

SOLUTION OF THE PROBLEM OF PLATEAU*

BY
JESSE DOUGLAS

1. Introduction. The problem of Plateau is to prove the existence of a minimal surface bounded by a given contour. This memoir presents the first solution of this problem for the most general kind of contour: an arbitrary Jordan curve in n -dimensional euclidean space. Topological complications in the contour, as well as the dimensionality n of the containing space, are without consequence for either method or result. Naturally, an arrangement of knots in the contour will produce corresponding complications in the minimal surface, such as self-intersections and branch points.

The method used is entirely novel, representing a complete departure from the classical modes of attack hitherto employed. In this introduction we shall outline three of the classic methods (wherein n is always 3) and, fourth, the method of the present paper, which we believe to furnish the key to the problem. That this is the fact will become even clearer when, in future papers, we apply this method to the case of several contours and of various topological structures of the minimal surface,† for instance, a Möbius leaf with a prescribed boundary.

It is to be signalized that the solution here given is strictly elementary, employing only the most simple and usual parts of analysis, and that the presentation is self-sufficient, requiring no special preliminary knowledge.

(1) First to be considered, in this introductory survey, is the method based on the ideas of Riemann, Weierstrass and Schwarz.‡ Here the given

* This work, in successive stages of its development, was presented to the Society at various meetings from December, 1926, to December, 1929. Abstracts appear in the Bulletin of the American Mathematical Society as follows: vol. 33 (1927), pp. 143, 259; vol. 34 (1928), p. 405; vol. 35 (1929), p. 292; vol. 36 (1930), pp. 49–50, 189–190. Received by the editors, Parts I–IV in August and Part V in December, 1930.

Various phases of the work were also presented in the seminars of Hadamard at Paris (January 18 and December 17, 1929), Courant and Herglotz at Göttingen (June 4, 11, 18; 1929), and Blaschke at Hamburg (July 26, 1929). The second presentation in Hadamard's seminar was in the present form.

† See the abstract *A general formulation of the problem of Plateau*, Bulletin of the American Mathematical Society, vol. 36 (1930), p. 50, which gives methods adequate to solve this most general form of the problem. The cases of two contours, a Möbius leaf, and three contours have already been worked out by the author and the results presented to this Society; abstracts are in the Bulletin of the American Mathematical Society, vol. 36 (1930), p. 797, and vol. 37 (1931).

‡ Riemann, *Werke*, Leipzig, 1892, pp. 301–337, 445–454; Weierstrass, *Werke*, Berlin, 1903, vol. 3, pp. 39–52, 219–220, 221–238; Schwarz, *Gesammelte Mathematische Abhandlungen*, Berlin, 1890, vol. 1. See also Darboux, *Théorie des Surfaces* (2d edition, Paris, 1914), pp. 490–601.

contour is a polygon Π . The problem is made to depend on a linear differential equation of second order

$$(1.1) \quad \frac{d^2\theta}{dw^2} + p \frac{d\theta}{dw} + q\theta = 0$$

where the coefficients p, q are rational functions of the complex variable w with, at first, undetermined coefficients. The monodromy group G of this equation (this is the group of linear transformations undergone by a fundamental set of solutions $\theta_1(w), \theta_2(w)$ when w performs circuits about the singular points of the equation) is known as soon as the polygonal contour is given. The monodromy group problem of Riemann is to determine the coefficients in p, q so that (1.1) shall have the prescribed monodromy group G . But the solution of this problem is not all that is required, for that gives only a minimal surface whose polygonal boundary Π_1 has its sides parallel to those of Π . It remains further to arrange that the sides of Π_1 shall have the same lengths as those of Π . All this is reduced by Riemann and Weierstrass to a complicated system of transcendental equations in the coefficients of p, q , which they and Schwarz succeed in solving only in special cases.

To these ideas attaches the solution given by R. Garnier for the problem of Plateau. In a preliminary memoir* he first gives his form of solution of the Riemann monodromy group problem, previously solved by Hilbert, Plemelj and Birkhoff. In a following memoir† he deals with the supplementary conditions relating to lengths of sides of the polygonal boundary, and concludes the existence of a solution of the above mentioned system of transcendental equations. The Plateau problem being thus solved for a polygon, Garnier passes to the case of a more general contour Γ by regarding Γ as a limit of polygons Π . He shows that the solution of the Riemann group problem with the supplementary conditions varies continuously with the data, so that the minimal surface determined by Π approaches to a minimal surface bounded by Γ . To insure the validity of the limit process, Γ is restricted to have bounded curvature by segments.

Subsequent to the presentation by the present writer to this Society‡ of a series of papers containing the substance of the memoir at hand, T. Radó published a note§ showing how the part of Garnier's work concerned with the passage from polygons to more general contours could be materially simpli-

* Annales Scientifiques de l'Ecole Normale Supérieure, (3), vol. 43 (1926), pp. 177-307.

† *Le problème de Plateau*, *ibid.*, vol. 45 (1928), pp. 53-144.

‡ Annual meeting, at Bethlehem, Pa., December 27, 1929.

§ Proceedings of the National Academy of Sciences, vol. 16 (1930), pp. 242-248.

fied, and rendering the restriction on Γ less stringent by requiring only rectifiability.

*(1') Later, after the dispatch to the editors of this journal of the manuscript of the present paper, two other papers by Radó† appeared in which he gives a solution of the Plateau problem, in the first for a rectifiable contour, and in the second for any contour capable of spanning a finite area (three-dimensional space). As its author states, this work is a continuation of the classic ideas, being based on the least-area characterization of a minimal surface and the theory of conformal mapping, especially the Koebe theory—relating to abstract Riemann manifolds—in the form of the theorem that it is possible to map any simply-connected polyhedral surface conformally on the interior of a circle, the map remaining one-one and continuous as between the boundary of the polyhedral surface and the circumference of the circle.

The present work, on the other hand, apart from its advantage of complete generality of the contour, breaks completely with the hitherto classic methods, replacing the area functional by an entirely new and much simpler functional, and carrying through the existence proof without assuming any of the theory of conformal mapping even for ordinary plane regions; on the contrary, in Part IV our results are applied to give new proofs of the classic theorems concerning conformal mapping essentially simpler than the classic proofs, a demonstration of the superior fundamental character of the present mode of attack.

In Part V, after the existence of the minimal surface has been established, the Koebe mapping theorem plays a rôle beside the formulas of Part III in a brief proof of the least-area property. However, we regard this treatment only as a stop-gap, having under development a disposal of the least-area part of the problem not using the Koebe or any other conformal mapping theorem. Such an independent treatment is desirable because the Koebe theory of conformal mapping is comparable in difficulty with the Plateau problem. The avoidance of the former will bring the solution of the least area problem to rest directly on the axioms of analysis, as has already been done in this paper with the proof of the existence of the minimal surface.

(2) A second class of methods is based on the partial differential equation of minimal surfaces (Lagrange)

$$(1.2) \quad (1 + q^2)r - 2pqs + (1 + p^2)t = 0.$$

* Article (1') added in proof.

† *Annals of Mathematics*, (2), vol. 31 (1930), pp. 457–469. *Mathematische Zeitschrift*, vol. 32 (1930), pp. 762–796.

Here belongs the work of S. Bernstein* and Ch. H. Müntz.† The surface is assumed in the restricted form $z=f(x, y)$, and the problem is regarded as a generalized Dirichlet problem, with (1.2) replacing Laplace's equation. Besides the restriction on the representation of the surface, it is assumed that the contour has a convex projection on the xy -plane. The work of Müntz has been criticized by Radó.‡

(3) Minimal surfaces first presented themselves in mathematics, and were named, by their property of having least area among all surfaces bounded by a given contour:

$$(1.3) \quad \iint (1 + p^2 + q^2)^{1/2} dx dy = \text{minimum}.$$

It is in this way that minimal surfaces appear in the pioneer memoir of Lagrange§ on the calculus of variations for double integrals.

In recent years A. Haar|| has treated the Plateau problem from this point of view, using the direct methods of the calculus of variations introduced by Hilbert. Haar assumes the surface in the form $z=f(x, y)$ and the contour subject to the following restriction: any plane containing three points of the contour has a slope with respect to the xy -plane that is less than a fixed finite upper bound, an assumption occurring first in the work of Lebesgue.¶

(4) The method of the present memoir is as follows. The contour Γ being taken as any Jordan curve in euclidean space of n dimensions, we consider the class of all possible ways of representing Γ as topological image of the unit circle C :

$$x_i = g_i(\theta) \quad (i = 1, 2, \dots, n).$$

This class forms an L -set in the sense of Fréchet's thesis,** and is compact and closed after "improper" topological representations of Γ have been adjoined: these are limits of proper ones and cause arcs of Γ to correspond to single points of C , or vice-versa (§3). The principal idea is then to introduce the functional (§5)

* *Mathematische Annalen*, vol. 69 (1910), pp. 82-136, especially § 18.

† *Mathematische Annalen*, vol. 94 (1925), pp. 53-96.

‡ *Mathematische Annalen*, vol. 96 (1927), pp. 587-596.

§ *Miscellanea Taurinensia*, vol. 2 (1760-61); also *Oeuvres*, vol. 1, p. 335.

|| *Ueber das Plateausche Problem*, *Mathematische Annalen*, vol. 97 (1927), pp. 124-158.

¶ *Intégrale, longueur, aire*, *Annali di Matematica*, (3), vol. 7, pp. 231-359; see chapter VI, especially p. 348.

** *Sur quelques points du calcul fonctionnel*, *Rendiconti del Circolo Matematico di Palermo*, vol. 22 (1906), pp. 1-74.

$$(1.4) \quad A(g) = \frac{1}{4\pi} \int_c \int_c \frac{\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2}{4 \sin^2 \frac{\theta - \phi}{2}} d\theta d\phi,$$

where the integrand has the following simple geometric interpretation: square of chord of contour divided by square of corresponding chord of the unit circle. This improper double integral has a determinate positive value, finite or $+\infty$, for every representation g . It is readily shown (§9) that $A(g)$ is lower semi-continuous; therefore, by a theorem of Fréchet† to the effect that a lower semi-continuous functional on a compact closed set attains its minimum value, the minimum of $A(g)$ is attained for a certain representation $x_i = g_i^*(\theta)$. If

$$(1.4_1) \quad x_i = H_i(u, v) = \Re F_i(w) \quad (w = u + iv)$$

are the harmonic functions in the interior of the unit circle determined according to Poisson's integral by the boundary functions

$$(1.4_2) \quad x_i = g_i^*(\theta)$$

it is then proved (§§11–16) that

$$(1.4_3) \quad \sum_{i=1}^n F_i'^2(w) = 0;$$

briefly speaking, this condition expresses that the first variation of $A(g)$ vanishes for the minimizing $g = g^*$. According to the formulas of Weierstrass, the condition (1.4₃) expresses that (1.4₁) defines a minimal surface. After it has been shown (§§17, 18) that g^* is a proper representation of Γ , it follows that this minimal surface is bounded by Γ , since by the properties of Poisson's integral the functions (1.4₁) then attach continuously to the boundary values (1.4₂).

One consideration is necessary to validate the preceding argument: we must be sure that $A(g)$ is not identically $= +\infty$, that it takes a finite value for some g . This is what makes it necessary to divide the discussion into two parts. In Part I we assume that there exists a parametric representation g of the given contour such that $A(g)$ is finite. It will be seen from Parts III and V that this means, more concretely, that it is possible to span some surface of finite area in the given contour. A sufficient condition for the property "there exists a g for which $A(g)$ is finite" (which, in anticipation of the discussion of Parts III and V, we will call the finite-area-spanning property) is that

† Loc. cit., §11, p. 9.

the contour be rectifiable. For if the contour have length L , and we choose as parameter $\theta = 2\pi s/L$, s being arc-length reckoned from any fixed initial point, then it will be readily seen from the fact that a chord is not greater than its arc that the integrand in $A(g)$ stays bounded, hence $A(g)$ is finite for this parameter. In particular, every polygon has the finite-area-spanning property. That, however, a finite-area-spanning contour is superior in generality to a rectifiable contour may be seen by taking any simply-connected portion of a surface, having finite area, and drawing upon it any non-rectifiable Jordan curve (e.g., a non-rectifiable Jordan curve on a sphere).

Part II deals with the case of an arbitrary Jordan contour, where generally $A(g) \equiv +\infty$, meaning that no finite area whatever can be spanned in the given contour; an example of such a contour is given in §27. The existence theorem is extended to a contour of this type by an easy limit process, where in the given contour is regarded as a limit of polygons. In this case the minimal surface can be defined only by the Weierstrass equations, the least-area characterization becoming meaningless.†

The distinctive feature of the present work is the determination of the minimal surface by the minimizing of the functional $A(g)$, decidedly simpler of treatment than the classic area functional. $A(g)$ has a simple relation to the area functional, dealt with in Part III. If $S(g)$ denote the area of the surface $x_i = H_i(u, v)$, these being the harmonic functions in $u^2 + v^2 < 1$ determined by the boundary functions $x_i = g_i(\theta)$, then $A(g) \geq S(g)$, and the relation of equality holds when and only when the surface is minimal. Thus $A(g)$, not equal to area in general, is capable of giving information about area in the case of a minimal surface. Part III, moreover, provides the basis for the easy proof in Part V that the minimal surface whose existence is proved in Parts I and II has the least area of any surface bounded by Γ .

An interesting and important consideration, unremarked before the writer's work, is that *the Riemann conformal mapping problem is included as the special case $n=2$ in the problem of Plateau*. The Riemann mapping theorem relating to the interiors of two Jordan regions is supplemented by the theorem of Osgood‡ and Carathéodory§ to the effect that the conformal correspondence between the interiors induces a topological correspondence between the boundaries. In Part IV a proof is given of the combined theorems of Riemann and Osgood-Carathéodory, independent of any previous treatment, and more elementary and perspicuous.

† But see the footnote at the end of this paper.

‡ Osgood and E. H. Taylor, *Conformal transformations on the boundaries of their regions of definition*, these Transactions, vol. 14 (1913), pp. 277–298.

§ Carathéodory, *Ueber die gegenseitige Beziehung der Ränder bei der konformen Abbildung des Innern einer Jordanschen Kurve auf einen Kreis*, Mathematische Annalen, vol. 73 (1913), pp. 305–320.

Examples show that the solution of the Plateau problem may not be unique.† The question of the degree of multiplicity of the solution is not dealt with here. As indicated above, the minimal surface whose existence is here assured is the one which furnishes an absolute minimum for the area.

2. Formulation. For definition of a minimal surface we adopt the formulas given, for $n=3$, by Weierstrass:

$$(2.1) \quad x_i = \Re F_i(w) \quad (i = 1, 2, \dots, n)$$

with

$$(2.2) \quad \sum_{i=1}^n F_i'^2(w) = 0.$$

The problem of Plateau may then be formulated precisely as follows.

Given any contour Γ in the form of a Jordan curve in euclidean space of n dimensions. To prove the existence of n functions F_1, F_2, \dots, F_n of the complex variable w , holomorphic in the interior of the unit circle C , satisfying there the condition

$$\sum_{i=1}^n F_i'^2(w) = 0$$

identically, and whose real parts

$$x_i = \Re F_i(w)$$

attach continuously to boundary values on C

$$x_i = g_i(\theta)$$

which represent Γ as a topological image of C .

As defined by (2.1), (2.2), the minimal surface appears in a representation on the circular region $|w| < 1$ which is conformal except at those (necessarily isolated) points where simultaneously

$$F_1'(w) = 0, \quad F_2'(w) = 0, \quad \dots, \quad F_n'(w) = 0.$$

† The following example was communicated to the writer by N. Wiener. Two co-axial circles may be so placed that the area of the catenoidal segment determined by them is greater than the sum of the areas of the two circles (Goldschmidt discontinuous solution). Consider the contour formed by two meridians of the catenoid, very close together, and the arcs remaining on the circles after the small arcs intercepted between the meridians have been removed. One solution of the Plateau problem for this contour is the intercepted part of the catenoid. But the surface formed of the two circles and the narrow catenoidal strip between the meridians has a smaller area. Consequently, there will be a second minimal surface bounded by the given contour, varying slightly from the surface just described: this second surface will have the absolutely least area.

On this remark is based the inclusion of the Riemann mapping problem in the Plateau problem as the special case $n=2$. We show that for $n=2$ the conformality is free of singular points, but for $n>2$ their absence cannot be guaranteed.†

I. A FINITE-AREA-SPANNING CONTOUR

Hypothesis. Part I is based on the hypothesis that there exists a parametric representation g of the given contour for which $A(g)$ is finite.

3. Topological correspondences between Γ and C . Γ may be supposed

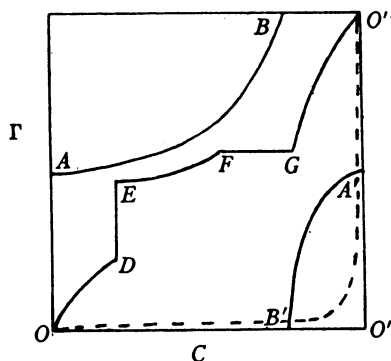


FIG. 1. Torus of representation $R = \Gamma C$

given in some initial representation $x_i = f_i(t)$, from which its most general representation may be derived by a relation $t = t(\theta)$ defining a one-one continuous transformation of C into itself. The two-dimensional manifold (t, θ) of pairs of points one on Γ , one on C , forms a torus ΓC , which will be called the *torus of representation* and denoted by R . This torus is depicted in the annexed figure as a square where points opposite one another on parallel sides, such as A and A' , B and B' , are to be regarded as identical. Rectilinear transversals of the square parallel to C will be termed *parallels*, those parallel to Γ *meridians*. A topological correspondence between Γ and C is represented by a continuous closed curve, such as $ABB'A'$, which is intersected in one and only one point by each parallel and by each meridian; such a curve may be described as *cyclically monotonic*. We will denote by \mathfrak{P} the totality of these curves, which, we will say, represent *proper* topological correspondences between Γ and C . In the corresponding equations $x_i = g_i(\theta)$ of Γ the functions g_i are continuous, and are not all constant on any arc of C .

† Example: $x_1 = \Re w^2$, $x_2 = \Re -iw^2$, $x_3 = \Re w^3$, $x_4 = \Re -iw^3$; $|w| \leq 1$. This is a minimal surface bounded by the contour $x_1 = \cos 2\theta$, $x_2 = \sin 2\theta$, $x_3 = \cos 3\theta$, $x_4 = \sin 3\theta$. Neither the minimal surface nor the contour is self-intersecting. The representation on $|w| < 1$ is conformal except at the origin, where angles are multiplied by 2.

A disadvantage in dealing with \mathfrak{P} is that it is not a closed set: a sequence of curves of \mathfrak{P} may converge to a limit not belonging to \mathfrak{P} . For instance, we may obtain as limit of curves of \mathfrak{P} a curve such as $ODEFGO''$, containing a segment of meridian DE or a segment of parallel FG , as well as properly monotonic arcs. An extreme case is that indicated by the dotted curve, where the limit is $OO'O''$, consisting of a parallel together with a meridian.

A correspondence between Γ and C whose graph contains, besides properly monotonic arcs, a meridian-segment less than an entire meridian, or a parallel-segment less than an entire parallel, will be called an *improper* topological correspondence.

In the graph $ODEFGO''$, the meridian-segment DE represents an arc $Q'Q''$ of Γ which corresponds to a single point P of C . In the corresponding equations of Γ , the functions g_i have one-sided limits at P equal respectively to the coördinates of Q' and Q'' ; and at least one of these functions is discontinuous (the vector g is discontinuous) since, Γ having by hypothesis no double points, Q' and Q'' are distinct. A monotonic function has at most a denumerable infinity of discontinuities, each in the form of distinct one-sided limits; therefore in an improper representation the functions g_i have at most a denumerable infinity of discontinuities, all of the so-called first kind, that is, where one-sided limits exist but are unequal. This observation will assure us in §5 of the existence for an improper representation of the Riemann integrals used in defining $A(g)$.

The parallel-segment FG represents an arc $P'P''$ of C which corresponds to a single point Q of Γ . The functions g_i are constant on the arc $P'P''$, where they are equal to the coördinates of Q .

The class of all improper representations of Γ will be denoted by \mathfrak{I} , and will be divided according to the above description into the two sub-classes \mathfrak{I}_1 and \mathfrak{I}_2 , *not mutually exclusive*:

\mathfrak{I}_1 , *improper representations of the first kind*, in which an arc of Γ less than all of Γ corresponds to a single point of C .

\mathfrak{I}_2 , *improper representations of the second kind*, in which an arc of C less than all of C corresponds to a single point of Γ .

Special attention must now be given to the correspondences between Γ and C whose graph consists of a parallel together with a meridian, such as $OO'O''$. Here the whole of Γ corresponds to a single point of C , and the whole of C to a single point of Γ . Such a representation will be termed *degenerate*; there are evidently ∞^2 degenerate representations, obtained by varying the distinguished points on Γ and C . In the corresponding equations of Γ the functions g_i reduce to constants. *The functional $A(g)$ will not be defined for the degenerate representations.*

Three fixed points. After having established in §6 a certain invariance property of $A(g)$, we shall be led to consider the class of those proper or improper representations of Γ wherein three distinct fixed points P_1, P_2, P_3 , of C , correspond to three distinct fixed points Q_1, Q_2, Q_3 , of Γ . These representations are pictured on the torus $R = \Gamma C$ by proper or improper cyclically monotonic curves passing through three fixed points no two of which lie on the same parallel or the same meridian.

The preceding discussion leads us to distinguish the following classes of representations of Γ on C .

- (1) The class of all representations: proper, improper, and degenerate,

$$\mathfrak{R} = \mathfrak{P} + \mathfrak{I} + \mathfrak{D}.$$

- (2) The class of all proper and improper representations:

$$\mathfrak{M} = \mathfrak{P} + \mathfrak{I}.$$

This is not a closed set, since a sequence of representations of \mathfrak{M} may tend to a degenerate representation as limit. \mathfrak{M} will serve as the range of the argument in the functional $A(g)$.

- (3) The class of all proper and improper representations whereby three distinct fixed points of Γ correspond to three distinct fixed points of C :

$$\mathfrak{M}' = \mathfrak{P}' + \mathfrak{I}'.$$

It is important to observe the following two properties of \mathfrak{M}' : it is closed; it does not contain any degenerate representation.

4. Harmonic surfaces. Each representation $x_i = g_i(\theta)$ of Γ determines a surface $x_i = \Re F_i(w)$, where the harmonic functions $\Re F_i(w)$ are those defined by Poisson's integral based on the respective boundary functions $g_i(\theta)$. We will refer to this surface as *the harmonic surface determined by the representation g* .

The limit of Poisson's integral when w approaches to a point θ of C where $g_i(\theta)$ is continuous is $g_i(\theta)$. If $g_i(\theta)$ has unequal one-sided limits at the point θ , then the limiting value of Poisson's integral in the approach of w to θ varies between these one-sided limits in a manner that depends linearly on the angle made by the direction of approach with the radius to the point θ .†

It follows that the harmonic surface determined by any proper representation of Γ is bounded by Γ . For an improper representation of the first kind, where the point P of C corresponds to the arc $Q'Q''$ of Γ , it is evident that the boundary points of the harmonic surface obtained by allowing w to approach

† A result due to Schwarz; cf. Picard, *Traité d'Analyse*, vol. 1 (3d edition, Paris, 1922), pp. 315-319.

to P in all the possible directions form the *chord* $Q'Q''$. In an improper representation of the second kind, the point w approaching to any point of an arc $P'P''$ of C gives the same boundary point Q for the harmonic surface.

It is to be observed from this that in the case of an improper representation of the first kind the corresponding harmonic surface will not be bounded by Γ , but by a curve derived from Γ by replacing certain of its arcs (at most a denumerable infinity) by their chords, which chords will correspond to single points of C . In the case of an improper representation of the second kind, the boundary point corresponding to each arc $P'P''$ lies on Γ , but then Γ is not in one-one relation with C .

Example 1. The graph of the correspondence between Γ and C may consist of k parallel-segments alternating with k meridian-segments. The boundary of the corresponding harmonic surface is a polygon of k sides and k vertices inscribed in Γ . The sides of the polygon correspond respectively to k points of C , and the vertices to the k arcs into which these points divide C . If $k=2$, the harmonic surface reduces to a chord of Γ . If $k=1$, the case of a degenerate representation, the harmonic surface reduces to a point of Γ .

Example 2. The correspondence $t=t(\theta)$ between Γ and C may be defined by the frequently cited monotonic function based on Cantor's perfect set.† Here the boundary of the harmonic surface consists of a denumerable infinity of chords of Γ together with the nowhere-dense perfect set of points of Γ which remain after the arcs of these chords have been removed. On C we have an everywhere-dense denumerable infinity of points of discontinuity of $x_i = g_i(\theta)$, corresponding respectively to the above-mentioned chords of Γ .

It will be seen from these examples that the harmonic surface determined by a given representation of Γ cannot be regarded as bounded by Γ unless this representation is proper. It is for this reason that after establishing the existence of a representation $x_i = g_i^*(\theta)$ such that the corresponding harmonic surface obeys the condition $\sum_{i=1}^n F_i'^2(w) = 0$, it is necessary (as is done in §§17, 18) to prove that the representation g^* is proper before we can say we have a minimal surface bounded by Γ .

5. The fundamental functional $A(g)$. The functional $A(g)$ is defined on the set $\mathfrak{M} = \mathfrak{P} + \mathfrak{I}$ of all proper and improper representations of Γ by the formula

$$(5.1) \quad A(g) = \frac{1}{16\pi} \int_c \int_c \frac{\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2}{\sin^2 \frac{\theta - \phi}{2}} d\theta d\phi.$$

† See Carathéodory, *Vorlesungen über reelle Funktionen*, Leipzig, 1918, §156, p. 159.

The domain of integration CC is a torus, which will be denoted by T . This *torus of integration* is to be carefully distinguished from the *torus of representation* R of §3.

The integrand is defined everywhere on T except on the *diagonal* $\theta = \phi$, where it takes the indeterminate form $0/0$. Let us isolate the diagonal from the rest of the torus by means of two regular curves† symmetrically disposed

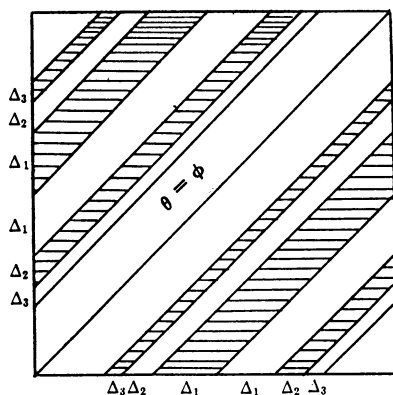


FIG. 2. Torus of integration $T = CC$

on either side of it, enclosing a region τ_1 which we delete from the torus, leaving $T_1 = T - \tau_1$.

In case g is proper, or improper only of the second kind, the integrand of (5.1) is defined and continuous on T_1 .

If g is improper of the first kind, the integrand is discontinuous at the points of certain parallels and meridians, symmetrically disposed with respect to the diagonal, and at most denumerably infinite in number. At a point belonging to a parallel but not to a meridian of discontinuity (or vice-versa), the discontinuity is in the form of distinct limits‡ according as the point is approached from one side or the other of the parallel (or meridian). At a point of intersection of a parallel of discontinuity with a meridian of discontinuity, there are four distinct limits‡ according to the quadrant within which the point is approached.

In any event, the discontinuities of the integrand in T_1 form at most a set of zero measure.

Let d denote the diameter (greatest chord) of Γ , and $\delta > 0$ the smallest value of $|\theta - \phi|$ in T_1 ; then the integrand of (5.1) is bounded on T_1 , being

† The only curves which will come into consideration in this connection will be straight lines parallel to the diagonal and images of them by the regular analytic transformations (6.1), (12.1).

‡ Some of these may accidentally be equal.

$$\leq \frac{d^2}{\sin^2 \frac{\delta}{2}}.$$

These two remarks insure the existence of the Riemann integral taken over T_1 of the integrand of (5.1).

Imagine now an infinite sequence of regions

$$\tau_1, \tau_2, \dots, \tau_r, \dots$$

each contained in the preceding and shrinking to the diagonal as limit, so that the complementary regions

$$T_1, T_2, \dots, T_r, \dots$$

swell continually and tend to the entire torus T as limit. Then, because every element of the integral (5.1) is positive (wide sense) the proper Riemann integrals

$$\iint_{T_1}, \iint_{T_2}, \dots, \iint_{T_r}, \dots$$

form a continually increasing† sequence of positive‡ numbers. Hence they approach either to a finite positive limit or to $+\infty$; and *this limit is by definition* $A(g)$, which thus appears as an improper integral. The same fact of the positivity of each element proves easily that the value obtained for $A(g)$ is independent of the particular sequence of regions τ used in its definition; in fact, $A(g)$ may be defined uniquely as the upper bound of the integral over any region of T to which the diagonal is exterior.

A(g) as an infinite series. For greater definiteness in determining $A(g)$, we proceed to divide the torus T into an infinite number of strips (Fig. 2) by means of the lines

$$|\theta - \phi| = \frac{\pi}{r} \quad (r = 1, 2, 3, \dots). \ddagger$$

The region defined by the inequality

$$(5.2) \quad \frac{\pi}{r+1} \leq |\theta - \phi| \leq \frac{\pi}{r},$$

† That “increasing” and “positive” may here be taken in the strict sense follows by the same proof given a little later on to show that $\Delta_r(g)$ is strictly positive. The only assumption to be made is that g is not a degenerate representation.

‡ By $|\theta - \phi|$ we shall understand the minor arc intercepted between the points θ and ϕ on the unit circle C .

consisting of a pair of strips symmetric with respect to the diagonal, will be denoted by Δ_r . We then define the functional

$$(5.3) \quad \Delta_r(g) = \frac{1}{16\pi} \iint_{\Delta_r} \frac{\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2}{\sin^2 \frac{\theta - \phi}{2}} d\theta d\phi.$$

This is a proper Riemann integral since the integrand stays bounded on Δ_r , being

$$\leq \frac{d^2}{\sin^2 \frac{\pi}{2(r+1)}}.$$

[Certainly $\Delta_r(g) \geq 0$; but it is interesting (though not necessary for the sequel) to prove the strict inequality $\Delta_r(g) > 0$.

First we see that we could have $\Delta_r(g) = 0$ only by having $\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2$ identically zero in Δ_r . For if $\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 = p > 0$ at some interior point† of Δ_r , it would be $> p/2$ in at least a sufficiently small square in the corner of one of the quadrants about this point; therefore this square makes a contribution > 0 to the value of the integral, which cannot be neutralized by the non-negative contribution of the other elements; hence $\Delta_r(g) > 0$, contrary to the hypothesis $\Delta_r(g) = 0$.

Thus $\Delta_r(g) = 0$ implies $g_i(\theta) = g_i(\phi)$ ($i = 1, 2, \dots, n$) for all (θ, ϕ) obeying (5.2). Now it is easy to see that, for every ϕ_1, ϕ_2 such that

$$(5.4) \quad |\phi_1 - \phi_2| \leq \frac{\pi}{r} - \frac{\pi}{r+1},$$

a θ exists such that (θ, ϕ_1) and (θ, ϕ_2) obey (5.2). On account of the transitivity of the relation $g_i(\theta) = g_i(\phi)$, it follows that $\Delta_r(g) = 0$ implies $g_i(\phi_1) = g_i(\phi_2)$ for all (ϕ_1, ϕ_2) obeying (5.4).

Now any two points whatever θ, ϕ of C can be made the first and last of a finite sequence $\theta, \phi_1, \phi_2, \dots, \phi_m, \phi$ any two consecutive terms of which obey (5.4). Consequently $\Delta_r(g) = 0$ implies $g_i(\theta) = g_i(\phi)$ for all θ, ϕ ; but then g would be a degenerate representation; however, such are excluded from the range of the argument in $A(g)$.]

The functional $A(g)$ may now be defined as the finite or positively infinite sum of the infinite series of positive terms

† At a point of discontinuity this will mean that one of the two or four limiting values is equal to p

$$(5.5) \quad A(g) = \Delta_1(g) + \Delta_2(g) + \cdots + \Delta_r(g) + \cdots.$$

6. The invariance of $A(g)$. This section is devoted to proving the important observation that $A(g)$ is invariant under the transformation

$$(6.1) \quad \tan \frac{\bar{\theta}}{2} = \frac{a \tan \frac{\theta}{2} + b}{c \tan \frac{\theta}{2} + d} \quad (a, b, c, d \text{ real; } ad - bc \neq 0)$$

of the circumference of the unit circle into itself.

In the expression (5.1) for $A(g)$ make the substitution

$$(6.2) \quad x = \tan \frac{\theta}{2}, \quad y = \tan \frac{\phi}{2} \cdot \dagger$$

We have

$$\begin{aligned} \frac{d\theta d\phi}{\sin^2 \frac{\theta - \phi}{2}} &= \frac{d\theta d\phi}{\left(\sin \frac{\theta}{2} \cos \frac{\phi}{2} - \cos \frac{\theta}{2} \sin \frac{\phi}{2} \right)^2} \\ &= \frac{\sec^2 \frac{\theta}{2} d\theta \cdot \sec^2 \frac{\phi}{2} d\phi}{\left(\tan \frac{\theta}{2} - \tan \frac{\phi}{2} \right)^2} = \frac{4dx dy}{(x - y)^2}. \end{aligned}$$

When θ and ϕ vary independently over C from $-\pi$ to $+\pi$, x and y vary independently from $-\infty$ to $+\infty$; so that, denoting by $h_i(x)$, $h_i(y)$ the functions which result from $g_i(\theta)$, $g_i(\phi)$ by the substitutions inverse to (6.2):

$$h_i(x) = g_i(2 \arctan x), \quad h_i(y) = g_i(2 \arctan y),$$

we have as transformed expression for $A(g)$:

$$(6.3) \quad A(g) = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \sum_{i=1}^n [h_i(x) - h_i(y)]^2 \frac{dx dy}{(x - y)^2}.$$

In terms of the variables (6.2), the transformation (6.1) is

$$(6.4) \quad \bar{x} = \frac{ax + b}{cx + d}, \quad \bar{y} = \frac{ay + b}{cy + d}.$$

\dagger This may be interpreted as replacing the unit circle by a half-plane, its circumference by the edge of the half-plane.

Let $\bar{g}_i(\bar{\theta})$, $\bar{h}_i(\bar{x})$ be the functions that result from $g_i(\theta)$, $h_i(x)$ by the respective transformations (6.1), (6.4); then

$$(6.5) \quad A(\bar{g}) = \frac{1}{4\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \sum_{i=1}^n [\bar{h}_i(\bar{x}) - \bar{h}_i(\bar{y})]^2 \frac{d\bar{x}d\bar{y}}{(\bar{x} - \bar{y})^2}.$$

The domain of integration in (6.3) is the entire real plane (x, y) , and the entire real plane (\bar{x}, \bar{y}) is the domain of integration in (6.5), because (6.4) sets up a one-one correspondence between all finite and infinite real values of x and of \bar{x} , and the same for y and \bar{y} .

The following simple calculations now lead to the desired result:

$$\begin{aligned} d\bar{x} &= \frac{ad - bc}{(cx + d)^2} dx, & d\bar{y} &= \frac{ad - bc}{(cy + d)^2} dy; \\ \bar{x} - \bar{y} &= \frac{(ad - bc)(x - y)}{(cx + d)(cy + d)}; \\ \frac{d\bar{x}d\bar{y}}{(\bar{x} - \bar{y})^2} &= \frac{dxdy}{(x - y)^2}. \end{aligned}$$

Therefore, by comparison of (6.5) with (6.3),

$$A(\bar{g}) = A(g).$$

Equivalent representations. Three fixed points. Let us designate as *equivalent* any two representations g , \bar{g} which are related to one another by a transformation (6.1). On account of the presence of three essential parameters in the transformation, it is possible to find for every non-degenerate representation g an equivalent \bar{g} which causes any chosen three distinct points of C to correspond to any chosen three distinct points of Γ . In the notation of §3: every element of \mathfrak{M} has an equivalent in \mathfrak{M}' .

Since $A(g)$ has the same value for two equivalent g 's, it follows that the lower bound M of $A(g)$ on \mathfrak{M} is equal to the lower bound M' of $A(g)$ on \mathfrak{M}' ; and if we can prove that M' is attained for a certain \bar{g} , then obviously M is attained for every g equivalent to \bar{g} . As already pointed out, the advantage of referring from \mathfrak{M} to \mathfrak{M}' is that the latter is a closed set not containing any degenerate representation, while the former is an open set, having the degenerate representations among its limit elements.

7. Fréchet's thesis; compact closed sets. We now have the problem of proving that $A(g)$ attains its minimum on \mathfrak{M}' . This will be accomplished by means of the simple and general ideas of Fréchet's thesis,[†] which concerns real-valued functions on sets of a very general nature.

[†] For reference, see the Introduction.

Suppose we have a set \mathfrak{A} of elements a of any nature. Fréchet terms this an L -set under the following conditions. Every infinite sequence

$$a_1, a_2, \dots, a_m, \dots$$

of elements of \mathfrak{A} is definitely designated as *convergent* or *divergent*, and with each convergent sequence there is associated a unique element of \mathfrak{A} called the *limit* of the sequence. Every sub-sequence of a convergent sequence must itself be convergent and have the same limit as the original sequence. Every infinite sequence all of whose elements are identical with the same element of \mathfrak{A} is convergent and has this element for limit.

An L -set is termed *compact* if it obeys the Bolzano-Weierstrass theorem: every infinite sub-set of the given set contains a convergent sequence of distinct elements, or has a limit element.

An L -set is termed *closed* provided every limit of an infinite sequence of elements of the set belongs to the set. Evidently the notion of closure has meaning only when the given set is considered as part of a larger set.

An L -set both compact and closed is termed by Fréchet an *extremal* set, because the fundamental theorem of Weierstrass that a continuous function on a closed interval in the ordinary real domain attains its extreme values, maximum and minimum, applies to such a set. This is easily proved in the cited thesis after the following definition of continuous function on an L -set has been given.

A real-valued function $U(a)$ of the elements of a set \mathfrak{A} is termed *continuous* if whenever the sequence of elements

$$a_1, a_2, \dots, a_m, \dots$$

converges to the element a as limit, then the sequence of functional values

$$U(a_1), U(a_2), \dots, U(a_m), \dots$$

converges in the ordinary sense to the value $U(a)$. This means, of course, that for every $\epsilon > 0$ there is an index m_ϵ such that, for every $m > m_\epsilon$,

$$U(a) - \epsilon < U(a_m) < U(a) + \epsilon.$$

If we require only the first of these two inequalities, we have the notion of lower semi-continuity: a function $U(a)$ is *lower semi-continuous* if, with the conventions of the preceding paragraph, we have

$$(7.1) \quad U(a_m) > U(a) - \epsilon$$

for $m > m_\epsilon$.

An alternative statement is:

$$(7.2) \quad \begin{array}{ll} \text{if} & a_1, a_2, \dots, a_m, \dots \rightarrow a, \\ \text{and} & U(a_1), U(a_2), \dots, U(a_m), \dots \rightarrow L, \\ \text{then} & U(a) \leq L, \end{array}$$

a condition also expressed in the form

$$U(a) \leq \liminf_{m \rightarrow \infty} U(a_m).$$

It is shown by Fréchet[†] that a lower semi-continuous function on a compact closed L -set attains its minimum value. Our proof that $A(g)$ attains its minimum will be a particular application of this general theorem. However, for the sake of completeness, we will not assume this easily proved theorem, but shall establish it in §10 with the actual set \mathfrak{M}' and functional $A(g)$ here under consideration.

8. The topological correspondences between Γ and C as an L -set. With a natural definition of limit, the set \mathfrak{R} of all (proper, improper, and degenerate) representations of Γ as topological image of C is an L -set, as are also its sub-sets \mathfrak{M} and \mathfrak{M}' . We will say, namely, that a sequence of representations

$$g^{(1)}, g^{(2)}, \dots, g^{(m)}, \dots$$

converges to a certain representation g as limit when the graphs of $g^{(1)}, g^{(2)}, \dots, g^{(m)}, \dots$ on the torus $R = \Gamma C$ (Fig. 1) converge in the ordinary sense to the graph of g ; this will mean that if R_ϵ denote the region covered by a circle of radius ϵ whose center describes the graph of g , then for all sufficiently large values of m the graph of $g^{(m)}$ lies within R_ϵ .

The L -set \mathfrak{R} (and, automatically, its sub-sets \mathfrak{M} and \mathfrak{M}') has the important property of *compactness*, the proof of which results directly from the following theorem:[‡]

An infinite set of curves contained in a finite domain is compact if the curves are rectifiable and their lengths less than a fixed finite upper bound.

The torus R which contains all the graphs of \mathfrak{R} is a finite domain.

Suppose a rectilinear polygon inscribed in the graph of any representation g . Then, resolving each side into its projections along a parallel and along a meridian, and adding, we obtain

$$\text{length of inscribed polygon} \leq \text{length of parallel} + \text{length of meridian},$$

where the cyclically monotonic character of the graph insures that each projection is counted once and only once. It follows from this inequality that each

[†] Loc. cit., §11, p. 9.

[‡] Fréchet, loc. cit., p. 65.

curve of \mathfrak{R} is rectifiable and has a length \leq the finite upper bound: length of parallel + length of meridian.

Thus \mathfrak{R} obeys all the conditions of the above theorem, and is compact.

\mathfrak{M}' is an extremal set: compact and closed.

9. **The lower semi-continuity of $A(g)$.** In (5.5) we have expressed $A(g)$ as the sum of an infinite series of positive terms. Defining the partial sums

$$(9.1) \quad A_r(g) = \Delta_1(g) + \Delta_2(g) + \cdots + \Delta_r(g) = \frac{1}{16\pi} \int \int_{T_r} \frac{\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2}{\sin^2 \frac{\theta - \phi}{2}} d\theta d\phi,$$

where T_r denotes the domain

$$\frac{\pi}{r+1} \leq |\theta - \phi| \leq \pi,$$

we may also express $A(g)$ as the limit of the sequence

$$A_1(g), A_2(g), \dots, A_r(g), \dots,$$

which, since each $\Delta_r(g) > 0$, is continually increasing:

$$A_1(g) < A_2(g) < \cdots < A_r(g) < \cdots.$$

It is easily seen that *each $A_r(g)$ is a continuous functional of g* : if a sequence of representations

$$g^{(1)}, g^{(2)}, \dots, g^{(m)}, \dots$$

tends to g as limit, then

$$A_r(g^{(1)}), A_r(g^{(2)}), \dots, A_r(g^{(m)}), \dots$$

tends to $A_r(g)$. For the integrand in (9.1) is *uniformly bounded*, being

$$\leq \frac{d^2}{\sin^2 \frac{\pi}{2(r+1)}},$$

where d is the diameter of Γ , and under this condition it is permissible to pass to the limit under the sign of integration.

The lower semi-continuity of $A(g)$ now results from the following general theorem.

THEOREM. *If a sequence of continuous† functions on an L -set tend, in increasing (wide sense), to a limit function (finite or infinite valued), then this limit function is lower semi-continuous.*

Let a denote an arbitrary element of the L -set, and

$$(9.2) \quad U_1(a) \leq U_2(a) \leq \cdots \leq U_r(a) \leq \cdots,$$

the increasing sequence of functions tending to the limit $U(a)$. Let

$$a_1, a_2, \cdots, a_m, \cdots$$

be any sequence of elements converging to a as limit.

Case 1: $U(a)$ finite. If $\epsilon > 0$ be assigned arbitrarily, there exists an r_ϵ such that for $r > r_\epsilon$

$$(9.3) \quad U_r(a) > U(a) - \epsilon/2.$$

We suppose that in this inequality r has a *fixed* value $> r_\epsilon$, for instance, $r = r_\epsilon + 1$.

By hypothesis, the function U_r is continuous; this implies the existence of an m_ϵ such that for $m > m_\epsilon$

$$(9.4) \quad U_r(a_m) > U_r(a) - \epsilon/2.$$

Combining the inequalities (9.3) and (9.4), we have

$$(9.5) \quad U_r(a_m) > U(a) - \epsilon$$

for $m > m_\epsilon$.

Now, by (9.2), each of the functions U_r is, for any fixed value of the argument, not greater than the limit function U ; thus

$$(9.6) \quad U(a_m) \geq U_r(a_m)$$

for every m , in particular for $m > m_\epsilon$.

From (9.5) and (9.6) it follows that

$$U(a_m) > U(a) - \epsilon$$

for $m > m_\epsilon$; but this is the definition of lower semi-continuity, according to (7.1).

Case 2: $U(a) = +\infty$. Here lower semi-continuity becomes identical with continuity: if $a_1, a_2, \cdots, a_m, \cdots$ is any sequence of elements converging to a , then

$$U(a_1), U(a_2), \cdots, U(a_m), \cdots$$

tends to $+\infty$.

† The theorem still remains valid if the functions of the sequence are merely lower semi-continuous. Cf. Carathéodory, *Vorlesungen über reelle Funktionen*, Leipzig, 1918, p. 175, where this theorem is proved for functions of n real variables.

For the proof, let G be an arbitrarily assigned finite positive number; then since by hypothesis

$$\lim_{r \rightarrow \infty} U_r(a) = +\infty,$$

an index r exists such that

$$(9.7) \quad U_r(a) > 2G.$$

Because of the continuity of the function U_r , there exists an index m_G such that for $m > m_G$

$$(9.8) \quad U_r(a_m) > U_r(a) - G.$$

Combining (9.7) and (9.8), we have

$$U_r(a_m) > G$$

for $m > m_G$. From this and (9.6), it follows that

$$U(a_m) > G$$

for $m > m_G$, that is,

$$\lim_{m \rightarrow \infty} U(a_m) = +\infty,$$

which was to be proved.

Since $A(g)$ obeys all the conditions of the theorem just proved, its lower semi-continuity is established.†

10. $A(g)$ attains its minimum. In this section we prove that $A(g)$, as a lower semi-continuous function on the compact closed set \mathfrak{M}' , must attain its minimum value on \mathfrak{M}' .

Since the values of $A(g)$ are all positive, and some are finite, they have a finite lower bound $M \geq 0$.‡ By definition of lower bound, $A(g)$ cannot take any value less than M , but can approach to M from above as closely as we please. On this basis we can construct a *minimizing sequence*

$$(10.1) \quad g^{(1)}, g^{(2)}, \dots, g^{(m)}, \dots,$$

that is, one such that

$$(10.2) \quad A(g^{(1)}), A(g^{(2)}), \dots, A(g^{(m)}), \dots$$

tends to the limit M ; the construction of such a sequence is the first step in the direct treatment of any calculus of variations problem.

† It is easy to prove that by making g approach suitably to g^0 , any number whatever $\geq A(g^0)$ (including $+\infty$) can be made the limit of $A(g)$.

‡ After it has been proved that M is attained, it results that $M > 0$.

Now the sequence (10.1) may not converge to a limit, but since the set \mathfrak{M}' is compact, we can select a sub-sequence

$$(10.1') \quad g^{(m_1)}, g^{(m_2)}, \dots, g^{(m_k)}, \dots$$

which converges to a limit g^* ; on account of the closure of \mathfrak{M}' , g^* belongs to \mathfrak{M}' . The sequence of corresponding functional values

$$(10.2') \quad A(g^{(m_1)}), A(g^{(m_2)}), \dots, A(g^{(m_k)}), \dots,$$

being a sub-sequence of (10.2) with the limit M , must tend to the same limit M .

Using the lower semi-continuity of $A(g)$ as expressed in (7.2), we therefore have

$$A(g^*) \leq M;$$

but the definition of lower bound makes $A(g^*) < M$ impossible, consequently

$$A(g^*) = M,$$

that is: the minimum of $A(g)$ on \mathfrak{M}' is attained for g^* .

By the discussion at the end of §6, it follows that the minimum of $A(g)$ on \mathfrak{M} is attained for g^* and all its equivalents.

11. Calculation of the power series $\sum_{i=1}^n F_i'^2(w)$. The rest of our argument is concerned with showing that the harmonic surface

$$x_i = \Re F_i(w)$$

determined by the minimizing representation g^* is minimal:

$$\sum_{i=1}^n F_i'^2(w) = 0.$$

This will be done by showing that, in a sense whose precise meaning appears in the sequel, the last condition expresses the vanishing of the first variation of $A(g)$ for $g = g^*$.

The functions $F_i(w)$ determined by the representation

$$x_i = g_i(\theta)$$

of Γ are given by the power series, convergent (at least) in the interior of the unit circle C ,

$$(11.1) \quad F_i(w) = \frac{a_{i0}}{2} + \sum_{p=1}^{\infty} (a_{ip} - ib_{ip}) w^p, \dagger$$

† Throughout, the symbol i will be used in two senses: i the index running from 1 to n , and i the square root of -1 . This notation should not lead to any confusion.

where a_{ip} , b_{ip} are the Fourier coefficients of $g_i(\theta)$:

$$(11.2) \quad a_{ip} = \frac{1}{\pi} \int_C g_i(\theta) \cos p\theta d\theta, \quad b_{ip} = \frac{1}{\pi} \int_C g_i(\theta) \sin p\theta d\theta.$$

Instead of $\sum_{i=1}^n F_i'^2(w)$ we shall find it more convenient to work with $w^2 \sum_{i=1}^n F_i'^2(w)$; it will be easy to dispose of the factor w^2 when we wish. The latter expression is representable in the interior of C by a power series, derivable from (11.1) by performing formally the operations indicated:

$$(11.3) \quad w^2 \sum_{i=1}^n F_i'^2(w) = \sum_{m=2}^{\infty} (A_m - iB_m) w^m.$$

The special object of this section is to calculate the coefficients A_m , B_m ; the results appear in (11.15).

Since the power series (11.1) is convergent in the interior of C , we may differentiate termwise, and find

$$(11.4) \quad wF_i'(w) = \sum_{p=1}^{\infty} p(a_{ip} - ib_{ip})w^p.$$

Rewriting this as

$$(11.4') \quad wF_i'(w) = \sum_{q=1}^{\infty} q(a_{iq} - ib_{iq})w^q,$$

we obtain, on multiplying together (11.4) and (11.4') and summing as to i from 1 to n , the formula (11.3) with

$$(11.5) \quad A_m - iB_m = \sum_{i=1}^n \sum_{p,q} p(a_{ip} - ib_{ip}) \cdot q(a_{iq} - ib_{iq})$$

where the range of the summation indices p, q is

$$(11.6) \quad p \geq 1, \quad q \geq 1, \quad p + q = m.$$

From (11.2),

$$a_{ip} - ib_{ip} = \frac{1}{\pi} \int_C g_i(\theta) e^{-pi\theta} d\theta, \quad a_{iq} - ib_{iq} = \frac{1}{\pi} \int_C g_i(\phi) e^{-qi\phi} d\phi.$$

The product of the two parts of this formula, after multiplying the first by p , the second by q , may be expressed as a double integral:

$$(11.7) \quad p(a_{ip} - ib_{ip}) \cdot q(a_{iq} - ib_{iq}) \\ = \frac{1}{\pi^2} \int_C \int_C g_i(\theta) g_i(\phi) \cdot p e^{-pi\theta} \cdot q e^{-qi\phi} \cdot d\theta d\phi.$$

We have obviously

$$(11.8) \quad \int_C \int_C g_i^2(\theta) \cdot p e^{-p i \theta} \cdot q e^{-q i \phi} \cdot d\theta d\phi = \int_C g_i^2(\theta) \cdot p e^{-p i \theta} d\theta \cdot \int_C q e^{-q i \phi} d\phi = 0$$

since the second factor vanishes; and similarly

$$(11.8') \quad \int_C \int_C g_i^2(\phi) \cdot p e^{-p i \theta} \cdot q e^{-q i \phi} \cdot d\theta d\phi = 0.$$

By the last three equations we may write

$$(11.9) \quad \begin{aligned} & p(a_{ip} - ib_{ip}) \cdot q(a_{iq} - ib_{iq}) \\ &= -\frac{1}{2\pi^2} \int_C \int_C [g_i(\theta) - g_i(\phi)]^2 \cdot p e^{-p i \theta} \cdot q e^{-q i \phi} \cdot d\theta d\phi, \end{aligned}$$

for, on expanding the bracket squared, this reduces to (11.7) when account is taken of (11.8), (11.8').

Substituting (11.9) in (11.5), we have

$$(11.10) \quad A_m - iB_m = -\frac{1}{2\pi^2} \int_C \int_C \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \cdot \sum_{p,q} p e^{-p i \theta} \cdot q e^{-q i \phi} \cdot d\theta d\phi.$$

Accordingly, we have to calculate

$$(11.11) \quad \sum_{p,q} p e^{-p i \theta} \cdot q e^{-q i \phi}$$

where the range of p, q is as defined in (11.6). Let

$$(11.12) \quad e^{-i\theta} = \zeta, \quad e^{-i\phi} = z,$$

so that (11.11) becomes

$$(11.11') \quad \begin{aligned} \sum_{p,q} p \zeta^p \cdot q z^q &= (m-1)\zeta^{m-1}z + (m-2)\zeta^{m-2} \cdot 2z^2 + \dots \\ &+ 2\zeta^2 \cdot (m-2)z^{m-2} + \zeta \cdot (m-1)z^{m-1}. \end{aligned}$$

The value of the last expression can be found by starting with the geometric progression

$$\zeta^m + \zeta^{m-1}z + \zeta^{m-2}z^2 + \dots + \zeta^2 z^{m-2} + \zeta z^{m-1} + z^m = \frac{\zeta^{m+1} - z^{m+1}}{\zeta - z}.$$

Applying to this the operator

$$\zeta z \frac{\partial^2}{\partial \zeta \partial z},$$

we get

$$(11.13) \quad \sum_{p,q} p \zeta^p \cdot q z^q = (m+1) \zeta z (\zeta^m + z^m) (\zeta - z)^{-2} - 2 \zeta z (\zeta^{m+1} - z^{m+1}) (\zeta - z)^{-3}.$$

The calculation of the first and second terms of this expression is as follows.

First term. Let

$$(11.14) \quad \sigma = \frac{\theta + \phi}{2}, \quad \delta = \frac{\theta - \phi}{2}.$$

Then

$$\begin{aligned} \zeta z &= e^{-i(\theta+\phi)} = e^{-2i\sigma}; \\ \zeta^m + z^m &= (\cos m\theta + \cos m\phi) - i(\sin m\theta + \sin m\phi) \\ &= 2 \cos m\sigma \cos m\delta - 2i \sin m\sigma \cos m\delta \\ &= 2 \cos m\delta e^{-mi\sigma}; \\ \zeta - z &= (\cos \theta - \cos \phi) - i(\sin \theta - \sin \phi) \\ &= -2 \sin \sigma \sin \delta - 2i \cos \sigma \sin \delta \\ &= -2i \sin \delta e^{-i\sigma}; \\ (\zeta - z)^{-2} &= -\frac{1}{4 \sin^2 \delta} e^{2i\sigma}; \\ \text{First term} &= -\frac{(m+1) \cos m\delta}{2 \sin^2 \delta} e^{-mi\sigma}. \end{aligned}$$

Second term.

$$\begin{aligned} \zeta z &= e^{-2i\sigma}; \\ \zeta^{m+1} - z^{m+1} &= -2i \sin (m+1)\delta e^{-(m+1)i\sigma}, \end{aligned}$$

found from the above formula for $\zeta - z$ by replacing θ, ϕ by $(m+1)\theta, (m+1)\phi$;

$$(\zeta - z)^{-3} = \frac{1}{8i \sin^3 \delta} e^{3i\sigma};$$

$$\text{Second term} = \frac{\sin (m+1)\delta}{2 \sin^3 \delta} e^{-mi\sigma}.$$

Substituting these results in (11.13), we obtain

$$\sum_{p,q} p \zeta^p \cdot q z^q = \left\{ -\frac{(m+1) \cos m\delta}{2 \sin^2 \delta} + \frac{\sin (m+1)\delta}{2 \sin^3 \delta} \right\} e^{-mi\sigma}.$$

The bracket is equal to

$$\begin{aligned} & - \frac{(m+1) \cdot 2 \cos m\delta \sin \delta + 2 \sin (m+1)\delta}{4 \sin^3 \delta} \\ & = \frac{-(m+1) \{ \sin (m+1)\delta - \sin (m-1)\delta \} + 2 \sin (m+1)\delta}{4 \sin^3 \delta} \\ & = - \frac{(m-1) \sin (m+1)\delta - (m+1) \sin (m-1)\delta}{4 \sin^3 \delta}; \end{aligned}$$

so that, referring to the notation (11.12), we have finally for the expression (11.11):

$$\sum_{p,q} p e^{-pi\theta} \cdot q e^{-qi\phi} = - \frac{(m-1) \sin (m+1)\delta - (m+1) \sin (m-1)\delta}{4 \sin^3 \delta} e^{-mi\sigma}.$$

Substituting this in (11.10), we get

$$\begin{aligned} A_m - iB_m = & \frac{1}{8\pi^2} \int_C \int_C \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \\ & \cdot \frac{(m-1) \sin (m+1)\delta - (m+1) \sin (m-1)\delta}{\sin^3 \delta} e^{-mi\sigma} d\theta d\phi. \end{aligned}$$

Writing $e^{-mi\sigma} = \cos m\sigma - i \sin m\sigma$, separating the real and imaginary parts, and referring to the notation (11.14), we arrive at the final expressions for A_m, B_m :

(11.15)

$$\begin{aligned} A_m = & \frac{1}{8\pi^2} \int_C \int_C \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \\ & \cdot \frac{(m-1) \sin \left[(m+1) \frac{\theta-\phi}{2} \right] - (m+1) \sin \left[(m-1) \frac{\theta-\phi}{2} \right]}{\sin^3 \frac{\theta-\phi}{2}} \cos \left[m \frac{\theta+\phi}{2} \right] d\theta d\phi, \\ B_m = & \frac{1}{8\pi^2} \int_C \int_C \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \\ & \cdot \frac{(m-1) \sin \left[(m+1) \frac{\theta-\phi}{2} \right] - (m+1) \sin \left[(m-1) \frac{\theta-\phi}{2} \right]}{\sin^3 \frac{\theta-\phi}{2}} \sin \left[m \frac{\theta+\phi}{2} \right] d\theta d\phi. \end{aligned}$$

It may be observed that these are *proper* Riemann integrals, for the fraction, which takes for $\theta = \phi$ the indeterminate form $0/0$, has the limiting value $-(2/3)m(m^2 - 1)$.

12. The functions $C_m(\lambda)$, $S_m(\lambda)$; $C_{mr}(\lambda)$, $S_{mr}(\lambda)$. *Hypothesis.* The work of §§12-15 is valid on the sole hypothesis that g is a fixed representation for which the functional $A(g)$ has a finite value.† In particular, we may put $g = g^*$, the minimizing representation of $A(g)$, since $A(g^*)$ is finite.

Consider the transformations

$$(12.1) \quad \bar{\theta} = \theta + \lambda \cos m\theta = c(\theta), \quad \bar{\theta} = \theta + \lambda \sin m\theta = s(\theta),$$

m any fixed positive integer,

of the unit circle C into itself \bar{C} . If λ is a real parameter restricted to the interval

$$(12.2) \quad -\frac{1}{m} < \lambda < \frac{1}{m},$$

then each of these transformations is one-one and continuous. This results from the fact that the respective derivatives

$$\frac{d\bar{\theta}}{d\theta} = 1 - m\lambda \sin m\theta, \quad \frac{d\bar{\theta}}{d\theta} = 1 + m\lambda \cos m\theta$$

are positive for all θ under the condition (12.2). Consequently the transformations (12.1) have one-one continuous inverses

$$(12.1') \quad \theta = c^{-1}(\bar{\theta}), \quad \theta = s^{-1}(\bar{\theta}),$$

and applying these to the representation g , we get the family of representations $g(c^{-1}(\bar{\theta}))$, $g(s^{-1}(\bar{\theta}))$, depending on the parameter λ .

The values of $A(g)$ for these representations are respectively

$$\frac{1}{16\pi} \iint_{\bar{T}} \frac{\sum_{i=1}^n [g_i(c^{-1}(\bar{\theta})) - g_i(c^{-1}(\bar{\phi}))]^2}{\sin^2 \frac{\bar{\theta} - \bar{\phi}}{2}} d\bar{\theta} d\bar{\phi},$$

$$\frac{1}{16\pi} \iint_{\bar{T}} \frac{\sum_{i=1}^n [g_i(s^{-1}(\bar{\theta})) - g_i(s^{-1}(\bar{\phi}))]^2}{\sin^2 \frac{\bar{\theta} - \bar{\phi}}{2}} d\bar{\theta} d\bar{\phi},$$

† It will follow from §17 that, in particular, g cannot be improper of the first kind.

where the torus $\bar{T} = \bar{C}\bar{C}$ corresponds to the torus $T = CC$ by the equations (12.1). Making the change of variables (12.1) in these double integrals, we obtain two functions of λ , which we will denote respectively by $C_m(\lambda)$, $S_m(\lambda)$:

$$(12.3) \quad \begin{aligned} C_m(\lambda) &= \frac{1}{16\pi} \iint_T \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \frac{(1 - m\lambda \sin m\theta)(1 - m\lambda \sin m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\cos m\theta - \cos m\phi}{2} \right)} d\theta d\phi, \\ S_m(\lambda) &= \frac{1}{16\pi} \iint_T \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \frac{(1 + m\lambda \cos m\theta)(1 + m\lambda \cos m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\sin m\theta - \sin m\phi}{2} \right)} d\theta d\phi. \end{aligned}$$

These are *improper* Riemann integrals (singular locus $\theta = \phi$).

As in §5, we express the torus T as the sum of an infinite number of strips:

$$T = \Delta_1 + \Delta_2 + \cdots + \Delta_r + \cdots,$$

defined by (5.2). Replacing the domain of integration T in (12.3) by Δ_r , we derive the functions

$$(12.4) \quad \begin{aligned} C_{mr}(\lambda) &= \frac{1}{16\pi} \iint_{\Delta_r} \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \frac{(1 - m\lambda \sin m\theta)(1 - m\lambda \sin m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\cos m\theta - \cos m\phi}{2} \right)} d\theta d\phi, \\ S_{mr}(\lambda) &= \frac{1}{16\pi} \iint_{\Delta_r} \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \frac{(1 + m\lambda \cos m\theta)(1 + m\lambda \cos m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\sin m\theta - \sin m\phi}{2} \right)} d\theta d\phi. \end{aligned}$$

These are *proper* Riemann integrals since the domain Δ_r does not contain any singular points $\theta = \phi$.

In this way we have for definition of $C_m(\lambda)$, $S_m(\lambda)$ the infinite series

$$(12.5) \quad \begin{aligned} C_m(\lambda) &= C_{m1}(\lambda) + C_{m2}(\lambda) + \cdots + C_{mr}(\lambda) + \cdots, \\ S_m(\lambda) &= S_{m1}(\lambda) + S_{m2}(\lambda) + \cdots + S_{mr}(\lambda) + \cdots. \end{aligned}$$

Complex values of λ . We shall wish to make use of certain classic convergence theorems valid only in the complex domain; for this reason we now allow λ to be a complex variable:

$$\lambda = \mu + \nu i,$$

subject to the restriction

$$(12.6) \quad |\lambda| < 1/m.$$

The (open) circle in the complex plane defined by this inequality will be denoted by \mathcal{C}_m . Evidently \mathcal{C}_m contains the real interval (12.2).

The expressions (12.4) are still proper Riemann integrals defining the now complex-valued functions $C_{mr}(\lambda)$, $S_{mr}(\lambda)$. To be assured of this, we must show that the denominators of the respective integrands do not vanish in the domain Δ_r . These denominators are

$$\begin{aligned} & \sin^2 \left(\frac{\theta - \phi}{2} + \mu \frac{\cos m\theta - \cos m\phi}{2} + i\nu \frac{\cos m\theta - \cos m\phi}{2} \right), \\ & \sin^2 \left(\frac{\theta - \phi}{2} + \mu \frac{\sin m\theta - \sin m\phi}{2} + i\nu \frac{\sin m\theta - \sin m\phi}{2} \right). \end{aligned}$$

Since the function $\sin^2(z/2)$ vanishes when and only when $z = 2k\pi$ ($k = 0, \pm 1, \pm 2, \dots$), it is evident that for these expressions to vanish we must have respectively

$$\begin{aligned} (12.7) \quad & \theta + \mu \cos m\theta = \phi + \mu \cos m\phi, \\ & \theta + \mu \sin m\theta = \phi + \mu \sin m\phi \quad (\text{mod } 2\pi). \end{aligned}$$

But $|\mu| \leq |\lambda| < 1/m$, and under this condition on the real parameter μ the equations (12.1), with μ in place of λ , define a one-one transformation. Hence each equation (12.7) implies

$$\theta = \phi \quad (\text{mod } 2\pi);$$

but this condition is never satisfied in the domain Δ_r .

The proof is thus complete that $C_{mr}(\lambda)$, $S_{mr}(\lambda)$ are well-defined by the formulas (12.4) for all values of λ in the interior of the circle \mathcal{C}_m .

We now define $C_m(\lambda)$, $S_m(\lambda)$ for complex λ by the infinite series (12.5) provided these series are absolutely convergent. We shall prove that the series in question are indeed absolutely convergent for all λ in \mathcal{C}_m , and then (what is of paramount importance) that $C_m(\lambda)$, $S_m(\lambda)$ are analytic functions of λ in \mathcal{C}_m .

Plan. The proof of the analyticity of $C_m(\lambda)$, $S_m(\lambda)$ will consist of a direct application of the Weierstrass double-series theorem.[†] We will show in §13 that each term $C_{mr}(\lambda)$, $S_{mr}(\lambda)$ of the series (12.5) is an analytic function of λ in \mathcal{C}_m , and then in §14 that the series (12.5) is uniformly convergent in every circle concentric with and smaller than \mathcal{C}_m . The result is to justify the formal operation of expanding in powers of λ the integrands in the formulas (12.3) for $C_m(\lambda)$, $S_m(\lambda)$, and then performing termwise the double integration as to θ , ϕ .

[†] Cf. Knopp, *Funktionentheorie*, Berlin and Leipzig, 1918, vol. 1, p. 89.

13. **Analyticity of $C_{mr}(\lambda)$, $S_{mr}(\lambda)$.** Considering the formulas (12.4) for $C_{mr}(\lambda)$, $S_{mr}(\lambda)$, we have just shown that the denominators of the fractions

$$(13.1) \quad \frac{(1 - m\lambda \sin m\theta)(1 - m\lambda \sin m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\cos m\theta - \cos m\phi}{2} \right)}, \quad \frac{(1 + m\lambda \cos m\theta)(1 + m\lambda \cos m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\sin m\theta - \sin m\phi}{2} \right)}$$

do not vanish for (θ, ϕ) in Δ_r and λ in \mathcal{C}_m ; therefore these fractions are holomorphic functions of λ in \mathcal{C}_m , depending on the parameters θ, ϕ . They are developable in a series of powers of λ convergent in \mathcal{C}_m , with coefficients functions of θ, ϕ . The developments may be obtained in the usual way as Taylor series, and are given by the following calculations, where we retain terms only as far as the first power of λ .

Returning to the notation (11.14),

$$(13.2) \quad \sigma = \frac{1}{2}(\theta + \phi), \quad \delta = \frac{1}{2}(\theta - \phi),$$

we have

$$\frac{\theta - \phi}{2} + \lambda \frac{\cos m\theta - \cos m\phi}{2} = \delta - \lambda \sin m\sigma \sin m\delta.$$

Making in the Taylor expansion

$$\frac{1}{\sin^2(x + h)} = \frac{1}{\sin^2 x} - \frac{2 \cos x}{\sin^3 x} h + \dots$$

the substitutions

$$x = \delta, \quad h = -\lambda \sin m\sigma \sin m\delta,$$

we get

$$(13.3) \quad \frac{1}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\cos m\theta - \cos m\phi}{2} \right)} = \frac{1}{\sin^2 \delta} + \lambda \frac{2 \cos \delta \sin m\sigma \sin m\delta}{\sin^3 \delta} + \dots$$

For the numerator of (13.1) we have

$$(13.4) \quad \begin{aligned} (1 - m\lambda \sin m\theta)(1 - m\lambda \sin m\phi) \\ = 1 + \lambda(-2m \sin m\sigma \cos m\delta) + \text{term in } \lambda^2. \end{aligned}$$

Multiplying together (13.3) and (13.4), we get for the development of the first of (13.1)

$$\frac{1}{\sin^2 \delta} + \lambda \left\{ \frac{2 \cos \delta \sin m\sigma \sin m\delta}{\sin^3 \delta} - \frac{2m \sin m\sigma \cos m\delta}{\sin^2 \delta} \right\} + \dots$$

The bracket reduces to

$$\begin{aligned} & \frac{2 \cos \delta \sin m\delta - 2m \cos m\sigma \sin \delta}{\sin^3 \delta} \sin m\sigma \\ &= \frac{\{\sin(m+1)\delta + \sin(m-1)\delta\} - m\{\sin(m+1)\delta - \sin(m-1)\delta\}}{\sin^3 \delta} \sin m\sigma \\ &= - \frac{(m-1) \sin(m+1)\delta - (m+1) \sin(m-1)\delta}{\sin^3 \delta} \sin m\sigma. \end{aligned}$$

Thus we have finally, referring to (13.2),

$$\begin{aligned} (13.5) \quad & \frac{(1 - m\lambda \sin m\theta)(1 - m\lambda \sin m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\cos m\theta - \cos m\phi}{2} \right)} = \frac{1}{\sin^2 \frac{\theta - \phi}{2}} \\ & - \lambda \frac{(m-1) \sin \left[(m+1) \frac{\theta - \phi}{2} \right] - (m+1) \sin \left[(m-1) \frac{\theta - \phi}{2} \right]}{\sin^3 \frac{\theta - \phi}{2}} \sin \left[m \frac{\theta + \phi}{2} \right] + \dots. \end{aligned}$$

An entirely similar calculation gives for the second fraction (13.1):

$$\begin{aligned} (13.5') \quad & \frac{(1 + m\lambda \cos m\theta)(1 + m\lambda \cos m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\sin m\theta - \sin m\phi}{2} \right)} = \frac{1}{\sin^2 \frac{\theta - \phi}{2}} \\ & + \lambda \frac{(m-1) \sin \left[(m+1) \frac{\theta - \phi}{2} \right] - (m+1) \sin \left[(m-1) \frac{\theta - \phi}{2} \right]}{\sin^3 \frac{\theta - \phi}{2}} \cos \left[m \frac{\theta + \phi}{2} \right] + \dots. \end{aligned}$$

The following observation is now of prime importance: if λ is given any fixed value in \mathcal{C}_m , so that

$$(13.6) \quad |\lambda| = \rho_0 < \frac{1}{m},$$

then the series (13.5), (13.5') are *uniformly convergent when* (θ, ϕ) *varies over* Δ_r . To prove this, denote by $F(\theta, \phi, \lambda)$ the first member of (13.5), and let ρ be any fixed positive number such that

$$(13.7) \quad \rho_0 < \rho < \frac{1}{m}.$$

Suppose θ, ϕ, λ to vary arbitrarily subject to the conditions

$$(13.8) \quad (\theta, \phi) \text{ in } \Delta_r, \quad |\lambda| = \rho.$$

The three-dimensional domain so defined is evidently closed and bounded, and the positive real-valued function $|F(\theta, \phi, \lambda)|$ is finite and continuous on this domain, for it has been shown that the denominator of $F(\theta, \phi, \lambda)$ is never zero in (13.8). By a fundamental theorem, $|F(\theta, \phi, \lambda)|$ has therefore a finite upper bound on (13.8):

$$|F(\theta, \phi, \lambda)| \leq B,$$

B being a positive real number independent of θ, ϕ, λ .

We now apply the appraisal formula of Cauchy† for the coefficients in the power series expansion of an analytic function. According to this formula, if $A_k(\theta, \phi)$ denote the coefficient of λ^k in the power series (13.5), then

$$|A_k(\theta, \phi)| \leq \frac{B}{\rho^k}$$

for all (θ, ϕ) in Δ_r . Hence, with (13.6),

$$|A_k(\theta, \phi)\lambda^k| \leq B\left(\frac{\rho_0}{\rho}\right)^k.$$

Since the series of constant positive terms

$$\sum_{k=0}^{\infty} B\left(\frac{\rho_0}{\rho}\right)^k$$

is convergent, being a geometric progression of ratio < 1 (by (13.7)), the conditions of a standard uniform convergence test of Weierstrass are satisfied, and we have proved the uniform convergence of (13.5) for λ fixed in \mathcal{C}_m and (θ, ϕ) varying over Δ_r . The same argument applies to (13.5').

That this uniform convergence is not disturbed when, to obtain the integrands in (12.4) of $C_{mr}(\lambda)$, $S_{mr}(\lambda)$, we multiply (13.5), (13.5') by $\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2$, results from the fact that the multiplying factor is bounded:

$$\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \leq d^2,$$

where d is the diameter of Γ .

Consequently, after introducing this factor in (13.5), (13.5'), we may integrate as to θ, ϕ over Δ_r term by term, and find that $C_{mr}(\lambda)$, $S_{mr}(\lambda)$ are equal to power series in λ :

† Cf. Knopp, loc. cit., vol. 1, p. 84.

$$(13.9) \quad \begin{aligned} C_{mr}(\lambda) &= \Delta_r(g) + C'_{mr}(0)\lambda + \cdots, \\ S_{mr}(\lambda) &= \Delta_r(g) + S'_{mr}(0)\lambda + \cdots, \end{aligned}$$

convergent for λ in \mathcal{C}_m .

We are especially interested in the values of the coefficients of the first power of λ , which, referring to (13.5), (13.5'), are seen to be

$$(13.10) \quad \begin{aligned} C'_{mr}(0) &= -\frac{1}{16\pi} \iint_{\Delta_r} \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \\ &\quad \frac{(m-1) \sin \left[(m+1) \frac{\theta-\phi}{2} \right] - (m+1) \sin \left[(m-1) \frac{\theta-\phi}{2} \right]}{\sin^3 \frac{\theta-\phi}{2}} \sin \left[m \frac{\theta-\phi}{2} \right] d\theta d\phi, \\ S'_{mr}(0) &= \frac{1}{16\pi} \iint_{\Delta_r} \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \\ &\quad \frac{(m-1) \sin \left[(m+1) \frac{\theta-\phi}{2} \right] - (m+1) \sin \left[(m-1) \frac{\theta-\phi}{2} \right]}{\sin^3 \frac{\theta-\phi}{2}} \cos \left[m \frac{\theta+\phi}{2} \right] d\theta d\phi. \end{aligned}$$

14. Analyticity of the functions $C_m(\lambda)$, $S_m(\lambda)$. The function $C_m(\lambda)$ is defined by the infinite series (12.5), each of whose terms has just been proved analytic in the circle \mathcal{C}_m . We proceed to prove that this series is *uniformly convergent* in every smaller concentric circle

$$(14.1) \quad |\lambda| \leq \rho, \quad \rho < \frac{1}{m}.$$

The analyticity of $C_m(\lambda)$ will then result immediately by the Weierstrass double-series theorem. The same argument applies to $S_m(\lambda)$.

Considering $C_{mr}(\lambda)$, we have from (12.4)

$$(14.2) \quad |C_{mr}(\lambda)| \leq \frac{1}{16\pi} \iint_{\Delta_r} \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \left| \frac{(1 - m\lambda \sin m\theta)(1 - m\lambda \sin m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\cos m\theta - \cos m\phi}{2} \right)} \right| d\theta d\phi.$$

Comparing this with the formula (5.3) for $\Delta_r(g)$, also a double integral over Δ_r with positive real elements, we wish to prove that the quotient of the integrand of (14.2) by that of $\Delta_r(g)$,

$$(14.3) \quad Q(\theta, \phi, \lambda) = \left| \frac{(1 - m\lambda \sin m\theta)(1 - m\lambda \sin m\phi)}{\sin^2 \left(\frac{\theta - \phi}{2} + \lambda \frac{\cos m\theta - \cos m\phi}{2} \right)} \right| \sin^2 \frac{\theta - \phi}{2},$$

has a finite upper bound when (θ, ϕ) varies over the entire torus T (*not merely over Δ_r*) and λ varies independently over the closed circle (14.1).

The expression (14.3) is not defined on the diagonal $\theta = \phi$ of the torus T , taking there the indeterminate form $0/0$; but it is seen directly that (14.3) has the limit value 1 for $\theta = \phi$:

$$(14.3') \quad Q(\phi, \phi, \lambda) = 1.$$

The function $Q(\theta, \phi, \lambda)$ is thus defined on the *closed* domain (of four real dimensions)

$$(14.4) \quad (\theta, \phi) \text{ on } T, \quad |\lambda| \leq \rho,$$

by (14.3) if $\theta \neq \phi$ and by (14.3') if $\theta = \phi$. As has been remarked several times, the denominator of (14.3) cannot vanish if $|\lambda| < 1/m$, $\theta \neq \phi$. Hence $Q(\theta, \phi, \lambda)$ is finite and continuous on the closed and bounded domain (14.4). It therefore has a finite positive upper bound K .

Consequently,

$$(14.5) \quad |C_{mr}(\lambda)| \leq K \Delta_r(g)$$

where it is to be emphasized that K is *independent of r* .

According to the hypothesis stated at the beginning of §12, g is a fixed representation for which $A(g)$ is finite. It follows that

$$K\Delta_1(g) + K\Delta_2(g) + \cdots + K\Delta_r(g) + \cdots$$

is a series of *constant, positive* terms, *convergent* to the sum $KA(g)$. By the inequality (14.5) this series dominates the series (12.5) for $C_m(\lambda)$. Therefore, by the same uniform convergence test of Weierstrass used in the preceding section, *the series (12.5) for $C_m(\lambda)$ is absolutely and uniformly convergent in the circle (14.1).*

The conditions of the Weierstrass double-series theorem being now satisfied, we may assert that $C_m(\lambda)$ is *analytic in the circle \mathcal{C}_m* .

According to the same theorem, the coefficients in the power series representing $C_m(\lambda)$ are to be found by adding together the coefficients of like powers of λ in the various terms $C_{mr}(\lambda)$ (summation of the double series by columns). By §13, the coefficient of any given power of λ in the expansion of $C_{mr}(\lambda)$ is a double integral taken over Δ_r of an integrand that is the same for all r . Therefore the coefficient of this power of λ in $C_m(\lambda)$ is the double

integral of the same integrand taken over the entire torus T . Consequently, referring to (13.9), (13.10), we have

$$(14.6) \quad C_m(\lambda) = A(g) + C'_m(0)\lambda + \dots$$

with

$$(14.7) \quad C'_m(0) = -\frac{1}{16\pi} \iint_T \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \cdot \frac{(m-1) \sin \left[(m+1) \frac{\theta-\phi}{2} \right] - (m+1) \sin \left[(m-1) \frac{\theta-\phi}{2} \right]}{\sin^3 \frac{\theta-\phi}{2}} \sin \left[m \frac{\theta+\phi}{2} \right] d\theta d\phi.$$

The entire preceding argument applies to $S_m(\lambda)$, and gives

$$(14.6') \quad S_m(\lambda) = A(g) + S'_m(0)\lambda + \dots$$

with

$$(14.7') \quad S'_m(0) = \frac{1}{16\pi} \iint_T \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \cdot \frac{(m-1) \sin \left[(m+1) \frac{\theta-\phi}{2} \right] - (m+1) \sin \left[(m-1) \frac{\theta-\phi}{2} \right]}{\sin^3 \frac{\theta-\phi}{2}} \cos \left[m \frac{\theta-\phi}{2} \right] d\theta d\phi.$$

15. Relations between A_m , B_m and $C'_m(0)$, $S'_m(0)$. The ultimate purpose of the calculations of §§11-14 was to establish the following relations, observable immediately by comparing (11.5) with (14.7), (14.7'):

$$(15) \quad A_m = \frac{2}{\pi} S'_m(0), \quad B_m = -\frac{2}{\pi} C'_m(0).$$

16. Existence of the minimal surface. The introduction of complex values of λ having served its purpose of establishing the power series (14.7), (14.7'), we now return to real values of λ in the interval (12.2),

$$(16.1) \quad -\frac{1}{m} < \lambda < \frac{1}{m},$$

where these power series remain valid, since this interval is part of the circle \mathcal{C}_m .

By applying the one-one continuous transformations (12.1) to the fixed representation g , we obtained in §12 a family of representations depending on

the parameter λ , and containing, for $\lambda=0$, the original representation g . Thus the function $C_m(\lambda)$ is a *part* of the functional $A(g)$ in the sense that its values are those of $A(g)$ on a certain part of the total range of g .

Suppose now $g=g^*$, the minimizing representation of $A(g)$. Then, a fortiori, $C_m(\lambda)$, considered as a function on the interval (16.1), has a minimum at $\lambda=0$, the value corresponding to g^* . Consequently,

$$(16.2) \quad C'_m(0) = 0.$$

Analogously,

$$(16.2') \quad S'_m(0) = 0.$$

Therefore, by (15),

$$(16.3) \quad A_m = 0, \quad B_m = 0.$$

Since m may have any integral value, it follows that every coefficient in $w^2 \sum_{i=1}^n F_i'^2(w)$ vanishes, where $F_i(w)$ are the power series, convergent within the unit circle, determined by g^* according to (11.1), (11.2). Hence, dividing out the non-identically vanishing factor w^2 , we have

$$(16.4) \quad \sum_{i=1}^n F_i'^2(w) = 0,$$

which expresses that the harmonic surface

$$x_i = \Re F_i(w)$$

determined by g^* is minimal.

That this minimal surface is bounded by Γ will follow after we have shown in the next two sections that g^* is a proper representation of Γ .

17. g^* cannot be improper of the first kind. We will rule out the possibility that g^* be improper of the first kind by proving that for a g of this type

$$A(g) = +\infty.$$

Since, by the hypothesis governing Part I, $A(g)$ takes a finite value for at least one representation g , the supposition that g^* , minimizing $A(g)$, could be improper of the first kind is thus reduced to an absurdity.

In an improper representation of the first kind, a point P of C corresponds to an arc $Q'Q''$ of Γ less than all of Γ . Since, by hypothesis, Γ has no double points, the end points Q' , Q'' are distinct and their distance l is not equal to zero.

The (vector) function g will have a discontinuity at P , where the distinct one-sided limits Q' , Q'' will exist. Therefore if two points of C approach to P

from opposite sides, the distance between the corresponding points of Γ tends to l as limit. Consequently, if f denote any fixed proper fraction,

$$(17.1) \quad \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 > fl^2$$

under the conditions

$$(17.2) \quad \alpha < \theta \leq \alpha + \delta, \quad \alpha - \delta \leq \phi < \alpha,$$

where α denotes the angular coördinate of P and $\delta > 0$ is fixed sufficiently small.

The domain (θ, ϕ) defined by (17.2) is a square S on the torus T , with one vertex $\theta = \phi = \alpha$ on the diagonal of T .

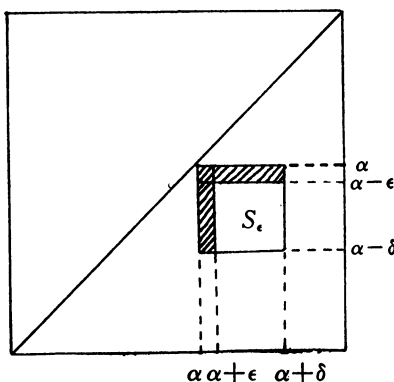


FIG. 3

Since the double integral (5.1) defining $A(g)$ is composed of non-negative elements, its value over T is not less than its value over S :

$$(17.3) \quad A(g) \geq \frac{1}{16\pi} \iint_S \frac{\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2}{\sin^2 \frac{\theta - \phi}{2}} d\theta d\phi;$$

therefore, by (17.1),

$$(17.4) \quad A(g) \geq m \iint_S \frac{d\theta d\phi}{\sin^2 \frac{\theta - \phi}{2}}$$

where $m = fl^2/(16\pi)$ is a positive constant.

The integral (17.4) is improper, since the integrand becomes infinite at the vertex of S lying on the diagonal $\theta = \phi$; hence this integral must be expressed as a limit:

$$(17.5) \quad \iint_S \frac{d\theta d\phi}{\sin^2 \frac{\theta - \phi}{2}} = \lim_{\epsilon \rightarrow 0} \iint_{S_\epsilon} \frac{d\theta d\phi}{\sin^2 \frac{\theta - \phi}{2}} = \lim_{\epsilon \rightarrow 0} \int_{\alpha+\epsilon}^{\alpha+\delta} \int_{\alpha-\delta}^{\alpha-\epsilon} \frac{d\theta d\phi}{\sin^2 \frac{\theta - \phi}{2}}.$$

Here S_ϵ is the square indicated in Fig. 3, obtained from S by removing strips of width ϵ along two of its sides.

We have for any double integral between *constant* limits the formula

$$\int_a^b \int_c^d f(\theta, \phi) d\theta d\phi = F(b, d) - F(b, c) - F(a, d) + F(a, c),$$

where $F(\theta, \phi)$ is any primitive of $f(\theta, \phi)$, that is, such that

$$\frac{\partial^2 F}{\partial \theta \partial \phi} = f(\theta, \phi).$$

In the case of (17.5),

$$F = 4 \log \sin \frac{\theta - \phi}{2},$$

and therefore

$$\begin{aligned} \int_{\alpha+\epsilon}^{\alpha+\delta} \int_{\alpha-\delta}^{\alpha-\epsilon} \frac{d\theta d\phi}{\sin^2 \frac{\theta - \phi}{2}} &= 4 \log \sin \frac{\delta - \epsilon}{2} \\ &\quad - 4 \log \sin \delta - 4 \log \sin \epsilon + 4 \log \sin \frac{\delta - \epsilon}{2}. \end{aligned}$$

If now $\epsilon \rightarrow 0$, the limit of the third term is $+\infty$, while the other terms stay finite; therefore the double integral over S is equal to $+\infty$; consequently, by (17.4),

$$A(g) = +\infty.$$

18. g^* cannot be improper of the second kind. That g^* cannot be improper of the second kind means that an arc $P'P''$ of C , less than all of C , cannot correspond by g^* to a single point Q of Γ . This will be established by showing that g^* cannot convert an arc $P'P''$ of C into a point Q of Γ without converting all of C into the point Q (degenerate representation).

The harmonic functions $\Re F_i(w)$ determined by g^* are given by

$$(18.1) \quad F_i(w) = \frac{1}{2\pi} \int_C \frac{e^{i\theta} + w}{e^{i\theta} - w} g_i^*(\theta) d\theta$$

(equivalent to the integral of Poisson). It is permissible to differentiate under the integral sign, and we obtain

$$(18.2) \quad wF'_i(w) = \frac{1}{2\pi} \int_C \frac{2we^{i\theta}}{(e^{i\theta} - w)^2} g_i^*(\theta) d\theta.$$

Writing $w = \rho e^{i\alpha}$, and taking the imaginary part of each side, we find

$$(18.3) \quad \Im wF'_i(w) = -\frac{1}{2\pi} \int_C \frac{2\rho(1 - \rho^2) \sin(\theta - \alpha)}{[1 - 2\rho \cos(\theta - \alpha) + \rho^2]^2} g_i^*(\theta) d\theta.$$

Except for the minus sign, the last expression is identical with one studied by Fatou in his thesis.[†] Fatou shows that at every point where the derivative $g_i^{*'}(\theta)$ exists and is continuous the expression in question has a unique limit equal to $g_i^*(\theta)$ when $(\rho, \alpha) \rightarrow (1, \theta)$.

Since in the present case $g_i^*(\theta)$ is supposed constant on a certain arc $P'P''$, $g_i^{*'}(\theta)$ has the continuous value zero on this arc; hence for θ any point interior to $P'P''$,

$$(18.4) \quad \lim_{w \rightarrow e^{i\theta}} \Im wF'_i(w) = 0 \quad (i = 1, 2, \dots, n).$$

We have proved in §16 that, $F_i(w)$ being determined by g^* ,

$$\sum_{i=1}^n w^2 F_i'^2(w) = \sum_{i=1}^n [\Re wF'_i(w) + i \Im wF'_i(w)]^2 = 0;$$

taking real parts, we have

$$\sum_{i=1}^n [\Re wF'_i(w)]^2 = \sum_{i=1}^n [\Im wF'_i(w)]^2.$$

Therefore, by (18.4),

$$\lim_{w \rightarrow e^{i\theta}} \sum_{i=1}^n [\Re wF'_i(w)]^2 = 0;$$

and since for each value of i

$$|\Re wF'_i(w)| \leq \left\{ \sum_{i=1}^n [\Re wF'_i(w)]^2 \right\}^{1/2},$$

we have

$$(18.4') \quad \lim_{w \rightarrow e^{i\theta}} \Re wF'_i(w) = 0 \quad (i = 1, 2, \dots, n).$$

By (18.4) and (18.4'),

$$(18.5) \quad \lim_{w \rightarrow e^{i\theta}} wF'_i(w) = 0 \quad (i = 1, 2, \dots, n),$$

for θ any interior point of $P'P''$.

[†] P. Fatou, *Séries trigonométriques et séries de Taylor*, Acta Mathematica, vol. 30 (1906), pp. 347-348.

Since the limiting value of the function $wF'_i(w)$ when w approaches to any point of the circular arc $P'P''$ is real, we may apply the symmetry principle of Riemann-Schwarz† to prolong the function analytically across this arc, attaching conjugate imaginary values to points inverse with respect to the arc. The function $wF'_i(w)$ will then be equal to zero on an arc *interior* to a domain of regularity, and is therefore identically equal to zero. Hence $F_i(w)$ is identically equal to a constant, $F_i(w) = a_i + ib_i$.

It follows that $x_i = g_i^*(\theta)$, boundary values of $\Re F_i(w)$, makes all of C correspond to the point of coördinates a_i ; but this is contrary to hypothesis.

By this and the preceding section, the representation g^* is proper; and therefore the minimal surface of §16, determined by g^* , is bounded by Γ .

With this, we have completed the solution of the problem of Plateau for any finite-area-spanning contour in n -dimensional euclidean space.

II. AN ARBITRARY JORDAN CONTOUR

19. **The Jordan contour as a limit of polygons.** The case of an arbitrary Jordan contour Γ , for which $A(g)$ is identically $+\infty$, will be dealt with by regarding Γ as the limit of a sequence of non-self-intersecting polygons‡

$$(19.1) \quad \Gamma^{(1)}, \Gamma^{(2)}, \dots, \Gamma^{(m)}, \dots$$

Let Γ , referred to a fixed initial parameter t , have the equations

$$(19.2) \quad x_i = f_i(t).$$

Then the polygons can be represented parametrically by equations

$$(19.3) \quad x_i = f_i^{(1)}(t), x_i = f_i^{(2)}(t), \dots, x_i = f_i^{(m)}(t), \dots,$$

so that $f_i^{(m)}(t)$ tends uniformly to $f_i(t)$ when $m \rightarrow \infty$.

Each polygon $\Gamma^{(m)}$ has, by Part I, a parameter θ such that the corresponding representation

$$(19.4) \quad x_i = g_i^{(m)}(\theta)$$

† Cf. G. Julia, *Principes Géométriques d'Analyse*, Paris, 1930, pp. 44–48, especially paragraph e.

‡ See C. Jordan, *Cours d'Analyse* (2d edition, Paris, 1893), p. 93. The polygons may be only partially inscribed in Γ , being derived from inscribed polygons by removing possible loops.

We could also proceed by expressing the continuous functions $x_i = f_i(t)$ representing Γ as the limits of their Fejér trigonometric polynomials $x_i = S_{im}(t)$. The contours $\Gamma^{(m)}$ thus represented might have multiple points. By referring back to Part I, it will be seen that this does not prevent the existence of a representation $t = \omega_m(\theta)$ of $\Gamma^{(m)}$ on C such that the parameter θ determines a minimal surface; only this representation might be improper in that a loop of $\Gamma^{(m)}$ could correspond to a single point of C , for the argument of §17 breaks down when $l=0$, and only then. The reader will readily see that the use of an improper representation $t = \omega_m(\theta)$ of $\Gamma^{(m)}$ would not at all complicate the proof which follows.

determines a minimal surface bounded by $\Gamma^{(m)}$:

$$(19.5) \quad \begin{aligned} x_i &= \Re F_i^{(m)}(w), \\ \sum_{i=1}^n [F_i^{(m)'}(w)]^2 &= 0. \end{aligned}$$

Each parameter θ of (19.4) is related to the corresponding parameter t of (19.3) by a proper topological transformation

$$(19.6) \quad t = \omega_1(\theta), \quad t = \omega_2(\theta), \quad \dots, \quad t = \omega_m(\theta), \quad \dots$$

of the unit circle C (considered as θ -locus) into itself (considered as t -locus). By adjoining a linear fractional transformation (6.1), which, according to §6, changes nothing essential in the preceding relations, we can secure that each transformation (19.6) converts three distinct fixed points $\theta_1, \theta_2, \theta_3$ into three distinct fixed points t_1, t_2, t_3 .

The sequence (19.6) is thus part of the set \mathfrak{M}' (§3) of topological transformations of C into itself. Since \mathfrak{M}' is compact and closed, we can select from (19.6) a sub-sequence converging to a limit

$$(19.7) \quad t = \omega(\theta),$$

which belongs to \mathfrak{M}' , and may be proper or improper but not degenerate, since \mathfrak{M}' contains no degenerate representations. To avoid complicating the notation, we will suppose that (19.6) already represents this convergent sub-sequence, and similarly in the formulas (19.1) to (19.5).

Let the (proper or improper) representation of Γ derived by applying the transformation (19.7) to (19.2) be

$$(19.8) \quad x_i = g_i(\theta) = f_i(\omega(\theta));$$

then we have

$$(19.9) \quad \lim_{m \rightarrow \infty} g_i^{(m)}(\theta) = g_i(\theta),$$

abstraction being made, in case the representation g is improper of the first kind, of the values of θ , at most denumerably infinite in number, where $g_i(\theta)$ is discontinuous. (19.9) rests on the fact (whose proof is trivial) that if $f_i^{(m)}(t)$ tends uniformly to the continuous $f_i(t)$ when $m \rightarrow \infty$, then if $t_m \rightarrow t$ as $m \rightarrow \infty$, we have $\lim f_i^{(m)}(t_m) = f_i(t)$ for $m \rightarrow \infty$.

The assertion is now easily proved that if

$$(19.10) \quad x_i = \Re F_i(w)$$

are the harmonic functions determined by $g_i(\theta)$, then

$$(19.11) \quad \sum_{i=1}^n F_i'^2(w) = 0,$$

so that the surface (19.10) is minimal.

For consider (18.2) without the factor w :

$$(19.12) \quad F_i^{(m)'}(w) = \frac{1}{2\pi} \int_C \frac{2e^{i\theta}}{(e^{i\theta} - w)^2} g_i^{(m)}(\theta) d\theta.$$

Since all the polygons $\Gamma^{(m)}$ are contained in a finite region of space, the functions $g_i^{(m)}(\theta)$ are uniformly bounded; and if w is any fixed point interior to the unit circle, the denominator $(e^{i\theta} - w)^2$ remains superior in absolute value to a fixed positive quantity when $e^{i\theta}$ describes C . Therefore the integrand in (19.12) remains uniformly bounded during the limit process (19.9); consequently the limit of the integral is equal to the integral of the limit:

$$(19.13) \quad \lim_{m \rightarrow \infty} F_i^{(m)'}(w) = F_i'(w).$$

It is evident that in case g is improper of the first kind this result is not affected by the circumstance that the points of discontinuity of $g_i(\theta)$ are not considered in the limit relation (19.9), since these points, being at most denumerably infinite in number, form a set of zero measure.

The result (19.11) now follows from (19.13) and the subsistence of (19.5) for every m .

20. The minimal surface is bounded by Γ . To show that the minimal surface whose existence is proved in the preceding section is bounded by Γ , we must prove that the representation (19.8) of Γ is proper. That it cannot be improper of the second kind is proved in §18, which, being based on the relation (19.11), applies here with full validity. We cannot however apply the argument of §17 to prove that (19.8) cannot be improper of the first kind. For although we would still have for a g of this kind $A(g) = +\infty$, it would not be true in the case of a general Jordan contour that $A(g)$ sometimes takes finite values.

We therefore use the following argument, based on the relation (19.11), to obtain the desired result. Suppose that under g the point P of C corresponds to the arc $Q'Q''$ of Γ . Since Γ is a Jordan curve, Q' and Q'' are distinct: and if a_i denote the coördinates of Q' , b_i of Q'' , the distance $Q'Q''$ or l with

$$(20.1) \quad l^2 = \sum_{i=1}^n (b_i - a_i)^2$$

is not equal to zero.

There is no loss of generality in supposing P to be at $w = 1$, for this may be achieved by a rotation of the unit circle, which changes nothing essential.

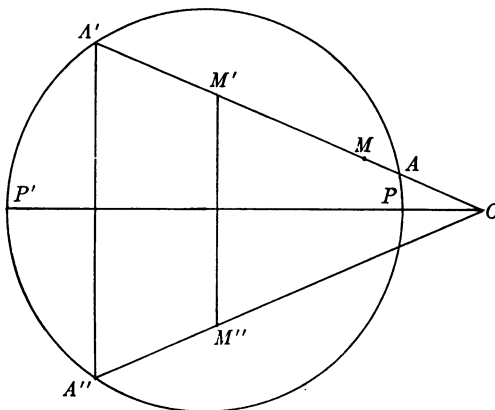


FIG. 4

Let O be any point on the diameter $P'P$ prolonged; then we define the point transformation $M \rightarrow M'$ (Fig. 4) by the condition

$$(20.2) \quad OM \cdot OM' = OA \cdot OA'.$$

If this be combined with a reflection $M' \rightarrow M''$ in $P'P$, the resulting transformation $M \rightarrow M''$, to be called \mathfrak{G} , is a conformal transformation converting the interior and circumference of the unit circle into themselves respectively. \mathfrak{G} has the linear fractional form

$$(20.3) \quad w' = \frac{aw + b}{cw + d}.$$

Suppose \mathfrak{G} to act on the functions $F_i(w)$ of (19.10) and on the boundary values $g_i(\theta)$ of $\Re F_i(w)$. Then $F_i(w)$ is transformed into

$$(20.4) \quad G_i(w) = F_i\left(\frac{aw + b}{cw + d}\right),$$

and the boundary values of $\Re G_i(w)$ are $h_i(\theta)$ defined by

$$(20.5) \quad h_i(A'') = h_i(A).$$

Differentiating (20.4), we find

$$G'_i(w) = \frac{ad - bc}{(cw + d)^2} F'_i\left(\frac{aw + b}{cw + d}\right),$$

whence

$$\sum_{i=1}^n G_i'^2(w) = \frac{(ad - bc)^2}{(cw + d)^4} \sum_{i=1}^n F_i'^2 \left(\frac{aw + b}{cw + d} \right).$$

Since, according to (19.11), $\sum_{i=1}^n F_i'^2(w) = 0$ identically in the interior of the unit circle, we have, also identically in the interior of the unit circle,

$$(20.6) \quad \sum_{i=1}^n G_i'^2(w) = 0.$$

Imagine now that O takes a sequence of positions tending to P as limit. Then we obtain a sequence of functions $G_i(w)$ constantly obeying (20.6). Consider the boundary values $h_i(\theta)$ of $\Re G_i(w)$ as defined by (20.5). It is evident that if A'' is any fixed point of the circumference, its image A by \mathfrak{G}^{-1} tends to P from below or above according as A'' lies on the upper or lower semi-circumference PP' . Hence, on account of the distinct limiting values, a_i and b_i , of $g_i(\theta)$ on the different sides of P , it follows that $h_i(\theta)$ tends to a function equal to the constant a_i on the upper semi-circumference, and to the constant b_i on the lower semi-circumference.

We have by (19.12)

$$G_i'(w) = \frac{1}{\pi} \int_C \frac{e^{i\theta}}{(e^{i\theta} - w)^2} h_i(\theta) d\theta.$$

Since $h_i(\theta)$, because it always represents Γ , remains uniformly bounded, we may, as in connection with (19.12), pass to the limit under the integral sign; and so find that with the approach of O to P ,

$$\begin{aligned} G_i'(w) &\rightarrow \frac{1}{\pi} \int_0^\pi \frac{e^{i\theta}}{(e^{i\theta} - w)^2} a_i d\theta + \frac{1}{\pi} \int_\pi^{2\pi} \frac{e^{i\theta}}{(e^{i\theta} - w)^2} b_i d\theta \\ &= \frac{2i(b_i - a_i)}{\pi} \cdot \frac{1}{1 - w^2}. \end{aligned}$$

Hence, with the notation (20.1),

$$(20.7) \quad \sum_{i=1}^n G_i'^2(w) \rightarrow \frac{-4l^2}{\pi^2} \cdot \frac{1}{(1 - w^2)^2}.$$

Here we have contradiction, for since $l \neq 0$, the last condition is incompatible with the fact that in the passage to the limit (20.6) is constantly obeyed.

Hence the representation $x_i = g_i(\theta)$ of Γ is proper, and the minimal surface determined by it is bounded by Γ .

III. THE RELATION OF $A(g)$ TO THE AREA FUNCTIONAL

It is customary to characterize a minimal surface by the property of rendering area a minimum. Instead, we have minimized the simpler functional $A(g)$. This leads to the question of the relationship between these two functionals, with which this Part is occupied. We shall develop certain formulas interesting in themselves, and useful in Parts IV and V. The reader who is interested in the application of the present theory to conformal mapping and in the proof of the least-area property of the minimal surface whose existence has just been established may go on at once to Parts IV and V, referring back to this Part for the necessary results.

21. Other forms of $A(g)$. In this section we derive two further formulas for $A(g)$. The first of these is

$$(21.1) \quad A(g) = \iint_D \frac{1}{2} \sum_{i=1}^n |F'_i(w)|^2 d\sigma,$$

where D denotes the interior of the unit circle, of which $d\sigma$ is the element of area. The second expresses $A(g)$ in terms of the Fourier coefficients of g :

$$(21.2) \quad A(g) = \frac{\pi}{2} \sum_{m=1}^{\infty} m \sum_{i=1}^n (a_{im}^2 + b_{im}^2).$$

Proof of (21.1). Since the integrand of (21.1) may not be bounded in D , we need the following limit process to define the double integral. Let D_ρ denote the interior of the circle C_ρ of radius $\rho < 1$ concentric with C . Then

$$\iint_{D_\rho} \frac{1}{2} \sum_{i=1}^n |F'_i(w)|^2 d\sigma$$

is evidently an increasing function of ρ , since the element of integration is positive. Hence the limit as $\rho \rightarrow 1$ of this double integral exists, finite or infinite, and we define

$$(21.3) \quad \iint_D \frac{1}{2} \sum_{i=1}^n |F'_i(w)|^2 d\sigma = \lim_{\rho \rightarrow 1} \iint_{D_\rho} \frac{1}{2} \sum_{i=1}^n |F'_i(w)|^2 d\sigma.$$

Suppose $F_i(w)$ separated into its real and imaginary parts:

$$F_i(w) = U_i(u, v) + iV_i(u, v) \quad (w = u + iv).$$

Then

$$(21.4) \quad F'_i(w) = \frac{\partial U_i}{\partial u} - i \frac{\partial U_i}{\partial v}.$$

Therefore

$$(21.5) \quad \sum_{i=1}^n |F'_i(w)|^2 = \sum_{i=1}^n \left[\left(\frac{\partial U_i}{\partial u} \right)^2 + \left(\frac{\partial U_i}{\partial v} \right)^2 \right].$$

Since U_i is a harmonic function, we have by Green's formula†

$$\iint_{D_\rho} \left[\left(\frac{\partial U_i}{\partial u} \right)^2 + \left(\frac{\partial U_i}{\partial v} \right)^2 \right] d\sigma = \int_{C_\rho} U_i \frac{\partial U_i}{\partial \rho} ds.$$

Whence, applying (21.5),

$$(21.6) \quad \iint_{D_\rho} \frac{1}{2} \sum_{i=1}^n |F'_i(w)|^2 d\sigma = \frac{1}{2} \int_0^{2\pi} \sum_{i=1}^n U_i \frac{\partial U_i}{\partial \rho} \rho d\theta.$$

For U_i there is the formula

$$(21.7) \quad U_i = \frac{a_{i0}}{2} + \sum_{p=1}^{\infty} \rho^p (a_{ip} \cos p\theta + b_{ip} \sin p\theta),$$

derived by taking the real part of (11.1) after writing $w = \rho e^{i\theta}$. This series is uniformly convergent for all θ and for $\rho \leq \rho_0$, with ρ_0 fixed and < 1 . This observation makes it legitimate to perform term by term the operations of differentiation, multiplication and integration indicated in the second member of (21.6). We get first

$$(21.8) \quad \rho \frac{\partial U_i}{\partial \rho} = \sum_{q=1}^{\infty} q \rho^q (a_{iq} \cos q\theta + b_{iq} \sin q\theta),$$

and then, observing the relations

$$\int_0^{2\pi} \cos p\theta \sin q\theta d\theta = 0, \quad \int_0^{2\pi} \cos p\theta \cos q\theta d\theta = \int_0^{2\pi} \sin p\theta \sin q\theta d\theta = \begin{cases} \pi & \text{if } p = q, \\ 0 & \text{if } p \neq q, \end{cases}$$

we obtain from (21.6), (21.7), (21.8):

$$(21.9) \quad \iint_{D_\rho} \frac{1}{2} \sum_{i=1}^n |F'_i(w)|^2 d\sigma = \frac{\pi}{2} \sum_{m=1}^{\infty} m \rho^{2m} \sum_{i=1}^n (a_{im}^2 + b_{im}^2).$$

This infinite series may be expressed as a double integral (formula (21.12)) in the following way. From the expressions (11.2) of the Fourier coefficients we easily derive in a manner analogous to (11.9):‡

† See Picard, *Traité d'Analyse*, vol. 2 (3d edition, Paris, 1925), p. 22.

‡ Cf. Picard, *Traité d'Analyse*, vol. 1 (3d edition, Paris, 1922), p. 341.

$$\begin{aligned}
 (21.10) \quad m\rho^{2m} \sum_{i=1}^n (a_{im}^2 + b_{im}^2) \\
 = -\frac{1}{2\pi^2} \int_0^{2\pi} \int_0^{2\pi} \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \cdot m\rho^{2m} \cos m(\theta - \phi) d\theta d\phi.
 \end{aligned}$$

Now by writing

$$z = \rho^2 e^{i(\theta - \phi)}$$

in the formula

$$\sum_{m=1}^{\infty} m z^m = \frac{z}{(1-z)^2}$$

and taking real parts, we obtain after some reduction

$$(21.11) \quad \sum_{m=1}^{\infty} m\rho^{2m} \cos m(\theta - \phi) = -\rho^2 \frac{(1+\rho^2)^2 \sin^2 \frac{\theta - \phi}{2} - (1-\rho^2)^2 \cos^2 \frac{\theta - \phi}{2}}{\left[(1+\rho^2)^2 \sin^2 \frac{\theta - \phi}{2} + (1-\rho^2)^2 \cos^2 \frac{\theta - \phi}{2} \right]^2}.$$

The convergence, moreover, is uniform for all θ, ϕ if ρ be regarded as fixed < 1 , since the convergent series of constant positive terms

$$\sum_{m=1}^{\infty} m\rho^{2m}$$

is majorant for (21.11). Hence, after multiplying (21.11) by the bounded factor

$$-\frac{1}{2\pi^2} \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2,$$

which does not disturb the uniform convergence, we may integrate term by term, and get, with attention to (21.10), (21.9),

$$\begin{aligned}
 (21.12) \quad \iint_{D_\rho} \frac{1}{2} \sum_{i=1}^n |F_i'(w)|^2 d\sigma \\
 = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{2\pi} \sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2 \rho^2 \\
 \frac{(1+\rho^2)^2 \sin^2 \frac{\theta - \phi}{2} - (1-\rho^2)^2 \cos^2 \frac{\theta - \phi}{2}}{\left[(1+\rho^2)^2 \sin^2 \frac{\theta - \phi}{2} + (1-\rho^2)^2 \cos^2 \frac{\theta - \phi}{2} \right]^2} d\theta d\phi.
 \end{aligned}$$

If in the last expression we write formally $\rho = 1$, there results

$$(21.13) \quad \frac{1}{4\pi} \int_0^{2\pi} \int_0^{2\pi} \frac{\sum_{i=1}^n [g_i(\theta) - g_i(\phi)]^2}{4 \sin^2 \frac{\theta - \phi}{2}} d\theta d\phi = A(g).$$

To prove that (21.13) is the limit of (21.12) when $\rho \rightarrow 1$, we first observe that if the second member of (21.12) and the first of (21.13) be abbreviated respectively as

$$\frac{1}{4\pi} \iint J(\rho; \theta, \phi) d\theta d\phi, \quad \frac{1}{4\pi} \iint I(\theta, \phi) d\theta d\phi,$$

then for all ρ and $\theta \neq \phi$

$$(21.14) \quad |J(\rho; \theta, \phi)| < I(\theta, \phi),$$

since the ratio

$$\begin{aligned} \frac{|J(\rho; \theta, \phi)|}{I(\theta, \phi)} &= \frac{4\rho^2 \sin^2 \frac{\theta - \phi}{2}}{(1 + \rho^2)^2 \sin^2 \frac{\theta - \phi}{2} + (1 - \rho^2)^2 \cos^2 \frac{\theta - \phi}{2}} \\ &\quad \cdot \frac{\left| (1 + \rho^2)^2 \sin^2 \frac{\theta - \phi}{2} - (1 - \rho^2)^2 \cos^2 \frac{\theta - \phi}{2} \right|}{(1 + \rho^2)^2 \sin^2 \frac{\theta - \phi}{2} + (1 - \rho^2)^2 \cos^2 \frac{\theta - \phi}{2}} \end{aligned}$$

is less than 1 (each denominator minus the corresponding numerator gives a positive quantity).

Suppose, as first case, that $A(g)$ is finite. Then (21.14) expresses that in the approach of ρ to 1 the absolute value of $J(\rho; \theta, \phi)$ remains less than a summable function, and hence a theorem of Lebesgue† permits us to pass to the limit under the integral sign:

$$\lim_{\rho \rightarrow 1} \iint J(\rho; \theta, \phi) d\theta d\phi = \iint I(\theta, \phi) d\theta d\phi,$$

† Lebesgue, *Leçons sur l'Intégration* (2d edition, Paris, 1928), p. 131. In the present case all the functions involved are Riemann integrable (proper or improper), and the result in question may easily be established without recourse to the notion of Lebesgue integral.

that is,

$$(21.15) \quad \begin{aligned} \lim_{\rho \rightarrow 1} \iint_{D_\rho} \frac{1}{2} \sum_{i=1}^n |F'_i(w)|^2 d\sigma \\ \equiv \iint_D \frac{1}{2} \sum_{i=1}^n |F'_i(w)|^2 d\sigma = A(g). \end{aligned}$$

Thus (21.1) is proved in the case $A(g)$ finite. If $A(g) = +\infty$, a simple way of seeing that the \iint_D is then also $= +\infty$ is to observe that \iint_D , as the limit of \iint_{D_ρ} which tends to it in increasing, is a lower semi-continuous functional of g , like $A(g)$. This remains true if g is not restricted to represent Γ but may be any parameterized contour. With this understanding, a g such that $A(g) = +\infty$ may always be approached by a sequence of g 's such that $A(g)$ is finite. It is to be seen immediately that two lower semi-continuous functionals which coincide whenever the first has a finite value must also coincide in case this value is $+\infty$.

Proof of (21.2). Returning to (21.4) and (21.9), we have

$$\iint_D \frac{1}{2} \sum_{i=1}^n |F'_i(w)|^2 = \lim_{\rho \rightarrow 1} \frac{\pi}{2} \sum_{m=1}^{\infty} m \rho^{2m} \sum_{i=1}^n (a_{im}^2 + b_{im}^2).$$

Now

$$\lim_{\rho \rightarrow 1} \frac{\pi}{2} \sum_{m=1}^{\infty} m \rho^{2m} \sum_{i=1}^n (a_{im}^2 + b_{im}^2) = \frac{\pi}{2} \sum_{m=1}^{\infty} m \sum_{i=1}^n (a_{im}^2 + b_{im}^2).$$

If the latter series is convergent, the justification for this is Abel's theorem asserting the continuity of the power series in the first member at the point of convergence $\rho = 1$. If the second member equals $+\infty$, it is easy to show, by taking account of the positive nature of all the terms, that this is also the value of the first member.

Combining the last two equations with (21.1), we have the desired result:

$$A(g) = \frac{\pi}{2} \sum_{m=1}^{\infty} m \sum_{i=1}^n (a_{im}^2 + b_{im}^2).$$

22. The area functional $S(g)$ and its relation to $A(g)$. We have seen how every representation g of Γ determines a harmonic surface

$$x_i = \Re F_i(w) = U_i(u, v).$$

The linear element of this surface is

$$ds^2 = E du^2 + 2F du dv + G dv^2$$

with

$$(22.1) \quad E = \sum_{i=1}^n \left(\frac{\partial U_i}{\partial u} \right)^2, \quad F = \sum_{i=1}^n \frac{\partial U_i}{\partial u} \frac{\partial U_i}{\partial v}, \quad G = \sum_{i=1}^n \left(\frac{\partial U_i}{\partial v} \right)^2,$$

and its area is a functional of g which we denote by $S(g)$:

$$(22.2) \quad S(g) = \iint_D (EG - F^2)^{1/2} d\sigma.$$

We have from (21.5)

$$\sum_{i=1}^n |F'_i(w)|^2 = \sum_{i=1}^n \left(\frac{\partial U_i}{\partial u} \right)^2 + \sum_{i=1}^n \left(\frac{\partial U_i}{\partial v} \right)^2 = E + G,$$

so that by (21.1)

$$(22.3) \quad A(g) = \iint_D \frac{1}{2}(E + G) d\sigma.$$

To the end of comparing the integrands in (22.2), (22.3), we observe

$$\frac{1}{4}(E + G)^2 - (EG - F^2) = \frac{1}{4}(E - G)^2 + F^2 \geq 0;$$

therefore

$$(22.4) \quad \frac{1}{2}(E + G) \geq (EG - F^2)^{1/2},$$

the equality holding when and only when

$$(22.5) \quad E - G = 0, \quad F = 0.$$

Since by (21.4) and (22.1),

$$\sum_{i=1}^n F_i'^2(w) = (E - G) - 2iF,$$

the conditions (22.5) are equivalent to

$$(22.6) \quad \sum_{i=1}^n F_i'^2(w) = 0,$$

characteristic of a minimal surface.

Consequently,

$$(22.7) \quad A(g) \geq S(g),$$

and the equality holds when and only when the harmonic surface determined by g is minimal.

IV. THE RIEMANN MAPPING THEOREM AND THE THEOREM OF OSGOOD AND CARATHÉODORY

23. The case $n=2$ of the problem of Plateau. Let the contour Γ be any Jordan curve in the plane (x_1, x_2) , Δ the region bounded by Γ , C the unit

circle, and D the interior of the unit circle. Then the classic Riemann mapping theorem states the existence of a one-one continuous and conformal correspondence between D and Δ . According to a theorem of Osgood[†] and Carathéodory,[†] it is possible to supplement this conformal map with a one-one continuous correspondence between C and Γ , so that the combination is one-one and continuous between $D+C$ and $\Delta+\Gamma$.

We show in this Part how by merely writing $n=2$ in the preceding work we have an immediate proof of the theorem of Riemann together with the theorem of Osgood-Carathéodory.

In the cited papers of the last two authors a sharp distinction is drawn between the "interior problem" and the "boundary problem." The existence of a conformal map of the interiors is supposed already established by the classic methods, and these authors then proceed to prove that this map of the interiors induces by continuity a topological correspondence between the boundaries.

It is characteristic of the method of the present paper to follow a directly opposite procedure: namely, we first distinguish a certain topological correspondence between the boundaries by the property of rendering $A(g)$ a minimum; this topological correspondence found, the conformal map of the interiors can be expressed immediately (see the theorem stated at the end of this Part).

The work of Parts I and II, with $n=2$, assures us of the existence of a certain proper representation of Γ ,

$$(23.1) \quad x_1 = g_1^*(\theta), \quad x_2 = g_2^*(\theta),$$

such that if

$$(23.2) \quad x_1 = \Re F_1(w), \quad x_2 = \Re F_2(w)$$

are the harmonic functions determined by the boundary values (23.1), we have

$$(23.3) \quad F_1'^2(w) + F_2'^2(w) = 0.$$

The functions F_1, F_2 are given by the formula

$$(23.4) \quad F_1(w) = \frac{1}{2\pi} \int_C \frac{e^{i\theta} + w}{e^{i\theta} - w} g_1^*(\theta) d\theta, \quad F_2(w) = \frac{1}{2\pi} \int_C \frac{e^{i\theta} + w}{e^{i\theta} - w} g_2^*(\theta) d\theta.$$

From (23.3),

$$F_1'(w) = \pm i F_2'(w),$$

[†] Reference in the Introduction.

and choosing the $+$ sign (the $-$ sign will lead to an inversely conformal transformation, easily discussed), we have by integration

$$F_1(w) = iF_2(w) + a + ib$$

where a, b are real constants. Separating F_1, F_2 into their real and imaginary parts:

$$F_1 = U_1 + iV_1, \quad F_2 = U_2 + iV_2,$$

this gives

$$U_1 = -V_2 + a, \quad V_1 = U_2 + b.$$

Consequently,

$$x_1 + ix_2 = U_1 + iU_2 = U_1 + iV_1 - ib = F_1(w) - ib = iF_2(w) + a.$$

Denote by $F(w)$ the common value of the last two expressions:

$$(23.5) \quad F(w) = F_1(w) - ib, \quad F(w) = iF_2(w) + a;$$

then

$$(23.6) \quad x_1 + ix_2 \equiv W = F(w),$$

a holomorphic function of w in the interior D of the unit circle. It will therefore be proved that the transformation defined by (23.6) or (23.2) is conformal in the domain D after we have shown a little later that $F'(w) \neq 0$ at any point of D .

We will first prove that (23.6) maps D in a one-one way on Δ . To this end, let W_0 be any point in the complex plane $x_1 + ix_2 \equiv W$ not on Γ ; what has to be shown is that the equation

$$(23.7) \quad F(w) = W_0$$

has exactly one solution w in D if W_0 belongs to Δ , and no solution w in D if W_0 does not belong to Δ .

Certainly there are only a finite number of solutions of (23.7) in any circle concentric with and smaller than C ; therefore we can construct a sequence of circles C_ρ concentric with C and with radii ρ increasing to 1 as limit, such that no solution of (23.7) lies on a circumference C_ρ . The number of solutions of (23.7) in the interior of C_ρ is given by the formula of Cauchy:

$$(23.8) \quad N_\rho = \frac{1}{2\pi i} \int_{C_\rho} \frac{F'(w)dw}{F(w) - W_0},$$

applicable here with full validity because C_ρ is *interior* to a simply connected domain of regularity of $F(w)$. The number of solutions of (23.7) in the interior of C is evidently

$$(23.9) \quad N = \lim_{\rho \rightarrow 1} N_\rho.$$

With $W = F(w)$, formula (23.8) gives

$$(23.10) \quad N_\rho = \frac{1}{2\pi i} [\log (W - W_0)]_{\Gamma_\rho} = \text{order of } W_0 \text{ with respect to } \Gamma_\rho.$$

Here Γ_ρ denotes the closed analytic curve† which is the image of C_ρ by the transformation $W = F(w)$, the bracket denotes the variation of $\log (W - W_0)$ when W describes Γ_ρ , and the order of W_0 with respect to Γ_ρ is an integer equal to $1/(2\pi)$ times the variation in the angle made by the vector W_0W with a fixed direction, followed continuously while W describes Γ_ρ .

Now when $\rho \rightarrow 1$, Γ_ρ tends uniformly to Γ , for the formulas (23.2), (23.4) are equivalent to Poisson's integral, and the boundary functions (23.1) are continuous. Evidently then, the order of W_0 with respect to Γ_ρ tends to the order of W_0 with respect to Γ ; indeed, for ρ near enough to 1 the former remains equal to the latter. Hence by (23.9), (23.10),

$$(23.11) \quad N = \text{order of } W_0 \text{ with respect to } \Gamma = \begin{cases} 1 & \text{if } W_0 \text{ is interior to } \Gamma, \\ 0 & \text{if } W_0 \text{ is exterior to } \Gamma. \end{cases}$$

According to this, the image of a point w of D is never a point exterior to Γ . But neither can it be a point Q of Γ . For then a neighborhood of w would go over into a neighborhood of Q ;‡ now every neighborhood of Q contains points exterior to Γ , so that we would have contradiction with the first statement of this paragraph. Hence the image of any point w of D is a point interior to Γ . Furthermore, by (23.11) every point interior to Γ is obtained, and exactly once, as image of a point w . Therefore the transformation $W = F(w)$ is one-one as between D and Δ .

To prove that this transformation is conformal without singular points, we must show that we cannot have $F'(w) = 0$ at any point w of D . If $F'(w) = 0$, then a neighborhood of w is mapped on a multiply-covered neighborhood of W ,§ but this contradicts the proof just given of the one-one nature of the correspondence between D and Δ .

That the conformal correspondence thus established between D and Δ attaches continuously to the topological correspondence (23.1) between C and Γ is an immediate consequence of the remark, already made, that the

† It will result from the sequel that Γ_ρ does not intersect itself but this fact is not necessary for the present argument.

‡ This is by the region-preserving property (Gebietstreu) of transformations $W = F(w)$; cf. L. Bieberbach, *Lehrbuch der Funktionentheorie*, Leipzig and Berlin, 1921, pp. 187-188.

§ Bieberbach, loc. cit., p. 188.

formulas (23.2), (23.4) are equivalent to Poisson's integral based on the continuous boundary functions (23.1).

In sum, we have proved the combined theorems of Riemann and Osgood-Carathéodory.

Expression for $F(w)$. An expression for $F(w)$ in the Cauchy form, more elegant than (23.2), (23.4), where real and imaginary parts are separated, may be obtained as follows.

Let w be a fixed point of D ; take $\rho > |w|$ and < 1 ; then by the formula of Cauchy,

$$F(w) = \frac{1}{2\pi i} \int_{C_\rho} \frac{F(z)dz}{z-w}.$$

If now $\rho \rightarrow 1$, then $F(z)$ tends uniformly to

$$f^*(z) = g_1^*(\theta) + ig_2^*(\theta)$$

and $1/(z-w)$ (z on C_ρ) tends uniformly to $1/(z-w)$ (z on C), wherein corresponding points on C_ρ and C are those with the same angular coördinate. Therefore

$$(23.12) \quad F(w) = \frac{1}{2\pi i} \int_C \frac{f^*(z)dz}{z-w}.$$

24. Range of values of $A(g)$. For a Jordan curve *in the plane*, we can easily obtain the exact range of values of $A(g)$, and see that finite values always occur among them.

For since the functions $F_1(w)$, $F_2(w)$ determined by the representation g^* (23.1) obey the condition (23.3) $F_1'^2(w) + F_2'^2(w) = 0$, we have by the final statement of Part III:

$$A(g^*) = \text{inner area of } \Gamma.$$

The *inner* area[†] must be taken because $S(g)$, as defined by (22.2), is the limit of the area bounded by Γ_ρ , which approaches to Γ from its interior.

To see that $A(g)$ takes every value in the interval

$$\text{inner area of } \Gamma \leq A(g) \leq +\infty$$

(and, of course, no other values), consider a continuous series of representations g connecting g^* with a representation g^0 such that $A(g^0) = +\infty$ (example: g^0 improper of the first kind); it is easy to arrange that $A(g)$ be continuous on this series of g 's.

[†] The region bounded by a Jordan curve has in general distinct inner and outer areas, differing by an amount called the exterior content of the curve. The first example of a Jordan curve of positive exterior content was given by Osgood, these Transactions, vol. 4 (1903), pp. 107-112.

25. The combined interior-boundary conformal mapping theorem. The results of this Part are summarized in the following theorem, combining the theorems of Riemann and Osgood-Carathéodory.

THEOREM. *Let Γ denote any Jordan curve in the plane, and*

$$x_1 = g_1(\theta), \quad x_2 = g_2(\theta),$$

or

$$Z = f(z)$$

with

$$Z = x_1 + ix_2, \quad z = e^{i\theta},$$

an arbitrary representation of Γ as topological image of the unit circle C .

The range of values of the functional

$$\begin{aligned} A(g) &= \frac{1}{16\pi} \int_C \int_C \frac{\sum_{i=1}^2 [g_i(\theta) - g_i(\phi)]^2}{\sin^2 \frac{\theta - \phi}{2}} d\theta d\phi \\ &= \frac{1}{4\pi} \int_C \int_C \frac{|f(z) - f(\xi)|^2}{|z - \xi|^2} |dz| |d\xi| \end{aligned}$$

when g or f varies over all possible representations of Γ consists of all positive real numbers (including $+\infty$) greater than or equal to the inner area of the region enclosed by Γ . This minimum value is attained for a certain proper representation

$$Z = f^*(z)$$

(determined up to linear fractional transformation of C into itself,

$$z' = (az + b)/(\bar{b}z + \bar{a}).$$

Then the transformation $w \rightarrow W$ defined by

$$W = \frac{1}{2\pi i} \int_C \frac{f^*(z) dz}{z - w}$$

produces a one-one, continuous and conformal map of the interior of Γ on the interior of C , which attaches continuously to the topological correspondence $Z = f^(z)$ between Γ and C .*

The implications of this theorem for the Dirichlet problem are apparent. The Dirichlet problem for any continuous distribution of assigned values on a Jordan boundary is reduced to the same problem for a circle, and therefore solved by Poisson's integral. The Dirichlet functional

$$D(\phi) = \iint_{\Delta} \left[\left(\frac{\partial \phi}{\partial x_1} \right)^2 + \left(\frac{\partial \phi}{\partial x_2} \right)^2 \right] dx_1 dx_2$$

is replaced by the functional $A(g)$, much simpler to deal with chiefly because the range of g is a compact closed set while the range of ϕ is not. It was the latter fact alone which rendered valid the criticism of Weierstrass against Riemann's treatment of the Dirichlet problem based on minimizing the Dirichlet functional.

V. ABSOLUTE MINIMUM OF AREA

26. **Proof of the least-area property.** We conclude this paper with the following brief proof, based on the formulas of Part III, that the minimal surface whose existence is proved in Part I has the least area of any surface bounded by the given contour.

Let the contour Γ be, first, a polygon. Then let Π be any polyhedral surface bounded by Γ . By the conformal mapping theorem of Koebe,[†] Π can be mapped conformally on the interior of the unit circle, the map attaching continuously to a topological correspondence between Γ and the circumference. Let the parametric equations of Π determined by this map be

$$x_i = x_i(u, v)$$

and of Γ

$$(26.1) \quad x_i = g_i(\theta).$$

The element of length of Π is

$$ds^2 = Edu^2 + 2F du dv + G dv^2$$

with

$$E = \sum_{i=1}^n \left(\frac{\partial x_i}{\partial u} \right)^2, \quad F = \sum_{i=1}^n \left(\frac{\partial x_i}{\partial u} \right) \left(\frac{\partial x_i}{\partial v} \right), \quad G = \sum_{i=1}^n \left(\frac{\partial x_i}{\partial v} \right)^2;$$

and by the conformality

$$E = G, \quad F = 0;$$

so that the area of Π is

$$(26.2) \quad \begin{aligned} S(\Pi) &= \iint (EG - F^2)^{1/2} du dv = \frac{1}{2} \iint (E + G) du dv \\ &= \frac{1}{2} \sum_{i=1}^n \iint \left\{ \left(\frac{\partial x_i}{\partial u} \right)^2 + \left(\frac{\partial x_i}{\partial v} \right)^2 \right\} du dv. \end{aligned}$$

[†] Cf. T. Radó, *Annals of Mathematics*, (2), vol. 31 (1930), pp. 458-460.

Consider the harmonic surface determined by the representation (26.1) of Γ , and denote by \bar{E} , \bar{F} , \bar{G} its fundamental quantities; then since a harmonic function gives the least value to the Dirichlet functional for fixed boundary values, we have

$$(26.3) \quad \frac{1}{2} \iint (\bar{E} + \bar{G}) du dv \leq \frac{1}{2} \iint (E + G) du dv.$$

By formula (22.3) of Part III,

$$(26.4) \quad A(g) = \frac{1}{2} \iint (\bar{E} + \bar{G}) du dv;$$

and by the minimum property of g^* ,

$$(26.5) \quad A(g^*) \leq A(g).$$

By the chain (26.2)–(26.5),

$$(26.6) \quad A(g^*) \leq S(\Pi):$$

the minimum value of $A(g) \leq$ the area of any polyhedral surface bounded by the polygon Γ .

Let now Γ denote an arbitrary Jordan curve and Σ any simply-connected surface bounded by Γ . According to the Lebesgue definition, the area $S(\Sigma)$ of Σ is the lower limit (finite or $+\infty$) of the areas of polyhedral surfaces which tend to Σ . There exists, then, a sequence of polyhedral surfaces Π_m tending to Σ such that

$$(26.7) \quad S(\Pi_m) \rightarrow S(\Sigma);$$

the bounding polygons $\Gamma^{(m)}$ tend to Γ .

Each polygon $\Gamma^{(m)}$ has a representation

$$(26.8) \quad x_i = g_i^{(m*)}(\theta)$$

minimizing $A(g)$ for that polygon, and by the procedure of §19 we can select a sub-sequence

$$(26.9) \quad m = m_1, m_2, \dots, m_k, \dots$$

so that (26.8) tends to a proper representation of Γ

$$(26.10) \quad x_i = g_i^*(\theta).$$

The harmonic surface determined by g^* will be minimal; consequently, by (22.7),

$$(26.11) \quad S(g^*) = A(g^*).$$

From the sequence (26.9) we can in turn select a sub-sequence

$$(26.12) \quad m = m'_1, m'_2, \dots, m'_k, \dots$$

so that $A(g^{(m)*})$ tends to a limit (finite or $+\infty$):

$$(26.13) \quad A(g^{(m)*}) \rightarrow L.$$

By the lower semi-continuity of $A(g)$, which holds for any sort of approach of one parameterized contour to another, as well when the contour itself is allowed to vary as when we have merely different parameterizations of the same contour, it follows (see (7.2)) that

$$(26.14) \quad A(g^*) \leq L.$$

By (26.6),

$$A(g^{(m)*}) \leq S(\Pi_m);$$

hence

$$\lim A(g^{(m)*}) \leq \lim S(\Pi_m),$$

that is

$$(26.15) \quad L \leq S(\Sigma).$$

Combining this with (26.14) and (26.11), we obtain

$$(26.16) \quad S(g^*) \leq S(\Sigma),$$

which was to be proved.

It is easy to see that the g^* here obtained from approaching polygons is the same as the g^* which minimizes $A(g)$ for Γ . For let g^{**} minimize $A(g)$ for Γ ; then

$$(26.17) \quad A(g^{**}) \leq A(g^*).$$

By (26.16),

$$(26.18) \quad S(g^*) \leq S(g^{**}),$$

and by (22.7)

$$A(g^{**}) = S(g^{**}) \text{ as well as } A(g^*) = S(g^*),$$

so that (26.18) may be expressed as

$$(26.19) \quad A(g^*) \leq A(g^{**}).$$

From (26.19) and (26.17) it follows that

$$A(g^*) = A(g^{**}).$$

27. A non-finite-area-spanning Jordan contour. It is important to note that if the Jordan contour is sufficiently tortuous, the least-area property of the minimal surface may become vacuous through the impossibility of spanning any finite area whatever within the given contour. The following example was constructed by the author in collaboration with P. Franklin.

Consider a broken line, broken at right angles and in the form of a spiral, whose successive segments have the lengths $1, 1/2^{1/2}, 1/3^{1/2}, \dots$. On each segment construct a square lying towards the outside of the spiral. In each square let there be a Peano curve starting at the initial point and ending at the terminal point of the corresponding segment. Let the unit interval $0 \leq t \leq 1$ be divided by the points $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots$ into a denumerable infinity of sub-intervals, and let the n th Peano curve be represented as continuous image of the n th interval. If to the equations $x = \phi(t), y = \psi(t)$, representing all these Peano curves laid end to end, we adjoin $z = t$, then we have a Jordan arc in space, being a one-one continuous image of the unit t -interval, end point $t = 1$ included. The desired example is formed of the four Jordan arcs $x = \phi(t), y = \psi(t), z = t; x = \phi(t), y = \psi(t), z = 2 - t; x = -\phi(t), y = -\psi(t), z = t; x = -\phi(t), y = -\psi(t), z = 2 - t$.

The proof results from the fact that the content of the orthogonal projection of this Jordan curve on the xy -plane, counting each point with its proper multiplicity, is four times the sum of the harmonic series. A fortiori, the content of the orthogonal projection of any surface spanned with the curve is $+\infty$, and this is, a fortiori, the area of the surface.

It thus appears that the separation of the existence proof into Parts I, II is inherent in the very nature of the problem, since it is futile to try to create distinctions with a functional which is identically $+\infty$. The limit process is absolutely essential for a non-finite-area-spanning Jordan contour, and the minimal surface can then be characterized only by the Weierstrass equations, the minimum-area property becoming meaningless.†

The corresponding situation in the Dirichlet problem is well known, having been pointed out by Hadamard‡ with the following example. If boundary values on the unit circle are defined by

$$f(\theta) = \sum_{p=1}^{\infty} \frac{\cos 2^p \theta}{2^p},$$

then the Dirichlet functional is identically $+\infty$, but a harmonic function as defined by Laplace's equation exists, being

$$\sum_{p=1}^{\infty} \rho^{2^p} \frac{\cos 2^p \theta}{2^p}.$$

† Bulletin de la Société Mathématique de France, vol. 34 (1906), pp. 135-139.

‡ However, a good sense in which the least-area property continues to hold will be given in a supplementary note.