# A proof of the regularity everywhere of the classical solution to Plateau's problem[1]

By ROBERT OSSERMAN

CONTENTS



The existence of a minimal surface spanning a curve was first proved about 40 years ago by Douglas [6] and Radó [16]. (For a precise statement, see § 1 below.) Since that time there have been many modifications, simplifications, and generalizations of their methods, but all of them have suffered from one major defect. In order to obtain a solution, the surfaces considered were allowed to have certain singularities, called *branch points*, and it has remained an open question whether or not these branch points actually occur. Our purpose here is to show that the solution surfaces do *not* have branch points, so that the classical solution provides an everywhere regular surface spanning the given curve.

There is a series of remarks at the end of the paper concerning the scope of our results and their relation to other work in the field. A brief reading of these remarks in advance would help to provide perspective, and to motivate some of the constructions used in our proof.

## 1. Statement of the problem

We start by giving a precise statement of the simplest and perhaps most basic form of Plateau's problem.

THEOREM A (Douglas and Radó). *Let $\Gamma$ be a rectifiable Jordan curve in* $\mathbf{R}^3$. *Then there exists a map*

$$x = h(u) \colon \Delta \to \mathbf{R}^3$$

*where $\Delta$ is the closed unit disk in $\mathbf{R}^2$, satisfying*

( i ) *$h$ is continuous in $\Delta$;*

(ii) *$h$ maps the boundary of $\Delta$ one-to-one onto $\Gamma$;*

[1] Preparation of this article was supported in part by the Army Research Office (Durham).

(iii) *$h \in C^\infty$ (in fact, $h$ is harmonic) in the interior of $\Delta$, and with the possible exception of isolated points, $h$ maps the interior of $\Delta$ conformally onto a regular minimal surface;*

(iv) *the image of $\Delta$ under $h$ has minimum area among all maps of $\Delta \to \mathbf{R}^3$ which are piecewise smooth in the interior and satisfy* (i) *and* (ii).

For a proof, see [17, p. 90], or [5, p. 100].

There are several comments to be made on the statement of this basic result.

Conditions (i) and (ii) together interpret the statement that the surface defined by $h$ "spans $\Gamma$". Condition (iv) says that "$h$ minimizes the area among all piecewise smooth surfaces spanning $\Gamma$". By "piecewise smooth," we mean continuously differentiable except for possibly a finite number of isolated points or smooth arcs. (Actually, $h$ minimizes area in a much larger class of maps, but this simplified formulation is sufficient for our purposes.)

Condition (iii) may be put in analytic form as follows. We introduce coordinates

$$x = (x_1, x_2, x_3) \qquad \text{in } \mathbf{R}^3 ,$$
$$u = (u_1, u_2) \qquad \text{in } \mathbf{R}^2$$

and the complex coodinate

$$w = u_1 + iu_2 .$$

Then $\Delta$ is the disk $|w| \leqq 1$, and $h(u)$ harmonic means that each of the coordinate functions $h_k(u_1, u_2)$, $k = 1, 2, 3$, is harmonic. This is equivalent to the statement that each of the functions

$$\varphi_k(w) = \frac{\partial h_k}{\partial u_1} - i\frac{\partial h_k}{\partial u_2} \tag{1.1}$$

is complex analytic. In terms of the basic expressions of differential geometry,

$$g_{ij} = \frac{\partial h}{\partial u_i}\cdot\frac{\partial h}{\partial u_j} ,$$

we have

$$\textstyle\sum_{k=1}^{3} \varphi_k^2 = g_{11} - g_{22} - 2ig_{12} \tag{1.2}$$

and

$$\textstyle\sum_{k=1}^{3} |\varphi_k|^2 = g_{11} + g_{22} . \tag{1.3}$$

We may now give the precise formulation of condition (iii).

(iii′) The functions $\varphi_k$ defined by (1.1) are analytic in $|w| < 1$ and

$$\varphi_1^2 + \varphi_2^2 + \varphi_3^2 \equiv 0 . \tag{1.4}$$

By virtue of (1.2), equation (1.4) is equivalent to

$$g_{11} \equiv g_{22},\ g_{12} \equiv 0 .$$

This implies that (1.3) takes the form

$$\textstyle\sum_{k=1}^{3} |\varphi_k|^2 = 2\sqrt{\det (g_{ij})} .$$

Thus the condition $\det (g_{ij}) \neq 0$ which guarantees the regularity of the surface $h(u)$ can fail only at points where all three functions $\varphi_k$ vanish simultaneously. Such points are necessarily isolated unless all the $\varphi_k$ vanish identically which would mean all the $h_k$ are constant, violating condition (ii). At all other points, $h(u)$ defines a regular surface in $\mathbf{R}^3$, equation (1.4) tells us that $(u_1, u_2)$ are isothermal parameters on this surface, and a classical result states that a surface given in isothermal parameters is *minimal* if and only if the coordinate functions are harmonic. This shows the equivalence of (iii) and (iii′).

The points where $\varphi_k$ vanish simultaneously are called *branch points* of the surface. Since $\det(g_{ij}) = 0$ at these points (in fact, all $g_{ij} = 0$), the methods of classical differential geometry fail to give any information on the surface in the vicinity of such a point. However, it is possible to obtain a complete picture in the case of branch points of a minimal surface.

We proceed in the usual manner by defining functions

$$f = \frac{\varphi_1 - i\varphi_2}{2}, \quad g = \frac{\varphi_3}{\varphi_1 - i\varphi_2} .$$

Then $f$ is analytic and $g$ is meromorphic in $|w| < 1$. (The only exception is when $\varphi_1 \equiv i\varphi_2$ in which case (1.4) implies $\varphi_3 \equiv 0$, so that $h_3$ is constant and the surface lies in a horizontal plane. We shall assume throughout that the surface does not lie in a plane. In this connection, see Remark 5.3.) It follows from (1.4) that

$$\begin{aligned} \varphi_1 &= f(1 - g^2) , \\ \varphi_2 &= if(1 + g^2) , \\ \varphi_3 &= 2fg . \end{aligned} \tag{1.5}$$

The function $g$ has the interpretation at regular points that it is the value under stereographic projection of the image of the unit normal to the surface. (See, for example, [14, § 8].) The assumption that the surface does not lie in a plane means that $g$ is non-constant. Since $g$ is meromorphic in $|w| < 1$, we see that the unit normal tends to a well-defined limit at each branch point. By a rotation of the surface in space, we may assume that $g = 0$ at the point we wish to consider. Then the point is a branch point if and only if $f = 0$ also. We thus obtain a representation of an arbitrary minimal surface in the neighborhood of a branch point. We shall use this representation in § 3 to obtain a geometric description of the surface in this neighborhood from which we derive our main result: *any surface satisfying conditions* (i)–(iv) *of Theorem* A *is free of branch points*. The idea of the proof will be seen most clearly by following the reasoning for a particular, easily visualized example.

## 2. An example

The simplest example is obtained by choosing $f = w$ and $g = w$ in equations (1.5). Since

$$x_k = \operatorname{Re} \int \varphi_k(w) dw \,, \tag{2.1}$$

the resulting surface would be given (up to a constant multiple) by

$$x_1 + i x_2 = \bar{w}^2 - \frac{1}{2} w^4 \,, \quad x_3 = \operatorname{Re} \frac{4}{3} w^3 \,.$$

Since analytic functions are more familiar than anti-analytic functions, we may reflect in the $x_1$, $x_3$-plane to obtain the surface

$$x_1 + i x_2 = w^2 - \frac{1}{2} \bar{w}^4 \,, \quad x_3 = \operatorname{Re} \frac{4}{3} w^3 \,. \tag{2.2}$$

This simple explicit surface may be analyzed as closely as desired to obtain a complete picture of its behavior in the neighborhood of the branch point. The following facts are worth noting.

(a) The image of a circular sector

$$|w| \leqq R \,, \quad -\frac{\pi}{6} \leqq \arg w \leqq \frac{\pi}{6}$$

for small $R > 0$ is a surface $\Sigma$ with a one-one projection onto the $x_1$, $x_2$-plane. The two radii $\arg w = \pm\pi/6$ map onto straight line segments in the $x_1$, $x_2$-plane, making an angle of $2\pi/3$ at the origin. Thus $\Sigma$ is bounded by these two line segments and an arc $\gamma$ joining them.

(b) The image of the entire disk is obtained by successive reflections of $\Sigma$ in the boundary line segments. This follows from general facts about minimal surfaces, and can easily be seen in this case by explicit computation.

(c) The image of the disk $|w| < R$ has a double projection onto a domain $D$ in the $x_1$, $x_2$-plane. Each point of $D$ except the origin is the image of exactly two points in $|w| < R$, while the origin corresponds to the origin. The map of $0 < |w| < R$ into $D$ has positive jacobian everywhere, and the image of each circle $|w| = r$ with $0 < r < R$ is a curve in $D$ winding twice about the origin. Thus the behavior of this map is exactly like that of the map $w^2$. The surface therefore represents exactly what is usually pictured as a two-sheeted surface with a branch point at the origin.

Our next object will be to show that a surface of this sort cannot minimize area with respect to its boundary. For this purpose, consider the sequence of mappings indicated in Figure 1.

The map $\tilde{H}$ is the composition of the maps $F$, $G$, and $H$, which are defined as follows. The map $H$ is simply the restriction of (2.2) to the unit disk $|w| \leqq 1$.

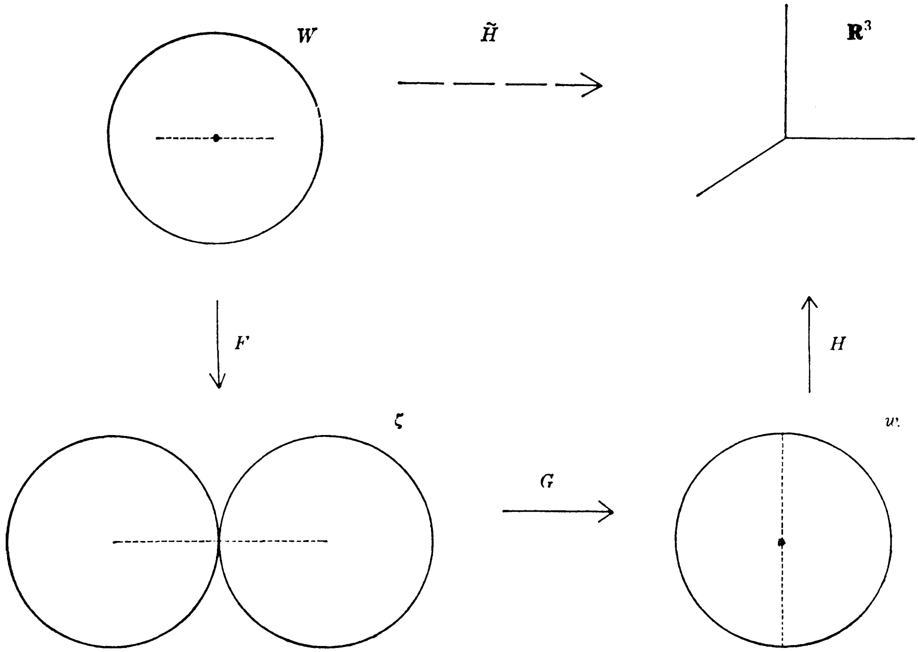


FIGURE 1

The map $G$ is defined in the union of the two disks

$$|\zeta - 1| \leqq 1 \, , \quad |\zeta + 1| \leqq 1 \, . \tag{2.3}$$

In the right-hand disk, we set

$$\zeta = 1 + re^{i\vartheta} \, , \qquad 0 \leqq r \leqq 1, \, -\pi < \vartheta \leqq \pi \, ,$$

and then

$$G(\zeta) = re^{i\vartheta/2} \, . \tag{2.4}$$

In the left-hand disk, we set

$$\zeta = -1 + re^{i\vartheta} \, , \qquad 0 \leqq r \leqq 1, \, 0 \leqq \vartheta < 2\pi$$

and

$$G(\zeta) = ire^{i\vartheta/2} \, . \tag{2.5}$$

Then $G$ maps the right-hand disk into the right half-disk $|w| \leqq 1$, Re $w \geqq 0$, and the left-hand disk into the left half-disk. It is differentiable everywhere except along the line segment

$$-1 < \operatorname{Re} \zeta < 1, \operatorname{Im} \zeta = 0 . \tag{2.6}$$

Approaching a point $t$ of this segment may yield two different limit values, $\pm it$, for the image under $G$. However, both of these points have the same image under $H$. It follows that the composed map $H \circ G$ is continuous everywhere, and differentiable except on the line segment (2.6).

Finally the map $F$ is to be a continuous map of the disk $|W| \leqq 1$ onto the union of the two disks (2.3). It is to have the following properties.

1. $F$ maps the right half-disk $|W| \leqq 1$, $\operatorname{Re} W \geqq 0$ onto the right-hand disk, and the left half-disk onto the left-hand disk.
2. $F$ is differentiable except along the line $\operatorname{Re} W = 0$.
3. $G \circ F$ is the identity on the boundary $|W| = 1$.

It is easy to construct such a map. For example,

$$F(\cos \vartheta + it \sin \vartheta) = \begin{cases} 1 + \cos 2\vartheta + it \sin 2\vartheta\,, \\ \qquad\qquad -1 \leqq t \leqq 1,\ 0 \leqq \vartheta \leqq \dfrac{\pi}{2} \\ -1 - \cos 2\vartheta - it \sin 2\vartheta\,, \\ \qquad\qquad -1 \leqq t \leqq 1,\ \dfrac{\pi}{2} \leqq \vartheta \leqq \pi\,. \end{cases} \tag{2.7}$$

The only difficulty is at the point $z = -i$, where the composed map $G \circ F$ is not the identity. However, we may summarize the situation as follows.

The map $\tilde{H} = H \circ G \circ F$ is a piecewise differentiable map of $|W| \leqq 1$ into $\mathbf{R}^3$. The maps $H$ and $\tilde{H}$ coincide on the unit circle $e^{i\vartheta}$, $0 \leqq \vartheta \leqq 2\pi$. The restrictions of $H$ and $\tilde{H}$ to the right half-disk are simply different parametrizations of the same surface $S_1$ in $\mathbf{R}^3$. Similarly, the restrictions of $H$ and $\tilde{H}$ to the left-hand disk map onto the same surface $S_2$. It happens here that $S_2$ is the reflection of $S_1$ in the $x_1$, $x_2$-plane, and that both $S_1$ and $S_2$ are bounded by Jordan curves and have one-one projections onto the simply-connected domain $D$ mentioned earlier. However, the only important fact is that $S_1$ and $S_2$ each contain in their interior a straight line segment along which the original surface had a self-intersection (namely, the negative $x_1$-axis). If $W_0$ is a point such that $F(W_0)$ lies on the line segment (2.6) and is different from zero, then the map $\tilde{H}$ will map a neighborhood of $W_0$ onto a surface which consists of two halves having distinct tangent planes. It is then obvious that we can modify the map $\tilde{H}$ in a neighborhood of $W_0$ to obtain an image surface with strictly smaller area. The new surface is again given by a piecewise smooth map $\hat{H}$

of $|W| \leqq 1$ into $\mathbf{R}^3$, and it coincides on the boundary with $H$.

The above construction will serve as a model for the procedure to be used for arbitrary branch points.

## 3. The behavior of a minimal surface near an arbitrary branch point

We now return to equations (1.5), and use them to obtain an essentially complete description of the behavior of a minimal surface in the neighborhood of a branch point. Our results partially overlap those obtained by Chen [3], who also analyzed behavior at branch points, using a somewhat different reasoning.

We may clearly assume that the point to be considered lies at the origin, and is the image of the origin in the $w$-plane. We introduce the functions

$$F(w) = \int f(w)dw = \sum_{n=l}^{\infty} a_n w^n , \tag{3.1}$$

$$G(w) = \int f(w)g(w)^2 dw = \sum_{n=l+2m}^{\infty} b_n w^n , \tag{3.2}$$

$$H(w) = 2\int f(w)g(w)dw = \sum_{n=l+m}^{\infty} c_n w^n . \tag{3.3}$$

By our normalizations, including the assumption $g(0) = 0$ mentioned at the end of § 1, we have

$$a_l \neq 0,\ b_{l+2m} \neq 0 , \quad c_{l+m} \neq 0 , \qquad \text{where } l \geqq 2,\ m \geqq 1 . \tag{3.4}$$

It then follows from (2.1) and (1.5) that the surface is represented in a neighborhood of $w = 0$ by

$$x_1 + ix_2 = \overline{F(w)} - G(w) , \quad x_3 = \operatorname{Re} H(w) . \tag{3.5}$$

As in the previous section, we bow to our greater familiarity with analytic functions by reflecting the surface (3.5) in the $x_1, x_3$-plane, obtaining the surface

$$x_1 + ix_2 = F(w) - \overline{G(w)} , \quad x_3 = \operatorname{Re} H(w) . \tag{3.6}$$

We introduce the complex variable

$$z = x_1 + ix_2 ,$$

and start by investigating the map $w \to z$.

PROPOSITION 3.1. *The map*

$$z = F(w) - \overline{G(w)} \tag{3.7}$$

*has the property that for all sufficiently small $\varepsilon > 0$, the inverse image of $|z| < \varepsilon$ is a domain $D$ such that the map* (3.7) *has positive jocobian in $D$*

*except at the origin. Each point of* $0 < |z| < \varepsilon$ *is the image of* $l$ *distinct points in* $D$. *For each* $\delta$, $0 < \delta < \varepsilon$, *the pre-image of* $|z| = \delta$ *is a Jordan curve in* $D$ *enclosing the origin.*

PROOF. Using

$$\frac{\partial(x_1, x_2)}{\partial(u_1, u_2)} = |z_w|^2 - |z_{\bar{w}}|^2 = |F'(w)|^2 - |G'(w)|^2 ,$$

we find that this quantity is positive for all small $w \neq 0$ by virtue of (3.1), (3.2), and (3.4). Similarly, the fact that $|G(w)| < |F(w)|$ for all small $w \neq 0$ implies first of all that $z \neq 0$ for all small $w \neq 0$, and second, that the winding number about the origin of the image of all small circles $|w| = \eta$ is equal to $l$. This means that the degree of the map is $l$, and that $l$ distinct points in the neighborhood of $w = 0$ map onto each small $z \neq 0$. Finally, the fact that

$$\lim_{w \to 0} \frac{F(w)}{a_l w^l} = 1$$

shows that the inverse image of $|z| = \delta$ lies in an annulus about $|w| = (\delta/|a_l|)^{1/l}$, and that as $z$ traverses the circle $|z| = \delta$ once in the positive direction, each branch of the inverse will describe an arc whose argument increases by approximately $2\pi/l$. Combining these facts, we obtain the result stated.

Note that the description of the mapping (3.7) given in Proposition 3.1 shows that the image of the domain $D$ on the surface (3.6) is an $l$-sheeted covering of the disk $|z| < \varepsilon$. In other words, a minimal surface in the neighborhood of a branch point provides a realization of what we picture abstractly as a neighborhood of the branch point on the Riemann surface $z^{1/l}$.

Our next object is to show the existence of a "branch line" on the surface. This will be a curve of self-intersection extending from the origin. In order to do this, and to better study the surface in a neighborhood of the branch point, we make a special choice of the parameter $w$, which has up to now been arbitrary. We note that

$$g(w) = w^m g_1(w),\ g_1(0) \neq 0 .$$

Choosing a single-valued branch of $[g_1(w)]^{1/m}$ in a neighborhood of the origin, we set

$$\tilde{w} = w[g_1(w)]^{1/m} . \tag{3.8}$$

Then

$$g(w) = \tilde{w}^m .$$

Since the map (3.8) is a conformal map in the neighborhood of the origin, we may use $\tilde{w}$ as a local complex parameter, and we obtain corresponding functions $\tilde{\varphi}_k(\tilde{w})$, $\tilde{f}(\tilde{w})$, $\tilde{g}(\tilde{w})$. But since the function $g$ is a point function on the surface determined by the unit normal and independent of parameter, it follows that

$$\tilde{g}(\tilde{w}) = g(w) = \tilde{w}^m .$$

We shall drop the tildes and simply assume from now on that a parameter has been chosen in a neighborhood of the origin so that

$$g(w) = w^m . \tag{3.9}$$

We retain the notation (3.1)–(3.4), from which it follows that

$$f(w) = F'(w) = \sum_{n=l}^{\infty} n a_n w^{n-1} , \tag{3.10}$$

$$G(w) = \int \sum_{n=l}^{\infty} n a_n w^{n+2m-1} dw = \sum_{n=l}^{\infty} \frac{n}{n+2m} a_n w^{n+2m} , \tag{3.11}$$

and

$$H(w) = 2\int \sum n a_n w^{n+m-1} dw = \sum_{n=l}^{\infty} \frac{n}{n+m} a_n w^{n+m} . \tag{3.12}$$

Thus (3.6) takes the form

$$\begin{aligned} z = x_1 + ix_2 &= a_l w^l + \cdots + a_{l+2m-1} w^{l+2m-1} \\ &\quad + \left(a_{l+2m} w^{l+2m} - \frac{l}{l+2m} \bar{a}_l \bar{w}^{l+2m}\right) + \cdots , \end{aligned} \tag{3.13}$$

$$x_3 = \operatorname{Re}\left\{\frac{l}{l+m} a_l w^{l+m} + \cdots\right\} . \tag{3.14}$$

The main purpose of this particular choice of parameter is that it allows us to recognize immediately a "false branch point." If a minimal surface is regular at $w = 0$ and we set $w = \zeta^2$, we obtain a representation with an apparent branch point at $\zeta = 0$. It will indeed be a branch point in the sense of minimal surfaces, and Proposition 3.1, for example, will remain valid. However, to understand the geometric behavior of the surface itself, we must separate the trivial case of branch points arising from such a parametrization from what we may call "true branch points." This will be done in Proposition 3.2.

LEMMA 3.1. *Let $\Phi(w)$ be a complex function of the complex variable $w$ in a neighborhood of the origin, and let $L$, $M$ be positive integers. Suppose*

$$\lim_{w \to 0} \frac{\Phi(w)}{w^L} = c \neq 0 .$$

*Then $\exists \delta > 0$ such that if $w_j = R_j e^{i\varphi_j}$, $j = 1, 2$, where*

$$0 < R_j \leqq \delta\ , \quad \varphi_1 - \varphi_2 = 2\pi\frac{k}{M} \qquad \textit{for some } k = 1, \cdots, M-1\ ,$$

*then* $\arg\Phi(w_1) \neq \arg\Phi(w_2)$ *unless $M$ divides $kL$.*

PROOF. Choose $\delta > 0$ such that for $0 < |w| \leqq \delta$ we have

$$\left|\frac{\Phi(w)}{w^L} - c\right| < |c| \sin\frac{\pi}{M}\ .$$

Then $|\Phi(w) - cw^L| < |cw^L| \sin(\pi/M)$, or geometrically, $\Phi(w)$ lies inside a circle of radius $|cw^L| \sin(\pi/M)$ centered at the point $cw^L$. This implies that

$$|\arg\Phi(w) - \arg cw^L| < \frac{\pi}{M}\ .$$

But

$$\begin{aligned}|\arg\Phi(w_1) - \arg\Phi(w_2)| &\geqq |\arg cw_1^L - \arg cw_2^L| \\ &\quad - |\arg\Phi(w_1) - \arg cw_1^L| - |\arg\Phi(w_2) - \arg cw_2^L|\ .\end{aligned}$$

The last two terms are strictly less than $\pi/M$, while the first term on the right is at least equal to $2\pi/M$ unless $M$ divides $kL$. This proves the lemma.

PROPOSITION 3.2. *Let a minimal surface $h(w)$ given in normalized form* (3.13), (3.14) *have a branch point at the origin* (i.e., $l \geqq 2$). *Then either*

(a) $\exists\delta > 0$ *such that* $0 < |w_1| < \delta$, $0 < |w_2| < \delta$, $h(w_1) = h(w_2)$, *and* $w_1 \neq w_2 \Rightarrow g(w_1) \neq g(w_2)$, *or*

(b) $h(w) \equiv h(e^{2\pi i/\nu}w)$ *for some integer* $\nu \geqq 2$.

PROOF. Suppose (a) does not hold. Then choose sequences $w_n$, $w'_n$ such that $w_n \to 0$, $h(w_n) = h(w'_n)$ and $g(w_n) = g(w'_n)$. By (3.9), $w'_n = \alpha_n w_n$, where $\alpha_n$ is an $m^{\text{th}}$ root of unity. By choosing a subsequence we may assume all $\alpha_n$ to be the same, say $\alpha$. Thus, in particular

$$z(w_n) = z(\alpha w_n)\ , \quad \alpha = e^{2\pi ik/m}\ , \qquad 1 \leqq k \leqq m-1\ .$$

Using expansion (3.13), it follows from Lemma 3.1 that $m\,|\,kl$. Hence $(\alpha w_n)^l = w_n^l$. Let $\nu$ be the smallest positive integer such that $m\,|\,k\nu$. Then $2 \leqq \nu \leqq l$. Hence $m \nmid kn$ for $n = l+1, \cdots, l+\nu-1$. We next apply Lemma 3.1 to the function

$$\Phi(w) = z(w) - a_l w^l$$

which takes on the same values at $w_n$ and $\alpha w_n$, and we deduce inductively that

$$a_n = 0 \qquad \text{for } n = l+1, \cdots, l+\nu-1\ .$$

We may then apply Lemma 3.1 to the function

$$\Phi(w) = z(w) - a_l w^l - a_{l+\nu} w^{l+\nu}$$

which also takes on the same values at $w_n$ and $\alpha w_n$, and repeating this process, we find

$$a_n = 0 \qquad \text{for } n = l+1, \cdots, l+2m-1$$
$$\textit{except} \text{ when } n = l + p\nu,\ p \text{ an integer}\ .$$

Examining the expansion (3.13) of $z(w)$ once more, we note first that $w^{l+2m}$ and $\bar{w}^{l+2m}$ take on the same values at $w_n$ and $\alpha w_n$. Furthermore, the terms involving

$$\bar{w}^{l+2m+1}, \cdots, \bar{w}^{l+2m+\nu-1}$$

all vanish, since their coefficients are multiples of

$$\bar{a}_{l+1}, \cdots, \bar{a}_{l+\nu-1}\ .$$

We may therefore re-apply Lemma 3.1 to show that

$$a_n = 0 \qquad \text{for } n = l+2m+1, \cdots, l+2m+\nu-1\ .$$

Proceeding inductively yields

$$a_n = 0 \qquad \text{except when } n = l + p\nu,\ p = 0, 1, 2, \cdots\ . \tag{3.15}$$

Finally, let us determine $\nu$ precisely. First of all, writing $l = p\nu + q$, $0 \leqq q < \nu$, and using the defining property of $\nu$ together with the fact $m \mid kl$, we see that $q = 0$, or $\nu \mid l$. Next, if $d = \text{g.c.d.}\,(k, m)$, then $k = dk_1$, $m = dm_1$, where $k_1$, $m_1$ are relatively prime. For any integer $r$,

$$m \mid kr \Longleftrightarrow m_1 \mid k_1 r \Longleftrightarrow m_1 \mid r\ .$$

This means, in view of the definition of $\nu$, that

$$\nu = m_1 = \frac{m}{d}\ .$$

In particular, $\nu$ divides both $l$ and $m$. Comparing the expansions (3.1), (3.11), (3.12) with (3.15), we see that the functions $F$, $G$, and $H$ are in fact power series in $w^\nu$. This means that (b) holds, and the proposition is proved.

PROPOSITION 3.3. *Every branch point of a minimal surface possesses at least one "branch line"*; i.e., *an analytic arc emanating from the point along which the surface intersects itself.*

PROOF. (See also Remark 5.7.) In case (b) of Proposition 3.2, the statement is trivially true. The map from the parameter plane covers the same image set $\nu$ times, and any two rays from the origin making an angle of $2\pi/\nu$ map onto the same analytic curve on the surface.

Suppose now that case (a) holds. The fact that $g(w_1) \neq g(w_2)$ at $x =$

$h(w_1) = h(w_2)$ implies that the two branches of the surface $h(w)$ which intersect at $x$ have distinct tangent planes there. This means that the intersection in a neighborhood of $x$ consists of an analytic arc, the image of analytic arcs through $w_1$ and $w_2$. Thus the image of $|w| < \delta$ intersects itself only along analytic arcs.

Choose $\varepsilon > 0$ so that the inverse image of $|z| < \varepsilon$ is a domain $D$ lying in $|w| < \delta$, where $D$ is bounded by a Jordan curve $C$. The image of $C$ on the surface is a curve $\Gamma$ which projects onto $|z| = \varepsilon$ in such a way that each point of $|z| = \varepsilon$ is the image of $l$ distinct points on $C$. (See Proposition 3.1.) The curve $\Gamma$ must clearly intersect itself, and the same is true for every curve $\Gamma'$ corresponding to $|z| = \varepsilon'$, for all positive $\varepsilon' < \varepsilon$. It follows that $\Gamma$ can have only a finite number of self-intersections. Namely, if there were an infinite number, then they would have a point of accumulation. The surface would intersect itself along a curve through this point whose projection on the $x_1, x_2$-plane would have to coincide with the circle $|z| = \varepsilon$, by virtue of the fact that both curves are analytic curves. But then for all $\varepsilon'$ sufficiently near $\varepsilon$, the surface would have no self-intersections over the circle $|z| = \varepsilon'$, a contradiction.

Consider now the finite number of points on $|z| = \varepsilon$ where the surface has self-intersections. Through each of these points passes an analytic curve over which the surface intersects itself, and we may follow each of these curves into the interior of $|z| < \varepsilon$. There are only two possibilities for a given curve. Either it leads to the origin or else it again intersects the circle $|z| = \varepsilon$. We assert that the former must hold for at least one curve. Suppose not. Then we have a finite number of arcs over which the surface intersects itself, joining pairwise the points on $|z| = \varepsilon$ over which the surface intersects itself. Each such arc divides the disk $|z| < \varepsilon$ in two parts, one of which contains the origin. The intersection of the parts containing the origin is a domain $E$ bounded by a finite number of arcs, some lying in $|z| = \varepsilon$, and others interior to $|z| < \varepsilon$ over which the surface intersects itself. About each boundary point we may find a disk over which the surface has no self-intersections other than possibly over the boundary itself. Covering with a finite number of such disks, we get a strip along the boundary of $E$ in which we can find a closed curve $\gamma$ about the origin, such that the surface has no self-intersections over $\gamma$. However, the inverse image of $\gamma$ in $D$ is a closed curve about the origin mapping $l$-to-1 onto $\gamma$ (as with $|z| = \varepsilon$), and there must obviously be at least one point of self-intersection of the image of this curve. Thus we obtain a contradiction, and the result is proved.

## 4. The main result

THEOREM 4.1. *Let $h(u)$ satisfy the conditions* (i)–(iv) *of Theorem* A *in* § 1. *Then $h$ is an immersion throughout the interior of* $\Delta$; i.e., *the corresponding surface has no branch points.*

PROOF. Suppose there were a branch point. We may assume that it lies at the origin. Either condition (a) of Proposition 3.2 will hold (a statement independent of normalization) or else, after normalization, condition (b) will hold. Let us first rule out the latter. Choose a curve $C_1$ joining $w = 0$ to a point $w = w_1$ where $|w_1| = 1$, such that $h$ has no branch points along $C_1$ other than the origin. Condition (b) of Proposition 3.2 implies that there is a well-defined branch of $h^{-1} \circ h$ in a neighborhood of the origin, mapping $C_1$ onto a curve $C_2$ making an angle with $C_1$ at the origin of $2\pi/\nu$. (This fact is invariant under the local conformal mapping used to normalize $g(w)$ at the origin.) Since $h$ has no branch points along $C_1$, the map $h^{-1} \circ h$ may be extended to map all of $C_1$ onto a curve $C_2$ joining $w = 0$ to a point $w_2$ on the boundary. But property (ii) of Theorem A implies that $w_2 = w_1$. Then $C_1$ and $C_2$ together form a closed curve $\tilde{C}$ which separates the disk $\Delta$ in such a manner that one component of the complement contains the whole boundary. But if $w_3$ is a boundary point of $\Delta$ different from $w_1$, we can find a curve on the surface joining $h(0)$ to $h(w_3)$ that is disjoint from $h(C_1) = h(C_2)$ except at the origin, which is its only branch point. Again by condition (b) of Proposition 3.2, this curve is the image of a pair of curves $C_3, C_4$ that may be chosen so that they lie on opposite sides of $\tilde{C}$ near the origin. But this is impossible, since only one component of the complement of $\tilde{C}$ can contain the point $w_3$. This contradiction shows that condition (b) of Proposition 3.2 is incompatible with property (ii) of Theorem A.

It remains to rule out condition (a) of Proposition 3.2. More precisely, we show that it is incompatible with property (iv) of Theorem A. The argument follows exactly the lines used in the special example of § 2.

Assume that the branch point satisfies condition (a). By Propositions 3.1 and 3.3, we can find a simply-connected domain $D$ about $w = 0$ which maps under (3.7) onto $|z| < \varepsilon$, and a pair of analytic arcs $C_1, C_2$ from $w = 0$ to points $w_1, w_2$ on the boundary of $D$ such that $h(C_1) = h(C_2)$. The points $w_1, w_2$ divide the boundary into two arcs $\gamma_1, \gamma_2$, each of which is mapped by $h$ onto a closed curve projecting onto $|z| = \varepsilon$. The arcs $C_1$ and $C_2$ divide $D$ into two subdomains $D_1, D_2$ bounded by $C_1, C_2, \gamma_1$ and $C_1, C_2, \gamma_2$ respectively. (See Fig. 2.)

We next introduce a sequence of plane mappings $F_1, F_2, F_3$ defined as follows. The map $F_1$ coincides with the map $F$ defined in (2.7). The map $F_2$

coincides with the map $G$ defined by (2.4) and (2.5). The map $F_3$ is a continuous one-one map of the unit disk onto the domain $D$ such that the left-half disk is mapped diffeomorphically onto $D_1$, the right half-disk is mapped diffeomorphically onto $D_2$, and for each real $t$, $0 \leqq t \leqq 1$, the points $it$, $-it$ map onto those points of $C_1$, $C_2$ respectively, having the same image under $h$. Finally, we let $\tilde{h} = h \circ F_3 \circ F_2 \circ F_1$. Then $\tilde{h}$ and $h \circ F_3$ are both maps of the unit disk onto a portion of our original surface. They agree on the boundary, and the area of the image is the same. However, just as in the special example (2.2), a point on the line segment (2.6) has the property that a neighborhood of it maps onto two portions of the original surface which intersect along $h(C_1)$. By virtue of condition (a) in Proposition 3.2, these two portions

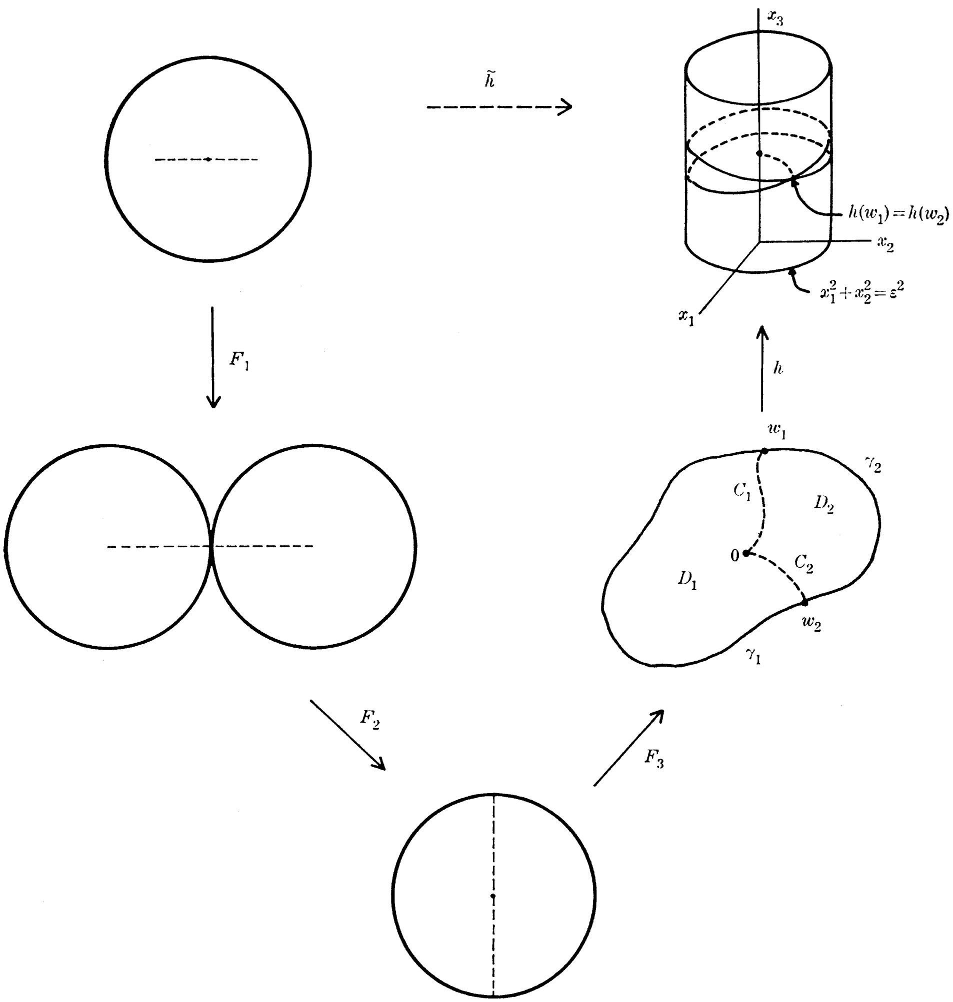


FIGURE 2

are not tangent to each other, so that they meet at an angle. Thus, we may again modify the map $\tilde{h}$ in an arbitrarily small neighborhood of a point $W_0$ so that the resulting map $\hat{h}$ has the same boundary values but strictly smaller area. Then the map of $|w| < 1$ which coincides with $h$ outside of $D$, and with $\hat{h} \circ F_3^{-1}$ in $D$, is a piecewise smooth map of $\Delta$ satisfying conditions (i) and (ii) of Theorem A, but contradicting (iv). This eliminates the possibility of branch points satisfying condition (a) of Proposition 3.2, and the theorem is proved.

## 5. Remarks

5.1. Plateau's problem has been generalized in various ways, chiefly by Douglas. He showed that if $\Gamma$ is an arbitrary Jordan curve (not necessarily rectifiable), then Theorem A still holds, provided property (iv) is replaced by the statement that each interior portion of the image has minimum area with respect to its own boundary curve. (See [17, p. 94].) Clearly, the reasoning of Theorem 4.1 still applies, and again we can assert that there are no branch points. Douglas further formulated and proved existence theorems for minimal surfaces bounded by a finite number of given Jordan curves, and he showed that the solution surfaces had the least area property. (See [8, p. 284]; also Courant [5, p. 166], for rectifiable boundary curves.) The arguments of Theorem 4.1 show that in all cases the solution surfaces are free of branch points.

5.2. One point emphasized by Douglas is that his approach, in contrast to others that had been used (such as Radó's), worked without change in any number of dimensions. In particular, Theorem A holds without change if we replace $\mathbf{R}^3$ by $\mathbf{R}^n$ for any $n \geqq 2$. However, for $n > 3$ it is *not* true that the solution surface must be free of branch points. A simple example is the surface defined by

$$x_1 + ix_2 = w^2, \; x_3 + ix_4 = w^3 . \tag{5.1}$$

Equations (5.1) define a one-one map of $\mathbf{R}^2$ into $\mathbf{R}^4$ whose image is a minimal surface with a branch point at the origin. The image of $|w| = 1$ is a Jordan curve $\Gamma$, and the area of the image of $|w| \leqq 1$ is a *strict* minimum; i.e., any other surface bounded by $\Gamma$ has greater area. (See [9, Th. 4.2 and Remark 4.3].) Thus for this curve $\Gamma$ the Douglas solution surface satisfying Theorem A is unique, and it has an interior branch point. This is a special case of the face that complex analytic curves considered as real two-dimensional surfaces are always minimal surfaces, and as Federer has shown [9], these surfaces always provide a strict minimum of area for their boundary curves.

5.3. The case $n = 2$ in the previous remark arises when $\Gamma$ is a plane curve. Douglas showed in this case that his solution was free of branch points and hence provided a conformal map of the interior of $\Delta$ onto the interior of $\Gamma$, continuous in the closure and one-to-one on the boundary. Throughout the present paper we have assumed that the surface did not lie in a plane. However, if it does, then the map $h$ is a complex analytic map from $w$ to $z$ (or its conjugate), and branch points are the classical branch points of complex function theory. In other words, they satisfy condition (b) of Proposition 3.2, and the argument of Theorem 4.1 still applies.

5.4. A further generalization of Theorem A was proved by Morrey [13] replacing euclidean space $\mathbf{R}^3$ (or $\mathbf{R}^n$) by a riemannian manifold. Morrey's results have been used recently by Lawson [12] to derive a number of results on minimal surfaces in $S^3$, including the existence of compact minimal surfaces of arbitrary genus. For this purpose, Lawson obtains conditions on a curve in $S^n$ which guarantee that Morrey's solution has no branch points [11]. It seems clear geometrically that the arguments used in Theorem 4.1 should extend to minimal surfaces in a riemannian manifold, but the absence of a representation such as (1.5) would necessitate finding suitable replacements for Propositions 3.1–3.3. It would certainly be of interest to know whether Morrey's solutions are also free of branch points for all three-dimensional manifolds.

5.5. In the past decade an intensive study of Plateau's problem has been made using measure-theoretic rather than parametric methods. This has led to the first significant results for higher-dimensional manifolds. But even in the classical case of two-dimensional surfaces in $\mathbf{R}^3$, Reifenberg [18, 20] and Fleming [10] were able to give the first existence proofs for branch-point free minimal surfaces spanning a given boundary in a certain prescribed sense. (See also the discussion in Almgren [1].) Their methods do not allow one to prescribe the topological structure of the solution surface, and the boundary behavior cannot in general be described in the simple strong fashion of Theorem A. It would appear that each method has its advantages and disadvantages, and that the results obtained are complementary to each other.

5.6. It is clear from equations (3.1)–(3.4) that although the surface (3.6) does not have a tangent plane in the sense of classical differential geometry, the $x_1$, $x_2$-plane is a tangent plane in any reasonable generalized sense. For example, the tangent cone to this surface at the origin in the sense of geometric measure theory consists of the $x_1$, $x_2$-plane with multiplicity $l$. Thus, the normalization $g(0) = 0$ amounts to a rotation making the tangent plane horizontal, and Proposition 3.1 describes the way a minimal surface projects onto its tangent plane in the neighborhood of an arbitrary branch point, using an arbitrary local parameter.

5.7. Proposition 3.3 may be replaced "in general" by a much stronger statement with a simpler proof. To see this, we note that equation (3.13) can be inverted in the form

$$w = \left(\frac{z}{a_l}\right)^{1/l} + O(|z|^{2/l})$$

which implies by (3.14) that

$$x_3 = \operatorname{Re}\left\{\frac{l}{l+m} a_l^{-m/l} z^{1+m/l}\right\} + O(|z|^{1+(m+1)/l}) \,. \tag{5.2}$$

This equation gives a very good first-order description of a branch point in non-parametric form. The fact mentioned in the previous remark that the $x_1, x_2$-plane may be considered as the tangent plane is most perspicuous from (5.2), since the exponent of $z$ is strictly greater than 1. By introducing polar coordinates $z = re^{i\vartheta}$, and examining the variation of $x_3$ for small $r$ as $\vartheta$ goes from 0 to $2\pi l$, one is able to show that *in the case that $l$ and $m$ are relatively prime, there must be $(l-1)(l+m)$ curves of self-intersection, each of which approaches the origin at a definite, easily-calculated angle.* This fact was first pointed out by Chen [3, p. 796]. In particular, this is always true if $m = 1$, which we may consider to be the "general case." Namely, the function $g$ defining the Gauss map will be locally one-one except at isolated points where $m > 1$ and the Gauss map has a branch point. In general, even if the surface has a branch point, the Gauss map will be locally one-one, as is the case for the example in § 2. Thus, it is only when *both* the surface *and* the Gauss map have branch points that any difficulty can occur. In other words, when both $l \geqq 2$ and $m \geqq 2$. Even then there is no trouble if $l$ and $m$ are relatively prime. More generally, we may deduce from (5.2) the existence of at least one branch line approaching the origin at a definite angle in every case except where $m$ is actually a multiple of $l$. In that case the lowest order terms in (5.2) define a finite number of copies of the same surface, and the self-intersections are determined by the higher-order terms. It would seem worthwhile to determine the precise behavior of the branch lines in this case also, and replace Proposition 3.3 by a more precise description.

Let us note that the representation (5.2) is a special case of results of Bers [2] on isolated singularities of the minimal surface equation. Bers does not seem to have noticed that his results include the behavior of branch points of parametric surfaces, for which he refers to the above-mentioned paper of Chen.

5.8. We conclude with some historical remarks concerning earlier discussions of branch points. Of the three main contributors to the classical work on Plateau's problem, Douglas, Radó, and Courant, only Radó seems to have avoided major misconceptions about branch points.

Courant discusses the subject in [4]. He makes the statement (p. 46), "Apparently no instance of a minimal surface with branch point has been pointed out explicitly in the literature". This is remarkable in view of the fact that all the examples one wants are obtained by substituting functions $f$ and $g$ in the classical representation formulas (1.5). In fact, Radó specifically notes [17, p. 30] that the representation (1.5) has the advantage that it represents locally *all* branch points, in contrast to another "Weierstrass representation" that does not.

Courant goes on to give a geometric construction for a class of surfaces with branch points, all of which have the same general structure as the surface (2.2). In fact, that surface is included as a special case. He further states that at least some of the surfaces so constructed furnish an absolute minimum of area with respect to their boundary curve, but the special argument given in § 2, without going to the general case, shows that to be false.

Similarly, Douglas was convinced that branch points must occur in his solution surfaces for certain boundary curves [7, pp. 733, 739, 753]. Namely, he considers a curve $\Gamma$ in the form of a double loop (Fig. 3). He asserts that "the surface $M$ determined by that contour is readily visualized to have the general form of the Riemann surface for $z^{1/2}$". Radó [17, p. 109] discusses this example and gives several good reasons for questioning Douglas' statement. We should like to add some further geometric reasons here that show, we believe, precisely where Douglas' intuition went astray, and that further provided the main idea for the argument given in Theorem 4.1.

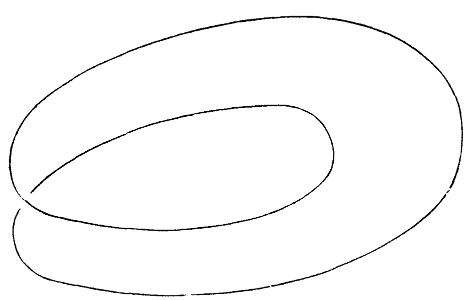

FIGURE 3

Consider two planes intersecting along a line, and from some point of the line of intersection draw two loops, one in each plane. Then separate them slightly at the point of intersection so that they form a single curve like that pictured in Figure 3 which almost crosses itself. When the angle between the planes is fairly large, it is quite clear that the surface of minimum area consists essentially of two plane regions, filling out the loops, joined by a narrow "bridge". The branched surface considered by Douglas would be essentially the union of two curvilinear cone-like surfaces bounding the two loops, and would clearly have greater area. As the angle between the planes tends to zero, the areas of the two types of surface would tend to the same limit, and it may seem less clear which has the smallest area. However, if the curve actually crosses itself, then clearly the two plane regions have less total area than any branched surface spanning the curve. One gets a parametrization of the union of the two plane domains by taking a disk, squeezing it together in the center (as in the map $F$ of Figure 1) and then twisting it about the point of intersection. It is *this* surface which gives the least area, and *not* the branched surface, no matter how small the angle between the two planes. Since a minimal surface always intersects itself near a branch point, the same general procedure can always be used to find a surface of smaller area: map the boundary curve by "squeezing and twisting" and then fill in each loop by a separate minimal surface. Although the actual reasoning we use is somewhat different, this idea was the main motivation for our proof.

*Added in proof.* The author would like to thank H. B. Lawson and J. C. C. Nitsche for the following comments concerning Proposition 3.3 and Theorem 4.1 respectively.

In the proof of Proposition 3.3, the domain $E$ should be defined as the connected component containing the origin of the complement with respect to $|z| < \varepsilon$ of the union of the finite number of arcs intersecting $|z| = \varepsilon$ over which the surface intersects itself. Furthermore, for the subsequent part of the proof, it is necessary to know that no curve over which the surface intersects itself can lie entirely in the interior of $|z| < \varepsilon$. However, such a curve would either have to tend in one direction to $z = 0$, which would prove the proposition, or else must be a closed curve. But it is not hard to prove that the surface cannot intersect itself over a closed curve lying in $|z| < \varepsilon$ unless it satisfies case (b) of Proposition 3.2.

In the first paragraph of the proof of Theorem 4.1, one should choose the point $w_1$ in such a manner that $h(w_1)$ does not coincide with the image under $h$ of any interior point of $\Delta$. It is then clear that $C_2$ must go to the boundary.

The same condition should be imposed on $w_3$.

STANFORD UNIVERSITY

## REFERENCES

[1] F. J. ALMGREN, JR. *Some interior regularity theorems for minimal surfaces and an extension of Bernstein's theorem*, Ann. of Math. **84** (1966), 277-292.

[2] L. BERS, *Isolated singularities of minimal surfaces*, Ann. of Math. **53** (1951), 364-386.

[3] Y-W. CHEN, *Branch points, poles and planar points of minimal surfaces in $R^3$*, Ann. of Math. **49** (1948), 790-806.

[4] R. COURANT, *On a generalized form of Plateau's problem*, Trans. Amer. Math. Soc. **50** (1941), 40-47.

[5] ———, Dirichlet's Principle, Conformal Mapping, and Minimal Surfaces. Interscience, New York, 1950.

[6] J. DOUGLAS, *Solution of the problem of Plateau*, Trans. Amer. Math. Soc. **33** (1931), 263-321.

[7] ———, *One-sided minimal surfaces with a given boundary*, Trans. Amer. Math. Soc. **34** (1932), 731-756.

[8] ———, *Minimal surfaces of higher topological structure*, Ann. of Math. **40** (1939), 205-298.

[9] H. FEDERER, *Some theorems on integral currents*, Trans. Amer. Math. Soc. **117** (1965), 43-67.

[10] W. FLEMING, *On the oriented Plateau problem*, Rend. Circ. Mat. Palermo **11** (1962), 69-90.

[11] H. B. LAWSON, JR., *The global behavior of minimal surfaces in $S^n$*, Ann. of Math. to appear.

[12] ———, *Complete minimal surfaces in $S^3$*, Ann. of Math. to appear.

[13] C. B. MORREY, JR., *The problem of Plateau in a riemannian manifold*, Ann. of Math. **49** (1948), 807-851.

[14] R. OSSERMAN, A Survey of Minimal Surfaces, Van Nostrand-Reinhold. New York, 1969.

[15] T. RADÓ, *On Plateau's problem*, Ann. of Math. **31** (1930), 457-469.

[16] ———, *The problem of the least area and the problem of Plateau*, Math. Z. **32** (1930), 763-796.

[17] ———, On the Problem of Plateau, Ergebnisse der Math. und ihrer Grenzgebiete, Vol. 2, Springer, Berlin, 1933.

[18] E. R. REIFENBERG, *Solution of the Plateau problem for m-dimensional surfaces of varying topological type*, Acta Math. **104** (1960), 1-92.

[19] ———, *An epiperimetric inequality related to the analyticity of minimal surfaces*, Ann. of Math. **80** (1964), 1-14.

[20] ———, *On the analyticity of minimal surfaces*, Ann. of Math. **80** (1964), 15-21.

(Received August 5, 1969)