A Real-Analytic Jordan Curve Cannot Bound Infinitely Many Relative Minima of Area

Michael Beeson
Department of Computer Science
San Jose State University
San Jose, California 95192
USA
email: beeson@cs.sjsu.edu

August 4, 2005

Abstract

Let Γ be a real-analytic Jordan curve in R^3. Then Γ cannot bound infinitely many disk-type minimal surfaces which provide relative minima of area in the C^n topology. This was already known for absolute minima of area, but not for relative minima, even in the C^0 topology.

1 Introduction

Branch points of minimal surfaces have been of interest ever since Douglas and Rado solved Plateau's problem in the 1930's, because their solution methods proved the existence of minimal surfaces which might in general have branch points. It was not until the seventies that the regularity (lack of branch points) of solutions providing an absolute minimum of area with given real-analytic boundary was proved, and in 1980 the interior regularity of solutions providing a C^1 relative minimum was proved in [3].

Branch points of minimal surfaces not furnishing a relative minimum of area are also of considerable interest, because they are intimately connected with the "finiteness problem". It follows from the work of Tomi and Tromba [18] that if a C^k Jordan curve Γ bounds infinitely many disk-type minimal surfaces, then it bounds a one-parameter family of minimal surfaces, all with the same Dirichlet's integral, which can be continued until it either loops or terminates in a branched minimal surface. In [16], it is shown that a loop of disk-type minimal surfaces contains one which is not a relative minimum of area. Therefore, to prove that Γ does not bound infinitely many relative minima of area, we only have to rule out the possibility that Γ bounds a one-parameter family of minimal surfaces, terminating in a branched minimal surface when the parameter t is zero, and furnishing a relative minimum of area for each positive t. Tomi solved the finiteness problem for absolute minima of area and real-analytic boundaries [16] in this way, since the limit of absolute minima of area is again an absolute minimum, and hence has no branch points.

The finiteness problem for relative minima of area was attacked by this method in [1], where interior branch points arising as the limit of one-parameter families of relative minima were shown not to exist, and in [2], such boundary branch points were eliminated under special conditions on Γ, but not in general; so the finiteness problem for relative minima remained open until now, except for special boundaries. Indeed, in [2] a one-parameter family of minimal surfaces is constructed, bounded not by a Jordan curve but by a straight line, which terminates in a boundary branch point and satisfies the local conditions considered in [2], so it appeared that there was no hope of a local argument (in the neighborhood of the branch point) to solve this problem.

In this paper, we show that this is not the case: the requirement that the boundary curve be real-analytic *and not lie in a plane* in the vicinity of the branch point is quite strong. We first study the case of a single minimal surface with a boundary branch point, bounded in part by a real-analytic arc which does not lie in a plane. There are four integers involved: The order and index of the branch point, and the integers p and q such that the curve Γ satisfies $\Gamma' = (1, t^q + \ldots, t^p + \ldots)$. We are able to derive some necessary relations between these four integers by a local analysis inspired by the proof of Lewy's regularity theorem.

This same Lewy-style analysis can be applied, not just to a single minimal surface, but to the entire one-parameter family of minimal surfaces. Here the Weierstrass function g (the stereographic projection of the unit normal) has poles for $t > 0$, called a_i, which go to zero as t approaches zero. We show that asymptotically we have $a_i = \alpha_i t^\gamma$ for some γ. Focusing attention on the region $z \le Rt^\gamma$, we are able to perform the Lewy-style analysis as a series in t, whose coefficients are functions of $w = t^\gamma/z$, instead of just numbers as in the case of a single surface. This is good, because terms that could cancel out in the case of a single surface cannot cancel out, even if their powers of t agree, when their coefficients do not agree as functions of w. We calculate $N \cdot \Gamma'$, where N is the unit normal and Γ' the tangent vector to the boundary curve Γ. Of course this must be zero, but calculating it directly, using the Weierstrass representation of the surfaces, we reach a contradiction, except in a certain case.

The main line of attack in [1] is an eigenvalue argument, that in the case of an interior branch point, the Gaussian image of u for small positive t must include more than a hemisphere, and hence u cannot be a relative minimum of area.[1] In this paper we first push this argument as far as it can be taken for boundary branch points: for small positive t, the Gaussian image contains m "extra hemispheres", which "pinch off" when t goes to zero. Their contribution of $2\pi m$ to

[1]Readers not familiar with the connection between the second variation of area and an eigenvalue problem should consult [1] or [14].

the Gaussian area is made up, in the limit, by the branch point. The point is that the "extra" Gaussian area of $2\pi m$ is made up of *hemispheres*, not spheres, since otherwise the Gaussian area would include a region with eigenvalue less than 2, and the surface would therefore not furnish a relative minimum of area. This requirement implies certain conditions on the unit normal N–in particular, the image of the boundary of the parameter domain under N "wraps around" the unit sphere m times monotonically (without reversing direction), near the origin in the w-plane for each small positive t. This corresponds intuitively to a helicoid-like surface spiraling around the boundary near the origin. The $2m$ sheets of this "helicoid" would have to somehow twist around from being perpendicular to the boundary curve, to being parallel, intersect the rest of the surface and produce the branch point, with its $2m+1$ sheets. Note that the unit normal on a helicoid alternates directions on successive sheets, so this picture makes sense: two sheets of surface per revolution of N about the Riemann sphere. As we have remarked, this can happen (locally) when the boundary is a straight line, but when the boundary is a real-analytic curve not lying in a plane, we will show that it can't happen unless the unit normal violates the monotonicity condition.[2]

The main line of attack in [2] is the calculation of the eigenfunction corresponding to the first eigenvalue. The starting point of this calculation is the relation between the second variations of area and of Dirichlet's integral; namely, the condition for ϕ to belong to the kernel of the second variation of area is given by $\phi = k \cdot N$, where k belongs to the kernel of the second variation of area.[3] In the case at hand, k can be shown to be a "forced Jacobi direction" associated with the branch point; these are purely tangential, so as a function of the parameter t, the eigenfunction goes to zero. We calculate how the eigenfunction behaves on the w-plane, that is, on a region near the origin that shrinks to zero as t^γ. The leading term in t of this function is a harmonic function in the w-plane which inherits from the eigenfunction the property that it must have only one sign in the upper half plane. Using this property, we are able to derive additional information about the behavior of N near the origin. But the apparent leading term might, it initially seems, sometimes be identically zero. There is a gap in the treatment of this case in [2], but here it is correctly treated. Namely, we show that it is not possible for the normal to oscillate wildly enough in the w-plane to cover the required hemispheres, and at the same time stay perpendicular to the same boundary curve Γ. The Lewy-style analysis is used here to prove the uniform convergence of certain expressions as $t \to 0$, in regions of the w-plane where the unit normal is taking on the "extra hemispheres". We calculate dZ/dY (where X, Y, and Z are the components of the surface u) on the boundary; this can be given in terms of arc length, but it can also be cal-

[2]Whether it can actually happen in any case at all is of course not known. The finiteness question for general disk-type minimal surfaces is completely open; we only deal with relative minima.

[3]Note that ϕ is a scalar, because with area it is enough to consider normal variations given by $u + \phi N$. On the other hand k is a vector since the second variation of Dirichlet's integral is defined as a Frechet derivative in the space of (vectors defining) surfaces u.

culated directly, using the Weierstrass representation. Since the extra Gaussian area has to be made up of hemispheres, we are able to show in a "Monotonicity Theorem" that the unit normal N (restricted to the boundary near the branch point) very nearly traces a great circle on the Riemann sphere several times, for small t. Therefore there are points $eq_1(t)$ and $eq_2(t)$ on the real axis where the normal N crosses the equator, with opposite signs of the second coordinate. We find an expression for dZ/dY at $eq_i(t)$ as a function of t which contradicts the expression in terms of arc length, in the case ("H constant") that could not be treated directly by the Lewy-style calculation of $N \cdot \Gamma'$. In the paper, the elimination of the case "H constant" comes first, so that the computation of $N \cdot \Gamma'$ in the last section is the *coup de grace*.

We note in passing that as of yet, nobody has produced an example of a real-analytic Jordan curve bounding a minimal surface with a boundary branch point, although Gulliver has exhibited in [9] a surface with a C^∞ boundary and a boundary branch point. The experts I have asked all believe that they can exist, but an example is still missing. Also, it is not known whether Gulliver's example is area-minimizing or not.[4]

I am grateful to Daniel Wienholtz, Fritz Tomi, Tony Tromba, and Joel Spruck for fruitful discussions of the subject, and to Wienholtz and Tomi for their careful attention to my earlier work on this problem. None of them, however, has reviewed this manuscript.

2 Geometric bounds on the index

In this section we consider a single minimal surface $u(z)$ parametrized at least in the upper half of the unit disk, with a branch point at the origin, and bounded in part by a real-analytic arc Γ defined on the unit interval, which does not lie in a plane. We suppose that Γ passes through the origin tangent to the X-axis, and that u takes the portion of the real axis near origin onto Γ, with $u(0) = 0$. We still are free to orient the Y and Z axes. We do this in such a way that the normal at the branch point (which is well-defined) points in the positive Z-direction. With Γ oriented in this way, there will be two integers p and q such that Γ has a parametrization in the form

$$\Gamma(\tau) = \begin{bmatrix} \tau \\ \frac{C_1 \tau^{q+1}}{q+1} + O(\tau^{q+2}) \\ \frac{C_2 \tau^{p+1}}{p+1} + O(\tau^{p+2}) \end{bmatrix}$$

for some nonzero real constants C_1 and C_2. We therefore have

$$\Gamma'(\tau) = \begin{bmatrix} 1 \\ C_1 \tau^q + O(\tau^{q+1}) \\ C_2 \tau^p + O(\tau^{p+1}) \end{bmatrix}$$

[4]Wienholtz has used the solution of Björling's problem ([5], page 120) with boundary values $\Gamma(t^3)$ where Γ is a regular parametrization of an analytic arc, and a suitably specified normal, to produce examples of a minimal surface partially bounded by a real-analytic arc which is not a straight line segment, with a branch point on the arc.

We write $u(z) = (X(z), Y(z), Z(z))$. We make use of the Enneper-Weierstrass representation of u (see e.g. [5], p. 108)

$$u(z) = Re \begin{bmatrix} \frac{1}{2}\int f - fg^2 dz \\ \frac{i}{2}\int f + fg^2 dz \\ \int fg dz \end{bmatrix}$$

where f is analytic and g is meromorphic in the upper half-disk.

Definition 1 *The* order *of the branch point is the order of the zero of f. The* index *of the branch point is the order of the zero of g.*

The order of a boundary branch point of a solution of Plateau's problem must be even (since the boundary is taken on monotonically). It is customary to write it as $2m$, and to use the letter k for the index. Thus $f(z) = z^{2m} + O(z^{2m+1})$ and $g(z) = cz^k + O(z^{k+1})$ for some constant c.

Let $\tau(z) = X(z) = \text{Re}\frac{1}{2}\int f - fg^2 dz$. Then on the boundary we have

$$u(z) \quad = \quad \Gamma(\tau(z))$$

This parametrization and function $\tau(z)$ were inspired by Lewy's equation (see [6], p. 38). We will expand everything as power series in z and compare coefficients. The main result we want is this:

Theorem 1 Geometric bounds on the index. *Let u have a boundary branch point of order M and index k on a real-analytic boundary segment Γ parametrized as $\Gamma' = (1, C_1\tau^q + \ldots, C_2\tau^p + \ldots)$. Then $k \le (M+1)p$.*

Remark: For this theorem, the boundary does not have to be a Jordan curve; it just has to be a non-planar arc.

Proof: We have $f(z) = \sum a_i z^i$ and $g(z) = \sum c_i z^i$. [Note: the letters a_i and c_i will be re-used with another meaning later in the paper. This notation is only for this proof.] The first nonzero a_i is a_M, where M is the order of the branch point. If u is monotonic on the boundary, as is the case in Plateau's problem, then M must be even, and we have $M = 2m$; but it is not necessary to assume that in this section, so we continue to write M for the order. Because of the orientation of Γ, we have a_M real. The first nonzero c_i is c_k.

If not all the coefficients of g are pure imaginary, define δ to be the least integer such that $c_{k+\delta}$ has a nonzero real part. If not all the coefficients of f are real, we define ν to be the least integer such that $a_{M+\nu}$ has a nonzero imaginary part. It cannot be that the coefficients of g are all imaginary and the coefficients of f are all real, for in that case the coefficients of fg^2 would be all real, and so $dY/dx = -\frac{1}{2}\text{Im}(f + fg^2)$ would be identically zero on the real axis, so Y would be constant on the real axis and Γ would lie in a plane. Since a_M is real, we have $\nu > 0$ if ν is defined, but we may have $\delta = 0$ if c_k has a nonzero real part.

We have

$$\begin{aligned}
\tau(z) &= X(z) \\
&= \mathrm{Re}\frac{1}{2}\int f - fg^2 dz \\
&= \frac{a_M}{2(M+1)}\mathrm{Re}\ z^{M+1} + O(z^{M+2})
\end{aligned}$$

Plugging this into the equation for Γ' above, we have for real z:

$$\Gamma'(\tau(z)) = \begin{bmatrix} 1 \\ C_3 z^{(M+1)q} + O(z^{(M+1)q+1}) \\ C_4 z^{(M+1)p} + O(z^{(M+1)p+1}) \end{bmatrix}$$

where

$$\begin{aligned}
C_3 &= \frac{C_1 a_M^q}{2^q (M+1)^q} \\
C_4 &= \frac{C_2 a_M^p}{2^p (M+1)^p}
\end{aligned}$$

We have $u(z) = \Gamma(\tau(z))$ for real z. To emphasize that we are considering real z we write x instead of z. Differentiating with respect to x we have $du/dx = \Gamma'(\tau(x))\tau'(x)$. Since $\tau(z) = X(z)$ we have

$$\frac{du}{dx} = \Gamma'(X(x))\frac{dX}{dx}$$

Writing out the third component of this vector equation, we have (on the real axis)

$$\begin{aligned}
\frac{dZ}{dx} &= (C_4 x^{(M+1)p} + \ldots) dX/dx \\
&= (C_4 x^{(M+1)p} + \ldots)(\frac{a_M x^M}{2} + O(x^{M+1})) \\
&= \frac{C_2 a_M}{2} x^{(M+1)p+M} + \ldots
\end{aligned}$$

The point here is the exponent: the leading term in dZ/dx starts with $x^{(M+1)p+M}$. On the other hand, using the Weierstrass representation directly,

$$\frac{dZ}{dx} = \mathrm{Re}\ fg$$

The real part of fg begins either with the term in $z^{M+k+\delta}$, or with the term $z^{M+k+\nu}$, depending on which is smaller, δ or ν; or conceivably, these two terms could cancel and the first term would be an even higher power of z. If one of δ or ν is not defined because the coefficients of f are all real or those of g

are all imaginary, then the real part of fg begins with whichever term does exist. Comparing these possibilities for the exponent of the leading term to the one derived above, we have proved that one of the following alternatives holds: Either

$$(M+1)p = k + \nu$$

or

$$(M+1)p = k + \delta$$

or

$$\nu = \delta \text{ and } k + \nu < (M+1)p.$$

If all the coefficients of f are real, so ν is not defined, only the second alternative is possible, while if all the coefficients of g are imaginary, only the first alternative is possible. In all three cases, we have the desired conclusion $k \leq (M+1)p$.

Remark: Wienholtz has used the solution to Björling's problem (see [5], p. 120) to construct a minimal surface u bounded by a portion of a real-analytic arc Γ with p and M given arbitrarily, and k taking any value satisfying the bound of the theorem, including $p(M+1)$. If k is taken larger than the bound permits, the construction does not work, and in fact Wienholtz has used the equations for Björling's problem to give another proof of Theorem 1. Wienholtz's construction shows that c_k can sometimes have a real part, so that $\delta = 0$.

Corollary 1 *If a real-analytic Jordan arc has nonzero curvature and torsion at a certain point, then a boundary branch point at that point satisfies* $k \leq 2(M+1)$.

Proof: Nonzero curvature at a point implies that the smaller of p and q is 1; and nonzero torsion implies $|p-q| = 1$. This can be proved by starting from the standard formula for torsion (see for example [11], p. 38). We omit this elementary computation, as the result is not used for our main theorem.

3 On the zero set of real-analytic functions

Several times in this paper we have reason to consider the zero set of a function that is analytic in two variables (one real and one complex). The following lemma applies to this situation.

Lemma 1 *Let A be an open subset of $R \times C$ containing the point $(0,p)$. Let f be an analytic function from A to R^k for some k. Suppose $f(0,p) = 0$ and that for each sufficiently small $t > 0$, $f(t,\cdot)$ is not constant. Then there exists a neighborhood V of $(0,p)$ and finitely many functions $c_i(t)$ such that $f(t,c_i(t)) = 0$, every zero of f in V has the form $(t,c_i(t))$, and the $c_i(t)$ are analytic in some rational power of t, with $c_i(0) = p$.*

Proof. A classical theorem tells us that the zero set S is locally analytically triangulable. That is, there exist a neighborhood V of $(0,p)$ and finitely many closed simplexes and real-analytic homeomorphisms from those simplexes into S such that $V \cap S$ is contained in the range of those those simplexes. Six references for this theorem are given in [1], where the theorem is discussed on page 117. The theorem refers to *closed* simplexes, which means that the homeomorphisms in the conclusion are analytic even at the endpoints of intervals or boundary points of 2-simplexes.

If f does not depend on t at all, the theorem is trivial. Otherwise, if f is divisible by some power of t, we can divide that power out without changing the zero set of f, so we can assume that $f(t, \cdot)$ is not constant for any sufficiently small t, including $t = 0$. Then the dimension of S is one, and S is composed of finitely many analytic 1-simplexes, i.e. analytic paths given by $t = d_i(\tau)$, $w = c_i(\tau)$, where τ is the parameter in which c_i is analytic. If any of these simplexes do not pass through the $(0, p)$, decrease V to exclude them. The ones that do pass through the $(0, p)$ should be divided into two simplexes, each ending at $(0, p)$. Then we can assume that $\tau = 0$ corresponds to $(0, p)$, i.e. $d_i(0) = 0$ and $c_i(0) = p$. We have

$$f(d_i(\tau), c_i(\tau)) = 0.$$

Since d_i is analytic and not constant, it has only finitely many critical points. Decrease the size of V if necessary to exclude all nonzero critical points of d_i. Then either d_i is increasing in some interval $[0, b]$ or decreasing in some interval $[0, b]$. In either case it has an inverse function $\phi = d_i^{-1}$, so $d_i(\phi(t)) = t$. If the leading term of d_i is τ^n, then ϕ is real-analytic in $t^{1/n}$. We can parametrize the zero set of f by $\tilde{c}_i(t) = c_i(\phi(t))$, which is real-analytic in $t^{1/n}$. That completes the proof.

Remark: In applications it will usually be possible to replace the original parameter t by $t^{1/n}$, enabling us to assume that the zero set is analytic in t.

4 Eigenvalues of coverings of the sphere

In this section we present a result which will be applied to the Gauss map of a one-parameter family of minimal surfaces. However, we present it here as a result about conformal mappings from a plane domain to the Riemann sphere. We take the "Riemann sphere" to be the unit sphere $\{(x, y, z) : x^2+y^2+z^2 = 1\}$. Stereographic projection **St** is defined by

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

and its inverse is given by

$$\mathbf{St}^{-1}(z) = \frac{1}{1 + |z|^2} \begin{bmatrix} 2\mathrm{Re}\ z \\ 2\mathrm{Im}\ z \\ |z|^2 - 1 \end{bmatrix}.$$

Thus the equator projects onto the unit circle, and the "north pole" $(0, 0, 1)$ projects onto ∞, while the southern hemisphere projects onto the unit disk, with the south pole going to origin. The picture then has the plane passing through the equator of the sphere. We follow [15], p. 46, in these details, and we mention them because some other authors use a sphere of radius 1/2, with the picture having the sphere entirely above the plane, tangent to the plane where the south pole of the sphere touches the origin of the plane.

Let D^+ be the upper half of the unit disk. We will consider a one-parameter family of analytic mappings from D^+ to the Riemann sphere. (These arise in our work as the Gauss maps of a one-parameter family of minimal surfaces, but here we consider a more abstract situation.) We write t for the parameter and $N(t, z)$ for the value of the mapping. For brevity we often omit the explicit t-dependence and write $N(z)$, or $N^t(z)$. Thus N^t is the map $N(t, \cdot)$ from D^+ to the Riemann sphere, for a fixed t. We define $g = \mathbf{St} \circ N$. The map N is thus conformal except at the zeros of g' and poles of g of order more than one. (These are sometimes known as *ramification points* of g.)

We will suppose that for sufficiently small t, the map g is a quotient of functions B/A, where B and A are real-analytic jointly in t and z, and that when $t = 0$, A has a zero of order m, and B has a zero of order $m + k$, for some positive integers m and k. (In our work, these arise from a boundary branch point of order $2m$ and index k, but here it is not necessary to be so specific.)

By Lemma 1, near the origin there exist paths $a_i(t)$ and $b_i(t)$ describing the poles and zeroes of g, respectively, so that each a_i and b_i is analytic in a rational power of t. Let s be the greatest common divisor of the denominators of these rational powers, and replace t by $t^{1/s}$. Then a_i and b_i will be analytic in t. We suppose that the parameter t has been chosen such that a_i and b_i are analytic in t. We also may suppose without loss of generality that for all sufficiently small positive t, we have $a_i(t) \neq b_j(t)$ for all i, j, since otherwise by analyticity we would have, for some i and j, $a_i(t) = b_j(t)$ for all sufficiently small t, and then the factor $z - a_i$ could be cancelled out of both A and B. The partial derivatives of a quotient of real analytic functions are again quotients of real-analytic functions, so the structure of the set of critical points of the real and imaginary parts of such a function is also known: it is the union of a finite set of real-analytic arcs.

Let $A(t)$ be the area of the spherical image of D^+ under N, counting multiple coverings. Specifically, we have

$$\begin{aligned} A(t) &= \int\int_{D^+} \frac{1}{2}|\nabla N|^2 \\ &= \int\int_{D^+} \left[\frac{4|g'|}{(1+|g|^2)^2}\right]^2 dxdy \end{aligned}$$

Associated with a map N defined on a region Ω in the plane, there is a natural eigenvalue problem:[5]

$$\triangle\phi + \frac{1}{2}\lambda|\nabla N|^2\phi = 0 \qquad \text{in } \Omega$$

[5] As is discussed in [1] and [14], p. 103, among many other sources, when N is the Gauss map of a minimal surface, this eigenvalue problem is connected to the second variation of area. This does not concern us directly in this section.

$$\phi = 0 \qquad \text{on } \partial\Omega$$

We denote the least eigenvalue λ of this problem by $\lambda_{N,\Omega}$, or simply by λ_Ω when N is clear from the context. (In this work, we have no occasion to consider eigenvalues other than the least one.) The factor (1/2) arises because the area element on the sphere is $(1/2)|\nabla N|^2 dx\, dy$. The least eigenvalue is well-known to be equal to the infimum of the Rayleigh quotient

$$R[\phi] = \frac{\int\int_\Omega |\nabla \phi|^2 dxdy}{\int\int_\Omega (1/2)|\nabla N|^2 \phi^2 dxdy}.$$

When we speak of the least eigenvalue λ_Ω of a region Ω on the Riemann sphere, we mean the following: Let Δ be the stereographic projection of Ω and N the inverse of stereographic projection. Then $\lambda_\Omega := \lambda_{N,\Delta}$. The eigenvalue problem $\triangle\phi - \frac{1}{2}\lambda|\nabla N|^2\phi$ on Δ is equivalent to the problem $\triangle\phi = \lambda\phi$ on Ω, where now $\triangle$ is the Laplace-Beltrami operator on the sphere. If Ω contains the north pole, we should use stereographic projection from some point not contained in Ω. We do not need to discuss the case when Ω is the entire sphere.

Example. We compute the least eigenvalue when $N(\Omega)$ is a hemisphere. In this case the eigenfunction in the lower hemisphere is minus the Z-component of N. For example with Ω equal to the unit disk and $g(z) = z$, we have $N(z)$ the inverse of stereographic projection. With $|z| = r$ and $z = x + iy$ we have

$$N(z) = \frac{1}{1+r^2}\begin{bmatrix} 2x \\ 2y \\ r^2 - 1 \end{bmatrix}.$$

The eigenfunction ϕ is given, with $|z| = r$, by

$$\phi(z) = \frac{1-r^2}{1+r^2}.$$

A few lines of elementary computations (or a couple of commands to a computer algebra program) show that

$$\begin{aligned} \triangle\phi &= \phi_{rr} + \frac{1}{r}\phi_r \\ &= \frac{8(r^2-1)}{(1+r^2)^3} \end{aligned}$$

and

$$\begin{aligned} \frac{1}{2}|\nabla N|^2 &= \frac{1}{2}(N_x^2 + N_y^2) = \frac{4r}{(1+r^2)^2} \\ \frac{1}{2}|\nabla N|^2\phi &= \frac{4(1-r^2)}{(1+r^2)^3} \end{aligned} \qquad (1)$$

Hence, $\triangle\phi + |\nabla N|^2\phi = 0$, which means the eigenvalue of a hemisphere is 2.

Lemma 2 *Let Ω be a connected open set on the sphere with least eigenvalue λ_Ω. Suppose Δ is a region (open set) in the plane and $\Omega \subseteq N(\Delta)$ and $N(\partial\Delta) \cap \Omega = \phi$. Then $\lambda_{N,\Delta} \le \lambda_\Omega$, and strict inequality holds if $N^{-1}\partial\Omega$ contains an interior point of Δ.*

Proof. Let ϕ be the least eigenfunction of Ω and define ψ on Δ by setting $\psi(z) = \phi(N(z))$ if $N(z) \in \Omega$ else $\psi(z) = 0$. Then ψ is admissible in the Rayleigh quotient for Δ, since ϕ is zero outside Ω and $N(\partial\Delta) \cap \Omega = \phi$. The Rayleigh quotient in question is

$$\frac{\int\int_\Delta |\nabla\psi|^2 dxdy}{\int\int_\Delta \frac{1}{2}|\nabla N|^2\psi^2}$$

Since $\Omega \subseteq N(\Delta)$, on the support of ψ, N is a covering map, except at the points of ramification of N, which are isolated. Since Ω is connected and $N(\partial\Delta) \cap \Omega = \phi$, the number of sheets over each non-ramification point is the same. Hence each of the integrals in the Rayleigh quotient is the number of sheets times the corresponding integral on the Riemann sphere, with ϕ in place of ψ. That is, the Rayleigh quotient for ψ on Δ equals the Rayleigh quotient for ϕ on Ω, which is λ_Ω. Since λ_Δ is the minimum of such Rayleigh quotients, $\lambda_\Delta \le \lambda_\Omega$. Now suppose there is an interior point p of Δ in $N^{-1}(\partial\Omega)$. By Hopf's lemma, $\nabla\phi$ is never zero at a point on $\partial\Omega$, and by analyticity, ∇N is zero only at isolated points; so there is a point q near p which is still on $N^{-1}(\partial\Omega)$ at which ∇N is not zero and $\nabla\psi$ is not zero. Hence near q, the set $N^{-1}(\partial\Omega)$ is a smooth arc, and $\nabla\psi$ is zero on one side of it and bounded away from zero on the other side. Hence we can "smooth out the edge" near q to obtain a function ψ' which is admissible for the Rayleigh quotient and has smaller Rayleigh quotient than ψ. Hence $\lambda_\Delta < \lambda_\Omega$. This completes the proof of the lemma.

Definition 2 *A* slotted sphere *is a sphere minus an arc of a great circle (of angle strictly less than 2π.) The* slot *is the portion of the great circle that has been removed. The* bridge width *of a slotted sphere is the length of the portion of this great circle that has not been removed, that is, 2π minus the length of the slot. A* fat slotted sphere *is formed by deleting from a slotted sphere, all points at distance ϵ or less from the slot, for some ϵ less than half the bridge width of the slotted sphere. We also require ϵ to be less than half the radius of the sphere. The bridge width of the fat slotted sphere is the bridge width of the slotted sphere minus 2ϵ. The* slot width *of a fat slotted sphere is 2ϵ. The slot width of a slotted sphere is zero.*

Lemma 3 (Continuity of eigenvalues) *Let Ω_n be a nested sequence of regions (on the sphere or in the plane) converging to a region Ω, in the sense that $\Omega_n \subset \Omega_m$ when $n < m$ and Ω is the union of the Ω_n. Let λ_n be the least eigenvalue of Ω_n and ϕ_n the corresponding eigenfunction, normalized by $\int_{\Omega_n} \phi_n^2 = 1$. Then the sequence λ_n converges to the least eigenfunction λ of Ω, and ϕ_n converges (in H_0^1 and C^0 norms) to the corresponding eigenfunction ϕ.*

Remark. The continuity of the least eigenvalue with respect to the domain is well-known. It is, however, not easy to find a citation to this "well-known" result that covers the case of a fat slotted sphere approaching a slotted sphere (as the slot width goes to zero), and one might worry about it because the behavior of the eigenfunction at the endpoint of the slot is not like its behavior at smooth boundary points.[6] It may be a special case of one of the theorems in [10], but we can give a short direct proof.

Proof. Extend ϕ_n to be zero outside Ω_n. Since $\Omega_n \subset \Omega$, ϕ_n so extended belongs to the Sobolev space $H^1_0(\Omega)$. Let any subsequence of ϕ_n be given; then by compactness of $H^1_0(\Omega)$, some further subsequence converges. Let ϕ be the limit function. Then $\phi \geq 0$ in Ω. By the Sobolev embedding theorem (see e.g. [8], Corollary 7.11, p. 151), we have convergence in the C^0 sense as well as the H^1_0 sense, as well as a uniform bound on the ϕ_n. Since the regions Ω_n increase with n, the eigenvalues λ_n decrease, and since they are all positive, they converge to a limit α. For n indexing the convergent subsequence we have

$$\begin{aligned}\lim \triangle \phi_n &= -\lim \lambda_n \phi_n \\ &= -\alpha\phi\end{aligned}$$

The limit on the left exists because those on the right exist. Then ϕ is a weak solution of the eigenvalue problem $\triangle\phi + \alpha\phi = 0$, as is seen by passing to the limit in $\int_\Omega \chi \triangle \phi_n$, where χ is a test function. Then by elliptic regularity theory, ϕ is smooth in the interior and hence is an actual solution of $\triangle\phi + \alpha\phi$, i.e. an eigenfunction for eigenvalue α. Since ϕ has one sign, α is the least eigenvalue λ for the limit region Ω. That is, λ_n converges to λ, which is the first claim of the lemma. Since the eigenspace of the first eigenvalue is one-dimensional, and we have normalized the ϕ_n to have H^1_0 norm 1, the limit function ϕ is independent of which subsequence of ϕ_n was originally chosen. Therefore the entire sequence ϕ_n converges to ϕ. This completes the proof of the lemma.[7]

Lemma 4 (Eigenvalue of a slotted sphere) *Let Ω be a slotted sphere, or a fat slotted sphere. Then $\lambda_\Omega < 2$, provided the slot width is small enough compared to the bridge width.*

Proof. First consider the case of a slotted sphere. Orient the sphere so the slot is in the YZ-plane, with the bridge near the south pole, as if the slot were made by a horizontal knife descending through the north pole. The function on the sphere whose value is the X-coordinate is an eigenfunction for eigenvalue 2. Since this function changes sign on the slotted sphere, it is not the first eigenfunction. Hence 2 is an eigenvalue, but not the least eigenvalue. Hence $\lambda_\Omega < 2$ for a slotted sphere. Now consider a fat slotted sphere Ω_ϵ of slot width

[6] Asymptically it has the form $\sqrt{r}\sin(\theta/2)$ where r and θ are local polar coordinates with $r = 0$ at the endpoint of the slot and $\theta = 0$ along the slot.

[7] One can also show that $\triangle\phi_n$ converges in L^2, whence ϕ_n converges in H^2_0, whence by Sobolev embedding the ϕ_n converge in $C^{0,1/2}$; this is the best convergence one could expect since the limit has behavior $\sqrt{r}\sin(\theta/2)$ near the endpoint of the slot when Ω is a slotted sphere.

ϵ, but with a fixed bridge width ρ as ϵ varies. The least eigenvalue is a function of both ρ and ϵ; call it $\lambda(\rho,\epsilon)$. We have proved $\lambda(\rho,0) < 2$. The result therefore follows from the previous lemma. This completes the proof of the lemma.

This proof, however, does not easily generalize to ramified coverings of the sphere, so we give a second, more direct and computational argument, which does generalize.

Second Proof. Let Ω be a fat slotted sphere with slot width at most ϵ, or a slotted sphere. Let ρ be the bridge width of Ω. Orient the sphere so that the slot is in the YZ plane and centered at the south pole. Define ϕ on the sphere by $\phi = \min(|X| - \epsilon, 0)$. Then ϕ is zero on $\partial\Omega$ and has its support in Ω. The Rayleigh quotient $R[\phi]$ satisfies $R[\phi] = R[X] + O(\epsilon) = 2 + O(\epsilon)$, as the following calculation shows in detail:

$$\begin{aligned}
R[\phi] &= \frac{\int_\Omega |\nabla\phi|^2\,dA}{\int_\Omega \phi^2\,dA} \\
&= \frac{\int_{S^2} |\nabla X|^2\,dA + 2\pi\epsilon + O(\epsilon^2)}{\int_\Omega (X-\epsilon)^2\,dA} \\
&= \frac{\int_{S^2} |\nabla X|^2\,dA + 2\pi\epsilon + O(\epsilon^2)}{\int_\Omega X^2\,dA - \epsilon\int_\Omega X\,dA + O(\epsilon^2)} \\
&= \frac{\int_{S^2} |\nabla X|^2\,dA + 2\pi\epsilon + O(\epsilon^2)}{\int_{S^2} X^2\,dA - \epsilon\int_\Omega X\,dA + O(\epsilon^2)} \\
&= R[X] + O(\epsilon) \\
&= 2 + O(\epsilon)
\end{aligned}$$

Now we modify ϕ to produce a new function ψ, which agrees with ϕ except near the "bridge". By the "bridge", we mean the portion of the great circle in the XY plane which lies outside the slot. This is an arc of length ρ. Let B be the domain on the sphere consisting of points (X,Y,Z) with $(0,Y,Z)$ on the bridge and $|X| \le \delta$, where we take $\delta = \epsilon^{1/2}$.[8] Outside B, and on ∂B, we will define $\psi = \phi$. Inside B, we define ψ to (1) have the value δ except within δ of the horizontal boundaries of B, and (2) near those boundaries, to not depend on X and have gradient 1. Informally, then, ψ crosses the bridge at height δ, except near the horizontal boundaries of B (the sides of the bridge) where it tapers off at gradient 1, until it meets ϕ, which will be at ∂B. Since ϕ does not exceed δ on B, psi is constant on B except within δ of the horizontal boundaries of B. How does $R[\psi]$ differ from $R[\phi]$? We only need to consider the changes to the numerator and denominator due to the integrals over B. First consider the numerator in $R[\psi]$. The numerator $\int |\nabla\psi|^2\,dA$ has contributions only from the strip of width δ along the horizontal part of ∂B, since $\nabla\psi$ is zero on B except near the horizontal boundaries. The contribution to the numerator is then at

[8] So B has four boundary arcs, two of them with Z and Y constant, the "horizontal" boundaries, and two circular arcs with X constant. Since $\delta > \epsilon$, the region B has both its curved boundaries in the regions where ϕ is nonzero, and it "bridges" those two regions by passing over the no-man's-land $|X| < \epsilon$. This bridge remains of fixed width no matter how small we take ϵ and δ; the length of the bridge, 2δ, thus becomes small compared to its width.

most $4\delta^2$, since there are two strips, one near each boundary, of length (in the X-direction) 2δ and width (in the Y-direction) δ, on which $\nabla\psi$ is $1 + O(\epsilon)$.

Now consider the numerator in $R[\phi]$. Over B, we have $\nabla\phi = 0$ at points with $|X| < \epsilon$, but outside this region, on most of B, $\nabla\phi = 1 + O(\epsilon)$. The area of B where $|X \geq \epsilon$ (and hence $\nabla\phi = 1 + O(\epsilon)$ is approximately $2\rho(\delta - \epsilon)$, where ρ is the bridge width, so the contribution from B to the numerator of $R[\phi]$ is $2\rho\delta(1 + O(\epsilon))$.

We now can estimate the Rayleigh quotient $R[\psi]$ as follows:

$$
\begin{aligned}
R[\psi] &= \frac{\int_\Omega |\nabla\psi|^2 \, dA}{\int_\Omega \psi^2 \, dA} \\
&\leq \frac{\int_\Omega |\nabla\phi|^2 \, dA + (4\delta^2 - 2\rho\delta)(1 + O(\epsilon))}{\int_\Omega \psi^2 \, dA} \\
&\leq \frac{\int_\Omega |\nabla\phi|^2 \, dA + (4\delta^2 - 2\rho\delta)(1 + O(\epsilon))}{\int_\Omega \phi^2 \, dA} \qquad \text{since } |\psi| \geq |\phi| \text{ on } B \\
&= R[\phi] - \frac{2(\rho - 2\delta)\delta}{\int_\Omega \phi^2 \, dA} \\
&= R[\phi] - c\rho\epsilon^{1/2} + O(\epsilon) \qquad \text{since } \delta = \epsilon^{1/2}
\end{aligned}
$$

for a positive constant c. Since $R[\phi] = 2 + O(\epsilon)$ we have

$$
R[\psi] \leq 2 - c\rho\epsilon^{1/2} + O(\epsilon)
$$

which is less than 2 for small ϵ. That completes the second proof of the lemma.

More generally, we consider the case of a "ramified covering" $N : D^+ \to S^2$ as described above, in which the image of $[0,1]$ lies on the YZ plane. The ramification points are the zeroes of ∇N. They may occur on the boundary or in the interior of D^+. A ramification point on the boundary of order 2 looks like the end of a slot in a slotted sphere. Ramification points in the interior or ramification points of order more than two on the boundary result in a multi-sheeted covering. If there is a ramification point on the boundary, then the pre-image of the YZ plane consists of a finite number of analytic arcs emanating from the ramification point. These must either lie on the boundary or go through the interior to another boundary ramification point; their images are segments of the great circle C in which S^2 meets the YZ plane. The entire pre-image of the great circle C is composed of such arcs. The ones passing through the interior of $D^+$are "bridge arcs" and the minimum of the lengths of the images of these arcs is the "bridge width" of the covering N. If there are no such interior arcs, then the pre-image of C lies entirely on the boundary, and hence there are no boundary ramification points, so the map N is one-to-one on the boundary. It follows that N maps the entire boundary onto C and the interior of D^+ onto one hemisphere.

Lemma 5 (Eigenvalue of a ramified slotted sphere) *Let N be an analytic map from D^+ to the Riemann sphere such that $N(\partial D^+)$ is contained in the YZ*

plane. Then either $N(D^+)$ lies in one hemisphere and N has no boundary ramification points, or the least eigenvalue satisfies $\lambda_{\min} < 2$.

Proof. Let $\phi(z) = {}^1N(z)$, the X-component of the covering. Then ϕ is an eigenfunction for eigenvalue 2. If 2 is the least eigenvalue, then ϕ has one sign, so $N(D^+)$ lies in one hemisphere. In that case N has no boundary ramification points. That completes the proof.

Definition 3 *A* ramified fat slotted sphere *of slot width ϵ is an analytic map N from D^+ to the Riemann sphere such that $N(\partial D^+)$ lies within $\epsilon/2$ of the YZ plane. A* bridge *of radius ρ is given by an open set $U \subset D^+$ on which ∇N is nonzero and such that $N(U)$ contains a neighborhood on the Riemann sphere of a point lying on the YZ-plane of radius $\rho > \epsilon/2$.*

Lemma 6 (Eigenvalue of a ramified fat slotted sphere) *Let ρ and ϵ be positive numbers with $\rho > \epsilon/2$. Let N be a ramified fat slotted sphere containing a bridge of radius ρ. Then if ϵ is sufficiently small compared to ρ, we have $\lambda_{\min} < 2$.*

Remarks: The situation seems similar to the unramified case considered above, but the proof by continuity of the eigenvalue is difficult to generalize. First, the limit covering is not *a priori* defined, and it requires some effort to prove that there does exist a Riemann surface which in a suitable sense is the limit of N as ϵ goes to zero, provided N depends real-analytically on ϵ. The limit Riemann surface may well be disconnected– several multi-sheeted coverings of the Riemann sphere may "break off" from the rest of the covering as ϵ approaches zero. Secondly, the eigenfunctions of course are in $H^1(D^+)$, but considered as functions on the Riemann sphere, they may not be in H^1 in the vicinity of a ramification point of order more than 2. Therefore, it is easier to generalize the direct construction of a function ψ to put into the Rayleigh quotient, as in the second proof of Lemma 4 above.
Proof. Let U be the bridge of radius ρ given by hypothesis. We may suppose that the sphere has been rotated so that the bridge is centered at a point p such that $N(p)$ is the north pole. We may also suppose that $N(U)$ is exactly a neighborhood of radius ρ of the north pole, since we can replace U by the preimage under N of this neighborhood. Then the "rectangular" region on the Riemann sphere V defined by $|Y| \le \rho/2$ and $|X| \le \delta$ lies within $N(U)$, so on this region N is invertible and has no ramification points, since ∇N is nonzero on a bridge. As in the second proof of Lemma 4, we let $\phi = |X| - \epsilon$, and construct ψ by modifying ϕ in $N^{-1}(V) \cap U$. The only difference is that now we regard ϕ and ψ as defined in D^+ rather than on the Riemann sphere. Because there are no ramification points in the bridge, there is a unique sheet of the covering N defined over U and the estimates of the Rayleigh quotient $R[\psi]$ given in the proof of Lemma 4 are valid in this case too.

Theorem 2 [Hemispheric covering theorem] *Let D be a disk in R^2 centered at the origin, $\bar{D}$ its closure, and D^+ the intersection of D with the upper half* *plane. Suppose that* $N : [0,b] \times \bar{D} \to S^2$ *for some* $b > 0$*. Let* g *be* N *followed by stereographic projection of the Riemann sphere from the north pole. Suppose that* g *is a quotient of functions jointly analytic in* t *and* w *for* t *in some interval including* $t = 0$ *and* $w \in \bar{D}$*. By* N^t *we mean the function* $N(t, \cdot) : D \to S^2$*. Suppose that* g *has exactly* m *poles in* D^+*, for sufficiently small* t*, which we call* $a_i(t)$*, and* g *has exactly* $m+k$ *zeroes, which we call* $b_i(t)$*, and for* $t > 0$ *we have* $a_i(t) \neq b_j(t)$*. Suppose also that* $N(t) = (0,0,-1) + O(t)$ *on the boundary of* D*.*

Suppose that $N(w)$ *always lies close to the plane* $X = 0$ *when* w *is on the boundary of* D^+*, in the sense that* ${}^1N(w) = O(t)$*. Choose* ϵ *depending on* t *such that* $|{}^1N(z)| < \epsilon = O(t)$*; let* S_+ *be the subset of* S^2 *with* X*-coordinate more than* ϵ *and* S_- *the subset with* X*-coordinate less than* $-\epsilon$*.*

Let $\lambda_{\min}(t)$ *be the least eigenvalue of* N^t *on* D^+*. Suppose* $\lambda_{\min}(t) \geq 2$ *for* $t > 0$*. Then:*

(i) Let the ramification points of N *(where* $\nabla N = 0$*) in* D *be called* $c_i(t)$*. There are finitely many of these and they depend analytically on a rational power of* t*.*

(ii) For sufficiently small positive t*, the map* N *is, apart from its finitely many ramification points, an* m*-sheeted covering of one of the two near-hemispheres* S_+ *or* S_-*, and does not cover the other near-hemisphere at all.*

(iii) Let $A(t)$ *be the area of* $N(D^+)$*, counting multiplicities (the "Gaussian area"). Then for sufficiently small positive* t *we have* $A(t) = A(0) + 2m\pi + O(t)$*.*

Proof. Let g be the stereographic projection of N; g is meromorphic in D. The *ramification points* of N are the zeroes of g' and the double poles, if any, of g. We are interested only in the ramification points that lie in D^+ and go to zero as $t \to 0$. Since g is meromorphic, the same is true of g'. The ramification points (in D^+) are thus the zeroes of g' or the poles of g', so in either case the ramification points are the zeroes of some analytic functions of w and t. It follows from Lemma 1 that the ramification points are finite in number and depend real-analytically on a rational power of t. If necessary, we could replace the parameter t by a new parameter, a suitable rational power of the old one, to ensure that the ramification points are actually real-analytic in t.

Enumerate those ramification points of N that, for sufficiently small t, stay inside D by real-analytic functions $w = c_i(t)$. Consider the points $\tilde{q}_i(t) = N(c_i(t))$ on the Riemann sphere. Since g, the stereographic projection of N, is complex analytic in w and real-analytic in t, and $c_i(t)$ is real-analytic in t, the $\tilde{q}_i(t)$ depend real-analytically on a rational power of t, and hence have limits q_i as $t \to 0$. This completes the proof of part(i).

Working towards part (ii), we will first prove the following claim: If $N(D^+)$ meets S_- at all, then N is a covering map of S_- away from the ramification points. Spelling the claim out more explicitly: N is a covering map from $N^{-1}(S_-) - \cup_i\{c_i(t)\}$ to $S_- - \cup_i\{q_i(t)\}$. Similarly for S_+ in place of S_-. The definition of being a covering map is that the map is surjective, a local homeomorphism, and each function value has only finitely many pre-images. Suppose $p \in D$ and $N(p) = q \in S_-$, and suppose p is not a ramification point, so $\nabla N(p) \neq 0$. Then (to reduce matters to the complex plane) let g be the composition of N followed by stereographic projection, so g is meromorphic. If q happens to be the north pole then for this argument take the stereographic projection from another point rather than from the north pole, so that g is analytic at p. Let Q be the stereographic projection of q. Then the implicit function theorem says we can find a one-to-one analytic function η on some neighborhood of p such that $g(\eta(z)) = Q + z$ and $\eta(0) = p$. Composing η with inverse stereographic projection, we have the desired local homeomorphism. The fact that there are only a finite number of pre-images of a given point Q follows from Lemma 1, since $g(w) - Q$ is analytic jointly in t and w. This number, the number of "sheets" over Q, does not change with t except when one of the pre-images meets the boundary of $N^{-1}(S_-)$. For sufficiently small t, that does not happen since $|{}^1N(w)| < \epsilon(t)$ on the boundary. That completes the proof that N is a covering map of S_- away from the ramification points.

We now discuss "path lifting" (for a fixed value of t). Let π be any path on the unit sphere that does not pass through the $q_i(t)$, with $\pi : [0,1] \to S^2$. Let q be a point on π, and let p be a point in D such that $N(p) = q$. To "lift the path" π means to construct a path $\eta : [0,1] \to R^2$ such that $N(\eta(s)) = \pi(s)$ and $\eta(0) = p$. The point p is not one of the $c_i(t)$ by hypothesis, so $\nabla N(p) \neq 0$. Hence by the implicit function theorem, we can solve the equation $N(\eta(s)) = \pi(s)$, in a neighborhood of $s = 0$ in $[0,1]$. The size of this neighborhood is bounded below in terms of a lower bound on ∇N. Let Δ be any open set in D whose closure does not include any of the $c_i(t)$. Then there is a lower bound for ∇N on Δ, and we can extend η to larger and larger values of s until we reach either $s = 1$ or the boundary of Δ. Now let Δ_n be a sequence of open sets in D, obtained by cutting out from D a disk of radius $1/n$ about each $c_i(t)$. Then for each n let b_n be the supremum of numbers b in $[0,1]$ such that η can be defined mapping $[0,b] \to \Delta_n$ and satisfying $N(\eta(s)) = \pi(s)$. If b_n is not 1, then it is because the path η runs into the boundary of Δ_n; hence the path can be continued further using Δ_{n+1}, so the b_n form a monotone increasing sequence of points, with $\eta(b_n)$ at a distance $1/n$ from the set of ramification points. Since the sequence b_n is bounded and monotone, it has a limit b. Since there are finitely many ramification points, some subsequence of the $\eta(b_n)$ converges to a ramification point $c_i(t)$. Re-index the sequence b_n to include only this subsequence. Then

$$\begin{aligned} \pi(b) &= \lim_{n\to\infty} \pi(b_n) \\ &= \lim_{n\to\infty} N(\eta(b_n)) \\ &= N(\lim_{n\to\infty} \eta(b_n)) \\ &= N(c_i(t)) \\ &= q_i \end{aligned}$$

which contradicts the assumption that π does not meet the q_i. Hence some $b_n = 1$ and we can lift the path π to $\eta : [0,1] \to R^2$.

We now discuss the dependence of η on t. We would like to find a function of two variables $\eta(t,s)$ such that $N^t(\eta(t,s)) = \pi(s)$. To apply the inverse function theorem we would need to know that not only $\nabla N \neq 0$ but also $N_t \neq 0$. The former are the ramification points $c_i(t)$, which depend analytically on t. The latter we denote by $\zeta_i(t)$. These points also are the zeroes of some function analytic in t and w (namely the numerator of g_t) so they are given by finitely many functions $\zeta_i(t)$ analytic in a rational power of t (unless g_t is identically zero, in which case the t-dependence of η is not an issue). We assume π is a real-analytic function from $[0,1]$ into S^2 such that, for t sufficiently small, π avoids the ramification points $c_i(t)$ and the points $N(\zeta_i(t))$. (π itself does not depend on t.) The argument above, carried out for t fixed, applies under these assumptions on π to define η real-analytically in any interval of *positive* sufficiently small values of t. We cannot prove that η is analytic at $t = 0$ since N_t may go to zero there. Indeed, the size of the neighborhood of (t, s) in which $\eta(t, s)$ is analytic may shrink as $t \to 0$ and we have not proved that η is analytic, or even C^1, or for that matter even continuous, at $t = 0$ for any s.

Now we will prove (ii). Assume, for proof by contradiction, that $N(D^+)$ meets both S_- and S_+. Since these three sets are all open, and there are only finitely many q_i and $N(\zeta_i(0))$, we can choose a point P in $S_- \cap N(D^+)$ and a point V in $S_+ \cap N(D^+)$, distinct from any of the q_i and $N(\zeta_i(0))$. Then for t sufficiently small, P and V are also not equal to any of the $\tilde{q}_i(t)$ or $N(\zeta_i(t))$. Fix a particular t, say t_0, at least that small. Let p and v be points in D^+ such that $N(p) = P$ and $N(v) = V$. Then there is a path $\eta : [0,1] \to D^+$ connecting p and v, i.e. $\eta(0) = p$ and $\eta(1) = v$, which avoids all the $c_i(t)$, $a_i(t)$, and $\zeta_i(t)$ for $t < t_0$. Its image $\pi(t) = N(\eta(t))$ is a path on S^2 between P and V. This path crosses the great circle G at some point T, which is not the north pole since η avoids the $a_i(t)$. Let τ be such that $\pi(\tau) = T$, and let $c = \eta(\tau)$ so $N(c) = T$. As proved above, we can now extend η to smaller positive values of t in such a way that $N(\eta(t, s)) = \pi(s)$. (The points P,Q, and path π do not depend on t.) Let $p(t) = \eta(t, 0)$ and $v(t) = \eta(t, 1)$, so $N(p(t)) = P$ and $N(q(t)) = V$. Let $c(t) = \eta(t, \tau(t))$ so that $N(c(t)) = T$ lies on G.[9]

Let Ω be a neighborhood on S^2 centered at T, disjoint from all the ramification points $\tilde{q}_i(t)$ and all the $N(\zeta_i(t))$ for t sufficiently small. Choose the radius ρ of Ω small enough that the north pole is not in Ω, and P and V are not in Ω. Recall that $\epsilon(t)$ is a number so small that ${}^1N(x) < \epsilon(t)$ for real x on the boundary of D^+, and $\epsilon(t) = O(t)$. According to Lemma 6, there exists an $\epsilon_0 > 0$ such that a ramified fat slotted sphere with bridge width ρ and slot width less than ϵ has least eigenvalue less than 2. Fix t so small that $\epsilon(t) < \epsilon_0$. Let Δ be the component of $N^{-1}(\Omega)$ containing $c(t)$. Then the pair (N, Δ) is a ramified fat slotted sphere of bridge width ρ and slot width less than ϵ_0. By Lemma 6, the least eigenvalue $\lambda_{N,\Delta}$ is less than 2. By the monotonicity of eigenvalues, $\lambda_{N,D^+} < \lambda_{N,\Delta} < 2$. But by hypothesis $\lambda_{N,D^+} \geq 2$. This contradiction shows that P and Q cannot exist with $N(P)$ in S_- and $N(Q)$ in S_+. That completes the proof of part (ii) of theorem.

[9] We cannot choose π first on the sphere, because we have no control over the possible "wiggles" in the image under N of the boundary. We first choose the path η in the parameter domain, for some positive t; then choose π as its image under N; then use path-lifting to extend η to smaller and smaller values of t. This order of path selection is the key to this proof.

Each of the a_i thus contributes approximately one hemisphere to the Gaussian area, i.e. $2\pi + \mathbf{O}(t)$. This proves $A(t) + O(t) \geq 2m\pi$. Now observe that near the north pole, there are exactly m sheets in the covering, since the only pre-images of the north pole are the m values a_i. But the number of sheets over s is constant for s in S_- or s in S_+, because any path in S_- or S_+ that avoids the finitely many ramification points can be lifted to the parameter domain, since the image of the boundary of D^+ lies outside S_- and outside S_+. Hence $A(t) = 2m\pi + O(t)$, proving part (iii) of the theorem.
Remarks. One would like to show that the path on S^2 defined by N restricted to the interval J on the x-axis on the boundary of D^+ approaches, in the C^0 sense, a map that traces out the great circle G exactly m times without reversing its direction. However, as far as I can see, it might have critical points, which for small t might bifurcate, so that for small t this map might reverse its direction for a short distance, then reverse it again; since for small t it can deviate slightly to the sides of G, it can accomplish these reversals and still be analytic in x. It is intuitively appealing to think that N traces out the boundary of the hemispheres in an "essentially monotonic" manner, i.e. monotonic in some sense except for these "small reversals", or perhaps even without small reversals, but I could not prove that. Luckily, we do not need to prove that, and I mention it only to help the reader have a sufficiently complicated mental picture of what might happen. Another thing that I could not prove is that the ramification points stay bounded away from the great circle G as $t \to 0$. You might think that if they did not, you could get a "bridge" from one hemisphere to another because of the branched nature of the the covering at a ramification point. Indeed this argument works if you assume the $a_i(0)$ are all different from all the $b_j(0)$ (because in that case the $a_i(t)$ can be shown to be distinct from the $c_j(t)$. But in general you could have several $a_i(t)$ and $c_j(t)$ converging to the same $c_i(0)$, and that situation is very complicated.

5 Basic Setup and Weierstrass Representation

We suppose we are in the following situation: a real-analytic Jordan curve, not lying in a plane, bounds a one-parameter family of minimal surfaces $u(t, z)$. The surfaces are parametrized by z in the upper half plane, so that on the real line each $u(t, \cdot)$ is a reparametrization of Γ. The surface $u(0, \cdot)$ has a boundary branch point of order $2m$ and index k at the origin. The surfaces $u(t, \cdot)$ for $t > 0$ are immersed. All our arguments will be local; we shall only be concerned with what happens near the origin.[10]

We suppose that Γ passes through the origin tangent to the X-axis. Then since Γ is a Jordan curve, it is not contained in a line. We still are free to orient the Y and Z axes. We do this in such a way that the normal to $u(0, \cdot)$ at the

[10]We do not need to assume that the branch point is a true branch point. We can rule out the existence of a one-parameter family of the type considered without that assumption. This does not, however, provide a new proof of the non-existence of true branch points which are not the terminus of a one-parameter family of relative minima of area.

branch point (which is well-defined) points in the positive Z-direction.

The surfaces $u(t,z)$ have an Enneper-Weierstrass representation

$$u(t,z) = \mathrm{Re}\begin{bmatrix} \frac{1}{2}\int f - fg^2 dz \\ \frac{i}{2}\int f + fg^2 dz \\ \int fg dz \end{bmatrix}$$

The branch points of $u(t,\cdot)$ are the places where f and fg^2 vanish simultaneously. In [1] we were concerned with interior branch points only. In that case we can argue that since there are no branch points for $t > 0$, the zeroes of f are of even order and coincide with the poles of g, so that f and g have the forms stated in Lemma 3.3 of [1], namely:

$$\begin{aligned} f(z) &= f_0(z)\prod_{i=1}^{m}(z-a_i)^2 \\ g(z) &= \frac{h(z)\prod_{i=1}^{m+k}(z-b_i)}{\prod_{i=1}^{m}(z-a_i)} \end{aligned}$$

However, this lemma was also used in [2], where boundary branch points are considered, and the formula was not established in this case. In the next section we will show that this formula is indeed correct, by ruling out the possibility of branch points in the lower half plane for $t > 0$ that converge to the origin as t goes to zero.

If the a_i are the zeroes of f, and each one is a double zero, and the $b_i(t)$ are the zeroes of fg^2, and each one is a double zero, then we can define analytic functions A and B such that

$$\begin{aligned} f(z) &= A^2 \\ f(z)g^2(z) &= -B^2 \\ g(z) &= iB/A \end{aligned}$$

We can find analytic functions A_0 and B_0, not vanishing at the origin, such that

$$\begin{aligned} A &= A_0\prod_{i=1}^{m}(z-a_i) \\ B &= B_0\prod_{i=1}^{m+k}(z-b_i) \end{aligned}$$

It is shown in [1] that A_0 and B_0 are real-analytic functions of z and t. When $t = 0$ we can assume $A_0 = 1$, since we can always change z to a constant times z. When $t = 0$, B_0 is some nonzero complex constant C.

In [1] it is proved that the a_i and b_i depend analytically on a rational power of the parameter t. Replacing t by a rational power of t, we can assume that the a_i and b_i depend analytically on t. Although they all approach zero as t goes to zero, some may go to zero faster than others. Let γ be the least integer such that one of the a_i or b_i goes to zero as t^γ. Among the a_i and b_i, the ones that go to zero as t^γ are called the *principal roots.* There are thus complex constants α_i, β_i such that

$$\begin{aligned} a_i &= \alpha_i t^\gamma + O(t^{\gamma+1}) \\ b_i &= \beta_i t^\gamma + O(t^{\gamma+1}) \end{aligned}$$

The principal roots are the ones with $\alpha_i \neq 0$ or $\beta_i \neq 0$.

An important theme is to study the behavior of the minimal surfaces u^t in the region where z is approximately t^γ. To this end we introduce

$$\begin{aligned} z &= t^\gamma w \\ \mathbf{A} &= \prod_{i=1}^{m}(w-\alpha_i) \\ \mathbf{B} &= C\prod_{i=1}^{m+k}(w-\beta_i) \end{aligned}$$

Note that

$$\begin{aligned} z - a_i &= t^\gamma(w-\alpha_i) \\ A &= t^{m\gamma}\mathbf{A} + O(t^{m\gamma+1}) \\ B &= t^{(m+k)\gamma}\mathbf{B} + O(t^{(m+k)\gamma+1}) \end{aligned}$$

Recall that $A = A_0\prod(z-a_i(t))$, where A_0 is analytic in t and z and $A_0(0,0)=1$.

Eventually, we will need to analyze the situation in the w-plane for small positive t, as well as for $t=0$. Generally we try to use boldface type to indicate functions of w rather than z, and then to indicate t-dependence, we add a tilde. Thus we have

$$\begin{aligned} \tilde{\mathbf{A}} &= A(t, z/t^\gamma) \\ \tilde{\mathbf{B}} &= B(t, z/t^\gamma) \\ \tilde{\mathbf{a}}_i &= a_i/t^\gamma \\ \tilde{\mathbf{b}}_i &= b_i/t^\gamma \\ \tilde{\mathbf{A}}_0(t,w) &= A_0(t, t^\gamma w) \\ \tilde{\mathbf{B}}_0(t,w) &= B_0(t, t^\gamma w) \\ \mathbf{A}_0(w) &= \tilde{\mathbf{A}}_0(0,w) \\ \mathbf{B}_0(w) &= \tilde{\mathbf{B}}_0(0,w) \end{aligned}$$

The exception to this typographical scheme is the use of α_i and β_i instead of $\mathbf{a}_i$ and $\mathbf{b}_i$. We shall also need to examine the way in which the $\tilde{\mathbf{a}}_i$ and $\tilde{\mathbf{b}}_i$ converge to α_i and β_i. We define

$$\begin{aligned} \tilde{\alpha}_i &= \text{Re } \tilde{\mathbf{a}}_i \\ \tilde{\beta}_i &= \text{Re } \tilde{\mathbf{b}}_i \end{aligned}$$

In case α or β is real, we then define the nonzero coefficients $\hat{a}_i$ and $\hat{b}_i$ and exponents η_i and δ_i as follows:

$$\begin{aligned} \tilde{\mathbf{a}}_i &= \tilde{\alpha}_i + i\hat{a}_i t^{\eta_i} + \text{Im } O(t^{\eta_i+1}) \\ \tilde{\mathbf{b}}_i &= \tilde{\beta}_i + i\hat{b}_i t^{\delta_i} + \text{Im } O(t^{\delta_i+1}) \end{aligned}$$

Note that $\mathbf{a}_i$ lies in the upper half plane for $t > 0$ so that there must be a nonzero imaginary term in $\tilde{\mathbf{a}}_i$. Should it happen that $\tilde{\mathbf{b}}_i$ is real for all t, then $\hat{b}_i := 0$ and $\delta_i := 1$, so that terms $\hat{b}_i t^{\delta_i}$ are zero.

In case β_i is not real, we shall also need to consider the rate at which $\tilde{\mathbf{b}}_i$ converges to Re β_i. Actually, it turns out that what is important is the *difference* of the rates at which $\tilde{\mathbf{b}}_i$ and the real part of the conjugate root $\tilde{\mathbf{b}}_j$ (where $\beta_i = \bar{\beta}_j$) approach β. We define $\hat{e}_i$ and δ'_i so that

$$\tilde{\mathbf{b}}_i - \tilde{\mathbf{b}}_j = \hat{e}_i t^{\delta'_i} + O(t^{\delta'_i+1}).$$

(Thus $\hat{e}_i = -\hat{e}_j$ when $\beta_i = \bar{\beta}_j$.) Again, if it should happen that $\tilde{\mathbf{b}}_i = \beta$ for all small t, then $\hat{e}_i := 0$ and $\delta'_i := 1$. If β_i is not real, we define $\hat{b}_i$ and δ_i so that

$$\text{Im } \tilde{\mathbf{b}}_i - \text{Im } \beta_i = \hat{b}_i t^{\delta_i}.$$

Then just to summarize, we have

$$\tilde{\mathbf{b}}_i = \tilde{\beta}_i + \text{Im}(\beta_i) + i\hat{b}_i t^{\delta_i}$$

Suppose that $\beta_i = \bar{\beta}_j$. We have for real w,

$$(w - \tilde{\beta}_i)(w - \tilde{\beta}_j) \quad \cong \quad (w - \tilde{\beta}_i - i \text{ Im}(\beta_i) - i\hat{b}_i t^{\delta_i})(w - \tilde{\beta}_j + i \text{ Im}(\beta_i) - i\hat{b}_j t^{\delta_j})$$

Therefore, for real w,

$$\begin{aligned} &\text{Im}((w - \tilde{\mathbf{b}}_i)(w - \tilde{\mathbf{b}}_j)) \\ &= \text{Im}(\beta_i)(\tilde{\beta}_j - \tilde{\beta}_i) - \hat{b}_i t^{\delta_i}(w - \tilde{\beta}_j) - \hat{b}_j t^{\delta_j}(w - \tilde{\beta}_i) + \mathbf{O}(t) \\ &= -\text{Im}(\beta_i)\hat{e}_i t^{\delta'_i} - \hat{b}_i t^{\delta_i}(w - \tilde{\beta}_j) - \hat{b}_j t^{\delta_j}(w - \tilde{\beta}_i) + \mathbf{O}(t) \qquad (2) \end{aligned}$$

while

$$\text{Re}((w - \tilde{\mathbf{b}}_i)(w - \tilde{\mathbf{b}}_j)) = w^2 + \text{Im}(\beta_i)^2 + \mathbf{O}(t) \qquad (3)$$

Note the use of $\mathbf{O}(t)$ instead of $O(t)$, which indicates that the convergence as $t \to 0$ is uniform in compact subsets of the w-plane, since no denominators are involved. The distinction between $\mathbf{O}(t)$ and $O(t)$ is made more generally in Section 7 below, where we have to consider other expressions that have denominators that may go to zero with t.

6 The branch point is isolated

By "isolated" we mean that the branch point of $u(0,\cdot)$ is not the limit of branch points of $u(t,\cdot)$ which lie in the lower half-plane (outside the portion of the surface bounded by Γ). Until this is proved, we are not entitled to assume that the roots of the function f in the Weierstrass representation are double. The necessity of proving this was overlooked in [2]; the analysis there assumed that the roots of f are double. We will prove in this section that such branch points for $t > 0$ cannot occur.

As discussed above, we are assuming that we have an analytic one-parameter family $u(t,z)$ of minimal surfaces, with parameter t, and when $t = 0$, there is a boundary branch point of u at the origin. The parameter domain is the upper half plane. The branch points of a minimal surface are the common zeroes of f and fg^2; it is these points which are of interest in this section.

In general we might, *a priori*, have some common zeroes $c_i(t)$ of f and fg^2, say $2N$ of them. The number must be even since the total number of roots of f is the even number $2m$, and the rest of the roots of f are double. We must have $\mathrm{Im}(c_i(t)) < 0$ for $t > 0$, since there will be branch points at the c_i, and $c_i(t)$, like $a_i(t)$ and $b_i(t)$, depends analytically on some power of t called t^γ and converging to 0 as t approaches 0. The formulas for f, fg^2, and g are then

$$\begin{aligned} f(z) &= A^2 S \\ f(z)g^2(z) &= -\lambda^2 B^2 S \\ g(z) &= \frac{fg^2}{f} = \frac{i\lambda B}{A} \\ f(z)g(z) &= i\lambda ABS \end{aligned}$$

where

$$A(z) = A_0(z,t)\prod_{i=1}^{m-N}(z - a_i)$$

$$B(z) = i\lambda B_0(z,t)\prod_{i=1}^{m+k-N}(z - b_i)$$

$$S(z) = \prod_{i=1}^{2N}(z - c_i)$$

and $A_0(0,0) = 1$, $B_0(0,0)$ is real, and $|\lambda| = 1$ with $\lambda \neq -1$. Note that the common zeroes c_i do not enter into the formula for g. In [2], in effect we assumed $N = 0$ so that S does not occur.

In [1], we proved (Lemma 3.3) that each a_i and b_i, as zeroes of functions analytic jointly in z and t, depend analytically on t^r for some rational number r. The same applies to the c_i. Taking the greatest common divisor of the denominators of these numbers r (one r for each a_i, b_i, and c_i, as i varies), we can assume that all the a_i, b_i, and c_i are analytic in t^r (for the same r).

Replacing t by t^r in the original parametrization, we can assume they are all analytic in t. Hence, each a_i is asympotically a constant times t^{γ_i}, where γ_i is an integer. Similarly, each b_i and c_i is asymptotic to a power of t. Define γ to be the least of these exponents of t, that is, the smallest exponent in the asymptotic form of any a_i, b_i, or c_i.

Near each a_i, the Gauss map of the minimal surface covers most of the sphere, as shown in the proof of Theorem 4.1 of [1]. If any α_i has positive imaginary part, we can do the same thing:

Lemma 7 $\text{Im}(\alpha_i) \leq 0$

Proof. Let $w = z/t^\gamma$ and $\alpha_i(t) = a_i/t^\gamma$, and $\alpha_i = \alpha_i(0)$. Let C_1 be any circle around α_i in the w-plane that excludes all the other α_j (except of course if $\alpha_j = \alpha_i$). The formula for the Gauss map in the w-plane is

$$G(w) = g(z/t^\gamma) = t^{k\gamma} \frac{\prod(w - \beta_i)}{\prod(w - \alpha_i)} + O(t^{(k+1)\gamma})$$

The factor of $t^{k\gamma}$ guarantees that $G(w)$ tends to zero on C_1, so the Gauss map covers nearly the entire sphere, $m-N$ times, on the interior of C_1. In particular, near each a_i the Gauss map covers more than the upper hemisphere. Therefore, if the circle C_1 were contained in the upper half-plane, the eigenvalue of the second variation (explained in [1], or in [14], p. 103) would be less than the critical value 2, contradiction. If α_i had positive imaginary part, then for small enough t, the circle C_1 would be contained in the upper half plane. Therefore, $\text{Im}(\alpha_i) \leq 0$. This proves the lemma.

Lemma 8 *The c_i do not occur; that is, the branch point is isolated.*

Proof. We will apply the Hemispheric covering theorem (Theorem 2) to the Gauss map $N(t, z)$. There are $M - N$ of the a_i in the upper half plane. We know that the eigenvalue $\lambda_{\min} \geq 2$ for $t > 0$, since u is a relative minimum for $t > 0$. Therefore, the theorem implies that the Gaussian area of u when $t > 0$ is $2\pi(M - N) + O(t)$ greater than the Gaussian area when $t = 0$.

Now consider the Gauss-Bonnet-Sasaki-Nitsche formula:

$$\int_\Gamma \kappa_g ds = \int\int -KW dxdy + 2\pi + 2M\pi$$

where κ_g is the geodesic curvature, the double integral on the right is the total curvature, and M is the sum of the orders of interior branch points and half-orders of boundary branch points. We consider what happens in this equation as the parameter t approaches zero. The left-hand side is continuous in t, since the integrand is (pointwise) bounded by the curvature of Γ. Since for positive t there are no branch points, the total curvature of $u(t, \cdot)$ for $t > 0$ must be approximately $2M\pi$ larger than the total curvature of $u(0, \cdot)$. This extra curvature comes from extra coverings of the sphere by the Gauss map near the a_i.

We now have two formulas for the excess Gaussian area: $2\pi(M-N)$ from the hemispheric covering theorem (Theorem 2) and $2\pi M$ from the Gauss-Bonnet-Sasaki-Nitsche formula. Therefore $N = 0$ and the c_i do not occur. This completes the proof of the lemma.

Remark: It is also possible to show that the c_i do not occur without using the hemispheric covering theorem (Theorem 2), by carrying them through the calculation of the eigenfunction given in a subsequent section. Eventually they can be eliminated. We think the proof is conceptually clearer using the hemispheric covering theorem, even though the proof of that theorem is complicated.

7 The Gauss map in the w-plane

In this section we exploit the hemispheric covering theorem (Theorem 2) to limit the possibilities for the configurations of the α_i and β_i in the w-plane. After eliminating the possibility of branch points in the lower half plane for positive t, we have the following basic formulae:

$$\begin{aligned} A(z) &= A_0(z,t)\prod_{i=1}^{m}(z-a_i(t)) \\ B(z) &= iB_0(z,t)\prod_{i=1}^{m+k}(z-b_i(t)) \\ f(z) &= A^2 \\ f(z)g^2(z) &= -\lambda^2 B^2 \\ g(z) &= \frac{i\lambda B}{A} \\ f(z)g(z) &= i\lambda AB \end{aligned}$$

and $A_0(0,0) = 1$, $B_0(0,0)$ is real, and $|\lambda| = 1$ with $\lambda \neq -1$.

Each $a_i(t)$ is analytic in t and so can be written as

$$\begin{aligned} a_i(t) &= \alpha_i t^{\gamma_i} + O(t^{\gamma_i+1}) \\ b_i(t) &= \beta_i t^{\gamma_{2m+i}} + O(t^{\gamma_{2m+i}+1}) \end{aligned}$$

Let γ be the least of the γ_i. The "principal roots" are those a_i and b_i with asymptotic behavior t^γ. We will make no use of the α_i and β_i for the non-principal roots, and henceforth regard them as undefined. We needed them only to define γ. Thus, henceforth, there are at most as many of the α_i as there are principal roots, and there may be fewer if two or more $a_i(t)$ have the same limit α_i. The possibility is not ruled out that two or more of the $a_i(t)$ might be identical for all (sufficiently small) t, but in that case, they still get different subscripts.

We introduce $\mathbf{B_0} = B_0(0,0)$, so $\mathbf{B_0}$ is a real constant, and introduce

$$\tilde{\mathbf{A}}(t,w) = A(t^\gamma w) = A(z)$$

$$\tilde{\mathbf{B}}(t,w) \quad = \quad B(t^\gamma w) = B(z)$$

Then when t goes to zero subject to $w = t^\gamma z$ we have $\tilde{\mathbf{A}}(0,w) = \mathbf{A}(w)$ where

$$\mathbf{A} \quad = \quad w^{m-b}\prod_{i=1}^{b}(w-\alpha_i)$$

$$\mathbf{B} \quad = \quad \lambda \mathbf{B_0} w^{m+k-p}\prod_{i=1}^{p}(w-\beta_i)$$

There is no need for a constant $\mathbf{A_0}$ since $A(0,0) = 1$, because we scaled the surface so that $a_m = 1$ to begin with. Here $|\lambda| = 1$ and $\lambda \neq -1$.

Lemma 9 *(i) The α_i are all real.*
(ii) The β_i are real and/or occur in complex-conjugate pairs.
(iii) $\mathbf{B}$ is real on the real axis and $\lambda = 1$.

Proof.
On the real axis, since Γ is tangent to the X-axis at origin, in the w-plane we have $Y = 0$. Therefore,

$$Y(w) = \lim_{t\to 0} {}^2u(t^\gamma w) = \operatorname{Im}\int_0^w \mathbf{A}^2\,dw = 0$$

Therefore the polynomial $\int_0^w \mathbf{A}^2 dw$ is real on the real axis. It follows that its coefficients are real. Hence the coefficients of its derivative $\mathbf{A}^2$ are real. Hence the roots of $\mathbf{A}^2$, which are the α_i, come in complex-conjugate pairs. But we know from the eigenvalue argument that α_i cannot have positive imaginary part. It follows that all the α_i are real.

Similarly, on the real axis we have

$$Z \cong \operatorname{Im}\,\lambda\int_0^w \mathbf{AB}\,dw = 0.$$

That is, Z is real on the real axis. Therefore also the derivative $dZ/dw = \lambda\mathbf{AB}$ is real on the real axis. But $\lambda\mathbf{AB}$ is a polynomial; since it is real on the real axis, its coefficients are real. Since the α_i are real, $\mathbf{A}$ is real on the real axis; hence $\lambda\mathbf{B}$ is real on the real axis. Hence its roots, which are the β_i, are real or occur in complex-conjugate pairs. Hence $\mathbf{B} = \mathbf{B}_0\prod(w-\beta_i)$ is real on the real axis. Hence λ is real. Since $|\lambda| = 1$ and $\lambda \neq -1$ by construction, we have $\lambda = 1$. That completes the proof of the lemma.

It follows that

$$A \quad \cong \quad t^{m\gamma}\mathbf{A}$$

$$B \quad \cong \quad t^{(m+k)\gamma}\mathbf{B}$$

The function g in the Weierstrass representation, which is the stereographic projection of the unit normal N, is given by $g(z) = iB/A$, and the normal itself is given by

$$N = (2\operatorname{Re}(g), 2\operatorname{Im}(g), |g|^2 - 1)/(|g|^2 + 1). \tag{4}$$

We write $\bar{A}$ for the complex conjugate of A. Substituting $g = iB/A$, we have

$$N = \frac{1}{|A|^2 + |B|^2} \begin{bmatrix} 2\,\mathrm{Re}(i\bar{A}B) \\ 2\,\mathrm{Im}(i\bar{A}B) \\ B\bar{B} - A\bar{A} \end{bmatrix} \tag{5}$$

Substituting $z = t^\gamma w$, we find

$$\begin{aligned} N &= \frac{1}{|\mathbf{A}|^2 + t^{2k\gamma}|\mathbf{B}|^2} \begin{bmatrix} -2\,\mathrm{Im}(\bar{\mathbf{A}}\mathbf{B})t^{k\gamma} \\ 2\,\mathrm{Re}(\bar{\mathbf{A}}\mathbf{B})t^{k\gamma} \\ -A\bar{A} + t^{2k\gamma}\mathbf{B}\bar{\mathbf{B}} \end{bmatrix} \\ &= \begin{bmatrix} -2\,\mathrm{Im}(\mathbf{B}/\mathbf{A})t^{k\gamma} \\ 2\,\mathrm{Re}(\mathbf{B}/\mathbf{A})t^{k\gamma} \\ -1 + t^{2k\gamma}|\mathbf{B}/\mathbf{A}|^2 \end{bmatrix} \end{aligned}$$

Thus on compact sets bounded away from α_i, $N(w)$ converges uniformly to $(0, 0, -1)$.

For small positive t, we know that the normal N covers m hemispheres, and it does so near the a_i. For the principal a_i, this happens in the w-plane near the α_i. These hemispheres "pinch off" when $t \to 0$, and this "lost" Gaussian area is compensated for by the branch point that appears when $t = 0$. The introduction of $w = t^\gamma z$ and the explicit formula above for N as a function of w and t allows us to see what happens to those m hemispheres. Specifically, the stereographic projection of N, namely g, is given by $g(t, z) = iB/A = t^{k\gamma} i\mathbf{B}/\mathbf{A}$, so $\mathbf{B}/\mathbf{A}$ can be used to analyze the "extra hemispheres".

The formula for $N(w)$ shows us the extra "hemispheres" (which are within $O(t)$ of being hemispheres), with the north pole taken on at α_i. However, if the α_i are multiple roots of $\mathbf{A}$, which is to say that several $a_j(t)$ approach the same α_i as $t \to 0$, then we may not necessarily have m hemispheres in the w-plane when $t = 0$. For small positive t, we do have m hemispheres, but some of the original m hemispheres may be "pinched off" at a rate faster than t^γ, and disappeared before the w-plane could capture them when $t = 0$. This also could happen for any non-principal a_i, near the origin of the w-plane.

Lemma 10 *If $\alpha_i < \alpha_j$ and no other α_ℓ occurs between α_i and α_j, and neither α_i nor α_j is equal to any β_ℓ, then*

(i) exactly one $\tilde{\mathbf{a}}_i$ converges to α_i, and

(ii)the number of β_ℓ with $\alpha_i \le \beta_\ell \le \alpha_j$ is 1 (counting multiplicity as zeroes of g)

(iii) exactly one $\mathbf{b}_\ell$ converges to this β_ℓ as $t \to 0$.

Proof. First we consider the behavior of $N(w)$ near α_i. Consider a circle C_ρ of small radius ρ about $\tilde{\mathbf{a}}_i(t) = t^{-\gamma} a_i(t)$ in the w-plane. Here ρ does not depend on t but is less than half the distance from α_i to any other α_j. Then C_ρ meets the real axis in two points $p(t)$ and $q(t)$, converging as $t \to 0$ to points p, q, since α_i is real. At p and q, and indeed on the entire circle C_ρ, N is within $O(t)$ of the south pole. We can therefore apply the hemispheric covering theorem (Theorem

2) to conclude that on the portion of the upper half plane inside C_ρ (call it D^+) N is a covering map with as many sheets as there are $\tilde{\mathbf{a}}_j$ converging to this particular α_i, with each sheet covering the same near-hemisphere S_+ or S_-. In particular, we will prove that N does pass one or more times "over the top" on $[p, q]$, taking a value within $O(t)$ of the north pole somewhere on $[p, q]$. Let $\epsilon(t)$ be such that $\epsilon(t) = O(t)$ and ${}^1N([p, q]) < \epsilon(t)$. Suppose that $N[p, q]$ does not contain a point within $2\epsilon(t)$ of the north pole. Since $N(\partial C_\rho) = O(t)$, for sufficiently small t, $N(\partial D^+)$ lies outside the neighborhood of the north pole of radius $2\epsilon(t)$. Hence the number of sheets of N is constant on that neighborhood. But that neighborhood overlaps both S_+ and S_-, so the number of sheets over each of those near-hemispheres is at least the number over the north pole, which is not zero. This contradicts the hemispheric covering theorem. Hence, for every sufficiently small t we can find a point r_i in $[p, q]$ such that $N(r_i)$ is within $2\epsilon(t)$ of the north pole. Similarly, for each index j there is a point r_j near α_j where $N(w)$ is within $O(t)$ of the north pole. (We have not proved that the $r_j(t)$ can be chosen to depend continuously on t, but that will not be needed.)

Now we fix i and consider α_i. There may be many $\tilde{\mathbf{a}}_j$ converging to α_i as $t \to 0$, but by hypothesis there are no $\tilde{\mathbf{b}}_j$ converging to α_i. Fix a point p less than α_i but greater than any $\alpha_j < \alpha_i$, and a point q greater than α_i but less than any $\alpha_j > \alpha_i$. Let C be a circle passing through the x-axis at right angles at p and again at q, and let D^+ be the half-disk inside C above the x-axis. We can apply the hemispheric covering theorem to D^+. The conclusion is that N is a covering map in D^+, away from any ramification points of N that lie in D^+. The number of sheets is the total number M of $\tilde{\mathbf{a}}_j$ that converge to α_i. On the other hand, since no $\tilde{\beta}_j$ is in D^+ (for sufficiently small t), the fact that the $\tilde{\alpha}_j$ are the poles of g, the stereographic projection of N, and the $\tilde{\beta}$ are its zeroes, implies that in the interior of C, N is a covering map whose number of sheets is M. As $t \to 0$ this approaches, on an annulus inside C, the covering map defined by the inverse stereographic projection of w^M. If $M > 1$ this covering extends into both hemispheres by an amount that does not go to zero with t, contradicting the conclusion obtained from the hemispheric covering theorem. Hence $M = 1$, proving part (i) of the theorem.

Now we consider $\alpha_i < \alpha_{i+1}$. We may suppose the indices have been chosen so that no other α_j lies between these two. Let $p < \alpha_i$, with no other α_ℓ between p and α_i. Let $q > \alpha_{i+1}$, with no other α_ℓ between α_{i+1} and q. As before define the circle C to pass through the x-axis at right angles at both p and q, and the half-disk D^+ to be bounded by the upper half of C and the x-axis. We know that exactly one $\tilde{a}_i$ converges to α_i and exactly one $\tilde{a}_j$, namely $\tilde{a}_{i+1}$, converges to α_{i+1}. There may be, so far as we know now, many $\tilde{b}_j$ converging to one or more β_j between α_i and α_{i+1}. Let M be the number of these $\tilde{\beta}_j$; so M is also the total number of the β_j counting multiplicities as zeros of g. As before, apply the hemispheric covering theorem (Theorem 2) to N and D^+; the conclusion is that in D^+, away from any ramification points of N, the map N defines a two-sheeted covering. On the other hand, on any compact subset of D^+ disjoint from α_i, α_{i+1}, and the β_j, the map converges to the inverse stereographic projection of

$$\frac{\prod_{j=1}^{M}(w-\beta_j)}{(w-\alpha_i)(w-\alpha_{i+1})}$$

For this to satisfy the conclusion of the hemispheric covering theorem, the winding number of N around C must be exactly 1 or -1; otherwise, on the upper half of C, both hemispheres will be partly covered, i.e. some values of N on the upper half of C will have X-coordinate more than ϵ and some will have X-coordinate less than $-\epsilon$, for some positive ϵ independent of t. This winding number, however, has absolute value $N-2$. Hence $N=1$. This completes the proof of the lemma.

8 Calculation of the eigenfunction

In this section, we compute the eigenfunction of the variational problem associated with the second variation of area. This is shown in [2] to be $\phi = k \cdot N$, where $k = u_t = du/dt$ is the "tangent vector" to the one-parameter family, and N is the unit normal. These calculations are essentially repeated from [2], but with some simplifications, and more important, the isolation of an important case that was overlooked in [2], which will be treated in the next section. We start with a formula for N:

$$N = \frac{1}{(|g|^2+1)}\begin{bmatrix} 2\mathrm{Re}(g) \\ 2\mathrm{Im}(g) \\ |g|^2-1 \end{bmatrix}. \tag{6}$$

We write $\bar{A}$ for the complex conjugate of A. Introduce $Q = |A|^2 + |B|^2 = A\bar{A} + B\bar{B}$. By (6) and the equation $g = iB/A$, we have

$$NQ = \begin{bmatrix} 2\,\mathrm{Re}(i\bar{A}B) \\ 2\,\mathrm{Im}(i\bar{A}B) \\ B\bar{B} - A\bar{A} \end{bmatrix} \tag{7}$$

Recall the Weierstrass representation:

$$u = \mathrm{Re}\begin{bmatrix} \frac{1}{2}\int f - fg^2\,dz \\ \frac{i}{2}\int f + fg^2\,dz \\ \int fg\,dz \end{bmatrix} \tag{8}$$

Using the formulas for f and g in terms of A and B, we have

$$u = \mathrm{Re}\begin{bmatrix} \frac{1}{2}\int A^2 + B^2\,dz \\ \frac{i}{2}\int A^2 - B^2\,dz \\ i\int AB\,dz \end{bmatrix} \tag{9}$$

Differentiate with respect to t:

$$k = \mathrm{Re}\begin{bmatrix} \frac{1}{2}\int (A^2)_t + (B^2)_t\,dz \\ \frac{i}{2}\int (A^2)_t - (B^2)_t\,dz \\ i\int (AB)_t\,dz \end{bmatrix} \tag{10}$$

Using (7),(10), and the definition $\phi = k \cdot N$, we find

$$\phi Q = \mathrm{Re} \begin{bmatrix} \frac{1}{2}\int (A^2)_t + (B^2)_t dz \\ \frac{i}{2}\int (A^2)_t - (B^2)_t dz \\ i\int (AB)_t dz \end{bmatrix} \cdot \begin{bmatrix} 2\,\mathrm{Re}(i\bar{A}B) \\ 2\,\mathrm{Im}(i\bar{A}B) \\ B\bar{B} - A\bar{A} \end{bmatrix} \tag{11}$$

Using the identities

$$(\mathrm{Re}(X), \mathrm{Im}(X), 0) \cdot (\mathrm{Re}(Y), -\mathrm{Im}(Y), 0) = \mathrm{Re}(XY)$$

$$(\mathrm{Re}(X), \mathrm{Im}(X), 0) \cdot (\mathrm{Re}(Y), \mathrm{Im}(Y), 0) = \mathrm{Re}(X\bar{Y})$$

we can rewrite (11) in the form

$$\begin{aligned} \phi Q \quad = \quad & \mathrm{Re}(i\bar{A}B \int (A^2)_t dz) \\ & -\mathrm{Re}(iA\bar{B} \int (B^2)_t dz) \\ & -(A\bar{A} - B\bar{B})\mathrm{Re}(i \int (AB)_t dz) \end{aligned} \tag{12}$$

This is our basic formula for ϕ. When $t = 0$, it is possible that ϕ is identically zero as a function of z, and so for some integer n we have $\phi = t^n \psi$, where ψ is not identically zero in z when $t = 0$. The function $\psi(0, z)$ is written ψ^0. We have $Q = r^{2m-2N} + O(r^{2m-2N+1})$, where $r = |z|$, so that any singularity in ψ^0 would arise from terms r^j in (12) with $j < 2m - 2N$.

We have already defined the γ_i to be the exponents such that $a_i \cong \alpha_i t^{\gamma_i}$; the principal roots are those a_i or b_i with $\gamma_i = \gamma$, the least of these exponents. We now assume that the principal roots are listed first among the a_i and b_i. Let b be the number of principal a_i, and p the number of principal b_i. Then $q = p + b$ is the total number of principal roots.

We introduce the complex variable $w = z/t^\gamma$, and calculate the behavior of the eigenfunction ϕ in the w-plane, as t goes to zero. This captures the behavior in a region of the z-plane of shrinking diameter of order t^γ. Our most important calculations will be in the w-plane. We will expand expressions in powers of t and w, and look at the leading term in t, which will in several important cases be a rational function or polynomial in w.

Definition 4 *For quantities U and V depending on t and z and/or w, we define $U \cong V$ to mean that $U - V$ converges uniformly to zero in compact subsets of the w-plane disjoint from the zeroes of V as $t \to 0$.*

In practice we only apply this definition to functions that depend analytically on t and w or quotients of such functions. For functions analytic in t and w, we can describe the notion $U \cong V$ as follows: first replace z by $t^\gamma w$ in both U and V so that U and V depend on t and w but not z. Then, if U is real-analytic in w and t, $U \cong V$ means that, when expressed as series in t, the leading terms of U and V are equal as functions of w (or of w and $\bar{w}$). If U and V are quotients of real-analytic functions, then $U \cong V$ if and only if we can write $U = U_0/U_1$ and $V = V_0/V_1$, where $U_0 \cong V_0$ and $U_1 \cong V_1$ and the U_i and V_i are real-analytic.

When U and V are already explicitly functions of w rather than z, we often just write $U = V + O(t)$ instead of $U \cong V$. Note that the $O(t)$ in such a situation implies uniform convergence in compact subsets of the w-plane away from the zeroes of V. Later we will also use a boldface $\mathbf{O}(t)$ to imply uniform convergence in compact subsets of the w-plane (without the requirement of avoiding any points).

We use the notation $U \cong V$ when the expressions for U and V may involve z. Consider how this works in calculations: in comparing two terms of the form $z^n t^{m\gamma}$, the z^n will eventually become $t^{n\gamma} w^n$, so $z^n t^{m\gamma} = w^n t^{(n+m)\gamma}$, and we can neglect the term with the greater sum $n + m$. Note that, after eliminating z in favor of w and t, the leading terms shown in approximate equations written with $\cong$ will have the same power of t, so that when we take the limit as t approaches 0, we will get exact equations involving w. The rule for differentiating with respect to t in this situation is found by applying the chain rule as follows, where d/dt means the partial derivative with respect to t, holding z fixed, and $\partial/\partial t$ means the partial derivative with respect to t, holding w fixed.

$$\begin{aligned} dH/dt &= \frac{\partial H}{\partial t} + \frac{dH}{dw}\frac{dw}{dt} \\ &= \frac{\partial H}{\partial t} + \frac{dt^{-\gamma} z}{dt}\frac{dH}{dw} \\ &= \frac{\partial H}{\partial t} - \gamma t^{-\gamma-1} z \frac{dH}{dw} \\ &= \frac{\partial H}{\partial t} - \gamma t^{-\gamma-1} t^{\gamma} w \frac{dH}{dw} \end{aligned}$$

The final result for differentiating with respect to t is then

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} - t^{-1}\gamma w \frac{dH}{dw} \tag{13}$$

We have

$$2k = \frac{d}{dt}\mathrm{Re}\begin{bmatrix} \int_0^z (A^2 + B^2)dz \\ i\int_0^z (A^2 - B^2)dz \\ 2i\int_0^z ABdz \end{bmatrix} \tag{14}$$

Recall that

$$\begin{aligned} \mathbf{A} &= w^{m-b}\prod_{i=1}^{b}(w - \alpha_i) \\ \mathbf{B} &= \mathbf{B_0} w^{m+k-p}\prod_{i=1}^{p}(w - \beta_i) \\ A &\cong t^{(m)\gamma}\mathbf{A} \\ B &\cong t^{(m+k)\gamma}\mathbf{B} \end{aligned}$$

Expressing (14) in terms of w instead of z, we have

$$2k \cong \frac{d}{dt}\mathrm{Re}\begin{bmatrix} t^{(2m+1)\gamma}\int_0^w \mathbf{A}^2 dw + t^{(2m+2k+1)\gamma}\int_0^w \mathbf{B}^2 dw \\ it^{(2m+1)\gamma}\int_0^w \mathbf{A}^2 dw - it^{(2m+2k+1)\gamma}\int_0^w \mathbf{B}^2 dw \\ 2it^{2m+k+1)\gamma}\int_0^w \mathbf{AB}dw \end{bmatrix} \tag{15}$$

Now that we have expressed N in terms of t and w, equations written with $\cong$ simply mean that higher powers of t can be neglected compared to lower powers. Therefore, even before differentiating, we can drop the terms in the first two components of (15) with higher powers of t.

$$2k \cong \frac{d}{dt}\mathrm{Re}\begin{bmatrix} t^{(2m+1)\gamma}\int_0^w \mathbf{A}^2 dw \\ it^{(2m+1)\gamma}\int_0^w \mathbf{A}^2 dw \\ 2it^{(2m+k+1)\gamma}\int_0^w \mathbf{AB}dw \end{bmatrix} \tag{16}$$

Applying (13), we find

$$\begin{aligned} &2k/\gamma \cong \\ &\mathrm{Re}\begin{bmatrix} t^{(2m+1)\gamma-1}(2m+1)\int_0^w \mathbf{A}^2 dw \\ t^{(2m+1)\gamma-1}(2m+1)i\int_0^w \mathbf{A}^2 dw \\ 2it^{(2m+k+1)\gamma-1}(2m+k+1)\int_0^w \mathbf{AB}dw \end{bmatrix} \\ &-\mathrm{Re}\begin{bmatrix} t^{(2m+1)\gamma-1}\mathbf{A}^2 w \\ it^{(2m+1)\gamma-1}\mathbf{A}^2 w \\ 2it^{(2m+k+1)\gamma-1}\mathbf{AB}w \end{bmatrix} \end{aligned} \tag{17}$$

Define

$$\begin{aligned} \Lambda_1 &= |\mathbf{A}|^2/(t^{2k\gamma}|\mathbf{B}|^2 + |\mathbf{A}|^2) \\ \Lambda_2 &= |\mathbf{B}|^2/(t^{2k\gamma}|\mathbf{B}|^2 + |\mathbf{A}|^2). \end{aligned} \tag{18}$$

From (6) we obtain

$$N \cong \begin{bmatrix} 2t^{k\gamma}\mathrm{Re}(i\mathbf{B}/\mathbf{A})\Lambda_1 \\ 2t^{k\gamma}\mathrm{Im}(i\mathbf{B}/\mathbf{A})\Lambda_1 \\ t^{2k\gamma}\Lambda_2 - \Lambda_1 \end{bmatrix} \tag{19}$$

Since Λ_1 and Λ_2 have the same denominator, and no factors of t in the numerator, in the last component of (19) we can replace $t^{2k\gamma}\Lambda_2 - \Lambda_1$ by $-\Lambda_1$; and since $t^{2k\gamma}|\mathbf{B}|^2 + |\mathbf{A}|^2 \cong |\mathbf{A}|^2$, we can replace Λ_1 by 1, so that (19) becomes

$$N \cong \begin{bmatrix} 2t^{k\gamma}\mathrm{Re}(i\mathbf{B}/\mathbf{A}) \\ 2t^{k\gamma}\mathrm{Im}(i\mathbf{B}/\mathbf{A}) \\ -1 \end{bmatrix} \tag{20}$$

We are going to take the dot product of (20) and (17) to compute $\phi = k \cdot N$.

$$\begin{aligned} 2\phi/\gamma \cong 2t^{(2m+k+1)\gamma-1} \quad \{ \quad & \mathrm{Re}(i\mathbf{B}/\mathbf{A})\mathrm{Re}((2m+1)\int_0^w \mathbf{A}^2 dw - \mathbf{A}^2 w) \\ + \quad & \mathrm{Im}(i\mathbf{B}/\mathbf{A})\mathrm{Re}(i((2m+1)\int_0^w \mathbf{A}^2 dw - \mathbf{A}^2 w)) \\ - \quad & \mathrm{Re}(i((2m+k+1)\int_0^w \mathbf{AB}dw - \mathbf{AB}w)\} \end{aligned} \tag{21}$$

Incidentally, this formula makes it clear that h^0 (where $k = t^n h$) is a forced Jacobi direction, since $\phi = k \cdot N$ goes to zero as $t^{(2m+k+1)\gamma-1}$ while k goes to zero only as $t^{(2m+1)\gamma-1}$, according to (16). Thus $h^0 \cdot N = 0$, which implies h^0 is forced Jacobi. This is the conclusion of Theorem 7.2 of [2], here reached without the complicated calculations given there. Define

$$\chi = \frac{\phi}{\gamma t^{(2m+k+1)\gamma-1}}. \tag{22}$$

We write χ^0 for the limit of χ as t goes to zero. Dividing (21) by $t^{(2m+k+1)\gamma-1}$ and taking the limit as t goes to zero, we obtain:

$$\begin{aligned} \chi^0 &= \mathrm{Re}(i\mathbf{B}/\mathbf{A})\mathrm{Re}((2m+1)\int_0^w \mathbf{A}^2 dw - \mathbf{A}^2 w) \\ &+ \mathrm{Im}(i\mathbf{B}/\mathbf{A})\mathrm{Re}(i((2m+1)\int_0^w \mathbf{A}^2 dw - \mathbf{A}^2 w)) \\ &- \mathrm{Re}(i((2m+k+1)\int_0^w \mathbf{A}\mathbf{B} dw - \mathbf{A}\mathbf{B} w) \end{aligned} \tag{23}$$

The first two terms have the form $\mathrm{Re}(iu)\mathrm{Re}(v) + \mathrm{Im}(iu)\mathrm{Re}(iv)$, which is equal to $\mathrm{Re}(iu)\mathrm{Re}(v) - \mathrm{Im}(iu)\mathrm{Im}(v)$, which in turn is $\mathrm{Re}(iuv) = -\mathrm{Im}(uv)$. We can therefore rewrite (23) as

$$\chi^0 = -\mathrm{Im}\ H(w) \tag{24}$$

where

$$\begin{aligned} H(w) &= \frac{\mathbf{B}}{\mathbf{A}}((2m+1)\int_0^w \mathbf{A}^2\, dw - \mathbf{A}^2 w) \\ &\quad -(2m+k+1)\int_0^w \mathbf{A}\mathbf{B}\, dw - \mathbf{A}\mathbf{B} w \end{aligned} \tag{25}$$

Multiplying the factor $\mathbf{B}/\mathbf{A}$ through in the first term, we obtain

$$\begin{aligned} H(w) &= \frac{\mathbf{B}}{\mathbf{A}}(2m+1)\int_0^w \mathbf{A}^2 dw - \mathbf{A}\mathbf{B} w \\ &\quad -(2m+k+1)\int_0^w \mathbf{A}\mathbf{B}\, dw + \mathbf{A}\mathbf{B} w \end{aligned}$$

Now, the two terms $\mathbf{A}\mathbf{B}w$ cancel out, leaving our final formula for H:

$$H(w) = \frac{\mathbf{B}}{\mathbf{A}}(2m+1)\int_0^w \mathbf{A}^2\, dw - (2m+k+1)\int_0^w \mathbf{A}\mathbf{B}\, dw \tag{26}$$

This is the formula obtained in [2], equation (23).

Lemma 11 *If H is not constant, then all the a_i and b_i are principal.*

Proof. We first fix some notation. Let Q_a be the number of non-principal a_i, and let Q_b be the number of non-principal b_i. Let $P = Q_a + Q_b$ be the total number of non-principal roots. Let α be the product of the α_i (all of which are nonzero and real); let β be the product of the β_i (all of which are nonzero, but have not yet been proved real). Note however that the product β is real since the β_i are real or come in complex-conjugate pairs.

We wish to analyze the asymptotic behavior of H near the origin. We will show that, if H is not constant, its leading term is a nonzero real constant times βw^{P+1}. Once we have done that, we can proceed as follows: Since H is not constant, χ^0 is not identically zero, so it must have only one sign in the upper half plane, and this is possible only if $P = 0$. Therefore, all the roots a_i and b_i are principal.

It remains to prove that H has the stated asymptotic form. We find it more convenient to analyze the form (25) rather than the cancelled form (26).

Consider the second term in (25), namely

$$(2m + k + 1) \int_0^w \mathbf{AB}\, dw - \mathbf{AB}w$$

By definition of γ, at least one of the a_i or b_i is a principal root. Therefore it is not the case that $\mathbf{AB} = O(w^{2m+k})$. Indeed, asymptotically $\mathbf{AB}$ is a constant times w^P, where P is the total number of non-principal roots among the a_i and b_i, and $P < 2m + k$. Hence, asymptotically, this term has the form $\beta c w^{P+1}$ for some nonzero real constant c. The actual value of c is $B_0\alpha\,[(2m + k + 1)/(P + 1) - 1]$.

Now consider the behavior of the first term in (25) near the origin, namely

$$\frac{\mathbf{B}}{\mathbf{A}}((2m + 1) \int_0^w \mathbf{A}^2\, dw - \mathbf{A}^2 w). \tag{27}$$

Suppose, for purposes of contradiction, that there are no principal a_i, that is, all the principal roots are among the b_i. In that case, we have

$$(2m + 1) \int_0^w \mathbf{A}^2\, dw = w^{2m+1} = \mathbf{A}^2 w,$$

so the entire first term in (25) would vanish, leaving the nonzero term in w^{P+1} contributed by the second term as the asymptotic form of H. Since in that case we are finished, we have proved that we may assume there are some principal roots among the a_i. Putting the matter another way, in that case we have a contradiction as in the second paragraph of the proof, since, under the assumption that there are no principal a_i, we will not have $P = 0$, and hence we will have $H(w) = O(r^2)$.

In that case, the leading terms of the two parts of (27) do *not* cancel, and the asymptotic form of (27) is a constant times w^{P+1}. The constant is $\beta B_0\alpha\,[(2m + 1)/(Q + R + 1) - 1]$. It is not zero since $Q + R < 2m$, because we have proved there do exist some principal roots among the a_i.

Putting together the two leading terms we have calculated, we find that the asymptotic form of H is

$$H(w) = \beta\alpha c w^{P+1} \tag{28}$$

where

$$c = \frac{2m+k+1}{P+1} - \frac{2m+1}{Q_a+1} \tag{29}$$

Since we cannot immediately exclude that $c = 0$, we cannot conclude that H is not constant. In any case we have proved $H(w) = O(w^{P+1})$. By hypothesis, H is not constant. Therefore, unless $P = 0$, we have $H(w) = O(r^2)$, and χ^0 will not be of one sign in the upper half-plane, contradiction. Therefore, $P = 0$, proving that all the a_i and b_i are principal. This completes the proof of the lemma.

Lemma 12 *If H is not constant, then at each α_i, either $\mathbf{B}/\mathbf{A}$ has a simple zero or a simple pole.*

Proof. By Lemma 10, if we prove there are two distinct real α_i, then the existence of a real β_j follows. Also, we have proved that all the α_i are real.

Fix a value of i such that α_i is real, and suppose that P of the numbers $\alpha_1, \ldots, \alpha_m$ are equal to α_i. Suppose also that Q of the β_j are equal to α_i. Let β be the product of $\alpha_i - \beta_j$ over the remaining β_j, so $\beta \neq 0$, and let α be the product of $\alpha_i - \alpha_j$ over j such that $\alpha_j \neq \alpha_i$. We consider the behavior of the terms in (26) for w near α_i. Let $w = \alpha_i + \xi$. We have

$$\begin{aligned}
\int_0^w \mathbf{A}^2\,dw &= \int_0^{\alpha_i} \mathbf{A}^2\,dw + O(\xi^{2P+1}) \\
\frac{\mathbf{B}}{\mathbf{A}} &= \frac{\beta}{\alpha}\xi^{Q-P}(1+O(\xi)) \\
\frac{\mathbf{B}}{\mathbf{A}}\int_0^w \mathbf{A}^2\,dw &= \frac{\beta}{\alpha}\xi^{Q-P}\left[\int_0^{\alpha_i} \mathbf{A}^2\,dw + \alpha^2\frac{\xi^{2P+1}}{2P+1} + O(\xi^{2P+2})\right] \\
&= \frac{\beta}{\alpha}\int_0^{\alpha_i} \mathbf{A}^2\,dw + \frac{\alpha\beta}{2P+1}\xi^{P+Q+1} + O(\xi^{P+Q+2}) \\
\int_0^w \mathbf{A}\mathbf{B}\,dw &= \int_0^{\alpha_i} \mathbf{A}\mathbf{B}\,dw + \alpha\beta\frac{\xi^{P+Q+1}}{P+Q+1} + O(\xi^{P+Q+2})
\end{aligned}$$

The last two lines can be combined using (26) to yield the following

$$H(w) = c_1 + c_2\xi^{Q-P} + \left[\frac{2m+1}{2P+1} - \frac{2m+k+1}{P+Q+1}\right]\xi^{P+Q+1} + O(\xi^{P+Q+2})$$

where $c_1 = \int_0^{\alpha_i} \mathbf{A}\mathbf{B}\,dw$ and $c_2 = (2m+1)(\beta/\alpha)\int_0^{\alpha_i} \mathbf{A}^2\,dw$. Since $\mathrm{Im}(H)$ is of one sign in the upper half plane, and H is not constant by hypothesis, we must have $Q = P$ or $Q - P = \pm 1$. In fact $Q = P$ is impossible, since in that case the ξ^{Q-P} term would be constant, and the leading term would be the ξ^{P+Q+1} term, whose coefficient would be $k/(2P+1)$. Since this coefficient is not zero, Im(H) would not have one sign in the upper half-plane (near α_i), contradiction. Therefore $Q = P \pm 1$ are the only two possibilities. That means that at each α_i, either $\mathbf{B}/\mathbf{A}$ has a simple zero, or a simple pole. That completes the proof of the lemma.

9 The case when H is constant

We cannot rule out the possibility that H is constant by analysis of the eigenfunction alone, because it is possible to choose α_i and β_i so that $H(w)$ defined as above is constant. For example, take $m = 2$ and $\alpha_1 = 1/\sqrt{2}$, $\alpha_2 = -1/\sqrt{2}$, and $k = 2m + 1 = 3$. Take $\beta_1 = \alpha_1$ and $\beta_2 = \alpha_2$. The polynomial $\int_0^w \mathbf{A}^2 dw$ is then $\int_0^w (w^2 - 1/2)^2 dw$. Evaluating this integral we find $(1/5)w^5 - (1/3)w^3 + (1/4)w$. The nonzero roots of this satisfy $w^2 = 5/6 \pm (5/2)\sqrt{((1/9) - (1/5))}$. These roots are not real since the expression under the square root is negative. Indeed we can construct a family of minimal surfaces bounded by a straight line (instead of a Jordan curve) that exhibit this behavior.

However, the assumption that H is constant leads to very specific forms for $\mathbf{A}$ and $\mathbf{B}$, as we will show in this section. Using that specific information, the hemispheric covering theorem, and the fact that the boundary curve is real-analytic (and not a straight line), we will be able to rule out the case that H is constant.

Lemma 13 *Let $\tilde{u}$ be a one-parameter family of minimal surfaces bounded by the analytic Jordan curve Γ, such that for $t > 0$, $\tilde{u}$ furnishes a relative minimum of area, and u has a boundary branch point when $t = 0$. Let H be defined as above. If $H(w)$ is constant, then:*

(i) Each α_i is equal to some β_j (with the same multiplicity). That is, $\mathbf{B}/\mathbf{A}$ has neither a zero nor a pole at α_i.

(ii) Each real β_i is equal to some α_j.

(iii) The non-real β_i are the nonzero roots of the polynomial $\int_0^w \mathbf{A}^2\, dw$.

(iv) $k/(2m+1)$ is an integer.

(v) Each non-real β_i, as well as zero, has multiplicity $k/(2m+1)$ as a root of $\mathbf{B}$.

(vi) $\mathbf{B}/\mathbf{A}$ and hence H are polynomials.

(vii) $c(\int \mathbf{A}^2\, dw)^{k/(2m+1)} = \mathbf{B}/\mathbf{A}$ for some constant $c \neq 0$.

(viii) All of the a_i are principal; that is, $\mathbf{A}$ does not vanish at the origin.

(ix) There are exactly $2m$ distinct non-real β_i.

Proof. Let $\tilde{\alpha}$ be one of the principal roots, i.e. either an α_i, or a β_i. We will analyze the asymptotic behavior of H near $\tilde{\alpha}$. Suppose that L of the α_j are equal to $\tilde{\alpha}$, and M of the β_j are equal to $\tilde{\alpha}$. Define

$$\begin{aligned} \xi &:= w - \tilde{\alpha}_i \\ \alpha &:= \prod_{j:\alpha_j \neq \tilde{\alpha}} (\tilde{\alpha} - \alpha_j) \end{aligned}$$

$$\beta := \mathbf{B}_0 \prod_{j:\beta_j \neq \tilde{\alpha}} (\tilde{\alpha} - \beta_j)$$

Then

$$\begin{aligned}
\mathbf{A} &= \alpha\xi^L + O(\xi^{L+1}) \\
\mathbf{B} &= \beta\xi^M + O(\xi^{M+1}) \\
\frac{\mathbf{B}}{\mathbf{A}} &= \frac{\beta}{\alpha}\xi^{M-L} + O(\xi^{M-L+1})
\end{aligned}$$

Now calculate the asymptotic form of H near $\tilde{\alpha}$, breaking each of the two integrals in the formula for H into an integral from 0 to $\tilde{\alpha}$, plus an integral from $\tilde{\alpha}$ to w. Define

$$\begin{aligned}
e_1 &:= \int_0^{\tilde{\alpha}} \mathbf{A}^2 dw \\
e_2 &:= \int_0^{\tilde{\alpha}} \mathbf{AB} dw
\end{aligned}$$

Then

$$\begin{aligned}
H(w) &= (2m+1)\frac{\mathbf{B}}{\mathbf{A}}\int_0^w \mathbf{A}^2 dw - (2m+k+1)\int_0^w \mathbf{AB} dw \\
&= (2m+1)\left[\frac{\beta}{\alpha}\xi^{M-L} + O(\xi^{M-L+1})\right]\left[e_1 + \frac{\alpha^2\xi^{2L+1}}{2L+1} + O(\xi^{2L+2})\right] \\
&\quad -(2m+k+1)\left[e_2 + \frac{\alpha\beta\xi^{L+M+1}}{L+M+1} + O(\xi^{L+M+2})\right] \\
&= e_1(2m+1)\frac{\beta}{\alpha}\xi^{M-L} - (2m+k+1)e_2 + O(\xi^{M-L+1})
\end{aligned}$$

Now suppose $\tilde{\alpha}$ is real, which is the case if $\tilde{\alpha} = \alpha_i$ or one of the real β_i. We have $e_1 \neq 0$ since $\mathbf{A}^2 \geq 0$ on the real axis, and $\mathbf{A}^2$ is not identically zero on the real axis. Therefore, in case $M \neq L$, we have

$$H(w) = e_1(2m+1)\frac{\beta}{\alpha}\xi^{M-L} + O(\xi^{M-L+1})$$

and hence $H(w)$ is not constant unless $M = L$, that is, at $\tilde{\alpha}$, the order of the zero of $\mathbf{B}$ is the same as the order of the zero of $\mathbf{A}$. We have thus proved (i) and (ii), since at a real β_i which is not equal to any α_i, we would have $L = 0$ and $M > 0$, and at a real α_i which is not equal to any β_i, we would have $M = 0$ and $L > 0$; and similarly if the multiplicities of the roots of $\mathbf{A}$ and $\mathbf{B}$ were different.

The same analysis applies even if $\tilde{\alpha}$ is not real, as long as $e_1 \neq 0$. Therefore, if H is constant we must have $e_1 = 0$ at each non-real β_i, which proves (iii).

We now claim that there must exist a β_i at which the order of the zero of $\mathbf{B}$ is greater than the order of the zero of $\mathbf{A}$. Suppose that, at each β_i, the order of the zero of $\mathbf{B}$ is less than or equal to the order of the zero of $\mathbf{A}$. That is, the number of β_j equal to β_i is less than or equal to the number of α_j equal to β_i. Then the total number of principal b_i is less than or equal to the total number of principal a_i. Since there are m of the a_i and $m+k$ of the b_i, there are k more of the b_i than of the a_i. Therefore, the number of non-principal b_i must be at least k more than the number of non-principal a_i. That is, $Q_b \geq k + Q_a$. But we have shown above that $Q_b < k$ if $c = 0$, and if $c \neq 0$ then H is not constant, but we have assumed H is constant, so $Q_b < k$. Since $Q_a \geq 0$, this is a contradiction. This contradiction shows that there must be some β_i at which the order of the zero of $\mathbf{B}$ exceeds the order of $\mathbf{A}$. Take this β_i as $\tilde{\alpha}$ in the above analysis; then $M \neq L$.

We have

$$\begin{aligned} H(w) &= (2m+1)\left[\frac{\beta}{\alpha}\xi^{M-L} + O(\xi^{M-L+1})\right]\left[\frac{\alpha^2\xi^{2L+1}}{2L+1} + O(\xi^{2L+2})\right] \\ &\quad -(2m+k+1)\left[e_2 + \alpha\beta\frac{\xi^{L+M+1}}{L+M+1} + O(\xi^{L+M+2})\right] \\ &= -(2m+k+1)e_2 + (2m+1)\alpha\beta\frac{\xi^{M+L+1}}{2L+1} \\ &\quad -(2m+k+1)\alpha\beta\frac{\xi^{L+M+1}}{L+M+1} + O(\xi^{L+M+2}) \end{aligned}$$

Note that the terms $O(\xi^{L-M+1})$ do not occur since $e_1 = 0$. The first non-constant term is ξ^{M+L+1}. The coefficient of this term is

$$\alpha\beta\left[\frac{2m+1}{2L+1} - \frac{2m+k+1}{L+M+1}\right].$$

We know β_i is not real (otherwise $M = L$ as shown above), and that $e_1 = 0$. Hence $L = 0$, since all the α_j are real, and hence are not equal to β_i. Therefore, the coefficient is

$$\alpha\beta\left[(2m+1) - \frac{2m+k+1}{M+1}\right].$$

Since we have assumed H is constant, this coefficient is zero, which means $M = k/(2m+1)$. This proves (iv).

The preceding analysis applies at each non-real β_i. Since the α_i are real, we have $L = 0$ and $M = k/(2m+1)$ is the multiplicity of the root as a zero of $\mathbf{B}$. This proves (v) for the nonzero roots. Counting the total number of roots proves it for the zero root also.

Since H is constant, we have $c = 0$, so

$$P + 1 = (Q_a+1)\frac{2m+k+1}{2m+1}.$$

Since $P = Q_b + Q_a$, we have

$$\begin{aligned} Q_a + Q_b + 1 &= (Q_a+1)\frac{2m+k+1}{2m+1} \\ &= (Q_a+1)(1 + \frac{k}{2m+1}) \end{aligned}$$

and hence

$$Q_b = \frac{k(Q_a+1)}{2m+1}$$

Since we have already proved $k/(2m+1)$ is an integer, it follows that $Q_b \geq Q_a+1$. Since Q_b is the order of $\mathbf{B}$ at zero and Q_a is the order of $\mathbf{A}$ at zero, and since the nonzero roots of $\mathbf{A}$ are matched by nonzero roots of $\mathbf{B}$ of equal multiplicity, we have proved that $\mathbf{B}/\mathbf{A}$ is a polynomial. This proves (vi). The degree of this polynomial (of course) is k, since the degree of $\mathbf{B}$ is $m+k$ and that of $\mathbf{A}$ is m.

To prove (vii): $(\int_0^w \mathbf{A}^2\,dw)^{k/(2m+1)}$ and $\mathbf{B}/\mathbf{A}$ are both polynomials, by (vi). They have the same non-real zeroes, with the same multiplicities, by (iii) and (v). $\mathbf{B}/\mathbf{A}$ has no real zeroes except 0, by (i) and (ii), and since $\mathbf{A}^2 \geq 0$ on the real axis, the polynomial $(\int_0^w \mathbf{A}^2\,dw)^{k/(2m+1)}$ also has no real zeroes except 0. The zero at $w=0$ of the two polynomials has the same multiplicity $k/(2m+1)$, by (v). Hence the ratio of the two polynomials is constant, proving (vii).

To prove (viii): we have shown that $\mathbf{B}/\mathbf{A}$ is a constant times $(\int \mathbf{A}^2\,dw)^{k/(2m+1)}$. Suppose there are d non-principal a_i. Then asymptotically for small w, we have

$$\begin{aligned}\frac{\mathbf{B}}{\mathbf{A}} &= w^{k/(2m+1)-d}(1+O(w)) \\ \left(\int_0^w \mathbf{A}^2\,dw\right)^{k/(2m+1)} &= w^{(2d+1)k/(2m+1)}(1+O(w))\end{aligned}$$

Equating the exponents of the leading terms, we have

$$\begin{aligned}\frac{k}{2m+1} - d &= \frac{k(2d+1)}{2m+1} \\ k - d(2m+1) &= k(2d+1)\end{aligned}$$

which implies $d=0$, proving (viii).

To prove (ix): Let d be the number of non-real β_i. We count the $\tilde{\mathbf{b}}_i$. There are $m+k$ of them altogether. Of these, m approach some α_j, $k/(2m+1)$ approach 0, and $k/(2m+1)$ approach each of the non-real β_i. Therefore

$$m+k = m + \frac{k}{2m+1} + \frac{dk}{2m+1}.$$

It follows that $d = 2m$, proving (ix). That completes the proof of the lemma.

The following lemma reverses the above calculations, giving a sufficient condition for H to be constant.

Lemma 14 *Suppose $k=(2m+1)(p-q)$ and $c(\int_0^w \mathbf{A}^2\,dw)^{p-q} = \mathbf{B}/\mathbf{A}$ for some constant c. Then H is constant.*

Proof. We have

$$H = (2m+1)\frac{\mathbf{B}}{\mathbf{A}}\int_0^w \mathbf{A}^2 dw - (2m+k+1)\int_0^w \mathbf{AB}\,dw$$

$$
\begin{aligned}
&= (2m+1)c\left(\int \mathbf{A}^2\,dw\right)^{p-q}\int \mathbf{A}^2\,dw - (2m+k+1)\int_0^w \mathbf{AB}\,dw \\
&= (2m+1)c\left(\int \mathbf{A}^2 dw\right)^{p-q+1} - (2m+k+1)\int_0^w \mathbf{AB}\,dw
\end{aligned}
$$

Differentiating with respect to w, we have

$$
\begin{aligned}
\frac{dH}{dw} &= (2m+1)c(p-q+1)\left(\int \mathbf{A}^2\,dw\right)^{p-q}\mathbf{A}^2 - (2m+k+1)\mathbf{AB} \\
&= (2m+1)(p-q+1)\left(\frac{\mathbf{B}}{\mathbf{A}}\right)\mathbf{A}^2 - (2m+k+1)\mathbf{AB} \\
&= (2m+1)(p-q+1)\mathbf{AB} - (2m+k+1)\mathbf{AB} \\
&= ((2m+1)(p-q)+2m+1)\mathbf{AB} - (2m+k+1)\mathbf{AB} \\
&= (k+2m+1)\mathbf{AB} - (2m+k+1)\mathbf{AB} \\
&= 0
\end{aligned}
$$

We have proved $dH/dw = 0$; hence H is constant. That completes the proof of the lemma.

We next recall the basic formulas for the unit normal N. We have

$$
\begin{aligned}
N(z) &= \frac{1}{|g|^2+1}\begin{bmatrix} 2\ \mathrm{Re}\ g \\ 2\ \mathrm{Im}\ g \\ |g|^2-1 \end{bmatrix} \\
&= \frac{1}{|\frac{B}{A}|^2+1}\begin{bmatrix} 2\ \mathrm{Re}(iB/A) \\ 2\ \mathrm{Im}(iB/A) \\ |B/A|^2-1 \end{bmatrix} \\
&= \frac{|\tilde{\mathbf{A}}|^2}{|\tilde{\mathbf{A}}|^2+t^{2k\gamma}|\tilde{\mathbf{B}}|^2}\begin{bmatrix} -2t^{k\gamma}\mathrm{Im}(\tilde{\mathbf{B}}/\tilde{\mathbf{A}}) \\ 2t^{k\gamma}\mathrm{Re}(\tilde{\mathbf{B}}/\tilde{\mathbf{A}}) \\ -1+t^{2k\gamma}|\tilde{\mathbf{B}}/\tilde{\mathbf{A}}|^2 \end{bmatrix} \\
&= \frac{1}{|\tilde{\mathbf{A}}|^2+t^{2k\gamma}|\tilde{\mathbf{B}}|^2}\begin{bmatrix} -2t^{k\gamma}\mathrm{Im}(\tilde{\mathbf{B}}\bar{\tilde{\mathbf{A}}}) \\ 2t^{k\gamma}\mathrm{Re}(\tilde{\mathbf{B}}\bar{\tilde{\mathbf{A}}}) \\ -|\tilde{\mathbf{A}}|^2+2t^{k\gamma}|\tilde{\mathbf{B}}|^2 \end{bmatrix} \\
&= \frac{1+\mathbf{O}(t)}{\mathbf{A}^2+t^{2k\gamma}\mathbf{B}^2}\begin{bmatrix} -2t^{k\gamma}\mathrm{Im}(\tilde{\mathbf{A}}\bar{\tilde{\mathbf{B}}}) \\ 2t^{k\gamma}\mathbf{AB} \\ -\mathbf{A}^2+2t^{k\gamma}|\mathbf{B}|^2 \end{bmatrix} \qquad (30)
\end{aligned}
$$

Lemma 15 *(i) There exists a point $eq(t)$ on the real w-axis such that $|\tilde{\mathbf{A}}|^2 = t^{2k\gamma}|\tilde{\mathbf{B}}|^2$, and $eq(t)$ is bounded as t goes to 0.*

(ii) In fact there exist two such points, say $eq_1(t)$ and $eq_2(t)$, such that $N_2(eq_1(t)) = 1+\mathbf{O}(t)$ and $N_2(eq_2(t)) = -1+\mathbf{O}(t)$.

(iii) In addition, if all the α_i are nonzero, we can ensure that $eq_1(t)$ and $eq_2(t)$ have the same sign.

(iv) $eq_i(t)$ can be found in such a way that $eq_i(t)$ has a limit as t goes to zero, and even depends analytically on a rational power of t. Therefore such

expressions as $\tilde{A}[w := eq_i(t)]$ are analytic in a rational power of t, and have limits as $t \to 0$.

Proof. The stereographic projection of the unit normal N is given by $\tilde{\mathbf{B}}/t^{k\gamma}\tilde{\mathbf{A}}$. The condition demanded in the lemma therefore has the geometric meaning that N should be horizontal at $eq(t)$. This can also be seen directly from equation (30). Recall that the zeroes of $\tilde{\mathbf{A}}$ are denoted by $\tilde{\mathbf{a}}_i$, which converge to α_i as t goes to zero. Recall also that we may have several $\tilde{\mathbf{b}}_j$ and several $\tilde{\mathbf{a}}_j$ converging to the same α_i. Fix a value of i. Pick a larged fixed value R (say, greater than twice the maximum magnitude of all the α_j and β_j. ${}^3N$ is $-1 + O(t^{2k})$ at the endpoints of the interval $[-R, R]$, and indeed on the entire circle of radius R, as can be seen from (30). Assume, for proof by contradiction, that no such point $eq(t)$ with the property demanded in the lemma exists. Then ${}^3N$ is never zero on $[-R, R]$. By continuity, ${}^3N < 0$ on that interval. Let D be the part of the disk of radius R lying in the upper half-plane. Then the image $N(D)$ covers the upper half of the Riemann sphere, since $N(\tilde{a}_i)$ is the north pole $(0, 0, 1)$, but $N(\partial D)$ lies in the lower half-sphere. Hence the least eigenvalue of the second variation of area is less than 2 for $t > 0$, contradiction. That completes the proof of (i).

Ad (ii): Since N is continuous on the w-axis (for fixed t), the above argument yields at least two different values of w where N is horizontal. To prove (ii), however, a deeper argument is required. Fix a value of R greater than all the $|\alpha_i$, and let $D+$ be a half-disk in the upper half of the w-plane centered at origin. Apply the hemispheric covering theorem (Theorem **??**), to the map N on this D^+. Then N is an m-sheeted covering of the set S_- or S_+, where $S_- = \{(X, Y, Z) \in S^2 : X \leq -\epsilon\}$ and $S_+ = \{(X, Y, Z) \in S^2 : X \geq \epsilon\}$, and $\epsilon = O(t)$. Say it is S_- that is covered m times. Let p and q be points on the boundary of S_- with Z-coordinate 0, and Y-coordinate $= 1 + O(t)$ for p and $-1 + O(t)$ for q. Since N is a covering map, there are not just 1 but m pre-images of p and q under N. Now consider a path π on S^2 intersect the plane $Z = 0$ that connects p to the boundary of $N(D^+)$; the length of this path is $O(\epsilon) = O(t)$ since the covering does not extend into S_+. So $\pi(0) = p$ and $\pi(1)$ is on the boundary of that sheet of $N(D^+)$. Since N is a covering, we can find a path η in the w-plane such that $N(\eta(s)) = \pi(s)$. Take $eq_1(t)$ to be $\eta(1)$. Then $N(eq_1(t))$ lies on the plane $Z = 0$, and since $\pi(1) - \pi(0) = O(t)$ we have $|N(eq_1(t)) - (0, 1, 0)| = O(t)$ as required. If it is S_+ that is covered m times, instead of S_-, the same argument applies with the roles of S_+ and S_- switched.

Ad (iii): In the argument for (ii), the circle of radius R can be replaced by any circle avoiding the α_i, since all we need is for $\mathbf{A}$ to be bounded below on that curve to obtain the estimate ${}^3N = -1 + \mathbf{O}(t)$. We know that all the α_i are real, and by hypothesis we can assume they are all nonzero. Fix an $\epsilon > 0$ with $|\alpha_i| > 2\epsilon$ for all α_i. If there is at least one positive α_i, then take the circle C centered at $(R + \epsilon, 0)$ of radius R; if all the α_i are negative, then take the circle C centered at $(-R - \epsilon, 0)$ of radius R. Apply the hemispheric covering theorem to N in the half-disk in the upper half plane bounded by C and the x-axis. As before, we obtain $eq_1(t)$ and $eq_2(t)$, but now the both are on the interval $[\epsilon, 2R+\epsilon]$ or both on the interval $-2R-\epsilon, -\epsilon]$, so they both have the same sign.

Ad (iv) The zero set of ${}^3N$ is defined by a quotient of functions analytic jointly in w and t. Its zero set therefore consists of a finite number of curves parametrizable as analytic functions of a rational power of t, as discussed with references in [1]. That completes the proof.

Lemma 16 *Let $u(t,z)$ be an analytic one-parameter family of minimal surfaces bounded by the same Jordan curve Γ, such that u is a relative minimum of area when $t > 0$, and u has a boundary branch point when $t = 0$, with the normal at the branch point pointing in the negative Z-direction. Suppose $\Gamma'(0) = (1, C_1 s^q, C_2 s^p)$ asymptotically near the branch point, where s is arc length. Let H be as defined above, so that $\phi = -\mathrm{Im}(H(z))$ is the apparent leading term in t of the first eigenfunction. Then H is not constant.*

Proof. Suppose, for proof by contradiction, that H is constant. Then $k = (2m+1)(p-q)$ and for some constant c we have

$$c\left(\int_0^w \mathbf{A}^2\,dw\right)^{p-q} = \frac{\mathbf{B}}{\mathbf{A}}$$

By (30) we have

$${}^2N = \frac{(1+\mathbf{O}(t))2t^{2k\gamma}\mathbf{A}\mathbf{B}}{\mathbf{A}^2 + t^{2k\gamma}\mathbf{B}^2}.$$

We now substitute $w := eq_i(t)$ into this equation, where $eq_1(t)$ and $eq_2(t)$ are the two points of Lemma 15. Then the two terms in the denominator are equal, and we obtain

$$\begin{aligned} {}^2N[w := eq_i(t)] &= (1+\mathbf{O}(t))\frac{2t^{2k\gamma}\mathbf{A}\mathbf{B}}{2t^{2k\gamma}\mathbf{B}^2} \\ &= (1+O(t))\frac{\mathbf{A}}{\mathbf{B}} \end{aligned}$$

Since when H is constant, we have $\mathbf{B}/\mathbf{A} = \int_0^w \mathbf{A}^2\,dw$, it follows that

$${}^2N[w := eq_i(t)] = \frac{1+O(t)}{c\int_0^{eq_i(t)} \mathbf{A}^2\,dw}$$

Because H is constant, by Lemma 13 none of the α_i are zero. According to Lemma 15, we can then choose $eq_1(t)$ and $eq_2(t)$ to have the same sign. Therefore the right hand side of this equation has the same sign for $i = 1$ and $i = 2$. This contradicts Lemma 15, which says that the left-hand side takes values near 1 and -1 (in some order) for $i = 1, 2$ and t small but positive. That completes the proof.

10 The imaginary parts of A, B, AB, and N

In order to complete the proof of the finiteness theorem, it will be necessary to consider not only the eigenfunction, but also the details of the convergence of $\tilde{\mathbf{A}} = A/t^{m\gamma}$ to $\mathbf{A}$ as t goes to zero, and of $\tilde{\mathbf{B}}$ to $\mathbf{B}$, and to calculate the behavior of the unit normal N for small positive t and w as well. These calculations will be carried out in this section, and the results used in the next section.

We have been using $O(t)$ to stand for a term which is bounded by a constant times t, uniformly in subsets of the w-plane bounded away from the α_i. However, now we will need to consider convergence even near the α_i, so we use a different typeface, $\mathbf{O}(t)$ to stand for a term which is bounded by a constant times t, uniformly in compact subsets of the w-plane.

We have

$$\text{Im } \tilde{\mathbf{A}} \quad = \quad \text{Re } \tilde{\mathbf{A}}_0 \text{ Im } \prod(w - \tilde{\mathbf{a}}_i) + \text{Im } \tilde{\mathbf{A}}_0 \text{ Re } \prod(w - \tilde{\mathbf{a}}_i)$$

Since $\mathbf{A}_0$ is real on the real axis when $t = 0$, the first imaginary coefficient (if there is one) in $\tilde{\mathbf{A}}_0$ is $\mathbf{O}(t)$. Let us give names to the coefficients and exponents in the leading term of Im $\tilde{\mathbf{A}}_0$:

$$\begin{aligned} \text{Im } \tilde{\mathbf{A}}_0 &= Et^{\nu\gamma} w^{n_\nu} + \mathbf{O}(t^{\nu\gamma+1}) + \mathbf{O}(t^{\nu\gamma} w^{\nu\gamma+1}) \\ &= Et^{\nu\gamma} w^{n_\nu}(1 + \mathbf{O}(t) + O(w)) \end{aligned}$$

where $E \neq 0$. That is, the "leading term" is the term with the lowest power of t, and for that power of t, the lowest power of w. Note that ν here might or might not be the same as the least power such that the coefficient of z^ν in f (when $t = 0$) is not real. But since A_0 is analytic in z and t, the leading term is a power of $z = t^\gamma w$ times another power of t, so we have

$$n_\nu \leq \nu\gamma \tag{31}$$

Similarly, we write

$$\text{Im } \tilde{\mathbf{B}}_0 = Dt^{\delta\gamma} w^{n_\delta}(1 + \mathbf{O}(t) + O(w))$$

for some constant $D \neq 0$, and we have

$$n_\delta \leq \delta\gamma. \tag{32}$$

Recall that the α_i are real, i.e. $\mathbf{A}$ is real on the real axis, whether or not H is constant, and all the a_i are principal, so that $\mathbf{A} = \prod(w - \alpha_i)$; and that if H is not constant, all the b_i are principal, so $\mathbf{B} = \prod(w - \beta_i)$. If H is constant, we index the b_i so that the ones which go to zero have the largest indices. In that case $\mathbf{B} = w^{k/(2m+1)} \prod(w - \beta_i)$. We review the notation: β_i is defined only for the i such that b_i is principal, i.e. $\lim_{t \to 0} b_i = \beta_i$ is nonzero. But for all i in the range 1 to $m + k$, $\tilde{\mathbf{b}}_i = b_i/t^{k\gamma}$ is defined.

We have for real w

$$\begin{aligned}\text{Im } \prod(w - \tilde{\mathbf{a}}_i) &= \text{Im } \prod(w - \text{Re}(\tilde{a}_i) - i\text{Im}(\tilde{\mathbf{a}}_i)) \\ &= (1 + O(t))\text{Im } \prod(w - \alpha_i - i\hat{a}_i t^{\eta_i}) \\ &= -(1 + O(t))\mathbf{A} \sum_{i=0}^{m} \frac{\hat{a}_i t_i^{\eta}}{w - \alpha_i}\end{aligned}$$

We then have for real w

$$\begin{aligned}\text{Im } \tilde{\mathbf{A}} &= \text{Im } (\tilde{\mathbf{A}}_0 \prod(w - \tilde{\mathbf{a}}_i)) \\ &= \text{Im } \tilde{\mathbf{A}}_0 \text{ Re} \prod(w - \tilde{\mathbf{a}}_i) + \text{Re } \tilde{\mathbf{A}}_0 \text{ Im} \prod(w - \tilde{\mathbf{a}}_i) \\ &= Et^{\nu\gamma} w^{n_\nu} \mathbf{A}(1 + \mathbf{O}(t) + O(w)) + \text{Im } \prod(w - \tilde{\mathbf{a}}_i) \\ &= Et^{\nu\gamma} w^{n_\nu} \mathbf{A}(1 + \mathbf{O}(t) + O(w)) - (1 + \mathbf{O}(t))\mathbf{A} \sum_{i=0}^{m} \frac{\hat{a}_i t^{\eta_i}}{w - \alpha_i} \\ &= \mathbf{A} \left[Et^{\nu\gamma} w^{n_\nu} - \sum_{i=0}^{m} \frac{\hat{a}_i t^{\eta_i}}{w - \alpha_i} \right] (1 + \mathbf{O}(t) + O(w))\end{aligned}$$

$$\begin{aligned}\text{Im } \tilde{\mathbf{A}}^2 &= 2\text{Re } \tilde{\mathbf{A}} \text{ Im } \tilde{\mathbf{A}} \\ &= 2\mathbf{A}^2(1 + \mathbf{O}(t) + O(w)) \left[Et^{\nu\gamma} w^{n_\nu} - \sum_{i=0}^{m} \frac{\hat{a}_i t^{\eta_i}}{w - \alpha_i} \right] \qquad (33)\end{aligned}$$

The computation of Im $\tilde{\mathbf{B}}$ is similar, but somewhat more complicated, since the β_i are not necessarily real, and moreover, if H is constant, not all the b_i are principal. We make the convention that if b_i is not principal then $\beta_i = 0$. That convention was not convenient earlier in the paper, but many of the subsequent lengthy formulas would be even lengthier without it.

We refer to equation (2) for notation and the start of the computation. Recall also the notation R and S for sets of indices such that β_i is real for $i \in R$ and the non-real β_i occur in complex-conjugate pairs with one member of each pair having an index in S. We include the i for which $\beta_i = 0$ in R. Then

$$\begin{aligned}\text{Im}(\tilde{\mathbf{B}}^2) &= 2\mathbf{B}^2 \Big[Dt^{\delta\gamma} w^{n_\delta} - \sum_{i \in R} \frac{\hat{b}_i t^{\delta_i}}{w - \beta_i} \\ &\quad - \sum_{i \in S} \text{Im}(\beta_i) \frac{\hat{e}_i t^{\delta'_i}}{w^2 + (\text{Im } \beta_i)^2)} \Big] (1 + \mathbf{O}(t) + O(w)) \qquad (34) \\ \text{Im}(\tilde{\mathbf{A}}\tilde{\mathbf{B}}) &= \mathbf{AB} \Big[Et^{\nu\gamma} w^{n_\nu} - \sum_{i=0}^{m} \frac{\hat{a}_i t^{\eta_i}}{w - \alpha_i} + Dt^{\delta\gamma} w^{n_\delta} - \sum_{i \in R} \frac{\hat{b}_i t^{\delta_i}}{w - \beta_i} \\ &\quad - \sum_{i \in S} \text{Im}(\beta_i) \frac{\hat{e}_i t^{\delta'_i}}{w^2 + (\text{Im}\beta_i)^2)} \Big] (1 + \mathbf{O}(t) + O(w)) \qquad (35)\end{aligned}$$

We have now calculated the imaginary parts of $\mathbf{A}^2$, $\mathbf{B}^2$, and $\mathbf{AB}$.

11 Geometric Considerations

In this section we combine the study of N as a covering map of the Riemann sphere with the extra information that N is perpendicular to an analytic curve Γ. We have already given a sample of information obtained in this way in the geometric bound on the index. Here we obtain additional information by considering the fact that each member of the one-parameter family of minimal surfaces $u(t,z)$ is bounded by the same Jordan curve Γ. As always, we assume that u is a relative minimum of area when $t > 0$, and u has a boundary branch point at origin when $t = 0$. Our method is to calculate dZ/dY in two ways: first as a function of arc length, using the parametrization of Γ; and second, using the Weierstrass representation of the surfaces u, expressed in terms of $\tilde{A}$ and $\tilde{B}$, or in other words, in terms of $\tilde{\mathbf{a}}_i$ and $\tilde{b}_i$.

We have, for real w,

$$\begin{aligned}\frac{dZ}{dY} &= \frac{dZ/dx}{dY/dx} \\ &= \frac{2\mathrm{Re}(fg)}{\mathrm{Im}(f+fg^2)} \\ &= t^{k\gamma}\frac{-2\mathrm{Im}(\tilde{\mathbf{A}}\tilde{\mathbf{B}})}{\mathrm{Im}(\tilde{\mathbf{A}}^2 - t^{2k\gamma}\tilde{\mathbf{B}}^2)} \qquad (36)\end{aligned}$$

Recall that p and q are integers such that

$$\Gamma'(\tau) = \begin{bmatrix} 1 \\ C_1\tau^q + O(\tau^{q+1}) \\ C_2\tau^p + O(t^{p+1}) \end{bmatrix}$$

We use s for arc length from origin. Since s is the integral of $|\Gamma'|$ by definition, and since $\tau = X = u_1$ by our choice of parametrization, we have

$$s = X + O(X^{2p}) + O(X^{2q})$$

On the real axis in the w-plane we have

$$s = t^{(2m+1)\gamma}\int \mathbf{A}^2 dw + O(t^{2m+2}).$$

Writing $u = (X,Y,Z)$ on the boundary (the real axis), we calculate

$$\begin{aligned}\frac{dZ}{dY} &= \frac{dZ/ds}{dY/ds} \\ &= \frac{C_2 s^p + O(s^{p+1})}{C_1 s^q + O(s^{q+1})} \\ &= C_5 s^{p-q} + O(s^{p-q+1}) \\ &= (1+\mathbf{O}(t))C_5 t^{(2m+1)(p-q)\gamma}(\int_0^w \mathbf{A}^2 dw)^{p-q} \qquad (37)\end{aligned}$$

on the real axis in the w-plane, with $C_5 \neq 0$.

Comparing (37) and (36), we have

$$t^{(2m+1)(p-q)} C_5 (1 + \mathbf{O}(t)) \left(\int_0^w \mathbf{A}^2 \, dw \right)^{p-q} = t^{k\gamma} \frac{-2\mathrm{Im}(\tilde{\mathbf{A}}\tilde{\mathbf{B}})}{\mathrm{Im}(\tilde{\mathbf{A}}^2 - t^{2k\gamma}\tilde{\mathbf{B}}^2)} \tag{38}$$

12 Finiteness

Theorem 3 *Let Γ be a real-analytic Jordan curve in R^3. Then Γ cannot bound infinitely many disk-type minimal surfaces which are relative minima of area in the C^n topology for any sufficiently large n.*

Remarks: Note that the larger n is, the stronger the theorem; so there is no point in worrying about the minimum value of n. As usual in Plateau's problem, we are counting two surfaces that differ only by a conformal reparametrization as the same surface.

Proof: Let Γ be a real-analytic Jordan curve in R^3; fix n large enough for the results of [18] to apply. Without loss of generality we may assume Γ does not lie in a plane. Assume Γ bounds infinitely many minimal surfaces of disk type. Then (see [1]), Γ bounds a one-parameter family of minimal surfaces, terminating in a surface with a boundary branch point.

Since Γ is real-analytic, and does not lie in a plane, we can choose coordinates so that the boundary branch point is at origin, oriented with its normal in the positive Z-direction, and Γ tangent to the X-axis at origin. There exist integers p and q and a parametrization of Γ in the form used above. Let $\tau(z) = X(z) = \mathrm{Re}\frac{1}{2}\int f - fg^2 dz$. Then on the boundary we have

$$u(z) \quad = \quad \Gamma(\tau(z))$$

The plan in this section is to compute the inner product

$$\langle \Gamma'(\tau(z)), N(z) \rangle$$

where $N(z)$ is the unit normal to $u(t, \cdot)$. This product must be zero for z real, since the unit normal must be normal to the boundary. This will connect the integers p and q, determined by the geometry of the boundary, to the coefficients in the Weierstrass representation of the surface. This normality condition, in the region where z is a constant times t^γ, produces equations not just between numbers, as a similar analysis for a single branched minimal surface does, but equations between functions of $w = z/t^\gamma$. We will show that these equations lead to a contradiction.

We make use of all the notation defined in previous sections. Fix a value of R much larger than all the $|\alpha_i|$ and $|\beta_i|$, and restrict attention to the domain $|w| \leq R$, that is, $|z| \leq Rt^\gamma$.

$$\tau(z) \quad = \quad \mathrm{Re}\frac{1}{2}\int f - fg^2 \, dz$$

$$\begin{aligned} &= \mathrm{Re}\frac{1}{2}\int A^2 + B^2\,dz \\ &= t^{(2m+1)\gamma}\mathrm{Re}\int \mathbf{A}^2\,dw + O(t^{(2m+1)\gamma+1}) \end{aligned}$$

since the B^2 term contains $t^{(m+k)\gamma}$, and $dz = t^\gamma dw$. Therefore on the boundary (that is, on the real axis) we have

$$\Gamma'(\tau(z)) = \begin{bmatrix} 1 \\ C_1 t^{(2m+1)q\gamma}(\int \mathbf{A}^2 dw)^q + O(t^{(2m+1)q\gamma+1}) \\ C_2 t^{(2m+1)p\gamma}(\int \mathbf{A}^2 dw)^p + O(t^{(2m+1)p\gamma+1}) \end{bmatrix}$$

We now need to give a name to the leading imaginary term, say $Dw^r t^\delta$, so that

$$B_0(z,t) \cong C + \mathrm{Re}(O(t)) + iDt^{\delta - r\gamma} z^r$$

We then have

$$\begin{aligned} B &\cong B_0(z,t)\prod(z-b_i) \\ &\cong (C + iDt^{\delta-r\gamma}z^r)\prod(z - \beta_i t^\gamma - i\hat\beta_i t^{\gamma+\delta_i}) \\ &\cong C\prod(z-\beta_i t^\gamma) + i\prod(z-\beta_i t^\gamma)t^\gamma\sum\frac{\hat\beta_i t^{\delta_i}}{z-\beta_i t^\gamma} \\ &\quad + iDt^{\delta-r\gamma}z^r\prod(z-\beta_i t^\gamma) \\ &\cong t^{(m+k)\gamma}\prod(w-\beta_i) + it^{(m+k)\gamma}\prod(w-\beta_i)\sum\frac{\hat\beta_i t_i^\delta}{w-\beta_i} \\ &\quad + iDt^{(m+k)\gamma+\delta-r\gamma}z^r\prod(w-\beta_i) \\ &\cong t^{(m+k)\gamma}\mathbf{B} + it^{(m+k)\gamma}\mathbf{B}\left(Dt^\delta w^r + \sum\frac{\hat\beta_i t^{\delta_i}}{w-\beta_i}\right) \end{aligned}$$

Similarly, we recall that $A = A_0(z,t)\prod(z-\alpha_i)$, where $A_0(0,0) = 1$; we introduce $\mathbf{A}_0(w,t) = A_0(t^\gamma w, t)$. We now need a name for the leading imaginary term of $\mathbf{A}_0$, say $Et^\nu w^s$, so that

$$A_0(z,t) \cong 1 + iEt^{\nu - s\gamma}z^s + \mathrm{Re}(O(t))$$

Then we have

$$A \cong t^{m\gamma}\mathbf{A} + it^{m\gamma}\mathbf{A}\left(Et^{\nu-s\gamma}z^s + \sum\frac{\hat\alpha_i t^{\eta_i}}{w-\alpha_i}\right)$$

Since the imaginary part is smaller than the real part by a factor of $O(t)$, we have

$$1/A \cong \frac{t^{-m\gamma}}{\mathbf{A}}\left(1 - i\left(Et^{\nu-s\gamma}z^s + \sum\frac{\hat\alpha_i t^{\eta_i}}{w-\alpha_i}\right)\right)$$

and hence

$$
\begin{aligned}
\mathrm{Re}(iB/A) &= -\mathrm{Im}(B/A) \\
&= -\mathrm{Im}(B)\mathrm{Re}(1/A) - \mathrm{Re}(B)\mathrm{Im}(1/A) \\
&\cong -t^{k\gamma}\left(\frac{\mathbf{B}}{\mathbf{A}}\right)\left(Dt^{\delta}w^{r} + \sum \frac{\hat{\beta}_i t^{\delta_i}}{w-\beta_i}\right) \qquad (39) \\
&\quad +t^{k\gamma}\left(\frac{\mathbf{B}}{\mathbf{A}}\right)\left(Et^{\nu}w^{s} + \sum \frac{\hat{\alpha}_i t^{\eta_i}}{w-\alpha_i}\right) + O(t^{(k+1)\gamma})
\end{aligned}
$$

$$
\cong t^{k\gamma}\left(\frac{\mathbf{B}}{\mathbf{A}}\right)\left(Et^{\nu}w^{s} - Dt^{\delta}w^{r} + \sum \frac{\hat{\alpha}_i t^{\eta_i}}{w-\alpha_i} - \sum \frac{\hat{\beta}_i t^{\delta_i}}{w-\beta_i}\right) \qquad (40)
$$

Let η be the minimum of all the exponents η_i,δ_i,δ, and ν. Define $\sum'$ to mean taking the sum only over terms i with powers of t equal to η in the previous equation. When we multiply the previous equation by a suitable power of t to expose the leading term, and take the limit as t goes to zero, Σ will become Σ'. Define

$$
\begin{aligned}
J(w) &:= \lim_{t\to 0} t^{-\eta}\left(Et^{\nu}w^{s} - Dt^{\delta}w^{r} + \sum \frac{\hat{\alpha}_i t_i^{\eta}}{w-\alpha_i} - \sum \frac{\hat{\beta}_i t_i^{\delta}}{w-\beta_i})\right) \\
&= E'w^{v} + \sum{}' \frac{\hat{\alpha}_i}{w-\alpha_i} - \sum{}' \frac{\hat{\beta}_i}{w-\beta_i}
\end{aligned}
$$

What will happen to the terms $Et^{\nu}w^{s}$ and $Dt^{\delta}w^{r}$? That will depend on whether $\eta < \xi_i$. If either of δ or ν is equal to η, we could get a term constant in t, containing w^r or w^s. However, if this happens, there is a restriction on the value of η. The following lemma summarizes the possibilities:

Lemma 17 *Denote the constant (in t) term in $J(w)$ by $E'w^v$. Here $v = r$ or s. The possible values of E', v, and the corresponding values of η are as follows:*
(1) $E' = E - D$ if $\delta = \nu = \eta$ and $r = s$. Then $v = r = s$.
(2) $E' = E$ and $v = s$ if $\nu = \eta$ but $\delta > \eta$, or if $\nu = \delta = \eta$ and $s < r$.
(3) $E' = -D$ and $v = r$ if $\delta = \eta$ but $\nu > \eta$, or if $\nu = \delta = \eta$ and $r < s$.
(4) $E' = 0$ if both $\delta > \eta$ and $\nu > \eta$. Then v is irrelevant.

Proof. Examination of (40).

The following two lemmas concerning E' term will be used at the end of the proof, but this is logical place to prove them.

Lemma 18 *$\delta > r\gamma$ and $\nu > s\gamma$.*

Proof. Since the leading term $Et^{\nu}w^{s}$ arose from a term analytic in z and t, when we set $w = zt^{-\gamma}$ the resulting exponent of t, namely $\nu - s\gamma$, must be nonnegative. It has to be strictly positive since it is an imaginary term and there is no imaginary term when $t = 0$. Similarly, $Dt^{\delta}w^{r}$ arose from some term analytic in z and t, hence $\delta - r\gamma > 0$.

Lemma 19 *It is impossible that $E' \neq 0$ and $\eta = v\gamma$.*

Proof. Since $E' \neq 0$, we are in one of the first three cases of Lemma 17. Suppose it is the first case or the third case: then $\eta = \delta$ and $v = r$. By Lemma 18 we have $\delta > r\gamma = v\gamma$. Since $\delta = \eta$ we have $\eta > v\gamma$, contradiction, since by hypothesis $\eta = v\gamma$. Hence the first and third cases of Lemma 17 are impossible. Therefore the second case holds. Then $\eta = \nu$ and $v = s$. By Lemma 18 we have $\nu > s\gamma = v\gamma$. Since $\nu = \eta$ we have $\eta > v\gamma$, contradicting $\eta = v\gamma$. That completes the proof of the lemma.

Multiply (40) by $1 = t^{\eta\gamma}t^{-\eta\gamma}$. Combine the $t^{-\eta\gamma}$ with the difference of sums in (40) and write it as $H(w) + O(t)$. The left side, $\mathrm{Re}(iB/A)$, can be written as $-\mathrm{Im}(B/A)$. Then we have

$$\begin{aligned} -\mathrm{Im}(B/A) &= -t^{k\gamma+\eta}\left(\frac{\mathbf{B}}{\mathbf{A}}\right)(J(w) + O(t)) \\ -\mathrm{Im}(B/A) &= -t^{k\gamma+\eta}\left(\frac{\mathbf{B}}{\mathbf{A}}\right)J(w) + O(t^{k\gamma+\eta+1}) \end{aligned}$$

Since $\mathbf{B}/\mathbf{A} \cong t^{k\gamma}B/A$, we also have

$$-\mathrm{Im}(\mathbf{B}/\mathbf{A}) \cong t^{\eta}J(w)\mathbf{B}/\mathbf{A} \tag{41}$$

The asymptotic behavior of $J(w)$ at infinity is controlled by the term $E'w^v$. If $E' = 0$, for instance, $J(w)$ goes to zero at infinity, since it is a sum of reciprocals of linear functions. More generally, $J(w)$ has the asymptotic behavior $E'w^v$. Hence, we have

Lemma 20 *$\frac{\mathbf{B}}{\mathbf{A}}J(w)$ has the following asymptotic behavior for large w: w^{k-1} if $E' = 0$, or w^{k+v} if E' is not zero.*

Proof. Follows from the asymptotic form of $J(w)$ given above, since $\mathbf{B}/\mathbf{A}$ has the asymptotic form w^k.

Finally we compute the inner product of the unit normal to the surface $u(t,\cdot)$ at a real value of z with the boundary curve at that same point, given by $\Gamma'(\tau(z))$, since $u(z) = \Gamma(\tau(z))$ on the boundary.

$$\begin{aligned} 0 &= \langle \Gamma'(\tau(z)), N(z)\rangle \\ &= \begin{bmatrix} 1 \\ C_1 t^{(2m+1)q\gamma}(\int \mathbf{A}^2 dw)^q + O(t^{(2m+1)q+1)\gamma}) \\ C_2 t^{(2m+1)p\gamma}(\int \mathbf{A}^2 dw)^p + O(t^{(2m+1)p+1)\gamma}) \end{bmatrix} \cdot \begin{bmatrix} -2t^{k\gamma}\mathrm{Im}(\mathbf{B}/\mathbf{A}) \\ 2t^{k\gamma}\mathrm{Re}(\mathbf{B}/\mathbf{A}) \\ -1 \end{bmatrix} \\ &= -2t^{k\gamma}\mathrm{Im}(\mathbf{B}/\mathbf{A}) \\ &+ 2C_1 t^{((2m+1)q+k)\gamma}\mathrm{Re}(\mathbf{B}/\mathbf{A})(\int \mathbf{A}^2 dw)^q + O(t^{(2m+1)q+k+1)\gamma}) \\ &- C_2 t^{(2m+1)p\gamma}(\int \mathbf{A}^2 dw)^p + O(t^{((2m+1)p+1)\gamma}) \end{aligned}$$

Applying the result (41) obtained above for $-\mathrm{Im}(\mathbf{B}/\mathbf{A})$ to the first term, we have

$$
\begin{aligned}
0 &= -2t^{k\gamma+\eta}(\mathbf{B}/\mathbf{A})J(w) \\
&+ 2C_1 t^{((2m+1)q+k)\gamma}\mathrm{Re}(\mathbf{B}/\mathbf{A})(\int \mathbf{A}^2 dw)^q + O(t^{(2m+1)q+k+1)\gamma}) \\
&- C_2 t^{(2m+1)p\gamma}(\int \mathbf{A}^2 dw)^p + O(t^{((2m+1)p+1)\gamma})
\end{aligned}
$$

The first term must cancel with one of the other terms, or be absorbed in the error terms. There are several possibilities, which we consider case by case.

Case 1, All three terms shown explicitly have the same leading power of t. Then we have

$$(2m+1)q + k = (2m+1)p.$$

That is, $k = (2m+1)(p-q)$. Then $k/(2m+1) + q = p$, so $\eta = (2m+1)q\gamma$. Now look at the leading term in w, which is $w^{(2m+1)p}$ in the third term, and $w^{k+(2m+1)q}$ in the second term, and in the first term is either w^{k-1} if $E' = 0$ or w^{k+v} if $E' \neq 0$. If it is w^{k-1} then we have $k-1 = k+(2m+1)q$, which is impossible. Hence it is w^{k+v} and we have $k+v = k+(2m+1)q$, whence $v = (2m+1)q$, and in view of $\eta = (2m+1)q\gamma$, we have $\eta = v\gamma$, contradicting Lemma 19.

Case 2, the second and third terms have the same leading power of t, but the first term has a higher power. Then the second and third terms must cancel. Then again we have

$$(2m+1)q + k = (2m+1)p. \tag{42}$$

Hence $k = (2m+1)(p-q)$. Looking at the coefficients as functions of w and noting that $\mathbf{B}$ and $\mathbf{A}$ are real on the real axis, we have for real w,

$$\frac{C_2}{2C_1}\left(\int_0^w \mathbf{A}^2\, dw\right)^{p-q} = \frac{\mathbf{B}}{\mathbf{A}}$$

Since both sides are meromorphic, this equation is valid in the upper half-plane. By Lemma 14, H must be constant. But this case has already been ruled out in Lemma 16. That completes the proof in Case 2.

Case 3, the first term cancels with the second term. Then

$$k + \eta/\gamma = (2m+1)q + k$$

so $\eta = (2m+1)q\gamma$. There is nothing wrong with this, but then also the coefficients of these terms must cancel:

$$(\mathbf{B}/\mathbf{A})J(w) = \mathrm{Re}(\mathbf{B}/\mathbf{A})(\int \mathbf{A}^2 dw)^q$$

Then since $\mathbf{B}$ and $\mathbf{A}$ are real on the real axis, we have

$$J(w) = (\int \mathbf{A}^2 dw)^q$$

If $E' = 0$, we have an immediate contradiction, since the left hand side tends to zero at infinity, but the right hand side is asymptotic to $w^{(2m+1)q}$.

We may therefore assume E' is not zero. Then $J(w)$ is asymptotic to w^v, so if this is to equal $w^{(2m+1)q}$, we must have $v = (2m+1)q$. Since $\eta = (2m+1)q\gamma$, we have $\eta = v\gamma$. But this is impossible, by Lemma 19. That completes Case 3.

Case 4, the first term cancels with the third term. Then $k+\eta/\gamma = (2m+1)p$. Otherwise put, we have $\eta = ((2m+1)p - k)\gamma$. Again, there is nothing wrong with this equation, but the coefficients must also cancel as functions of w. We obtain

$$-(\mathbf{B}/\mathbf{A})J(w) = (\int \mathbf{A}^2 dw)^p$$

Considering the asymptotic behavior for large w: If $E' \neq 0$, we have w^{k+v} on the left (by Lemma 20) while on the right the function is asymptotic to $w^{(2m+1)p}$. Therefore $k + v = (2m+1)p$, whence $v = (2m+1)p - k$. Hence $\eta = v\gamma$. It follows from Lemma 19 that $E' = 0$. In that case, the asymptotic behavior on the left is w^{k-1}, by Lemma 20. But by Theorem 1, we have $k \leq (2m+1)p$, so the two sides of the equation have different asymptotic behavior, and hence cannot be equal. [11]

Case 5, the first term is absorbed into one of the error terms. Then, however, the second or the third term would have the lowest power of t, and since they do not cancel (that was Case 1), either the second or third term would be the leading term, and since the coefficient is not zero, that is contradictory.

But the first term must either cancel one or both of the other terms or be absorbed in an error term, and since we have ruled out all the possibilities, we have reached a contradiction. QED.

References

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

[2] Beeson, M. Some results on finiteness in Plateau's problem, Part II, *Math Zeitschrift* **181** (1982) 1-30.

[3] Beeson, M., On interior branch points of minimal surfaces, *Math Zeitschrift* **171** (1980) 133-154.

[11] It is an interesting question whether we actually need to appeal to Theorem 1 in Case 4. The right side is analytic, while the left side has a double pole at α_i, providing $\eta_i = \eta$. So if any $\eta_i = \eta$, we are finished with this case. If not, then H is a sum of terms with $w - \beta_i$ in the denominator (and no $w - \alpha_i$ terms), and for each α_i, $H(\alpha_i) = 0$, to cancel the factor of $w - \alpha_i$ in the denominator $\mathbf{A}$. This gives m equations in the $m + k$ numbers β_i, in addition to those already discovered in [2]. We do not know if these equations are new ones or the old ones in a new context. However, we do not need to analyze this complicated situation because we do have Theorem 1.

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

[5] Dierkes, U., Hildebrant, S., Küster, A., Wohlrab, O., *Minimal Surfaces I: Boundary Value Problems*, Springer-Verlag, Berlin/Heidelberg/New York (1992).

[6] Dierkes, U., Hildebrant, S., Küster, A., Wohlrab, O., *Minimal Surfaces II: Boundary Regularity*, Springer-Verlag, Berlin/Heidelberg/New York (1992).

[7] Evans, L., *Partial Differential Equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, R. I. (1998).

[8] Gilbarg, D., and Trudinger, N. S., *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin/Heidelberg/New York (1977).

[9] Gulliver, R. A minimal surface with an atypical boundary branch point, in: B. Lawson and K. Tenenblat (eds.), *Differential Geometry: a Symposium in honor of Manfredo P. do Carmo*, pp. 211-228, Longman, Harlow (1991).

[10] Kato, T., *Perturbation Theory for Linear Operators*, second edition. Springer-Verlag, Berlin/Heidelberg/New York (1980).

[11] Kreyszig, E., *Differential Geometry*, Dover, New York (1991). Originally published: University of Toronto Press, Toronto (1959).

[12] Lewy, H. On the boundary behaviour of minimal surfaces, Proc. Natl. Acad. Sci. USA **37**, 102-110 (1951).

[13] Morrey, C., Jr., *Multiple Integrals in the Calculus of Variations*, Springer-Verlag, Berlin/Heidelberg/New York (1966).

[14] Nitsche, J. C. C. *Lectures on Minimal Surfaces, Volume 1*, Cambridge University Press, Cambridge (1988).

[15] Osserman, R., *A Survey of Minimal Surfaces*, Dover Publications, New York (1986).

[16] Tomi, F., On the local uniqueness of the problem of least area, Arch. Ration. Mech. Anal. **52**, 312-381 (1978).

[17] Tomi, F., and Tromba, A. J., On the structure of the set of curves bounding minimal surfaces of prescribed degeneracy, J. für die reine und angewandte Mathematik **316** (1980) 31-43.

[18] Tromba, A. J., On the number of simply connected minimal surfaces spanning a curve, Mem. Am. Math. Soc. No. 194, **12** (1977).