a compact manifold must have some limit points. It follows from the piecewise smooth convergence that, for large values of i, $S_{i,n}$ is transverse to the fibers of the normal bundle in a tubular neighborhood of S. Thus $S_{i,n}$ covers S and it follows that S must be a sphere or a one-sided projective plane. The convergence must be with multiplicity one or two and Theorem 4.1 follows.

The previous arguments combined with the sphere theorem [**P**, **Wh**] (see also [**He**]) from 3-dimensional topology now easily imply the existence of least area maps of spheres into a 3-manifold among all those which are homotopically nontrivial.

THEOREM 4.2 (SACKS-UHLENBECK, MEEKS-YAU). *Let M be a closed Riemannian 3-manifold such that $\pi_2(M)$ is nonzero. Let F_0 be the set of all smooth maps of S^2 to M which represent a nonzero element of $\pi_2(M)$, and let $I_0 = \inf\{\text{Area}(f) \colon f \text{ in } F_0\}$. Then there is a map f_0 in F_0 with area I_0. Further f_0 is either a smooth embedding or a double cover of a smoothly embedded projective plane.*

PROOF. First assume that M is orientable. Then the sphere theorem tells us that F_0 contains an embedding. Let F denote the set of embeddings in F_0, and let I denote $\inf\{\text{Area}(f) | f \text{ in } F\}$. As F is nonempty we want to apply the arguments used in Theorem 4.1 to find either an embedded 2-sphere in F of area I or an embedded one-sided projective plane of area $I/2$ and with the boundary of a regular neighborhood a member of F. The only difference in the situations is that now we want to work with 2-spheres which are nonzero in $\pi_2(M)$, which may be a stronger condition than simply insisting that the spheres do not bound balls. These conditions on embedded spheres in a 3-manifold are equivalent if and only if the Poincaré Conjecture is true. However, we simply need to note that, when doing surgery on an embedded 2-sphere which is a member of F, one of the two new spheres must again be a member of F. Now the arguments of Theorem 4.1 apply. It remains to show that $I_0 = I$, so that the least area 2-sphere we have obtained in F is also least area in F_0. This is true because given a singular 2-sphere in F, the proof of the sphere theorem actually supplies one with an embedded sphere in F which has less area. This completes the proof of Theorem 4.2 in the case when M is orientable. When M is nonorientable, one first finds a least area sphere (or projective plane) Σ in the orientable double covering $\tilde{M}$ of M, and then notes that Σ must be equivariant under the covering involution τ of $\tilde{M}$, i.e. either $\tau\Sigma = \Sigma$ or $\tau\Sigma \cap \Sigma$ is empty. For otherwise one could cut and paste Σ and $\tau\Sigma$ to obtain an essential sphere or projective plane in $\tilde{M}$ with area less than that of Σ (note that τ is an isometry of $\tilde{M}$). It follows that Σ covers an embedded surface in M and this is the required sphere or projective plane in M. It is easy to extend this result to the following result first proved by Meeks and Yau.

THEOREM 4.3. *Let M be a closed Riemannian 3-manifold such that $\pi_2(M)$ is nonzero. Then there exist disjoint smooth immersions f_1, $f_2, \ldots, f_k$, from S^2 into M such that*

(1) *f_1 represents a nonzero element of $\pi_2(M)$ and has area I_0,*

(2) *$[f_i] \in \pi_2(M)$ is not in the $\pi_1(M)$-submodule generated by $[f_1], [f_2], \ldots, [f_{i-1}]$ and f_i has least area among all such maps,*

(3) *$[f_1], [f_2], \ldots, [f_k]$ generate $\pi_2(M)$ as a $\pi_1(M)$-module,*

(4) *each f_i is either an embedding or double covers an embedded projective plane.*

PROOF. Theorem 4.2 provides us with f_1. This map generates a $\pi_1(M)$-submodule M_1 of $\pi_2(M)$. Let F_2 denote the set of embeddings of S^2 in M which represent an element of $\pi_2(M) - M_1$, and let $I_2 = \inf\{\text{Area}(f) \colon f \text{ in } F_2\}$. The proof of Theorem 4.2 provides an embedded sphere in F_2 of area I_2 or an embedded projective plane of area $I_2/2$ whose regular neighborhood has boundary in F_2. The key point here is that doing surgery on a sphere in F_2 produces two spheres of which at least one must also lie in F_2. Now, as in Theorem 4.2, this sphere is least area among all maps which represent elements of $\pi_2(M) - M_1$. We repeat this process to obtain a sequence of maps $f_1, f_2, f_3, \ldots$, where f_i is least area among all maps which represent elements of $\pi_2(M) - M_{i-1}$. These spheres must be disjoint, as otherwise one could cut and paste to produce a sphere of still smaller area. Now it follows that this process must eventually stop, i.e. there is k such that $M_k = \pi_2(M)$, as there is a bound on the number of nonparallel disjoint 2-spheres and projective planes which M can contain [**He**]. This completes the proof of Theorem 4.3.

5. Higher genus surfaces. In this section, we consider the problem of finding least area incompressible surfaces of higher genus. First we consider the special case when M is P^2-irreducible and the surface is embedded. Theorem 5.1 states that there is a least area surface in an isotopy class. Theorem 5.2 extends this to reducible 3-manifolds. In Theorem 5.3, we use these results on the existence of embedded least area surfaces to prove a result on the existence of least area maps. If $\pi_2(M)$ is zero, this result yields the existence of a least area map in an incompressible homotopy class.

THEOREM 5.1. *Let M be a P^2-irreducible closed Riemannian 3-manifold and let F be an incompressible closed surface embedded in M. Let $\mathscr{F}$ denote the set of piecewise-smooth surfaces in M which are isotopic to F, and let*

$$I = \inf\{\text{Area}(G) |\ G \text{ in } \mathscr{F}\}.$$

Then either there is a surface F' in $\mathscr{F}$ and with area I or there is a one-sided surface F'' of area $I/2$ and the boundary of a regular neighborhood of F'' is in $\mathscr{F}$.

PROOF. Let $\{F_i\}$, $i \geq 1$, be a sequence of surfaces in F whose areas tend to I as $i \to \infty$, and let $B_1, B_2, \ldots, B_n$ be a finite collection of 3-balls chosen as in the proof of Theorem 4.1 to cover M. As before, we can assume that F_i is transversal to ∂B_j for all i and j. Any curve of $F_i \cap \partial B_1$ is clearly null-homotopic in M and so must bound a 2-disk in F_1, as F is incompressible in M. As in the proof of Theorem 4.1, we replace these disks by disks of least area with the same boundary. We need only carry this out for those curves which are outermost in F_i. We obtain a new surface

$F_{i,1}$ which is isotopic to F_i, as M is irreducible, and has less area than F_i. As before, some subsequence of the $F_{i,1}$'s must converge on $\operatorname{int}(B_1)$, and, by repeating this argument for $B_2, \ldots, B_n$, we will obtain a minimizing sequence of surfaces in $\mathscr{F}$ which converges to a smooth surface F^* in M. Since F_i naturally covers F^* for large enough values of i, it follows that either F^* is in $\mathscr{F}$ or F^* is one-sided and the boundary of a regular neighborhood lies in $\mathscr{F}$ [**J**, **Sc**]. This completes the proof of Theorem 5.1.

If M is a reducible manifold, the above argument works except that the surfaces $F_{i,n}$ need not be isotopic to F. Thus we obtain

THEOREM 5.2. *Let M be a closed Riemannian 3-manifold and let F be an incompressible closed surface embedded in M. Let $\mathscr{F}$ denote the set of surfaces embedded in M which are isotopic to the connected sum of F with some 2-spheres, and let $I = \inf\{\text{Area}(G) | G \text{ in } \mathscr{F}\}$. Then either there is a surface F' in $\mathscr{F}$ with area I or there is a one-sided surface F'' with area $I/2$ such that the boundary of a regular neighborhood of F'' is in $\mathscr{F}$.*

When one considers extending these results to the nonembedded case, one cannot simply argue as in §4. We overcome this by passing to a certain covering of M. However, as this covering is not compact, we need to pass back and forth between M and this covering in order to guarantee the convergence of our surfaces.

THEOREM 5.3. *Let M be a closed Riemannian 3-manifold, let F be a closed surface other than the 2-sphere, let $f: F \to M$ be an injective, two-sided map and let $\mathscr{S}$ denote the set of maps of F into M which induce the same action on $\pi_1(F)$ as f, modulo conjugation in $\pi_1(M)$. If $I = \inf\{\text{Area}(g) | g \text{ in } \mathscr{S}\}$, then there is a smooth immersion f' of area I which is in $\mathscr{S}$.*

PROOF. As in the proof of Theorem 4.1, we cover M by a finite collection of 3-balls $B_1, B_2, \ldots, B_n$. Let M_F be the cover of M with $\pi_1(M_F) = f_*(\pi_1(F))$, so that f lifts to a map $\tilde{f}: F \to M_F$, which induces an isomorphism of fundamental groups. M_F is covered by the lifts of $B_1, \ldots, B_n$ which we label $B_{1,i}, \ldots, B_{n,i}$, $i \geq 1$. In §2 of [**F-H-S**], it is shown that if g is a 2-sided immersion of F into a 3-manifold such that g induces an isomorphism of fundamental groups, then either g is an embedding or there is an embedding g' obtained by cut and paste techniques from g such that g' has less area than g and also induces an isomorphism of fundamental groups. It follows that we can find a sequence of embeddings $\{\tilde{f}_i\}$, $i \geq 1$, of F into M_F, each inducing an isomorphism of fundamental groups, such that $\text{Area}(\tilde{f}_i)$ tends to I as $i \to \infty$.

Let $p: M_F \to M$ denote the covering projection and let $f_i = p \circ \tilde{f}_i$. Thus f_i lies in $\mathscr{S}$. It would be convenient to be able to show that some subsequence of the $\tilde{f}_i$ must converge to a map $\tilde{f}$, for then $f = p\tilde{f}$ would be the required map of F to M of least area. However, the noncompactness of M_F rules out this approach (see Example 6.1), so we consider the convergence of the sequence of singular maps $f_i: F \to M$. As usual, we can suppose that f_i is transverse to ∂B_j for all i and j. Thus $\tilde{f}_i$ is transverse to $\partial B_{j,k}$ for all i, j and k. We replace $\tilde{f}_i(F) \cap B_{1,k}$ with least area disks to obtain $\tilde{f}_{i,1}(F)$. Thus if $f_{i,1} = p\tilde{f}_{i,1}$, then $f_{i,1}(F)$ intersects B_1 in a collection of embedded least area disks which may not be disjoint. Now Lemma 3.6 shows that there is a subsequence $\overline{f}_i$ which converges on $\operatorname{int}(B_1)$ to a collection of embedded least area disks, possibly empty. We can repeat this procedure by lifting $\overline{f}_i$ to a map of F into M_F. Again, we call these lifts $\tilde{f}_i$. Note that each $\tilde{f}_i$ is an embedding. Now we alter these new embedded surfaces by replacing their intersections with the $B_{2,k}$'s by least area disks. As before the projection of this new sequence of surfaces into M has a subsequence which converges on $\operatorname{int}(B_2)$. As in the proof of Theorem 4.1, it follows that it converges on $\operatorname{int}(B_1) \cup \operatorname{int}(B_2)$. Repeating, we eventually obtain a minimizing sequence of embeddings of F in M_F whose projections into M converge to an immersion f' of F into M, whose area is equal to I. This is the required map.

REMARK. This result was first proved by Schoen and Yau [**S-Y**] and by Sacks and Uhlenbeck [**S-UII**]. They assume that F is orientable but do not assume that F is two-sided. A topological argument extends their result to the case when F is nonorientable and two-sided [**F-H-S**]. However, their technique can be extended to cover all cases.

The condition of incompressibility for F in Theorems 4.1, 4.2, 4.3, 5.1, 5.2, and 5.3 can be replaced by a weaker condition, the "condition of cohesion" [**C**]. This condition essentially states that no essential curve on F can have very short image in M. More precisely, if γ is an essential simple closed curve on F and the area of F is less than $I + \varepsilon$, then there is a constant C, depending on ε, such that γ has length greater than C. In the proof of Theorem 5.1, this ensures that when we replace $F_i \cap B_i$ by least area disks to obtain $F_{i,1}$, then $F_{i,1}$ must be isotopic to F_i.

6. The bounded and noncompact cases. In this section, we discuss how the preceding sections need to be modified to handle the existence results when one or both of the 3-manifold and the surface have boundary, and when the 3-manifold may be noncompact.

When one considers surfaces with boundary, there are two natural ideas of a least area map. In the first case, which we call the fixed boundary case, one takes a proper map $g\colon F \to M$ and considers all maps homotopic to g rel ∂F. In the second case, which we call the free boundary case, one considers all maps of F into M which are properly homotopic to g. In this paper, we will restrict our attention to the fixed boundary case, as our methods naturally extend to handle this case.

For noncompact 3-manifolds one can obtain some results similar to those of §§4 and 5 so long as one imposes conditions on the curvature of the 3-manifold and conditions which prevent a minimizing sequence from going to infinity. However, the following example shows that even in a well-behaved noncompact 3-manifold, there are isotopy classes of closed surfaces which contain no least area surface, so that Theorems 4.1 and 5.1 fail to hold when the 3-manifold is noncompact.

EXAMPLE 6.1. Let F be a closed Riemannian surface of area A, and let M be $F \times R$ with the product metric scaled in the directions tangent to F by the function $\varphi(t) = 1 + e^{-t}$, $t \in R$. Then there are no least area surfaces in the isotopy (or homotopy) class of $F \times \{\mathrm{pt}\}$, as pushing any nontrivial surface in M in the direction of increasing t will decrease its area. Note that the curvatures of M are bounded and the injectivity radius is bounded from below. Note also that the infimum of area in the isotopy class of surfaces isotopic to $F \times \{\mathrm{pt}\}$ is equal to A, and this is not achieved by any surface in the isotopy class.

To modify the results in §1, we need to introduce the following condition on an open manifold which was first discussed by Morrey [**MoI**] and was essential for his solution of the Plateau problem in a Riemannian manifold.

A Riemannian 3-manifold M is *homogeneously regular* if there exist positive constants k, K such that every point of the manifold lies in the image of a chart φ with domain the unit ball $B(0,1)$ in R^3 such that

$$k\|v\|^2 \leq g_{ij}(\varphi(x))v_i v_j \leq K\|v\|^2 \quad \text{for all } x \text{ in } B(0,1),$$

where v is any tangent vector to x, g is the metric on M and g_{ij} its components. This can be shown to be equivalent to M having sectional curvature bounded above and injectivity radius bounded away from zero. Note that any closed manifold is automatically homogeneously regular and hence so is any manifold which covers a closed manifold. Note also that the manifold in Example 6.1 is homogeneously regular. This condition was introduced by Morrey [**MoI**] and was essential for his solution of the Plateau problem in a general Riemannian manifold. Morrey showed that if M is a homogeneously regular manifold (of any dimension) and if γ is a simple closed curve in M which bounds a disk of finite area, then there is a disk of least possible area bounded by γ. Further such a disk is smooth. In [**MoI**], Morrey gives an example to explain why the condition of homogeneous regularity is needed. In this example γ bounds a noncompact surface of finite area but bounds no compact surface. We will describe an example where γ bounds a disk of finite area in M, but no disk whose area equals the infimum of all possible areas of disks. The example gives a complete metric on R^3 with this property.

EXAMPLE 6.2. We take M to be R^3 with a metric which we will describe below, and we take γ to be the circle $x^2 + y^2 = 2$, $z = 0$. We will start with the standard metric g on R^3 (which is, of course, homogeneously regular) and will alter it in the region where $x^2 + y^2 \leq 3$, preserving the circular symmetry of the metric about the z-axis. As a first step, we alter g in the region where $1 \leq x^2+y^2 \leq 3$, so that the half cylinder $x^2 + y^2 = 2$, $z \geq 0$ has finite area $1/2$ and the disk $x^2 + y^2 \leq 2$, $z = 0$ still has area 2π. Such a metric can, for example, be constructed by making the above half cylinder isometric to a hyperbolic cusp, and smoothing off using a partition of unity. We can do this in such a way that the lengths of vertical curves remain unchanged, thus making it clear that the resulting metric is complete. We call this new metric g_0. As the metric g_0 is equal to g on the region where $x^2 + y^2 \leq 1$, it is clear that any disk bounded by γ has area greater than π. If γ does not bound a disk of least possible area, this metric will serve as our example. Otherwise, we will describe a sequence of alterations to g_0 which introduce disks of less and less area spanning γ, which will eventually produce the required example. Let D_n denote the 2-disk $x^2 + y^2 \leq 2$, $z = n$. We will inductively construct a metric g_n from g_{n-1} by altering g_{n-1} in a neighborhood of D_n so as to reduce the area of D_n while preserving the metric in a small neighborhood of the half cylinder $x^2 + y^2 \leq 2$, $z \geq 0$. To construct g_1, we choose a 1-parameter family of metrics g_t, t in $[0, \infty)$, in which the area of D_1 decreases as $t \to \infty$. For each such metric, consider the infimum, $I(g_t)$, of the areas of all 2-disks spanning γ. If γ always bounds a disk of area equal to $I(g_t)$, then it is easy to see that $I(g_t)$ is a continuous function of t. By construction, it is decreasing and we can arrange that, for large values of t, it is less than 1, by considering the disk spanning γ formed by D_1 and the annulus $x^2+y^2 = 2$, $0 \leq z \leq 1$. We pick the metric g_t for which $I(g_t)$ equals $1+1/2$ and call this g_1. We choose g_n similarly so that the infimum of areas of all 2-disks spanning γ is $1 + 1/2n$, $n \geq 1$. The limit metric g_∞ is the required metric on R^3. For, by construction, the infimum of the areas of all 2-disks spanning γ is 1 but there is no 2-disk of area equal to 1 which spans γ. Note that there is a punctured disk of area 1/2 spanning γ.

In [**MoI**] Morrey only considered the case of manifolds without boundary. Clearly, some condition analogous to convexity will be needed for ∂M in order to obtain existence theorems for least area surfaces lying properly in M. The appropriate condition was introduced by Meeks and Yau [**M-YIII**]. They called it Condition C, but we prefer the more descriptive phrase sufficiently convex. We use the convention that the unit sphere in R^n has positive mean curvature with respect to the inward pointing normal. A Riemannian manifold M has *sufficiently convex* boundary if the following conditions hold:

1. ∂M is piecewise smooth,
2. Each smooth subsurface of ∂M has nonnegative mean curvature with respect to the inward normal,
3. There exists a Riemannian manifold N such that M is isometric to a submanifold of N and each smooth subsurface S of ∂M extends to a smooth embedded surface S' in N such that $S' \cap M = S$.

We will also need the following definitions. A subsurface Y of the boundary of a Riemannian 3-manifold M is *convex* if given y in the interior of Y, there is β such that whenever $0 < \varepsilon < \beta$, any two points of $B(y, \varepsilon) \cap \partial M$ can be joined by a geodesic λ of M which lies in $B(x, \varepsilon)$. Y is *strictly convex* if $\lambda \cap M = \partial\lambda$.

As we hope the name suggests, being sufficiently convex is a much weaker condition on ∂M than being convex. For example, ∂M is sufficiently convex if it is a minimal surface. Thus if M_1 and M_2 denote the closures of the two complementary components of a catenoid in R^3, then both M_1 and M_2 have sufficiently convex boundary, but clearly neither has convex boundary. In [**M-YIII**], Meeks and Yau generalize Morrey's result to the case of homogeneously regular manifolds with sufficiently convex boundary. A noncompact manifold with boundary will be called *homogeneously regular* if the manifold N described above can be chosen to be homogeneously regular.

THEOREM 6.3. *Let M be a compact Riemannian 3-manifold with sufficiently convex boundary, and let γ be a simple closed curve in M which is null-homotopic in M. Then γ bounds a least area disk in M. If γ is embedded in ∂M, then any such least area disk D is either properly embedded in M or is embedded in ∂M. Moreover, if D' is another such disk with boundary γ in ∂M, then* $\operatorname{int}(D) \cap D'$ *is empty or $D = D'$. If $\overline{\gamma}$ is a simple closed curve disjoint from γ and bounding a least area disk $\overline{D}$, then either $D \cap \overline{D}$ is empty or D and $\overline{D}$ are embedded in ∂M with one contained in the interior of the other.*

REMARK. The methods of Meeks and Yau can be used to establish the same result when M is noncompact, so long as M is homogeneously regular.

The other lemmas in §2 also need to be modified. Lemma 2.1 is replaced by the following.

LEMMA 6.4. *Let M be a homogeneously regular 3-manifold with sufficiently convex boundary. There exist constants $\alpha > 0$ and $\delta > 0$ such that for any closed* *curve* Γ *contained in a ball of radius* α, *there is a disk* D *spanning* Γ *with* $\mathrm{Area}(D) \leq \delta\,\mathrm{Length}(\Gamma)^2$.

PROOF. In the case where M is compact with sufficiently convex boundary, the techniques of [**M-YIII**] show how to embed M isometrically in an open homogeneously regular 3-manifold N so that for any simple closed curve Γ in M, some least area disk in N which spans Γ lies in M. Now the proof of Lemma 2.1, and the observation that the area of a least are disk is smaller than the area of the disk constructed in Lemma 2.1, completes the argument in the case when M is compact.

When M is not compact, picking α sufficiently small gives that a ball of radius α in M has sufficiently convex boundary, as its frontier will be convex. The above argument now applies for this compact ball.

Lemma 2.2 needs no change. Lemma 2.3 is replaced by the following result whose proof is the same except that Lemma 6.4 is used in place of Lemma 2.1.

LEMMA 6.5. *Let* M *be a homogeneously regular Riemannian 3-manifold with sufficiently convex boundary. There exists* $\alpha > 0$ *and a function* $\beta(r)$ *defined on the interval* $(0, \alpha)$ *such that if* D *is a least area disk in* M, x *a point on* D, r *less than the distance of* x *from* ∂M, *and if* B *is a ball of radius* r *in* M *centered at* x *with* $\partial D \cap B = \emptyset$, *then the area of* $D \cap B$ *exceeds* $\beta(r)$.

Lemma 2.4 needs no change. Lemma 2.6 is replaced by the following.

LEMMA 6.6. *Let* M *be a Riemannian 3-manifold with sufficiently convex boundary, and let* F_1 *and* F_2 *be minimal surfaces immersed in* M. *Suppose that* F_1 *and* F_2 *are tangent at a point* P *in the interiors of* F_1 *and* F_2. *Then either* F_1 *and* F_2 *agree on an open neighborhood of* P *or there is a* C^1 *coordinate chart* (x^1, x^2, x^3) *about* P *in which* F_1 *is given by* $x^3 = 0$ *and* F_2 *is given by* $x^3 = \mathrm{Real}(x^1 + ix^2)^n$ *for some* $n \geq 2$. *In particular, if* F_1 *is tangent to* ∂M *at an interior point of* F_1, *then* F_1 *lies in* ∂M.

In §3, we need to restate all the results because we want to consider 3-manifolds with boundary and so will also have to consider 3-balls in these 3-manifolds which meet the boundary. Recall that, in the proof of Lemma 3.1, we used the fact that small balls in a closed manifold have strictly convex boundary. When the ambient manifold has boundary, we will use the fact that small balls have strictly convex frontier. Lemma 3.1 is replaced by the following.

LEMMA 6.7. *Let* M *be a homogeneously regular Riemannian 3-manifold with sufficiently convex boundary. Then there is* $\varepsilon > 0$ *such that for any point* x *in* M, *the ball of radius* ε *about* x, $B(x, \varepsilon)$, *has the property that if* $\gamma \subset \partial B(x, \varepsilon)$ *is a simple closed curve and if* D *is a least area disk spanning* γ, *then either* D *is properly embedded in* $B(x, \varepsilon)$ *or* D *lies in* ∂M.

Lemma 3.2 is replaced by the following. The proof is the same but uses Lemma 6.4 instead of Lemma 2.1.

LEMMA 6.8. *Let* M *be a homogeneously regular Riemannian 3-manifold with sufficiently convex boundary. Then there exists* $r > 0$ *such that if* $\varepsilon < r$ *and* Γ *is a closed curve of length less than* ε *contained in* $B(x, \varepsilon)$, *then any least area disk* D *spanning* Γ *lies in the ball* $B(x, \varepsilon)$.

The replacement for our key convergence result, Lemma 3.3, needs to be more complicated because, when considering the convergence of surfaces in a manifold with boundary, we will need to consider 3-balls which meet the boundary. As we are only going to consider the existence problem in the fixed boundary case, we need the following result, which follows by combining the arguments of Lemmas 3.3 and 3.5.

LEMMA 6.9. *Let X be a compact Riemannian 3-manifold with sufficiently convex boundary, and let Y be a compact subsurface of ∂X which is strictly convex. Let Y' denote the closure of $\partial X - Y$. Let Γ be a compact 1-manifold properly embedded in Y', and let $\{D_i\}$ be a sequence of properly embedded least area disks in X with $D_i \cap Y' = \Gamma$, and with uniformly bounded area. Then there is a subsequence $\{D_j\}$ which converges smoothly in $X - Y$ to an embedded minimal surface T with boundary* $\operatorname{int}(\Gamma)$. *This surface is the union of a minimal surface properly embedded in $X - Y$ and a minimal subsurface of Y', either of which may be empty.*

Let X_ε denote X with an open ε-neighborhood of Y removed. Then there is a $\beta > 0$ such that the intersection of T with X_ε consists of a collection of disjoint embedded least area disks for each ε with $0 < \varepsilon < \beta$.

Lemma 3.5 needs to be replaced by the following.

LEMMA 6.10. *Let X be a 3-ball with a Riemannian metric such that ∂X is sufficiently convex. Let Y be a strictly convex subsurface of ∂X which contains H^+ in its interior. Let $\{D_i\}$ be a sequence of properly embedded least area disks in X which converge to T as in Lemma 6.8. Suppose there is a smooth simple arc Γ_0 properly embedded in H^+ and suppose there is a sequence of simple arcs $\Gamma_i \subset \partial D_i$ such that $D_i \cap H^+ = \Gamma_i$ and $\{\Gamma_i\}$ converges to Γ_0. Then a component D of T extends smoothly to Γ_0. That is, there is a smooth map $f_0\colon \operatorname{int}(D) \cup S^+ \to \operatorname{int}(M) \cup H^+$ and $\{f_i\}$ converges on $\operatorname{int}(D) \cup S^+$ to f_0.*

Lemma 3.6 is replaced by the following.

LEMMA 6.11. *Let X be a compact Riemannian 3-manifold with sufficiently convex boundary, and let Y be a compact subsurface of ∂X which is strictly convex. Let Y' denote the closure of $\partial X - Y$. Let Γ denote a countable collection of 1-manifolds each properly embedded in Y' and let T_i be a countable collection of properly embedded least area disks in X with $T_i \cap Y' = \Gamma$ for all i, and with the area of T_i uniformly bounded. Then there is a subsequence $\{T_j\}$ of $\{T_i\}$ which converges in $X - Y$ to a countable collection T of simply connected minimal surfaces with boundary* $\operatorname{int}(\Gamma)$. *Each surface is either properly embedded in $X - Y$ or is embedded in Y'.*

Armed with the preceding results to replace the results of §§2 and 3, it is easy to extend the results of §§4 and 5 on the existence of least area closed surfaces to the case where M is a compact Riemannian 3-manifold with sufficiently convex boundary. In Lemmas 6.9 and 6.11, the 1-manifold Γ will be empty for these applications. It is also easy to extend these results to the case of surfaces with boundary in 3-manifolds which need not be compact.

THEOREM 6.12. *Let M be a P^2-irreducible homogeneously regular Riemannian 3-manifold with sufficiently convex boundary and let F be an incompressible compact* *surface with nonempty boundary properly embedded in M. Let $\mathcal{F}$ denote the set of piecewise smooth surfaces in M which are isotopic to F rel ∂F, and let $I = \inf(\text{Area}(G) | G$ in $\mathcal{F})$. Then there is a surface F' in $\mathcal{F}$ with area I.*

REMARK. Unlike the situation of Theorem 5.1, one cannot obtain a surface F'' double covered by F, as $\partial F''$ must equal ∂F. We do not need to assume compactness of M as the fixed boundary of F stops the minimizing sequence from converging to the empty surface (see Example 6.1). The same applies to the following result, a version of which was proved by Lemaire [**Le**].

THEOREM 6.13. *Let M be a homogeneously regular Riemannian 3-manifold with sufficiently convex boundary, and let F be a compact surface with nonempty boundary. Let f be a proper two-sided injective map of F into M, such that f embeds ∂F in ∂M, and let S denote the set of all maps of F into M which are homotopic to f rel ∂F. Let I denote $\inf\{\text{Area}(g) | g \in S\}$. Then there is a smooth immersion f' of F into M with $f' \in S$ and such that f' has area equal to I.*

REMARK. One can weaken the assumption that ∂F is embedded to the assumption that F lifts to an embedding in M_F as in §5.

REFERENCES

[**A-S**] F. J. Almgren, Jr. and L. Simon, *Existence of embedded solutions of Plateau's problem*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **6** (1979), 447–495.

[**C**] R. Courant, *Dirichlet's principle, conformal mappings and minimal surfaces*, Interscience, New York, 1950.

[**Ch**] S. S. Chern, *An elementary proof of the existence of isothermal parameters on a surface*, Proc. Amer. Math. Soc. **6** (1955), 771–782.

[**Do**] J. Douglas, *Solution of the problem of Plateau*, Trans. Amer. Math. Soc. **34** (1932), 731–756.

[**F-H-S**] M. Freedman, J. Hass and P. Scott, *Least area incompressible surfaces in 3-manifolds*, Invent. Math. **71** (1983), 609–642.

[**G**] R. D. Gulliver, *Regularity of minimizing surfaces of prescribed mean curvature*, Ann. of Math. (2) **97** (1973), 275–305.

[**H**] J. Hass, *The existence of minimal surfaces in 3-manifolds*. II, in preparation.

[**He**] J. Hempel, *Three-manifolds*, Ann. of Math. Stud., Vol. 86, Princeton Univ. Press, Princeton, N.J., 1976.

[**Hi**] S. Hildebrandt, *Boundary behavior of minimal surfaces*, Arch. Rational Mech. Anal. **35** (1969), 47–82.

[**J**] W. Jaco, *Finitely presented subgroups of 3-manifold groups*, Invent. Math. **13** (1971), 335–346.

[**L**] H. B. Lawson, *Lectures on minimal submanifolds*, Publish or Perish, 1980.

[**Le**] L. Lemaire, *Boundary value problems for harmonic and minimal maps of surfaces into manifolds*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **9** (1982), 91–103.

[**M-YI**] W. H. Meeks III and S. T. Yau, *The classical Plateau problem and the topology of three dimensional manifolds*, Topology **21** (1982), 409–442.

[**M-YII**] ——, *Topology of three dimensional manifolds and the embedding theorems in minimal surface theory*, Ann. of Math. (2) **112** (1980), 441–484.

[**M-YIII**] ——, *The existence of embedded minimal surfaces and the problem of uniqueness*, Math. Z. **179** (1982), 151–168.

[**M-S-Y**] W. H. Meeks III, L. Simon and S. T. Yau, *Embedded minimal surfaces, exotic spheres and manifolds with positive Ricci curvature*, Ann. of Math. (2) **116** (1982), 621–659.

[**MoI**] C. B. Morrey, *The problem of Plateau in a Riemannian manifold*, Ann. of Math. (2) **49** (1948), 807–851.

[**MoII**] ——, *Multiple integrals in the calculus of variations*, Springer-Verlag, New York, 1966.

[**O**] R. Osserman, *A proof of the regularity everywhere of the classical solution to Plateau's problem*, Ann. of Math. (2) **91** (1970), 550–569.

[**P**] C. D. Papakyriakopoulos, *On Dehn's lemma and the asphericity of knots*, Ann. of Math. (2) **66** (1957), 1–26.

[**Ra**] T. Rado, *On Plateau's problem*, Ann. of Math. (2) **31** (1930), 457–469.

[**S-UI**] J. Sacks and K. Uhlenbeck, *The existence of minimal immersions of 2-spheres*, Ann. of Math. (2) **113** (1981), 1–24.

[**S-UII**] ——, *Minimal immersions of closed Riemann surfaces*, preprint.

[**S**] R. Schoen, *Estimates for stable minimal surfaces in three dimensional manifolds*, Seminar on Minimal Submanifolds (E. Bombieri, ed.), Ann. of Math. Stud., Vol. 103, Princeton Univ. Press, Princeton, N.J., pp. 111–126.

[**S-Y**] R. Schoen and S. T. Yau, *Existence of incompressible minimal surfaces and the topology of 3-dimensional manifolds with nonnegative scalar curvature*, Ann. of Math. (2) **110** (1979), 127–142.

[**Sc**] P. Scott, *On sufficiently large 3-manifolds*, Quart. J. Math. Oxford (2) **23** (1972), 159–172.

[**Wh**] J. H. C. Whitehead, *On 2-spheres in 3-manifolds*, Bull. Amer. Math. Soc. **64** (1958), 161–166.

DEPARTMENT OF MATHEMATICS, HEBREW UNIVERSITY OF JERUSALEM, JERUSALEM, ISRAEL

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF LIVERPOOL, LIVERPOOL, ENGLAND

Current address (Joel Hass): Department of Mathematics, University of California, Davis, California 95616

Current address (Peter Scott): Department of Mathematics, University of Michigan, Ann Arbor, Michigan 48109