Minimal Surfaces

Michael Beeson

Aug. 9, 2007

Contents

Chapter 1

Introduction

These lectures contain some of the fundamental results needed to read papers on minimal surfaces. I have attempted to make them as self-contained as possible, rather than just a list of references to books containing these results. The title doesn't mention the word "Introduction", although the background assumed here is only calculus and the most elementary parts of the theory of functions of a complex variable. The necessary differential geometry and theory of harmonic functions is introduced and proved. However, a true introduction to minimal surfaces would involve more pictures of examples, and discussion of other results not presented here. For such books, see the list of references. These lectures have a different purpose: to supply proofs that don't constantly refer you to some other place for the details. At present that aim has not been completely achieved; for example Lichtenstein's theorem is not completely proved here.

The first chapters accompanied lectures given at San José in November 2001. Later chapters were added in July, 2007, including the unpublished material from [2]. The bibliography lists a few reference books on the subject. There are many more, which explore different aspects of the theory of minimal surfaces.

1.1 Notation and Basic Concepts

The open unit disk is D; the closed unit disk is $\bar{D}$; the unit circle is S^1.

C^n means possessing n continuous derivatives. C^0 means continuous.

Surfaces of disk type are given by maps $u : \bar{D} \mapsto R^3$. We will suppose they are at least C^3 in D and C^1 on the boundary. Partial derivatives will be denoted by subscripts u_x and u_y. A surface is *regular* at a point (x, y) in D if the tangent plane is well-defined there, i.e. u_x and u_y have nonzero cross product. Surfaces are required to be regular except at isolated points. A "regular point" of u is a point where u is regular. Another way of expressing regularity is that the Jacobian matrix $\nabla u = \langle u_x, u_y \rangle$ has maximal rank two.

A Jordan curve is a continuous one-one map Γ from S^1 into R^3. A reparametrization of Γ is another Jordan curve of the form $\Gamma \circ \phi$, where ϕ is a one-one map of S^1 into S^1.

A surface u is said to be bounded by Γ in case u restricted to S^1 is a reparametrization of Γ.

Plateau's Problem is this: Given a Jordan curve Γ, find a surface of least area bounded by Γ.

The space of all vectors in R^3 which are tangent to the surface u at a regular point (x, y) is a vector space T_p, called the *tangent space*. A basis for the space is formed by u_x and u_y.

The unit normal $N = N(x, y)$ at p is given by

$$N = \frac{u_x \times u_y}{|u_x \times u_y|}$$

We claim that N_x and N_y are tangent vectors. *Proof:* $N \cdot N = 1$. Differentiating, we have $N \cdot N_x = 0$ and $N \cdot N_y = 0$, so N_x and N_y are tangent vectors.

1.2 Weingarten map and fundamental forms

In this section, we fix a point (x, y) at which u is regular, i.e. $u_x \times u_y$ does not vanish. We have assumed that non-regular points of a surface are isolated, by definition.

The *Weingarten map* $S = S(x, y)$ is a linear map of T_p into itself, defined as follows: If $v = v^1 u_x + v^2 u_y$ then

$$Sv = -v^1 N_x - v^2 N_y = -v^i N_i$$

where the repeated subscript implies summation, and v^i means v_i, but considered as a column vector ("contravariant"). (We always sum one raised index times one lowered index.) The letter S is used because this is also known as the "shape operator". In this section we write u_i for the derivative of u with respect to x_i, risking confusion with the components of u, but avoiding double subscripts as in u_{x_i}.

Lemma 1 *The Weingarten map is self-adjoint. That means $Sv \cdot w = v \cdot Sw$.*

Proof: Differentiate $N \cdot u_i = 0$ with respect to x_j. We get

$$N_j \cdot u_i + N \cdot u_{ij} = 0.$$

Therefore

$$N_i \cdot u_j = N_j \cdot u_i$$

since both are equal to $N \cdot u_{ij}$. Hence

$$\begin{aligned} Sv \cdot w &= -N_i v^i \cdot u_j w^j \\ &= -N_i \cdot u_j v^i w^j \\ &= -u_i \cdot N_j v^i w^j \\ &= -u_i v^i \cdot N_j w^j \\ &= v \cdot Sw \end{aligned}$$

That completes the proof of the lemma.

Thus we can define the following three symmetric bilinear forms:

$$\begin{aligned} I(v,w) &:= v \cdot w \\ II(v,w) &:= Sv \cdot w \\ III(v,w) &:= Sv \cdot Sw \end{aligned}$$

These are called the first fundamental form, second fundamental form, and third fundamental form of u.

1.3 Mean Curvature and Gauss Curvature

The geometric meaning of the second fundamental form can best be seen by finding an orthonormal basis in which the Weingarten map S is diagonal. Let κ_1 and κ_2 be the eigenvalues of the matrix of S. Since S is self-adjoint, standard linear algebra tells us that it can be diagonalized, and that the eigenvalues are the minimum and maximum of the Rayleigh quotient $II(v,v)/I(v,v)$. Let p and q form a basis in which S is diagonal. If $\kappa_1 \neq \kappa_2$ then p and q are automatically orthogonal, since $\kappa_1 p \cdot q = Sp \cdot q = p \cdot Sq = \kappa_2 p \cdot q$. If $\kappa_1 = \kappa 2$, however, then any orthogonal unit vectors p and q will do.[1]

Thinking geometrically instead of algebraically, we write $\nabla p N = \kappa_1 p$ instead of the equivalent $S(p) = \kappa_1 p$. The values κ_1 and κ_2 are called the *principal curvatures* of the surface u at the point (x, y). We will next explain the reason for this terminology, but we must first review the basic facts about space curves.

A continuous, locally one-to-one map from an interval to R^3 is called a *space curve*, or just a *curve*. A *reparametrization* of such a curve is of the form $\eta(t) = \gamma(\phi(t))$ for some monotonic function ϕ. (The interval of definition of the reparametrized curve may be different.) If γ is a C^2 space curve, it has an *arc length parametrization* in which the parameter t equals the arc length $\int_0^t |\gamma_t(\xi)|^2 \, d\xi$. Let $T(t)$ be the unit tangent to $\gamma(t)$ in such a parametrization. Then the *curvature* of γ is defined to be $|T_t|$.

Now to explain the connection between curvature and the Weingarten map. Consider planes passing through $P = u(x, y)$ whose normal at P lies in the tangent space, i.e. planes which contain the unit normal N. Each such plane intersects the surface u (technically, the range of u) in a curve. If the plane's normal is $\cos\theta p + \sin\theta q$ then the curvature of the curve turns out to be (as a calculation shows)

$$\kappa(\theta) = \cos^2\theta\kappa_1 + \sin^2\theta\kappa_2.$$

Averaging over all angles θ between 0 and 2π we find that the average curvature of such curves is $(\kappa_1 + \kappa_2)/2$. Accordingly this quantity is called the *mean curvature* of u at (x, y). It is always denoted by H:

$$H = \frac{\kappa_1 + \kappa_2}{2}.$$

[1] This proof is non-constructive, and the consequence is that it does not show that p and q can be chosen to depend continuously on the point in the parameter domain, near a point where $\kappa_1 = \kappa_2$.

The Gauss curvature K is defined to be $\kappa_1\kappa_2$. Note that H and K are the two elementary symmetric functions of κ_1 and κ_2.

We define $W := |u_x \times u_y|$. We have

$$W = \sqrt{|u_x|^2|u_y|^2 - |u_x \cdot u_y|^2}.$$

The *area element* is $W dx\, dy$ and the area is given by

$$A[u] = \int_D W\, dx\, dy.$$

The formulas in this section apply to any surface at any regular point. They are basic to the subject of differential geometry, rather than being specific to minimal surfaces. They are necessary to connect the variation of area with curvature, which is basic to the theory of minimal surfaces.

There are several systems of notation for the coefficients of the fundamental forms. We have

$$\begin{aligned} g_{ij}(z) &:= u_i \cdot u_j \\ b_{ij}(z) &:= -N_i \cdot u_j \\ c_{ij}(z) &:= N_i \cdot N_j \end{aligned}$$

Remember $b_{ij} = b_{ji} = N \cdot u_{ij}$, and of course $g_{ij} = g_{ji}$.

The older (nineteenth-century) notation uses $E, F, G, \mathsf{L}, \mathsf{M}$, and N, defined by

$$G := \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} E & F \\ F & G \end{bmatrix}$$

$$B := \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} \mathsf{L} & \mathsf{M} \\ \mathsf{M} & \mathsf{N} \end{bmatrix}$$

Hopefully, when you see this notation (in old books), you won't confuse N with the unit normal. Here we have used a different font to make the distinction. We have

$$W = \sqrt{g} = \sqrt{EG - F^2} = \sqrt{\det(g_{ij})}$$

We define g^{ij} to be the coefficients of the inverse of G, so

$$\begin{bmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{bmatrix} = \frac{1}{W^2} \begin{bmatrix} G & -F \\ -F & E \end{bmatrix}$$

The *Weingarten equations* tell how to compute N_x and N_y in terms of u_x and u_y:

$$N_i = -b_i^j u_j$$

where $b_i^j = b_{ik}g^{kj}$, and as usual repeated indices imply summation. To find this formula for the b_i^j, first write, with unknown coefficients a_i^j, $N_i = a_i^j u_j$. Now take the dot product with u_k:

$$N_i \cdot u_k = a_i^j u_j \cdot u_k$$

That is,

$$-b_{ik} = a_i^j g_{jk}$$

Now we can solve for the a_i^j by using the inverse matrix of G:

$$\begin{aligned} a_i^j g_{jk} g^{kp} &= a_i^j \delta_j^p \\ &= a_i^p \\ &= -b_{ik} g^{kp} \end{aligned}$$

This is the formula for the coefficients in the Weingarten equations.

Recall that the principal curvatures κ_1 and κ_2 are the eigenvalues of S. In the eigenvalue equation $Sv = \kappa v$, write $v = v^i u_i$, and $w = w^j u_j$, and write out the equation $Sv \cdot w = \kappa v \cdot w$ as

$$\begin{aligned} Sv \cdot w &= -b_{ij} v^j w^i \\ &= \kappa g_{ij} v^i w^j \end{aligned}$$

Since this is true for all $w = (w^1, w^2)$, we have

$$b_{ij} v^i = \kappa g_{kj} v^i.$$

Since $b_{ij} = b_{ji}$, the eigenvalue equation is

$$Bv = \kappa G v$$

(reading $v = (v^1, v^2)$ as a column vector). Thus, κ_1 and κ_2 are the roots of

$$\begin{aligned} \det(b - \kappa G) &= \begin{bmatrix} b_{11} - \kappa g_{11} & b_{12} - \kappa g_{12} \\ b_{21} - \kappa g_{21} & b_{22} - \kappa g_{22} \end{bmatrix} \\ &= \det(g_{ij})\kappa^2 - (b_{11}g_{22} + b_{22}g_{11} - b_{12}g_{21} - b_{21}g_{12})\kappa + \det(b_{ij}) \\ &= \det(g_{ij}) \left[\kappa^2 - (\kappa_1 + \kappa_2)\kappa + \kappa_1\kappa_2\right] \end{aligned}$$

Comparing coefficients of the powers of κ, we obtain

$$\kappa_1\kappa_2 = \frac{\det B}{\det G} = \det G^{-1}B = \det b_i^j$$
$$\kappa_1 + \kappa_2 = trace(G^{-1}B) = b_{ij}g^{ij} = b_1^1 + b_2^2$$

Since we already worked out the Weingarten equations for b_i^j and the formula for the coefficients of G^{-1}, we are finished. Using the nineteenth-century terminology we have

$$H = \frac{\mathsf{L}G + \mathsf{N}E - 2\mathsf{M}F}{2W^2}$$

The Gauss curvature $K = \kappa_1\kappa_2$ is given by

$$K = \frac{\mathsf{LN} - \mathsf{M}^2}{W^2} = \frac{\det(b_{ij})}{\det(g_{ij})} = \det(b_i^j)$$

The *Gauss map* is the unit normal, considered as a map from D to the sphere S^2. The determinant of the Jacobian of this map is K, so the area element is $K\,dx\,dy$. The "Gaussian area" of a surface is $\int KW\,dx\,dy$, the area of its "Gaussian image" on the sphere, counting multiplicities. The Gauss map plays an important role in the theory of minimal surfaces.

1.4 The definition of minimal surface

Surfaces as well as Jordan curves can be reparametrized; if φ is a C^3 diffeomorphism of D we define $u \circ \varphi$ to be a reparametrization of the surface u. It is an exercise in calculus to show that $A[u \circ \varphi] = A[u]$.

A *critical point* of the functional A is a surface u such that the Frechet derivative $DA[u]$ is zero. In less fancy language, this means the following. Let φ be any function from $\bar{D}$ to R, vanishing on S^1, C^3 in D and C^1 in the closed unit disk $\bar{D}$. Consider $u^t = u + t\varphi N$. This is called a *normal variation.* The parameter t is written as a superscript because we we use subscripts for differentiation. The *first variation of area in direction* φ is defined to be

$$DA[u](\varphi) = (d/dt)A[u^t]|_{t=0}.$$

Then u is a critical point of A if and only if $DA[u](\varphi) = 0$ for all φ satisfying the conditions mentioned. The terminology "stationary point" is also used to mean the same as "critical point".[2]

A surface of least area bounded by Γ would be a critical point of A, but not necessarily conversely. There could be relative minima of area which are not absolute minima of area. There can also be "unstable" critical points of area which are not even relative minima.

We now come to the first theorem in the subject of minimal surfaces.

Theorem 1 *The surface u is a critical point of A if and only if its mean curvature H is identically zero.*

Proof. The proof depends on the following formula for the first variation of area:

$$DA[u](\varphi) = -2\int_D HW\varphi\,dx\,dy$$

[2]To use the Frechet derivative we technically need to specify a function space, and prove that the area functional is Frechet-differentiable on that space. We have not done that here, but the reader worried about it can just use the words "first variation of area" instead of "Frechet derivative" for $DA[u]$. Because we are about to derive a simple formula for this first variation, the exact function space we use does not matter much.

Once we have established this formula, the result follows from the so-called "fundamental lemma of the calculus of variations, which says that if f is continuous and $\int_D f(x,y)\varphi(x,y)\,dx\,dy = 0$ for all φ vanishing on S^1 and C^3 in D, then $f(x,y)$ is identically zero. This lemma is itself easy to prove: if f is nonzero at (x,y) in D, say $f(x,y) > 0$, then by continuity there is a neighborhood of (x,y) in which f is positive; and we can choose φ to be positive in that region and vanish outside it (it takes some argument to do this in a C^3 way, but it isn't deep), but this leads immediately to a contradiction. So the proof boils down to proving the stated formula for the first variation.

Let φ vanish on S^1 and be C^3 in $\bar{D}$. Define

$$\tilde{u} = u + t\varphi N.$$

Then

$$A[\tilde{u}] = \int_D \sqrt{\tilde{E}\tilde{G} - \tilde{F}^2}\, dxdy$$

and we have, neglecting terms which are $O(t^2)$,

$$\begin{aligned}
\tilde{E} &= \tilde{u}_x^2 \\
&= (u_x + t\varphi_x N + t\varphi N_x)^2 \\
&= u_x^2 + 2t\varphi N_x u_x \\
&= u_x^2 + 2t\varphi b_{11}
\end{aligned}$$

Similarly

$$\begin{aligned}
\tilde{G} &= \tilde{u}_y^2 \\
&= u_y^2 + 2t\varphi b_{22}
\end{aligned}$$

Then

$$\begin{aligned}
\tilde{E}\tilde{G} &= EG + t\varphi[u_x^2 b_{22} + u_y^2 b_{11}] \\
&= EG + t\varphi[g_{11}b_{22} + g_{22}b_{11}]
\end{aligned}$$

We have

$$\begin{aligned}
\tilde{F} &= \tilde{u}_x \cdot \tilde{u}_y \\
&= (u_x + t\varphi_x N + t\varphi N_x) \cdot (u_y + t\varphi_y N + t\varphi N_y) \\
&= F + t\varphi(N_x u_y + N_y u_x) \\
&= F + t\varphi(b_{12} + b_{21}) \\
\tilde{F}^2 &= F^2 + 2t\varphi[g_{12}b_{12} + g_{21}b_{21}]
\end{aligned}$$

Thus

$$\begin{aligned}
\tilde{E}\tilde{G} - \tilde{F}^2 &= EG - F^2 + 2t\varphi[g_{11}b_{22} + g_{22}b_{11} - g_{12}b_{12} + g_{21}b_{21}] \\
&= EG - F^2 + 2t\varphi W^2 g^{ij} b_{ij}
\end{aligned}$$

in view of the formula for the g^{ij}. But now we recognize the formula for the mean curvature which we calculated using the Weingarten equations!

$$\tilde{E}\tilde{G} - \tilde{F}^2 \quad = \quad EG - F^2 + 2t\varphi W^2 H$$

We have thus proved, remembering $W^2 = EG - F^2$, that

$$\begin{aligned} A[\tilde{u}] &= \int_D \sqrt{W^2(1 + 2t\varphi H)}\, dx\, dy + O(t^2) \\ &= \int_D W\,[1 + t\varphi H]\; dx\, dy + O(t^2) \\ &= \int_D W\, dx\, dy + t \int_D HW\, dx\, dy + O(t^2) \end{aligned}$$

It follows that

$$DA[u](\varphi) \quad = \quad \int_D \varphi HW\, dx\, dy$$

We could have finished this proof another way, recognizing $HW = LG + NE - 2MF$ using the nineteenth-century notation, without using the formula for g^{ij}.

Definition 1 *u is called a* minimal surface *if it has zero mean curvature.*

1.5 Non-parametric minimal surfaces

The surfaces we have defined are sometimes called parametric surfaces, because they are given by a vector function u of parameters x, y in D. By contrast, a surface given by a scalar function $Z = f(x, y)$ is said to be in non-parametric form. Of course, every surface in non-parametric form can be given parametrically by the vector function $(x, y, f(x, y))$, but not conversely.[3]

We now consider surfaces given by $z = f(x, y)$ where (x, y) ranges over some open set Ω in the plane, usually bounded by a Jordan curve. One can calculate the mean curvature of such a surface, using the explicit formulas for H for parametric surfaces, and the parametrization $(x, y, f(x, y))$. One finds

$$H = \frac{(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy}}{W^2}$$

where

$$W^2 = 1 + f_x^2 + f_y^2.$$

Therefore, the equation for the surface to be minimal is

$$(1 + f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1 + f_x^2)f_{yy} = 0.$$

[3]When dealing with parametric surfaces, we use lower-case variables in the parameter domain, e.g. $z = x + iy$, and upper-case variables for the coordinates X,Y, and Z in R^3. When dealing with non-parametric surfaces, we usually use lower-case x and y in place of X and Y, and either z or Z for the third coordinate.

This is the *non-parametric minimal surface equation.* It is nonlinear and elliptic (for those who know something about differential equations).

Alternately, one can derive this equation by considering the first variation of area among non-parametric surfaces. We have $A_\Omega = \int W\,dx\,dy$. Considering the variation in the direction φ, where φ vanishes on the boundary $\partial\Omega$ and is C^3 in Ω, and setting $DA[f](\varphi) = 0$, we find after a calculation that

$$\operatorname{div}\left(\frac{\nabla f}{W}\right) = 0$$

which is another way to write the minimal surface equation.

1.6 Examples of minimal surfaces

Plane. The "trivial" minimal surface is a plane.

Catenoid. This is a surface obtained by rotating a catenoid around the z-axis:

$$r = \alpha \cosh\left(\frac{Z}{\alpha}\right)$$

Taking the parameter x to be Z/α, and y to be the polar angle often written as θ, it can be parametrized by

$$\begin{bmatrix} \alpha \cosh x \cos y \\ -\alpha \cosh x \sin y \\ \alpha x \end{bmatrix}$$

with $-\infty < x < \infty$ and $0 \leq y < 2\pi$.

Helicoid. This can be written in the form $Z = \alpha\theta$, where θ is a polar angle in the XY-plane. Taking $y = \theta$ and $r = \sinh x$, we have the parametrization

$$\begin{bmatrix} \alpha \sinh x \sin y \\ \alpha \sinh x \cos y \\ \alpha y \end{bmatrix}$$

A portion of the helicoid can be written in non-parametric form as $Z = \alpha cosh^{-1}(r\alpha)$, provided the domain Ω does not include the origin, where the boundary values are not continuous and the gradient is not bounded.

Scherk's (first) surface.

$$Z = \ln \frac{\cos y}{\cos x}$$

is defined on the square of side π centered at the origin, and on all "black squares" of the infinite checkerboard containing that square as one of its black squares. It is only defined on those squares since $\cos y$ and $\cos x$ must have the same sign.

Enneper's surface. We give this surface using polar coordinates in the parameter domain, which can be any disk about the origin.

$$u(r, \theta) = \begin{bmatrix} r\cos\theta - \frac{1}{3}r^3\cos 3\theta \\ -r\sin\theta - \frac{1}{3}r^3\sin 3\theta \\ 2r^2\cos 2\theta \end{bmatrix}$$

1.7 Calculus review

In this section we remind the reader of some results in two-dimensional vector calculus. The reader is assumed to have studied vector calculus, which usually includes both two and three dimensional calculus, but there is a gap between completing the course and having all the formulas at your fingertips without having to think about them. Here we concentrate on the two-dimensional formulas.

Using subscript notation for differentiation we have

$$\begin{aligned} \nabla u &= (u_x, u_y) \\ \nabla(u, v) &= (u_x, v_y) \\ \triangle u &= u_{xx} + u_{yy} \\ &= \nabla^2 u \qquad \text{when } u \text{ is a scalar function} \end{aligned}$$

These operators can be applied either to a scalar function u or a vector function u. When u is a scalar, ∇u is called the *gradient* of u, sometimes written grad u. When u is a vector, ∇u is called the *divergence* of u, sometimes written div u. Whether u is a vector or scalar, $\triangle u$ is called the Laplacian of u.

When u is a vector, we do not have $\triangle$u= $\nabla^2 u$. Indeed if $u = (p, q)$ we have $\triangle u = (p_{xx} + p_{yy}, q_{xx} + q_{yy})$ while $\nabla^2 u = (p_{xx}, q_{yy})$. In three dimensions we have $\triangle u = \nabla^2 u - \nabla \times (\nabla \times u)$, but this equation, the cause of many difficulties in vector calculus, is not used in the theory of minimal surfaces, so we will not even trouble to explain the meaning of $\nabla \times u$.

We assume that the reader knows what "continuous" and "differentiable" mean. The unit circle S^1 is $\{\langle x, y\rangle : x^2 + y^2 = 1\}$. Its interior is the unit disk D. The closed unit disk $\bar{D}$ $\{\langle x, y\rangle : x^2 + y^2 \le 1\}$. A function is said to be (of class) C^k on a set Ω if it has first, second, and up to k-th derivatives on Ω, and all those derivatives are continuous. Thus C^1 means that the function is differentiable and the derivative is continuous. In case the set is not open, being differentiable at a boundary point does not imply that the function is even defined off the set. For example, $|x|$ is differentiable on $[0, 1]$, but not on $[-1, 1]$. As another example: $\sqrt{x}$ is C^1 in $(0, 1)$ but not in $[0, 1]$.

A Jordan curve is a continuous, one-to-one image of S^1. A famous theorem called the Jordan curve theorem says that a Jordan curve lying in a plane divides its complement into two open sets, one bounded (the "interior") and one unbounded (the "exterior"). A *plane domain* is the interior of a C^1 Jordan curve.

One form of Green's theorem says that if F is any vector function defined in a plane domain Ω bounded by a C^1 Jordan curve C, then

$$\int_C F \cdot n\, ds = \int_\Omega \nabla F\, dA \tag{1.1}$$

where the integral on the right is a two-dimensional integral. Sometimes one uses two integral signs to indicate an area integral, probably because when one wishes to evaluate such an integral, it is reduced to two one-dimensional integrations; but we shall usually just use one integral sign, since the dA and the subscript on the integral already contain the dimension information.

Applying this version of Green's theorem to the vector function ∇u, when u is a scalar function, we find

$$\int_C \nabla u \cdot n\, ds = \int_\Omega \triangle u\, dA$$

The integrand on the left, $\nabla u \cdot n$, is the *outward normal derivative* of u, often written u_ν. With that notation, Green's theorem takes the form

$$\int_C u_\nu\, ds = \int_\Omega \triangle u\, dA \tag{1.2}$$

In the special case that the domain Ω is the unit disk D, we obtain

$$\int_{S^1} u_r\, ds = \int_D \triangle u\, dA \tag{1.3}$$

There is a nice formula for the Laplacian of a product (derived using the chain rule):

$$\triangle(uv) = v\triangle u + 2\nabla u \nabla v + u\triangle v$$

Applying (1.2) we have

$$\begin{aligned} \int_C (uv)_\nu\, ds &= \int_\Omega \triangle(uv) \\ \int_C uv_\nu + vu_\nu\, ds &= \int_\Omega u\triangle v + 2\nabla u \nabla v + v\triangle u\, dA \end{aligned}$$

In fact this easily-remembered formula combines two copies of another form of Green's theorem:

$$\int_C uv_\nu\, ds = \int_\Omega u\triangle v + \nabla u \nabla v\, dA \tag{1.4}$$

Mathematicians often refer to applications of Green's theorem as "integration by parts".

Chapter 2

Harmonic Functions

Recall the definition of the Laplacian operator: $\triangle u := u_{xx} + u_{yy}$. A function is called *harmonic* if $\triangle u = 0$. This lecture is devoted to deriving some basic facts about harmonic functions, as these are indispensable tools in the study of minimal surfaces. The real and imaginary parts of a complex-analytic function are harmonic, since if $f(z) = u + iv$ then the Cauchy-Riemann equations are $u_x = v_y$ and $u_y = -v_x$. Hence $u_{xx} + v_{yy} = u_{yx} - u_{xy} = 0$.

2.1 Complex differentials

It is often convenient to use complex differentials defined as follows, where $z = x + iy$:

$$\begin{aligned} \frac{\partial}{\partial z} &= \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \\ \frac{\partial}{\partial \bar{z}} &= \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) \end{aligned}$$

Any real-analytic of x and y can be written as a function of z and $\bar{z}$ by replacing x by $\frac{1}{2}(z+\bar{z})$ and y by $(z-\bar{z})/(2i)$ in a power series for the function. This extends the function to a function of two complex variables. On the two-dimensional subspace of C^2 defined by requiring the variable $\bar{z}$ to be the complex conjugate of z, the extended function agrees with the original function.

We often write differentiation using a subscript; for example, u_z instead of $\frac{\partial u}{\partial z}$.

These complex differentials simplify many calculations. For example: The Cauchy-Riemann equations for f are just $f_{\bar{z}} = 0$. A function is analytic if it depends only on z, not on $\bar{z}$. We have

$$\frac{\partial^2 u}{\partial z \partial \bar{z}} = \frac{1}{4}\triangle u$$

so a function u is harmonic if and only if $u_{z\bar{z}} = 0$.

Here is our first application of complex differentials:

Lemma 2 *If u is harmonic, then there exists a* conjugate harmonic function v *such that $f(z) = u + iv$ is complex-analytic. The function u_z is complex analytic and $f_z = 2u_z$.*

Proof. By hypothesis $\triangle u = u_{z\bar{z}} = 0$, so u_z satisfies the Cauchy-Riemann equations and hence is complex-analytic. Integrating it with respect to z we define $f(z) = 2\int u_z\,dz$, choosing the constant of integration so that $f(z)$ agrees with $u(z)$ at some point z_0. Then f is analytic. The real part of f is u since

$$\begin{aligned} \Re \int u_z\,dz &= 2\,\Re \int \frac{1}{2}(u_x - i\,u_y)(dx + i\,dy) \\ &= \int u_x\,dx + u_y\,dy \\ &= \int du \\ &= u \end{aligned}$$

The imaginary part of f is v.

When working with complex differentials, one must remember to treat $\bar{z}$ and z as independent variables while differentiating. Only after finishing the differentiations can we return to the submanifold of C^2 where $\bar{z}$ is the complex conjugate of z. What if we want to apply complex differentials to functions that are only assumed to be C^2 rather than real-analytic? Is this legitimate? When one asks this question one is usually told that the use of complex differentials is just a formal device, abbreviating more complex expressions evaluating real differentials. I have never seen this claim justified to a logician's satisfaction.

The following relations between complex differentials and polar coordinates are often useful.

$$\begin{aligned} z &= re^{i\theta} \\ \frac{dz}{z} &= \frac{dr}{r} + i\,d\theta \\ dz &= e^{i\theta}\,dr + ire^{i\theta}\,d\theta \\ d\theta &= \frac{dz}{iz} \qquad \text{when integrating on } S^1 \\ dz &= ie^{i\theta}d\theta \qquad \text{when integrating on } S^1 \end{aligned}$$

We have $r_z = \sqrt{z\bar{z}}_z = \bar{z}/(2r) = e^{-i\theta}/2$ and θ_z is calculated as follows:

$$\begin{aligned} i\theta &= \log(z/r) \\ &= \log z - \log r \\ i\theta_z &= 1/z - r_z/r \end{aligned}$$

$$
\begin{aligned}
&= 1/z - \bar{z}/(2r^2) \\
&= 1/z - 1/(2z) \\
&= -1/(2z) \\
\theta_z &= i/(2z) \\
&= \frac{ie^{-i\theta}}{2r}
\end{aligned}
$$

Similarly $r_{\bar{z}} = e^{i\theta}/2$ and $\theta_{\bar{z}} = -i/\bar{z} = -ie^{i\theta}/(2r)$.

Converting complex derivatives to polar coordinates is done as follows:

$$
\begin{aligned}
u_z &= u_r r_z + u_\theta \theta_z \\
&= u_r \frac{\bar{z}}{2r} + u_\theta \frac{i}{2z} \\
&= \frac{1}{2} u_r e^{-i\theta} + \frac{ie^{-i\theta}}{2r} u_\theta \\
u_{\bar{z}} &= u_r \frac{z}{2r} - u_\theta \frac{i}{2\bar{z}} \\
&= \frac{1}{2} u_r e^{i\theta} - \frac{ie^{i\theta}}{2r} u_\theta
\end{aligned}
$$

We illustrate the use of these techniques by calculating $\triangle u$ in polar coordinates:

$$
\begin{aligned}
\frac{1}{4}\triangle u &= u_{z\bar{z}} \\
&= \left(u_r \frac{\bar{z}}{2r} + u_\theta \frac{i}{2z} \right)_{\bar{z}} \\
&= u_{r\bar{z}} \frac{\bar{z}}{2r} + u_r \left(\frac{\bar{z}}{2r} \right)_{\bar{z}} + u_{\theta\bar{z}} \frac{i}{2z} \\
&= u_{r\bar{z}} \frac{\bar{z}}{2r} + u_r \frac{2r - 2\bar{z} r_{\bar{z}}}{4r^2} + u_{\theta\bar{z}} \frac{i}{2z} \\
&= u_{r\bar{z}} \frac{\bar{z}}{2r} + u_r \frac{2r - 2re^{-i\theta} e^{i\theta}/2}{4r^2} + u_{\theta\bar{z}} \frac{i}{2z} \\
&= u_{r\bar{z}} \frac{\bar{z}}{2r} + u_r \frac{2r - r}{4r^2} + u_{\theta\bar{z}} \frac{i}{2z} \\
&= u_{r\bar{z}} \frac{\bar{z}}{2r} + \frac{u_r}{4r} + u_{\theta\bar{z}} \frac{i}{2z} \\
&= \left(\frac{1}{2} u_{rr} e^{i\theta} - \frac{ie^{i\theta}}{2r} u_{r\theta} \right) \frac{\bar{z}}{2r} + \frac{1}{4r} u_r + u_{\theta\bar{z}} \frac{i}{2z} \\
&= \frac{1}{4} u_{rr} - \frac{ie^{i\theta}\bar{z}}{4r^2} u_{r\theta} + \frac{1}{4r} u_r + \frac{i}{2z} \left(\frac{1}{2} u_{\theta r} e^{i\theta} - \frac{ie^{i\theta}}{2r} u_{\theta\theta} \right) \\
&= \frac{1}{4} u_{rr} + \frac{1}{4r} u_r + \frac{1}{4r^2} u_{\theta\theta}
\end{aligned}
$$

Note the miraculous cancelation of the $u_{r\theta}$ terms. Multiplying by 4 we obtain our final result,

$$\triangle u = u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta}.$$

2.2 The Poisson integral

The boundary-value problem for the Laplacian is this: given continuous boundary values f on S^1, find a continuous function $u : \bar{D} \mapsto \mathbf{R}$ such that u restricted to S^1 is f and u is harmonic in the interior.

The solution of this boundary-value problem is given by Poisson's integral,

$$u(z) = \Psi[f] = \int_0^{2\pi} f(e^{i\varphi})P(z,\varphi)\,d\varphi$$

where $P(z,\varphi)$ is the *Poisson kernel*, given as follows, with $\zeta = e^{i\varphi}$ and $z = re^{i\theta}$:

$$\begin{aligned} P(z,\varphi) &= \frac{1}{2\pi}\Re\frac{\zeta+z}{\zeta-z} \\ &= \frac{1}{2\pi}\Re\frac{1+re^{i(\theta-\varphi)}}{1-re^{i(\theta-\varphi)}} \end{aligned}$$

Multiplying the integrand's numerator and denominator by the complex conjugate of the numerator and then simplifying, we obtain the following two forms of the Poisson kernel:

$$\begin{aligned} P(z,\varphi) &= \frac{1}{2\pi}\frac{1-r^2}{1-2r\cos(\theta-\varphi)+r^2} \\ &= \frac{1}{2\pi}\frac{1-r^2}{|e^{i\cos(\theta-\varphi)}-r|^2} \end{aligned}$$

While the Poisson formula is reminiscent of Cauchy's formula for analytic functions, it cannot be proved by simply resolving Cauchy's formula into real and imaginary parts. It can be proved in several ways, each of which casts light on the situation. First, we observe that $P(z,\varphi)$ (as a function of z) is harmonic in the open disk D, since it is the real part of an analytic function. Hence if we can differentiate under the integral sign in the definition of $\Psi[f]$, then u is harmonic; that will be justified (for z in the open disk D) if f is bounded, for example. Hence it only remains to show that $\Psi[f]$ does take on the boundary values f. The following classical result is due to Schwartz:

Theorem 2 (Continuity of Poisson's integral) *Let f be a continuous function on S^1, and let u be defined by the Poisson integral $\Psi[\varphi]$ in the open disk D, and let $u(z) = f(z)$ on S^1. Then u is continuous in the closed unit disk.*

Proof. It suffices to establish the continuity by radial limits, i.e.

$$\lim_{r\to 1} u(re^{i\theta}) = f(e^{i\theta})$$

since if this is known, then we have

$$|u(re^{i\varphi}) - f(e^{i\theta})| \quad \leq |u(re^{i\varphi}) - f(e^{i\varphi})| + |f(e^{i\varphi}) - f(e^{i\theta})|$$

Given $\epsilon > 0$, the first term can be made less than $\epsilon/2$ by taking r near 1, if we have radial-limit continuity, and the second term can be made less than $\epsilon/2$ by the continuity of f. We proceed now to prove the radial-limit continuity. Change the variable of integration in the Poisson integral from φ to $\varphi + \theta$; since the integrand is periodic, we can leave the limits of integration unchanged. Using the last form of the Poisson kernel derived above, we have

$$u(re^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 - r^2}{|e^{i\varphi} - r|^2} f(e^{i(\theta+\varphi)})\, d\varphi$$

From the Poisson formula with constant boundary values 1, we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 - r^2}{|e^{i\varphi} - r|^2}\, d\varphi = 1.$$

Technically we haven't yet established the validity of the Poisson formula even for constant boundary values, so this formula should be independently derived. That can be done by Cauchy's residue theorem using the first form of the Poisson kernel given above. Here are the details: The preceding integral is equal to

$$\begin{aligned}
\frac{1}{2\pi} \int_{-\pi}^{\pi} \Re \frac{\zeta + z}{\zeta - z}\, d\varphi &= \frac{1}{2\pi} \Re \int_{-\pi}^{\pi} \frac{\zeta + z}{\zeta - z}\, d\varphi \\
&= \frac{1}{2\pi} \Re \int_{S^1} \left(\frac{\zeta + z}{\zeta - z} \right) \frac{d\zeta}{i\zeta} \qquad \text{since } \zeta = e^{i\varphi} \\
&= \frac{1}{2\pi} \Im \int_{S^1} \frac{(\zeta + z)}{\zeta(\zeta - z)}\, d\zeta \\
&= \frac{1}{4\pi} \Im\, (2\pi i(1 + 1)) \qquad \text{by Cauchy's residue theorem} \\
&= 1 \qquad \text{as claimed.}
\end{aligned}$$

Hence

$$u(re^{i\theta}) - f(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 - r^2}{|e^{\imath\varphi} - r|^2} f(\theta\varphi)[f(\theta + \phi) - f(\theta)]\, d\varphi.$$

Fix an $\epsilon > 0$. Since f is continuous, there exists $\delta > 0$ such that $|f(\phi) - f(\theta)| < \epsilon$ when $|\phi - \theta| < \delta$. Now we will estimate the boundary integral in three pieces: one is the integral from $-\delta$ to δ, and the other two are from $-\pi$ to $-\delta$ and from δ to π. Thus

$$u(re^{i\theta}) - f(\theta) = I_1 + I_2 + I_3$$

where

$$I_2 = \frac{1}{2\pi} \int_{-\delta}^{\delta} \frac{1 - r^2}{|2^{\imath\varphi} - r|^2} f(\theta\varphi)[f(\theta + \phi) - f(\theta)]\, d\varphi.$$

and I_1 and I_3 differ only in the limits of integration. We have

$$\begin{aligned} I_2 &\leq \epsilon \frac{1}{2\pi} \int_{-\delta}^{\delta} \frac{1-r^2}{|2^{\imath\varphi} - r|^2}\, d\varphi \\ &\leq \epsilon \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1-r^2}{|2^{\imath\varphi} - r|^2}\, d\varphi \qquad \text{since the integrand is positive} \\ &\leq \epsilon \end{aligned}$$

Now to estimate I_1 and I_3. Let M be the maximum of $|f|$ on S^1, and observe that for $|\phi| \geq \delta$ we have

$$\begin{aligned} |e^{i\phi} - r| &= |\cos(\phi) + i\, \sin(\phi) - r| \\ &\geq \sin\delta \end{aligned}$$

since if $|\phi| \geq \pi/2$ then $\cos(\phi)$ is negative (so the expression is at least $|r$ and otherwise $\sin\phi \geq \sin\delta$.

$$\begin{aligned} I_1 &\leq \frac{1}{2\pi} 2M(1-r^2) \int_{-\pi}^{-\delta} \frac{1}{|2^{\imath\varphi} - r|^2}\, d\varphi \\ &\leq \frac{1}{2\pi} 2M(1-r^2) \int_{-\pi}^{-\delta} \int_{-\pi}^{\delta} \frac{1}{\sin^2\delta}\, d\varphi \\ &\leq \frac{2M(1-r)}{\sin^2\delta}. \end{aligned}$$

Similarly I_3 is bounded by the same quantity. Putting the three estimates together we have

$$u(re^{i\theta}) - f(\theta) \leq \epsilon + \frac{4M}{\sin^2\delta}(1-r)$$

and taking the limit as $r \to 1$ we obtain the desired result.

2.3 The maximum principle

Theorem 3 (Maximum principle) *A non-constant harmonic function cannot have an interior maximum or minimum.*

Proof. By the Poisson representation, a harmonic function's value at a point is the average of its values on any circle about that point.

The proof of the following useful theorem illustrates the typical use of the Poisson representation and the maximum principle.

Theorem 4 *Suppose a harmonic function u is bounded in a punctured disk. Then the singularity is removable, i.e. there is a function harmonic in the entire unit disk that agrees with u on the punctured disk.*

Remark. The function $\log(1/r)$ is harmonic in the punctured disk, so the boundedness hypothesis is not superfluous.

Proof. Without loss of generality we can assume that the disk in question is the unit disk D. Let P be the punctured disk $D - \{0\}$. Let f be the harmonic extension of the boundary values of u; let $r = |z|$. For each $\epsilon > 0$, define

$$\phi_\epsilon(z) = u(z) - f(z) + \epsilon \log\left(\frac{1}{r}\right).$$

Then ϕ_ϵ is harmonic in the punctured disk. As $r \to 1$ we have $\phi_\epsilon \to 0$, by Theorem 2. Because u is bounded, $\phi_\epsilon \to \infty$ as $z \to 0$. By the maximum principle, $\phi_\epsilon(z) \geq -1/m$ in the punctured disk, for each m; hence $\phi_\epsilon(z) \geq 0$. Now let $\epsilon \to 0$; we find $u(z) - f(z) \geq 0$ in the punctured disk. Now define

$$\psi_\epsilon(z) = u(z) - f(z) - \epsilon \log\left(\frac{1}{r}\right)$$

and repeat the argument with ψ_ϵ in place of ϕ_ϵ, and $+1/m$ in place of $-1/m$. We find $u(z) - f(z) \leq 0$ in the punctured disk. Combining the two results we have $u(z) = f(z)$ in the punctured disk. That completes the proof.
Remark. Another interesting theorem (which we do not prove here) about the boundary behavior of a harmonic function is this: if φ has a step discontinuity, then its harmonic extension behaves like a helicoid asymptotically near the discontinuity, i.e. it has radial limits along rays approaching the boundary point.

2.4 Poisson's formula in the upper half-plane

We can express the Poisson formula over the upper half plane H^+, instead of over the disk. That can be useful when we want to study the boundary behavior of a minimal surface; then the boundary is parametrized along the real axis instead of a circle. The only book in which I have seen this discussed is [1], p. 145, where it is done for n dimensions instead of just $n = 2$. Here we take a simpler, complex-variables approach.

Lemma 3 *Let U be a function harmonic in the open upper half plane, continuous in the closed upper half plane, and bounded at infinity. Let*

$$K(x, z) = \Im \, \frac{1}{x - z}$$

Then for $\Im(z) > 0$ we have

$$U(z) = \frac{1}{\pi} \int_{-\infty}^{\infty} U(x) K(x, z) \, dx.$$

Moreover if U is $C^{0,\mu}$ on the real axis for some $\mu > 0$, then the formula is valid for real z too.

Remark. Note that the constant in front of the integral is $1/\pi$, not $1/2\pi$. Axler's formula (*op. cit.*) is

$$K(x,z) = c_2 \frac{y}{|x-p|^2 + q^2}$$

where $z = p + iq$, and c_2 is given (p. 144) as $2/nV(B_n)$ with $n = 2$, which works out to $1/\pi$, making our answer and Axler's agree.

Proof. Let $V(z)$ be the conjugate harmonic function of U, and let $F(z) = U(z) + iV(z)$, so F is complex analytic where U is harmonic. We first establish the theorem in case $\overline{F(z)} = F(\bar{z})$. Then we have, for z in the upper half plane, and R larger than $|z|$,

$$F(z) \quad = \quad \frac{1}{2\pi i} \int_{-R}^{R} \frac{F(x)}{x-z}\, dx + \frac{1}{2\pi i} \int_0^{\pi} \frac{F(Re^{i\theta})}{Re^{i\theta} - z} Rie^{i\theta}\, d\theta$$

by Cauchy's integral formula. Since $\bar{z}$ is in the lower half plane, we have

$$0 \quad = \quad \frac{1}{2\pi i} \int_{-R}^{R} \frac{F(x)}{x-\bar{z}}\, dx + \frac{1}{2\pi i} \int_0^{\pi} \frac{F(Re^{i\theta})}{Re^{i\theta} - \bar{z}} Rie^{i\theta}\, d\theta$$

Subtracting the two equations we get

$$\begin{aligned} F(z) \quad = \quad & \frac{1}{2\pi i} \int_{-R}^{R} F(x) \left(\frac{1}{x-z} - \frac{1}{x-\bar{z}} \right) dx \\ & + \frac{1}{2\pi i} \int_0^{\pi} \frac{F(Re^{i\theta})}{Re^{i\theta} - z} Rie^{i\theta}\, d\theta - \frac{1}{2\pi i} \int_0^{\pi} \frac{F(Re^{i\theta})}{Re^{i\theta} - \bar{z}} Rie^{i\theta}\, d\theta \end{aligned}$$

We will show in a moment that the θ integrals disappear in the limit as $R \to \infty$. That leaves us with

$$F(z) \quad = \quad \frac{1}{2\pi i} \int_{-\infty}^{\infty} (U(x) + iV(x)) 2i\, \Im \frac{1}{x-z}\, dx$$

Taking real parts we have

$$U(z) \quad = \quad \frac{1}{2\pi} \int_{-\infty}^{\infty} U(x) K(x,z)\, dx$$

as claimed. Now we return to show that the θ integrals disappear as $R \to \infty$. We have

$$\begin{aligned} & \int_0^{\pi} \frac{F(Re^{i\theta})}{Re^{i\theta} - z} Rie^{i\theta}\, d\theta - \int_0^{\pi} \frac{F(Re^{i(\theta})}{Re^{i\theta} - \bar{z}} Rie^{i\theta}\, d\theta \\ = \quad & \int_0^{\pi} F(Re^{i\theta}) \left(\frac{1}{Re^{i\theta} - z} - \frac{1}{Re^{i\theta} - \bar{z}} \right) Rie^{i\theta}\, d\theta \\ = \quad & \int_0^{\pi} F(Re^{i\theta}) Rie^{i\theta} \left(\frac{z - \bar{z}}{R^2 e^{2i\theta} - 2Re^{i\theta} \Re(z) + z\bar{z}} \right) d\theta \end{aligned}$$

$$\begin{aligned}
&= O(1/R)\int_0^\pi F(Re^{i\theta})ie^{i\theta}\,d\theta \\
&= O(1/R)\int_0^\pi ie^{i\theta}\,d\theta \qquad \text{since } F \text{ is bounded} \\
&= O(1/R)
\end{aligned}$$

as claimed. That completes the proof under the assumption $\overline{F(z)} = F(\bar{z})$.

Now suppose that U is harmonic in some neighborhood of 0, as well as in the upper half plane. Then $F(z) = \sum_{n=0}^{\infty} a_n z^n$ in some neighborhood of the origin, and in case all the a_n are real, we have $F(\bar{z}) = \bar{F}(z)$, which then holds in the upper half plane by real-analytic continuation. So the theorem applies in this case. More generally we have $F(z) = \sum (a_n + ib_n)z^n$ and if we define $G(z) = \sum a_n z^n$ and $H(z) = \sum b_n z^n$ then $F = G + iH$ and the theorem holds for G and for H separately, and hence for F:

$$\begin{aligned}
F(z) &= G(z) + iH(z) \\
&= \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{G(x)\,dx}{x-z}\,dx + \frac{i}{2\pi i}\int_{-\infty}^{\infty} \frac{H(x)\,dx}{x-z}\,dx \\
&= \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{F(x)\,dx}{x-z}\,dx
\end{aligned}$$

Finally we must eliminate the assumption that U is harmonic in a neighborhood of zero. Let $U_n(z) = U(z + 1/n)$. Then each U_n is harmonic in a neighborhood of zero, so the theorem applies to it:

$$U_n(z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} U_n(x)K(x,z)\,dx.$$

Since the U_n converge to U by the continuity of U, we can pass the limit under the integral sign by the bounded convergence theorem. That completes the proof of the lemma.

Remark. This is essentially the first proof of the Poisson formula from [16], transplanted from the disk to the upper half plane. In the disk, the complex conjugate has to be replaced by the "reflection" of a point in the unit circle, which is less intuitive.

The following lemma expresses the complex derivative F_z in terms of the boundary values of a harmonic function. In particular the normal derivative F_y is thus expressed in terms of the boundary values.

Lemma 4 *Let F be harmonic in the open upper half-plane, continuous in the closed upper half plane, C^2 on the real line, and bounded at infinity. Then for $\Im(z) > 0$ we have*

$$\begin{aligned}
F_z(z) &= \frac{-1}{2\pi i}\int_{-\infty}^{\infty} \frac{F(x)}{(x-z)^2}\,dx \\
&= \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{F_x(x)}{x-z}\,dx
\end{aligned}$$

Proof. By Lemma 3 (applied separately to the real and imaginary parts of F) we have

$$\begin{aligned} F(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \Im\left(\frac{1}{x-z}\right) F(x)\,dx \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \Re\left(\frac{-i}{x-z}\right) F(x)\,dx \end{aligned}$$

Differentiating with respect to z we have (for z in the upper half plane)

$$F_z(z) = \frac{1}{2\pi}\frac{d}{dz}\int_{-\infty}^{\infty} \left(\frac{-i}{x-z}\right) F(x)\,dx$$

Since $\Im(z) > 0$, we can push the derivative through the integral sign. Then

$$\begin{aligned} F_z(z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{d}{dz}\left(\frac{-i}{x-z}\right) F(x)\,dx \\ &= \frac{-1}{2\pi i}\int_{-\infty}^{\infty} \frac{F(x)}{(x-z)^2}\,dx \end{aligned}$$

proving the first formula of the lemma. The second formula of the lemma is obtained by integration by parts, as follows:

$$\begin{aligned} F_z(z) &= \frac{-1}{2\pi i}\lim_{R\to\infty}\int_{-R}^{R} \frac{F(x)}{(x-z)^2}\,dx \\ &= \lim_{R\to\infty}\int_{-R}^{R} \frac{F_x(x)}{x-z}\,dx - \frac{1}{2\pi i}\lim_{R\to\infty}\left(\frac{F(R)}{R-z} - \frac{F(-R)}{-R-z}\right) \\ &= \frac{1}{2\pi i}\int_{-\infty}^{\infty} \frac{F_x(x)}{x-z}\,dx \end{aligned}$$

The second limit term vanishes since F is bounded at infinity, by hypothesis.

2.5 Poisson formula for the half-plane, reprise

Since the unit disk and the upper half plane are conformally equivalent, one could either use the above proof, plus a linear fractional transformation, to prove the Poisson formula for the disk, or one could go the other way, and derive the Poisson formula for the half-plane from the Poisson formula for the disk. This is a good exercise, and since the spirit of these lectures is to carry out all the details, we will give this proof too.

Linear fractional transformations are conformal maps of the form

$$z \mapsto \frac{az+b}{cz+d}.$$

They take lines or circles into lines or circles. (Another good exercise.) Hence if we want a linear fractional transformation that takes the unit disk to the upper half plane, we choose a, b, c, and d so that i goes to ∞, $-i$ goes to 0, and 1 goes to 1. Then the unit circle will go to the real line. The transformation in question is

$$w = \frac{iz-1}{-z+i}.$$

It takes the disk onto the upper half plane, not the lower half plane, since 0 goes to the interior point i of the upper half plane. One can also calculate directly to show that this map takes S^1 onto the real line:

$$\begin{aligned}
\Im \frac{iz-1}{-z+i} &= \Im \frac{(iz-1)(-\bar{z}-i)}{(-z+i)(-\bar{z}-i)} \\
&= \frac{1}{|-z+i|^2} \Im\,(-iz\bar{z}+\bar{z}+z+i) \\
&= \frac{1}{|-z+i|^2} \Im\,(-iz\bar{z}+i) \qquad \text{since } z+\bar{z} \text{ is real} \\
&= 0 \qquad \text{for } z \text{ on } S^1, \text{ since then } z\bar{z}=1.
\end{aligned}$$

The inverse of this transformation is $z = (iw+1)/(w+i)$. (To invert a linear fractional transformation, we invert the matrix of its coefficients.) Let w be in the upper half plane and let $F : \mathbf{R} \to \mathbf{R}$ be bounded. Let $z = (iw+1)/(w+i)$ be in the unit disk and let $f(z) = F(w) = f((iz-1)/(-z+i))$ and let u be the harmonic extension of f in the unit disk. Then the harmonic extension U of F into the upper half plane is given by $U(w) = u(z)$. By the Poisson formula we have, with $\zeta = e^{i\varphi}$,

$$\begin{aligned}
u(z) &= \frac{1}{2\pi} \int_0^{2\pi} u(\zeta) \Re \frac{\zeta+z}{\zeta-z}\, d\varphi \\
&= \frac{1}{2\pi} \Re \int_{S^1} u(\zeta) \frac{\zeta+z}{\zeta-z} \frac{d\zeta}{i\zeta}
\end{aligned}$$

In the integral we shall make the substitution $\zeta = (i\xi+1)/(\xi+i)$ to transform the variable ζ, which ranges over S^1, to ξ ranging over the real axis. We have

$$\begin{aligned}
d\zeta &= \frac{i(\xi+i)-(i\xi+1)}{(\xi+i)^2}\, d\xi \\
&= \frac{-2\, d\xi}{(\xi+i)^2} \\
\frac{d\zeta}{i\zeta} &= \frac{\frac{-2\, d\xi}{(\xi+i)^2}}{\frac{i(i\xi+1)}{\xi+i}} \\
&= \frac{2\, d\xi}{(\xi^2+1)}
\end{aligned}$$

Substituting $z = (iw+1)/(w+i)$ as well as making the given substitution for ζ we have

$$U(w) = u(z)$$

$$= \frac{1}{2\pi}\Re\int_{\infty}^{\infty} U(\xi)\frac{\frac{i\xi+1}{\xi+i}+\frac{iw+1}{w+i}}{\frac{i\xi+1}{\xi+i}-\frac{iw+1}{w+i}}\frac{2\,d\xi}{(\xi^2+1)}$$

Simplifying the first compound fraction we have

$$\begin{aligned}
U(w) &= \frac{1}{2\pi}\Re\int_{\infty}^{\infty} U(\xi)\frac{\xi w+1}{i(\xi-w)}\frac{2\,d\xi}{(\xi^2+1)} \\
&= \frac{1}{2\pi}\Re\int_{\infty}^{\infty} U(\xi)\frac{2}{i}\left(\frac{1}{\xi-w}-\frac{\xi}{\xi^2+1}\right)d\xi \\
&= \frac{1}{2\pi}\Re\int_{\infty}^{\infty} U(\xi)\frac{2}{i}\frac{1}{\xi-w}\,d\xi - \frac{1}{2\pi}\Re\int_{\infty}^{\infty} U(\xi)\frac{\xi}{\xi^2+1}\,d\xi
\end{aligned}$$

The second integral is zero, since it is equal to the limit as $R\to\infty$ of the contour integral of the analytic function $(U+iV)w/(w^2+1)$ around the boundary of the upper half-disk of radius R (where V is the harmonic conjugate of U). Hence

$$\begin{aligned}
U(w) &= \frac{1}{2\pi}\Re\int_{\infty}^{\infty} U(\xi)\frac{2}{i}\frac{1}{\xi-w}\,d\xi \\
&= \frac{1}{2\pi}\Im\int_{\infty}^{\infty} U(\xi)\frac{2}{\xi-w}\,d\xi \\
&= \frac{1}{\pi}\int_{\infty}^{\infty} U(\xi)\Im\frac{1}{\xi-w}\,d\xi
\end{aligned}$$

That is the correct form for the Poisson kernel on the upper half-plane, as derived in the lemma above, and now re-derived from the Poisson kernel on the disk by the use of a linear fractional transformation.

2.6 Harmonic functions and Dirichlet's integral

Another consequence of the Poisson representation is

Theorem 5 (Harnack's theorem) *Suppose the sequence of harmonic functions u_n converges uniformly in the closed unit disk D to a limit u. Then u is harmonic and the derivatives of u_n converge uniformly in compact subdomains to the derivatives of u.*

Proof. First, for simplicity, assume that the limit function u is continuous.

$$u_n(z) = \frac{1}{2\pi}\int_0^{2\pi} u_n(e^{i\theta})\Re\frac{\zeta+z}{\zeta-z}\,d\varphi$$

where $\zeta = e^{i\varphi}$ and $z = re^{i\theta}$. Passing the limit through the integral sign, which is justified since u is continuous. we obtain

$$u(z) = \frac{1}{2\pi}\int_0^{2\pi} u(e^{i\theta})\Re\frac{\zeta+z}{\zeta-z}\,d\varphi$$

so u is harmonic. Differentiating the first equation, and using the fact that for analytic F, we have $F_z = (d/dz)\Re F$, we obtain

$$\begin{aligned}\frac{d}{dz}u_n(z) &= \frac{1}{2\pi}\int_0^{2\pi} u_n(e^{i\theta})\frac{d}{dz}\frac{\zeta+z}{\zeta-z}\,d\varphi \\ &= \frac{1}{2\pi}\int_0^{2\pi} u_n(e^{i\theta})\frac{2\zeta}{(\zeta-z)^2}\,d\varphi\end{aligned}$$

Since u_n converges uniformly on the boundary to u, we can take the limit under the integral sign, obtaining

$$\begin{aligned}\lim \frac{d}{dz}u_n(z) &= \frac{1}{2\pi}\int_0^{2\pi} u(e^{i\theta})\frac{2\zeta}{(\zeta-z)^2}\,d\varphi \\ &= \frac{du}{dz}\end{aligned}$$

Now, if the limit function u is not assumed to be continuous on the boundary, we restrict the u_n and u to the disk of radius R. Then u is continuous on the boundary (the circle of radius R) and hence u is harmonic and the derivatives of u_n converge to those of u on the disk of radius R. That completes the proof.

Since the derivatives of a harmonic function are themselves harmonic, Harnack's theorem applies as well to the higher derivatives, not just the first derivatives.

The solution $u = \Psi[\varphi]$ is known as the *harmonic extension* of the boundary values φ. We now consider the map Ψ as a map from one function space to another. We have just observed that Ψ maps $C^0(S^1)$ into $C^0(\bar{D})$, and it can be shown that Ψ is continuous. In general Ψ does *not* map $C^n(S^1)$ into $C^n(\bar{D})$. We "lose one derivative". If we use the Lipschitz-condition spaces $C^{n,\alpha}$, we find better behavior: Ψ does map $C^{n,\alpha}$ continuously into $C^{n,\alpha}$. This is the theorem of Korn and Privalov. You can find a proof in [11], page 17. But if we stick to C^n spaces, we need one higher derivative on the boundary: if we want to know that $\nabla\phi$ is small, for example, we would need to estimate $f_{\theta\theta}$.

Definition 2 *Let $f : D \to R$. Dirichlet's integral is given by*

$$\begin{aligned}E[f] &= \frac{1}{2}\int_D |\nabla f|^2\,dx\,dy \\ &= \frac{1}{2}\int_D f_x^2 + f_y^2\,dx\,dy\end{aligned}$$

Lemma 5 (Semicontinuity of E) *Suppose the sequence of harmonic functions u_n converges (uniformly in compact subdomains of the unit disk D) to a limit u. Then*

$$E[u] \le \liminf E[u_n].$$

Proof. By Harnack's theorem, ∇u_n converges on compact subsets to ∇u. If we integrate $|\nabla u_n|^2$ over the disk of radius $R < 1$, we can take the limit under the integral sign. Let $E_R[u]$ denote (half of) this integral. Then

$$\lim_{n\to\infty} E_R[u_n] = E_R[u].$$

Fix $\epsilon > 0$. We must show that for n sufficiently large we have $E[u_n] \geq E[u] - \epsilon$. Pick R so large that $E_R[u] > E[u] - \epsilon/2$. Pick k so large that for $n \geq k$ we have $|E_R[u_n] - E_R[u]| < \epsilon/2$. Then

$$E[u_n] - E[u] = E[u_n] - E_R[u_n] + E_R[u_n] - E_R[u] + E_R[u] - E[u]$$

Since $E[u_n] - E_R[u_n] \geq 0$ we have

$$E[u_n] - E[u] \geq E_R[u_n] - E_R[u] + E_R[u] - E[u]$$

The right hand side has absolute value bounded by ϵ:

$$\begin{aligned} |E_R[u_n] - E_R[u] + E_R[u] - E[u]| &\leq |E_R[u_n] - E_R[u]| + |E_R[u] - E[u]| \\ &\leq \epsilon/2 + \epsilon/2 = \epsilon \end{aligned}$$

Hence

$$E[u_n] - E[u] \geq -\epsilon$$

as required. That completes the proof.

Theorem 6 (Harmonic functions and Dirichlet's integral) *Let f be in the Sobolev space $W^{1,2}$ of the closed unit disk, continuous except possibly at finitely many boundary points, and bounded. Suppose that $E[f]$ is a minimum among functions from $\bar{D}$ to R with the same boundary values. Then f is harmonic.*

Proof. Let u be the harmonic extension of the boundary values of f. Let $\phi = f - u$ so $f = u + \phi$ with u harmonic. Suppose for the moment that f is C^2 in the closed disk, so that Green's theorem is applicable where we need it below. Then calculate:

$$\begin{aligned} E[f] - E[u] &= \frac{1}{2}\int_D \nabla(u+\phi)\nabla(u+\phi) - |\nabla u|^2\,dx\,dy \\ &= \int_D \nabla u \nabla\phi + \frac{1}{2}|\nabla u|^2\,dx\,dy \\ &= \int_0^{2\pi} \phi u_r\,d\theta + \int_D \phi\triangle u + \frac{1}{2}|\nabla\phi|^2\,dx\,dy \quad \text{by Green's theorem} \\ &= \frac{1}{2}\int_D |\nabla\phi|^2\,dx\,dy \quad \text{since } \phi = 0 \text{ on the boundary and } \triangle u = 0 \\ &\geq 0 \end{aligned}$$

On the other hand, since we have assumed $E[f]$ is a minimum, we have $E[f] - E[u] \leq 0$. Hence $\int_D |\nabla\phi|^2\,dx\,dy = 0$. Since the integrand is non-negative, we have $\nabla\phi = 0$ almost everywhere, so ϕ is constant. But since ϕ is zero on the boundary, and continuous at all but finitely many boundary points, ϕ is identically zero. That completes the proof in case Green's theorem is applicable, in particular in case f is C^2.

To prove the theorem for more general f, let u_n be a sequence of harmonic polynomials converging to u, defined by the truncated Fourier series of u, and let $\phi_n = f - u_n$. Repeat the above calculation, using u_n instead of u. Then Green's theorem is applicable to $\int_D \nabla u \nabla \phi$, since u is C^2. (It doesn't matter that ϕ is not C^2 since we don't need derivatives of ϕ in this application of Green's theorem.) We conclude that $E[f] - E[u_n] \geq 0$. That is, $E[u_n] \leq E[f]$ Taking the limit as $n \to \infty$, we find $\liminf E[u_n] \leq E[f]$. Applying the semicontinuity of E (proved in the previous lemma) we have $E[u] \leq E[f]$ as desired. That completes the proof.

Chapter 3

Harmonic Surfaces

3.1 Isothermal coordinates

If $u_x \cdot u_y = 0$ and $|u_x^2| = |u_y^2|$ then we say u is given in *isothermal coordinates.* These are also called "conformal coordinates". There is a general theorem that any C^1 regular surface has an isothermal parametrization. But, we need the existence of isothermal parameters for a minimal surface, which is allowed to have isolated non-regular points. Luckily, there is an explicit construction of (local) isothermal parameters for minimal surfaces. This theorem is attributed to Riemann and Beltrami in Rado's book.

Theorem 7 (Isothermal parameters) *A sufficiently small portion of a minimal surface admits an isothermal parametrization.*

Proof. In a sufficiently small neighborhood, we can orient the X, Y, Z axes so that the surface can be written in non-parametric form, $Z = f(x, y)$. We write X for the vector (x, y, z). Then f satisfies the non-parametric minimal surface equation. We claim that $N \times dX$ is a complete differential. That means that for some (vector) function ω, we have $d\omega = N \times dX$. From calculus we know that $Pdx + Qdy$ is a complete differential if and only if $P_y = Q_x$. Here we have $dX = (dx, dy, dZ)$ and $dZ = f_x dx + f_y dy$, and with $W^2 = 1 + f_x^2 + f_y^2$ we have

$$
\begin{aligned}
N &= \frac{1}{W}\begin{bmatrix} -f_x \\ -f_y \\ 1 \end{bmatrix} \\
N \times dX &= \frac{1}{W}\begin{bmatrix} -f_y\, dZ - dy \\ dx + f_x\, dZ \\ -f_x\, dy + f_y\, dx \end{bmatrix} \\
&= \frac{1}{W}\begin{bmatrix} -f_y(f_x\, dx + f_y\, dy) - dy \\ dx + f_x(f_x\, dx + f_y\, dy) \\ -f_x\, dx + f_y\, dy \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{W}\begin{bmatrix} -f_y f_x\,dx - (1+f_y^2)dy \\ (1+f_x^2)\,dx + f_x f_y dy \\ -f_x\,dx + f_y\,dy \end{bmatrix} \\
&= dx\begin{bmatrix} f_x f_y/W \\ (1+f_x^2)/W \\ -f_x/W \end{bmatrix} + dy\begin{bmatrix} -(1+f_y^2)/W \\ f_x f_y \\ f_y \end{bmatrix}
\end{aligned}
$$

Now we apply the cross-wise differentiation test. Let

$$T = (1+f_y^2)f_{xx} - 2f_x f_y f_{xy} + (1+f_x^2)f_y y$$

so that $T = 0$ if and only if the surface is minimal. After a few lines of computation we find

$$
\begin{aligned}
\begin{bmatrix} ((f_x f_y)/W)_y - (-(1+f_y^2)/W)_x \\ (1+f_x^2)/W)_y - (f_x f_y/W)_x \\ (f_y/W)_y - (-f_x/W)_x \end{bmatrix} &= \frac{T}{W}\begin{bmatrix} -f_x \\ -f_y \\ 1 \end{bmatrix} \\
&= \frac{T}{W}N
\end{aligned}
$$

Hence $N \times dX$ is a complete differential if and only if T is minimal. Here we only need one direction: since we have assumed u is minimal, $N \times dX$ is a complete differential. Therefore for some function ω we have $d\omega = N \times DX$. Introduce $\alpha = x$, $\beta = \omega_1(x, y)$, the first component of ω. The map $(x, y) \mapsto (\alpha, \beta)$ is a local diffeomorphism, so we can consider $u = (x, y, f(x, y))$ as a function of α and β. We claim that α and β are local isothermal parameters. We have to show that $u_\alpha \cdot u_\beta = 0$ and $u_\alpha^2 = u_\beta^2$. We have

$$
\begin{aligned}
u_\alpha &= u_x x_\alpha + u_y y_\alpha \\
u_\beta &= u_x x_\beta + u_y y_\beta
\end{aligned}
$$

3.2 Uniformization

The main theorem in this section is the existence of global isothermal coordinates for minimal surfaces. The proof will only be sketched. For details see [5], chapters 2 and 5.

Theorem 8 *If u is a minimal surface then there exists a reparametrization $\tilde{u}$ of u which is in isothermal coordinates.*

Proof Sketch. We have already shown the existence of *local* isothermal coordinates. We can therefore find a triangulation of the surface u and local isothermal coordinates defined in each triangle of the triangulation. The problem is to "uniformize" these coordinates. The method involves the existence of a Green's function, or "dipole potential", on the surface. One uses the local coordinates to define the meaning of "harmonic"; namely, a function defined on the surface is harmonic if it is harmonic when expressed as a function of the local isothermal coordinates. Now we can define the concept of a Green's function: it is a function $G(z, \zeta)$ which for fixed ζ on the boundary is harmonic on the surface as a function of z and is zero on the boundary. We will not explain the proof that G exists. Once it is known to exist, then fix a point ζ and let $g(z) = G(z, \zeta)$, and g^* the harmonic conjugate of g, and define $F(z) = g(z) + ig^*(z)$. Then F maps the surface conformally onto the upper half plane. Since the upper half plane is conformally equivalent to the unit disk, the surface can be mapped conformally onto the unit disk, which is what we were trying to prove.

3.3 Minimal surfaces as harmonic conformal surfaces

Since the 1930's, it has been customary to study minimal surfaces as harmonic isothermal surfaces. Here is the basic theorem that justifies this practice.

Theorem 9 *A surface in isothermal parameters is minimal if and only if it is harmonic.*

Proof. Suppose u is given in isothermal parameters. Then $E = G = W$ and $F = 0$, and the mean curvature H, which we proved is given by

$$H = \frac{\mathsf{L}G + \mathsf{N}E - 2\mathsf{M}F}{2W^2}$$

reduces to

$$H = \frac{\mathsf{L} + \mathsf{N}}{2W}.$$

Recall that $\mathsf{L} = u_{xx} \cdot N$ and $\mathsf{N} = u_{yy} \cdot N$, where we have used N in a slightly different font than N to minimize confusion, we have

$$\triangle u \cdot N = 2WH.$$

The conformality conditions are $u_x^2 = u_y^2$ and $u_x u_y = 0$. Differentiating with respect to x and y we find

$$\begin{aligned} u_x u_{xx} &= u_y u_{yx} \\ u_y u_{yy} &= u_x u_{xy} \\ u_{xx} u_y + u_x u_{xy} &= 0 \\ u_{yy} u_x + u_y u_{xy} &= 0 \end{aligned}$$

Therefore

$$\begin{aligned} \triangle u u_x &= 0 \\ \triangle u u_y &= 0. \end{aligned}$$

This means that $\triangle u$ is a normal vector. Since N is a unit vector, and we proved that $\triangle u \cdot N = HW$, it follows that

$$\triangle u = HWN.$$

Hence H is identically zero if and only if $\triangle u$ is identically zero, which is what we had to prove.

Theorem 10 *A harmonic surface is minimal if and only if it is in isothermal parameters.*

Proof. Suppose $\triangle u = 0$. If u is also in isothermal parameters then u is minimal by the previous theorem. Suppose then that u is minimal; we must show u is in isothermal parameters. The proof of this will be postponed until the section on Dirichlet's integral below.

Corollary 1 *A harmonic surface u is minimal if and only if $u_z^2 = 0$.*

Proof. Suppose u is harmonic. Then u_z is complex analytic. Consider $u_z^2 = u_z \cdot u_z$. The real part is $u_x^2 - u_y^2$ and the imaginary part is $2u_x \cdot u_y$. Thus the parameters are isothermal if and only if $u_z^2 = 0$.

3.4 The Weierstrass representation

Start with a minimal surface u in harmonic isothermal parameters. Then u_z is complex analytic. That is, each of its three components is complex analytic. The minimal surface equation says that $u_z^2 = 0$. That means that the three components of u_z are complex analytic functions ϕ_i such that $\phi_1^2 + \phi_2^2 + \phi_3^2 = 0$, and conversely, any such triple of analytic functions can be integrated with respect to z to yield a minimal surface. This establishes an important and fundamental connection between minimal surfaces and analytic function theory.

Enneper and Weierstrass independently observed that such triples of functions can be written in terms of *two* analytic functions. Given such a triple, in which none of the ϕ_i is identically zero, the equation $\phi_1^2 + \phi_2^2 + \phi_3^2 = 0$ implies

$$(\phi_1 - i\phi_2)(\phi_1 + i\phi_2) = -\phi_3^2$$

which implies that $\phi_1 - i\phi_2$ is not identically zero. Define

$$\begin{aligned} f(z) &= \phi_1 - i\phi_2) \\ g(z) &= \frac{\phi_3}{f(z)} \end{aligned}$$

Then neither f nor g is identically zero, and they are both analytic except possibly at zeroes of ϕ_3. It follows that

$$\begin{aligned} \phi_1 &= \frac{1}{2}(f - fg^2) \\ \phi_2 &= \frac{i}{2}(f + fg^2) \\ \phi_3 &= fg \end{aligned}$$

Conversely, if f and g are given analytic functions, then ϕ_i as defined by these equations will satisfy $\phi_1^2 + \phi_2^2 + \phi_3^2 = 0$. We have proved:

Theorem 11 (Enneper-Weierstrass Representation) *Let u be a minimal surface. Define $f(z) = {}^1u_z - i\,{}^2u_z$, and $g(z) = {}^3u_z/f(z)$. Then we have*

$$u(z) = \Re \left[\begin{array}{l} \frac{1}{2}\int f - fg^2\,dz \\ \frac{i}{2}\int f + fg^2\,dz \\ int fg\,dz \end{array} \right]$$

To put the matter equivalently, we have

$$u_z = \left[\begin{array}{l} \frac{1}{2}(f - fg^2) \\ \frac{i}{2}(f + fg^2 \\ fg \end{array} \right].$$

This is a wonderful theorem, because it enables us to produce an example of a minimal surface from any pair of analytic functions f and g, and moreover to draw pictures of them whenever we can actually compute the integrals involved.

It is also a wonderful theorem, because it enables us to study complicated questions about minimal surfaces by writing the (unspecified) minimal surface in Weierstrass representation and reasoning about f and g.

3.5 Branch points

Definition 3 *The minimal surface u has a* branch point *at z if $u_x = u_y = 0$ at z.*

That is, the branch points are the points of non-regularity of u. In case u is harmonic, we can equally well describe the branch points as the places where u_z vanishes.

What do branch points imply about f and g in the Weierstrass representation? First note that f is always analytic, but g can be meromorphic. Since fg^2 is also analytic, if g does have poles, they are matched by zeroes of f of at least half the order of the zero of g. For the third component of u_z to be zero, both f and g must vanish, and for the first component also to be zero, fg^2 must vanish too. Therefore,*the branch points of u are the simultaneous zeroes of f and fg^2.*

If the surface u has a branch point at $z = a$, we will have $f(z) = c(z-a)^m + O((z-a)^{m+1})$ for some m. This number m is called the *order* of the branch point (assuming $c \neq 0$). If the branch point occurs on the boundary, m will have to be even for the boundary to be taken on monotonically.

We can simply put $f(z) = z^m$ and $g(z) = z$ into the Weierstrass representation to produce examples of minimal surfaces with branch points.

If $g(z-a) = cz^k + O(z^{k+1}$ (with $c \neq 0$), then k is called the *index* of the branch point. We can also use the Weierstrass representation to draw minimal surfaces with any desired index.

Branch points have been important in almost all work on minimal surfaces since the solution of Plateau's problem seventy years ago. In particular, my work in 2000-2001 depends on detailed analysis of one-parameter families of surfaces, one member of which has a branch point.

3.6 The Dirichlet Integral

Dirichlet's integral can be considered for a surface as well as for a scalar function.

Definition 4 *Dirichlet's integral is given by*

$$\begin{aligned} E[u] &= \frac{1}{2}\int_D |\nabla u|^2 \, dx\, dy \\ &= \frac{1}{2}\int_D u_x^2 + u_y^2 \, dx\, dy \end{aligned}$$

The letter E stands for "energy". We do not use D for "Dirichlet" because D is needed for Frechet derivatives.

It is sometimes useful to express $E[u]$ as an integral over S^1:

Lemma 6

$$E[u] = \frac{1}{2}\int_{S^1} u u_r d\theta.$$

Proof.

$$= \frac{1}{2}\int_{S^1} u u_r d\theta$$

as is seen using Green's theorem:

$$\begin{aligned} E[u] &= \frac{1}{2}\int_D \nabla u \cdot \nabla u \, dx\, dy \\ &= \frac{1}{2}\int_D u\triangle u + \nabla u \cdot \nabla u \, dx\, dy \qquad \text{since } \triangle u = 0 \\ &= \frac{1}{4}\int_D \triangle(u^2)\, dx\, dy \\ &= \frac{1}{4}\int_{S^1} (u^2)_r d\theta \\ &= \frac{1}{2}\int_{S^1} u u_r d\theta \end{aligned}$$

The first thing to calculate is the Frechet derivative of E, that is, the first variation. There are several interesting spaces in which we might try to calculate this derivative. We first consider the effect of varying the parametrization. The following theorem shows that, among all parametrizations of the same surface, those parametrizations that minimize E are in isothermal parameters. This theorem justifies restricting attention to harmonic surfaces, since if we find a harmonic surface minimizing E, then it must be in isothermal parameters, and hence minimal; and no other (possibly not harmonic) parametrization can further decrease E.

We set

$$\begin{aligned}\tilde{x} &= x + t\lambda(x,y)\\ \tilde{y} &= y + t\mu(x,y)\\ \tilde{u}(x,y) &= u(\tilde{x},\tilde{y})\end{aligned}$$

where λ and μ are the real and imaginary parts of $k = \lambda + i\mu$. The following formula is valid without restricting λ and μ so that (λ,μ) is tangent to the parameter domain at the boundary.

Theorem 12 (First variation of Dirichlet's integral) . *The Frechet derivative of E is given by*

$$DE[u](k) \quad = \quad \int_D (u_x^2 - u_y^2)(\lambda_x - \mu_y) + 2u_x u_y(\lambda_y + \mu_x)\,dx\,dy$$

Proof. We calculate

$$\begin{aligned}E[\tilde{u}] &= \int_D (\tilde{u}_x^2 + \tilde{u}_y^2)\,dx\,dy\\ &= \int_D (u_x\tilde{x}_x + u_y\tilde{y}_x)^2 + (u_x\tilde{x}_y + u_y\tilde{y}_y)^2\,dx\,dy\\ &= \int_D [u_x(1+t\lambda_x) + u_y t\mu_x]^2 + [u_x t\lambda_y + u_y(1+t\mu_y)]^2\,dx\,dy\end{aligned}$$

Differentiating with respect to t and then setting $t = 0$ we find the formula given in the theorem.

Now we restrict attention to $C^{k,\alpha}$ surfaces u with the same boundary Γ. Then the functions λ and μ must be such that (λ,μ) is tangential to D on ∂D. In fact, as the following proof shows, it is enough if $DE[u] = 0$ when λ and μ actually vanish on ∂D.

Corollary 2 *$DE[u] = 0$ if and only if u is in isothermal parameters.*

Proof. (from [5], p. 112) Suppose $DE[u] = 0$. Let g be any C^n function from $\bar{D}$ to R, with ∇g vanishing on S^1. and let ϕ be a solution of Poisson's equation $\triangle\phi = g$ in $\bar{D}$. Define $\lambda = \phi_y$ and $\mu = \phi_x$. Then λ and μ solve the differential equations

$$\begin{aligned}\lambda_x - \mu_y &= 0\\ \lambda_y + \mu_x &= g(x,y)\end{aligned}$$

and vanish on the boundary. It follows from the theorem that

$$\int_D (u_x^2 - u_y^2)g(x,y)\,dx\,dy = 0.$$

Since g was arbitrary, it follows from the fundamental lemma of the calculus of variations that $u_x^2 = u_y^2$. Similarly, we can solve the differential equations

$$\begin{aligned} \lambda_x - \mu_y &= f(x,y) \\ \lambda_y + \mu_x &= 0 \end{aligned}$$

by taking $\triangle\phi = f$ and $\lambda = \phi_x$, $\mu = -\phi_y$. Then by the theorem we have

$$\int u_x u_y f(x,y)\, dx\, dy =$$

for all f. Hence by the fundamental lemma of the calculus of variations we have $u_x u_y = 0$. Thus u is in isothermal parameters.

Chapter 4

Dirichlet's Integral and Plateau's Problem

4.1 Dirichlet's integral and area

Dirichlet's integral is much nicer to work with than area, since it doesn't have the ugly square root:

$$\begin{aligned} E[u] &= \frac{1}{2}\int_D u_x^2 + u_y^2 \, dx\, dy \\ A[u] &= \int_D \sqrt{u_x^2 u_y^2 - (u_x u_y)^2} \, dx\, dy \end{aligned}$$

On the other hand, area is invariant under reparametrizations of the surface, while Dirichlet's integral is not.

We have the inequality

$$A[u] \le E[u]$$

which follows from the algebraic inequality $\sqrt{a^2b^2 - c^2} \le (a^2 + b^2)/2$. Equality holds in this algebraic inequality if and only if $c = 0$. Similarly, $A[u] = E[u]$ if and only if the surface u is in isothermal parameters. Indeed, we have

$$E[u] - A[u] = \int_D \frac{1}{2}(u_x^2 + u_y^2 - \sqrt{u_x^2 u_y^2 - (u_x u_y)^2} \, dx\, dy$$

Since the integrand is continuous and nonnegative, if the integral is zero, then the integrand must be identically zero.

For the following theorem, we need Lichtenstein's theorem, that any C^2 surface has an isothermal parametrization. This is the only place in the subject where we really need this theorem (which we have not proved).

Theorem 13 *Suppose u minimizes $E[u]$ among harmonic surfaces bounded by Γ. Then u minimizes $A[u]$ also, not only among harmonic surfaces but among surfaces C^2 in D and spanning Γ.*

Proof. First, observe that if u minimizes Dirichlet's integral among harmonic surfaces, it also minimizes Dirichlet's integral among C^2 surfaces, since the harmonic extension of the boundary values of u has Dirichlet integral at most $E[u]$.

Since u minimizes $E[u]$, the first variation $DE[u]$ is zero and hence u is in isothermal coordinates. Hence $A[u] = E[u]$. Suppose u does not minimize area. Let v be another surface bounded by Γ with $A[v] < A[u] = E[u]$. Let w be an isothermal parametrizaton of v, so $E[v] = A[v] < A[u] = E[u]$; but this contradicts the assumption that $E[u]$ is a minimum.

Remark: In [5], p. 116, a proof is given which avoids Lichtenstein's theorem, by using a class of piecewise continuous surfaces that includes polyhedra, proving that polyhedra have isothermal parametrizations, and then letting v be the limit of polyhedra, and using the lower semicontinuity of E. This is also not quite simple.

4.2 Plateau's Problem

A surface u defined and continuous in the closed unit disk $\bar{D}$ is said to span a Jordan curve Γ, or to be bounded by Γ, if u restricted to ∂D is a reparametrization of Γ. That is, for some α mapping the unit circle monotonically to itself we have $u(z) = \Gamma(\alpha(z))$ for z on ∂D.

Plateau's problem is this: *Given a Jordan curve* Γ*, find a minimal surface spanning* Γ*, preferably an absolute minimum of area among surfaces spanning* Γ*, and preferably without branch points.*

The basic idea of the solution to Plateau's problem is to find a surface minimizing Dirichlet's integral in the class S of harmonic surfaces spanning a given Jordan curve Γ. There are, however, many details. We give a sketch of the proof.

The plan is to let d be the infimum of values $E[u]$ for u in S, and then let u_n be a sequence of surfaces in S with $E[u_n]$ decreasing monotonically to d. If we can arrange that S is a compact space, we can then pass to a convergent subsequence, converging to a surface u. If we can show that E is continuous, or even lower semicontinuous, we can conclude that $E[u] = d$, so u is an absolute minimum of Dirichlet's integral. Then in particular it is a critical point of E, so it is in isothermal parameters. Being harmonic and isothermal, it is minimal.

With respect to this plan we note the following difficulties:

(1) It is not obvious that there is even one harmonic surface spanning Γ whose Dirichlet integral is finite.

(2) It is not obvious why the space S should be compact.

(3) The condition of spanning Γ is not closed under uniform convergence. The limit surface u might be only weakly monotonic on S^1, and might even have "arcs of constancy".

(4) Regularity will not be preserved in the limit. Even if all the u_n are regular, the limit u_n might have branch points.

(5) The solution is certainly not unique as the problem is posed, and convergence will not work right, because of the existence of the conformal group: there is a three-parameter group of conformal transformations of D to D, and one can always repametrize a harmonic surface by a member of this group without changing $E[u]$.

These problems are solved as follows:

(1) We restrict to rectifiable curves, and show that in that case, there is a harmonic surface with finite Dirichlet integral spanning Γ. Afterwards, we approximate any Jordan curve by a convergent sequence of rectifiable curves, solve Plateau's problem for each of these, and find a convergent subsequence.

(2) The key to compactness is the Courant-Lebesgue lemma, which says that the boundary values of functions in S with Dirichlet integral bounded by M are equicontinuous.

(3) The Courant-Lebesgue lemma solves this too.

(4) This was not solved until the seventies, and for boundary branch points, is still open.

(5) This is easily fixed by restricting S to surfaces satisfying a "three-point condition", in which three fixed points on S^1 are required to correspond to three fixed points on Γ.

Chapter 5

The Second Variation of Area

5.1 The second variation of area defined

Let u be any surface defined on some plane domain Ω. We assume that u is C^3 in the interior and at all but a finite number of "exceptional" boundary points, and that the unit normal N extends in a C^1 fashion to the boundary, even at the exceptional points. [1] We do not assume that u is harmonic. The unit normal N is defined in the interior and at all but a finite number of boundary points. In [10] and [12], it is assumed that u is regular when calculating the second variation of area. (Here "regular" means that $g = \det g_{ij}$ is never zero, i.e. $u_x^2 u_y^2 - u_x u_y \neq 0$.) We want to allow for the case when u is a branched minimal surface, or possibly even a harmonic surface that is not regular. What we assume instead of regularity is that the coefficients of the first, second, and third fundamental forms of u are bounded in Ω. That will follow, for example, if u_x, u_y, N_x, and N_y are bounded, as they certainly are for a branched minimal surface.

The second variation of area is a bilinear functional $D^2A[u]$ operating on two functions ϕ and ψ in the same "tangent space" that we used when calculating the first variation of area. Given ϕ and ψ, we construct a two-parameter family

$$\tilde{u}(s,t) = u + t\phi + s\psi$$

and consider the area $A[\tilde{u}]$ as a function of t and s. A calculation, similar to the one we shall make below, shows that the second derivatives of $A[\tilde{u}]$ with respect to t and s are *intrinsic*, in the sense that they only depend on the tangent vectors $\phi = u_t \cdot N$ and $\psi = u_s \cdot N$ (the derivatives are evaluated at $t = s = 0$). Because of the intrinsic nature of these second derivatives, we can

[1] This condition holds, for example, if u is a minimal surface bounded by a polygonal boundary curve.

define $D^2A[u]$ as a bilinear form operating on tangent vectors ϕ and ψ. This form can be diagonalized. If it is positive definite, the minimal surface u is called *stable*. In that case, u is a relative minimum of area in the C^0 topology. However, u might be a relative minimum without being stable, if $D^2A[u]$ has a non-trivial kernel; and of course $D^2A[u]$ will have a kernel when $A[u]$ is not a relative minimum. We write $D^2A[u](\phi) = D^2A[u](\phi, \phi)$; the kernel consists of those ϕ for which $D^2A[u](\phi) = 0$.

We consider variations of the form

$$\tilde{u} = u + t\phi N.$$

These are called "normal variations" because they are in the direction of the unit normal. One can more generally consider variations with a tangential component, and shall do so at the end of this lecture.

The question arises as to what kind of function ϕ can be. In order to use the fundamental lemma of the calculus of variations, C^3 is enough. But we need to check that our formulas are valid for more general ϕ. For example, one case of interest to us is when $\phi = \max(0, \psi \circ N)$ for some smooth function ψ on the unit sphere. In fact, our calculations will work if ϕ is in the Sobolev space of functions in $W^{1,2}(\Omega)$ with generalized boundary values 0 on $\partial\Omega$.[2]

We assume that ϕ is piecewise C^3 in the closed unit disk. That means that the unit disk can be decomposed into a finite number of domains, separated by a finite number of closed C^1 arcs meeting at a finite number of points, such that ϕ is C^3 on each domain. That will cover the example, since the zero set of $\psi \circ N$ is C^1 by the implicit function theorem, since N is C^1 up to the boundary.

The calculation given in Lecture 1 shows that

$$\tilde{E}\tilde{G} - \tilde{F}^2 \quad = \quad EG - F^2 + 2t\varphi W^2 H$$

at every point where ϕ is C^1, that is, except on a finite number of C^1 arcs. Hence the integration of this expression can still be performed, yielding the standard formula for the first variation,

$$DA[u](\varphi) \quad = \quad \int_D \varphi H W \, dx \, dy$$

now established for piecewise C^1 normal variations.

We write g_{ij}, b_{ij}, and c_{ij} for the coefficients of the first, second, and third fundamental forms I, II, and III (see Lecture 1). H and K are the mean and Gauss curvature (also defined in Lecture 1).

Lemma 7 $c_{ij} - 2Hb_{ij} + Kg_{ij} = 0$. *This equation is sometimes written as an equation between bilinear forms on the tangent space:* $III - 2HII + KI = 0$.

[2] For readers not familiar with the Sobolev spaces: $W^{1,2}(\Omega)$ is the space of real-valued functions ϕ on Ω such that ϕ and its first derivatives are square-integrable on Ω. In general $W^{k,p}$ is the space of functions whose derivatives up to the k-th order have integrable p-th powers. Note that [10] uses H instead of W.

Proof. The principal curvatures κ_1 and $\kappa 2$ are the eigenvalues of the Weingarten map. The mean curvature $H = \frac{1}{2}(\kappa 1 + \kappa 2)$ and the Gauss curvature $K = \kappa 1 \kappa 2$ are the elementary symmetric functions of κ_1 and $\kappa 2$. Therefore κ_1 and κ_2 are roots of the polynomial

$$(x - \kappa_1)(x - \kappa_2) = x^2 - 2Hx + K.$$

According to the Cayley-Hamilton theorem, the Weingarten map S satisfies this same polynomial:

$$S^2 - 2HS + K\mathrm{ID} = 0$$

where ID is the identity map on the tangent space. Applying this operator equation to u_i and then taking the inner product of the result with u_j we obtain

$$\begin{aligned} 0 &= S(S(u_i))u_j - 2HS(u_i)u_j + Ku_iu_j \\ &= S(u_i)S(u_j) - 2HS(u_i)u_j + Ku_iu_j \qquad \text{since } S \text{ is self-adjoint} \\ &= c_{ij} - 2Hb_{ij} + Kg_{ij} \end{aligned}$$

That completes the proof of the lemma.

5.2 The second variation of area for a normal variation

This calculation is a fundamental result, vital for many results about minimal surface theory. It is taken up in [12], section 102, page 95, where the general case of a variation with both normal and tangential components is considered. However, only the result is given–Nitsche says, "By a direct but lengthy calculation (which we omit owing to lack of space)". This in a book of more than 560 pages. More details can be found in [10], pp. 83-84, but it is still a bit difficult to follow at equation (14). The following calculation proceeds along Hildebrandt's lines, but fills in more details.

We write u_1 for u_x and u_2 for u_y. Then

$$\begin{aligned} \tilde{u} &= u + t\phi \\ \tilde{u}_i &= u_i + t\phi_i N + t\phi N_i \qquad \text{where } u \text{ is } C^1 \end{aligned}$$

Hence, at points where ϕ is C^1 and which are not the exceptional points on the boundary,

$$\begin{aligned} \tilde{g}_{ij} &= \tilde{u}_i\tilde{u}_j \\ &= (u_i + t(\phi_i N + \phi N_i))(u_j + t(\phi_j N + \phi N_j)) \\ &= g_{ij} - 2t\phi b_{ij} + t^2(\phi_i\phi_j + \phi^2 N_iN_j) \\ &= g_{ij} - 2t\phi b_{ij} + t^2(\phi_i\phi_j + \phi^2 c_{ij}) \end{aligned}$$

Note that if we had started by allowing $\tilde{u}$ to depend in a more complicated, but still C^2, way on t, we would get these same formulas, but multiplied by $1+O(t)$, which would not make any difference below; and similarly, considering $D^2[A]$ as a bilinear functional complicates the calculation only notationally; so for simplicity we stick to one tangent vector ϕN.

We write g for $\det g_{ij} = g_{11}g_{22} - g_{12}^2 = W^2$. Recall from Lecture 1 that

$$2Hg = 2HW^2 = b_{ij}g^{ij} = b_{11}g_{22} + b_{22}g_{11} - 2b_{12}g_{12}.$$

$$Kg = b = \det b_{ij} = b_{11}b_{22} - b_{12}^2.$$

Now we calculate $\det \tilde{g}_{ij}$. We have

$$\begin{aligned}
\det \tilde{g}_{ij} &= \tilde{g}_{11}\tilde{g}_{22} - \tilde{g}_{12}^2 \\
&= (g_{11} - 2t\phi b_{11} + t^2(\phi_x^2 + \phi^2 c_{11}))(g_{22} - 2t\phi b_{22} + t^2(\phi_y^2 + \phi^2 c_{22})) \\
&\quad -(g_{12} - 2t\phi b_{12} + t^2(\phi_x\phi_y + \phi^2 c_{12}))^2 \\
&= g - 2t\phi(b_{11}g_{22} + b_{22}g_{11} - 2b_{12}g_{12}) \\
&\quad +t^2(\phi_x^2 g_{22} + \phi_y^2 g_{11} - 2\phi_x\phi_y g_{12}) \\
&\quad +t^2\phi^2(4b_{11}b_{22} - 4b_{12}^2 + g_{11}c_{22} + g_{22}c_{11} - 2g_{12}c_{12}) + O(t^3) \\
&= g - 4t\phi Hg + t^2\phi_i\phi_j g^{ij} + t^2\phi^2(4Kg + g^{ij}c_{ij}) + O(t^3)
\end{aligned}$$

By the lemma, $c_{ij} = 2Hb_{ij} - Kg_{ij}$, so $g^{ij}c_{ij} = 2Hg^{ij}b_{ij} - Kg^{ij}g_{ij} = 4H^2g - 2Kg$, where we used $g^{ij}g_{ij} = 2g$ in the last step. Hence

$$\begin{aligned}
\det \tilde{g}_{ij} &= g - 4t\phi Hg + t^2\phi_i^2 g^{ij} + t^2\phi^2(4Kg + 4H^2g - 2Kg) + O(t^3) \\
&= g - 4t\phi Hg + t^2\phi_i\phi_j g^{ij} + t^2\phi^2(4H^2g + 2Kg) + O(t^3)
\end{aligned}$$

The term $\phi_i\phi_j g^{ij}$ can be written in terms of the "first Beltrami operator" as $g|\nabla_u\phi|^2$. Our final result for $\det \tilde{g}_{ij}$ is thus

$$\det \tilde{g}_{ij} = g\{1 - 4t\phi H + t^2[|\nabla_u\phi|^2 + \phi^2(4H^2 + 2K)]\} + O(t^3)$$

We have

$$\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \dots$$

for small x, and hence

$$\sqrt{1 + t\alpha + t^2\beta} = 1 + \frac{\alpha}{2}t + \left(\frac{\beta}{2} - \frac{\alpha^2}{8}\right)t^2 + O(t^3)$$

Hence

$$\begin{aligned}
\sqrt{\det \tilde{g}_{ij}} &= \sqrt{g}[1 = 2t\phi H + t^2\{\frac{1}{2}|\nabla_u\phi|^2 + K\phi^2 + 2H^2\phi^2 - 2H^2\phi^2\}] + O(t^3) \\
&= \sqrt{g}[1 - 2tH\phi + t^2\{\frac{1}{2}|\nabla_u\phi|^2 + K\phi^2\}] + O(t^3)]
\end{aligned}$$

Note the miraculous cancellation of the H terms in the t^2 part! This means that we will get the same formula for the second variation of area, whether u is minimal or not.

The formula for area is

$$A(\tilde{u}) = \int_\Omega \sqrt{\det \tilde{g}_{ij}}\, dx\, dy$$

Note that the $O(t^3)$ term is actually of the form $t^3F + t^4G$ for some functions F and G that are explicitly known, even though we haven't written them out, and are integrable over Ω, so the integral of the $O(t^3)$ term is still $O(t^3)$. This is obvious if g is bounded below, as it is when u is regular (as assumed in [10]), but still true even if u is a branched minimal surface. The functions F and G are defined in terms of the first, second, and third fundamental forms of u. Specifically, we have from the equation just preceding the first mention of $O(t^3)$,

$$F = 2\phi b_{11}(\phi_y^2 + \phi^2 c_{22}) + 2\phi b_{22}(\phi_x^2 + \phi^2 c_{11}) - 4\phi b_{12}(\phi_x\phi_y + \phi^2 c_{12})$$

$$G = (\phi_x^2 + \phi^2 c_{11})(\phi_y^2 + \phi^2 c_{22}) + (\phi_x\phi_y + \phi^2 c_{12})^2$$

Since we have assumed that the three fundamental forms are bounded, these functions are integrable, and hence the integral of the $O(t^3)$ term is still $O(t^3)$.

The second variation of area in the direction ϕ is thus given by the t^2 term:

$$\begin{aligned}
D^2A[u](\phi) &= \frac{d^2}{dt^2}A[\tilde{u}] \\
&= \int_\Omega \{|\nabla_u\phi|^2 + 2K\phi^2\}\sqrt{g}\, dx\, dy \\
&= \int_\Omega \{|\nabla_u\phi|^2 + 2K\phi^2\}\, dA \\
&= \int_\Omega \{|\nabla\phi|^2 + 2KW\phi^2\}\, dx\, dy
\end{aligned}$$

The last two lines express the second variation as an integral on the surface, and then as an integral on the parameter domain. As Hildebrandt *et. al.* point out (p. 84 of [10]), $D^2A[u]$ can be considered as a functional defined on the Sobolev space of functions ϕ in $H^{1,2}(\Omega)$ with (generalized) boundary values 0 on $\partial\Omega$. This formula is valid whether u is harmonic or not, and whether u is minimal or not.

5.3 Non-normal variations

Here we consider *harmonic* variations k mapping the unit disk into R^3 and tangential on the boundary, i.e. $k(e^{i\theta})$ is tangent to Γ at $u(e^{i\theta})$. The main result is that first and second variations depend only on the normal component.

Theorem 14 *Let u be a harmonic surface defined in the unit disk and bounded by a Jordan curve Γ. Let N be the unit normal.*

(i) Suppose HW is integrable on the unit disk. Then the area functional A is Frechet-differentiable at u, and $DA[u](k) = DA[u](k \cdot N)$

(ii) If u is a minimal surface, the second Frechet derivative $D^2A[u]$ is a well-defined bilinear form. Then we have

$$D^2A[u](h,k] = D^2A[u][h \cdot N, k \cdot N];$$

that is, the second variation depends only on the normal component.

Proof. *Ad (i).* Let u be harmonic with HW integrable. Let u^t be any C^1 one-parameter family of $C^{k,\beta}$ harmonic surfaces with $u^0 = u$. Write the partial derivative u_t^0 in the form $A + \phi N$. Represent A on the boundary in the form $A_1 e + A_2 n$, where e is a unit vector tangent to Γ and n is a unit vector $e \times N$. Then a straightforward calculation, imitating our earlier calculation for normal variations, shows that

$$DA[u](u_t^0) = -\int_D HW\phi \, dx\, dy + \int_\Gamma A_2 \, ds$$

where ds is the element of arc length along the boundary. In the case of tangential variations we have $A_2 = 0$, so the first variation depends only on the normal component ϕ.

Ad (ii). To calculate the second variation, we consider a two-parameter family u^{st} with $u^{00} = u$, and compute

$$\frac{\partial^2 A[u^{st}]}{\partial s \partial t}.$$

Let $h = u_t^{00}$ and $k = u_s^{00}$, and $\phi = h \cdot N$, and $\psi = k \cdot N$. Then

$$\begin{aligned} D^2A[u^t](h,k) &= \frac{\partial^2 A[u^{st}]}{\partial s \partial t} \\ &= \int_D \psi(-\triangle\phi + 2KW\phi)\, dx\, dy \end{aligned}$$

These computations are not (yet) supplied here; see [12], p. 94, for another case of stating the result without including the computations–at least I am in good company.

Chapter 6

Eigenvalues and the Gauss Map

6.1 The Gauss map of a minimal surface

KW is the Jacobian of the Gauss map N, considered as a map from D to the Riemann sphere S^2. This map is conformal when u is a minimal surface (in harmonic isothermal form). The easiest way to see this is to consider the composition of the Gauss map with stereographic projection. We have an explicit formula for stereographic projection, and we can then work out an explicit formula for the stereographic projection of N. This turns out to be nothing but the function g in the Weierstrass representation of u. That is, if S is stereographic projection from S^2 to R^2, then $S \circ N$ is a meromorphic map from D to R^2, with poles where N points in the positive Z-direction. Here are the details:

Theorem 15 *Let u be a minimal surface, N the unit normal to u, and f and g the functions in the Weierstrass representation of u. Then g is the stereographic projection of N, and hence N is conformal.*

Proof. We calculate the basic differential-geometric quantities of u in terms of f and g in the Weierstrass representation of u. First, the tangent vectors u_x and u_y determine $f = u_z = u_x - iu_y$. Then

$$\begin{aligned} W^2 &= |u_x|^2 = |u_y|^2 \\ &= \frac{1}{2}|u_z|^2 \\ &= \frac{1}{8}\left|\begin{bmatrix} f(1-g^2) \\ if(1+g^2) \\ 2fg \end{bmatrix}\right|^2 \\ &= \frac{1}{8}[f\bar{f}(1-g^2)(1-\bar{g}^2) + f\bar{f}(1+g^2)(1+\bar{g})^2 + 4f\bar{f}g\bar{g}] \end{aligned}$$

$$\begin{aligned} &= \frac{1}{8}|f|^2[1+|g|^4+1+|g|^4+4|g|^2] \\ &= \frac{1}{4}|f|^2[1+|g|^4+2|g|^2] \\ W^2 &= \left[\frac{|f(1+|g|^2)}{2}\right]^2 \end{aligned}$$

Next we calculate N. We calculate

$$u_x \times u_y = \frac{|f|^2(1+|g|^2}{4}\begin{bmatrix} 2\Re\, g \\ 2\Im\, g \\ |g|^2-1 \end{bmatrix}$$

It follows that

$$|u_x \times u_y| = W^2$$

and hence

$$\begin{aligned} N &= \frac{u_x \times u_y}{|u_x \times u_y|} \\ &= \frac{1}{|g|^2+1}\begin{bmatrix} 2\Re\, g \\ 2\Im\, g \\ |g|^2-1 \end{bmatrix} \end{aligned}$$

Stereographic projection S maps the sphere S^2 to the plane.[1] The map is defined by

$$S((x_1,x_2,x_3)) = \frac{x_1+ix_2}{1-x_3}.$$

The equation for the inverse of stereographic projection is

$$S^{-1}(z) = \frac{1}{|z|^2+1}\begin{bmatrix} 2\Re\, z \\ 2\Im\, z \\ |z|^2-1 \end{bmatrix}$$

Thus g is the stereographic projection of N.

6.2 Eigenvalue problems

This section reviews the basic facts about eigenvalues and eigenfunctions. Two classical references are [8], Chapters 10 and 11, and [6], Chapter V, especially pp. 297 *ff.*

[1] Geometrically, we picture the sphere (or radius 1) lying with its north pole on the Z-axis, and its equator is the unit circle in the XY plane. The image $S((x_1,x_2,x_3))$ is the point where the line joining the north pole to (x_1,x_2,x_3) meets the XY plane. It is also possible to visualize a sphere of radius $1/2$ with its south pole at origin, and its center at $(0,0,1/2)$. The equations are the same.

We consider the problem

$$\triangle\phi + \lambda f(x,y)\phi = 0$$

where $f(x,y)$ is a given function, nonnegative in $\bar{D}$, and positive except at isolated points, and ϕ is to map $\bar{D}$ into R^3, vanish on the boundary, and satisfy the given equation in D. Values λ for which a nonzero solution ϕ can be found are called *eigenvalues*.

A suitable set of functions ϕ defined on $\bar{D}$ and vanishing on S^1 can be turned into a Hilbert space with the inner product

$$\langle \phi, \psi \rangle = \int_D \nabla\phi(x,y)\nabla\psi(x,y)\, dx\, dy$$

The *Rayleigh quotient* is defined by

$$R[\phi] = \frac{\int_D |\nabla\phi|^2\, dx\, dy}{\int_D f\phi\psi\, dx\, dy}$$

The infimum of $R[\phi]$ over all ϕ exists and is the least eigenvalue λ_{min}.

There exist infinitely many eigenvalues, and the corresponding eigenfunctions form an orthonormal basis for the Hilbert space mentioned. The least eigenvalue has only a one-dimensional eigenspace. That is, it is non-degenerate (an eigenvalue is called *degenerate* if the corresponding eigenspace is of dimension greater than one). Each of these eigenfunctions is smooth in the interior of the domain–at least as smooth as $f(x,y)$ is. Moreover, their zero sets are a union of piecewise smooth arcs; the gradient is zero only at isolated points. In [6] there are pictures of these arcs for several examples.

One can also consider eigenvalue problems over other plane domains than the disk. One can also consider eigenvalue problems on a surface. If Ω is a plane domain, or a domain on a surface, we let $\lambda_{\min}(\Omega)$ be the least eigenvalue of the equation $\triangle\phi - f(x,y)\phi = 0$ in Ω, $\phi = 0$ on $\partial\Omega$.

Theorem 16 (Monotonicity of the least eigenvalue) . *If $\Omega_1 \subset \Omega_2$, we have*

$$\lambda_{\min}(\Omega_1) > \lambda_{\min}(\Omega_2).$$

Proof. The least eigenvalue is given by the infimum of the Rayleigh quotient. But competitors for the Rayleigh quotient can be allowed to be continuous but only piecewise differentiable, so the least eigenfunction of the smaller domain Ω_1, extended to be zero outside Ω_1, is eligible to count in the infimum of Rayleigh quotients for $\lambda_{\min}(\Omega_2)$. Hence $\lambda(\Omega_1) \le \lambda(\Omega_2)$. But if Ω_1 is strictly contained in Ω_2, then this function is not smooth in the interior of Ω_2, contradiction.

Theorem 17 (Properties of the least eigenfunction) *The least eigenfunction, i.e. the eigenfunction corresponding to the least eigenvalue, has only one sign.*

Proof. Let ϕ be the least eigenfunction. Define $\psi = |\phi|$. Then the Rayleigh quotient $R[\psi]$ has the same value as $R[\phi]$. If ϕ does not have one sign, then the zeroes of ϕ include an arc, and along that arc ψ is not smooth, since $\nabla\phi$ only has isolated zeroes. Hence we can round off ψ slightly, decreasing the Rayleigh quotient. Hence $R[\phi] = R[\psi]$ is not a minimum, contradiction.

Lemma 8 (Hopf boundary-point lemma) *Let ϕ be any solution of $\triangle\phi + \lambda f(x, y)\phi = 0$ which is nonnegative in a neighborhood of a boundary point p. Then the normal derivative $\phi_n u$ is not zero at p.*

Proof. See [7], p. 519 (or any other good textbook on PDE). This is a property of second-order linear elliptic equations, and the proof takes us too far afield.

Corollary 3 *If ϕ is the first eigenfunction then ϕ_r does not vanish on S^1.*

6.3 Eigenvalues and the Gauss Map

The connection of these classical results to minimal surfaces arises when we take the function $f(x, y)$ to be $-KW$, where K is the Gaussian curvature $\kappa_1\kappa_2$, and W is the area element $\det g_{ij} = \sqrt{EF - G^2}$. Since for minimal surfaces we have $\kappa_2 = -\kappa_1$, the Gaussian curvature K is always negative, or at least not positive, so $-KW$ is nonnegative. It might be zero at branch points (where W is zero), and it might be zero at *umbilic points*, where K is zero.

KW is the Jacobian of the Gauss map N. Therefore, in case the Gauss map is one-to-one, the eigenvalue problem $\triangle\phi - \lambda KW\phi = 0$ in D is equivalent to the eigenvalue problem for the Laplace-Beltrami operator on the sphere:

$$\triangle\phi + \lambda\phi = 0$$

on the spherical domain $N(D)$. In general the Gauss map is not one-to-one, so the eigenvalue problem corresponds to a "multiply-covered" domain on the sphere, intuitively. This can be made precise but it is not worth the trouble; one may use the intuition, but formally one just works with $\triangle\phi - \lambda KW\phi$ in D.

6.4 The eigenvalue problem associated with $D^2A[u]$

Theorem 18 *Let u be a minimal surface (not necessarily in harmonic isothermal form). Then the kernel of $D^2A[u]$ is exactly the space of solutions of the eigenvalue problem*

$$\triangle\phi - 2KW\phi = 0.$$

In particular $D^2A[u]$ has a nontrivial kernel if and only if 2 is an eigenvalue of $\triangle\phi - \lambda KW\phi = 0$.

6.5 The Gauss-Bonnet theorem

The *total curvature* of a surface is by definition $\int_D KW\,dx\,dy$. This is the area (counting multiplicities) of the "spherical image" of the surface, that is, the range of the unit normal N.

The *geodesic curvature* of the boundary κ_g is the component of the curvature vector of the boundary in the direction tangent to the surface. In more detail: at each point of the Jordan curve Γ, there is a unit tangent $T = \Gamma_s$, where s is arc length along Γ. The rate of change T_s is the curvature vector of Γ, which is normal to T, and can be broken into a component in the direction of the surface normal N, and a component orthogonal to that (hence tangent to the surface), whose magnitude is defined as the geodesic curvature, κ_g, of Γ (relative to the specific surface u bounded by Γ).

The *total curvature* of a Jordan curve Γ is the integral of the magnitude of the curvature vector around the curve. The magnitude of the curvature vector is usually written κ. (Curiously, there is no standard notation for the curvature vector itself.) Thus, the total curvature is $\int_\Gamma \kappa ds$. This quantity depends only on Γ, not on some surface bounded by Γ. Of course the geodesic curvature for any surface u bounded by Γ is bounded above by the total curvature of Γ.

These quantities are connected by the Gauss-Bonnet theorem, stated below. The proof of that theorem requires a pretty formula for the geodesic curvature, which we give in the next lemma. We assume u is given in isothermal coordinates. (Use Lichtenstein's theorem if necessary, or if u is minimal, use uniformization.)

Lemma 9 (Minding's formula) *The geodesic curvature of a surface defined in the unit disk, in isothermal coordinates, is given by*

$$\kappa_g\sqrt{W} = 1 + (\log\sqrt{W})_r.$$

Proof. See [10], p. 33. The proof is a straightforward calculation, but it involves the Christoffel symbols, which we did not introduce in Lecture 1.

Theorem 19 (Gauss-Bonnet) *If u is a regular surface of class C^2 bounded by a C^3 Jordan curve Γ then*

$$\int_D KW\,dx\,dy + \int_0^{2\pi} \kappa_g(\theta)\sqrt{W}d\theta = 2\pi$$

This can equally well be written as

$$\int_u K\,dA + \int_\Gamma \kappa_g ds = 2\pi$$

Remarks. In case u is a minimal surface, KW is negative. In case u lies in a plane, K is zero and the equation is obvious. In case u is a spherical cap, the geodesic curvature integrates to less than 2π, but K is positive.

Proof. In isothermal coordinates, the Gauss curvature has a simple and remarkable formula:

$$-KW = \triangle \log \sqrt{W}.$$

This is a special case of Gauss's *Theorema Egregium*, which expresses the Gauss curvature K in terms of the first fundamental form g_{ij}, in spite of the fact that it was defined using the second fundamental form. In particular, if we write out $\triangle \log \sqrt{W}$ in terms of derivatives, we find:

$$\begin{aligned} \triangle \log \sqrt{W} &= (\log \sqrt{W})_{xx} + (\log \sqrt{W})_{yy} \\ &= \frac{\partial}{\partial x} \frac{(\sqrt{W})_x}{\sqrt{W}} + \frac{\partial}{\partial y} \frac{(\sqrt{W})_y}{\sqrt{W}} \end{aligned}$$

which is the Theorema Egregium for the case of isothermal coordinates. See for example [10], formula (27), page 30; and the proof of the Theorema Egregium can be found on the preceding pages.

Proceeding to the proof of the Gauss-Bonnet theorem, we integrate this formula for $-KW$:

$$\begin{aligned} -\int KW\,dx\,dy &= \int \triangle \log \sqrt{W}\,dx\,dy \\ &= \int_0^{2\pi} (\log \sqrt{W})_r\,d\theta \qquad \text{by Green's theorem} \\ &= \int_0^{2\pi} \kappa_g \sqrt{W} - 1\,d\theta \qquad \text{by Minding's formula} \\ &= \int_0^{2\pi} \kappa_g \sqrt{W}\,d\theta - 2\pi \end{aligned}$$

which proves the theorem.

6.6 The Gauss-Bonnet theorem for branched minimal surfaces

There is a beautiful extension of the Gauss-Bonnet theorem to branched minimal surfaces. The formula can be thought of this way: Each interior branch point of order m, and each boundary branch point of order $2m$, counts as $2m\pi$ of total curvature. In other words, if the surface is perturbed so as to "split" or "break" the branch point, there will necessarily be a lot of curvature created.

Theorem 20 *If u is a branched minimal surface defined and C^2 in the closure of a simply-connected domain, and M is the sum of the orders of the interior branch points plus half the orders of the boundary branch points, then the*

$$\int_D KW\,dx\,dy - 2M\pi + \int_0^{2\pi} \kappa_g(\theta)\sqrt{W}d\theta = 2\pi$$

Remark. If u spans a Jordan curve, then the boundary branch points must be of even order. However, this is not an assumption of the theorem; boundary branch points of odd order are also allowed, even if the boundary cannot be taken on monotonically in that case.
Remark. The theorem can also be generalized to surfaces defined in multiply-connected domains (see [11] p. 121) and to surfaces with piecewise C^1 boundaries, for example polygonal boundaries (see [10], p. 37).
Proof. One cuts the branch points out of the domain, preserving the simply-connectedness of the domain, by first connecting each branch point to the boundary (by a set of non-intersecting arcs, one per branch point) and then "fattening" each arc a tiny bit. If we think of the domain as an island and its exterior as the sea, we are running a river from the sea to each interior branch point, which is the source of that river. At the branch point, we make a small circle around the branch point, so the river connects to a "pond" containing the branch point. Now the Gauss-Bonnet theorem applies to the (regular) surface defined in the island minus the rivers. The contributions to the geodesic curvature along the banks of the rivers very nearly cancel out, and do cancel out when we take the limit as the river width goes to zero, since the inward normals are in opposite directions on opposite banks. Near each interior branch point there is an almost-closed circle; this maps to almost $m + 1$ circles on the surface and hence contributes $(m + 1)2\pi$ to the geodesic curvature. At the river mouths, however, there are two ninety-degree turns in the opposite direction, and where the river joins the "pond" containing the branch point, there are two more ninety-degree turns, so the net contribution from each interior branch point is $(m+1)2\pi - 2\pi = 2m\pi$. Similarly, at a boundary branch point of order $2m$ there is a semicircle that contributes $(2m + 1)\pi$, and two ninety-degree turns in the opposite direction, so the net contribution is $(2m + 1)\pi - \pi = 2m\pi$. Taking the limit as the river width goes to zero, the proof is completed.

6.7 Laplacian of the Gauss map of a minimal surface

Theorem 21 *Let u be a minimal surface. Then the Laplacian of its unit normal is given by the following formula:* $\triangle N = 2KWN$

Proof. To prove this elegantly, we make use of the general fact that the Laplace-Beltrami operator of any surface S, applied to the position vector of S, is exactly twice the mean curvature of S. Apply this facto to the Riemann sphere, whose position vector $h(w)$ coincides with its unit normal. Thus $\triangle h = -2h$. Next, note that the map from the disk D to the Riemann sphere induced by the Gauss map of u is a conformal map with Jacobian $-KW$. Under a conformal map, the Laplace-Beltrami operator changes to the Laplace-Beltrami operator on the range surface, multiplied by the Jacobian of the mapping. Hence $\triangle N = 2KWN$, and the theorem is proved.

Chapter 7

Second Variation of Dirichlet's Integral

7.1 Tangent vectors and the weak inner product

Consider the space of all $C^{k,\beta}$ surfaces bounded by a Jordan curve Γ in the non-monotonic sense, i.e. we do not require that the surface take the boundary monotonically. The first variation of Dirichlet's integral is easily computed from the formula $E(u) = (1/\pi) \int u u_r \, d\theta$. When we write a subscript r by a function of θ, it means the partial derivative of the harmonic extension, evaluated at $r = 1$. The first variation is given by the formula

$$DE(u)[k] = \frac{1}{\pi} \int k u_r \, d\theta.$$

Here k is a "tangent vector" to u in the manifold of harmonic surfaces; that is, a function from S^1 to R^n such that $k(\theta)$ is tangent to Γ at $u(e^{i\theta}$, for each θ, and k is $C^{k,\beta}$. It follows from the fundamental lemma of the calculus of variations that the minimal surface equation can be written as $u_\theta u_r = 0$ on S^1.

In case u has no boundary branch points, every tangent vector has the form λu_θ for some scalar function λ. If u has boundary branch points, this is not so. In that case the tangent vectors may be nonzero at the boundary branch point. Let Γ be given in arc-length parametrization; suppose $u = \Gamma \circ h$. Then all tangent vectors have the form $\lambda \Gamma_\theta \circ h$ for some scalar function λ.

There is an inner product on the space of tangent vectors to a given minimal surface u, namely that defined by $(h, k) = \int h_r k \, d\theta$. Tromba calls this the "weak inner product"; we use the terms "orthogonal" and "weak orthogonal" interchangeably. Note that the space of tangent vectors is not complete under the metric induced by this inner product. (It is "weak" because the topology induced by this norm is weaker than the norm inherited from $C^{k,\beta}(S^1, R^n)$.)

If we extend h and k to the disk by harmonic extension, then $(h, k) =$

$\int_D \nabla h \nabla k \, dx \, dy$, since

$$\begin{aligned}
\int_D \triangle(hk) &= \int_D h\triangle k + 2\nabla h \nabla k + k \triangle h \, dx \, dy \\
\int (hk)_r \, d\theta &= \int_D 2\nabla h \nabla k \, dx \, dy \\
\int h_r k + k_r h \, d\theta &= \int_D 2\nabla h \nabla k \, dx \, dy
\end{aligned}$$

and $\int h_r k \, d\theta = \int k_r h \, d\theta$. Alternately one can apply Green's theorem directly to observe that both (h,k) and (k,h) are equal to $\int_D \nabla h \nabla k \, dx \, dy$.

7.2 The conformal group

The conformal group of the disk (the group of conformal transformations of the disk) acts on the space of harmonic surfaces by composition (i.e. by reparametrization), and preserves the property of being a minimal surface. One calculates that the tangent directions to u introducted by this action, which we call the *conformal directions*, are of the form λu_θ, where λ has the form $a + b\cos\theta + c\sin\theta$. Each minimal surface is part of a three-parameter family of minimal surfaces differing only by the action of the conformal group. Therefore we wish to impose some restriction on the class of harmonic surfaces considered, so that only one member of each such family will be allowed. The traditional way to do this has been to impose a "three-point condition", requiring that three given points on S^1 be transformed to three given boundary points. Another way, more suited to the global-analytic approach of [17], is to restrict attention to the family E of harmonic surfaces defined in [17]. The definition of E depends on a fixed minimal surface u, and it is defined to be a co-dimension 3 submanifold of the space of harmonic surfaces bounded by Γ (in the non-monotonic sense) such that the three conformal directions at u are orthogonal to the tangent space of E at u. For our purposes, it is not important exactly how the conformal group action is disposed of–it is only necessary to realize the necessity of doing so.

We sometimes have to work with a one-parameter family of minimal or harmonic surfaces; we denote the dependence on the parameter t using a superscript u^t, since there is little chance of confusing that t with an exponent. Sometimes the superscript t is omitted, as in u_t, which means the partial derivative of u^t with respect to t. We call such a family *non-trivial* if $u_t = t^a h$ for some positive number a and some tangent vector h (depending on t) such that h is not a conformal direction when $t = 0$.

It should be noted that $a > 0$ is allowed; that is, u_t can vanish when $t = 0$. When we use the tools of global analysis to prove the existence of one-parameter families under certain conditions, those families are real-analytic in t; so if $a > 0$, we can reparametrize using a t^a as the new parameter; but the new family is only C^1, since it is real-analytic in a rational power of t rather than in t. The families constructed by glogal analysis have the property that, when so reparametrized, their tangent vectors are nowhere conformal. In fact, if the surfaces u^t are not all conformal reparametrizations of the same surface, then we can "project" each u^t onto its sole representative in E , obtaining as a result a non-trivial family, each member of which is a conformal reparametrization of the corresponding surface in the original family.

7.3 Calculation of the second variation of E

Theorem 22 (Tromba) *The second variation of Dirichlet's integral is given by*

$$D^2E[u](h,k) = \int k(h_r - \tilde{h}_\theta)\, d\theta$$

where $k = \lambda u_\theta$ and $\tilde{h} = \eta u_r$. The tangent vector k to the minimal surface u belongs to Ker $D^2E[u]$ if and only if

$$k(k_r - \tilde{k}_\theta) = 0$$

or equivalently

$$k_z u_z = 0$$

Proof. Let $k = \lambda u_\theta$ and $h = \eta u_\theta$ be two tangent vectors to the minimal surface u. Then

$$D^2E[u](h,k) = \frac{\partial^2 E[\tilde{u}]}{\partial s \partial t}$$

where $\tilde{u}$ is defined by

$$\tilde{u}(\theta) = u(\theta + t\lambda + s\eta).$$

Define

$$\begin{aligned} f^{11} &= \lambda^2 u_{\theta\theta} \\ f^{22} &= \eta^2 u_{\theta\theta} \\ f^{12} &= \lambda\eta u_{\theta\theta} \end{aligned}$$

Expand $\tilde{u}$ to second order in s and t:

$$\tilde{u} = u + kt + hs + \frac{1}{2}f^{11}t^2 + \frac{1}{2}f^{22}s^2 + f^{12}st + \dots$$

Differentiating, we have

$$\begin{aligned} \tilde{u}_x &= u_x + k_x t + h_x s + \frac{1}{2}f^{11}_x t^2 + \frac{1}{2}f^{22}_x s^2 + f^{12}_x st + \dots \\ \tilde{u}_y &= u_y + k_y t + h_y s + \frac{1}{2}f^{11}_y t^2 + \frac{1}{2}f^{22}_y s^2 + f^{12}_y st + \dots \end{aligned}$$

Thus

$$\tilde{u}_x^2 = ux^2 + tk_x u_x + sh_x u_x + st(f^{12}_x u_x + k_x h_x) + t^2(k_x^2 + \frac{1}{2}u_x f^{11}_x) + s^2(h_x^2 + \frac{1}{2}u_x f^{22}_x) + \dots.$$

Adding this with the corresponding expression for u_y^2, we get

$$\begin{aligned} |\nabla \tilde{u}|^2 &= \tilde{u}_x^2 + \tilde{u}_y^2 \\ &= |\nabla u|^2 + t\nabla k \nabla u + s\nabla h \nabla u + 2st(\nabla f^{12} + \nabla k \nabla h) \\ &\quad + t^2(|\nabla k|^2 + \nabla u \nabla f^{11}) + s^2(|\nabla h|^2 + \nabla u \nabla f^{22}) + \dots \end{aligned}$$

Integrating this expression, on the left side we get $2E[\tilde{u}] = \int_D |\nabla \tilde{u}|^2 \, dx \, dy$. Looking at the st term on the right, we find

$$\frac{\partial^2 E[\tilde{u}]}{\partial s \partial t} = \int_D \nabla f^{12} \nabla u + \nabla k \nabla h \, dx \, dy$$

Applying Green's theorem, we have

$$\int_0^{2\pi} f^{12} u_r + k h_r \, d\theta = \int_0^{2\pi} \lambda \eta u_r u_{\theta\theta} + k h_r \, d\theta.$$

Integrating the first term by parts, we get

$$\begin{aligned} \frac{\partial^2 E[\tilde{u}]}{\partial s \partial t} &= \int_0^{2\pi} (-\lambda \eta u_r)_\theta u_\theta + k h_r \, d\theta \\ &= \int_0^{2\pi} -k(\eta u_r)_\theta - \lambda_\theta \eta u_r u_\theta + k h_r \, d\theta \\ &= \int_0^{2\pi} -k(\eta u_r)_\theta + k h_r \, d\theta \qquad \text{since } u_r u_\theta = 0. \\ &= \int_0^{2\pi} k(h_r - \tilde{h}_\theta) \, d\theta \end{aligned}$$

That is the formula of the lemma. If the integrand vanishes for all k, the fundamental lemma of the calculus of variations yields formula (i) of the lemma.

Now to prove formula (ii). Note that (on S^1)

$$u_z k_z = z^2 (u_r - i u_\theta)(k_r - i k_\theta).$$

This will be identically zero if $\Im\left((u_r - i u_\theta)(k_r - i k_\theta\right)$ is identically zero; that is, if $u_\theta k_r + u_r k_\theta$ is identically zero. From the penultimate line in the preceding calculation we have

$$D^2 E[u](h, k) = \int_0^{2\pi} -k(\eta u_r)_\theta + k h_r \, d\theta$$

Integrating the first term by parts, we have

$$\begin{aligned} D^2 E[u](h, k) &= \int_0^{2\pi} k_\theta u_r + k h_r \, d\theta \\ &= \int_0^{2\pi} k_\theta u_r + k) r h \, d\theta \\ &= \int_0^{2\pi} \eta(k_\theta u_r + k_r u_\theta) \, d\theta \end{aligned}$$

Applying the fundamental lemma of the calculus of variations, we see that h is in the kernel if and only if $k_\theta u_r + k_r u_\theta$ vanishes identically; but we have proved this is equivalent to formula (ii). That completes the proof of the theorem.

7.4 Forced Jacobi fields

Consider the kernel equation $k_\theta(k_r - \tilde{k}_t heta) = 0$. One way in which this could be satisfied is if $k_r - \tilde{k}_\theta = 0$; vectors k satisfying this condition and not induced by the conformal group are called "forced Jacobi fields" or "forced Jacobi directions". Tromba proved that they do not occur in the absence of branch points, and that in the presnse of branch points there are two for each interior banch point (counting multiplicities) and one for each boundary branch point, so that the space of forced Jacobi fields is finite dimensional. (There can be at most finitely many branch points, even if the boundary is not real-analytic, as long as the total curvature of the surface is finite, thanks to the Gauss-Bonnet formula for branched minimal surfaces.) The forced Jacobi directions are just the directions k such that the function $K = k + i\tilde{k}$ is complex analytic, i.e. such that $\tilde{k}$ is the conjugate harmonic function of k.

Another important characterization of the forced Jacobi fields is this: they are exactly the tangent vectors of the form

$$k = \Re\,(i\omega z\, u_z)$$

where $i\omega z$ is a function meromorphic in the parameter domain, and having a pole of order at most m at each branch point of order m. Any function ω with suitable behavior on the boundary, and poles of the right orders at the branch poitns, will produce a tangent vector by this equation. The reason for writing the equation with ωz instead of with ω is that in case the parameter domain is the unit disk, the appropriate boundary condition is that ω be real on S^1. In case the parameter domain is the upper half plane, the condition is that $i\omega z$ be real on the x-axis. The Appendix of [4] contains Tromba's treatment of the forced Jacobi fields.

Lemma 10 (Tromba [17]) *Suppose u is a minimal surface, and k is a tangent vector belong to $KerD^2E[u]$ whose harmonic extension is everywhere tangent to u. Then k is a forced Jacobi direction or a direction induced by the conformal group.*

Proof. Since k is everywhere tangent to u we have $k = \alpha u_x + \beta u_y$ for some functions α and β defined in the disk. Define $\omega = -i(\alpha - \beta i)$. Then $k = \Re\,(i\omega u_z)$. We must show $\omega_{\bar{z}} = 0$, so ω is meromorphic, and also we must show that ω is analytic except for poles at the branch points of order at most the order of the branch point. Calculate:

$$\begin{aligned} k &= \Re\,(i\omega u_z) \\ &= i\omega u_z - i\bar{\omega}\bar{u}_z \end{aligned}$$

$$\begin{aligned} &= i\omega u_z - i\bar{\omega}\bar{u}_{\bar{z}} \\ k_z &= i\omega_z u_z + i\omega u_{zz} - i\bar{\omega}_z - i\bar{\omega}\bar{u}_{\bar{z}z} \\ &= i\omega_z u_z + i\omega u_{zz} - i\bar{\omega}_z \qquad \text{since } u_{z\bar{z}} = 0 \end{aligned}$$

Now take the dot product with u_z. On the left we get zero, since $k_z u_z = 0$ is the kernel equation and k is in the kernel of $D^2E[u]$ by hypothesis.

$$\begin{aligned} 0 &= i\omega_z u_z^2 + i\omega u_{zz} u_z - i\bar{\omega}_z u_z \\ &= -i\bar{\omega}_z u_z \qquad \text{since } u_z^2 = 0 \text{ and } u_z u_{zz} = (u_z^2)_z = 0 \end{aligned}$$

But $\bar{\omega}_z$ is the complex conjugate of $\omega_{\bar{z}}$. Hence $\omega_{\bar{z}} u_{\bar{z}} = 0$. That is,

$$(\alpha_x - \beta_y)u_x + (\alpha_y + \beta_x)u_y = 0.$$

This is a vector equation; since $u_x u_y = 0$, taking the dot products with u_x and u_y respectively shows that $\alpha u_x = \beta_y$ and $\alpha_y = -\beta u_z$, i.e. $\omega_{\bar{z}} = 0$ as desired. Then $K = i\omega u_z$ is analytic except perhaps at the branch points, and $k = \Re(K)$ except at the branch points. Since k is harmonic in the unit disk, K is analytic in the unit disk. Hence ω is meromorphic and has poles only at the branch points and of order not greater than the order of the branch point. That completes the proof of Tromba's lemma.

7.5 From the kernel of D^2E to the kernel of D^2A

Theorem 23 *Suppose that the minimal surface u has no (interior or boundary) branch points, and a C^k boundary. If $D^2A[u]$ is positive definite on normal variations, then $D^2E[u]$ is positive definite (on the space E , i.e. in directions not induced by the conformal group).*

Proof. For tangent vectors k to u, let $F(k)$ be the normal component of the harmonic extension of k; thus $F(k) = k \cdot N$. According to the result given at the end of Chapter 4, we have $D^2A[u](h,k) = D^2A[u](F(h), F(k)]$. Fix a tangent vector k, and let $\tilde{u}(\theta) = u(\theta + t\lambda$, so that $\tilde{u}_t = k$ when $t = 0$, where $k = \lambda u_\theta$. In view of the general inequality $E(u) \geq A(u)$, and the fact that u is a critical point of both E and A, we have

$$\frac{d^2E[u]}{dt^2}|_{t=0} \geq \frac{d^2A(u)}{dt^2}|_{t=0}.$$

Writing ϕ for $F(k)$, we have

$$D^2E[u](k,k) \geq D^2A[u](\phi, \phi).$$

By hypothesis, the right-hand side is positive for all non-zero ϕ. Hence $D^2E[u](k,k) > 0$ unless ϕ is identically zero, i.e. unless k is tangential. Then, by Tromba's lemma, k is a forced Jacobi or conformal direction. But by hypothesis, u has no branch points, so k is not a forced Jacobi direction. Hence it is a conformal direction. That completes the proof.

Note that the previous theorem works in R^n. Our next theorem is only for R^3:

Theorem 24 *Let u be a minimal surface in R^3 with C^k boundary, and unit normal N. Let k be in $KerD^2E[u]$. Then $\phi = k \cdot N$ belongs to $KerD^2A[u]$.*

Corollary 4 *If $KerD^2A[u]$ has no kernel among normal variations, $KerD^2E[u]$ contains only the conformal and forced Jacobi directions.*

Proof. The Corollary follows immediately from the theorem and Tromba's lemma. We now prove the theorem. Suppose k is in $KerD^2E[u]$; we shall show $\phi = k \cdot N$ satisfies $\triangle\phi - 2KW\phi = 0$. We have

$$\triangle\phi = (\triangle k) \cdot N + 2\nabla k \nabla N + k \triangle N.$$

The first term vanishes because k is harmonic. We claim the second term vanishes also. To prove this, fix a point z in the unit disk, and choose coordinates a and b in a neighborhood of z that diagonalize the first fundamental form at z, so that $N_a = \kappa_1 u_a$ and $N_b = \kappa_2 u_b$, where κ_1 and κ_2 are the principal curvatures of u at z. (Note: if these equations hold in a whole neighborhood, then a and b are called "local curvature coordinates"; it costs some trouble to prove they exist, and we do not need them; we need the first fundamental form to be diagonalized and one point only, which is easy by taking a and b to be a certain linear combination of x and y.) Because u is a minimal surface, we have $\kappa_1 = -\kappa_2$. Thus

$$\nabla k \nabla N = \kappa_1(k_a u_a - k_b u_b) = \kappa_1 \Re\,(e^{2i\nu} k_z u_z)$$

where ν is the angle between the positive x-direction and the positive a-direction (so ν is a function of z). Since k is assumed to be in $KerD^2E[u]$, we have $k_z u_z = 0$. Hence the term $\nabla k \nabla N$ vanishes, and we have proved $\triangle\phi = k \cdot N$.

The proof of the theorem is thus reduced to proving $\triangle N = 2KWN$. But this is Theorem 21. That completes the proof.

7.6 From the kernel of D^2A to the kernel of D^2E

In the previous section we proved that every tangent vector in the kernel of D^2E, except for the forced Jacobi and conformal vectors, has its normal component in the kernel of D^2A. In this section we address the converse question, whether every normal variation in the kernel of D^2A arises in this way, as $k \cdot N$ for some harmonic tangent vector k. This is answered positively in part (i) of the theorem below; k is found by solving a certain differential equation given in part (v). Part(ii) addresses the question of the uniqueness of k. Part (iv) characterizes the kernel of the map $F : k \mapsto k \cdot N$; in "Tromba's lemma" we identified the kernel of F restricted to $KerD^2E[u]$, but that left open the question whether it might have additional kernel not in $D^2E[u]$.

The following theorem was printed in [2], which was not a journal publication. Parts (i) and (v) were obtained independently by Schüffler in his dissertation [15], at least for the case when there are no boundary branch points. A requirement for the German Ph. D. is original publication, and the editor to whom I submitted [3] gave me a choice: omit this theorem, or delay publication until after Schüffler's. I chose to omit it, and never subsequently published the result, but it seems logically to belong in this chapter.

Theorem 25 *Let u be a minimal surface in R^3 with real-analytic Jordan boundary, not lying in a plane. Let N be the unit normal to u. D is the unit disk.*

(i) Let ϕ be a nontrivial solution of $\triangle\phi - 2KW\phi = 0$ in D, $\phi = 0$ on the boundary S^1. Then there exists a tangent vector k to u such that $\phi = k \cdot N$ and k is in $KerD^2E[u]$.

(ii) Any two vectors k as in part (i) differ by a forced Jacobi or conformal direction.

(iii)$DimKerD^2E[u] = 3 + M + DimKerD^2A[u]$, where M is the number of forced Jacobi fields, namely M is the sum of the orders of boundary branch points plus twice the orders of the interior branch points.

(iv) If $k \cdot N$ is identically zero for some tangent vector k, then k is forced Jacobi or conformal.

(v) All solutions k of the problem in part (i) may be characterized as follows: They are $k = \Re(hu_z) + \phi N$, where h is a complex valued, real analytic function in $\bar{D}$ minus the branch points of u, satisfying the following system of two partial differential equations in the two unknowns $\Re h$, $\Im h$:

$$\begin{aligned} h_{\bar{z}} &= G \qquad \textit{in } D \textit{ minus the branch points} \\ \Re(\bar{z}h) &= 0 \qquad \textit{on } S^1 \\ \textit{where } G &= \phi\frac{d^2u}{d\bar{z}^2} \cdot \frac{N}{W} \end{aligned}$$

Proof. Let G be as defined in part (v), and suppose h satisfies the equation given there. Define

$$k = \Re(hu_z) + \phi N.$$

We shall prove that $\triangle k = 0$, that $k_z u_z = 0$, and that $k(\theta)$ is tangent to the boundary of u at $u(\theta)$, except possibly at the boundary branch points. (The exception applies to all three equations.) We first check the boundary condition, which can as well be expressed in the form $k \cdot N = 0$ and $k \cdot u_r = 0$; we shall derive this from the boundary condition $\Re(\bar{z}h) = 0$ satisfied by h. First, we automatically have $k \cdot N$ on the boundary, since $u_x \cdot N = u_y \cdot N = 0$ and $\phi = 0$ on the boundary. We now compute on S^1:

$$\begin{aligned} k \cdot u_r &= \Re(hu_z) \cdot \Re(zu_z) \\ &= (Re\,(h)u_x - \Im(h)u_y)(xu_x - yu_y) \end{aligned}$$

$$
\begin{aligned}
&= (x\Re(h) + y\Im(h))W \\
&= W Re(\bar{z}h) \\
&= 0 \qquad \text{on } S^1 \text{ by hypothesis}
\end{aligned}
$$

Next we prove $k_z \cdot u_z = 0$. As usual, we may omit explicit mention of the dot product when two vectors are written side by side. We compute $k_z u_z = 0$ as follows.

$$
\begin{aligned}
2k &= 2\Re(hu_z) + 2\phi N \\
&= hu_z + \bar{h}u_{\bar{z}} + 2\phi N.
\end{aligned}
$$

Differentiating with respect to z, and using the fact that $u_{z\bar{z}} = \triangle u = 0$, we find

$$2k_z = h_z u_z + hu_{zz} + \bar{h}_z u_{\bar{z}} + 2(\phi N)_z. \tag{7.1}$$

Note that $uz^2 = 0$ (this is the minimal surface equation); and hence, differentiating with respect to z, $u_{zz}u_z = 0$. Also, observe that $u_z u_{\bar{z}} = 2W$. With these observations, take the dot product of equation (7.1) with u_z. We find

$$k_z u_z = \bar{h}_z W + (\phi N)_z u_z.$$

Since $N \cdot u_z = 0$, we can rewrite the last term, obtaining

$$k_z u_z = \bar{h}_z W + \phi(N_z u_z).$$

Since $\bar{h}_z$ is the complex conjugate of $h_{\bar{z}}$, we can substitute $\bar{G}$ for $\bar{h}_z$, obtaining

$$
\begin{aligned}
k_z u_z &= \phi u_{zz} N + \phi N_z u_z \\
&= \phi(u_z N)_z
\end{aligned}
$$

Since $u_z N = 0$, also $(u_z N)_z = 0$. We have proved $k_z u_z = 0$.

Next we prove $\triangle k = 0$. It will suffice to prove $\triangle k \cdot N$ and $\triangle k \cdot u_z$ are zero. The latter of these is easily proved: Differentiate $k_z u_z = 0$ with respect to $\bar{z}$. Since $u_{z\bar{z}} = 0$, we find $\triangle k \cdot u_z = 0$. We now set out to prove $\triangle k \cdot N = 0$.

Differentiate (7.1) with respect to $\bar{z}$. We find

$$2\triangle k = (\triangle h)u_z + h_{\bar{z}} u_{zz} + (\triangle \bar{h})u_z + \bar{h}_z u_{\bar{z}\bar{z}} + 2\triangle(\phi N) \tag{7.2}$$

Take the dot product with N. We find

$$\triangle k \cdot N = \Re(h_{\bar{z}} u_{zz} \cdot N) + \triangle(\phi N) \cdot N.$$

Substituting $h_{\bar{z}} = \phi u_{\bar{z}\bar{z}} \cdot N/W$, we have

$$\triangle k \cdot N = |u_{zz}N|^2 \phi/W + \triangle(\phi N) \cdot N. \tag{7.3}$$

Now $\triangle(\phi N) = N\triangle\phi + \phi\triangle N + \phi_x N_x + \phi_y N_y$, and since $N_x \cdot N = N_y \cdot N = 0$, we have

$$\triangle(\phi N) \cdot N = \triangle\phi + \phi\triangle N \cdot N.$$

By Theorem 21, we have $\triangle N = 2KWN$. By hypothesis we have $\triangle\phi = 2KW\phi$. Hence $\triangle(\phi N)\cdot N = 4KW\phi$, and equation (7.3) becomes

$$\triangle k \cdot N = |u_{zz}\cdot N|^2\phi/W + 4KW\phi \tag{7.4}$$

We now claim

$$|u_{zz}\cdot N|^2 = -4KW^2 \tag{7.5}$$

Once (7.5) is proved, (7.4) immediately implies that $\triangle k\cdot N = 0$.

To establish (7.5), it will be convenient to use again coordinates a and b such that at the fixed point z, we have $N_a = \kappa_1 u_a$ and $N_b = \kappa_2 u_b$. In these coordinates we have

$$\frac{\partial}{\partial a}\left(\frac{u_a}{\sqrt{W}}\cdot N\right) = -\kappa_1\sqrt{W} \qquad \text{and} \qquad \frac{\partial}{\partial b}\left(\frac{u_b}{\sqrt{W}}\cdot N\right) = 0 \tag{7.6}$$

Let ν be the angle between the positive x-direction and the positive a-direction. Then

$$\frac{d}{dz} = e^{-i\nu}\left(\frac{\partial}{\partial a} - i\frac{\partial}{\partial b}\right) \tag{7.7}$$

From (7.6) we obtain

$$u_{aa}\cdot N = \kappa_1 W \qquad \text{and} \qquad u_{ab}\cdot N = 0 \tag{7.8}$$

A straightforward calculation using (7.7) and (7.8) shows

$$u_{zz}\cdot N = -2e^{-i2\nu}\kappa_1 W \tag{7.9}$$

which establishes (7.5), and completes the proof that $\triangle k = 0$.

We have now proved that if h satisfies the equation in part (v) of the theorem, then k is harmonic (except possibly at the singularites of h), is tangent to the boundary (except possibly at the boundary singularities of h), and satisfies the kernel equation of $D^2E[u]$, again with possible exceptions at the singularities of h. In order to show that the singularities of h do not pose a serious problem, we define

$$H(z) = \prod_{i=1}^{n}(z - z_i)_i^m \tag{7.10}$$

where the z_i are the branch points of u, and m_i is the order of the branch point z_i. The equation $h_{\bar z} = G$ is then equivalent to $(hH)_{\bar z} = HG$. The right-hand side HG is real-analytic, since

$$\begin{aligned} HG &= (H\phi/W)(u_{\bar z\bar z}\cdot N) \\ &= \frac{\phi|H|^2}{W}\frac{d}{d\bar z}\left(\frac{u_{\bar z}}{\bar H}\right)\cdot N \end{aligned}$$

in view of $u_{\bar z}\cdot N = 0$; and the functions $|H|^2/W$ and $u_{\bar z}/H$ are real-analytic.

In order to prove part (i) of the theorem, we now have only to solve the Riemann-Hilbert system

$$(hH)_{\bar{z}} = HG \qquad \text{in } \bar{D} \tag{7.11}$$

$$\Re(\bar{z}\bar{H}hH) = 0 \qquad \text{on } S^1 \tag{7.12}$$

and verify that the solution is analytic up to the boundary. (We write the solution as hH to conform with the notation of the theorem.) There are two minor problems in solving this system. First, if there are boundary branch points, H can vanish on the boundary, and Riemann-Hilbert systems are usually considered only with a non-vanishing function in the place occuped by $\bar{z}\bar{H}$ in (7.12). Secondly, the boundary regularity. The first difficulty is removed by writing (7.12) in the form

$$\Re(\bar{z}\bar{\sigma}hH) = 0 \qquad \text{where } \sigma = H/|H| \tag{7.13}$$

We then must prove

$$\frac{H}{|H|} \qquad \text{is real-analytic on } S^1 \tag{7.14}$$

It suffices to verify that $H/|H|$ is real-analytic at the z_i, which are the zeroes of the denominator. By a rotation, we may assume without loss of generality that $z_i = 1$. Now let θ be as usual (instead of as above) so that $e^{i\theta}$ parametrizes S^1. With $z = e^{i\theta}$ and $m_i = 2m$ (remember boundary branch points have even order), we have

$$\begin{aligned}
H/|H| &= \frac{z^{2m}(1+O(\theta)}{r^{2m}(1+O(\theta))} \\
&= \frac{z^{2m}}{z^m\bar{z}^m}(1+O(\theta)) \\
&= \frac{(1-e^{i\theta})^{2m}}{(1-e^{i\theta})^m(1-e^{-i\theta})^m}(1+O(\theta)) \\
&= \frac{(1-e^{i\theta})^m}{(1-e^{-i\theta})^m}(1+O(\theta)) \\
&= \left(\frac{e^{-i\theta/2}2i\sin(\theta/2)}{e^{i\theta/2}2i\sin(\theta/2)}\right)^m(1+O(\theta)) \\
&= e^{im\theta}(1+O(\theta))
\end{aligned}$$

which is a real-analytic function of θ. Thus (7.6) is proved.

We shall now construct directly a solution of the Riemann-Hilbert pfoblem (7.11), (7.13), rather than appeal to a reference at once. By first reducing the problem to the existence of a suitable complex-analytic function, we reduce the boundary regularity problem to a simple application of the reflection principle.

Lemma 11 *Let α be a complex-valued function, real-analytic and never zero on S^1. Let σ be a real-valued, real-analytic function on S^1. Let N be the "characteristic" of α; that is, the (algebraic) number of counterclockwise revolutions of the vector $(\Re\,\alpha, \Im\,\alpha)$ as S^1 is traversed once counterclockwise. Then there exists a complex-analytic function ω, analytic in $\bar{D}$, satisfying the boundary condition $\Re\,(\bar{\alpha}\omega) = \sigma$ on S^1, provided that $N \geq 0$.*

Remark. The family of all solutions ω forms a manifold of dimension $2N + 1$.

Proof of lemma. This lemma (including the Remark) is almost a special case of the theorem on p. 236 of [9]. The only difference is that the solutions ω whose existence is there asserted are only guaranteed to be continuous in $\bar{D}$ and C^1 in D. Of course, the interior analyticity of solutions of the Cauchy-Riemann equations is classical; but it remains for us to prove the boundary analyticity. As is usual with boundary regularity theorems, we can prove a local boundary regularity theorem. Namely, if ω is analytic in a neighborhood $V \cap D$ of a point z_0 on S^1, and satisfies $\Re\,(\bar{\alpha}\omega) = \sigma$ on $S^1 \cap V$, where D is the unit disk and V is a small disk about z_0, then ω is analytic in $W \cap \bar{D}$ for some disk W about z_0. To prove this, we first show that σ and $\bar{\alpha}$ can be extended to complex-analytic functions defined in some neighborhood W of z_0. To do this, let F be a conformal tranformation from D to the upper half plane, with $F(z_0) = 0$. Then σF^{-1} is real-analytic on the x-axis, say $\sigma F^{-1}(x) = \sum a_n x^n$. Then $\sum a_n z^n$ defines a complex-analytic extension of σF^{-1}, and $\sum a_n (F(z))^n$ defines a complex-analytic extension of σ. Similarly for $\bar{\alpha}$. We denote these extensions by the same letters as the original functions. Now $\Re\,(\bar{\alpha}\omega - \sigma) = 0$ on $S^1 \cap W$. Since $\bar{\alpha}\omega - \sigma$ is complex-analytic in $D \cap W$, we can apply the Schwarz reflection principlt. Hence $\bar{\alpha}\omega - \sigma$ is analytic up to the boundary. Since α does not vanish on S^1, ω is also analytic up to the boundary. That completes the proof of the lemma.

We now return to the proof of the theorem. By $\int_0^z G(z, \bar{z})\, d\bar{z}$ we mean V_z/H, where H is as above, a complex-analytic function such that HG is real-analytic, and V is a function such that $\triangle V = HG$, for example,

$$V(z, \bar{z}) = \int_D \frac{1}{|z - \xi|} H(\xi) G(\xi_1, \xi_2)\, d\xi_1\, d\xi_2 \qquad \text{where } \xi = \xi_1 + \xi_2$$

Thus $\int_0^z G(z, \bar{z})\, d\bar{z}$ is some function whose derivative with respect to $\bar{z}$ is G.

We shall show that, in order to produce a tangent vector k as required in part (i) of the theorem, it suffices to find a function A such that

$$\Re\,(\bar{z}A) = -\Re\left(\bar{z}\int_0^{\bar{z}} G(z, \bar{z})\, d\bar{z}\right) \qquad \text{on } S^1 \tag{7.15}$$

and A is meromorphic in $\bar{D}$ with HA analytic in $\bar{D}$. Here H is as defined in (7.10). For suppose we have an A as in (7.15). Then we define

$$h(z, \bar{z}) = \int_0^z G(z, \bar{z})\, d\bar{z} + A(z). \tag{7.16}$$

Then $h_{\bar z} = G$ and

$$\Re(\bar z h) = \Re\left(\bar z \int_0^{\bar z} G(z,\bar z)\, d\bar z\right) + \Re(\bar z A) = 0 \qquad \text{on } S^1$$

by the boundary conditions on A. Define $k = \Re(hu_z) + \phi N$. As we have proved, $\triangle k = 0$ and $k_z u_z = 0$, away from the singularities of h. By (7.15) and the analyticity of GH, hH is real-analytic in $\bar D$. Now, in the vicinity of the branch point z_i, we have $|hu_z| \le |h|cr^m$ for some constant c, and $|hH| = |h|r^m(1+O(r))$, so $|h|r^m = |hH|(1+O(r)) \le C|hH|$ in some neighborhood of z_i. Hence $|hu_z| \le C|hH|$. It follows that k is bounded in the vicinity of z_i. According to Theorem 4, if a function is harmonic and bounded in a punctured disk, then the singularity is removable; hence k extends to a harmonic (vector-valued) function defined in $\bar D$. That completes the reduction of part (i) of the theorem to the problem of finding an A as in (7.15).

Since $B(z,\bar z) = H(z)\bar z \int_0^{\bar z} G(z,\bar z)\, d\bar z$ is real-analytic in $\bar D$, the problem of finding A as in (7.15) is a special case of the following problem:

> Let B be real-analytic in $\bar D$. Find a function A meromorphic in $\bar D$ such that HA is complex-analytic and $\Re(\bar z A) = \Re(B/H)$ on S^1.

We now show how to solve this problem. Define $\alpha = zH/|H|$. The characteristic of α (defined in Lemma 11) is easily computed to be 1 plus the sum of the m_i for z_i in D plus half the sum of the m_i for z_i on S^1. Note that this number is N such that $2N + 1 = M + 3$, where M is the number of forced Jacobi fields. By (7.14), α is real-analytic and non-vanishing on S^1. Now apply Lemma 11 with $\sigma = \Re(z\bar\alpha B)$ on S^1. The result is an analytic function ω such that $\Re(\bar\alpha\omega) = \Re(z\bar\alpha B)$ on S^1. Now define $A = \omega/H$. Then A is meromorphic and AH is analytic. We now verify that A satisfies the boundary condition in (7.16). Compute

$$\begin{aligned}
\Re(\bar z A) &= \Re(\bar z\omega/H) \\
&= \Re\left(\frac{\bar z\bar H\omega}{|H|^2}\right) \\
&= \Re\left(\frac{\bar\alpha\omega}{|H|}\right) \\
&= \Re\left(\frac{z\bar\alpha B}{|H|}\right) \\
&= \Re\left(\frac{\bar H B}{|H|^2}\right) \\
&= \Re(B/H) \\
&= \Re\left(\bar z \int_0^{\bar z} G(z,\bar z)\, d\bar z\right)
\end{aligned}$$

which is the value required in (7.16). This completes the proof of part (i) of the theorem.

Part (ii) follows immediately from Tromba's Lemma, since the difference of two solutions of $k \cdot N = 0$ and $k_z u_z = 0$ satisfies $k \cdot N = 0$ and $k_z u_z = 0$, so by Tromba's Lemma it is a forced Jacobi or conformal direction. It is interesting to note, however, how this comes out of the above analysis as well, since as we have remarked, the family of possible solutions A has exactly the dimension $2N + 1 = M + 3$.

Now we prove part (v) of the theorem. We have already proved half of it, namely that any solution h of $h_{\bar{z}} = G$ in D and $\Re(\bar{z}h) = 0$ on S^1, such that $k = \Re(hu_z) + \phi N$ is bounded, gives rise to a solution k of $k \cdot N = \phi$ with k in $KerD^2E[u]$. Moreover, we have proved there exists such a solution h_0. Now let k be any tangent vector in $KerD^2E[u]$ with $k \cdot N = \phi$. We must prove $k = \Re(hu_z) + \phi N$ for some h satisfying $h_{\bar{z}} = G$ in D and $Re(\bar{z}h) = 0$ on S^1. Let $k_0 = \Re(h_0 u_z) + \phi N$. By part (ii), we have $k = k_0 + k_1$, where k_1 is forced Jacobi or conformal. Thus $k_1 = \Re(h_z u_z)$, for h_1 meromorphic, have poles of order at most m_i at z_i. Then $k = \Re(hu_z) + \phi N$, where $h = h_0 + h_1$. The boundary condition on h will be satisfied, since h_0 and h_1 separately satisfy it, since k_0 and k_1 are tangent vectors. That proves part (v).

Now part (iii) is immediate. We have established that the map $F : k \mapsto k \cdot N$ is a surjective linear map from $KerD^2E[u]$ to $KerD^2A[u]$, and has a kernel of dimension $M + 3$.

Ad (iv). Write k in the form $\Re(hu_z)$; any tangent vector k with $k \cdot N = 0$ can be put in this form for some $C^{k,\beta}$ function h. We wish to prove $h_{\bar{z}} = 0$, i.e. h is meromorphic, for then k is a forced Jacobi or conformal direction. Take the dot product of equation (7.2) with u_z, remembering $\triangle k = 0$, $\phi = 0$, $u_z^2 = 0$, and $u_{zz}u_z = 0$. We get

$$0 = h_{\bar{z}} u_{\bar{z}\bar{z}} u_z. \tag{7.17}$$

We wish to show that $h_{\bar{z}}$ vanishes identically. In that case u is conformally equivalent to a surface with $u_z = \alpha$ for some complex constant vector α, and $u = \Re(\alpha)x + \Im(\alpha)y$.

We begin by showing that the second factor in (7.17), namely $u_{\bar{z}\bar{z}}u_z$, does not vanish identically in any neighborhood V unless u lies in a plane. Suppose to the contrary that it does vanish in some neighborhood V. Then $(|u_z|^2)_{\bar{z}} = (u_z u_{\bar{z}})_{\bar{z}} = u_z u_{\bar{z}\bar{z}}$ since $u_{z\bar{z}} = 0$. Hence if $u_z u_{\bar{z}\bar{z}}$ is zero, then $|u_z|^2$ is a constant; hence $|u_z|$ is a real constant, say R. That is not enough to conclude that u_z is constant, since u_z is a vector. We use the functions of the Weierstrass representation, namely $f = {}^1u_z - i^2u_z$ and $g = {}^3u_z/f$. We have

$$\begin{aligned} 4W &= f\bar{f}(1 + g\bar{g})^2 \\ 4W_z &= (1 + g\bar{g})\bar{f}(f_z(1 + g\bar{g}) + 2f\bar{g}g_z). \end{aligned}$$

If this vanishes identically then $f_z(1 + g\bar{g}) + 2fgg_z$ vanishes identically. Since u does not lie in a plane, g is not constant, so we can divide by fg_z, obtaining $f_z/(fg_z) = -2\bar{g}/(1 + g\bar{g})$. But the left-hand side is meromorphic, while the right-hand side definitely depends on $\bar{z}$. To prove this rigorously, differentiate the right-hand side with respect to $\bar{z}$. We get a fraction whose numerator is $\bar{g}_{\bar{z}}$. Hence if the right-hand side is meromorphic, $\bar{g}_{\bar{z}}$ is identically zero; but then g

is constant; and by analytic continuation, it is constant not only in V but in the whole unit disk. Since g is the stereographic projection of the unit normal, u lies in a plane, contradiction. Hence the zero set of the real-analytic function $u_{\bar{z}\bar{z}}u_z$ does not contain any neighborhood.

Now consider $h_{\bar{z}}$, which we want to prove is identically zero. Suppose to the contrary that it is nonzero at some point z_0. Since h is at least C^1, $h_{\bar{z}}$ is continuous. Therefore, $h_{\bar{z}}$ is nonzero in some neighborhood V of z_0. Since $u_{\bar{z}\bar{z}}u_z$ does not vanish in V, there is a point z_1 in V where $u_{\bar{z}\bar{z}}u_z$ is nonzero; but then (7.17) yields a contradiction, since both factors on the right are nonzero at z_1. That completes the proof of (iv), and the proof of the theorem.

Bibliography

[1] Axler, S., Bourdon, P., and Ramey, W., *Harmonic Function Theory, second edition*, Springer (2000).

[2] Beeson, M., Some results on finiteness in Plateau's problem, Part I, Preprint No. 286, Sonderforschungsbereich 72, Universität Bonn, 1972. The last half was later published as [3]; the results of the first part are in Chapter 6 of these lectures.

[3] Beeson, M., Some results on finiteness in Plateau's problem, Part I, *Math Zeitschrift* **175** (1980) 103–123.

[4] Böhme, R., and Tromba, A. J., The index theorem for classical minimal surfaces, *Annals of Mathematics* **113** 2, 447-499 (1981).

[5] Courant, R., *Dirichlet's Principle, Conformal Mapping, and Minimal Surfaces*, Interscience (1950), reprinted by Springer-Verlag (1977).

[6] Courant, R., and Hilbert, D., *Methods of Mathematical Physics, volume 1*, Interscience (1962).

[7] Evans, L., *Partial Differential Equations*, AMS (1991).

[8] Garabedian, P. R., *Partial Differential Equations*, Wiley, New York (1964).

[9] Hellwig, G., *Partielle Differential-Gleichungen*, B. G. Teubner Verlaggesellschaft, Stuttgart (1960).

[10] Dierkes, U., Hildebrandt, S., Küster, A., and Wohlrab, O., *Minimal Surfaces I*, Springer-Verlag (1991).

[11] Dierkes, U., Hildebrandt, S., Küster, A., and Wohlrab, O., *Minimal Surfaces II*, Springer-Verlag (1994).

[12] Nitsche, J.C.C., *Lectures on Minimal Surfaces, Vol. 1*, Cambridge Univ. Press (1989).

[13] Osserman, R., *A Survey of Minimal Surfaces*, Dover (1986). The first edition was in 1969, but the Dover edition has a chapter on later developments.

[14] Rado, T., *On the Problem of Plateau*, Springer-Verlag, Berlin (1933). [out of print].

[15] Die Variation des konformen Typs mehrfach zusammenhängender Minimalfl"achen,: Anwending auf die Isoliertheit und Stabilit"at des Plateauproblems, Thesis, Saarbr"ucken (1978).

[16] Titchmarsh, E. C., *The Theory of Functions*, Second edition, Oxford Univ. Press (1939).

[17] Tromba, A. J., On the number of simply connected minimal surfaces spanning a curve, *Memoirs of the AMS* **12**, No. 194 (1977).