AUTOMORPHIC FUNCTIONS

BY

LESTER R. FORD, PH.D.,
Assistant Professor of Mathematics in the Rice Institute

This volume was written during the tenure of a National Research Fellowship

FIRST EDITION

McGRAW-HILL BOOK COMPANY, INC.
NEW YORK: 370 SEVENTH AVENUE
LONDON: 6 & 8 BOUVERIE ST., E. C. 4
1929

AUTOMORPHIC FUNCTIONS

PRINTED IN THE UNITED STATES OF AMERICA

THE MAPLE PRESS COMPANY, YORK, PA.

PREFACE

In 1915, the author published in the series of *Edinburgh Mathematical Tracts* a brief introduction to "Automorphic Functions." This booklet has been long out of print. Except for this little volume, no book on the subject has ever appeared in English. This is regrettable, in view of the importance of the subject to those whose interests lie in the field of Functions of a Complex Variable and of its numerous contacts with other domains of mathematical thought.

It has been the author's aim in the earlier chapters to lay the foundations of the theory with all possible rigor and simplicity. The introduction and use of the isometric circle (the name was suggested by Professor Whittaker) have given the theory of linear groups a simplicity it has not had hitherto. The fundamental region which results was given by J. I. Hutchinson, in 1907, and independently by G. Humbert, in 1919. But the interesting properties of the circle itself and its utility in the derivation of the theory seem to have escaped the attention of workers in the field. It may be stated here, that much of the material involving the isometric circle embodies researches of the author which have not appeared elsewhere in published form.

In the later chapters, the author's task has been largely a matter of selection of material and method of treatment. Here, also, the use of the isometric circle has often led to a simplification of the proofs. The material included has been, perhaps, a matter of the author's personal taste. The classical elliptic modular functions deserve a place as the best-known examples of non-elementary automorphic functions. The theory of conformal mapping has been presented at some length—from the modern function-theory point of view—as a preparation for the theories of uniformization which follow. The chapter on Conformal Mapping can be read independently of the rest of the book. In the final chapter, the connection between automorphic functions and differential equations is brought out.

This treatment, which is necessarily brief, leads up to the triangle functions.

The connection between groups and non-Euclidean geometry has not been treated, as it now seems of less importance, in view of the way the foundations have been laid. A chapter on Abelian Integrals, treated in the light of uniformization, would have been of interest; however, the subject has not been worked out to any extent. This chapter should probably not be written until some geometer sets up a product for the prime function for the Fuchsian group of the first kind. Finally, the book might have been improved by a more extensive use of the theory of normal families of functions.

It is now almost fifty years since Poincaré created the general theory of automorphic functions, in a brilliant series of papers in the early volumes of *Acta Mathematica*. Since that time, the subject has had a steady growth. The material in the present volume will be found to spring very largely from researches of the past twenty years, either in content or in method of treatment. The theory of uniformization rests on the papers of Koebe and Poincaré, published in 1907. The first rigorous proof of the possibility of mapping one plane simply connected region upon another was presented by Osgood, in 1901, but the treatment of this and similar problems by function-theory methods came a dozen years later. Area theorems and other aspects of mapping are of still later date and are active subjects of investigation today. The foundations of the theory of groups, as previously stated, are based on the author's studies during the past few years.

The author wishes to thank Prof. E. T. Whittaker for various kindnesses while at the Mathematical Laboratory of the University of Edinburgh. He is indebted similarly to Profs. Otto Hölder and Paul Koebe, of the University of Leipzig. He received much profit from the lectures of Professor Hölder on "Elliptic Modular Functions" and from those of Professor Koebe on "Uniformization." So much of the material in the latter half of the book was gleaned from the published papers of Professor Koebe that specific references have often not been given and attention is called here to that fact. The author acknowledges a debt of long standing to Prof. W. F. Osgood, whose inspiring teaching first aroused an interest in the subject dealt with in this volume and whose "Funktionentheorie" has been a mine of usefulness.

The preceding gentlemen are in no wise responsible for such defects as the book possesses. Their encouragement and their suggestions have been valuable; but they are busy men, and the author has not presumed to ask them to read the manuscript. The only assistance in the task of writing the book was from Mr. Jack Kronsbein, a student of the University of Leipzig, who, as a labor of friendship, typed most of the manuscript and checked many of the formulæ.

The author's chief debts are to the National Research Council, whose grant of a fellowship made the writing of the book possible, and to the Rice Institute, which gave him leave of absence.

L. R. F.

HOUSTON, TEXAS.
June, 1929.

CONTENTS

CHAPTER IV

AUTOMORPHIC FUNCTIONS

CHAPTER V

THE POINCARÉ THETA SERIES

CHAPTER VI

THE ELEMENTARY GROUPS

I. The Finite Groups

II. The Groups with One Limit Point

III. The Groups with Two Limit Points

CHAPTER VII

THE ELLIPTIC MODULAR FUNCTIONS

CHAPTER VIII

CONFORMAL MAPPING

CHAPTER IX

UNIFORMIZATION. ELEMENTARY AND FUCHSIAN FUNCTIONS

AUTOMORPHIC FUNCTIONS

CHAPTER I

LINEAR TRANSFORMATIONS

1. The Linear Transformation.—Let z and z' be two complex numbers connected by some functional relation, $z' = f(z)$. Let the values of z be represented in the customary manner on an Argand diagram, or z-plane, and the values of z' be represented on a second Argand diagram, or z'-plane. To each point z of the first plane for which the function is defined there correspond one or more values of z' by virtue of the functional relation. To points, curves, and areas of the z-plane there correspond, usually, points, curves, and areas in the z'-plane. We shall speak of the configurations in the z-plane as being *transformed* by the functional relation into the corresponding configurations in the z'-plane.

We shall find it convenient to represent z' and z on the same Argand diagram, rather than on different ones. Then the functional relation transforms configurations in the z-plane into other configurations in the z-plane. In what follows but one plane will be used unless the contrary is stated.

The whole theory of automorphic functions depends upon a particular type of transformation, defined as follows:

DEFINITION.—*The transformation*

$$z' = \frac{az + b}{cz + d}, \tag{1}$$

where a, b, c, d are constants and $ad - bc \neq 0$, is called a linear transformation.[1]

The present chapter will be devoted to a study of this fundamental transformation.

[1] This is more properly called a "linear fractional transformation"; but we shall use the briefer designation. It is also called a "homographic transformation."

If $ad - bc = 0$, the equation reduces to $z' =$ constant; but this case is without interest.

The quantity $ad - bc$ is called the *determinant* of the transformation. It will be convenient to have always

$$ad - bc = 1, \tag{2}$$

the determinant in the general case becoming 1 if the numerator and denominator of the fraction in the second member be divided by $\pm\sqrt{(ad - bc)}$.

The second member of (1) is an analytic function of z. The linear transformation has, therefore, the property of conformality; that is, when a figure is transformed, angles are preserved both in magnitude and in sign.

We note that for each value of z, equation (1) gives one and only one value of z'. There is no exception to this statement if we introduce the point at infinity. Thus if $c \neq 0$, $z = -d/c$ is transformed into $z' = \infty$, and $z = \infty$ into $z' = a/c$; if $c = 0$, $z = \infty$ is transformed into $z' = \infty$.

Let equation (1) be solved for z:

$$z = \frac{-dz' + b}{cz' - a}. \tag{3}$$

This transformation which, applied after the transformation (1) has been made, carries each configuration back into its original position is called the *inverse* of the transformation (1). We note that (3) is a linear transformation; hence,

THEOREM 1.—*The inverse of a linear transformation is a linear transformation.*

We note that (3) is formed from (1) by interchanging a and d with a change of sign. When formed in this way the determinant is the same as in (1).

We see from (3) that to each value of z' there corresponds one and only one value of z. We have, then, the following result:

THEOREM 2.—*The z-plane is transformed into itself in a one-to-one manner by a linear transformation.*

Moreover, the linear transformation is the most general analytic transformation which has the property stated in Theorem 2. We shall prove first the following theorem:

THEOREM 3.—*If, except for a finite number of points, the plane is mapped in a one-to-one and directly conformal manner upon a plane region, the mapping function is linear.*

Let $z' = f(z)$ be such a mapping function; and let $q_1, q_2, \ldots, q_n (= \infty)$ be the excepted points. Owing to the conformality, $f(z)$ is analytic except at the isolated points $q_1, \ldots, q_n$. Now q_i

is not an essential singularity, else the function takes on certain values an infinite number of times in the neighborhood of the point, which is contrary to hypothesis. Hence, $f(z)$ either remains finite in the neighborhood of q_i, and hence is analytic there if properly defined, or has a pole. So $f(z)$ is a rational function of z.

A rational function which is not a constant takes on every value m times, where m is the number of its poles. Since $f(z)$ takes on no value twice, it has a single pole of the first order. If the pole is at a finite point q_k we may write

$$z' = \frac{A_1}{z - q_k} + A_0 = \frac{A_0 z + A_1 - A_0 q_k}{z - q_k}, \quad A_1 \neq 0. \tag{4}$$

If the pole is at infinity, we have

$$z' = A_1 z + A_0, \quad A_1 \neq 0. \tag{4'}$$

In either case the function is linear.

COROLLARY 1.—*The most general one-to-one and directly conformal (where conformality has a meaning) transformation of the plane into itself is a linear transformation.*

We have not defined conformality when one of the points involved is the point at infinity. Excepting the point $z = \infty$ and the point of the z-plane which is carried into $z' = \infty$, the transformation is to be conformal. Theorem 3 then applies.

COROLLARY 2.—*The most general one-to-one and directly conformal transformation of the finite plane into itself is the linear transformation* $z' = A_1 z + A_0$.

We shall now consider the successive performance of linear transformations. After subjecting the z-plane to the transformation (1) let a second linear transformation

$$z'' = \frac{\alpha z' + \beta}{\gamma z' + \delta} \tag{5}$$

be made. Expressing z'' as a function of z, we have

$$z'' = \frac{\alpha \dfrac{az + b}{cz + d} + \beta}{\gamma \dfrac{az + b}{cz + d} + \delta} = \frac{(\alpha a + \beta c)z + \alpha b + \beta d}{(\gamma a + \delta c)z + \gamma b + \delta d}. \tag{6}$$

Making the transformation (1) and then making (5) is equivalent to making the single transformation (6). Now, (6) is a linear transformation; its determinant, in the form in which the fraction is written, is $(ad - bc)(\alpha\delta - \beta\gamma)$. It is worth noting that if the determinants of (1) and (5) are each unity, that of (6) is also unity without further change.

If z'' be subjected to a linear transformation, we conclude on combining the new transformation with (6) that the succession of three linear transformations is equivalent to a single linear transformation, and so on. We have then the following result:

THEOREM 4.—*The successive performance of a finite number of linear transformations is equivalent to a single linear transformation.*

A further well-known property of the linear transformation is expressed in the following theorem:

THEOREM 5.—*The linear transformation leaves invariant the cross-ratio of four points.*

Let z_1, z_2, z_3, z_4 be four distinct points and let z_1', z_2', z_3', z_4' be the points into which they are transformed by (1). We shall suppose all the points are finite. We have

$$z_1' - z_2' = \frac{az_1 + b}{cz_1 + d} - \frac{az_2 + b}{cz_2 + d} = \frac{(ad - bc)(z_1 - z_2)}{(cz_1 + d)(cz_2 + d)} = \frac{z_1 - z_2}{(cz_1 + d)(cz_2 + d)}; \tag{7}$$

whence

$$\frac{(z_1' - z_2')(z_3' - z_4')}{(z_1' - z_3')(z_2' - z_4')} = \frac{(z_1 - z_2)(z_3 - z_4)}{(z_1 - z_3)(z_2 - z_4)}. \tag{8}$$

If one of the points is at infinity, we make the necessary change in (8) by a limiting process. Thus, if $z_2 = \infty$ and $z_1' = \infty$, (8) becomes

$$\frac{z_3' - z_4'}{z_2' - z_4'} = -\frac{z_3 - z_4}{z_1 - z_3}.$$

2. Symbolic Notation.—For brevity in writing and for convenience in combination, we shall represent the second member of a transformation such as (1) by the functional notation, using capital letters for the function; thus,

$$T(z) = \frac{az + b}{cz + d},$$

so that (1) becomes

$$z' = T(z).$$

We shall speak of this as the transformation T, the argument z being omitted unless ambiguity might arise without it. If two transformations are the same, $T_1(z) \equiv T(z)$, this is indicated by the equation $T_1 = T$.

Let S be the transformation (5), so

$$z'' = S(z').$$

Then, (6) is

$$z'' = S[T(z)] = ST(z).$$

Thus, the succession of two transformations is written as a product, the enclosing brackets of the functional notation being omitted. It should be noted that ST is the single linear transformation resulting from making first the transformation T and then the transformation S, the order of performance being from right to left. TS is, in general, different from ST.

It is easily seen from the meaning of the symbols that the associative law of multiplication holds,

$$U(ST) = (US)T,$$

and there is no ambiguity in writing simply UST. In a product, any sequence of factors may be combined into a single linear transformation.

The transformations equivalent to performing T twice, thrice, etc., are represented by T^2, T^3, etc. Thus, $T^2(z)$ means $T[T(z)]$. The inverse of T is written T^{-1}; hence, from (3)

$$T^{-1}(z) = \frac{-dz + b}{cz - a}.$$

The result of applying the inverse n times is represented by T^{-n}. If we represent the identical transformation, $z' = z$, by 1 so that $1(z)$ means z, we observe that positive and negative integral powers of T together with unity combine in accordance with the law of the addition of exponents in multiplication.

The inverse of a sequence of transformations can now be written down. To find the inverse of ST we make on the plane transformed by ST the transformation S^{-1} followed by T^{-1}; we have

$$(T^{-1}S^{-1})(ST) = T^{-1}(S^{-1}S)T = T^{-1}T = 1.$$

Thus, $T^{-1}S^{-1}$ is the transformation which, applied after ST, carries each point back to its original position; so $T^{-1}S^{-1}$ is the inverse of ST. In a similar manner we have for any number of transformations

$$(ST \cdots UV)^{-1} = V^{-1}U^{-1} \cdots T^{-1}S^{-1}.$$

The rule for the transposition of factors by division is easily found. Let $UST = V$; then,

$$(UST)T^{-1} = VT^{-1}, \text{ or } US = VT^{-1},$$

and

$$U^{-1}(UST) = U^{-1}V, \text{ or } ST = U^{-1}V.$$

Then, in an equation connecting two products, the first (last) factor of one member can be transferred to the beginning (end) of the other member by changing the sign of its exponent. Thus, symbolic division is permissible, provided the proper order of performing the operations is followed. For example, the inverse W of $ST \cdots UV$ is a transformation such that

$$WST \cdots UV = 1.$$

By repeated division on the right we get the result given above.

3. The Fixed Points of the Transformation.—The points which are unchanged by the transformation (1) are found by setting $z' = z$ in (1) and solving the resulting equation.

$$z = \frac{az + b}{cz + d}, \quad \text{or} \quad cz^2 + (d - a)z - b = 0. \tag{9}$$

Suppose, first, that $c \neq 0$. Then, (9) has the two roots

$$\xi_1, \xi_2 = \frac{a - d \pm \sqrt{M}}{2c} \tag{10}$$

where

$$M = (d - a)^2 + 4bc = (a + d)^2 - 4. \tag{11}$$

The second expression for M is derived on the assumption that $ad - bc = 1$. We see from (1) that ∞ is not transformed into itself, so there are at most two fixed points. If $M = 0$, that is if $a + d = \pm 2$, there is but one fixed point,

$$\xi = \frac{a - d}{2c}. \tag{12}$$

If $c = 0$, we must have $a \neq 0$, $d \neq 0$ since, otherwise, the determinant would be zero. We see from (1) that ∞ is then a fixed point. Solving (9) we get a finite fixed point provided $a \neq d$. The fixed points are

$$\xi_1 = \frac{b}{d - a}, \quad \xi_2 = \infty. \tag{13}$$

If $c = 0$ and $a = d$, (1) takes the form

$$z' = z + b',$$

a translation with the single fixed point, $\xi = \infty$. In the case $c = 0$, we see from (11) that we have, as before, two fixed points if $M \neq 0$, and one fixed point if $M = 0$.

There cannot be more than two fixed points unless (9) is identically zero; that is, $c = 0$, $d = a$, and $b = 0$. Equation (1) then takes the form $z' = z$. Hence, THEOREM 6.—*The only linear transformation with more than two fixed points is the identical transformation* $z' = z$.

By means of this theorem we are able to prove the following important proposition:

THEOREM 7.—*There is one and only one linear transformation which transforms three distinct points,* z_1, z_2, z_3, *into three distinct points,* z_1', z_2', z_3'.

We shall prove first that there is not more than one such transformation. Let T be one transformation carrying z_1, z_2, z_3 into z_1', z_2', z_3', and let S be any other such transformation. Consider the transformation $T^{-1}S$. We have $S(z_1) = z_1'$ and $T^{-1}(z_1') = z_1$; so

$$T^{-1}S(z_1) = T^{-1}(z_1') = z_1.$$

Hence, z_1, and similarly z_2 and z_3, are fixed points of the transformation $T^{-1}S$. It follows from Theorem 6 that

$$T^{-1}S = 1;$$

whence applying the transformation T to both members,

$$S = T.$$

There is, thus, not more than one transformation of the kind required.

We shall prove that there is always one such transformation by actually setting it up. If none of the six values is infinite, consider the transformation defined by

$$\frac{(z' - z_1')(z_2' - z_3')}{(z' - z_2')(z_1' - z_3')} = \frac{(z - z_1)(z_2 - z_3)}{(z - z_2)(z_1 - z_3)}, \tag{14}$$

an equation which expresses the equality of the cross-ratios $(z'z_1', z_2'z_3')$ and (zz_1, z_2z_3). This is of the form (1) when solved for z' in terms of z. It obviously transforms z_1, z_2, z_3 into z_1', z_2', z_3'; for both members of (14) are equal to zero when $z = z_1$, $z' = z_1'$; they are both infinite when $z = z_2$, $z' = z_2'$; and they are both 1 when $z = z_3$, $z' = z_3'$.

If one of the given points is at infinity, we have but to replace the member of (14) in which that point occurs by its limiting value when the required variable becomes infinite. If $z_1 = \infty$, $z_2 = \infty$, or $z_3 = \infty$, we replace the second member of (14) by

$$-\frac{z_2 - z_3}{z - z_2}, \quad -\frac{z - z_1}{z_1 - z_3}, \quad \text{or } \frac{z - z_1}{z - z_2},$$

respectively; and a similar change is necessary in the first member for an infinite value of z_1', z_2', or z_3'. In any case, there is one transformation with the desired property, and the theorem is established.

Equation (14) is a convenient form for use in actually setting up the transformation carrying three given points into three given points. Theorem 7 will be of great utility in our subsequent work; to prove that two transformations are identical, we shall have merely to show that they transform three points in the same way.

4. The Linear Transformation and the Circle.—Since we are operating on the complex number z it will be convenient to have the equations of curves expressed directly in terms of z. If x is the real part of z and iy is its imaginary part, and if we represent by $\bar{z}$ the conjugate imaginary of z, we have

$$z = x + iy, \quad \bar{z} = x - iy. \tag{15}$$

From these we have

$$x = \tfrac{1}{2}(z + \bar{z}), \quad y = \frac{i}{2}(\bar{z} - z), \quad z\bar{z} = x^2 + y^2. \tag{16}$$

From the first two equations of (16) we can readily express the equation of any curve in terms of z and $\bar{z}$.

We shall now get the general equation of the circle and of the straight line. The equation

$$A(x^2 + y^2) + b_1x + b_2y + C = 0,$$

where the constants are real, is the general equation of the circle (possibly of imaginary or zero radius) if $A \neq 0$, and is the general equation of the straight line if $A = 0$ and b_1 and b_2 are not both zero. Substituting from (16) we have

$$Az\bar{z} + \tfrac{1}{2}(b_1 - ib_2)z + \tfrac{1}{2}(b_1 + ib_2)\bar{z} + C = 0.$$

Putting $B = \tfrac{1}{2}(b_1 - ib_2)$, this takes the form

$$Az\bar{z} + Bz + \bar{B}\bar{z} + C = 0 \tag{17}$$

where A and C are real. Equation (17) is the general equation of the circle if $A \neq 0$ and of the straight line if $A = 0$, $B \neq 0$.

The center and radius of the circle are easily found. Writing (17) in the form

$$\left(z + \frac{\bar{B}}{A}\right)\left(\bar{z} + \frac{B}{A}\right) = \frac{B\bar{B} - AC}{A^2},$$

we see that the first member is the square of the distance of z from $-\bar{B}/A$. Hence, (17) is a circle with center $-\bar{B}/A$ and radius $\sqrt{\dfrac{B\bar{B} - AC}{A^2}}$. Since we shall be interested only in real circles, we shall require that $B\bar{B} > AC$.

Now let us see what the circle or straight line (17) becomes when the linear transformation (1) is applied. Substituting from (3)

$$z = \frac{-dz' + b}{cz' - a}, \quad \bar{z} = \frac{-\bar{d}\bar{z}' + \bar{b}}{\bar{c}\bar{z} - \bar{a}},$$

we have

$$A\frac{(-dz' + b)(-\bar{d}\bar{z}' + \bar{b})}{(cz' - a)(\bar{c}\bar{z}' - \bar{a})} + B\frac{-dz' + b}{cz' - a} + \bar{B}\frac{-\bar{d}\bar{z}' + \bar{b}}{\bar{c}\bar{z}' - \bar{a}} + C = 0. \tag{18}$$

Clearing of fractions and collecting terms, we have

$$\begin{aligned} Ad\bar{d} - B\bar{c}d - \bar{B}c\bar{d} + Cc\bar{c}]z'\bar{z}' \\ + [-A\bar{b}d + B\bar{a}d + \bar{B}\bar{b}c - C\bar{a}c]z' \\ + [-Ab\bar{d} + Bb\bar{c} + \bar{B}a\bar{d} - Ca\bar{c}]\bar{z}' \\ + Ab\bar{b} - B\bar{a}b - \bar{B}a\bar{b} + Ca\bar{a} = 0. \end{aligned} \tag{19}$$

In this equation the coefficient of $z'\bar{z}'$ is real, for $d\bar{d}$ and $c\bar{c}$ are real, being each the product of a number by its conjugate; and $B\bar{c}d + \bar{B}c\bar{d}$ is real, being the sum of a number and its conjugate. Similarly the constant term is real. Also the coefficient of $\bar{z}'$ is the conjugate of the coefficient of z'. Therefore (19) is of the form (17). We have then the following result:

THEOREM 8.—*The linear transformation carries a circle or straight line into a circle or straight line.*

It will often be convenient to consider the straight line as a circle of infinite radius, in which case we say briefly that a circle is carried into a circle.

It is easy to see when the transform will be a straight line. The straight line is characterized by the fact that it passes through the point ∞. Hence, if the point which is carried to ∞ lies on the original circle or straight line, the transform will be a straight line; otherwise the transform will be a circle. This is easily shown analytically. For the coefficient of $z'\bar{z}'$ will vanish if (dividing by $c\bar{c}$)

$$A\left(-\frac{d}{c}\right)\left(-\frac{\bar{d}}{\bar{c}}\right) + B\left(-\frac{d}{c}\right) + \bar{B}\left(-\frac{\bar{d}}{\bar{c}}\right) + C = 0;$$

that is, if $-d/c$ lies on the original circle or straight line.

Since three points determine a circle, we can set up a transformation which carries three distinct points of the first circle into three distinct points of the second circle. Having chosen the first three points, the transformed points can be selected in an infinite variety of ways (∞^3 ways, in fact); and each different selection gives a different transformation. Hence,

THEOREM 9.—*There exist infinitely many linear transformations which transform a given circle into a second given circle.*

In particular, we may choose the second circle to be the same as the first. Hence, there are ∞^3 linear transformations which transform a given circle into itself.

5. Inversion in a Circle.—There is an intimate relation, as we shall now show, between the linear transformation of the complex variable and the geometrical transformation known as "inversion in a circle."

Consider a circle Q with center at K and radius r. Let P be any point of the plane and construct the half line KP, beginning at K and passing through P. Let P_1 be a point on the half line KP such that $KP_1 \cdot KP = r^2$; then P_1 is called the "inverse of P with respect to the circle Q." The relation is a reciprocal one; P is the inverse of P_1. We speak of P and P_1 as points inverse with respect to Q.

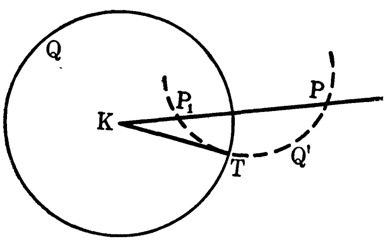


FIG. 1.

Inverse points have the property that any circle passing through P and P_1, the inverse of P with respect to Q, is orthogonal to Q. For, let Q' be any circle through P and P_1 and draw KT tangent to Q', T being the point of tangency. We have

$$\overline{KT}^2 = KP_1 \cdot KP = r^2,$$

whence, T lies on Q. The radii to T are perpendicular and the circles are orthogonal.

We shall now get an analytic expression for the transformation. Let P, P_1, and K be the points z, z_1, and k in the Argand diagram. The equations of the transformations are

$$|(z_1 - k)(z - k)| = r^2, \quad \arg(z_1 - k) = \arg(z - k).$$

The first equation expresses the condition $KP_1 \cdot KP = r^2$; the second expresses the collinearity of K, P, and P_1. Since $\arg(z - k) = -\arg(\bar{z} - \bar{k})$, the two equations are satisfied if, and only if,

$$(z_1 - k)(\bar{z} - \bar{k}) = r^2. \tag{20}$$

This is the equation of inversion in terms of complex variables.

If Q is the circle,

$$Az\bar{z} + Bz + \bar{B}\bar{z} + C = 0, \tag{17}$$

equation (20) becomes, on substituting the center and radius previously found,

$$\left(z_1 + \frac{\bar{B}}{A}\right)\left(\bar{z} + \frac{B}{A}\right) = \frac{B\bar{B} - AC}{A^2},$$

which on simplification is

$$Az_1\bar{z} + Bz_1 + \bar{B}\bar{z} + C = 0, \tag{21}$$

We thus get the relation between z and its inverse z_1 from the equation of Q by substituting z_1 for z and leaving $\bar{z}$ unchanged. Solving (21) we have the explicit form of the transformation,

$$z_1 = \frac{-\bar{B}\bar{z} - C}{A\bar{z} + B}. \tag{22}$$

When $A = 0$ so that (17) is a straight line, we shall still use formulæ (21) and (22) for the inversion. It is not difficult to show geometrically that when A approaches zero, P and P_1 attain positions such that Q is the perpendicular bisector of the segment $P\,P_1$. Inversion then becomes a reflection in the line Q. To show this analytically, let z_2 be a point on Q. Then, $|z_2 - z|$ is the distance from z_2 to z. The distance from z_2 to z_1, using the equation of the transformation and the equation of Q (with $A = 0$) which z_2 satisfies, is

$$|z_2 - z_1| = \left|\frac{-\bar{B}\bar{z}_2 - C}{B} + \frac{\bar{B}\bar{z} + C}{B}\right| = \left|\frac{\bar{B}}{B}(\bar{z} - \bar{z}_2)\right| = |z_2 - z|.$$

Thus all points of the line Q are equidistant from P and P_1.

We shall now prove the following proposition:

THEOREM 10.—*The linear transformation carries two points which are inverse with respect to a circle into two points which are inverse with respect to the transformed circle.*

Let z and z_1 be inverse with respect to the circle (17); then (21) is satisfied. Make the transformation (1) and let z', z_1', be the transformed points. We have

$$z_1 = \frac{-dz_1' + b}{cz_1' - a}, \quad \bar{z} = \frac{-\bar{d}\bar{z}' + \bar{b}}{\bar{c}\bar{z}' - \bar{a}};$$

whence, substituting in (21),

$$A\frac{(-dz_1' + b)(-\bar{d}\bar{z}' + \bar{b})}{(cz_1' - a)(\bar{c}\bar{z}' - \bar{a})} + B\frac{-dz_1' + b}{cz_1' - a} + \bar{B}\frac{-\bar{d}\bar{z}' + \bar{b}}{\bar{c}\bar{z}' - \bar{a}} + C = 0.$$

This equation is the same as (18) except that z' is replaced by z_1'; hence, on simplifying we shall get (19) with $\bar{z}'$ unchanged and z' replaced by z_1'. But this is the condition that z' and z_1' be inverse points with respect to the transformed circle (19).

Let us return to our study of the inversion. We see from (22) that to each z there corresponds one and only one z_1. Likewise to each z_1 corresponds one and only one z; hence, the transformation is one-to-one. The fixed points of the inversion are seen, from the geometrical construction, to be the points of Q itself. This appears analytically if we set $z_1 = z$ in (21).

The inversion (22) can be written as the succession of the two transformations

$$z_2 = \bar{z}, \quad z_1 = \frac{-\bar{B}z_2 - C}{Az_2 + B}.$$

The first of these is a reflection in the real axis; the second is a linear transformation. The first preserves the magnitudes of angles but reverses their signs; the second makes no alteration. Hence, the inversion is inversely conformal. This follows also from the fact that the second member of (22) is an analytic function of $\bar{z}$.

The reflection in the real axis obviously carries circles into circles, and the succeeding linear transformation does likewise. Again the reflection carries a circle and two points which are inverse with respect to it into a circle and two inverse points; and this relation, by Theorem 10, is preserved by the linear transformation.

We summarize the results in the following theorem:

THEOREM 11.—*Inversion in a circle is a one-to-one inversely conformal transformation which carries circles into circles, and carries two points inverse with respect to a circle into two points inverse with respect to the transformed circle.*

Since the inversion is a one-to-one transformation of the plane which preserves the magnitude of the angle but changes its sign, the result of performing two inversions, or any even number of inversions, is a one-to-one transformation which preserves both the magnitude and the sign of the angle. According to Theorem 3, Corollary 1, such a transformation is a linear transformation. Hence, we have the theorem:

THEOREM 12.—*The successive performance of an even number of inversions is equivalent to a linear transformation.*

We shall prove that, conversely, any linear transformation is so constituted. First, let us examine some of the simpler linear transformations and find equivalent pairs of inversions.

(*a*) *The translation*, $z' = z + b$.

By this transformation each point of the plane is translated parallel to the line Ob (Fig. 2) a distance equal to the length of Ob. Let L_1 and L_2 be two lines perpendicular to the line Ob and at a distance apart equal to half the length of Ob. A reflection in L_1 followed by a reflection in L_2, the lines being designated as in the figure, is equivalent to the given translation. It is sufficient to note that three points are transformed in the proper manner (Theorem 7). We observe at once that

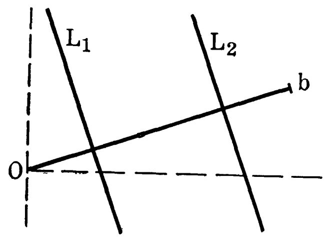


Fig. 2.

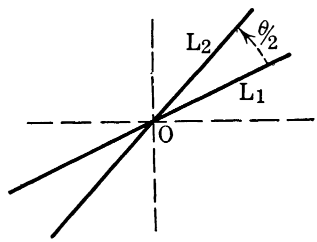


Fig. 3.

the points of L_1, which are unchanged by the first reflection, are translated in the desired manner by the second reflection.

(*b*) *The rotation*, $z' = e^{i\theta}z$.

Each point is rotated about the origin through an angle θ. A reflection in L_1, followed by a reflection in L_2, arranged as in Fig. 3, clearly rotates the points of L_1 as required; hence, the two reflections are equivalent to the desired rotation.

(*c*) *The stretching from the origin*, $z' = Az$, $A > 0$.

Each point is transformed into a point with the same argument, but with the modulus multiplied by A. There is a stretching from the origin, or, if $A < 1$, a contraction toward the origin. This is equivalent to an inversion in a circle Q_1 with center at the origin and radius r_1 (Fig. 4) followed by an inversion in a circle Q_2 with center at the origin and radius $r_2 = r_1\sqrt{A}$. For, if z_1, z' are the successive transforms of z, we have

$$z_1\bar{z} = r_1^2, \quad z'\bar{z}_1 = r_1^2A;$$

whence,

$$z' = \frac{r_1^2A}{\bar{z}_1} = r_1^2A \cdot \frac{z}{r_1^2} = Az.$$

(*d*) *The transformation* $z' = -1/z$.

This can be written

$$z_1 = -\bar{z}, \quad z' = \frac{1}{\bar{z}_1}.$$

It is, thus, a reflection in the imaginary axis, $z + \bar{z} = 0$, followed by an inversion in the unit circle, $z\bar{z} = 1$ (Fig. 5).

Consider now the general transformation (1). If $c \neq 0$, we can write this

$$z' - \frac{a}{c} = \frac{bc - ad}{c(cz + d)} = -\frac{1}{c^2\left(z + \frac{d}{c}\right)}, \tag{22'}$$

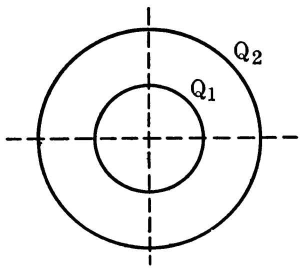


FIG. 4.

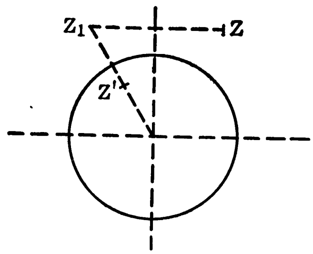


FIG. 5

supposing $ad - bc = 1$. This can be written as the following sequence of transformations:

$$z_1 = z + \frac{d}{c}, \quad z_3 = c^2 z_1, \quad z_4 = -\frac{1}{z_3}, \quad z' = z_4 + \frac{a}{c}.$$

The first, third, and fourth of these transformations have each been shown to be equivalent to a pair of inversions. The second can be broken into two, putting $c^2 = Ae^{i\theta}$, $A > 0$,

$$z_2 = e^{i\theta} z_1, \quad z_3 = A z_2,$$

each of which is equivalent to a pair of inversions.

If $c = 0$, the transformation has the form

$$z' = \alpha z + \beta.$$

Putting $\alpha = Ae^{i\theta}$, this is equivalent to the following sequence:

$$z_1 = e^{i\theta} z, \quad z_2 = A z_1, \quad z' = z_2 + \beta.$$

We have proved the converse of Theorem 12.

THEOREM 13.—*Any linear transformation is equivalent to the successive performance of an even number of inversions in circles.*

A linear transformation can be expressed as the sequence of inversions in an infinite number of ways. Later, we shall show that any linear transformation is equivalent to four suitably chosen inversions, and that a transformation of the class subsequently called "non-loxodromic" can be expressed as a succession of two inversions.

Let us consider now the most general one-to-one inversely conformal transformation of the plane,

$$z' = \overline{V}(z),$$

where, of course, $\overline{V}$ is not an analytic function of z. If we first make a reflection in the real axis, $z_1 = \bar{z}$, and then apply the preceding, we have a one-to-one directly conformal transformation, and hence, by Theorem 3, Corollary 1, a linear transformation of z. That is,

$$\overline{V}(z_1) = \overline{V}(\bar{z}) = \frac{az + b}{cz + d},$$

and

$$z' = \overline{V}(z) = \frac{a\bar{z} + b}{c\bar{z} + d}. \tag{23}$$

Inversion is seen to be a special case of this general transformation.

Equation (23) is a reflection followed by a linear transformation. So we can state the following general result:

THEOREM 14.—*The most general one-to-one conformal transformation of the plane into itself is equivalent to a succession of inversions in circles. The transformation is directly or inversely conformal according as the number of inversions is even or odd.*

6. The Multiplier, K.—We have already noted a separation of linear transformations into two classes. Putting aside the identical transformation $z' = z$, a transformation has either two fixed points or one fixed point. The number of fixed points and the behaviour of the transformation with reference to the fixed points furnish a useful basis of classification of linear transformations. We treat, first, the larger class with two fixed points.

Suppose, first, that in the transformation (1) $c \neq 0$. The finite points ξ_1, ξ_2 [Equation (10)] and the point ∞ are carried into ξ_1, ξ_2, and a/c, respectively. Hence, from (14), the transformation may be written

$$\frac{(z' - \xi_1)\left(\dfrac{a}{c} - \xi_2\right)}{(z' - \xi_2)\left(\dfrac{a}{c} - \xi_1\right)} = \frac{z - \xi_1}{z - \xi_2},$$

or

$$\frac{z' - \xi_1}{z' - \xi_2} = K\frac{z - \xi_1}{z - \xi_2}, \tag{24}$$

where

$$K = \frac{a - c\xi_1}{a - c\xi_2}. \tag{25}$$

K is called the "multiplier" of the transformation; its value, as we shall see, determines the character of the transformation.

To get an expression for K in terms of the coefficients of the transformation, we form the following symmetric function of ξ_1 and ξ_2:

$$K + \frac{1}{K} = \frac{a - c\xi_1}{a - c\xi_2} + \frac{a - c\xi_2}{a - c\xi_1}$$
$$= \frac{2a^2 - 2ac(\xi_1 + \xi_2) + c^2(\xi_1{}^2 + \xi_2{}^2)}{a^2 - ac(\xi_1 + \xi_2) + c^2\xi_1\xi_2}.$$

Since ξ_1 and ξ_2 are the roots of

$$cz^2 + (d - a)z - b = 0, \tag{9}$$

we have

$$\xi_1 + \xi_2 = \frac{a - d}{c}, \quad \xi_1\xi_2 = -\frac{b}{c}, \quad \xi_1{}^2 + \xi_2{}^2 = \frac{(a - d)^2 + 2bc}{c^2}.$$

Making these substitutions and simplifying, we have

$$K + \frac{1}{K} = \frac{(a + d)^2 - 2ad + 2bc}{ad - bc}.$$

or, if $ad - bc = 1$,

$$K + \frac{1}{K} = (a + d)^2 - 2. \tag{26}$$

We observe from (26) that the value of K depends solely upon the value of $a + d$. If we replace K by $1/K$, equation (26) is unaltered; hence, the two roots of (26) are reciprocals. The particular root to be used in (24) depends upon which fixed point is called ξ_1 and which ξ_2.

Another simple equation satisfied by K is got from (26) by transposing the 2 and extracting the square root:

$$\sqrt{K} + \frac{1}{\sqrt{K}} = a + d. \tag{27}$$

We now make the change of variables

$$Z = G(z) = \frac{z - \xi_1}{z - \xi_2}, \qquad Z' = G(z') = \frac{z' - \xi_1}{z' - \xi_2}, \tag{28}$$

transformations which carry ξ_1 and ξ_2 to 0 and ∞, respectively.

Then, (24) takes the form

$$Z' = KZ. \tag{29}$$

Call this transformation K so that $K(Z) \equiv KZ$, and we have for the original transformation

$$z' = G^{-1}(Z') = G^{-1}K(Z) = G^{-1}KG(z).$$

Writing for the original transformation $z' = T(z)$, we have

$$T = G^{-1}KG, \text{ whence } K = GTG^{-1}. \tag{30}$$

Let F be any configuration (point, circle, region, or what not) and let F be carried into F' by K. Operating with T on $G^{-1}(F)$, we have

$$TG^{-1}(F) = G^{-1}KGG^{-1}(F) = G^{-1}K(F) = G^{-1}(F');$$

that is, T carries $G^{-1}(F)$ into $G^{-1}(F')$. We shall use this fact in the following manner: We shall investigate the simple transformation (29) and, having found how configurations are transformed, we shall carry the results back to the case of z and z' by applying G^{-1}.

The transformation with $c = 0$,

$$z' = \frac{az + b}{d}, \quad ad = 1$$

can be put in the form (29) and treated similarly. One fixed point is ∞, the other is $\xi_1 = \frac{b}{d - a}$. We find easily that

$$z' - \xi_1 = K(z - \xi_1) \tag{31}$$

where

$$K = \frac{a}{d} \tag{32}$$

Putting

$$Z = G(z) = z - \xi_1, \quad Z' = G(z') = z' - \xi_1,$$

transformations which carry ξ_1 and ∞ to 0 and ∞, respectively, we have (29) as before. We have also

$$K + \frac{1}{K} = \frac{a}{d} + \frac{d}{a} = \frac{a^2 + d^2}{ad} = \frac{(a + d)^2 - 2ad}{ad} = (a + d)^2 - 2;$$

hence, for this case also K satisfies equations (26) and (27).

An advantage in writing a transformation in terms of K lies in the ease with which powers of the transformation can be written down. If the transformation (1) be repeated n times, the equivalent single transformation becomes rapidly complicated if expressed in terms of a, b, c, d. But, if we use (24) or (31), we have, obviously, as the result of n applications of the transformation

$$\frac{z' - \xi_1}{z' - \xi_2} = K^n \frac{z - \xi_1}{z - \xi_2}, \text{ or } z' - \xi_1 = K^n(z - \xi_1). \tag{33}$$

Thus, we merely replace the multiplier K by K^n. Similarly, for the inverse we use the multiplier K^{-1}; and for n applications of the inverse we use K^{-n}.

Writing K in terms of its modulus A (>0) and its amplitude θ, so

$$K = Ae^{i\theta},$$

we distinguish the three classes of transformations treated in the following sections.

7. The Hyperbolic Transformation, K = A.—We assume that $A \neq 1$, since, otherwise, we have the identical transformation. The transformation $Z' = AZ$ is the stretching from the

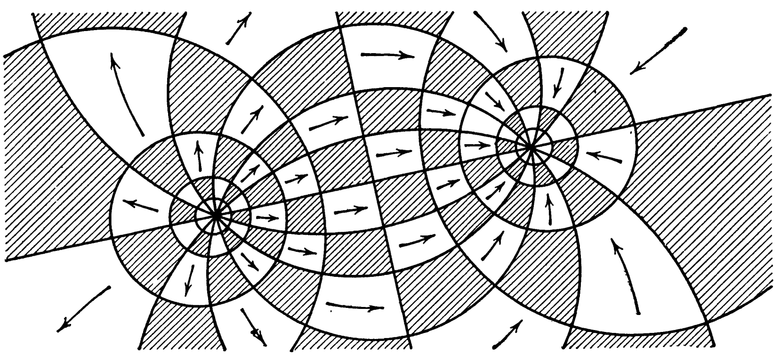


FIG. 6.

origin studied in Sec. 5(*c*). We observe at once the following facts concerning it: (1) a straight line through the origin (that is, a circle through the fixed points 0 and ∞) is transformed into itself, each half line issuing from the origin being transformed into itself; (2) the half plane on one side of a line through 0 is transformed into itself; (3) any circle with center at the origin (and hence orthogonal to the family of fixed lines) is transformed into some other circle with center at the origin; (4) the points 0 and ∞ are inverse points with respect to any circle with center at the origin.

We now make the transformation G^{-1} which carries 0 and ∞ to ξ_1 and ξ_2. We have, then, the following facts concerning the hyperbolic transformation: (1) any circle through the fixed points is transformed into itself, each of the two arcs into which the circle is separated by the fixed points being transformed into itself; (2) the interior of a circle through the fixed points is transformed into itself; (3) any circle orthogonal to the circles through the fixed points is carried into some other such circle; (4) the fixed points are inverse with respect to each circle of (3).

Figure 6 shows the two families of circles just mentioned. The way in which regions are transformed is indicated in the figure, each shaded region being transformed into the next in the direction of the arrow.

From (26), we get the condition that the transformation be hyperbolic in terms of $a + d$. The quantity $K + \frac{1}{K}$ has for real positive values of K the minimum value 2 when $K = 1$. Since $K \neq 1$, we have $K + \frac{1}{K} > 2$; whence, from (26), $(a + d)^2 > 4$. Hence, *in order that the transformation be hyperbolic it is necessary that* $a + d$ *be real and* $|a + d| > 2$. That this condition is sufficient will appear presently.

8. The Elliptic Transformation, $K = e^{i\theta}$.—Here, $\theta \neq 2n\pi$. The transformation $Z' = e^{i\theta}Z$ is the rotation about the origin

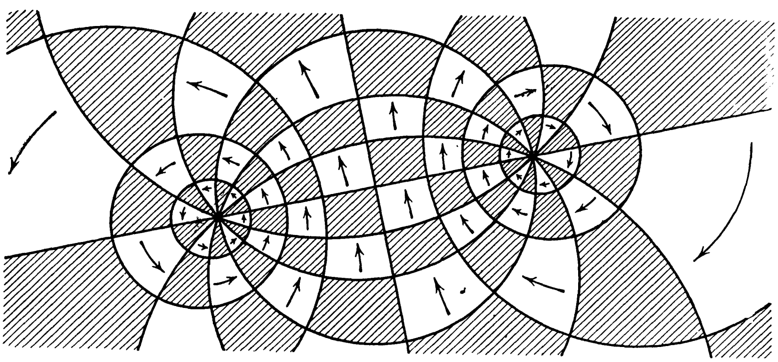


FIG. 7.

discussed in Sec. 5(*b*). The straight lines and circles of the preceding section have their rôles interchanged. The circle with center at the origin is transformed into itself, the interior of the circle being transformed into itself. The points 0 and ∞ are inverse with respect to each fixed circle. A line drawn through the origin is transformed into a line through the origin which makes an angle θ with the first line.

Applying the transformation G^{-1}, which carries 0 and ∞ to ξ_1 and ξ_2, respectively, we have the following facts: (1) an arc of a circle joining the fixed points is transformed into an arc of a circle joining the fixed points and making an angle θ with the first arc; (2) each circle orthogonal to the circles through the fixed points is transformed into itself; (3) the interior of each circle of (2) is transformed into itself; (4) the fixed points are inverse points with respect to each circle of (2).

The character of the transformation, with $\theta = \frac{1}{3}\pi$ is shown in Fig. 7. The shaded regions are transformed into shaded regions as indicated by the arrows.

For the elliptic transformation, (26) gives

$$(a + d)^2 = 2 + e^{i\theta} + e^{-i\theta} = 2 + 2\cos\theta. \tag{34}$$

The second member is positive or zero and less than 4. Hence, *$a + d$ is real and $|a + d| < 2$.* From (27) we have

$$a + d = \pm\left(e^{i\frac{\theta}{2}} + e^{-i\frac{\theta}{2}}\right) = \pm 2\cos\frac{\theta}{2}. \tag{35}$$

If θ is commensurable with π, there will exist an integer n such that $n\theta = 2m\pi$; and $K^n = e^{2m\pi i} = 1$. The result of applying the transformation n times is that each point is returned to its original position. The transformation is then said to be *of period n.* We shall find that the only transformations possessing this periodic property are certain of the elliptic transformations.

We illustrate with two or three useful cases. If $\theta = \pi$ the transformation is of period 2. Then $K = e^{\pi i} = -1$, and $a + d = \pm 2\cos\frac{\pi}{2} = 0$.

If $\theta = 2\pi/3$, the period is 3. Then $K = e^{2\pi i/3} = \frac{1}{2}(-1 + i\sqrt{3})$, and $a + d = \pm 2 \cos\frac{\pi}{3} = \pm 1$.

If $\theta = \pi/2$, the period is 4. Then $K = e^{\pi i/2} = i$, and $a + d = \pm 2\cos\frac{\pi}{4} = \pm\sqrt{2}$.

9. The Loxodromic Transformation, $K = Ae^{i\theta}$.—Here A is positive and unequal to 1 and $\theta \neq 2n\pi$. The transformation $Z' = Ae^{i\theta}Z$ can be written as the succession of the transformations

$$Z' = e^{i\theta}Z_1, \quad Z_1 = AZ,$$

of which one is hyperbolic, the other elliptic. There is a stretching from the origin followed by a rotation about the origin. Each circle with center at the origin is carried into another circle with center at the origin; and each half line through the origin is carried into a half line making an angle θ with the first.

For the original transformation, there is a combination of the motions shown in Figs. 6 and 7. Each circular arc joining the fixed points is carried into another such arc making an angle θ with the first. Each circle orthogonal to the circles through the fixed points is carried into another such orthogonal circle.

The loxodromic transformation has, in general, no fixed circles. There is an exception when $\theta = \pi$. Then, any circular arc joining the fixed points is carried into another arc joining the fixed points and making an angle π with the first, the two arcs thus forming a circle. Any circle through the fixed points is then carried into itself. There is, however, this difference from the preceding cases: *the interior of a fixed circle is transformed into its exterior.*

For the loxodromic transformation, (26) gives

$$\begin{aligned}(a + d)^2 &= 2 + Ae^{i\theta} + \frac{1}{A}e^{-i\theta} \\ &= 2 + \left(A + \frac{1}{A}\right)\cos\theta + i\left(A - \frac{1}{A}\right)\sin\theta. \quad (36)\end{aligned}$$

In general, the second member is not real. If, however, $\theta = \pi$ the second member becomes $2 - \left(A + \frac{1}{A}\right)$, which is real. But we found that $A + \frac{1}{A} > 2$; so in this case the second member is negative. Without exception *in the loxodromic transformation* $a + d$ *is a complex (non-real) number.* In the loxodromic transformation with fixed circles $a + d$ is a pure imaginary.

10. The Parabolic Transformation.—There remains the transformation with one fixed point, which is called a "parabolic transformation." The condition that there is a single fixed point (Sec. 3) is that $a + d = \pm 2$. Then the multiplier, if defined by (26), has the value $K = 1$.

If $c \neq 0$, the fixed point is $\xi = (a - d)/2c$. The transformation carries ∞, ξ, $-d/c$ into a/c, ξ, ∞, respectively; hence, by the use of the formula (14), it can be written

$$\frac{z' - \dfrac{a}{c}}{z' - \xi} = -\frac{\xi + \dfrac{d}{c}}{z - \xi}.$$

Subtracting 1 from each member,

$$\frac{\xi - \dfrac{a}{c}}{z' - \xi} = -\frac{\xi + \dfrac{d}{c}}{z - \xi} - 1.$$

Now,

$$\xi - \frac{a}{c} = \frac{a - d}{2c} - \frac{a}{c} = -\frac{a + d}{2c} = \mp\frac{1}{c},$$

$$\xi + \frac{d}{c} = \frac{a - d}{2c} + \frac{d}{c} = \frac{a + d}{2c} = \pm\frac{1}{c}.$$

Hence, the transformation can be written in the form

$$\frac{1}{z' - \xi} = \frac{1}{z - \xi} \pm c. \tag{37}$$

In (37), we have $+c$ if $a + d = 2$, and $-c$ if $a + d = -2$

Making the change of variable,

$$Z = G(z) = \frac{1}{z - \xi}, \quad Z' = G(z') = \frac{1}{z' - \xi}, \tag{38}$$

a transformation which carries ξ to ∞, we have

$$Z' = Z \pm c. \tag{39}$$

If $c = 0$, so that ∞ is the single fixed point, we have already found, in Sec. 3, that the transformation is of the form (39) without further change. We have, in fact, $a = d = \pm 1$, and

$$z' = z \pm b. \tag{39'}$$

The transformation (39) is the translation discussed in Sec. 5(a). The plane is translated parallel to the line joining the origin to the point $\pm c$. Any line parallel to this line is transformed into itself. The half plane on one side of a fixed line is carried into itself. Any other straight line is carried into a parallel line.

On applying G^{-1}, ∞ is carried to ξ. Parallel straight lines, intersecting at ∞ only, are carried into circles intersecting only at ξ, and, hence, tangent at ξ. Hence, in the parabolic transformation: (1) any circle through the fixed point is transformed into a tangent circle through the fixed point; (2) there is a one parameter family of tangent circles each of which is transformed into itself; (3) the interior of each fixed circle is transformed into itself.

The manner in which the plane is transformed is shown in Fig. 8. Each shaded region is carried in the direction of the arrow.

It is clear from the Figs. 6 to 8 and from the reasoning on which they are based that if a linear transformation is hyperbolic, elliptic, or parabolic there passes through each point of the plane, other than a fixed point, a unique fixed circle. In particular, there is in each case a single fixed circle through ∞; that is, there is one fixed straight line. This line is easily constructed; for it passes through the point $-d/c$ which is carried to ∞ and the point a/c into which ∞ is carried.

At this point we shall combine certain of the results of the latter sections into a theorem for reference. We exclude the identical transformation $z' = z$.

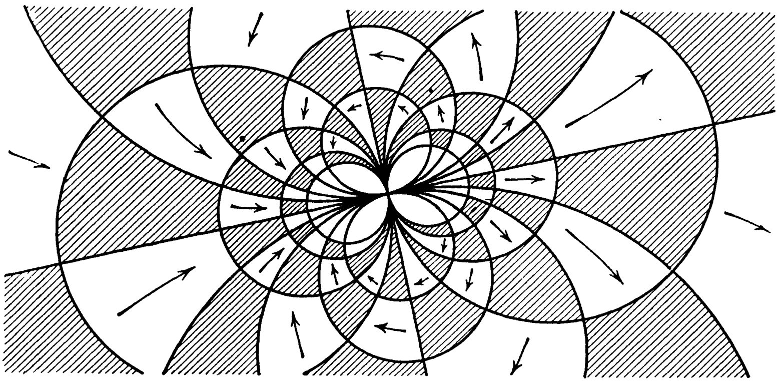


FIG. 8.

THEOREM 15.—*The transformation* $z' = (az + b)/(cz + d)$, *where* $ad - bc = 1$, *is of the type stated if, and only if, the following conditions on* $a + d$ *hold:*

Hyperbolic, if $a + d$ *is real and* $|a + d| > 2$.
Elliptic, if $a + d$ *is real and* $|a + d| < 2$.
Parabolic, if $a + d = \pm 2$.
Loxodromic, if $a + d$ *is complex.*

We have proved that these conditions are necessary. That they are sufficient follows, by elementary reasoning, from the fact that they are mutually exclusive. Thus, if $a + d$ is real and $|a + d| > 2$, the transformation can be neither elliptic, parabolic, nor loxodromic, so it must be hyperbolic; and so on.

11. The Isometric Circle.—In an analytic transformation $z' = f(z)$, a lineal element $dz = z_2 - z_1$ connecting two points in the infinitesimal neighborhood of a point z is transformed into the lineal element dz' in the neighborhood of z'. We have $dz' = f'(z)dz$; hence, the length of the element is multiplied by $|f'(z)|$, and the element is rotated through an angle $\arg f'(z)$. For the linear transformation

$$z' = T(z) = \frac{az + b}{cz + d}, \quad ad - bc = 1, \tag{1}$$

we have the following theorem:

THEOREM 16.—*When the transformation* (1) *is applied, infinitesimal lengths in the neighborhood of a point z are multiplied by $|cz + d|^{-2}$; infinitesimal areas in the neighborhood of z are multiplied by $|cz + d|^{-4}$.*

For we have

$$\frac{dz'}{dz} = T'(z) = \frac{1}{(cz + d)^2}, \tag{40}$$

whence, lengths are multiplied by $|T'(z)|$, or $|cz + d|^{-2}$. An infinitesimal region is carried into a similar region with corresponding lengths multiplied by $|T'(z)|$; hence, the area is multiplied by $|T'(z)|^2$, or $|cz + d|^{-4}$.

We get alternative forms for $T'(z)$ from (24) and (37). Differentiating (24) and simplifying, we obtain the first of the following results: inverting and differentiating, we find the second. For the transformation with two fixed points,

$$\frac{dz'}{dz} = T'(z) = K\left(\frac{z' - \xi_2}{z - \xi_2}\right)^2 = \frac{1}{K}\left(\frac{z' - \xi_1}{z - \xi_1}\right)^2. \tag{41}$$

For the parabolic transformation (37), we have

$$T'(z) = \left(\frac{z' - \xi}{z - \xi}\right)^2, \tag{42}$$

the same as (41) with the value $K = 1$.

From (41), we have

$$T'(\xi_1) = K, \quad T'(\xi_2) = \frac{1}{K}. \tag{43}$$

At ξ_1 which is fixed, $dz' = Kdz$. For the hyperbolic transformation, $K = A$, the infinitesimal neighborhood of ξ_1 undergoes a stretching from ξ_1; for the elliptic transformation, $K = e^{i\theta}$, there is a rotation about ξ_1 through the angle θ; for the loxodromic transformation, $K = Ae^{i\theta}$, there is a combination of stretching and rotation. Analogous remarks apply to the neighborhood of

ξ_2. Since $\frac{1}{K} = \frac{1}{A}e^{-i\theta}$, the stretching and rotation are in the opposite sense.

In the parabolic transformation, we find, on substituting $\xi = (a - d)/2c$ into (40), that $T'(\xi) = 1$. The infinitesimal neighborhood of ξ is unaltered.

Lengths and areas are unaltered in magnitude if, and only if, $|cz + d| = 1$. If $c \neq 0$, the locus of z is a circle. Writing $\left|z + \frac{d}{c}\right| = \frac{1}{|c|}$, we see that the center is $\frac{-d}{c}$, the radius is $\frac{1}{|c|}$.

DEFINITION.—*The circle* I,

$$I: \quad |cz + d| = 1, \quad c \neq 0, \tag{44}$$

which is the complete locus of points in the neighborhood of which lengths and areas are unaltered in magnitude by the transformation (1), *is called the isometric circle of the transformation.*

The isometric circle will play a fundamental part in many of our later developments. In this section we shall investigate some of its properties.

We note, first, that if $c = 0$, so that ∞ is a fixed point, there is no unique circle with the property of the isometric circle. The derivative $T'(z)$ is constant and equal to K (Equations (31) and (39')). Either $|K| \neq 1$, and all lengths are altered in magnitude; or $|K| = 1$, and all lengths are unaltered. The latter case comprises the rigid motions—the rotations (31) and translations (39').

THEOREM 17.—*Lengths and areas within the isometric circle are increased in magnitude, and lengths and areas without the isometric circle are decreased in magnitude, by the transformation.*

For, if z is within I, $\left|z + \frac{d}{c}\right| < \frac{1}{|c|}$, or $|cz + d| < 1$, and $|T'(z)| > 1$. A length or area within I is thus magnified in all its parts. Similarly, if z is without I, $|T'(z)| < 1$; and a length or an area without I is diminished in all its parts.

THEOREM 18.—*A transformation carries its isometric circle into the isometric circle of the inverse transformation.*

The inverse transformation, $z' = (-dz + b)/(cz - a)$, has the isometric circle

$$I': \quad |cz - a| = 1. \tag{45}$$

Its center is a/c, its radius $1/|c|$. Now T carries I into a circle I_0 without alteration of lengths in the neighborhood of any point, hence T^{-1} carries I_0 back to I without alteration. But I' is the complete locus of points in the neighborhood of which T effects no change of length; hence, I_0 coincides with I'.

(*a*) *Geometric Interpretation of the Transformation.*—The transformation T carries I into I' (Fig. 9) without alteration of any arc. Let a point P on I be carried into P'. Then, if I be set down upon I' so that P coincides with P', with proper orientation, corresponding points will coincide. Any sequence of an even number of inversions which will effect the proper transformation on I will be equivalent to T (Theorem 7).

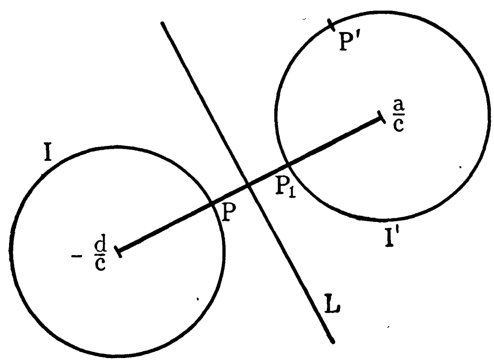


FIG. 9.

As a point moves from P counterclockwise around I, suppose that the corresponding point moves from P' counterclockwise around I'. Then, I can be carried into I' by a rigid motion so that corresponding points coincide. But ∞ is fixed for a rigid motion, so $c = 0$; hence, this case is impossible. Consequently, as a point moves counterclockwise around I the corresponding point moves clockwise around I'. The circle I must be turned over before being applied to I'.

An inversion in I, leaving the points of I invariant, followed by a reflection in L, the perpendicular bisector of the line segment joining the centers, carries I into I' with the desired change of order. P is carried into a point P_1. A rotation with a/c fixed will carry P_1 into P'. The two inversions together with the rotation are equivalent to T.

Since a rotation is equivalent to two reflections (Sec. 5(*b*)), four inversions at most are adequate for the representation of the transformation. If P' coincides with P_1 two inversions are sufficient.

Several alternative geometric transformations are possible. Thus, instead of inverting in I and then reflecting in L we may reflect in L and then invert in I'. Or we may rotate about $-d/c$ at the start; and so on.

The preceding construction fails if I and I' coincide, for then L is not defined. In this case $a = -d$, or $a + d = 0$; and T is an elliptic transformation of period two (Sec. 8). P' lies on I. An inversion in I followed by a reflection in L, the line joining $-d/c$ to the midpoint of the arc PP' is equivalent to T.

(*b*) *The Types of Transformations.*—The distance between the centers of I and I' is $\left|\frac{a}{c} + \frac{d}{c}\right|$; the sum of the radii is $2/|c|$. The circles will intersect, touch, or be totally exterior according as $|a + d|$ is less than, equal to, or greater than 2. Hence, applying Theorem 15, if T is hyperbolic, the isometric circles of T and T^{-1} are external; if T is elliptic, they intersect; if T is parabolic, they are tangent. If T is loxodromic, $|a + d|$ may have any value other than zero, and the isometric circles may have any relation to one another other than coincidence.

A distinction between the loxodromic and the three non-loxodromic transformations appears when we study the geometrical transformations which are equivalent to the transformation. Let P, P_1, P' (arranged as in Fig. 9) have the coordinates z, z_1, z'. Since z, $-d/c$, and a/c lie on a line, $z + \frac{d}{c}$ and $\frac{a}{c} + \frac{d}{c}$ have the same argument, the moduli being $1/|c|$ and $|a + d|/|c|$, respectively, hence,

$$z + \frac{d}{c} = \frac{a + d}{|a + d|c}.$$

Similarly,

$$z_1 - \frac{a}{c} = -\frac{a + d}{|a + d|c}.$$

Making the transformation T, using (22′),

$$z' - \frac{a}{c} = -\frac{1}{c^2(z + d/c)} = -\frac{|a + d|}{(a + d)c} = -\frac{\bar{a} + \bar{d}}{|a + d|c}.$$

From these equations we see that z' coincides with z_1 if, and only if, $\bar{a} + \bar{d} = a + d$; that is, if $a + d$ is real. The transformation is then non-loxodromic.

If the transformation is loxodromic, writing $a + d = re^{i\varphi}$, $\varphi \neq n\pi$, we have

$$z' - \frac{a}{c} = \frac{\bar{a} + \bar{d}}{a + d}\left(z_1 - \frac{a}{c}\right) = e^{-2i\varphi}\left(z_1 - \frac{a}{c}\right).$$

To carry z_1 to z', there is a rotation about a/c through the angle -2φ. We have the result:

THEOREM 19.—*If the transformation is hyperbolic, elliptic, or parabolic, it is equivalent to an inversion in I followed by a reflection in L; if it is loxodromic there is in addition a rotation about the center of I′ through the angle* $-2 \arg (a + d)$.

Consider now the fixed points. Since $T'(\xi_1) = K$, $T'(\xi_2) = 1/K$ (Equation (43)), we have, if $|K| \neq 1$, increase of lengths at one fixed point and decrease at the other; if $|K| = 1$, there is no alteration. Hence, for the hyperbolic and loxodromic transformations one fixed point is within I, the other without; for the elliptic transformation both fixed points, and for the parabolic transformation the single fixed point, are on I. Identical statements are true of I' for similar reasons.

In the elliptic transformation I and I' intersect and L is the common chord. The points of intersection are fixed for both the inversion in I and the reflection in L; hence, they are the fixed points. We found that the lineal elements issuing from the fixed point are rotated through an angle θ, where $K = e^{i\theta}$. Since an arc of I issuing from the fixed point is transformed into an arc of I' issuing from the point, it follows that I and I' intersect at the angle θ.

If $a + d = 0$, so that I and I' coincide, the line L is the line joining the fixed points, which are then at the ends of a diameter.

In the parabolic transformation, L is the common tangent to I and I' at their point of tangency. The point of tangency is then the fixed point.

(*c*) *The Fixed Circles.*—We consider now the non-loxodromic transformations. Each such transformation has a one-parameter family of fixed circles, including, as we found in Sec. 10, the line joining the centers of I and I'. The family of fixed circles is easily constructed. It consists of the circles with centers on L orthogonal to I. For, being orthogonal to I, such a circle is transformed into itself by an inversion in I; and a reflection in L, a diameter, transforms it again into itself. Each fixed circle is also orthogonal to I' from symmetry.

THEOREM 20.—*In a non-loxodromic transformation the isometric circle is orthogonal to the fixed circles.*

For use later we shall prove the following theorem:

THEOREM 21.—*Let Q be a fixed circle of a non-loxodromic transformation and I its isometric circle. Let h be the distance of a point z from q, the center of Q, and h' be the distance of the transformed point z' from q; then*

$h' = h$, if z is on I or Q;

$h' < h$, if z is within both I and Q, or without both;

$h' > h$, if z is within either I or Q, and without the other.

An inversion in I carries z to a point z_1 which is carried to z' by a reflection in L (Fig. 10). Obviously, the reflection does not alter distances from q. The proposition hinges, then, on what happens when z is inverted in I.

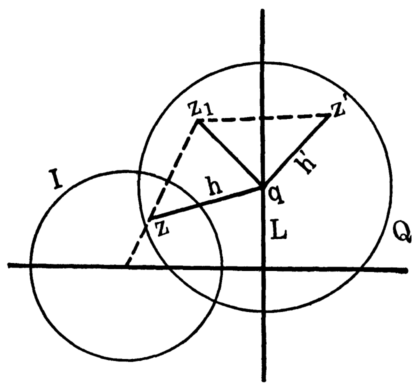


FIG. 10.

The distances of a point and its inverse from the center of a circle orthogonal to the circle of inversion is clearly independent of the orientation of the circles, and their relative magnitudes are independent of the scale used; hence, it will suffice to take for I the unit circle $z\bar{z} = 1$ and to take q on the real axis. The equation of Q is $(z - q)(\bar{z} - q) = r^2$, where, for orthogonality, $r^2 + 1 = q^2$; whence,

$$z\bar{z} - q(z + \bar{z}) + 1 = 0.$$

The expression in the first member of this equation is positive for points without Q and negative for points within. Now,

$$h^2 = (z - q)(\bar{z} - q) = z\bar{z} - q(z + \bar{z}) + 1 + r^2,$$

and, since $z_1 = 1/\bar{z}$,

$$h'^2 = z_1\bar{z}_1 - q(z_1 + \bar{z}_1) + 1 + r^2 = \frac{1 - q(\bar{z} + z) + z\bar{z}}{z\bar{z}} + r^2;$$

whence,

$$h^2 - h'^2 = \frac{[z\bar{z} - 1][z\bar{z} - q(z + \bar{z}) + 1]}{z\bar{z}}.$$

The theorem follows immediately from this equation. If z is within both circles or without both circles, the factors in the numerator of the second member are both negative or both positive, and $h' < h$; if z is within one circle and without the other, the factors differ in sign, and $h' > h$; if z is on one circle, one factor is zero, and $h' = h$.

The following theorem relative to the fixed straight line is easily seen to hold:

THEOREM 22.—*In a non-loxodromic transformation let k and k' be the distances of z and z', respectively, from the fixed straight line M; then $k' = k$ if z is on I or on M; otherwise, $k' > k$ if z is within I and $k' < k$ if z is without I.*

12. The Unit Circle.—We shall, subsequently, have much to do with sets of linear transformations which have one fixed circle in common. It will usually be convenient to take as the common fixed circle some simple circle such as the real axis or the unit circle with center at the origin. It is this latter circle, which we shall henceforth designate by Q_0, that we shall study in this section.

We proceed to find the conditions on the constants in (1) in order that Q_0 be a fixed circle. The equation of Q_0 is

$$z\bar{z} - 1 = 0. \tag{46}$$

The transform of Q_0 by (1) is, from Equation (19),

$$(d\bar{d} - c\bar{c})z'\bar{z}' + (-\bar{b}d + \bar{a}c)z' + (-b\bar{d} + a\bar{c})\bar{z}' + b\bar{b} - a\bar{a} = 0.$$

This circle is identical with Q_0 if, and only if,

$$-\bar{b}d + \bar{a}c = 0, \qquad -b\bar{d} + a\bar{c} = 0. \tag{a}$$

$$d\bar{d} - c\bar{c} = a\bar{a} - b\bar{b} \neq 0. \tag{b}$$

Each equation in (*a*) is a consequence of the other; from the second,

$$\frac{b}{\bar{c}} = \frac{a}{\bar{d}} = \lambda,$$

say; then

$$b = \lambda\bar{c}, \qquad a = \lambda\bar{d}.$$

Substituting in (*b*)

$$d\bar{d} - c\bar{c} = \lambda\bar{\lambda}(d\bar{d} - c\bar{c}) \neq 0,$$

whence, $\lambda\bar{\lambda} = 1$.

From $ad - bc = 1$, we have

$$\lambda(dd - c\bar{c}) = 1;$$

hence, λ is real, so $\lambda = \pm 1$. The sign of λ depends upon the sign of $d\bar{d} - c\bar{c}$. If the interior of Q_0 is transformed into its interior, the point $-d/c$ which is carried to ∞ is outside the circle and $|-d/c| > 1$; so $d\bar{d} - c\bar{c} > 0$ and $\lambda = 1$. We have, then,

$$b = \bar{c}, \qquad a = \bar{d}, \qquad d = \bar{a}.$$

These values obviously satisfy the conditions (a) and (b). We have the following result:

THEOREM 23.—*The most general linear transformation carrying Q_0 into itself and carrying the interior of Q_0 into itself is the transformation*

$$z' = \frac{az + \bar{c}}{cz + \bar{a}}, \qquad a\bar{a} - c\bar{c} = 1. \tag{47}$$

The most general linear transformation carrying Q_0 into itself and carrying the interior of Q_0 into its exterior is found similarly. Then, $-d/c$ is within Q_0 and $d\bar{d} - c\bar{c} < 0$; so that $\lambda = -1$. The most general transformation is the resulting loxodromic transformation

$$z' = \frac{az - \bar{c}}{cz - \bar{a}}, \qquad c\bar{c} - a\bar{a} = 1. \tag{48}$$

The transformation (47) maps the interior of Q_0 in a one-to-one and directly conformal manner on itself. It is a remarkable fact, which we shall now prove, that it is the most general such transformation. We first prove the following proposition:

The most general transformation which maps the interior of Q_0 in a one-to-one and directly conformal manner on itself and which leaves the origin fixed is a rotation about the origin.

Let $z' = f(z)$ be such a transformation. Owing to the conformality $f(z)$ is analytic in Q_0. Further, $|f(z)| < 1$ when $|z| < 1$, since an interior point is carried into an interior point. Since $z' = 0$ when $z = 0$, $f(z)$ has a zero at the origin; hence $f(z)/z$ is analytic in Q_0.

Consider now $|f(z)/z|$ in a circle Q' with center at the origin and radius $r < 1$. Since the absolute value of a function which is analytic in and on the boundary of a region takes on its maximum value on the boundary, we have, since $|z| = r$ on Q',

$$\left|\frac{f(z)}{z}\right| < \frac{1}{r} \text{ in } Q'.$$

Since r may be taken as near to 1 as we like, we have

$$\left|\frac{z'}{z}\right| = \left|\frac{f(z)}{z}\right| \leqslant 1 \text{ in } Q_0.$$

Considering the inverse function, $z = \varphi(z')$, we have by the same reasoning $|z/z'| \leqslant 1$ in Q_0. Consequently, $|z'/z| = 1$ in Q_0, whence $z'/z = e^{i\alpha}$. But, if the absolute value of an analytic function is constant, so also is its argument; hence α is constant. We have, thus,

$$z' = e^{i\alpha}z,$$

a rotation about the origin.

We shall now remove the restriction that the origin be fixed. Let $z' = f(z)$ map the interior of Q_0 on itself in a one-to-one and directly conformal manner, and let $f(0) = z_0$. Let $z' = S(z)$ be a linear transformation of the form (47) such that $S(0) = z_0$. (It is easy to determine the constants of (47) so that $\bar{c}/\bar{a} = z_0 < 1$.) If we make the transformation f and then make S^{-1}, the interior of Q_0 is carried into itself and the origin is fixed. Hence, $S^{-1}f = U$, a rotation; and $f = SU$. We thus have a linear transformation. Since it carries the interior of Q_0 into itself, it is of the form (47).

THEOREM 24.—*The most general transformation which maps the interior of Q_0 in a one-to-one and directly conformal manner on itself is the linear transformation* (47).

The proof of the following more general theorem is now easily made.

THEOREM 25.—*The most general transformation which maps the interior or exterior of one circle in a one-to-one and directly conformal manner upon the interior or exterior of another circle is a linear transformation.*

Let $z' = f(z)$ carry the interior or exterior of Q_1 into the interior or exterior of Q_2 in the manner stated. Let S_1 and S_2 be linear transformations carrying Q_1 and Q_2, respectively, into Q_0, the interior or exterior of each which is involved in the mapping being carried into the interior of Q_0. Then, the sequence of transformations, S_1^{-1}, followed by f, followed by S_2, carries the interior of Q_0 into itself, and is equivalent to a linear transformation T of the form (47).

$$S_2fS_1^{-1} = T, \text{ or } f = S_2^{-1}TS_1.$$

The transformation is thus a linear transformation.

CHAPTER II

GROUPS OF LINEAR TRANSFORMATIONS

13. Definition of a Group. Examples.—The automorphic function depends for its definition on a set of linear transformations called a "group." In the present chapter we shall make a study of groups of linear transformations, after which we shall be in a position to pass to the definition of the automorphic function and to a study of its properties.

DEFINITION.—*A set of transformations, finite or infinite in number, is said to form a group if,*

(*a*) *the inverse of each transformation of the set is a transformation of the set;*

(*b*) *the succession of any two transformations of the set is a transformation of the set.*

The definition applies to all kinds of transformations, but we shall be concerned only with sets of linear transformations. The two group properties, expressed in symbolic notation, are: (*a*) if T is any transformation of the set so also is T^{-1}; (*b*) if S is a transformation of the set, not necessarily different from T, so also is ST. It follows by a repeated application of (*b*) that the transformation equivalent to performing any sequence of transformations of a group belongs to the group. In particular, all positive and negative integral powers of a transformation T of the group belong to the group. Also $T^{-1}T(=1)$ belongs to the group; that is, *every group contains the identical transformation, $z' = z$.*

Given a set of linear transformations, we may test whether or not it constitutes a group by applying (*a*) and (*b*) to the transformations of which it is composed. There are, however, certain cases in which the group properties obviously hold. For example, if the set consists of *all* linear transformations which leave some configuration F in the z-plane invariant, then the set is a group; for, clearly, the inverse of any transformation or the successive performance of any two transformations will leave F invariant and, being themselves linear transformations, will then belong to the set. Thus, all linear transformations with a common fixed point constitute a group. All linear transformations of the form (47), Sec. 12, leaving Q_0 and its interior invariant, form a group. The set of all linear transformations which carry a given regular polygon into itself, consisting of certain rotations about its center, form a group.

Similarly, the set of all linear transformations which leave invariant some function of z constitute a group. For example, all linear transformations $z' = T(z)$ such that $\sin z' \equiv \sin z$ form a group. Such transformations as $z' = z + 2\pi$, $z' = z + 4\pi$, $z' = \pi - z$, etc., belong to this group. It is by virtue of this property, as we shall see later, that $\sin z$ is called an "automorphic function."

Given a set of linear transformations $T_1, T_2, \ldots, T_n$, we may form a group containing them in the following way: Let the set contain the given transformations, their inverses, and the transformations formed by combining the given transformations and their inverses into products in all possible ways. Then it is easily seen that the inverse of any transformation or the product of any two is itself some combination of the given transformations and their inverses and, consequently, is included in the set. Hence, the whole set forms a group. The group is said to be "generated" by the transformations $T_1, T_2, \ldots, T_n$, and the transformations are called "generating transformations" of the group.

Examples.—The following are a few examples of well-known groups, some of which will be discussed later.

1. *A Group of Rotations about the Origin.*—The m transformations, $z' = z$, $e^{2\pi i/m}z$, $e^{4\pi i/m}z$, $\ldots$, $e^{2(m-1)\pi i/m}z$, form a group. They are the rotations about the origin through multiples of the angle $2\pi/m$. The group is generated by the transformation $z' = e^{2\pi i/m}z$.

2. *The Group of Anharmonic Ratios.*—The six transformations

$$z' = z, \quad \frac{1}{z}, \quad 1 - z, \quad \frac{1}{1 - z}, \quad \frac{z - 1}{z}, \quad \frac{z}{z - 1}$$

form a group. It can be verified by forming the inverses and by combining the transformations that both group properties are satisfied. The group is so named for the reason that if z is any one of the anharmonic ratios of four points on a line, the six anharmonic ratios are given by the transformations of the group.

3. *The Group of the Simply Periodic Functions.*—The set $z' = z + m\omega$, where ω is a constant different from zero, and m is any positive or negative integer or zero, forms a group. The group is generated by the transformation $z' = z + \omega$.

4. *The Group of the Doubly Periodic Functions.*—The set $z' = z + m\omega + m'\omega'$, where ω and ω' are constants different from zero and the ratio ω'/ω is not real, and where m and m' are any positive or negative integers or zero, forms a group. It is generated by the transformations $z' = z + \omega$, $z' = z + \omega'$.

The restrictions on ω and ω' are not necessary for establishing the group properties.

5. *The Modular Group.*—The infinite set of transformations $z' = (az + b)/(cz + d)$, where a, b, c, d are real integers such that $ad - bc = 1$, constitutes a group. For, a reference to Equations (3) and (6) of Sec. 1 shows that the inverse of such a transformation and, also, the product of two such transformations are transformations with integral coefficients and of unit determinants. Since the coefficients are real, each transformation carries the real axis into itself.

6. *The Group of Picard.*—The set of transformations $z' = (az + b)/(cz + d)$, where a, b, c, d are either real or complex integers (*i.e.* of the form $m + ni$, where m and n are real integers) such that $ad - bc = 1$, constitutes a group. The proof is as in the preceding case.

7. *A Group Allied to Q_0.*—In a similar manner, the transformations $z' = (az + \bar{c})/(cz + \bar{a})$, where a and c are real or complex integers such that $a\bar{a} - c\bar{c} = 1$, form a group. The transformations of this group (Theorem 23, Sec. 12) have Q_0 as a fixed circle and carry the interior of Q_0 into itself.

14. Properly Discontinuous Groups.—If we compare the group of the simply periodic functions, $z' = z + m\omega$, with the group of all translations, $z' = z + b$, where b is any constant, we observe the following difference: In the former case there is no transform of a point z within the distance $|\omega|$ of z; in the latter group we get transforms of z as near to z as we like by taking b small enough. These two groups bring out an essential distinction.

DEFINITION.—*A group is called properly discontinuous in the z-plane if there exists a point z_0 and a region S enclosing z_0 such that all transformations of the group, other than the identical transformation, carry z_0 outside S.*

The automorphic functions are founded on the properly discontinuous groups, and these only will appear in our subsequent study. The groups whose transformations contain continuously varying parameters, which have given rise to so many and so profound researches, play no part in the theory to which this book is devoted and will not be considered further.

A group is said to contain *infinitesimal transformations* if there is, for some region A and any given $\epsilon > 0$, a transformation $z' = (az + b)/(cz + d)$, $ad - bc = 1$, such that for all points z of A we have $|z' - z| < \epsilon$. It is found without difficulty that a necessary and sufficient condition for this is that there be transformations for which c, $d - a$, and b are all arbitrarily small (but not all zero, for then we have the identical transformation).

Not all groups which are free of infinitesimal transformations are properly discontinuous. The group of Picard, for example, does not contain infinitesimal transformations, since c, $d - a$, and b are complex integers and cannot be made arbitrarily small without being all zero. It can be shown, however, that the points into which any point is carried are everywhere dense in the whole z-plane. Such a group, that is, one which does not contain infinitesimal transformations and yet which is not properly discontinuous in the z-plane, is called "improperly discontinuous" in the z-plane.

15. Transforming a Group.—From a given group of linear transformations infinitely many other groups can be derived by applying linear transformations to the plane in which z and its transforms are represented. Let T be any transformation of the given group, and let T carry z into z'. Let a transformation G be applied, z and z' being transformed into z_1 and z_1', respectively. Then z_1 is carried into z_1' by the transformation S where

$$S = GTG^{-1}; \tag{1}$$

for

$$GTG^{-1}(z_1) = GT(z) = G(z') = z_1'.$$

Let all the transformations of the original group be altered in this manner, so that to each T of the group there corresponds an S given by (1). We shall show that the new set of transformations forms a group. We have $S^{-1} = (GTG^{-1})^{-1} = GT^{-1}G^{-1}$, which belongs to the set since T^{-1} belongs to the original group. If $S_1 = GT_1G^{-1}$ is a second transformation of the set, $SS_1 = GTG^{-1}GT_1G^{-1} = GTT_1G^{-1}$, and SS_1 belongs to the set since TT_1 belongs to the original group. Thus, both group properties are satisfied. Two groups whose transformations can be made to correspond in a one-to-one manner, as the S and T transformations are paired by virtue of (1), so that the product of any number of transformations of one group corresponds to the analogous product of the corresponding transformations of the other, are said to be "isomorphic."

It will often facilitate the study of a group to transform it in the manner indicated. For example, an important point can be carried to ∞, or an important circle can be carried into the real axis or the unit circle Q_0. Having found how figures are transformed by the new group, we can then carry the results back to the old by applying G^{-1}. For if S carries a figure F into F', T carries $G^{-1}(F)$ into $G^{-1}(F')$, as we see at once from the equation $T = G^{-1}SG$.

It should be mentioned that the transformations S and T of (1) are of the same type, whatever G may be. Let

$$T = \frac{az + b}{cz + d}, \qquad G = \frac{\alpha z + \beta}{\gamma z + \delta}, \qquad \begin{matrix} ad - bc = 1, \\ \alpha\delta - \beta\gamma = 1, \end{matrix}$$

and form the product in (1), using the equations (3) and (6) of Sec. 1,

$$S = \frac{(-\alpha\delta a + \alpha\gamma b - \beta\delta c + \beta\gamma d)z + \alpha\beta a - \alpha^2 b + \beta^2 c - \alpha\beta d}{(-\gamma\delta a + \gamma^2 b - \delta^2 c + \gamma\delta d)z + \beta\gamma a - \alpha\gamma b + \beta\delta c - \alpha\delta d}, \quad (2)$$

the determinant of S being 1. Then,

$$(-\alpha\delta a + \alpha\gamma b - \beta\delta c + \beta\gamma d) + (\beta\gamma a - \alpha\gamma b + \beta\delta c - \alpha\delta d) = -(a + d).$$

It follows that K has the same value for S as for T (Equation (26), Sec. 6).

16. The Fundamental Region.—Before proceeding to the study of the general properly discontinuous group it is desirable to introduce the important concept of the fundamental region.

DEFINITION.—*Two configurations (points, curves, regions, etc.) are said to be congruent with respect to a group if there is a transformation of the group other than the identical transformation, which carries one configuration into the other.*

DEFINITION.—*A region, connected or not, no two of whose points are congruent with respect to a given group, and such that the neighborhood of any point on the boundary contains points congruent to points in the given region, is called a fundamental region for the group.*

The accompanying figures show fundamental regions for certain simple groups. The reader is probably already familiar with some of them.

For the group of rotations about a point through multiples of an angle θ, which is a submultiple of 2π, we draw two half lines from the fixed point forming an angle θ. The region R_0 within the angle is a fundamental region. In Fig. 11, R_0 is a fundamental region for the group $z' = e^{2n\pi i/6}z$. The neighborhood of any point on the boundary of R_0 contains points which can be carried into the interior of R_0 by a rotation through the angle $\pm 2\pi/6$.

In Fig. 12, R_0, whose construction is evident from the figure, is a fundamental region, or period strip, for the group of the simply periodic functions, $z' = z + m\omega$.

In Fig. 13, R_0 is a fundamental region, or period parallelogram, for the group of the doubly periodic functions $z' = z + m\omega + m'\omega'$.

In Fig. 14, R_0, which is bounded by circles with centers at the origin and with radii 1 and A, is a fundamental region for the group of stretchings from the origin, $z' = A^n z$.

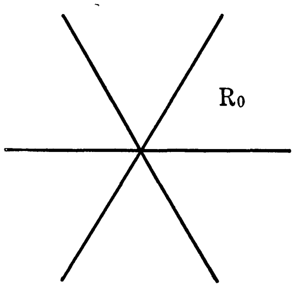


FIG. 11.

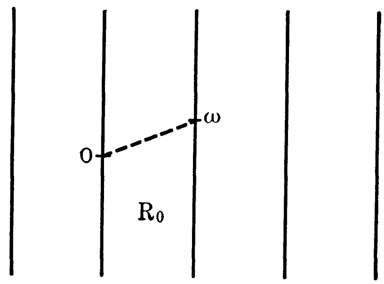


FIG. 12.

Attention may be called to certain properties that are common to the four fundamental regions constructed in the figures and which we shall find to be more or less generally true—to what extent will appear from later analysis—of the fundamental regions we shall use for less simple groups. We note first that the

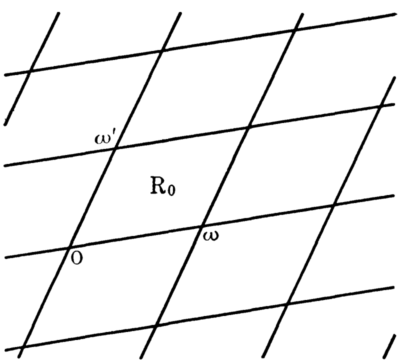


FIG. 13.

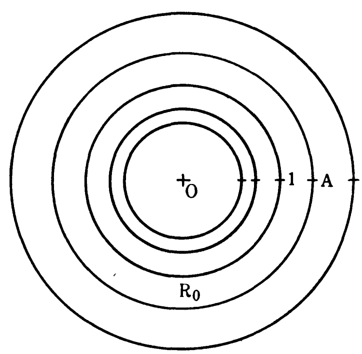


FIG. 14.

boundaries of R_0 in each case consist of congruent curves. In Figs. 11, 12, 14, each of the two boundaries can be carried into the other by a transformation of the group. In Fig. 13, the lower boundary can be carried into the upper by the translation $z' = z + \omega'$, and the left into the right by $z' = z + \omega$.

Further, the transformations connecting congruent boundaries are generating transformations of the group. The two translations just mentioned generate the group of doubly periodic functions. In Fig. 11, all transformations are formed by successive applications of the rotation $z' = e^{2\pi i/6}z$, which carries one boundary into the other. The like fact is true of the other examples.

We note that we can add to the open region R_0 one, but not both, of two congruent boundaries without getting two congruent points in the region. But R_0 must remain in part an open region.

The region R_0 and the regions congruent to it, some of which are shown in the figures, form a set of adjacent, non-overlapping regions covering practically the whole plane. The origin in Fig. 14, however, is not in any region congruent to R_0.

The angle at the vertex of R_0 in Fig. 11 is a submultiple of 2π. The sum of the angles at the four congruent vertices of R_0 in Fig. 13 is equal to 2π. These facts will reappear, suitably generalized.

It is clear that the fundamental region is in no wise unique. Any region congruent to R_0 will serve as a region. Furthermore, we can replace any part of R_0 by a congruent part and still have a fundamental region. Thus, we can subtract a part at one boundary and add a congruent part at another. In this way the character of the bounding curve can be altered freely.

THEOREM 1.—*If no two points of a region are congruent, the transforms of the region by two distinct transformations of the group do not overlap.*

Let A be a region containing no two congruent points. Suppose that two transformations of the group, S and T, carry A into two overlapping regions. Any point z_0 in the common part is the transform by S of a point z_1 of A and the transform by T of a point z_2 of A. If z_1 and z_2 are different for any z_0, then z_1 and z_2 are congruent points of A, which is impossible. If z_1 and z_2 coincide for every z_0 in the common part, S and T are the same transformation (Theorem 7, Sec. 3).

Since a fundamental region contains no two congruent points, we can state the following useful corollary:

COROLLARY.—*The transforms of a fundamental region by two distinct transformations of the group do not overlap.*

17. The Isometric Circles of a Group.—We shall now investigate the properties of the most general properly discontinuous group. For such a group there exists, by hypothesis, at least one point z_0 such that there are no transforms of z_0 in a suitably small region about z_0. Let G be a transformation carrying z_0 into ∞; and let the group be transformed by G as explained in Sec. 15. It is this transformed group which we shall study.

There is no point congruent to ∞ outside or on a circle Q_ρ with a given center and with radius ρ suitably large. In particular, ∞ is not a fixed point for any transformation of the group. Hence, in any transformation $T = (az + b)/(cz + d)$ we have $c \neq 0$, except in the case of the identical transformation. The center of an isometric circle is congruent to ∞ ($-d/c$ is carried to ∞ by T); hence, the centers of all isometric circles lie within Q_ρ.

(*a*) *The Isometric Circle of the Product of Two Transformations.* Certain relations between the isometric circles of two transformations and the isometric circle of their product will be of use now and subsequently. Consider any two transformations

$$T = \frac{az + b}{cz + d}, \qquad S = \frac{\alpha z + \beta}{\gamma z + \delta}, \qquad \begin{array}{l} ad - bc = 1, c \neq 0, \\ \alpha\delta - \beta\gamma = 1, \gamma \neq 0. \end{array}$$

Then,

$$ST = \frac{(\alpha a + \beta c)z + \alpha b + \beta d}{(\gamma a + \delta c)z + \gamma b + \delta d}. \tag{3}$$

In what follows we assume that $S \neq T^{-1}$ so that ST is not the identical transformation; then, the isometric circle of ST is

$$|(\gamma a + \delta c)z + \gamma b + \delta d| = 1. \tag{4}$$

Represent by I_s, I_s', I_t, I_t', I_{st} the isometric circles of S, S^{-1}, T, T^{-1}, ST, respectively; by g_s, g_s', g_t, g_t', g_{st} their respective centers; and by r_s, r_t, r_{st} their radii. We have

$$g_s = -\frac{\delta}{\gamma}, \quad g_s' = \frac{\alpha}{\gamma}, \quad g_t = -\frac{d}{c}, \quad g_t' = \frac{a}{c}, \quad g_{st} = -\frac{(\gamma b + \delta d)}{(\gamma a + \delta c)},$$

$$r_s = \frac{1}{|\gamma|}, \qquad r_t = \frac{1}{|c|}, \qquad r_{st} = \frac{1}{|\gamma a + \delta c|}. \tag{5}$$

From these values, we derive the following relations:

$$r_{st} = \frac{1}{|\gamma a + \delta c|} = \frac{1}{|\gamma c| \cdot \left|\frac{a}{c} + \frac{\delta}{\gamma}\right|} = \frac{r_s r_t}{|g_t' - g_s|} \tag{6}$$

and

$$g_{st} - g_t = -\frac{\gamma b + \delta d}{\gamma a + \delta c} + \frac{d}{c} = \frac{\gamma}{c(\gamma a + \delta c)}, \tag{7}$$

whence,

$$|g_{st} - g_t| = \frac{r_{st} r_t}{r_s} = \frac{r_t^2}{|g_t' - g_s|}. \tag{8}$$

(*b*) *The Arrangement of the Isometric Circles.*—By means of the preceding equations we can derive certain simple facts concerning the isometric circles of the transformations of the group.

The radii of the isometric circles are bounded.

Let T $(\neq 1)$ be any transformation of the group and S $(\neq 1)$ a transformation of the group different from T^{-1}. Then, from (8),

$$r_t^2 = |g_{st} - g_t| \cdot |g_t' - g_s|.$$

But each factor in the second member, being the distance between points of Q_ρ, is less than 2ρ; hence,

$$r_t^2 < 4\rho^2, \qquad r_t < 2\rho. \tag{9}$$

The number of isometric circles with radii exceeding a given positive quantity is finite.

Let I_s and I_t' be any two different isometric circles with radii greater than k, a positive quantity. Then ST is not the identical transformation, and, from (6),

$$|g_t' - g_s| = \frac{r_s r_t}{r_{st}} > \frac{k^2}{2\rho}. \tag{10}$$

The distance between the centers of two isometric circles with radii exceeding k has thus a positive lower bound. Since the centers of all such circles lie in the circle Q_ρ, their number must be finite.

It follows from this fact that the transformations of the group are denumerable.

Another consequence may be stated in the following manner:

Given any infinite sequence of distinct isometric circles I_1, I_2, I_3, . . . , of transformations of the group, the radii being $r_1, r_2, r_3, \ldots$, then $\lim_{n=\infty} r_n = 0$.

18. The Limit Points of a Group.—In this section and the remaining sections of the present chapter we suppose that no transformation of the group has a fixed point at infinity, so that the isometric circles exist for all transformations except the identical transformation; and that there are no points congruent to infinity in the neighborhood of infinity. This assumption involves no essential restriction since, as we have already noted, any properly discontinuous group can be transformed into one with the properties mentioned.

Consider the centers of the isometric circles. If the group contains an infinite number of transformations, the centers are infinite in number and, hence, have one or more cluster points. We lay down the following definitions:

DEFINITIONS.—*A cluster point of the centers of the isometric circles of the transformations of a group is called a limit point of the group.*

A point which is not a limit point is called an ordinary point.

It is clear that all limit points lie within or on the circle Q_ρ of Sec. 17, since the centers of all isometric circles lie within that circle. If the group contains only a finite number of transformations, there are, of course, no limit points.

THEOREM 2.—*In the neighborhood of a limit point P there is an infinite number of distinct points congruent to any point of the plane, with, at most, the exception of P itself and of one other point.*

Since only a finite number of isometric circles have radii exceeding a given positive quantity, there are isometric circles of arbitrarily small radius in the neighborhood of P. Let Q be a small circle about P and let I_1', I_2', . . . be an infinite sequence of isometric circles contained in Q, where the center g_n' approaches P as n becomes infinite. Let these be the transforms of the isometric circles I_1, I_2, . . . , and let S_n be the transformation carrying I_n into I_n'.

The centers g_n of I_n have at least one cluster point. Suppose first that there is such a point P' distinct from P. It will suffice to show that for any point P_1 distinct from P and P' there is a congruent point in Q distinct from P; for, by decreasing the region Q, we then have an infinite number of congruent points. Let I_n be near P' and of small enough radius that I_n encloses neither P_1 nor P. Then, since P_1 is outside I_n, $S_n(P_1)$ is inside I_n' and in Q. If $S_n(P_1)$ is different from P, the proposition is established. If $S_n(P_1)$ coincides with P, then P is not a fixed point for S_n and $S_n^2(P_1)$, or $S_n(P)$, is in Q and different from P.

There remains the case that the only limit point of the centers g_n is P itself. Let A and A' be any two points distinct from one another and from P. A and A' are outside an infinite number of circles I_n. For these circles the congruent points, $A_n = S_n(A)$ and $A_n' = S_n(A')$ are in I_n' and in Q. At least one of the points A_n and A_n' is distinct from P. It follows that at least one of the points A and A' has an infinite number of congruent points in Q which are distinct from P. Hence, there is not more than one point, distinct from P, which does not have an infinite number of congruent points distinct from P and in Q. This establishes the theorem.

THEOREM 3.—*The set of limit points is transformed into itself by any transformation of the group.*

The centers of the isometric circles consist of all points congruent to infinity. The transform of the center of an isometric circle is the center of another isometric circle or is ∞ itself.

Let P be a cluster point of the centers $g_1, g_2, \ldots$ Then a transformation S which carries P into P' carries $g_1, g_2, \ldots$ into $g_1', g_2', \ldots$ with P' as cluster point. The points of the latter set, with the possible exception of one point at ∞, are centers of isometric circles. Hence, P' is a limit point.

Furthermore, no point which is not a limit point is carried by S into a limit point, since otherwise S^{-1} would carry a limit point into a point not a limit point.

THEOREM 4.—*If the set of limit points contains more than two points, it is a perfect set.*

A set is perfect, by definition, if it has the following two properties: (1) each cluster point of the set belongs to the set; that is, the set is closed; and (2) each point of the set is a cluster point of points of the set; that is, the set is dense in itself.

That the set is closed follows at once. For, since each limit point contains an infinite number of centers of isometric circles in its neighborhood, a point at which limit points cluster has also an infinite number of centers of isometric circles in its neighborhood; hence, a cluster point of limit points is itself a limit point.

To establish the second property we must show that any limit point P has an infinite number of limit points in its vicinity. If P_1 and P_2 are two other limit points, at least one of them has an infinite number of transforms in the neighborhood of P (Theorem 2). As these transforms are limit points, the second property of the perfect set is established.

There are groups of transformations—the finite groups—with no limit points. Groups with a single limit point and groups with two limit points exist. A group other than these simple kinds has an infinite number of limit points. Furthermore, by a well-known property of perfect sets, the limit points are non-denumerable.

THEOREM 5.—*If a closed set of points Σ, consisting of more than one point, is transformed into itself by all transformations of the group, then Σ contains all the limit points of the group.*

Suppose, on the contrary, that there is a limit point P not belonging to Σ. Then, since Σ is closed, there is no point of Σ within a suitable neighborhood of P. Let P_1, P_2 be two points of Σ. At least one (Theorem 2) has transforms in the neighborhood of P. This contradicts the hypothesis that Σ is transformed into itself.

As an example of the use of the last theorem, suppose that all the transformations of the group are real. Then the real axis is always transformed into itself. It follows that the limit points all lie on the real axis.

19. Definition of the Region R.—R will consist of all that part of the plane which is exterior to the isometric circles of all the transformations of the group. More accurately, a point z will belong to R if a circle can be drawn with z as center which contains no point interior to an isometric circle. We thus rule out those limit points, if any, which are not themselves within or on an isometric circle but which have arcs of isometric circles in any neighborhood of them. Later, we shall adjoin to R a part of its boundary, but for the present it shall consist only of interior points. It may be a connected region, or it may comprise two or more disconnected parts. We see from (9) that it contains all of the plane lying outside a circle concentric with Q_ρ and of radius 3ρ.

It is clear that no two points of R are congruent. A transformation T carries all points exterior to I_t into the interior of I_t'. Any transformation of the group, except the identical transformation, carries a point of R into an isometric circle and, hence, outside R; so no point of R is congruent to another point of R.

20. The Regions Congruent to R.—If we apply to R the various transformations of the group, there results a set of congruent regions no two of which overlap (Theorem 1). Concerning the distribution of these regions, we have the following important theorem:

THEOREM 6.—*R and the regions congruent to R form a set of regions which extend into the neighborhood of every point of the plane.*

Suppose, on the contrary, that there is a point z_0 enclosed by a circle Q with z_0 as center and of radius r sufficiently small that Q contains neither points of R nor points congruent to points of R. Then, all transforms of Q contain neither points of R nor points congruent to points of R. In particular, Q and its transforms contain the centers of no isometric circles, since these are congruent to ∞, which is a point of R. The interior points of Q and of its transforms are ordinary points.

Since z_0 is not a point of R, z_0 is within or on the boundary of some isometric circle. Similarly, the center of each circle congruent to Q lies within or on the boundary of some isometric circle.

The proof consists in showing that there is a circle congruent to Q of arbitrarily large radius, which constitutes a contradiction. Let S be a transformation whose isometric circle I_s has z_0 for an interior or boundary point. The center of I_s is exterior to Q. Consider the circle Q_1 into which S carries Q. S is equivalent to an inversion in I_s followed by a reflection in a line and possibly a rotation. The magnitude of Q_1 is determined by the inversion.

It is a matter of simple algebra to show that if the center of Q is on I_s the radius of Q_1 is

$$r_1 = \frac{r}{1 - \frac{r^2}{r_s^2}},$$

where r_s is the radius of I_s. If z_0 lies within I_s, r_1 exceeds this value. Since $r_s < 2\rho$ (Equation (9)), and $r < r_s < 2\rho$, we have

$$r_1 > kr, \qquad k = \frac{1}{1 - \frac{r^2}{4\rho^2}} > 1.$$

If we apply to Q_1 a transformation whose isometric circle has the center of Q_1 as an interior or boundary point, we get a circle Q_2 of radius r_2 where, since $r_1 > r$,

$$r_2 > \frac{r_1}{1 - \frac{r_1^2}{4\rho^2}} > kr_1 > k^2 r.$$

Continuing in this manner, we prove the existence of a circle Q_n congruent to Q and of radius exceeding $k^n r$. By taking n large enough, Q_n will contain points of R exterior to the finite region in which the isometric circles lie. These points are congruent to points of Q, a contradiction which proves the theorem.

We can now establish the following result:

THEOREM 7.—*R constitutes a fundamental region for the group.*

We have already shown that no two points of R are congruent. We must show further that in any circle Q about a point P on the boundary of R there are points congruent to points of R. Let z_0 be a point of Q which lies in an isometric circle I. Then in the region common to Q and I, which contains no points of R, there are, by Theorem 6, points congruent to points of R. This establishes the theorem.

THEOREM 8.—*Any closed region not containing limit points of the group is covered by a finite number of transforms of R (including possibly R itself). These regions fit together without lacunæ.*

Let A be a closed region; for example, a region bounded by a simple closed curve, having no limit points of the group in its interior or on the boundary. Then there is a finite number of isometric circles containing points of A. For, if there is an infinite number, there are circles of arbitrarily small radius. Their centers then have a point of A as cluster point, contrary to hypothesis.

A transformation S carries R into a region R_s lying in I_s' the isometric circle of S^{-1}. If I_s' contains points of A, R_s may contain points of A; if A is exterior to I_s', then R contains no points of A. Hence, the number of regions congruent to R which lie wholly or in part in A is not greater than the number of isometric circles which contain points of A. This number is finite. Also, since there are points of R, or points congruent to points of R in the neighborhood of every point of A (Theorem 6), it follows that the regions fit together without lacunæ. A is completely covered, except, of course, for the boundaries separating the various regions.

THEOREM 9.—*Within any region enclosing a limit point of the group, there lie an infinite number of transforms of the entire region R.*

This theorem follows at once from the fact that there is an infinite number of isometric circles lying entirely within a given region enclosing a limit point. Each of these circles contains a region congruent to the entire region R, and the various transforms are different regions.

The preceding theorems furnish a picture of the transforms of R. R and the regions congruent to it fit together to fill up all that part of the plane which is composed of ordinary points. They cluster in infinite number about the limit points.

21 The Boundary of R.—A point on the boundary of R is a point P not belonging to R but such that in any circle with P as center there are points of R. P may be an ordinary point or a limit point. Obviously, P cannot lie within an isometric circle.

If P is an ordinary point, it lies on one or more isometric circles. Since there is but a finite number of isometric circles whose arcs lie in the neighborhood of an ordinary point, a circle Q can be drawn with P as center such that Q is exterior to all isometric circles other than those which pass through P.

In the most general case, a boundary point P belongs to one of the following three categories:

(α) P is a limit point of the group;

(β) P is an ordinary point and lies on a single isometric circle;

(γ) P is an ordinary point and lies on two or more isometric circles. P is then called a "vertex."

It is desirable to include under (γ) the following special case: If P is the fixed point of an elliptic transformation of period two, so that, although P lies on a single isometric circle, it separates two congruent arcs on the circle, we shall classify P under (γ) rather than (β). The advantages of this classification will appear subsequently.

Concerning the boundary points of category (α), there is nothing to be added to the theorems on limit points already derived in Sec. 18. We shall show subsequently (Sec. 25) that groups exist for which the boundary points of R are all limit points. The groups of interest for our present theory, however, possess ordinary boundary points also.

(*a*) *The Sides.*—Consider a boundary point of category (β). Let P lie on I_t, and let P' on I_t' be the point into which T carries P. We shall show that P' is also a boundary point of category (β). We put aside the case $P' = P$, a situation which can arise only if I_t and I_t' coincide and P is a fixed point of the resulting elliptic transformation of period two; for this case has been included in (γ).

First, P' is within no isometric circle. Suppose P' to be within I_s; then S magnifies lengths in the neighborhood of P'. But, since T carries P into P' without alteration of lengths, ST magnifies lengths in the neighborhood of P. Then P is within I_{st}, which is contrary to the hypothesis that P is a boundary point.

Second, P' does not lie on an isometric circle other than I_t'. For, if P' lies on I_s, the transformation ST effects no alteration in lengths at P. Then P is on I_{st}, which is contrary to the hypothesis that P lies on a single isometric circle. It follows from these facts that P' is a boundary point of category (β).

There is no isometric circle in the neighborhood of P other than I_t. It is clear, then, that the points on I_t in the neighborhood of P are likewise boundary points of category (β); so, consequently, are the congruent points on I_t'. We thus have as a part of the boundary an arc of I_t and the congruent arc on I_t'. These arcs may consist of the entire circles or they may terminate in points of category (α) or (γ). Since the arcs lie on isometric circles, they are of equal length. We have, then, the following theorem:

THEOREM 10.—*The boundary points of R of category (β) form a set of bounding circular arcs, or sides, which are congruent in pairs. Two such congruent sides are equal in length.*

(*b*) *The Vertices.*—There remain for consideration the boundary points of category (γ). Through a point P of this category there pass a finite number of isometric circles. Let Q be a circle about P sufficiently small that all isometric circles other than those through P are without Q and such that any points of intersection of the circles through P, other than P itself, lie without Q. The isometric circles through P divide Q into a finite number of parts. One of these parts A, owing to the assumption that P is a boundary point, belongs to R. The two arcs which bound A, on I_t and I_s, say, are a part of the boundary of R. The points of these arcs other than P belong to category (β). That is, at a vertex two sides of R meet.

Now make the transformation T, P being carried into P' on I_t'. By reasoning almost identical with that employed in the preceding case, we show that P' is a vertex. We can show, in fact, that P and P' lie on the same number of isometric circles. Let P lie on the isometric circles of $T, S, U, \ldots$; then P' lies on the isometric circle of T^{-1}, and also on those of ST^{-1}, $UT^{-1}, \ldots$, as we see on considering the way in which lengths in the neighborhood of P' are affected by the transformations. These transformations are different—thus, $ST^{-1} = UT^{-1}$ implies $S = U$—hence, their isometric circles are different. Hence, as many isometric circles pass through P' as through P. Interchanging the rôles of P and P', as many pass through P as through P'.

(*c*) *Extension of R*.—We shall find it convenient to add to the region R certain points of its boundary. Of two congruent sides, one, exclusive of the end points, may be added without including points congruent to points previously in R. A vertex where bounding arcs meet may be congruent to several other vertices, one of which may be adjoined. The resulting region is still a fundamental region.

22. Example.—A Finite Group.—The following example of a finite group illustrates parts of the preceding discussion. The transformations

$$T_0 = z, \qquad T_1 = \frac{2z - 1}{3z - 2}, \qquad T_2 = \frac{z}{5z - 1},$$

$$T_3 = \frac{3z - 1}{7z - 2}, \qquad T_4 = \frac{2z - 1}{7z - 3}, \qquad T_5 = \frac{3z - 1}{8z - 3},$$

constitute a group. This is, in fact, the group got by transforming the group of anharmonic ratios (Sec. 13 (2)) by $G = 1/(z + 2)$ so that ∞ shall not be a fixed point of any transformation. The isometric circles are

$$I_1: |z - \tfrac{2}{3}| = \tfrac{1}{3}, \qquad I_2: |z - \tfrac{1}{5}| = \tfrac{1}{5},$$

$$I_3: |z - \tfrac{2}{7}| = \tfrac{1}{7}, \qquad I_4: |z - \tfrac{3}{7}| = \tfrac{1}{7}, \qquad I_5: |z - \tfrac{3}{8}| = \tfrac{1}{8}.$$

These are shown in Fig. 15.

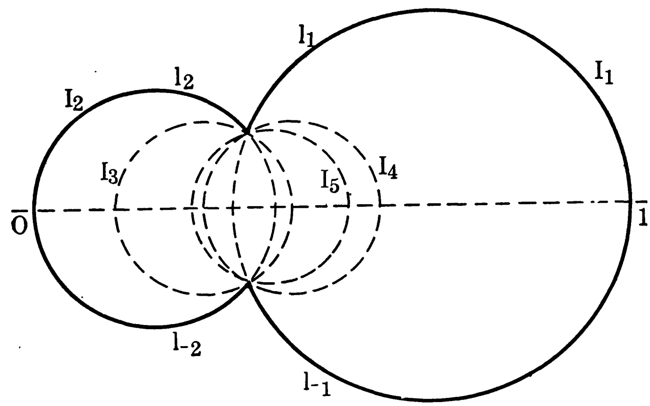


FIG. 15.

The fundamental region R is bounded by arcs of I_1 and I_2. T_1 is an elliptic transformation of period two $[a + d = 0]$, one of the fixed points being 1. T carries the upper half l_1 of the bounding arc on I_1 into the lower half l_{-1}. Similarly, T_2 is of period two, the upper and lower sides, l_2 and l_{-2} being congruent. There are thus two pairs of equal congruent sides.

Two congruent vertices are $\frac{5}{14} \pm i\sqrt{3}/14$, through each of which pass all five isometric circles. The points 0 and 1 are also vertices.

R and the five regions congruent to R fill up the whole plane without overlapping. The congruent regions are not drawn, but it is not difficult to verify that the six regions consist of the six parts into which the plane is divided by the complete isometric circles I_1, I_2, and I_5.

23. Generating Transformations.—In Sec. 13, mention was made of the group resulting from combining a finite or infinite number of linear transformations in all possible ways. We shall now raise the question whether a knowledge of the fundamental region R furnishes any information as to a set of generating transformations of the group.

Let $l_1, l_{-1}, l_2, l_{-2}, \ldots$ be the congruent pairs of sides of R. Let T_1 carry l_1 into l_{-1}; T_2 carry l_2 into l_{-2}; and so on. We shall show that under certain circumstances $T_1, T_2, \ldots$ form a set of generating transformations for the group.

It is clear that the group formed by all possible combinations of $T_1, T_2, \ldots,$ and their inverses contains only transformations of the original group. That is, the group generated by T_1, $T_2 \ldots$, which we shall designate by Γ_t, is a *subgroup* of the original group. The question at issue is whether Γ_t contains all the transformations of the group or only a part of them.

Let us consider the way in which R is transformed by the transformations of the group Γ_t. For convenience in notation, let us represent T_n^{-1} by T_{-n}. T_1 carries R into a region R_1 adjoining R along the side l_{-1}; $T_{-1}(= T_1^{-1})$ carries R into a region R_{-1} adjoining R along the side l_1. In general, T_n carries R into a region R_n adjoining R along the side l_{-n}. Hence, $T_1, T_2, \ldots$ and their inverses carry R into regions adjacent to R along all its sides.

Let S be any transformation of the original group, and let S carry R into R_s. R_s is bounded by circular arcs congruent to the sides l_n of R. The regions R_n adjoining R along its sides are carried by S into regions adjoining R_s along all its sides. That is, $R_s(= S(R))$ is surrounded along all its sides by the regions $ST_n(R)$, $n = \pm 1, \pm 2, \ldots$

In particular, let S belong to the group Γ_t. Then, ST_n belongs to Γ_t, and the regions surrounding R_s are transforms of R by transformations of the group Γ_t. The transforms of R by the group Γ_t are entirely surrounded by other such transforms, the regions fitting together, without lacunæ and with no free sides, to form one or more networks of regions. If these networks fill up the whole plane, Γ_t coincides with the original group; otherwise it does not.

Let A be a closed region containing no limit points of the group and having points in common with the region R. Then R and a finite number of its transforms by the group Γ_t cover A completely. For, along the sides of R which lie in A, we can adjoin regions of the network; along the sides of the new regions which lie in A, we can adjoin other regions of the network; and so on. We can continue this process as long as there are any free sides lying in A. But A is covered completely by a finite number of regions congruent to R (Theorem 8); hence, this process must end in a finite number of steps.

By similar reasoning, if the region A contains points of any region R_t congruent to R by a transformation of Γ_t, A can be covered by a finite number of regions congruent to R by the transformations of Γ_t.

We can now prove the following theorem:

THEOREM 11.—*If some point of R can be joined to all its congruent points by curves not passing through limit points, then the transformations T_1, T_2, . . . by which the sides of R are congruent constitute a set of generating transformations for the group.*

Let S be any transformation of the group. Let z_0 be a point of R which can be joined to any congruent point by a curve not passing through a limit point. Let C be such a curve joining z_0 to $z_0' = S(z_0)$. C can be embedded in a closed region A consisting entirely of ordinary points.

A is covered by R and a finite number of regions congruent to R by transformations of the group Γ_t. In particular, there is a transformation T of Γ_t which carries R into a region covering the neighborhood of z_0'. The transforms of R by S and T overlap, whence $S = T$. Any transformation of the original group belongs to Γ_t, which was to be proved.

There are certain cases in which we can state at once that T_1, T_2, . . . are generating transformations. The first is the finite group. Since there are no limit points, the conditions of the theorem hold. For example, in the group given in Sec. 22, T_1 and T_2 are generating transformations. The reader can verify that their combinations give the remaining transformations.

Again, the distribution of the limit points may be such that any two points which are not limit points may be joined by a curve not passing through the limit points. Then, obviously, the conditions of the theorem hold. The simplest examples are the groups with a finite number (one or two) of limit points.

24. Cyclic Groups. DEFINITION.—*A cyclic group is a group generated by a single transformation.*

If T is the transformation which generates the group, then the group consists of the transformations . . . T^{-2}, T^{-1}, 1, T, T^2, . . . The groups whose fundamental regions are shown in Figs. 11, 12, 14 are cyclic groups. In the examples, all the transformations have a fixed point at infinity. In the present section, we shall examine the cyclic groups arising when T is one or another of the various types of transformations and where the fixed points are finite. A knowledge of cyclic groups is important because of the fact that every group contains cyclic subgroups. For, the group generated by any transformation of the original group belongs to that group.

If there are two fixed points, ξ_1 and ξ_2, T can be written in the form (Sec. 6, Equation 24).

$$\frac{z' - \xi_1}{z' - \xi_2} = K\frac{z - \xi_1}{z - \xi_2}, \; K \neq 1, \tag{11}$$

and the general transformation T^n is

$$\frac{z' - \xi_1}{z' - \xi_2} = K^n\frac{z - \xi_1}{z - \xi_2}. \tag{12}$$

On solving for z', we have

$$z' = T^n(z) = \frac{(K^n\xi_2 - \xi_1)z + (1 - K^n)\xi_1\xi_2}{(K^n - 1)z + \xi_2 - K^n\xi_1}.$$

We find the determinant to be $K^n(\xi_1 - \xi_2)^2$, so we must divide numerator and denominator by $K^{n/2}(\xi_1 - \xi_2)$ to render the determinant 1. The isometric circle I_n of T^n is

$$\left|z + \frac{\xi_2 - K^n\xi_1}{K^n - 1}\right| = \left|\frac{K^{n/2}(\xi_1 - \xi_2)}{K^n - 1}\right| = \left|\frac{\xi_1 - \xi_2}{K^{n/2} - K^{-n/2}}\right|. \tag{13}$$

(*a*) *Hyperbolic and Loxodromic Cyclic Groups.*—If T is hyperbolic, K is real and $|K| \neq 1$; then the multiplier K^n of T^n, $n \neq 0$, is likewise real and in absolute value unequal to 1; whence all transformations of the group are hyperbolic. If K is not real and $|K| \neq 1$, then $|K^n| \neq 1$, $n \neq 0$, although in certain cases K^n may be real; hence, if T is loxodromic all the transformations of the group are loxodromic or hyperbolic.

We found in Sec. 11 (*b*) that a fixed point cannot lie on either of the isometric circles I_n, I_n' of a hyperbolic or loxodromic transformation T^n and its inverse. Clearly, there can be no fixed points outside both I_n and I_n', for T^n shifts the position of such a point. Each I_n contains a fixed point. As n increases, the radius of I_n approaches zero; so each fixed point is a limit point of the group. An application of Theorem 5, where we let the set Σ consist of the two fixed points, shows that there are no further limit points.

In studying the arrangement of the isometric circles we shall make use of the following proposition:

THEOREM 12.—*Let I_t, I_s, I_s', I_u be the isometric circles of T, S, S^{-1}, $U = ST$, respectively. If I_t and I_s' are exterior to one another, then I_u is contained in I_s.*

If I_t and I_s' are tangent externally, then I_u lies in I_s and is tangent internally.

The proof is simple. Suppose the circles not tangent; and let z be a point outside (or on) I_s. Then S carries z into a point z' within (or on) I_s' with decrease of lengths (or without alteration of lengths). Now z' is outside I_t; so T transforms z' with decrease of lengths. Hence, U transforms z with decrease of lengths, whence z is outside I_u. Since every point on or outside I_s is also outside I_u, the latter circle is contained in the former.

If I_t and I_s' are tangent externally at a, the preceding reasoning holds except at the point a_0 on I_s which S carries into a. TS makes no alteration of lengths at a_0, whence a_0 lies on I_u.

If, in the cyclic group, T is hyperbolic or if T is loxodromic and $|a + d| \geqslant 2$, I and I', the isometric circles of T and T^{-1}, are exterior to one another (possibly tangent). We now show that I encloses I_2, I_2 encloses I_3, etc., and, likewise, I' encloses I_2', I_2' encloses I_3', and so on. We see at once from Theorem 12, taking $S = T$, that I_2 is contained in I. Similarly, from the product $T^{-1}T^{-1}$, I_2' lies in I'.

We establish the general relation by induction. Suppose that the circles are arranged as stated up to I_n and I_n', and consider I_{n+1}. We write $T^{n+1} = TT^n$. By hypothesis I_n', the isometric circle of T^{-n}, lies in I', and hence is exterior to I. It follows from Theorem 12 that I_{n+1} lies within I_n. By identical reasoning I'_{n+1} is contained in I_n'.

It follows from the preceding that I and I' contain all other isometric circles, whence the fundamental region R is the region lying outside these two circles. In Fig. 16, the isometric circles are drawn for a group generated by a hyperbolic transformation with $K = 4$.

The situation is somewhat different for a loxodromic transformation for which $|a + d| < 2$. In Fig. 17 the isometric circles are shown for the group generated by $T = 1/(z + 1)$. Here $a + d = i$, whence T is loxodromic. The congruent boundaries of R are connected by the transformations T and T^2.

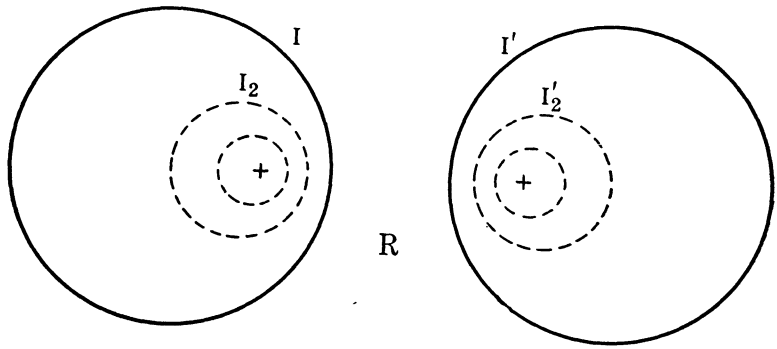


FIG. 16.

According to Theorem 11, these are generating transformations for the group. This example illustrates the fact that the generating transformations found by an application of Theorem 11 are not necessarily the best obtainable, that some of the transformations found in this way may be consequences of others.

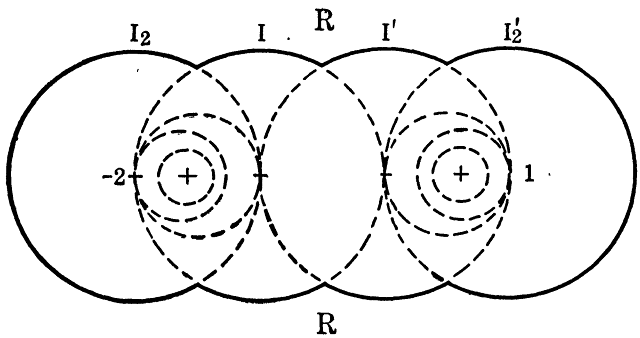


FIG. 17.

A simpler fundamental region for this case, although not bounded by isometric circles, is got as follows. The loxodromic transformation may be written $T = UV$, where V and U are hyperbolic and elliptic transformations, respectively, with the same fixed points as T. Let I and I' be the isometric circles of V and V^{-1}. Then, it is readily shown that the part of the plane exterior to I and I' is a fundamental region for the group generated by T.

(*b*) *Elliptic Cyclic Groups.*—If T is elliptic, all isometric circles pass through the fixed points. Here $K = e^{i\theta}$, and unless θ is commensurable with π, the group is continuous. In Fig. 18, the isometric circles are drawn for $K = e^{6\pi i/5}$. The group may be generated. as the figure shows, by T^2 which has the multiplier $e^{2\pi i/5}$.

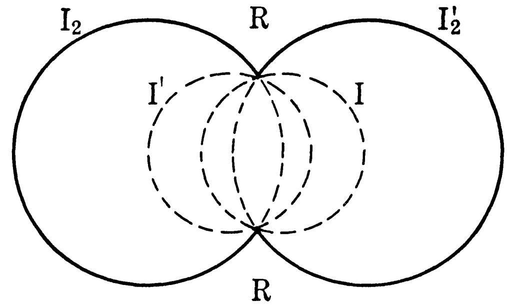


FIG. 18.

(*c*) *Parabolic Cyclic Groups.*—If T is parabolic, the transformation may be written (Sec. 10, Equation 37)

$$\frac{1}{z' - \xi} = \frac{1}{z - \xi} + c, \quad c \neq 0. \tag{14}$$

Then T^n is the transformation

$$\frac{1}{z' - \xi} = \frac{1}{z - \xi} + nc, \tag{15}$$

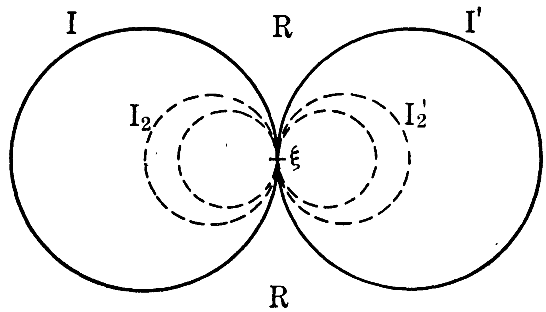


FIG. 19.

or

$$z' = \frac{(nc\xi + 1)z - nc\xi^2}{ncz + 1 - nc\xi}. \tag{16}$$

The determinant in (16) is 1 and the isometric circle I_n is

$$\left| z - \left(\xi - \frac{1}{nc} \right) \right| = \frac{1}{|nc|}. \tag{17}$$

We see from (17) that ξ lies on I_n, and that the center of I_n, namely $\xi - \frac{1}{nc}$, lies on the line joining ξ to $1/c$. Hence, all isometric circles have a common tangent at ξ. If n is positive, the center of I_n lies on one side of ξ; if negative, on the other. As n increases in absolute value, the radius steadily diminishes. From these facts it follows that the arrangement of the isometric circles is as shown in Fig. 19. This can also be shown from Theorem 12. The fixed point is the only limit point of the group.

25. The Formation of Groups by the Method of Combination. We here explain a method of forming properly discontinuous groups, by the use of which a great variety of groups with very diverse properties can be constructed. The reader will gain some idea of the intricate possibilities in the broad class of properly discontinuous groups.

Given a finite or infinite number of properly discontinuous groups $\Gamma_1, \Gamma_2 \dots$ Let the transformations of these groups be combined in all possible ways to form a group Γ. We get two kinds of transformations: (*a*) those belonging to the original groups; and (*b*) *cross-products* resulting from the combination of transformations not all belonging to the same original group. The resulting group may be continuous or discontinuous. In certain cases, however, we are able to state that the group is properly discontinuous and to specify its fundamental region R.

THEOREM 13.—*Let the R-regions $R_1, R_2, \dots$ of the groups $\Gamma_1, \Gamma_2, \dots$ all contain some neighborhood of infinity; and let the isometric circles of all transformations of each group be exterior (possibly tangent externally) to all isometric circles of all the transformations of the remaining groups. Then the group Γ formed by combining the given groups is properly discontinuous.*

The region containing all points common to $R_1, R_2, \dots$ is the region R for Γ. Here a point is not counted as belonging to R unless a region about the point lies in $R_1, R_2, \dots$

R is the region lying outside the isometric circles of the transformations of Γ of the category (*a*). The cross-products have not been taken into account. We now show that the isometric circles of the transformations (*b*) contain no points of R.

Consider the general cross-product

$$U = S_n S_{n-1} \cdots S_2 S_1.$$

Here we suppose that the two successive transformations S_i, S_{i+1} do not belong to the same one of the groups $\Gamma_1, \Gamma_2, \dots$;

otherwise, we combine them into one transformation. We apply the transformation U to a point z_0 of R. Let I_i, I_i' represent the isometric circles of S_i, S_i^{-1}. Since z_0 is exterior to I_1, S_1 carries z_0 into z_1 within I_1'. Since I_1' is exterior to I_2, z_1 is exterior to I_2 and S_2 carries z_1 into z_2 within I_2'; and so on. At each step, lengths in the neighborhood of the point are decreased. So U decreases lengths in the neighborhood of z_0; hence z_0 is exterior to the isometric circle of U. Thus the region R of the theorem is the region lying outside all the isometric circles belonging to the group Γ. The existence of R shows that Γ is properly discontinuous.

By the use of Theorem 12, employing reasoning similar to that of the preceding section, we can show that the isometric circles of

$$S_1,\ S_2S_1,\ S_3S_2S_1,\ \cdots,\ U$$

form a sequence such that each circle encloses the circle which follows it. It follows that the isometric circle of a cross-product lies within one of the isometric circles of the original groups, and hence has no bearing on the construction of R.

A few examples wherein we actually construct combination groups will throw some light on the various forms which the region R can take.

Given two equal circles, we can set up infinitely many transformations T such that the given circles are the isometric circles of T and T^{-1}. It is easily shown that the most general linear transformation such that I_t and I_t' are, respectively,

$$|z - q| = r, \quad |z - q'| = r,$$

is

$$T = \frac{q'z - (qq' + r^2e^{i\theta})}{z - q}, \tag{18}$$

where θ is any real quantity. If I_t and I_t' are exterior to one another, the region exterior to I_t and I_t' is the region R for the cyclic group generated by T. We shall use groups of this sort in applying Theorem 13.

(*a*) Given $2n$ circles $I_1, I_1'; \ldots; I_n, I_n'$, which are equal in pairs, and which are exterior to one another or are externally tangent. We set up the transformation T_i so that T_i and T_i^{-1} have the isometric circles I_i and I_i'. Then the group generated by $T_1, T_2, \ldots T_n$ has for a fundamental region the region outside the $2n$ circles.

If no circles are tangent, R is a connected region, although not simply connected. Its boundary consists entirely of sides, there being no boundary points of categories (α) or (γ) (Sec. 21).

By making the circles tangent in various ways, the region exterior to the circles can be separated into several regions. Then R will consist of several parts.

(b) In the same way, we can select an infinite number of circles equal in pairs and construct a group. Since we are assuming that the isometric circles lie in some finite domain, we shall have circles of arbitrarily small radius.

These circles can be put together in such a way that R is composed of an infinite number of separate parts.

(c) That R can be a region whose boundary points are all limit points is shown by the following example:

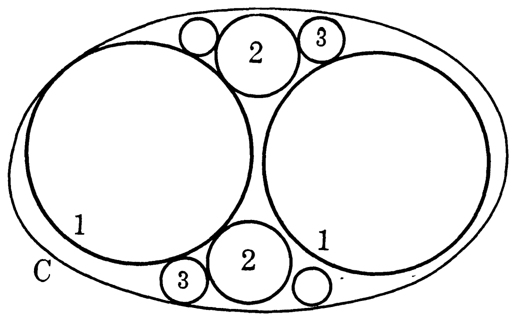


FIG. 20.

Let C be a closed curve (Fig. 20). Without going into detail, it is seen that pairs of isometric circles with which to form a combination group can be put into the interior of C in such a manner that every point within or on C which is not within a circle has an infinite number of circles in its neighborhood. Then R comprises the exterior of C. The boundary of R, namely C itself, is composed entirely of limit points.

Constructions of this sort for other and more complicated forms of the region R will occur to the reader.

An Extension. Schottky Groups.—Results analogous to Theorem 13 can be derived for other kinds of fundamental regions. Let $\Gamma_1, \ldots, \Gamma_m$ be groups with fundamental regions $F_1, \ldots, F_m$. Let F_i contain in its interior all points of the plane not interior to $F_j (i \neq j)$. If the combination group Γ be formed we can show that the region F consisting of all points common to $F_1, \ldots, F_m$ is a fundamental region for Γ. The boundary of F consists of the boundaries of $F_1, \ldots, F_m$.

In the neighborhood of a point P on the boundary of F_i there exist points not belonging to F_i which are congruent to points of F_i. Further, these points can be so chosen that the congruent points in F_i lie as near the boundary of F_i as we wish and so belong to F. The second requirement of the fundamental region is satisfied.

Obviously, no two points of F are congruent by transformations of the original group. We now consider the effect of a cross-product on a point z_0 of F. Let $U = S_n \cdots S_1$, where $S_i(\neq 1)$ belongs to the group Γ_{n_i}, and Γ_{n_i} and $\Gamma_{n_{i+1}}$ are different groups. S_1 carries z_0 into a point z_1 outside F_{n_1} and hence in F_{n_2}; S_2 carries z_1 into a point z_2 outside F_{n_2} and hence in F_{n_3}; etc. Finally, S_n carries z_{n-1} into a point z_n outside F_{n_n}; that is, U carries z_0 into a point outside F, and the first property of the fundamental region is established.

The *Schottky group*[1] is constructed as follows: Let Q_1, Q_1'; $\ldots$; Q_m, Q_m' be m pairs of circles external to one another. Let T_i be a linear transformation (loxodromic or hyperbolic) carrying Q_i into Q_i' in such a way that the exterior of Q_i is carried into the interior of Q_i'. T_i generates a cyclic group Γ_i for which all that part F_i of the plane exterior to Q_i and Q_i' is a fundamental region. Here F_i contains all that part of the plane not contained in $F_j(j \neq i)$. The Schottky group Γ is got by combining $\Gamma_1, \ldots, \Gamma_m$. It has as fundamental region all that part of the plane exterior to the $2m$ circles. Γ is generated by the transformations $T_1, \ldots, T_m$.

In a subsequent chapter (Chap. X) there will arise groups generated as is the Schottky group except that $Q_1, \ldots, Q_m'$ are closed curves which are not necessarily circles. Such a combination group is called a "*group of Schottky Type.*"

26. Ordinary Cycles.—Returning now to the fundamental region R for the general properly discontinuous group, we shall make a study of the vertices. Let A_1 be a vertex. If either of the sides which meet in A_1 is carried into its congruent side, A_1 is carried into a vertex at the extremity of this latter arc. These congruent vertices may be carried into others, with the result that A_1 may be congruent to several of the vertices of R.

[1] *Crelle's Jour.*, Vol. 101, pp. 227–272, 1887.

We may find the vertices congruent to A_1 by the following method: Let us think of the boundary of R as being traced in a positive sense; that is, with the region on the left. In passing through a vertex, we proceed along one side to the vertex and then proceed along a second side from the vertex. We consider the vertex as the end of the former side and the beginning of the latter. When a side is carried into its congruent side, the beginning and end of the former are carried into the end and beginning, respectively, of the latter; this results from the fact that the direction around the isometric circle is reversed (Sec. 11 (*a*)).

Let l_1 be the side beginning at A_1 (Fig. 21). This side is carried by a transformation T_1 into the congruent side l_{-1}, A_1 being carried into A_2 at the end of l_{-1}. There is a side l_2 beginning at A_2. The arc l_2 is carried by some transformation T_2 into the congruent side l_2, A_2 being carried into A_3 at the end of l_{-2}. We can continue in this manner, getting other congruent vertices until we return to A_1 and the side l_1.

We will return to A_1 in a finite number of steps. Suppose, on the contrary, that an infinite number of vertices $A_2, A_3, \ldots$ are congruent to A_1. The transformation S_n which carries A_1 to A_n has an isometric circle passing through A_1; for if A_1 is outside the isometric circle of S_n, A_n is within the isometric circle of S_n^{-1}, which is impossible. Then the isometric circles of $S_2, S_3, \ldots$, an infinite number, pass through A_1, which is contrary to the hypothesis that A_1 is an ordinary point. Hence, in applying the process just explained, we encounter a finite number of vertices $A_2, A_3, \ldots A_m$, congruent to A_1 and then return to A_1.[1]

DEFINITION.—*A complete set of congruent vertices of a fundamental region is called an ordinary cycle.*

We shall show presently that there are no vertices of R congruent to A_1 other than those just found; whence $A_1, A_2 \ldots A_m$ constitute a cycle.

Let $T_1, T_2, \ldots, T_m$ be the transformations in the preceding treatment by which we carry A_1 to A_2, A_2 to $A_3 \ldots A_m$ to A_1, respectively. Some of these transformations, it will be noted, may be inverses of the others, but each connects a pair of bound-

[1] We continue until we encounter the vertex A_1 *followed by the side* l_1. We may encounter A_1 once before this happens, in the special case that the sides at A_1 are tangent. See Fig. 24 p. 73.

ing arcs. The transformation $S = T_m \cdot \cdot \cdot T_2T_1$ carries A_1 into itself. It may happen that S is the identical transformation. If not, S is an elliptic transformation. For, A_1 is a fixed point of S; and the fixed points of hyperbolic or loxodromic transformations lie within isometric circles and the fixed point of a parabolic transformation (Fig. 19) is a limit point of the group.

Consider now the way in which the transforms of R fit together at A_1. The transformations T_m, T_mT_{m-1}, . . . , $T_m \cdot \cdot \cdot T_2$, $T_m \cdot \cdot \cdot T_1(= S)$ carry A_m, A_{m-1}, . . . A_2, A_1, respectively, into A_1. Further, the regions R_m, R_{m-1}, . . . , R_2, R_1, respectively, into which these transformations carry R fit together

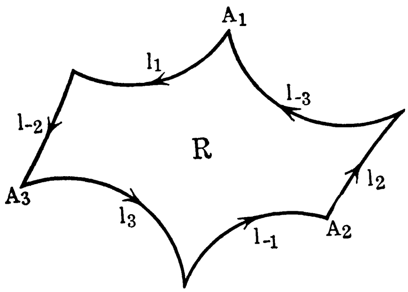


FIG. 21.

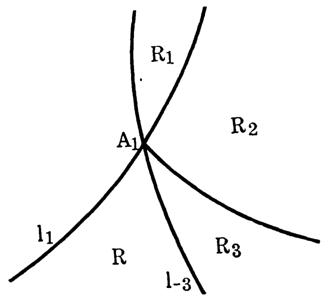


FIG. 22.

at A_1 (Fig. 22). Thus, T_m carries A_m into A_1, R_m being adjacent to R along the arc l_{-m} ending at A_1. In general, T_i carries A_i into A_{i+1} and carries R into a region abutting on R along a side ending in A_{i+1}, whence the transforms of these two regions by $T_m \cdot \cdot \cdot T_{i+1}$, namely R_i and R_{i+1}, are adjacent along an arc issuing from A_1. In proceeding counterclockwise around A_1, starting from R, we encounter in order the adjacent regions R_m, R_{m-1}, . . . , R_2, R_1. The curvilinear angle of R at A_n is carried into an equal angle of R_n at A_1.

Since there can be no overlapping of congruent regions, there are two possibilities. First, R_1 may coincide with R and the regions R, R_m, . . . , R_2 completely fill up the angle about A_1. Then, S is the identical transformation. The sum of the angles at the vertices A_1, A_2, . . . , A_m is equal to 2π.

If R_1 does not coincide with R, S is an elliptic transformation. Now, an elliptic transformation with multiplier $K = e^{i\theta}$ amounts to a rotation in the neighborhood of the fixed point through the angle $\theta \neq 0$. Carrying R into R_1 requires (Fig. 22) that θ be equal to the sum of the angles at the vertices. On applying S, R, R_m, . . . , R_2 are carried into adjacent regions filling out more of the region about A_1 counterclockwise from R_1. After a finite number, k, of applications of S, the region about A_1 is completely filled up and $S^k(R)$ coincides with R.

It is now clear that there are no vertices of R congruent to A_1 other than A_2, . . . , A_m. For, a transformation carrying any other vertex A to A_1 carries R into a region overlapping the regions which fill out the angle about A_1, which is impossible.

We state the preceding results in the following form:

THEOREM 14.—*The sum of the angles at the vertices of an ordinary cycle is $2\pi/k$, where k is an integer. If $k > 1$, each vertex of the cycle is a fixed point of an elliptic transformation of period k.*

THEOREM 15.—*Each ordinary cycle determines a relation of the form $(T_m T_{m-1} \cdots T_2 T_1)^k = 1$ satisfied by the transformations connecting congruent sides of R.*

As an illustration of the preceding results, we take the group of Sec. 22, the fundamental region of which is shown in Fig. 15. Consider the upper vertex in that figure. The side l_1 beginning at that vertex is carried by T_1 into l_{-1} ending at the lower vertex. The side l_{-2} beginning at the lower vertex is carried by T_2^{-1} into l_2 ending at the upper vertex. The upper and lower vertices thus constitute a cycle.

The transformation $S = T_2^{-1}T_1$ is clearly not the identical transformation, since the sum of the angles at the two vertices is less than 2π; hence, S is an elliptic transformation with the upper vertex as fixed point. We find readily that $S = T_4$, and that $S^3 = 1$. The sum of the angles at the two vertices is $2\pi/3$.

Again the origin is a vertex. The arc l_2 beginning at the origin is carried by T_2 into the arc l_{-2} ending at the origin. Hence, the origin alone constitutes a cycle. The angle there is π; whence $T_2^2 = 1$.

Similarly, the vertex at the point 1 constitutes a cycle. There are thus three cycles.

27. Parabolic Cycles.—If a side of R terminates in a limit point P_1 various situations may arise. It may happen that there is no other side terminating in P_1. Let us suppose that two sides meet in P_1; and let us apply the method given in the preceding section for getting points congruent to P_1.

Let l_1 be the side beginning at P_1. Then l_1 is carried by a transformation T_1 into a side l_{-1} ending at a limit point P_2. Let l_2 be the side, if any, beginning at P_2, and let T_2 carry l_2 into the side l_{-2} ending at P_3; and so on.

In the application of this process there are three situations that may arise: (1) the process may be terminated at some stage by arriving at a point at which no side begins; (2) the process may continue ad infinitum without a return to P_1; (3) after arriving at a finite number of congruent points $P_2, P_3, \ldots, P_m$, we may return to P_1 and the side l_1. It is not difficult to set up combination groups exemplifying each possibility.

If (3) holds, we say that $P_1, P_2, \ldots, P_m$ constitute a *parabolic cycle*, and each point of the cycle is called a "*parabolic point.*"

In many of the groups to be studied subsequently R has a finite number of sides and the only limit points on the boundary

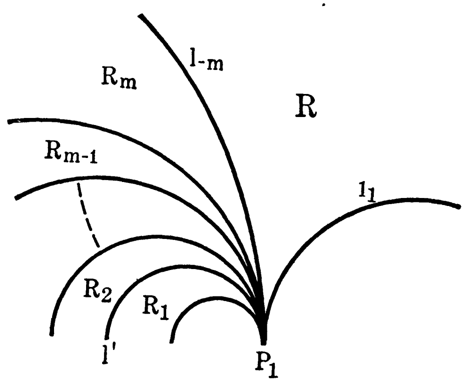


FIG. 23.

are points where two sides meet. These conditions rule out cases (1) and (2); whence the limit points on the boundary are parabolic points arranged in cycles.

The transformation $S = T_m T_{m-1} \cdots T_2 T_1$ carries P_1 into itself; whence S is either elliptic or parabolic. The reasoning of the preceding section can be repeated word for word to show that $T_m, T_m T_{m-1}, \ldots, T_m \cdots T_2, T_m \cdots T_1$ carry $P_m, P_{m-1}, \ldots, P_2, P_1$, respectively, into P_1 and carry R into regions $R_m, R_{m-1}, \ldots, R_2, R_1$ fitting together along arcs issuing from P_1 in the order shown in Fig. 23. The angle between the sides of R_i which meet at P_1 is equal to the angle between the sides of R which meet at P_i.

S carries R into R_1, the side l_1 being carried into the side l' which separates R_1 and R_2. If S is elliptic l_1 and l_1' meet at an angle different from zero. By repeating S a finite number of times, the neighborhood of P_1 is covered by a finite number of regions, which is contrary to Theorem 9. Hence S is parabolic.

Since S is parabolic, l_1 and l' are tangent at the fixed point P_1. Then, the arcs bounding the intervening regions are also tangent to l_1. The angle between the sides which meet at each point of the cycle is zero.

By repeated application of S to the regions in Fig. 23, we get infinitely many other regions with two sides meeting at P_1, the sides being tangent to l_1.

THEOREM 16.—*The sides of R which meet at a parabolic point are tangent. There is an infinite number of regions congruent to R each having two sides which meet at the parabolic point and are tangent to the sides of R.*

If the sides of R meet always at vertices or at parabolic points, the region is of a particularly simple kind.

THEOREM 17.—*If the boundary of R consists entirely of ordinary points, or if the only limit points on the boundary are parabolic points, then R has a finite number of sides.*

Suppose that R has an infinite number of sides. Let z_1, z_2, . . . be an infinite suite of points each lying on a side of R and no two lying on the same side. These points have at least one cluster point P, which is also a boundary point. In the neighborhood of P lie infinitely many sides of R. This is impossible either at an ordinary point or at a parabolic point; and the theorem is established.

28. Function Groups.—The ordinary points of a properly discontinuous group either form a single connected region, or two-dimensional continuum, Σ, or else are separated by the limit points into two or more two-dimensional continua Σ, Σ_1, Σ_2, . . .

DEFINITION.—*A properly discontinuous group will be called a function group if one of the connected regions Σ into which the limit points separate the plane is carried into itself by all the transformations of the group.*

In a finite group or a group with one or two limit points the ordinary points constitute a single connected region which is carried into itself by all transformations, whence the group is a function group. A group of the sort shown in Fig. 20 is not a function group. It is the function groups only that will play a part in our subsequent theory.

THEOREM 18.—*In the region Σ of a function group lies a part R_0 of the region R, and R_0 is a fundamental region.*

Not all of R can be exterior to Σ, for then the transforms of R would all be exterior to Σ, contrary to Theorem 6. Let P be a boundary point of R_0. If P is a limit point, there are points congruent to points of R_0 in the neighborhood of P by Theorem 2. If P is an ordinary point on the boundary of R_0, it lies in Σ and there are points congruent to points of R in its neighborhood. But these points are congruent to points of R_0 since the transforms of all other points of R are exterior to Σ. Since obviously no points of R_0 are congruent, R_0 is a fundamental region for the group.

Theorem 19.—*The sides of R_0 are congruent in pairs; and the transformations connecting congruent sides of R_0 form a set of generating transformations for the function group.*

It is clear that R_0 is bounded in part by sides. For, otherwise, R_0 is bounded entirely by limit points (or consists of the whole plane if there are no limit points) and coincides with Σ. Then, no two points of Σ are congruent, which is impossible. (We assume here that the function group does not consist solely of the identical transformation.) Each side of R_0 is congruent to some side of R. But the congruent side must lie in Σ and so is a side of R_0.

Since Σ is, by hypothesis, a connected region, any interior point of R_0 can be joined to any of its congruent points by a curve not passing through a limit point of the group. The latter part of the theorem then follows from Theorem 11.

Theorem 20.—*If a function group possesses a fundamental region F which, together with its boundary, consists of interior points of Σ, and whose transforms cover the neighborhood of each of its boundary points, then the boundary of R_0 consists of interior points of Σ, and R_0 has a finite number of sides.*

A finite number of transforms of R_0 cover F and its boundary completely. Let us carry the portion of F in each region R_i into R_0 by means of the transformation which carries R_i into R_0. These transforms of parts of F do not overlap, since no two points of F are congruent.

Furthermore, they fill R_0 completely without lacunæ. Suppose, on the contrary, that an interior point z_0 of R_0 is exterior to all the transforms of parts of F. Let z_1 be the nearest point of one of the parts in R_0. Then in any neighborhood of z_1 are points which are not covered and which are, therefore, not congruent to points of F. On carrying the particular part back to F, z_1 goes into a boundary point of F which has in its neighborhood points not congruent to points of F. This is impossible.

If R_0 has a limit point on its boundary, so has one of the parts covering R_0 and there is a congruent limit point on the boundary of F, contrary to hypothesis.

We may now imbed the region R_0 in a closed region A consisting entirely of ordinary points. Only a finite number of isometric circles enclose points of A. As isometric circles which are exterior to A cannot form part of the boundary of R_0, it follows that R_0 is bounded by a finite number of sides.

The group is then generated by a finite number of transformations.

Classification of Function Groups.—We shall separate the function groups into three major classes.

1. *Elementary Groups.*—These consist of the finite groups and the groups with one or two limit points.

2. *Fuchsian Groups.*—A group is called "Fuchsian" if its transformations have a common fixed circle and if each transformation carries the interior of the fixed circle into itself.

3. *Kleinian Groups.*—A function group is called "Kleinian" if it does not belong to one of the preceding classes.

Examples of elementary groups are the group discussed in Sec. 22 and the cyclic groups of Sec. 24.

There is a certain amount of overlapping between (1) and (2). Certain of the elementary groups possess fixed circles; for example, the non-loxodromic cyclic groups.

Most of the combination function groups are Kleinian groups.

CHAPTER III

FUCHSIAN GROUPS

29. The Transformations.—As defined in the preceding section, a Fuchsian group is a properly discontinuous group each of whose transformations carries a certain circle into itself and carries each of the parts into which the circle divides the plane into itself. The common fixed circle will be called the "principal circle." A point within the principal circle is carried into an interior point; an exterior point is carried into an exterior point. If the principal circle is a straight line, a point on one side of the line is carried into a point on the same side of the line.

Consider now the kinds of transformations that can belong to a Fuchsian group. We found in Sec. 9 that the only loxodromic transformations which have fixed circles carry the interior of each fixed circle into its exterior. Hence, there can be no loxodromic transformation. Referring to Secs. 7, 8, 10, where we studied the fixed circles of the non-loxodromic transformations, and to the accompanying figures, we see that a transformation of the group, other than the identical transformation, must be one of the following kinds:

(*a*) A hyperbolic transformation with its fixed points on the principal circle.

(*b*) An elliptic transformation with its fixed points inverse to one another with respect to the principal circle.

(*c*) A parabolic transformation with its fixed point on the principal circle and with its fixed straight line tangent to the principal circle.

Conversely, a transformation of any one of these three kinds carries the circle into itself.

An application of Theorem 20, Sec. 11(*c*), gives the following result:

THEOREM 1.—*The isometric circles of the transformations of a Fuchsian group are orthogonal to the principal circle.*

30. The Limit Points.—Since the principal circle is transformed into itself, it constitutes a set to which Theorem 5 of Sec. 18 applies. Hence, we have the following theorem:

Theorem 2.—*The limit points of a Fuchsian group lie on the principal circle.*

This theorem can also be proved directly from Theorem 1. For, the requirements that a limit point shall have isometric circles of arbitrarily small radius in its neighborhood and that these circles shall be orthogonal to the principal circle can be met only if the limit point is on the principal circle.

Theorem 3.—*If there are more than two limit points, either* (1) *the set of limit points consists of all points of the principal circle; or* (2) *the set of limit points is a perfect set which is nowhere dense on the principal circle.*

We have already found (Sec. 18, Theorem 4) that if there are more than two limit points the limit points form a perfect set of points. We must show further that unless every point of the principal circle is a limit point, then the limit points form a set which does not contain all the points of any arc of the principal circle.

Let z_0 be a point lying on the principal circle and which is not a limit point. Then the points in the neighborhood of z_0 are also ordinary points. In particular, the points on a suitably small arc h of the principal circle passing through z_0 are ordinary points.

Let P be a limit point. Then, applying Theorem 2, Sec. 18, there are points congruent to points of h in the neighborhood of P. These points are ordinary points and lie on the principal circle. Since there are ordinary points on the principal circle in the neighborhood of a limit point, it follows that the set of limit points is nowhere dense on the principal circle.

It is known from the theory of perfect sets that the set of limit points in (2) can be formed by the removal from the principal circle of an infinite number of open arcs, h_1, h_2, . . . These arcs do not overlap and have no common end points. Further between any two arcs lie infinitely many others. The perfect set consists of the points that remain after the arcs have been removed.

On the basis of the preceding theorem we shall classify Fuchsian groups as follows:

(*a*) *Fuchsian groups of the first kind,* or groups for which every point of the principal circle is a limit point.

(*b*) *Fuchsian groups of the second kind,* or groups whose limit points are nowhere dense on the principal circle.

To (*b*) belong the groups whose limit points form a non-dense perfect set and also the elementary Fuchsian groups, where the number of limit points is finite.

In (*a*) the limit points separate the plane into two règions each of which is carried into itself. In (*b*) the ordinary points form a single connected region.

We shall find that groups of both kinds exist, and that the two kinds of groups have strikingly different properties.

31. The Region R and the Region R_0.—When we make an inversion in a circle, any orthogonal circle is carried into itself and the interior of the orthogonal circle is carried into itself. Hence, from Theorem 1, if we make an inversion in the principal circle, each isometric circle is carried into itself and its interior and exterior go, respectively, into its interior and exterior. A point of R, being exterior to all the isometric circle, is carried into another point of R. Hence, we have the following result:

THEOREM 4.—*An inversion in the principal circle carries the region R into itself.*

In this inversion, infinity is carried into the center of the principal circle, providing, as we shall assume, that the principal circle is not a straight line. Since the points in the neighborhood of infinity lie in R, it follows that points in the neighborhood of the center of the principal circle lie in R.

The principal circle divides R into two parts which may or may not be connected with one another along the principal circle.

We shall designate by R_0 the part of R lying within the principal circle, by R_0' the part of R lying without the principal circle.

R_0' is the inverse of R_0 in the principal circle; its sides and vertices are the inverses of the sides and vertices of R_0. The sides which bound R_0' lie on the same isometric circles as the sides which bound R_0. Corresponding sides are connected by the same transformations.

Let a transformation of the group be made. R_0 is carried into a region in the interior of the principal circle and R_0' into a region on the exterior. As a consequence of Theorem 10, Sec. 5, we can state that the two transformed regions are inverse with respect to the principal circle.

It follows from the preceding remarks that, in a study of the fundamental region, of its sides and vertices, of its congruent regions, and the like, it will suffice to study the region R_0 within the principal circle. An inversion in the principal circle will then furnish the corresponding results for R_0'. We shall, therefore, limit our study to R_0.

THEOREM 5.—*The region R_0 is simply connected.*

It is clear that the region R_0 is connected. A straight line segment from the center of the principal circle to any point of R_0 lies entirely within R_0. Thus any two points of R_0 can be joined by a curve lying in R_0; for example, by combining the line segments joining each to the center.

R_0 is simply connected if any closed curve in R_0 can be shrunk continuously to an interior point without crossing the boundary. This can be done by the simple process of moving the points of the curve continuously along radii to the center.

The main facts concerning the sides of R_0 are summarized in the following theorem:

THEOREM 6.—*The sides of R_0 are circular arcs orthogonal to the principal circle. These arcs are congruent in pairs. Two congruent sides are equal in length and congruent points thereon are equidistant from the center of the principal circle.*

That the sides are orthogonal to the principal circle follows from Theorem 1. That the sides are arranged in congruent pairs which are equal in length follows from the fact (Sec. 21, Theorem 10) that the sides of R are so arranged. A side of R_0 is a side, or a portion of a side, of R, and is equal in length to its congruent side. This congruent arc lies also within the principal circle and is a side of R_0.

That congruent points on two congruent sides of R_0 are equidistant from the center of the principal circle is a consequence of Theorem 21, Sec. 11.

Attention should be called to the fact that an arc of the principal circle along which R_0 and R_0' are adjacent is not considered as a side of R_0. The term "side" is here limited to arcs of isometric circles.

THEOREM 7.—*Any closed region lying entirely within the principal circle is covered by a finite number of transforms of R_0. These regions fit together without lacunæ.*

This theorem follows from the fact that a region of the kind specified is covered without lacunæ by a finite number of transforms of R (Sec. 20, Theorem 8) and from the further fact that R_0 contains all the points of R whose transforms lie within the principal circle.

THEOREM 8.—*The transforms of R_0 fill up, without lacunæ, the whole interior of the principal circle. They cluster in infinite number about each limit point of the group.*

The first part of this theorem is a consequence of Theorem 7. A circle Q concentric with the principal circle and of smaller radius is covered by R_0 and a finite number of regions congruent to R_0. By taking the radius of Q near enough to that of the principal circle we can enclose in Q any given point interior to the principal circle. It follows that the interior of the principal circle is completely covered. The second part of the theorem follows from Theorem 9, Sec. 20.

THEOREM 9.—*An interior point of R_0 is nearer the center of the principal circle than any point congruent to it.*

An interior point of R_0 is outside the isometric circle of any transformation by which it is carried into a congruent point. The theorem then follows from Theorem 21, Sec. 11.

32. Generating Transformations.

THEOREM 10.—*The transformations by which the sides of R_0 are congruent form a set of generating transformations for the group.*

If the Fuchsian group is of the first kind, the theorem follows directly from Theorem 19, Sec. 28; for R_0 is that part of R lying in the interior of the principal circle, which may be taken as the region Σ of Sec. 28. If the group is of the second kind, Σ consists of all ordinary points in the plane, and the fundamental region of that theorem is R itself. But the sides of R which lie outside the principal circle are congruent by the same transformations that connect the sides of R_0 and so supply no new generating transformations.

Let Q be a circle lying within and concentric with the principal circle, and let us consider the network of regions by which Q is covered. We shall prove the following theorem:

THEOREM 11.—*A circle Q concentric with the principal circle and of smaller radius is completely covered by R_0 and by regions which are congruent to R_0 by transformations formed by combining those generating transformations which connect sides of R_0 lying wholly or in part in Q.*

The network of regions covering Q is found by adjoining regions congruent to R_0 along the sides of R_0 lying in Q, adjoining regions along such sides of these transformed regions as lie in Q, and so on (Sec. 23). In this process, we employ only those generating transformations connecting sides of R_0 which lie in Q or are congruent to sides lying in Q. For example, if one of these regions, $T(R_0)$, has a side l in Q which is congruent to l_k of R_0, then $TT_k(R_0)$ is the region adjacent to $T(R_0)$ along l. Hence, at each step we introduce only the generating transformations stated.

Now, when R_0 is carried into a congruent region, the distance of an interior point from the center of the principal circle is increased (Theorem 9); hence, the distance of a boundary point is not decreased. Then all transforms of all sides of R_0 which lie outside Q are themselves outside Q; and the transformations connecting these sides play no part in the formation of the network. This establishes the theorem.

Theorem 12.—*Let S be a transformation of the group; let P be the point on the isometric circle of S which is nearest the center of the principal circle; and let Q be the circle through P concentric with the principal circle. Then S can be expressed as a combination of those generating transformations which connect sides of R_0 which have at least one interior or end point within or on Q.*

The theorem is so worded as to include not only sides lying wholly or in part in Q but also sides which touch Q or have an end point lying on Q. Since there is but a finite number of sides in the neighborhood of Q we can draw a slightly larger circle Q', concentric with Q, which contains no new sides of R_0. Q' contains P on I_s and the congruent point P' on I_s', since P' and P are equidistant from the center of the principal circle.

It follows from the preceding theorem that Q' is completely covered by R_0 and the regions congruent to R_0 by transformations of the kind mentioned in the present theorem. These regions cover the neighborhoods of P and P'. S carries interior points of a region $T(R_0)$ in the neighborhood of P into interior points of some region $\bar{T}(R_0)$ in the neighborhood of P', where both T and $\bar{T}$ are combinations of generating transformations of the kind specified in the theorem. Then $ST(R_0)$ and $\bar{T}(R_0)$ overlap, whence (Theorem 1, Sec. 16) $ST = \bar{T}$; and $S = \bar{T}T^{-1}$. This is a combination of the kind required; and the theorem is established.

33. The Cycles.—Let $A_1, A_2, \ldots A_m$ be the vertices of an ordinary cycle of R, arranged in order as in Sec. 26. Then A_k and A_{k+1} are congruent points at the ends of congruent sides l_k, l_{-k} of R. We need consider only vertices of R_0, those of R_0' being got by an inversion. It follows from Theorem 6 that A_k and A_{k+1} are equidistant from the center of the principal circle.

The congruent points of a parabolic cycle are limit points and lie on the principal circle. We have, then, the following theorem:

Theorem 13.—*The congruent vertices of an ordinary cycle lie on a circle concentric with the principal circle. The points of a parabolic cycle lie on the principal circle.*

There arises the question whether there can be an ordinary cycle with vertices lying on the principal circle. Such cycles do exist for certain groups. Two isometric circles which meet on the principal circle, being both orthogonal to the principal circle, are tangent. Hence, the angle at each vertex is either 0 or π.

In Fig. 24 is shown the region R for a combination group arising from two hyperbolic cyclic groups. Two of the circles are made tangent. The order of the sides and vertices, according to the scheme of Sec. 26 is indicated. The congruent vertices are in order A_1, A_2, $A_3(=A_1)$, A_4.

This situation can arise only if the sum of the angles of the cycle is 2π. Otherwise, the transformation $S = T_m \cdot \cdot \cdot T_1$, which carries A_1 into itself, is an elliptic transformation and (Sec. 29 (*b*)) the fixed point A_1 does not lie on the principal circle. In the figure, the angles at A_2 and A_4 are each π and those at A_1 and A_3 are zero.

The cycle in this problem is not an essential one and we can remove it entirely if we wish. We can replace any part of R by a congruent part and still have a fundamental region. Let the sides l_{-1} and l_2 be slightly deformed in the neighborhood of A_2 so that A_2 becomes an interior point. The part of R which is removed in the neighborhood of A_2 can be replaced by the congruent part lying within I_1' in the neighborhood of A_1. Then, l_1 and l_{-2} are displaced slightly to the right and the sides no longer touch.

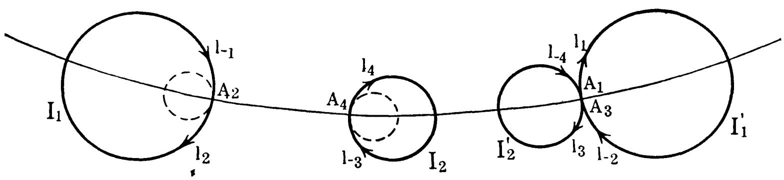


FIG. 24.

It is not difficult to see that this can be done in the most general case. In such a cycle, two of the vertices, as A_2 and A_4, have the angle π, and a finite number have the angle zero. By deforming the sides at one of the former vertices we can remove one of the points of tangency. By deforming the sides which now meet at the latter point we remove another point of tangency. By successive steps we can remove the remainder.

34. Fuchsian Groups of the First and Second Kinds.—Whether a Fuchsian group is of the first or the second kind (Sec. 30) depends upon the region R, as stated in the following theorem:

THEOREM 14.—*If R contains on its interior a point of the principal circle, the group is of the second kind; if not, it is of the first kind.*

The first part of the theorem is evident. If a point of the principal circle is on the interior of R, it is an ordinary point. Not all points of the principal circle are limit points; and the group is of the second kind.

To complete the proof we show that if the group is of the second kind, R contains a point of the principal circle on its interior. Let z_0 be an ordinary point on the principal circle. Then a circle Q can be drawn about z_0 such that all points within and on Q can be covered by a finite number of transforms of R (Theorem 8, Sec. 20), $R_1, \ldots, R_n$, say. Each of these regions has a finite number of sides lying in Q. For, an infinite number of sides would have a cluster point in or on Q; and this cluster point, being the transform of a point of R at which infinitely many isometric circles cluster, would itself be a limit point, which is impossible. Further, the sides of these regions, being congruent to the sides of R, are orthogonal to the principal circle. Let h be the arc of the principal circle lying in Q. Then, with the exception of the finite number of points where the sides of R_1, $R_2, \ldots R_n$ meet the principal circle, each point of h is interior

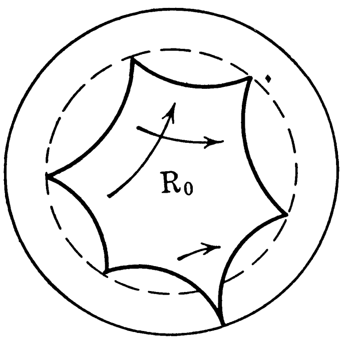


FIG. 25.

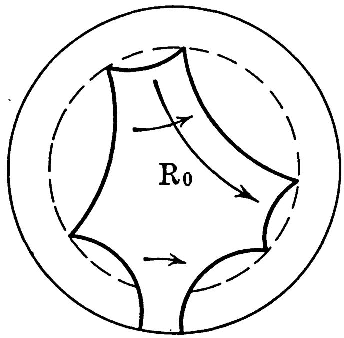


FIG. 26.

to one of the regions. Let z_1 be a point of h interior to R_k. Making the transformation which carries R_k into R, z_1 is carried into a point which is interior to R and which lies on the principal circle. The existence of this point establishes the theorem.

If R contains a point of the principal circle on its interior, it contains all points of an arc of the principal circle passing through the point. In the group of the second kind, then, R contains one or more arcs of the principal circle on its interior. The two regions R_0 and R_0' abut along these arcs.

The region R_0 for the two types of groups has the character illustrated in the accompanying figures. In the group of the first kind (Fig. 25) R_0 either lies, together with its boundary, within the principal circle; or, if there are points of the boundary of R_0 on the principal circle they are limit points of the group with sides of R_0 in the neighborhood. In the group of the second kind (Fig. 26) R_0 abuts on the principal circle along one or more arcs, and R contains these arcs on its interior.

THEOREM 15.—*R_0 constitutes a fundamental region for the Fuchsian group of the first kind.*

This theorem is a consequence of Theorem 18, Sec. 28.

THEOREM 16.—*If the boundary of R_0 lies within the principal circle, or if the only boundary points on the principal circle are parabolic points, then R_0 has a finite number of sides. The group is then generated by a finite number of transformations.*

If the only boundary points of R_0 which lie on the principal circle are parabolic points, the same is true of the boundary points of R_0'. The first part of the theorem then follows from Theorem 17, Sec. 27. The latter is then a consequence of Theorem 10.

In our later work, particularly in the study of uniformization, we shall come upon Fuchsian groups of the first kind where the fundamental region is found in quite other ways than that employed here. We shall give now some theorems concerning other fundamental regions.

THEOREM 17.—*If a Fuchsian group possesses a fundamental region F whose transforms cover the neighborhood of each of its boundary points and which lies within a circle Q concentric with the principal circle and of smaller radius, then R_0 lies within Q.*

A finite number of transforms of R_0 cover F completely. When we carry the portion of F in each region R_i into R_0 by means of the transformation which carries R_i into R_0 these transformed parts of F fill R_0 completely without lacunæ, as shown in the proof of Theorem 20, Sec. 28. Then, R_0 is in Q; for on transforming the parts of F into congruent parts in R_0 the distance of no point from the center of the principal circle is increased (Theorem 9).

THEOREM 18.—*The transforms of F fill up, without overlapping and without lacunæ, the whole interior of the principal circle.*

The transforms of F cover R_0 completely. The transforms of F, then, cover all transforms of R_0. These fill up the interior of the principal circle. There can be no overlapping of the transforms of F since no two points of F are congruent.

THEOREM 19.—*Of the fundamental regions lying within the principal circle, R_0 has the maximum area.*

A fundamental region different from R_0 is formed by replacing parts of R_0 by congruent parts. Since R_0 is exterior to all isometric circles, a shift of any part of R_0 to a congruent position effects a diminution of area.

35. Fixed Points at Infinity. Extension of the Method.—If a transformation has a fixed point at infinity, either there is no isometric circle or all circles are isometric. In the study of a group, we have supposed hitherto that it has been so transformed that there are no transformations with ∞ as a fixed point. This involves, of course, no loss of generality. There is, however, sometimes a distinct loss of simplicity in the definition of the group. In the present section, we shall discuss briefly the formation of the fundamental region for a properly discontinuous group certain of whose transformations have fixed points at infinity.

THEOREM 20.—*All the transformations of a given group which leave some configuration unaltered constitute a subgroup of the given group.*

For, the succession of two transformations which leave the configuration unaltered and also the inverse of any, leave the configuration unaltered, and they belong to the given group.

Taking the point at infinity as the configuration of the theorem, we have the following corollary:

COROLLARY.—*All those transformations of a given group which leave the point at infinity fixed constitute a subgroup of the given group.*

We shall represent by $U_0(=1)$, U_1, U_2, . . . the transformations with ∞ as fixed point and shall call the subgroup which they form Γ_u.

All transformations other than those of Γ_u possess isometric circles. It is no longer necessarily true that the radii of these circles are bounded, that their centers lie in a finite region, or that the centers are distinct. Concerning these isometric circles we shall prove the following theorem:

THEOREM 21.—*A transformation of the group Γ_u carries an isometric circle into an isometric circle.*

Let I_t be the isometric circle of T, and let a transformation U be applied. We shall show that $U(I_t)$ is the isometric circle of the transformation $S = UTU^{-1}$. U has the form

$$z' = U(z) = Kz + b,$$

where K is the multiplier, as we found in Secs. 6 and 10, Equations (31) and (39'). Since $U'(z) = K$, the transformation multiplies all lengths by $|K|$; the transformation U^{-1} multiplies all lengths by $1/|K|$.

Let P be a point on the circle $U(I_t)$ and let P be transformed by S. U^{-1} carries P to P', a point on I_t, lengths in the neighborhood being multiplied by $1/|K|$. T carries P' into P'' without alteration of lengths. U multiplies lengths in the neighborhood of P''

by $|K|$. The result is that UTU^{-1} has not altered lengths in the neighborhood of P; hence P is on the isometric circle of I_s.

Let us now construct a fundamental region F for the subgroup Γ_u. The existence of F follows (after a preliminary transformation of the group Γ_u) from the general theory of the preceding chapter—sides congruent in pairs, ordinary vertices arranged in cycles, etc. In many cases, however, as in Figs. 11 to 14, we know a fundamental region already, and we do not need to employ the general theory.

The principal theorem of the present section will now be proved.

THEOREM 22.—*Let F be a fundamental region for the subgroup Γ_u, where the sides of F are congruent in pairs and the transforms of F cover the finite plane,*[1] *and let R consist of all that part of F which is exterior to all isometric circles of the group. Then R is a fundamental region for the group.*

A transformation of Γ_u carries a point of R into a point of a region congruent to F and hence outside R. Any other transformation of the group carries a point of R into the interior of an isometric circle and hence outside R. So no two points of R are congruent.

On the assumption that the sides of F are congruent in pairs, the sides of R are congruent in pairs. That part of a side of F which is exterior to all isometric circles, and so is also a side of R, is, from Theorem 21, congruent to a side which is also exterior to all isometric circles. An ordinary boundary point P of R lying on an isometric circle I_t is carried by T into a point P' on I_t'. We can show in the usual way that P' is interior to no isometric circle. If P' lies in F, it is a boundary point of R; if not, it lies in a congruent region $U(F)$, and the congruent point $P'' = U^{-1}(P')$ lies in F and is a boundary point of R. It follows that the sides of R which lie on isometric circles are congruent in pairs.

A region about a boundary point of R on a side l_1 contains points congruent to points of R in the neighborhood of the congruent side l_{-1}; the neighborhood of a limit point contains points congruent to points of R by Theorem 2, Sec. 18. Hence, R is a fundamental region.

The further properties of R, such as the arrangement of the vertices in cycles, the theorems on the generating transformations,

[1] The fundamental regions of Figs. 11 to 14 have these properties. That the most general subgroup Γ_u has a fundamental region whose transforms cover the whole finite plane will appear in Chap. VI.

and the like, are established as in our previous treatment, and will not be repeated here.

36. Examples. *The Group of the Anharmonic Ratios.*—This group (Sec. 13 (2)) contains, in addition to the identical transformation, the following transformations:

$$U = 1 - z, \quad T_1 = \frac{1}{z}, \quad T_2 = \frac{1}{1-z},$$
$$T_3 = \frac{z-1}{z}, \quad T_4 = \frac{z}{z-1}.$$

Here U is a rotation through the angle π about the point $\frac{1}{2}$. A fundamental region for the subgroup Γ_u is the half plane bounded by any straight line through $\frac{1}{2}$; for instance, the upper half plane.

The isometric circles of the remaining transformations are

$$I_1, I_3\colon\ |z| = 1; \qquad I_2, I_4\colon\ |z-1| = 1.$$

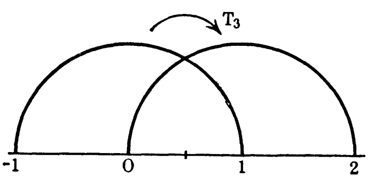


FIG. 27.

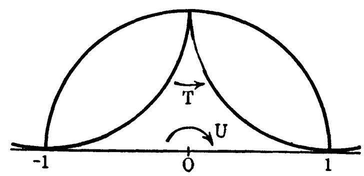


FIG. 28.

We have, then, as a fundamental region R for the group that part of the upper half plane lying outside these circles (Fig. 27).

The rectilinear sides of R are congruent by U, the circular sides by T_3. U and T_3 are then generating transformations. From the two cycles of angles $2\pi/3$ and π, we have the relations $T_3^3 = 1$ and $(UT_3)^2 = 1$.

A Group with Q_0 as Principal Circle.—As a further example, we shall consider the group mentioned in Sec. 13 (7):

$$z' = \frac{az + \bar{c}}{cz + \bar{a}}, \qquad a\bar{a} - c\bar{c} = 1,$$

where a and c are complex integers. This is a Fuchsian group with Q_0 as principal circle.

When $c = 0$, we have, $a\bar{a} = 1$; whence, $a = \pm 1$ or $\pm i$. This gives for Γ_u the two transformations,

$$z' = z, \qquad z' = U(z) = \frac{iz}{-i} = -z.$$

U is a rotation about the origin through the angle π. A fundamental region for Γ_u is the half plane above the real axis.

Among the remaining transformations, consider, first, those with the largest isometric circles. The smallest absolute value of c occurs when $c = \pm 1$ or $\pm i$. Then $a\bar{a} = 1 + c\bar{c} = 2$; and $a = \pm 1 \pm i$. This gives for the center $-\bar{a}/c$, the possible positions $1 + i$, $1 - i$, $-1 + i$, $-1 - i$. Only two circles with these centers and of unit radius lie in the upper half plane. They are the isometric circles of

$$T = \frac{(1+i)z + 1}{z + 1 - i}$$

and its inverse. They are shown in Fig. 28.

We shall now show that each other isometric circle in the upper half plane is contained in one of these. Each such circle is orthogonal to Q_0 and of radius less than 1. Unless it lies within or is tangent internally to one of the two circles already drawn, it must contain one of the points ± 1 or i on its interior. Its isometric circle is

$$|cz + \bar{a}| = 1, \qquad |c| > 1.$$

If one of the three points lies within this circle, we have

$$|\pm c + \bar{a}| < 1, \quad \text{or} \quad |ic + \bar{a}| < 1.$$

Since the term whose absolute value appears in the first member is an integer, its absolute value can be less than 1 only if it is zero. Then $|c| = |a|$, and $a\bar{a} - c\bar{c} = 0$, which is impossible.

R, then, is that part of the upper half plane which is exterior to the two circles of Fig. 28. R contains no points of Q_0; so the group is of the first kind. All points of Q_0 are limit points. R consists of two parts: R_0 within Q_0, and R_0', the inverse of R_0 in Q_0. The transforms of R_0 fill up the whole interior of Q_0. U and T are generating transformations.

There are three cycles. The origin constitutes an ordinary cycle of angle π, whence $U^2 = 1$. The point i constitutes a parabolic cycle. The points 1 and -1 form a second parabolic cycle.

37. The Modular Group.—This group (Sec. 13 (5)) consists of the transformations

$$z' = \frac{az + b}{cz + d}, \qquad ad - bc = 1,$$

where a, b, c, d, are real integers. The real axis is a fixed circle. Whether the upper half plane is carried into itself or into the lower half plane can be determined from a consideration of one point. When $z = i$, we find $z' = [ac + bd + (ad - bc)i]/(c^2 + d^2)$. The imaginary part of this quantity is positive; so z' is also in the upper half plane. The group is thus a Fuchsian group.

When $c = 0$, we have $ad = 1$. Then $a = d = \pm 1$, and b can have any integral value. This gives for the subgroup Γ_u (Sec. 35) the set of transformations

$$z' = z + n,$$

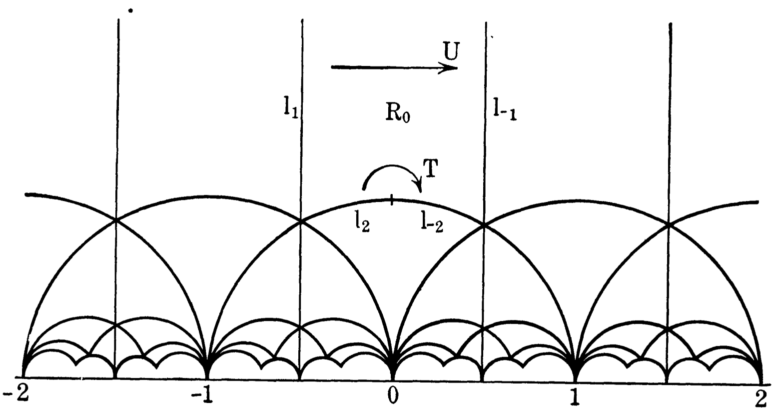


Fig. 29.

where n is any integer. This is the group of translations whose fundamental region—setting $\omega = 1$—is shown in Fig. 12. We shall take as fundamental region the strip enclosed by lines perpendicular to the real axis through the points $\pm\frac{1}{2}$. The subgroup is generated by the transformation

$$z' = U(z) = z + 1.$$

Consider the largest isometric circles. If $c = \pm 1$, we have $ad \mp b = 1$. For any integral value of a and d, we can determine from this equation an integral value of b. It follows that the center, $\mp d$, of the isometric circle $|\pm z + d| = 1$ can be any integer. These circles, then, are the unit circles with centers at the real integers. They are the large circles in Fig. 29.

These isometric circles enclose all points within a distance of $\frac{1}{2}\sqrt{3}$ of the real axis. For any other transformation $|c| \geqslant 2$; and the isometric circle is of radius not exceeding $\frac{1}{2}$. As the center is on the real axis, such a circle lies in the space enclosed by the unit isometric circles. These smaller circles, then, can form no part of the boundary of R.

The region R, lying in the period strip previously selected and exterior to the isometric circles, is bounded by the unit circle with center at the origin. It consists of the region R_0 shown in the figure together with the reflection of R_0 in the real axis. The circular boundary of R_0 consists of two sides l_2 and l_{-2} which are congruent by

$$z' = T(z) = -\frac{1}{z},$$

an elliptic transformation with fixed points $\pm i$. U and T are generating transformations for the group.

There are three cycles. The point at infinity, to which R_0 extends, is a parabolic point and constitutes a cycle. The point i constitutes a cycle of angle π; whence $T^2 = 1$. The remaining two vertices, namely, $\pm\frac{1}{2} + \frac{1}{2}\sqrt{3}i$, constitute a cycle of angle $2\pi/3$.

To get the relation connecting U and T which arises from the last cycle, we proceed as in Sec. 26. Starting from the right-hand vertex and the side l_{-1} beginning there, we get the transformation U^{-1} then T before returning. Then,

$$S = TU^{-1} = -\frac{1}{z - 1}$$

is a transforms of period three which carries the right-hand vertex into itself. The desired relation is $S^3 = 1$ or $TU^{-1}TU^{-1}$ $TU^{-1} = 1$. This can also be written, if we take the inverse and use the fact that $T^{-1} = T$, in the form

$$UTUTUT = 1.$$

The transforms of R_0 cover the whole upper half plane. A number of the regions congruent to R_0 are shown drawn to scale in the figure. The regions cluster in infinite number about each point of the real axis.

38. Some Subgroups of the Modular Group.—As a further exemplification of the method of forming the fundamental region, we shall now consider a particular set of the great variety of subgroups contained within he modular group. Let n be an integer greater than 1, and consider all those transformations of the modular group for which b and c are divisible by n. We then have all transformations of the form

$$T = \frac{az + nb'}{nc'z + d'}, \qquad ad - n^2b'c' = 1,$$

where a, b', c', d are real integers.

We prove first that these transformations form a group. The inverse, $T^{-1} = (-dz + nb')/(nc'z - a)$ is of the same form. Let $S = (\alpha z + n\beta')/(n\gamma'z + \delta)$ be a second transformation of the set, then

$$ST = \frac{(\alpha a + n^2\beta'c')z + n(\alpha b' + \beta'd)}{n(\gamma'a + \delta c')z + n^2\gamma'b' + \delta d}.$$

Since $n(\alpha b' + \beta'd)$ and $n(\gamma'a + \delta c')$ are divisible by n, this belongs to the set. Thus both group properties are satisfied.

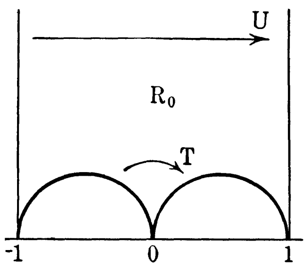


FIG. 30.

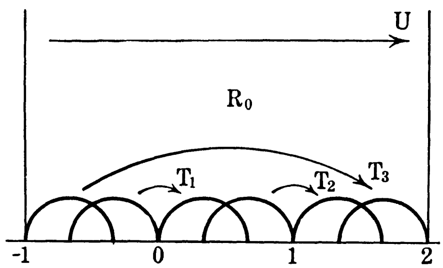


FIG. 31.

Setting $c' = 0$, we get for the group Γ_u the set of transformations

$$U_m = z + mn,$$

where m is any integer. Γ_u is generated by $U = z + n$.

In the accompanying figures, fundamental regions for two cases are shown. The reasoning follows the lines of the preceding sections and is left to the reader.

Figure 30 shows R_0 for the case $n = 2$. The group is generated by the two transformations

$$U = z + 2, \qquad T = \frac{z}{2z + 1}.$$

There are three parabolic cycles.

Figure 31 is for the case $n = 3$. The generating transformations are

$$U = z + 3, \qquad T_1 = \frac{z}{3z + 1}, \qquad T_2 = \frac{4z - 3}{3z - 2}, \qquad T_3 = \frac{5z + 3}{3z + 2}.$$

There are four parabolic cycles and one ordinary cycle of angle 2π.

CHAPTER IV

AUTOMORPHIC FUNCTIONS

39. The Concept of the Automorphic Function.—Automorphic functions are the generalization of the circular, hyperbolic, elliptic, and certain other functions of elementary analysis. A circular function, as $\sin z$, has the property that it is unchanged in value if z is replaced by $z + 2m\pi$, where m is any integer; that is, the function is unaltered in value if z be subjected to a transformation of the group $z' = z + 2m\pi$. A hyperbolic function, such as $\sinh z$, is unchanged in value if z be subjected to a transformation of the group $z' = z + 2m\pi i$. An elliptic function, as the Weierstrassian function $\wp(z)$, retains its value under transformations of a group of the form $z' = z + m\omega + m'\omega'$.

The automorphic function is an extension of this concept to the more general properly discontinuous group. Roughly speaking, a function is automorphic with respect to such a group if it has the same value at congruent points. We shall lay down a more precise definition.

By the *domain of existence* of a single-valued analytic function $f(z)$ we shall mean the set of points at which $f(z)$ is analytic or has poles. The domain of existence is a connected region consisting entirely of interior points,—a two dimensional continuum.

DEFINITION.—*A function $f(z)$ will be said to be automorphic with respect to a group of linear transformations T_1, T_2 . . . provided*

1. *$f(z)$ is a single-valued analytic function.*
2. *If z lies in the domain of existence of $f(z)$ so also shall $T_n(z)$.*
3. $f[T_n(z)] \equiv f(z)$.

Because of the first condition, the functions here defined might more properly be called "single-valued" automorphic functions. There exist many-valued analytic functions satisfying conditions (2) and (3). But we shall be concerned altogether with single-valued functions; and to avoid repeatedly calling attention to this fact we shall make single-valuedness a part of the definition.

We note that there are no functions, other than constants, which are automorphic, according to this definition, with respect to a continuous or an improperly discontinuous group. For, let $F(z)$ be such a function and let z_0 be a point at which the function is analytic. There are infinitely many points in the neighborhood of z_0 which are congruent to z_0. At each of these points $F(z) = F(z_0)$. It is a well-known fact that a function can take on the same value at an infinite number of points in the neighborhood of a point at which it is analytic only if it is constant.

It is observed from the definition, that if $f(z)$ is automorphic with respect to a group it is automorphic with respect to any subgroup.

In showing that a function is automorphic with respect to a group, it is not necessary to investigate conditions (2) and (3) for all points of the domain of existence of the function nor for all transformations of the group. As the following theorems show, it suffices to establish the conditions for some small region and for the generating transformations of the group.

THEOREM 1.—*Let $f(z)$ be a single-valued function analytic at z_0. Let $T(z_0)$, where T is a linear transformation, lie in the domain of existence of the function; and let*

$$f[T(z)] \equiv f(z) \tag{1}$$

be valid in the neighborhood of z_0. Then, if z is any point in the domain of existence of the function, so also is $T(z)$, and (1) *holds throughout the whole domain of existence.*

The transformation T carries the domain of existence of $f(z)$ into itself.

This theorem is an immediate consequence of the principle of analytic continuation. The two functions of z that appear in (1) are identical in a region surrounding z_0; they are, therefore, identical in any region to which either can be continued analytically. Let z_1 be a point at which $f(z)$ is analytic. Then $f(z)$ can be continued analytically from z_0 throughout a suitable region S surrounding z_1. Then in S $f(T(z))$ is analytic and (1) holds. That is, $f(z)$ is analytic in the neighborhood of $z_1' = T(z_1)$ and $f(z_1') = f(z_1)$.

In the neighborhood of a pole z_2, $f(z)$ is analytic and $f(z)$ becomes infinite as z approaches z_2. Then $f(z)$ is analytic in the neighborhood of $z_2' = T(z_2)$ and becomes infinite as z_2' is approached.

T carries any point of the domain of existence of $f(z)$ into another such point. The same is true of T^{-1}; since $f[T^{-1}(z)] \equiv f(z)$ holds in the neighborhood of $z_0' = T(z_0)$. Then T carries no point without the domain of existence into the domain. Hence, T carries the domain of existence into itself.

THEOREM 2.—*If $f(z)$ is a single-valued analytic function and if*

$$f[T_1(z)] \equiv f(z), \qquad f[T_2(z)] \equiv f(z), \quad \cdot \cdot \cdot ,$$

then $f(z)$ is automorphic with respect to the group generated by $T_1, T_2, \ldots$

Each transformation of the group carries the domain of existence of $f(z)$ into itself.

The group is formed by constructing all possible products by means of $T_1, T_2, \ldots$ and their inverses. It is clear that $f(z)$ is unaltered when z is replaced by $T_1^{-1}(z)$, $T_2^{-1}(z)$, . . . Since any product can be formed by repeatedly combining transformations two at a time it suffices, to prove the first part of the theorem, to point out that if $f[S(z)] \equiv f(z)$, $f[T(z)] \equiv f(z)$, then $f[ST(z)] \equiv f(z)$. But if z is in the domain of existence of the function so is $T(z)$, and hence $ST(z)$, and we have $f[ST(z)] \equiv f[T(z)] \equiv f(z)$.

The latter part of the theorem is an application of the latter part of Theorem 1.

Consider, as an example, the function $\cos z$. We have $\cos (z + 2\pi) \equiv \cos z$ and $\cos(-z) \equiv \cos z$. Then $\cos z$ is automorphic with respect to the group generated by $z' = z + 2\pi$ and $z' = -z$.

The existence of a non-constant single-valued analytic function which is unaltered when a set of linear transformations is applied to the independent variable is sufficient to show that the group generated by the transformations is properly discontinuous. Thus, the group mentioned in the preceding paragraph is necessarily properly discontinuous.

THEOREM 3.—*The domain of existence of an automorphic function extends into the neighborhood of every limit point of the group.*

For, in the neighborhood of a limit point lie points congruent to points in the domain of existence of the function. These points belong to the domain of existence of the function.

THEOREM 4.—*If the automorphic function is not a constant, each limit point of the group is an essential singularity of the function.*

In the neighborhood of a point at which a function is analytic or has a pole, the function can take on any value only a finite number of times. In the neighborhood of a limit point there is an infinite number of congruent points at which the function takes on the same value. Hence, the limit point is an essential singularity.

COROLLARY.—*All points of the domain of existence of a non-constant automorphic function are ordinary points.*

It is not true, of course, that all points on the boundary of the domain of existence of the function are necessarily limit points. The limit points lie on the boundary of the domain, but there may be further boundary points. For example, the function $e^{\wp(z)}$ is automorphic with respect to the group $z' = z + m\omega + m'\omega'$. Its domain of existence consists of all points except the point ∞ (the only limit point) and the points $m\omega + m'\omega'$. At these latter points $\wp(z)$ has poles and the function has essential singularities.

It results from the preceding discussion that not all properly discontinuous groups have non-constant automorphic functions.

THEOREM 5.—*If a group possesses a non-constant automorphic function, it is a function group.*

Let a group have a non-constant automorphic function existing in a domain S. Let Σ be the part of the plane, bounded by limit points, in which S lies. Σ consists of all ordinary points which can be joined to a point of S by curves not passing through limit points. Any point z of Σ and a curve C joining it to a point of S are carried by any transformation T of the group into a point z' and a curve C' joining z' to a point of S, where C' consists of ordinary points. Then z' belongs to Σ; whence Σ is carried into itself. The group is, therefore, a function group.

We shall find in the following chapter that every function group possesses non-constant automorphic functions.

40. Simple Automorphic Functions.—In the present chapter, and in our later work, we shall have much to do with automorphic functions of a somewhat restricted kind. We shall impose restrictions both on the character of the function and on the group with respect to which it is automorphic.

Let $f(z)$ have no essential singularity at an ordinary point of the group. Then the domain of existence of the function, provided it is not a constant, is one of the regions Σ into which the limit points of the group separate the plane.

That part R_0 of R which lies in Σ is a fundamental region for the group. We shall require that R_0 have a finite number of sides.

If R_0 possesses one or more parabolic points we shall impose a further condition on the function. As z approaches a parabolic point P from the region R_0 let the function approach a definite value, finite or infinite; that is,

$$\lim_{z=P} f(z) = C, \text{ or } \infty,$$

where z is restricted, in its approach, to lie within or on the boundary of R_0.[1]

To avoid long circumlocutions in the statement of theorems we shall call such a function a "simple automorphic function." A simple automorphic function then (1) belongs to a function group such that R_0 has a finite number of sides; (2) has the domain of existence Σ, provided it is non-constant; and (3) behaves in the manner specified at the parabolic points, if any.

If the group is finite, Σ consists of the whole plane. The simple automorphic function, then, has no other singularities than poles and is, therefore, a rational function.

If the group is Fuchsian, the simple automorphic function is called a "Fuchsian function." If the group is of the first kind, the domain of existence of the function, if not constant, is the interior or the exterior of the principal circle. If the group is of the second kind, the domain of existence consists of the whole plane exclusion of the limit points lying on the principal circle.

If the group is Kleinian, the simple automorphic function is called a "Kleinian function."

Extension of the Definition.—If $f(z)$ is a simple automorphic function belonging to a group T_n, we shall define

$$\psi(z) \equiv f[S(z)],$$

where S is a linear transformation to be a simple automorphic function belonging to the transformed group $S^{-1}T_nS$. It is clearly automorphic; for

$$\psi[S^{-1}T_nS(z)] \equiv f[SS^{-1}T_nS(z)] \equiv f[T_nS(z)] \equiv f[S(z)] \equiv \psi(z).$$

Here $\psi(z)$ has the domain of existence $S^{-1}(\Sigma)$, where Σ is the domain of existence of $f(z)$; and $S^{-1}(R_0)$ is a fundamental region with a finite number of sides for the transformed group.

[1] If P counts as two parabolic points of the region, as in Fig. 19, the approach shall be from one side only. There shall be a limit when the approach is from either side, but the two limits may be different.

By this means we extend the definition to groups with fixed points or limit points at infinity. Such a group can be represented in infinitely many ways as the transform of a group with infinity as a non-fixed ordinary point. The propositions derived in this chapter—on zeros, poles, algebraic relations, etc.—hold for the fundamental region $S^{-1}(R_0)$ when the group is transformed.

Also, for many of the commoner groups we know already fundamental regions, not based on isometric circles, to which the proofs of the following theorems apply (Figs. 11 to 14, for example). In the general case, however, we fall back on the region R_0, for the reason that the properties of R have been established with complete generality.

A familiar example of a simple automorphic function is the Weierstrassian $\wp$-function. Here the domain of existence is the finite plane, the period parallelogram is a fundamental region, and there are no parabolic points.

Likewise $\sin z$ is a simple automorphic function. The period strip is a fundamental region with parabolic points at the ends. We find readily that $\sin z$ approaches ∞ as z approaches either end of the strip.

41. Behavior at Vertices and Parabolic Points.—Of the fixed points of the transformations of a group, only those belonging to elliptic transformations can lie within the domain in which the automorphic function is analytic or has poles; all other fixed points are limit points. At a fixed point of an elliptic transformation, the function must behave in a particular way.

THEOREM 6.—*A non-constant automorphic function takes on its value k times, or some multiple thereof, at a fixed point of an elliptic transformation of period k within the domain of existence of the function.*

A function $f(z)$ is said to take on its value s times at a point z_0 at which it is analytic if it can be written, in the neighborhood of the point, in the form $f(z) = f(z_0) + (z - z_0)^s \varphi(z)$ where $\varphi(z)$ is analytic at z_0 and $\varphi(z_0) \neq 0$. It takes on the value infinity s times if $1/f(z)$ takes on the value zero s times. If $z_0 = \infty$, $z - z_0$ is replaced by $1/z$.

Let z_0 be the fixed point of an elliptic transformation of period k. We shall suppose that z_0 is finite. Then there is a transformation S of the form

$$\frac{z' - z_0}{z' - z_0'} = e^{2\pi i/k} \frac{z - z_0}{z - z_0'},$$

where z_0', the second fixed point, is distinct from z_0. Let $f(z)$ be analytic at z_0; then we can write, for z in the neighborhood of z_0,

$$f(z) = f(z_0) + (z - z_0)^s\varphi(z) = f(z_0) + \left(\frac{z - z_0}{z - z_0'}\right)^s \psi(z),$$

where $\psi(z) = (z - z_0')^s\varphi(z)$; so $\psi(z)$ is analytic at z_0 and $\psi(z_0) \neq 0$. If z is in the neighborhood of z_0 so also is $z' = S(z)$; so

$$f(z') = f(z_0) + \left(\frac{z' - z_0}{z' - z_0'}\right)^s \psi(z') = f(z_0) + \left(\frac{z - z_0}{z - z_0'}\right)^s e^{2\pi is/k}\psi(z').$$

Since the function is automorphic, $f(z') = f(z)$, whence

$$e^{2\pi is/k} = \frac{\psi(z)}{\psi(z')}.$$

The first member of this equation is constant; so, also, is the second. Letting z approach z_0, z' also approaches z_0, and we have

$$e^{2\pi is/k} = 1.$$

It follows that s, which is a positive integer, is a multiple of k, which was to be proved.

If $f(z)$ has a pole at z_0, $1/f(z)$ is an automorphic function with a zero at z_0. The order of the zero is a multiple of k; hence, $f(z)$ has a pole whose order is a multiple of k.

The proof for the case that $z_0 = \infty$ is not essentially different.

THEOREM 7.—*A non-constant automorphic function takes on its value k times, or some multiple thereof, at a vertex belonging to a cycle the sum of whose angles is $2\pi/k$.*

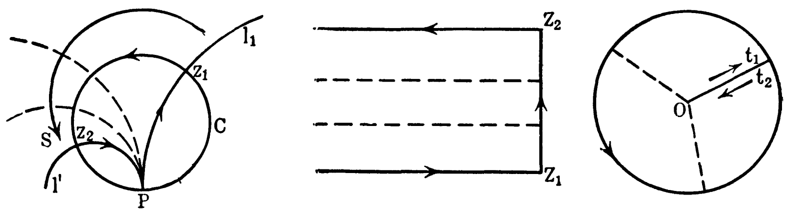
This is a corollary of the preceding theorem. If $k > 1$, the vertex is a fixed point of an elliptic transformation of period k (Sec. 26, Theorem 14); if $k = 1$, the proposition is trivial.

Consider now the behaviour of a simple automorphic function at a parabolic point P. We can carry the congruent points of the cycle to P, the transforms of the fundamental region fitting together at P as in Fig. 23. Then $f(z)$ approaches a definite finite or infinite value as z approaches P from within the regions of that figure. For, $f(z)$ approaches a limit as z approaches P within or on the boundary of each region; and, owing to the common boundaries, the limits are equal. In other words, $f(z)$ approaches the same value at all the points of the parabolic cycle.

As pointed out in Sec. 27, the transformation S which carries R of Fig. 23 into R_1, and the side l_1 into l', is parabolic. S has the form (Sec. 10, Equation 37)

$$\frac{1}{z' - P} = \frac{1}{z - P} + c.$$

We shall investigate the function in the triangular region z_1z_2P formed by l_1, l', and a small circle C through P orthogonal to the sides which meet at P (Fig. 32(a)). C is a fixed circle for S. By repeated applications of S, the transforms of the region mentioned fill up the circle C; and the values of $f(z)$ repeat themselves in the transformed regions.



(a) (b) (c)

FIG. 32.

We make the change of variable

$$t = e^{2\pi i/c(z-P)}$$

thus mapping the region under consideration on the t-plane. We write this in the form

$$Z = \frac{2\pi i}{c(z - P)}, \quad t = e^Z.$$

The first transformation is linear. It carries P into ∞ and the circles into straight lines. Let Z_1 and Z_2 be the transforms of z_1 and z_2. We have

$$Z_1 = \frac{2\pi i}{c(z_1 - P)},$$

$$Z_2 = \frac{2\pi i}{c(z_2 - P)} = \frac{2\pi i}{c}\left[\frac{1}{z_1 - P} + c\right] = Z_1 + 2\pi i.$$

The arc z_1z_2 of C is transformed into a straight line parallel to the imaginary axis (Fig. 32(b)). The sides Pz_1 and z_2P are carried into straight lines perpendicular to Z_1Z_2, and, hence, parallel to the real axis. The triangle of (a) is mapped on the region bounded by the three lines of (b). Congruent points of l_1 and l' are carried into points in (b) which differ by $2\pi i$.

These latter points are carried into coincident points by the transformation $t = e^Z$. The side Z_1Z_2 is carried into a circle with center at the origin. The original circular arc triangle is thus mapped on a circle slit along a radius, as in (c).

The function $f(z)$ becomes a function of t, $\varphi(t)$, analytic except for poles within and on the boundary of the region in (c) except possibly at 0. $\varphi(t)$ takes on the same value at a point on the radius along which the region is slit when approached from either side. This slit can be removed, the function being single-valued. Since $\varphi(t)$ approaches a finite value at 0 or becomes infinite, either it is analytic at 0, if properly defined there, or it has a pole at 0.

We have established the following result:

THEOREM 8.—*At a parabolic point a simple automorphic function is a function of t analytic or having a pole at* $t = 0$, *where*

$$t = e^{2\pi i/c(z-P)}.$$

If $P = \infty$, S has the form $z' = z + c$ and we put $t = e^{2\pi iz/c}$ in Theorem 8.

In the neighborhood of 0, Fig. 29 (c), we can write

$$f(z) = \varphi(t) = a_0 + a_1t + a_2t^2 + \cdots,$$

or

$$f(z) = t^{-n}(a_0 + a_1t + \cdots),$$

according as $f(z)$ approaches a finite value or becomes infinite as z approaches P. This expansion is valid in a circle with 0 as center and passing through the nearest singularity of the function in the t-plane. Carrying this back to the z-plane, the expansion is valid within the circle C through the nearest pole or limit point of $f(z)$.

42. The Poles and Zeros.—In counting the poles and zeros of a simple automorphic function which lie in the fundamental region, certain conventions are necessary in the cases of poles or zeros lying on the boundary. (1) If there is a pole or zero on the side l_n there is a pole or zero at the congruent point on the side l_{-n}. Only one of these shall be counted as belonging to the region. (2) The order s of a pole or zero at a vertex shall be partitioned equally among the regions which meet there. If there are m vertices in the cycle and the sum of the angles at the vertices is $2\pi/k$, then km regions meet at each vertex. Counting

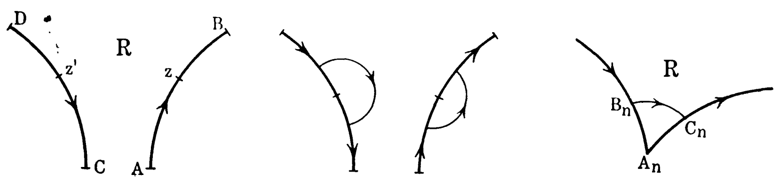
s/km poles or zeros at each vertex, we have s/k poles or zeros at all the m vertices of the cycle. This number (Theorem 7) is an integer. (3) If $f(z)$ becomes infinite or approaches zero at a parabolic point P we shall determine the number of poles or zeros from the behavior of $\varphi(t)$ at the origin. The number of poles or zeros of $\varphi(t)$ at $t = 0$ shall be the number of poles or zeros of $f(z)$ in the parabolic points, taken all together, of the cycle to which P belongs.

The number of times $f(z)$ takes on any other value C shall be the number of times $f(z) - C$ takes on the value zero.

THEOREM 9.—*A simple automorphic function which is not identically zero has an equal number of zeros and poles in the fundamental region.*

Suppose, first, that the function has neither poles nor zeros on the boundary. Then,

$$N - M = \frac{1}{2\pi i}\int d \log f(z),$$



(a) (b) (c)

FIG. 33.

where N is the number of zeros in the fundamental region, and M the number of poles, the integral being taken in a positive sense around the boundary. (If the region is multiply connected or disconnected, we integrate, of course, around the complete boundary.)

Consider the parts of the integral arising from two congruent sides AB and CD (Fig. 33(*a*)): $\int_A^B d \log f(z) + \int_D^C d \log f(z)$. At the congruent points z and z', we have $f(z') = f(z)$; then, log $f(z')$ and log $f(z)$ differ at most by a multiple of $2\pi i$, and d log $f(z) = d$ log $f(z')$. The second integral can then be written $\int_B^A d \log f(z)$; and the two integrals cancel. The integrals along each pair of congruent sides cancel, and we have $N - M = 0$, or $N = M$; which was to be proved.

If there is a zero or pole on the side AB, we deform the side slightly, as in Fig. 33(*b*), so as to include the zero or pole, and we make the corresponding alteration in the side CD. The integrals along the new congruent sides cancel as before. Only one of the two zeros or poles now lies within the contour. The theorem holds as before, since but one of the pair should be counted as belonging to the region.

Let $f(z)$ have a zero of order s at a vertex. We alter the path of integration to exclude each vertex of the cycle as in Fig. 33(c), the points B_n, C_n being at a distance d from A_n. The parts of the sides that remain are congruent in pairs; and the integrals over these sides cancel as before. At A_n we have $f(z) = (z - A_n)^s\varphi(z)$, where $\varphi(z)$ is analytic at A_n and does not vanish there; and

$$\int_{B_n}^{C_n} d\log f(z) = s\int_{B_n}^{C_n} d\log(z - A_n) + \int_{B_n}^{C_n}\frac{\varphi'(z)dz}{\varphi(z)}.$$

Letting d approach zero, the last integral approaches zero, since the integrand remains finite and the length of the path of integration approaches zero. The first integral of the second member approaches $s(-i\angle A)$. Summing for all the vertices of the cycle, we have

$$N - M = -\frac{s}{2\pi}\sum\angle A_n.$$

If the sum of the angles of the cycle is $2\pi/k$, we have

$$N - M = -\frac{s}{k}, \text{ or } N + \frac{s}{k} = M.$$

Here, N is the number of zeros within the contour. As s/k is precisely the number of zeros which we are to count at the vertices of the cycle, the number of zeros is equal to the number of poles.

A pole at a vertex is treated similarly, s being replaced by $-s$.

Finally, let $f(z)$ have a pole or zero at a parabolic point P. Draw a circle C through P, as in Fig. 32, sufficiently small that there are no poles or zeros other than at P within the region z_1z_2P of that figure. The arc z_1z_2 cuts off from the regions that lie in z_1z_2P certain parts $A_1, A_2, A_3, \ldots, A_m$. The congruent parts $A_1', A_2', \ldots, A_m'$ lying in the fundamental region make up the neighborhoods of the parabolic points of the cycle. We shall remove these parts from the fundamental region and integrate around the contour of the remainder. The integrals over pairs of congruent sides cancel. The integrals over the circular arcs cutting off the parabolic points may be replaced by the integrals over the congruent arcs on C:

$$N - M = \frac{1}{2\pi i}\int_{z_2}^{z_1} d \log f(z).$$

This last integral may be replaced by $1/2\pi i \int d \log \varphi(t)$ taken clockwise around the circle in Fig. 32(c). This has the value $-s$ if $\varphi(t)$ has a zero of order s at $t = 0$, and the value p if $\varphi(t)$ has a pole of order p there. Hence we have $N + s = M$ or $N = M + p$; and the theorem holds.

By combinations of the preceding methods of contour integration we dispose of all cases in which there is a finite number of poles and zeros on the boundary of the fundamental region. Now, $f(z)$ cannot have an infinite number of poles for the poles would then have a cluster point z'. If z' is an ordinary point, $f(z)$ has an essential singularity there; if z' is a parabolic point, $\varphi(t)$ has an essential singularity at the origin, both of which are contrary to hypothesis. Similarly $f(z)$ cannot have an infinite number of zeros unless it is identically zero, for a cluster point of zeros would be, likewise, an essential singularity. Hence, the theorem is established.

THEOREM 10.—*A simple automorphic function which has no poles in the fundamental region is a constant.*

Let $f(z)$ be a simple automorphic function having no poles in the fundamental region; and let its value at z_0, a point of the region, be C. Then $f(z) - C$ is a simple automorphic function with a zero at z_0 and having no poles. This is possible, according to Theorem 9, only if $f(z) - C \equiv 0$, or $f(z) \equiv C$.

THEOREM 11.—*A simple automorphic function which is not a constant takes on every value the same number of times in the fundamental region.*

If a simple automorphic function $f(z)$ is not constant, it has a certain finite number of poles in the fundamental region. The function $f(z) - C$, where C is any constant, is a simple automorphic function with the same poles as $f(z)$. The number of zeros is equal to the number of poles. That is, the number of times $f(z)$ takes on the value C is equal to the number of poles of $f(z)$, which establishes the theorem.

43. Algebraic Relations.—As a consequence of the preceding results, the following important theorem can now be established.

THEOREM 12.—*Between two simple automorphic functions belonging to the same group and having the same domain of existence, there exists an algebraic relation.*

Let $f_1(z)$ and $f_2(z)$ be two such simple automorphic functions, with k_1 and k_2 poles, respectively, in the fundamental region. We are to show that there exists a relation of the form

$$\Phi(f_1, f_2) \equiv A_1 f_1^m f_2^n + A_2 f_1^m f_2^{n-1} + \cdots + A_{(m+1)(n+1)} = 0, \quad (2)$$

where $A_1, A_2, \ldots$ are constants, the relation holding for all values of z in the domain of definition of the functions. Whatever values be given to the constants, the function Φ is a simple automorphic function. The degrees of Φ in f_1 and f_2, namely, m and n, respectively, will be determined later. The number of poles of Φ is not greater than $mk_1 + nk_2$.

The most general polynomial of degree m in f_1 and n in f_2 contains $(m + 1)(n + 1)$ constants. We can so choose these constants that Φ shall have zeros at $(m + 1)(n + 1) - 1$ assigned points in the fundamental region. For, let $c_1, c_2, \ldots, c_{(m+1)(n+1)-1}$ be distinct points of the region different from the poles of $f_1(z)$ and $f_2(z)$; and let $A_1, A_2, \ldots$ be determined to satisfy the equations

$$A_1 f_1^m(c_i) f_2^n(c_i) + \cdots + A_{(m+1)(n+1)} = 0, \quad (3)$$
$$i = 1, 2, \cdots, (m + 1)(n + 1) - 1.$$

Constants not all zero can always be found to satisfy these equations, since there is one more constant than equations to be satisfied. With these values of $A_1, A_2, \ldots, \Phi$ has zeros at the points $c_1, c_2, \ldots$

The function then has at least $(m + 1)(n + 1) - 1$ zeros and not more than $mk_1 + nk_2$ poles. Now, if m and n be large enough

$$(m + 1)(n + 1) - 1 > mk_1 + nk_2, \quad (4)$$

and Φ has more known zeros than possible poles. According to Theorem 9, Φ is identically zero. This establishes the theorem.

It will generally happen, if the algebraic relation be found in this manner, that Φ is reducible:

$$\Phi(f_1, f_2) \equiv \Phi_1(f_1, f_2)\Phi_2(f_1, f_2) \cdots \Phi_r(f_1, f_2),$$

where Φ_i is an irreducible polynomial in f_1 and f_2. Some one of the irreducible factors must vanish identically. This irreducible relation will contain both factors unless one of them is a constant;

for from a relation of the form $\Phi_i(f_1) = 0$, we deduce $f_1 = \text{const.}$ If neither function is a constant, there is essentially but one algebraic relation connecting them; for, from two independent relations $\Phi_i(f_1, f_2) \equiv 0$ and $\Phi_j(f_1, f_2) \equiv 0$, we have, on solving, $f_1 \equiv \text{const.}$ and $f_2 \equiv \text{const.}$

It is easy to see, in the general case, what the degrees of the irreducible equation in f_1 and f_2 will be. The degree in f_1 is the number of values of f_1 which satisfy the equation when f_2 is given a fixed value. There are k_2 points in the fundamental region at which $f_2 = C$, where, as before, k_2 is the number of poles of f_2. At each of these points f_1 has a value satisfying the irreducible equation. Hence, in general, this equation is of degree k_2 in f_1. Similarly, it is of degree k_1 in f_2, in general. It may happen, for particular functions, that some of the values of f_1 at the k_2 points are always coincident, in which case the degree in f_1 is less than k_2. The functions f_2 and $f_1 = f_2^2$ furnish a simple example.

There arises the question whether, conversely, each pair of values c_1, c_2 satisfying the irreducible equation $\Phi_i(c_1, c_2) = 0$ is taken on by the functions at some point of the fundamental region. This is, in fact, the case. The algebraic equation $\Phi_i(f_1, f_2) = 0$ determines f_1, say, as a function of f_2, $f_1 = \psi(f_2)$, and the Riemann surface of this function is connected. All pairs of values satisfying the equation are represented by points of the Riemann surface.

In the neighborhood of a point z_0 of the fundamental region let $f_1 = c_1'$, $f_2 = c_2'$, the points being so chosen that $c_1'c_2'$ is not at a branch point of the Riemann surface. In the neighborhood of z_0, $f_1(z)$ coincides with that branch of $\psi[f_2(z)]$ which takes on the value c_1' at z_0. These two functions are then equal wherever they can be extended analytically. Now, in the f_2-plane we can trace such a path that as f_2 moves from c_2' to c_2, $\psi(f_2)$ moves from c_1' to c_1. Along this path $z(f_2)$ is analytic provided we avoid certain singular positions, since the derivative $dz/df_2 = 1/f_2'(z)$ exists. Hence, z traces a path in the domain of existence of the functions. At the terminus z' of this path $f_1(z') = c_1$. In the point z_0' of the fundamental region congruent to z' we have $f_1(z_0') = c_1$, $f_2(z_0') = c_2$.

THEOREM 13.—*Any simple automorphic function can be expressed as a rational function of two simple automorphic functions which have the property that an arbitrary pair of values of the* *functions is taken on at but one point of the fundamental region, the domain of existence of the three functions being the same.*

Let f_1 and f_2 be two functions with the desired property, and let f_3 be a third function. To each pair of values of f_1 and f_2, there corresponds, in general, one value, and only one, of z in the region and, hence, one value of f_3. That is, f_3 is single valued on the Riemann surface of $f_1 = \psi(f_2)$. Also, f_3 is an analytic function of f_2, except for certain exceptional points where f_3 becomes infinite or $f_2'(z)$ is zero. At all exceptional points of the surface, f_3 approaches a finite value or becomes infinite. Hence, f_3 has no other singularities than poles on the Riemann surface. By a well-known theorem, f_3 is a rational function of f_1 and f_2:

$$f_3 = \frac{A_1 f_1^r f_2^s + \cdots}{B_1 f_1^p f_2^q + \cdots} = R(f_1, f_2). \tag{5}$$

THEOREM 14.—*If there exists a simple automorphic function $f_1(z)$ having a single pole in the fundamental region, then any simple automorphic function connected with the group, and having the same domain of definition, is a rational function of $f_1(z)$.*

Let $f_2(z)$ be a simple automorphic function with k_2 poles. If in equation (2) we take $m = k_2$, $n = 1$, the inequality (4) is satisfied: $2k_2 + 1 > 2k_2$. The identically vanishing polynomial (2) is then of the form

$$Q_{k_2}(f_1) \cdot f_2 + P_{k_2}(f_1) \equiv 0,$$

where P_{k_2} and Q_{k_2} are polynomials of degree k_2 at most. Not all the coefficients in Q_{k_2} are zero; for, otherwise, $P_{k_2}(f_1) \equiv 0$ and $f_1 \equiv$ const., contrary to hypothesis. We have, then,

$$f_2 = -\frac{P_{k_2}(f_1)}{Q_{k_2}(f_1)}, \tag{6}$$

which was to be proved.

COROLLARY.—*The most general simple automorphic function having a single pole in the fundamental region is the function*

$$f = \frac{Af_1 + B}{Cf_1 + D}, \quad AD - BC \neq 0.$$

It is clear that this function has a single pole in the fundamental region whatever the constants A, B, C, D may be, provided that $AD - BC \neq 0$. For, the values of f correspond in a one-to-one manner to those of f_1; hence, f takes on each value once in the fundamental region. That this is the most general such function follows from the proof of the preceding theorem. Any function f_2 having a single pole satisfies an equation of the form (6) where, since $k_2 = 1$, the numerator and denominator are linear, and where, since $f_2 \not\equiv$ const., the determinant is different from zero.

44. Differential Equations.—The derivative of an automorphic function is, in general, not automorphic. Differentiating the equation

$$f(z') \equiv f(z),$$

where

$$z' = \frac{az+b}{cz+d}, \quad \frac{dz'}{dz} = \frac{1}{(cz+d)^2},$$

we have

$$f'(z') = f'(z)\frac{dz}{dz'} = (cz+d)^2 f'(z). \tag{7}$$

The derivative is automorphic with respect to the group, provided it is not identically zero, only if $c = 0$, $d = \pm 1$ for all transformations of the group. The transformations are then all of the form $z' = z + b$; and the group, as we shall find later (Sec. 59), is simply or doubly periodic.

We observe from (7) that the quotient of the first derivatives of two automorphic functions belonging to a group is unaltered by the transformations of the group. This is apparent for simple automorphic functions when we differentiate the relation connecting them. We have $\Phi'_{f_1} f_1'(z) + \Phi'_{f_2} f_2'(z) = 0$; whence the quotient $f_1'(z)/f_2'(z)$ is rational in $f_1(z)$ and $f_2(z)$, and so is a simple automorphic function.

We can set up other combinations which are unaltered by the transformations of the group. Differentiating (7), we find

$$f''(z') \equiv (cz+d)^4 f''(z) + 2c(cz+d)^3 f'(z),$$

$$f'''(z') \equiv (cz+d)^6 f'''(z) + 6c(cz+d)^5 f''(z) + 6c^2(cz+d)^4 f'(z),$$

and

$$2f'(z')f'''(z') - 3f''(z')^2 = (cz+d)^8[2f'(z)f'''(z) - 3f''(z)^2]. \tag{8}$$

The quantity

$$D(y)_x = \frac{2\dfrac{dy}{dx}\dfrac{d^3y}{dx^3} - 3\left(\dfrac{d^2y}{dx^2}\right)^2}{2(dy/dx)^2} = \frac{d^2}{dx^2}\log\frac{dy}{dx} - \frac{1}{2}\left(\frac{d}{dx}\log\frac{dy}{dx}\right)^2 \tag{9}$$

is known as the "Schwarzian derivative"[1] of y with respect to

[1] SCHWARZ, H. A., "Gesammelte Mathematische Abhandlungen," Bd. 2, p. 78. Various notations have been used for this expression: $\psi(y, x)$ by Schwarz; $\{y, x\}$ by Cayley; $[y]_x$ by Klein; $D(y)_x$ by Koebe.

x. We have from (8) and (7) that the function

$$\frac{2f'(z)f'''(z) - 3f''(z)^2}{2f'(z)^4} = \frac{D(f)_z}{f'(z)^2} \tag{10}$$

is unaltered by the transformations of the group.

We now show that if $f(z)$ is a simple automorphic function so also is (10). If $f(z)$ is analytic or has a pole at a point, the same is true of its derivatives and the rational combination of derivatives in (10) is analytic or has a pole at the point. The function is then analytic except for poles throughout the domain of existence of $f(z)$. There remains the question of its behavior at the parabolic points, if any occur. At such a point $f(z)$ is analytic or has a pole at the origin when expressed as a function of t, where

$$t = e^Z,\ Z = \frac{2\pi i}{c(z - P)} \text{ or } Z = \frac{2\pi i z}{c},$$

according as the parabolic point is finite or infinite (Sec. 41).

In changing the variable, we make use of the following properties of the Schwarzian derivative, which are easily established:

$$D\left(\frac{ay + b}{cy + d}\right)_x = D(y)_x, \text{ whence } D\left(\frac{ax + b}{cx + d}\right) = D(x)_x = 0. \tag{11}$$

$$D(y)_x = D(y)_t\left(\frac{dt}{dx}\right)^2 + D(t)_x. \tag{12}$$

Equation (11) expresses the fundamental property of the Schwarzian derivative; in fact, the derivative was originally so set up as to be invariant when y is subjected to a linear transformation. Equation (12) is the formula for the change of variable.

Making the change of variable given above and noting that

$$D(Z)_z = 0, \quad D(t)_Z = -\tfrac{1}{2},$$

we have

$$D(f)_z = D(f)_Z\left(\frac{dZ}{dz}\right)^2 = \left(\frac{dZ}{dz}\right)^2\left[D(f)_t\left(\frac{dt}{dZ}\right)^2 - \frac{1}{2}\right];$$

whence,

$$\frac{D(f)_z}{(df/dz)^2} = \frac{D(f)_t}{(df/dt)^2} - \frac{1}{2(df/dZ)^2} = \frac{D(f)_t}{f'^2(t)} - \frac{1}{2t^2f'(t)^2}.$$

This function of t is analytic or has at most a pole at $t = 0$. Hence, the function (10) is a simple automorphic function.

We shall now prove the following remarkable theorem:

THEOREM 15.—*If $w(z)$ is a non-constant simple automorphic function of z, then z can be expressed as a function of w by the* *quotient of two solutions of a linear differential equation of th second order of the form*

$$\frac{d^2\eta}{dw^2} = u\eta, \tag{13}$$

where u is an algebraic function of w, $\Phi(u, w) = 0$.

If w has a single pole of the first order in the fundamental region, u is a rational function of w.

Consider the functions

$$\eta_1 = z\sqrt{\frac{dw}{dz}}, \quad \eta_2 = \sqrt{\frac{dw}{dz}}, \quad z = \frac{\eta_1}{\eta_2}. \tag{14}$$

We shall show that

$$\frac{1}{\eta_1}\frac{d^2\eta_1}{dw^2} = \frac{1}{\eta_2}\frac{d^2\eta_2}{dw^2} = \frac{D(w)_z}{2(dw/dz)^2}. \tag{15}$$

From the relations

$$\frac{dw}{dz} = \frac{\eta_1^2}{z^2} = \eta_2^2, \tag{16}$$

we proceed to find $D(w)_z$, using the last formula of (9). Taking logarithms and differentiating, we have

$$\frac{d}{dz}\log\frac{dw}{dz} = \frac{2\dfrac{d\eta_1}{dw}}{\eta_1}\frac{dw}{dz} - \frac{2}{z} = \frac{2\dfrac{d\eta_2}{dw}}{\eta_2}\frac{dw}{dz}.$$

Replacing dw/dz in the second and third members by its values in (16),

$$\frac{d}{dz}\log\frac{dw}{dz} = \frac{2\eta_1\dfrac{d\eta_1}{dw}}{z^2} - \frac{2}{z} = 2\eta_2\frac{d\eta_2}{dw}. \tag{17}$$

Differentiating again and substituting as before for dw/dz,

$$\begin{aligned}\frac{d^2}{dz^2}\log\frac{dw}{dz} &= \left[\frac{2\eta_1\dfrac{d^2\eta_1}{dw^2}}{z^2} + \frac{2\left(\dfrac{d\eta_1}{dw}\right)^2}{z^2}\right]\frac{\eta_1^2}{z^2} - \frac{4\eta_1\dfrac{d\eta_1}{dw}}{z^3} + \frac{2}{z^2} \\ &= \left[2\eta_2\frac{d^2\eta_2}{dw^2} + 2\left(\frac{d\eta_2}{dw}\right)^2\right]\eta_2^2.\end{aligned} \tag{18}$$

On subtracting half the square of (17) from (18), most of the terms cancel, and we have

$$D(w)_z = \frac{2\eta_1^3\dfrac{d^2\eta_1}{dw^2}}{z^4} = 2\eta_2^3\frac{d^2\eta_2}{dw^2}.$$

On dividing by $2(dw/dz)^2$ from (16) we have (15).

It follows from (15) that η_1 and η_2, whose ratio is z, are solutions of the differential equation

$$\frac{d^2\eta}{dw^2} = u\eta,$$

where we write u for the last member of (15). But we found in (10) that u is a simple automorphic function of z. Hence, (Theorem 12) u and w are connected by an algebraic relation, $\Phi(u, w) = 0$. In particular, if w has a single pole in the fundamental region (Theorem 14), u is a rational function of w. The theorem is thus established.

CHAPTER V

THE POINCARÉ THETA SERIES

45. The Theta Series.—In the preceding chapter, we assumed the existence of automorphic functions and studied their properties. In the present chapter, we shall demonstrate their existence by the process of actually setting them up by means of series.

Let the transformations of the group be

$$z_i = T_i(z) = \frac{a_i z + b_i}{c_i z + d_i},\ a_i d_i - b_i c_i = 1, \qquad (1)$$
$$i = 0, 1, 2, \ldots,$$

the identical transformation being $z_0 = T_0(z) = z$. As an aid to simplicity in the formulæ, we shall use the notation $z_{ij} = T_i(z_j) = T_iT_j(z)$, $z_{ijk} = T_iT_jT_k(z)$, etc.

We consider, first, a case whose treatment involves little difficulty, namely, the finite group. Let the group contain m transformations ($i = 0, 1, \cdots, m - 1$). Let $H(z)$ be any rational function of z and form the function

$$\varphi(z) = H(z) + H(z_1) + H(z_2) + \cdots + H(z_{m-1}). \qquad (2)$$

This function has no other singularities than poles. If we apply a transformation of the group to z, we have

$$\varphi(z_k) = H(z_k) + H(z_{1k}) + \cdots + H(z_{m-1,k}).$$

Now, $z_k, z_{1k}, \ldots, z_{m-1,k}$ are the set of transforms of z and, since z_k is congruent to z, this set coincides with $z, z_1, \ldots, z_{m-1}$. The terms in the sum are the same as before, their order being merely interchanged; hence, $\varphi(z_k) = \varphi(z)$. The function is thus automorphic. In fact, having no other singularities than poles, it is a simple automorphic function. In like manner, any rational symmetric function of $H(z), H(z_1), \ldots, H(z_{m-1})$ is simple automorphic function.

If the group contains an infinite number of transformations and we extend (2) to an infinite number of terms, the series will ordinarily fail to converge. For, the general term $H(z_n)$ will not even approach zero unless $H(z)$ is zero in the limit points about which the points z_n cluster. Poincaré got around the difficulty by the introduction of convergence factors. By the use of these factors we are led to a sum which is no longer automorphic but which behaves in a simple manner when a transformation of the group is applied.

Let $H(z)$ be a rational function none of whose poles is at a limit point of the group.[1] We consider the following series:

$$\theta(z) = \sum_{i=0}^{\infty}(c_i z + d_i)^{-2m}H(z_i). \tag{3}$$

This is the theta series of Poincaré.[2] We shall presently establish its convergence, under suitable circumstances, when m is an integer greater than 1. We shall assume its convergence and derive the basic property of the function which it defines.

If z be subjected to a transformation T_j of the group, the series becomes

$$\begin{aligned}\theta(z_j) &= \sum\left(c_i\frac{a_j z + b_j}{c_j z + d_j} + d_i\right)^{-2m} H(z_{ij}) \\ &= \sum\left[\frac{(c_i a_j + d_i c_j)z + c_i b_j + d_i d_j}{c_j z + d_j}\right]^{-2m} H(z_{ij}).\end{aligned}$$

The factor $(c_j z + d_j)^{2m}$ comes out of all the terms and we have, on replacing the numerator by an equivalent expression,

$$\theta(z_j) = (c_j z + d_j)^{2m}\Sigma(c_{ij}z + d_{ij})^{-2m}H(z_{ij}).$$

The series on the right is the series (3) with the terms rearranged. Our subsequent convergence proof will justify rearranging the terms, so that we have

$$\theta(z_j) = (c_j z + d_j)^{2m}\theta(z). \tag{4}$$

This equation expresses the fundamental property of the theta series.

By means of these series we can set up functions which are unaltered when a transformation of the group is applied. Let

[1] In the case of a function group which carries a region Σ into itself, it ufficies that $H(z)$ have no other singularities than poles in Σ and be bounded on the boundary of Σ.

[2] POINCARÉ, H., "Mémoire sur les fonctions fuchsiennes," *Acta Math.*, vol. 1, pp. 193–294; Oeuvres, vol. 2, pp. 169–257.

$\theta_1(z)$ and $\theta_2(z)$ be two theta series formed with the same integer m and consider their quotient, $F(z) = \theta_1(z)/\theta_2(z)$. We have

$$F(z_j) = \frac{\theta_1(z_j)}{\theta_2(z_j)} = \frac{(c_jz + d_j)^{2m}\theta_1(z)}{(c_jz + d_j)^{2m}\theta_2(z)} = F(z). \tag{5}$$

It will appear subsequently that, for a function group, z_j lies in the domain of existence of the theta functions. Then $F(z)$ is an automorphic function.

Poincaré called the series (3) a "theta-fuchsian series" or a "theta-kleinian series" according as the group to which it belongs is Fuchsian or Kleinian. He calls a function with the property (4) and which has no other singularities than poles at ordinary points of the group, whatever the manner of its formation may be, a "theta-fuchsian function" or a "theta-kleinian function." For example, the derivative of an automorphic function (Sec. 44, Equation 7) is a theta function with $m = 1$.

46. The Convergence of the Series.—The following proposition is fundamental for establishing the convergence of various series and products connected with the group:

THEOREM 1.—*If the point at infinity is an ordinary point of the group, the series* $\Sigma|c_n|^{-2m}$, *where in the summation the finite number of terms for which* $c_n = 0$ *are omitted, converges for* $m \geqslant 2$.

The series in the theorem can be written Σr_n^{2m}, where r_n is the radius of the isometric circle I_n of T_n. Suppose, first, that infinity is not a fixed point for an elliptic transformation. Then, there are no points congruent to infinity in the neighborhood of infinity, and we have the groups for which we developed the properties of the isometric circles in Chap. II.

It will suffice to prove the theorem for $m = 2$. Except, possibly, for a finite number of terms $r_n < 1$, so that if $m > 2$, $r_n^{2m} < r_n^4$ and the convergence follows from that of Σr_n^4.

The centers of the isometric circles lie in a finite region and their radii are bounded (Sec. 17). Hence there exists a positive constant h such that a circle Q_n of radius h concentric with any isometric circle I_n contains all the isometric circles on its interior. Let Q_n' be the transform of Q_n by T_n. Q_n' is got by inverting Q_n in I_n and making certain other transformations which do not alter magnitudes (Theorem 19, Sec. 11). By inversion, we find the radius of Q_n' to be r_n^2/h; its area is $\pi r_n^4/h^2$. Now, the exterior of Q_n, which lies entirely within R, goes into the interior of Q_n', which, therefore, lies within R_n, the transform of R by T_n. As the regions R_n fit together without overlapping, it follows that there is no overlapping of the circles Q_n'. Their areas then are less than the area of any one of the circles Q_n which encloses them. We have then

$$\sum \frac{\pi r_n^4}{h^2} < \pi h^2, \quad \sum r_n^4 < h^4,$$

and the series converges.

If infinity is a fixed point for an elliptic transformation, the reasoning is not essentially different. The isometric circles are confined to a finite region as before, and the constant h exists. There is the difference that a point outside Q_n may have $p - 1$ points exterior to Q_n congruent to it, where the elliptic cyclic subgroup with fixed point at ∞ contains p transformations. The circles Q_n' can overlap, but no point can be interior to more than p such circles. Hence,

$$\sum \frac{\pi r_n^4}{h^2} < p\pi h^2, \quad \sum r_n^4 < ph^4,$$

and the series is convergent.

By the use of the preceding theorem, we are able to establish the convergence of the theta series.

THEOREM 2.—*If $m \geqslant 2$ and if the point at infinity is an ordinary point of the group, then the theta series* (3) *defines a function which is analytic except possibly for poles in any connected region not containing limit points of the group in its interior.*

It will suffice to prove the theorem for a region S' such that there are no limit points of the group within or on the boundary of S', since such a region can be made large enough to include any given interior point of a region with limit points on its boundary.

We observe, first, that certain terms of (3) may have poles in S'. At $z = -d_i/c_i$, the center of the isometric circle I_i, the factor $(c_iz + d_i)^{-2m}$ becomes infinite. Again, if z is such that $z_i = a$, where a is the pole of $H(z)$,—that is, if $z = T_i^{-1}(a)$—then $H(z_i)$ has a pole. It is clear, however, that only a finite number of terms of the series have poles in S'; for S' contains in its interior and on its boundary only a finite number of centers of isometric circles and only a finite number of points congruent to each of the poles of $H(z)$.

We now put aside the finite number of terms having poles in and on the boundary of S', and we prove that the remainder of the series converges absolutely and uniformly in S'. Let $d > 0$ be the minimum distance from the boundary of S' to the centers of those isometric circles whose centers are exterior to S'. We have, then, for all the terms we are considering and for all z in S'

$$\left|z + \frac{d_i}{c_i}\right| \geqslant d.$$

Further, we can draw curves about the poles of $H(z)$ such that when z is in S' all points z_i, in the terms considered, lie outside the several curves. But in the regions outside these curves, $H(z)$ is bounded, so that we have

$$|H(z_i)| < M.$$

We have, then, excluding further the finite number of terms for which $c_i = 0$,

$$|(c_i z + d_i)^{-2m} H(z_i)| = \left|\frac{H(z_i)}{c_i^{2m}\left(z + \frac{d_i}{c_i}\right)^{2m}}\right| < \frac{M}{d^{2m}}|c_i|^{-2m}. \tag{6}$$

This inequality holds for all points of S'. Since the series of positive constant terms $\sum \frac{M}{d^{2m}}|c_i|^{-2m}$ converges (Theorem 1), the absolute and uniform convergence of the series in S' is established. The sum of the series is an analytic function in S'. It follows that (3) is analytic in S' except for the finite number of poles which arise from the terms which we put aside.

47. The Convergence for the Fuchsian Group of the Second Kind.—It can be shown that in general the theta series does not converge if $m = 1$. There are certain groups, however, for which the series does converge. An important case is the following:[1]

THEOREM 3.—*For a Fuchsian group of the second kind for which the point at infinity is an ordinary point, the series* $\Sigma|c_n|^{-2}$ *converges. Then, the theta series* (3) *converges for* $m = 1$.

We shall suppose, first, that the principal circle is a straight line—the real axis, for example. The region R contains a portion of the real axis in the neighborhood of infinity in its interior. The isometric circles lie in a finite region, and their centers are on the real axis.

We shall employ the constructions used in the proof of Theorem 1. The circle Q_n concentric with I_n and of suitable radius h

[1] BURNSIDE, W., "On a Class of Automorphic Functions," *Proc.* London Math. Soc., vol. 23, pp. 49–88, 1892.

lies in R. Its exterior is transformed into the interior of the circle Q_n', with center on the real axis. The portion of the real axis exterior to Q_n is transformed into the portion of the real axis within Q_n'. The length of this latter segment is $2r_n^2/h$, the diameter of Q_n'. Since the circles Q_n' are exterior to one another and all lie in any one of the circles Q_n, the segments of the real axis which they contain are non-overlapping, and the sum of their lengths is finite:

$$\sum \frac{2r_n^2}{h} < 2h, \quad \sum r_n^2 < h^2.$$

Hence, the series Σr_n^2, or $\Sigma |c_n|^{-2}$, converges.

The proof for the general Fuchsian group of the second kind rests on the following lemma:

LEMMA.—*If the point at infinity is an ordinary point for a group T_n and its transform $T_n' = GT_nG^{-1}$, then the series $\Sigma|c_n|^{-2m}$ and $\Sigma|c_n'|^{-2m}$ both converge or both diverge.*

Let the transformation which is applied to the group be $G = (\alpha z + \beta)/(\gamma z + \delta)$, $\alpha\delta - \beta\gamma = 1$. Then (Equation (2), Sec. 15)

$$c_n' = -\gamma\delta a_n + \gamma^2 b_n - \delta^2 c_n + \gamma\delta d_n. \tag{7}$$

Suppose, first, that $\gamma \neq 0$. Substituting for b_n the value $b_n = (a_n d_n - 1)/c_n$ and combining terms, we can write (7) in the form

$$c_n' = \gamma^2 c_n \left[\left(\frac{d_n}{c_n} - \frac{\delta}{\gamma}\right)\left(\frac{a_n}{c_n} + \frac{\delta}{\gamma}\right) - \frac{1}{c_n^2}\right]. \tag{8}$$

Now, the centers, $-d_n/c_n$ and a_n/c_n, of the isometric circles lie in a finite region; and their distances from the point $-\delta/\gamma$ are bounded

$$\left|\frac{d_n}{c_n} - \frac{\delta}{\gamma}\right| < K, \quad \left|\frac{a_n}{c_n} + \frac{\delta}{\gamma}\right| < K.$$

Also, $1/|c_n^2|$ is bounded: $1/|c_n^2| < K'$. Hence, we have

$$|c_n'| < |\gamma^2| \cdot |c_n|(K^2 + K') = K_1|c_n|,$$

and

$$|c_n|^{-2m} < K_1^{2m}|c_n'|^{-2m}.$$

It follows from this inequality that if $\Sigma|c_n'|^{-2m}$ converges, so also does $\Sigma|c_n|^{-2m}$.

If $\gamma = 0$, we have at once $|c_n'| = |\delta|^2 \cdot |c_n|$, δ being different from zero, and the conclusion follows as before.

Now the group T_n is also a transform of T_n'; thus $T_n = G^{-1}T_n'G$. Hence, if $\Sigma|c_n|^{-2m}$ converges, so, also, does $\Sigma|c_n'|^{-2m}$. This establishes the lemma.

We return now to the general Fuchsian group of the second kind. Let the group be transformed so that the principal circle is carried into the real axis and an ordinary point on the principal circle is carried to infinity. We found above that $\Sigma|c_n'|^{-2}$ converges for the transformed group. Hence, applying the lemma, $\Sigma|c_n|^{-2}$ converges for the original group.

The convergence of the theta series follows from Equation (6) with $m = 1$ since the series of positive terms $\sum \frac{M}{d^2}|c_i|^{-2}$ converges.

48. Some Properties of the Theta Functions.—Let Σ be a connected region of the z-plane bounded by limit points. Then, $\theta(z)$ in (3) is analytic except for poles in Σ. If the limit points separate the plane into two or more regions, the series (3) defines a function in each region, but, in general, the functions so defined are distinct.

Consider, for example, the Fuchsian group of the first kind. Here the limit points consist of all points of the principal circle. Then (3) defines a function analytic except for poles within the principal circle. The poles of $\theta(z)$ arise from the poles of the individual terms of the series. If $H(z)$ has a pole at a within the principal circle, then $H(z_i)$ becomes infinite when $z_i = a$. That is, $\theta(z)$ has poles at the poles of $H(z)$ within the principal circle and at points congruent to these poles—except that, in special cases, $H(z)$ may have poles at congruent points of such a character that the singularities arising from two or more terms of the series cancel. Putting this special case aside, the number of poles in the fundamental region R_0 is exactly equal to the number of poles of $H(z)$ within the principal circle. If $\theta(z)$ has poles within the principal circle, these poles cluster in infinite numbers about each point on the principal circle. The principal circle is thus a natural boundary of the function.

In a similar manner, (3) defines a function analytic except for poles on the exterior of the principal circle. These poles are the poles of $H(z)$ lying without the principal circle, the points congruent thereto, and the points $-d_i/c_i$, where a convergence factor becomes infinite; although here, also, there is a possibility of the singularities from different terms cancelling. If the function has poles, these poles cluster about the points of the principal circle, and the function cannot be extended analytically across the circle. The one formula (3) then defines two distinct functions, one existing within the principal circle, the other without.

In the Fuchsian group of the second kind the function defined by (3) is a single function, existing both within and without the principal circle.

The proof of the convergence of the theta series does not require that the group be a function group. The property (4) is a relation connecting two distinct functions, however, unless $\theta(z)$ can be continued analytically from z to z_j. In order that (4) express a property of a single one of the functions defined by the series, it is necessary that the domain of existence of the function be carried into itself by the transformations of the group. The group is then a function group.

In setting up functions by means of the series (3), the poles of $H(z)$ are at our disposal. By placing a pole at a desired point, we can be sure that $\theta(z)$ in a region Σ under consideration has a singularity and, hence, is not identically zero. Further, in forming automorphic functions for a function group by means of (5), we can place the poles of the numerator and denominator at different points of Σ and, thus, be assured that the automorphic function does not reduce to the trivial case of a constant.

Equation (4) exhibits a relation between the theta function and the isometric circle. If $\theta(z)$ is neither zero nor infinite at z, then $|\theta(z_j)|$ is greater than, less than, or equal to $|\theta(z)|$ according as $|c_jz + d_j|$ is greater than, less than, or equal to 1; that is, according as z lies without, within, or on the isometric circle of T_j. We note that $|\theta(z)|$, when not zero or infinity, has a smaller value at a point of R (exterior to all isometric circles) than at any congruent point within the domain of existence of the function.

Behavior at a Vertex.—Let ξ be the fixedpoint of an elliptic transformation lying within the domain of existence of a theta function. The transformations with ξ as fixed point form an elliptic cyclic group generated by $z' = T(z)$ of the form

$$\frac{z' - \xi}{z' - \xi'} = K\frac{z - \xi}{z - \xi'}, \quad K = e^{2\pi i/k}, \tag{9}$$

where ξ' is the second fixed point and k is an integer greater than 1. The fundamental property (4), for T, may be written

$$\theta(z') = \left(\frac{dT}{dz}\right)^{-m}\theta(z). \tag{10}$$

We find, on taking the logarithm of (9) and differentiating,

$$\frac{dT}{dz} = \frac{dz'}{dz} = \frac{(z' - \xi)(z' - \xi')}{(z - \xi)(z - \xi')}. \tag{11}$$

If we put

$$F(z) = (z - \xi)^m (z - \xi')^m \theta(z), \tag{12}$$

we can write (10) in the form

$$\begin{aligned} F(z') &= (z' - \xi)^m (z' - \xi')^m \theta(z') \\ &= (z - \xi)^m (z - \xi')^m \theta(z) \\ &= F(z). \end{aligned} \tag{13}$$

Hence, $F(z)$ is unaltered by the transformation T, and consequently is automorphic with respect to the cyclic subgroup generated by T.

Applying Theorem 6, Sec. 41, we conclude that $F(z)$ *takes on its value k times, or a multiple thereof, at ξ.*

Suppose that $\theta(z)$ is analytic at ξ. Then $F(z)$ has a zero at ξ of order m at least, owing to the factor $(z - \xi)^m$. Now, if m is not a multiple of k, then $\theta(z)$ must have a zero at ξ also, the order s of the zero being such that

$$m + s = lk, \tag{14}$$

where l is an integer.

If $\theta(z)$ is analytic at ξ, it is necessarily zero there unless m is a multiple of k. The order s of the zero satisfies an equation of the form (14).

If $\theta(z)$ has a pole of order p at ξ, then $F(z)$ has a pole of order $p - m$ or a zero of order $m - p$, unless $p = m$. An equation of the form (14) holds where we put $s = -p$, l being a positive or negative integer or zero.

Behavior at a Parabolic Point.—There is a parabolic cyclic subgroup generated by a transformation S with the parabolic point P as fixed point. S has the form

$$\frac{1}{z' - P} = \frac{1}{z - P} + c. \tag{15}$$

Here we have

$$\frac{dS}{dz} = \frac{dz'}{dz} = \left(\frac{z' - P}{z - P}\right)^2. \tag{16}$$

Writing

$$G(z) = (z - P)^{2m} \theta(z), \tag{17}$$

the fundamental relation (10) for the transformation S may be written in the form

$$G(z') = (z' - P)^{2m} \theta(z') = (z - P)^{2m} \theta(z) = G(z). \tag{18}$$

The function $G(z)$ is automorphic with respect to the subgroup generated by S.

On making the customary change of variable

$$t = e^{2\pi i/c(z-P)} \tag{19}$$

(See Sec. 41; here we use particularly Fig. 32), $G(z)$ is transformed into a function of t, $G(z) \equiv g(t)$, single valued in the neighborhood of $t = 0$. We have, then,

$$\theta(z) = \frac{G(z)}{(z-P)^{2m}} = \left(\frac{c \log t}{2\pi i}\right)^{2m} g(t). \tag{20}$$

The theta function thus has a logarithmic singularity at $t = 0$.

We consider now the form of $g(t)$ when $\theta(z)$ is defined by the theta series (3). We can take the circle C of Fig. 32(a) small enough that it contains no point $-d_i/c_i$ and no point congruent to a pole of $H(z)$. For C, when small enough, contains only points congruent to points of the fundamental region R which lie in the neighborhoods of the parabolic points of the cycle; and these neighborhoods can be sufficiently restricted to exclude ∞ and the finite number of points of R congruent to poles of $H(z)$. Taking C slightly smaller we have bounds for the following quantities:

$$\left|\frac{z-P}{z+\dfrac{d_i}{c_i}}\right| < K, \qquad |H(z_i)| < M,$$

where z lies within or on the boundary of the triangle z_1z_2P of Fig. 32(a), exclusive of P itself. In the series

$$G(z) = \sum\left(\frac{z-P}{z+\dfrac{d_i}{c_i}}\right)^{2m} \frac{1}{c_i^{2m}} H(z_i),$$

the general term is less in absolute value than the corresponding term of the convergent series of positive terms $K^{2m}M\Sigma|c_i|^{-2m}$. $G(z)$ thus remains finite as z approaches P from the interior of the triangle. In the t-plane, then, $g(t)$ is analytic in the neighborhood of $t = 0$ and is bounded; hence, *$g(t)$ is analytic at $t = 0$ if properly defined there.*

If we form an automorphic function as the quotient of two theta series, which may be written

$$f(z) = \frac{(z-P)^{2m}\theta_1(z)}{(z-P)^{2m}\theta_2(z)} = \frac{g_1(t)}{g_2(t)},$$

we observe that $f(z)$, as a function of t, is analytic or has a pole at $t = 0$. Then as z approaches P from the interior of R $f(z)$ approaches a definite finite or infinite limit. It thus satisfies the requirements laid down in Sec. 40 for the behavior of a simple automorphic function at a parabolic point.

49. Zeros and Poles of the Theta Functions.—We consider now function groups in which the region R_0 lying in Σ has a finite number of sides (Sec. 40). We consider theta functions having a finite number of poles in R_0, such as those defined by the theta series (3). We suppose that $g(t)$ in (20), as in the case of the series (3), is analytic at $t = 0$ or has a pole there; and we assume further that $\theta(z)$ is not identically zero.

We make the same conventions about counting zeros and poles at ordinary points on the boundary as were made in Sec. 42. We observe that this may lead to the count of a fractional number of zeros or poles in the region; thus, at the vertices of a cycle of angle $2\pi/k$ the zeros, from (14), amount to

$$\frac{s}{k} = l - \frac{m}{k},$$

which may be fractional. The order of the zero or pole of $g(t)$ at $t = 0$ will be counted as the number of zeros or poles of the function $\theta(z)$ in the points of the parabolic cycle to which P belongs.

We now remove the neighborhoods of the vertices and parabolic points of R_0 by small circles as explained in Sec. 42. Let N_0, M_0 be the number of zeros and poles, respectively, in the resulting region. Then,

$$N_0 - M_0 = \frac{1}{2\pi i}\int d \log \theta(z), \tag{21}$$

the integral being taken around the boundary—further altered as in Fig. 33(*b*) if a pole or zero lies on a side.

Before evaluating the integrals along the sides, we shall consider the integrals along the small circular arcs just constructed. When we let the radii of the arcs cutting off the vertices approach zero, we find, exactly as in Sec. 42, that the integrals along these arcs approach $m_0 - n_0$ where n_0 and m_0 are the number of zeros and poles, respectively, that we assign to the various vertices.

The parabolic points, as usual, are less simple to handle. The circular arcs $h, h_1, \ldots, h_n$ about the parabolic points of a cycle $P, P_1, \ldots, P_n$ are the transforms of $h, h_1', \ldots, h_n'$ which fit together to make up the arc z_1z_2 of Fig. 32(a). The arc h_k is the transform of h_k' by some transformation $\eta_k = T_k(z)$, and we have,

$$\begin{aligned}\int_{h_k} d\log\theta(\eta_k) &= \int_{h_k'} d\log(c_kz+d_k)^{2m}\theta(z)\\ &= \int_{h_k'} d\log\left(\frac{c_kz+d_k}{z-P}\right)^{2m} G(z)\\ = 2m\int_{h_k'}\left[d\log\left(z+\frac{d_k}{c_k}\right) - d\log(z-P)\right] &+ \int_{h_k''} d\log g(t), \quad (22)\end{aligned}$$

where h_k'' is the corresponding arc on the circle in the t-plane in Fig. 32(c).

We need consider here only the imaginary part of the integral $\int d\log\theta$, since the members of (21) are necessarily real. Now, when we let the radius of C in Fig. 32(a) approach zero, it is easy to see that the imaginary parts of

$$\int_{h_k'} d\log\left(z+\frac{d_k}{c_k}\right), \qquad \int_{h_k'} d\log(z-P)$$

both approach zero, being i times the angle through which the line from $-d_k/c_k$, or from P to z, turns as z moves along the arc h_k' lying on z_1z_2.

We have, then, only the last integral in (22) to consider. Adding the integrals arising from the several vertices, we have

$$\frac{1}{2\pi i}\int d\log g(t),$$

the integral being taken around the circle in Fig. 32(c) in a clockwise direction. Its value is $-n'$ or m', where $g(t)$ has a zero of order n' or a pole of order m' at $t = 0$.

On shifting the integrals evaluated to the first member of (21), and letting N, M represent the total number of zeros and poles, respectively, belonging to R_0, we have

$$N - M = \frac{1}{2\pi i}\sum\int_{l_i} d\log\theta(z), \qquad (23)$$

the integrals being taken along the *sides* l_i of R_0.

Consider the two integrals arising from congruent sides AB, CD (Fig. 33 (a) or (b)). Let $z' = (az+b)/(cz+d)$ be the transformation carrying AB to CD. Here, $-d/c$ is the center of the arc AB. We have

$$\int_{AB} d \log \theta(z) + \int_{DC} d \log \theta(z') = \int_{AB} d \log \theta(z) + \int_{BA} d \log (cz + d)^{2m}\theta(z) = 2m\int_{BA} d \log\left(z + \frac{d}{c}\right) = 2im\alpha,$$

where α is the angle subtended by the arc BA at its center. The equal arc CD also subtends the angle α at its center; and we can write the integral as mi times the sum of the angles subtended by the two arcs. We get similar results for the other sides, which gives the following theorem:

THEOREM 4.—*Let N, M be the number of zeros and poles, respectively, of the theta function in the region R_0; and let $\alpha_1, \alpha_2, \ldots, \alpha_k$ be the angles subtended by the sides of R_0 at the centers of the isometric circles on which these sides lie. Then*

$$N - M = \frac{m}{2\pi}\sum \alpha_i. \tag{24}$$

We observe that the function has always more zeros than poles in R_0. We note that the difference $N - M$ depends only upon m and upon the character of the group. It is independent of the function $H(z)$ used in the construction of the theta series (3). This independence could have been easily foreseen. The fact that the automorphic function in (5) has the same number of zeros as poles requires that the difference $N - M$ for the functions $\theta_1(z)$ and $\theta_2(z)$ be the same.

The poles and zeros in a region R_j congruent to R_0 are determined from (4). If R_0 does not contain the point ∞ the number of zeros and poles in R_j is the same as in R_0. If, however, R_0 contains ∞, $N - M$ is decreased by $2m$, owing to the factor $(c_jz + d_j)^{2m}$ appearing in (4).

Suppose R_0 is a fundamental region with n pairs of sides for the Fuchsian group of the first kind. Let $2\pi/k_1, 2\pi/k_2, \ldots$ be the sum of the angles in the ordinary cycles of R_0. Then the sum of the interior angles in the circular polygon is $2\pi\sum\frac{1}{k_i}$.

Consider the rectilinear polygon of $4n$ sides in Fig. 34. The sum of its interior angles is $2\pi(2n - 1)$. The sum of the angles at the centers of isometric circles is $\Sigma\alpha_i$; the sum at the remaining vertices is $2\pi\sum\frac{1}{k_i} + 2n\pi$. Hence,

$$2\pi(2n - 1) = \sum\alpha_i + 2\pi\sum\frac{1}{k_i} + 2n\pi.$$

Substituting the value of $\Sigma\alpha_i$ from this into (24), we have

$$N - M = m\left(n - 1 - \sum\frac{1}{k_i}\right).$$

This result is given by Poincaré in the memoir previously cited.

A particularly simple application of (24) is to the group of Sec. 25(*a*). Here, R is bounded by $2n$ complete circles, and $\alpha_i = 2\pi$. For R, then,

$$N - M = 2mn.$$

Here, R contains ∞; so for R_j, any congruent region,

$$N - M = 2m(n - 1).$$

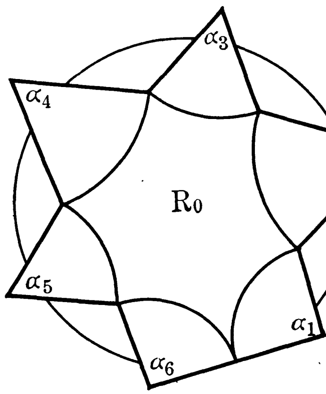


FIG. 34.

50. Series and Products Connected with the Group.—By the aid of Theorems 1 and 3 we can establish the convergence of numerous series and products connected with the group. We illustrate with a few examples.

Let $z_n, u_n, v_n, \ldots$ be the transforms of $z, u, v, \ldots$ by the transformation T_n of the group. We have

$$z_n - u_n = \frac{z - u}{(c_n z + d_n)(c_n u + d_n)} = \frac{z - u}{\left(z + \frac{d_n}{c_n}\right)\left(u + \frac{d_n}{c_n}\right)} \cdot \frac{1}{c_n^2}. \quad (25)$$

If z and u be restricted to lie in regions containing no limit points of the group and having no limit points on the boundary, then, excepting for a finite number of values of n, the factor preceding $1/c_n^2$ in (25) is bounded.

Suppose that $\Sigma|c_n|^{-2}$ converges; and consider the series

$$\Sigma(z_n - u_n),$$

the summation being extended to all transformations of the group. Except for a finite number of terms, this series converges absolutely and uniformly in the regions mentioned. It is thus analytic in each variable except possibly for poles arising from a finite number of individual terms. In a similar manner we establish the convergence of such functions as

$$\sum \frac{z_n - u_n}{z - w_n}, \qquad \sum (z_n - u_n) H(z_n),$$

where $H(z)$ is defined as for the theta series; etc.

In a subsequent chapter (Sec. 99), we shall employ products of the type

$$\prod \frac{z - u_n}{z - v_n} = \prod \left(1 + \frac{v_n - u_n}{z - v_n}\right) \tag{26}$$

in connection with Fuchsian groups of the second kind, where u and v are distinct from limit points. The convergence here depends upon the convergence of the series

$$\sum \frac{v_n - u_n}{z - v_n},$$

which is readily established from the known fact that $\Sigma |c_n|^{-2}$ converges. We conclude readily that (26) is an analytic function of z at all ordinary points of the z-plane except for poles at the points v_n, and that the function is different from zero except at the set of points u_n.

In the general case, where $\Sigma |c_n|^{-4}$ converges, we establish readily the convergence of such series as

$$\sum (z_n - u_n)^2, \quad \sum \frac{(z_n - u_n)(v_n - w_n)}{z - t_n}, \quad \sum \frac{z_n - u_n}{(c_n z + d_n)^2} H(z_n),$$

and the like.

From (26) we can build a convergent product for the general case by the insertion of convergence factors, thus

$$\prod \frac{z - u_n}{z - v_n} e^{\frac{u_n - v_n}{z - w_n}},$$

where w is an ordinary point. Another example in which the convergence is easily established is the product due to Whittaker[1]

$$(z - u) \prod{}' \left[\frac{(z - u_n)(u - z_n)}{(z - z_n)(u - u_n)}\right]^{\frac{1}{2}} e^{\frac{1}{2} \frac{(z-u)(z_n - u_n)}{(z - z_n)(u - u_n)}}$$

where the product extends to all transformations of the group excepting the identical transformation.

[1] *Mess. Math.*, vol. 31, p. 145, 1902.

CHAPTER VI

THE ELEMENTARY GROUPS

I. THE FINITE GROUPS

51. Inversion in a Sphere.—The study of finite groups will be much simplified by the introduction of transformations of three dimensional space. In this treatment, the space transformation known as "inversion in a sphere" plays a fundamental rôle. We shall now make a digression and derive the salient properties of this transformation.

Inversion in a sphere is a direct generalization of inversion in a circle (Sec. 5). Consider a sphere S with center K and radius ρ. Let P be any point in space and construct the half line beginning at K and passing through P. Let P_1 be a point on this half line such that $KP_1 \cdot KP = \rho^2$; then P_1 is called *the inverse of P with respect to S.* P is, also, obviously the inverse of P_1. S and K are called the "sphere of inversion" and the "center of inversion," respectively. The inversion carries a point within S into a point without S and leaves the points on the surface of S fixed.

It is clear also, as in Sec. 5, that *any sphere or circle through P and P_1 is orthogonal to S.* For, any tangent KT from K to the sphere or circle (Fig. 1) has the property that $\overline{KT}^2 = KP_1 \cdot KP = \rho^2$. That is, T lies on S; so that at T the sphere or circle is orthogonal to S. Further, any other secant through K, meeting the sphere or circle at P_1', P' has the property that $KP_1' \cdot KP' = \overline{KT}^2 = \rho^2$. Hence, by the inversion, any circle or sphere through two inverse points is transformed into itself.

We now derive the analytic expressions for the inversion. Let K, P, P_1 be the points (a, b, c), (ξ, η, ζ), (ξ_1, η_1, ζ_1), respectively, in rectangular space coordinates. Put

$$r^2 = (\xi - a)^2 + (\eta - b)^2 + (\zeta - c)^2,$$
$$r'^2 = (\xi' - a)^2 + (\eta' - b)^2 + (\zeta' - c)^2. \qquad (1)$$

Then it is required that

$$rr' = \rho^2. \qquad (2)$$

Projecting the segments KP and KP_1 on the three axes, we have, from the resulting similar triangles,

$$\frac{\xi' - a}{\xi - a} = \frac{\eta' - b}{\eta - b} = \frac{\zeta' - c}{\zeta - c} = \frac{r'}{r} = \frac{rr'}{r^2} = \frac{\rho^2}{r^2}. \tag{3}$$

These give the following equations for the transformations

$$\xi' - a = \frac{\rho^2(\xi - a)}{r^2},\ \eta' - b = \frac{\rho^2(\eta - b)}{r^2},\ \zeta' - c = \frac{\rho^2(\zeta - c)}{r^2}. \tag{4}$$

Expressions for ξ, η, ζ in terms of ξ', η', ζ' are got by interchanging the primed and unprimed variables in (4).

Consider, now, the magnification of an element ds. Let ds' be the length of the transformed element. We have from (4)

$$d\xi' = \frac{\rho^2}{r^2}d\xi - \frac{2\rho^2(\xi - a)}{r^3}dr,$$

with similar expressions for $d\eta'$ and $d\zeta'$. Squaring and adding,

$$\begin{aligned} ds'^2 &= d\xi'^2 + d\eta'^2 + d\zeta'^2 \\ &= \frac{\rho^4}{r^4}(d\xi^2 + d\eta^2 + d\zeta^2) - \frac{4\rho^4}{r^5}dr[(\xi - a)d\xi + (\eta - b)d\eta + \\ &\qquad (\zeta - c)d\zeta] + \frac{4\rho^4}{r^6}dr^2[(\xi - a)^2 + (\eta - b)^2 + (\zeta - c)^2] \\ &= \frac{\rho^4}{r^4}ds^2 - \frac{4\rho^4}{r^5}dr[rdr] + \frac{4\rho^4}{r^6}dr^2[r^2] = \frac{\rho^4}{r^4}ds^2, \end{aligned}$$

or

$$ds' = \frac{\rho^2}{r^2}ds. \tag{5}$$

The magnification depends only upon the position of the element and not upon its direction. A small triangle in the neighborhood of a point will have all its sides multiplied by the same quantity; hence, it is transformed into a similar triangle. Its angles are unaltered in magnitude. In other words, *inversion in a sphere is a conformal transformation.*

We consider now the transform of a sphere or plane Σ:

$$A(\xi^2 + \eta^2 + \zeta^2) + B\xi + C\eta + D\zeta + E = 0.$$

This is a sphere (possibly of imaginary radius) if $A \neq 0$; a plane if $A = 0$. This can be written equally well—thinking of the origin as translated to (a, b, c)—in the form

$$A[(\xi - a)^2 + (\eta - b)^2 + (\zeta - c)^2] + B'(\xi - a) + C'(\eta - b) + D'(\zeta - c) + E' = 0,$$

or

$$Ar^2 + B'(\xi - a) + C'(\eta - b) + D'(\zeta - c) + E' = 0. \tag{6}$$

Then, from (2) and (4), the transform is

$$A\frac{\rho^4}{r'^2} + B'\frac{\rho^2}{r'^2}(\xi - a) + C'\frac{\rho^2}{r'^2}(\eta' - b) + D'\frac{\rho^2}{r'^2}(\zeta' - c) + E' = 0,$$

or

$$E'r'^2 + B'\rho^2(\xi - a) + C'\rho^2(\eta' - b) + D'\rho^2(\zeta' - c) + A\rho^4 = 0. \tag{7}$$

This is the equation of a sphere or plane according as $E' \neq 0$ or $E' = 0$. We have proved that *by inversion in a sphere a sphere or plane is carried into a sphere or plane.*

In (6) $E' = 0$, if Σ passes through the center of inversion. Hence, a sphere or plane through the center of inversion is transformed into a plane. This could have been foreseen geometrically.

Since a circle or straight line is the intersection of two spheres or two planes, and these latter are carried into spheres or planes, it follows that *by an inversion in a sphere a circle or straight line is carried into a circle or straight line.* The transform is a straight line if, and only if, the original circle or straight line passes through the center of inversion.

Finally, *if P and Q are two points inverse with respect to a sphere Σ, the transformed points P_1 and Q_1 are inverse with respect to the transformed sphere Σ_1.* This is most easily established geometrically. Through P and Q pass a family of spheres orthogonal to Σ. The transforms of these spheres pass through P_1 and Q_1 and are orthogonal to Σ_1—since angles are preserved. On inverting in Σ_1, each of these spheres is transformed into itself. The point P_1, common to all these spheres, is transformed into the second common point Q_1. Hence, P_1 and Q_1 are inverse with respect to Σ_1.

If Σ_1 is a plane, two points P_1 and Q_1 such that all spheres through them are orthogonal to the plane must be equidistant from the plane and on a common perpendicular to it. That is, Q_1 is the reflection of P_1 in Σ_1. We thus extend the idea of inversion to the case in which the sphere of inversion is a plane. *Inversion in a plane is merely a reflection in a plane.* This transformation possesses the various properties of inversion in a sphere which are italicized in this section.

52. Stereographic Projection.—We now introduce a widely used method of representing the complex z-plane on a sphere. Let the complex plane be a plane in three dimensional space, and let an inversion be made in a sphere whose center K does not lie in the z-plane. Then the z-plane is carried into a sphere which passes through K. The points of the z-plane and the points of the resulting sphere correspond in a one-to-one manner. Corresponding points lie on a line through K. Corresponding angles are equal. This correspondence between the points of the plane and of the sphere is known as a *stereographic projection.*

For convenience we shall so choose the sphere of inversion that the transform of the z-plane is the unit sphere with center at the origin. Let the space axes be so placed that the ξ- and η-axes coincide with the real and imaginary axes in the z-plane. Let K be the point $(0, 0, 1)$; and let the radius of S, the sphere of inversion, be $\sqrt{2}$. Then S passes through the unit circle Q_0 with center at the origin in the z-plane. The points of Q_0 remain fixed on making the inversion. The point at infinity in the z-plane inverts into K, the center of S. The z-plane then inverts into a sphere Σ_0 through Q_0 and the point $(0, 0, 1)$. Hence, Σ_0 is the unit sphere

$$\xi^2 + \eta^2 + \zeta^2 = 1. \tag{8}$$

Corresponding points of the z-plane and of Σ_0 lie on a line through $(0, 0, 1)$. The interior of Q_0 is transformed into the lower half of Σ_0; the exterior into the upper half. Circles and straight lines in the z-plane are transformed into circles on Σ_0.

The equations connecting a point $z = x + iy$ and the corresponding point (ξ, η, ζ) of Σ_0 can be written down at once from (4). Putting $(x, y, 0)$ for (ξ, η, ζ) and (ξ, η, ζ) for (ξ', η', ζ') we have, since

$$r^2 = x^2 + y^2 + 1 = z\bar{z} + 1$$

and $\rho^2 = 2$, the equations

$$\xi = \frac{2x}{z\bar{z} + 1}, \quad \eta = \frac{2y}{z\bar{z} + 1}, \quad \zeta = \frac{z\bar{z} - 1}{z\bar{z} + 1}. \tag{9}$$

Interchanging primed and unprimed values in (4) and using the equation

$$r'^2 = \xi^2 + \eta^2 + (\zeta - 1)^2 = \xi^2 + \eta^2 + \zeta^2 - 2\zeta + 1 = 2(1 - \zeta),$$

we have

$$x = \frac{\xi}{1 - \zeta}, \quad y = \frac{\zeta}{1 - \zeta}, \quad z = \frac{\xi + i\eta}{1 - \zeta}. \tag{10}$$

53. Rotations of the Sphere.—If Σ_0 be transformed into itself in a directly conformal manner, the corresponding points of the z-plane undergo a one-to-one and directly conformal transformation. That is, the z-plane is subjected to a linear transformation. We shall be particularly interested in the rigid motions of space which carry Σ_0 into itself. Such a motion is, obviously, one-to-one and preserves angles.

The most general rigid motion which carries a sphere into itself is a rotation about an axis through the center. Thus, in a rigid motion carrying Σ_0 into itself, there is at least one fixed point P_1 on Σ_0, since the corresponding linear transformation of the z-plane has a fixed point. The center O of Σ_0 is also fixed in rigid motion. Hence, the line OP_1 is a fixed axis; and the motion is a rotation.

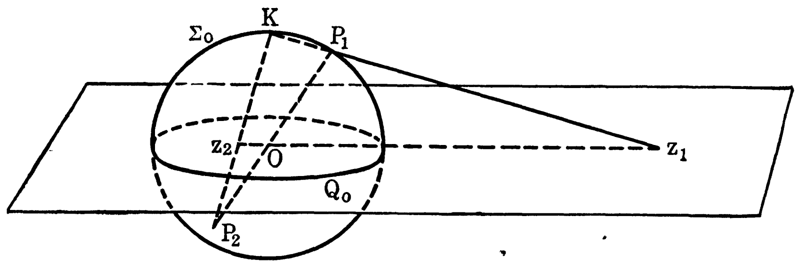


FIG. 35.

What is the character of the resulting linear transformation in the plane? Let P_1 and P_2 be the ends of the axis of rotation (Fig. 35). A great circle through P_1 and P_2 is rotated into another such circle making an angle θ (the angle through which the sphere is rotated) with the original circle. Any circle lying in a plane perpendicular to P_1P_2 is carried into itself. This latter circle is orthogonal to the circles through P_1 and P_2. We now make the stereographic projection. P_1 and P_2 project into two points z_1, z_2 which remain fixed as the sphere is rotated. The circles through P_1 and P_2 project into circles through z_1 and z_2. Then, each circle through z_1 and z_2 is carried into another such circle making an angle θ with its former position. The fixed circles on the sphere project into fixed circles on the plane orthogonal to the circles through z_1 and z_2. This arrangement of the fixed circles shows that the transformation is (Fig. 7) an *elliptic transformation.*

Conversely, any elliptic transformation of the z-plane whose fixed points are the projections of the ends of a diameter corresponds to a rotation of Σ_0. For, we can find a rotation of Σ_0

whose corresponding transformation of the z-plane has the same fixed points as the given transformation and which turns a circle through the fixed points through the same angle. The given transformation and the transformation corresponding to the rotation have the same fixed points and the same multiplier, so they are identical.

We now find the relation between two points z_1, z_2 in order that they be projections of two points P_1, P_2 which are at opposite ends of a diameter of Σ_0. If (ξ_1, η_1, ζ_1) are the coordinates of P_1, then $(-\xi_1, -\eta_1, -\zeta_1)$ are the coordinates of P_2. So, from (10),

$$z_1 = \frac{\xi_1 + i\eta_1}{1 - \zeta_1}, \quad z_2 = -\frac{\xi_1 + i\eta_1}{1 + \zeta_1}.$$

Forming the conjugate of z_1 and multiplying, we have, using (8),

$$\bar{z}_1 z_2 = -\frac{\xi_1^2 + \eta_1^2}{1 - \zeta_1^2} = -1,$$

or

$$z_2 = -\frac{1}{\bar{z}_1}. \tag{11}$$

Conversely, if (11) is satisfied the points on Σ_0 corresponding to z_1 and z_2 lie at opposite ends of a diameter.

We shall now prove the following theorem:

THEOREM 1.—*The necessary and sufficient condition that a transformation of the z-plane correspond to a rotation of Σ_0 is that it be of the form*

$$z' = \frac{az - \bar{c}}{cz + \bar{a}}, \quad a\bar{a} + c\bar{c} = 1. \tag{12}$$

Let $z' = (az + b)/(cz + d)$, $ad - bc = 1$, correspond to a rotation of Σ_0. The rotation has the property that points at the ends of a diameter remain at the ends of a diameter after the rotation. So, if z_1 and z_2 satisfy (11) so also do the transformed points $(az_1 + b)/(cz_1 + d)$ and $(az_2 + b)/(cz_2 + d)$. We have, then,

$$\frac{az_2 + b}{cz_2 + d} = -\frac{\bar{c}\bar{z}_1 + \bar{d}}{\bar{a}\bar{z}_1 + \bar{b}} = \frac{\bar{d}z_2 - \bar{c}}{-\bar{b}z_2 + \bar{a}},$$

the last expression resulting from setting $\bar{z}_1 = -1/z_2$ in the preceding. From this identity in z_2, we have

$$\frac{a}{\bar{d}} = -\frac{b}{\bar{c}} = -\frac{c}{\bar{b}} = \frac{d}{\bar{a}} = \lambda,$$

or

$$a = \lambda\bar{d}, \quad b = -\lambda\bar{c}, \quad c = -\lambda\bar{b}, \quad d = \lambda\bar{a}.$$

From these equations, we have

$$ad - bc = \lambda(\bar{a}\bar{d} - \bar{b}\bar{c}); \text{ or } 1 = \lambda.$$

Hence, $b = -\bar{c}$ and $d = \bar{a}$, and the transformation has the form (12).

We now assume that the transformation has the form (12) and prove that it corresponds to a rotation of the sphere. In the first place, (12) is elliptic, unless it is the identical transformation. If $c \neq 0$, we have $|a| < 1$, so that $|a + \bar{a}| < 2$. Since $a + \bar{a}$ is real, this is the condition that the transformation be elliptic (Theorem 15, Sec. 10). If $c = 0$, we have $|a| = 1$; then, $|a + \bar{a}| < 2$, unless $a = \pm 1$. In this latter case, we have the identity $z' = z$.

Again, the fixed points of (12) satisfy (11). Let z_1 be one of the fixed points; then z_1 satisfies the equation

$$cz_1^2 + (\bar{a} - a)z_1 + \bar{c} = 0.$$

Taking conjugates

$$\bar{c}\bar{z}_1^2 + (a - \bar{a})\bar{z}_1 + c = 0;$$

whence, dividing by $\bar{z}_1^2$,

$$\frac{c}{\bar{z}_1^2} + (\bar{a} - a)\left(-\frac{1}{\bar{z}_1}\right) + \bar{c} = 0.$$

That is, $-1/\bar{z}_1$ is also a root of the equation determining the fixed points. The fixed points thus satisfy (11), and, hence, are the projections of the ends of a diameter. It follows that the transformation corresponds to a rotation of the sphere.

54. Groups of the Regular Solids.—We propose, now, to construct finite groups of linear transformations by forming finite groups of rotations of the sphere Σ_0 and projecting stereographically on the z-plane. We proceed as follows: Let one of the regular solids,—a cube, for instance, or a regular tetrahedron—be placed with its center at the center of Σ_0. There exist certain rigid motions of space which carry the regular solid into itself. These motions, which interchange the faces in various ways, leave the center of the figure fixed. Hence, each such motion is a rotation about an axis through the center. By these motions, the sphere Σ_0 undergoes certain rotations about diametral axes; and, by stereographic projection, the z-plane is subjected to corresponding linear transformations.

It is clear that the set of all rigid motions which carry a body into itself constitutes a group of rigid motions. For, the succession of two such motions, as well as the inverse of any, is a rigid motion which also leaves the body invariant, and, hence, belongs to the set. The set of linear transformations of the z-plane corresponding to the group of rotations carrying the regular solid into itself constitutes a group of linear transformations isomorphic with the group of rotations.

It is evident that the number of rotations which carry a regular solid into itself is finite. There is but a finite number of ways, for example, in which a given face can be made to coincide with itself or with some other face of the solid. Hence, in the z-plane we have groups containing a finite number of linear transformations.[1]

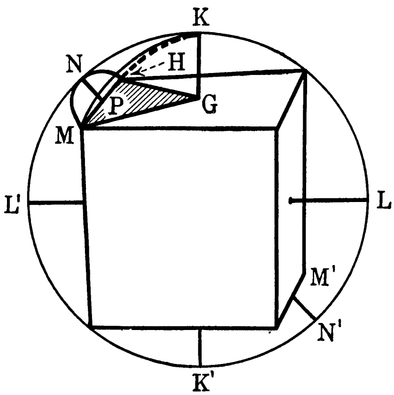


FIG. 36.

55. A Study of the Cube.—We begin with the most familiar of the regular solids—the cube. Let a cube be placed with its center at the origin (Fig. 36) and with its edges parallel to the coordinate axes. We may suppose that the cube is inscribed in Σ_0. We study the axes about which the cube can be rotated into itself. These are of three kinds.

There are, first, the axes joining midpoints of opposite faces as KK' in the figure. There are three such axes. Rotations about each of these axes through angles of 90, 180, and 270 degrees carry the cube into itself. There are, thus, nine rotations about the three axes which carry the cube into itself.

Second, there are the axes joining opposite vertices, as MM' in the figure. There are four of these axes. Rotations about each axis through 120 and 240 degrees carry the cube into itself. There are, thus, eight rotations arising from the four axes.

Finally, there are the axes joining the midpoints of opposite edges, as NN' in the figure. There are six of these axes, a rotation about each of them through 180 degrees carries the cube into itself.

[1] All these groups are treated in KLEIN'S "Vorlesungen über das Ikosaeder."

From the three types of axes we have $9 + 8 + 6 = 23$ rotations. To these we add the identical transformation; that is, the case of no rotation. *The group of rotations which carry the cube into itself consists of* 24 *rotations.*

The corresponding group in the z-plane consists of 24 linear transformations. The setting up of these transformations is a mere matter of algebra. We have, first, to find the two points where an axis of rotation meets Σ_0. We next project the two points on the z-plane (Equation (10)). Let z_1, z_2 be the projected points. We then set up the transformation

$$\frac{z' - z_1}{z' - z_2} = e^{i\theta}\,\frac{z - z_1}{z - z_2}, \tag{13}$$

where θ is the angle through which the cube is rotated.

We consider a few examples from Fig. 36. The points $L(1, 0, 0)$ and $L'(-1, 0, 0)$ project into $z_1 = 1$ and $z_2 = -1$. The angle θ is $\pi/2$, π, or $3\pi/2$; so $e^{i\theta} = i, -1, -i$. The three transformations corresponding to the rotations about LL' are

$$\frac{z' - 1}{z' + 1} = k\,\frac{z - 1}{z + 1}, \qquad k = i, -1, -i.$$

The points $M(-1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3})$ and $M'(1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3})$ project into, from (10), $z_1 = -\dfrac{1 + i}{\sqrt{3} - 1}$, $z_2 = \dfrac{1 + i}{\sqrt{3} + 1}\cdot$ Here $\theta = 2\pi/3$ or $4\pi/3$; $e^{i\theta} = \frac{1}{2}(-1 + i\sqrt{3})$ or $\frac{1}{2}(-1 - i\sqrt{3})$. On substituting in (13) we have the desired transformations.

The points $N(-1/\sqrt{2}, 0, 1/\sqrt{2})$ and $N'(1/\sqrt{2}, 0, -1/\sqrt{2})$ project into $z_1 = -(\sqrt{2} + 1)$ and $z_2 = \sqrt{2} - 1$. Here $\theta = \pi$ and $e^{i\theta} = -1$. On substituting into (13) and simplifying we have

$$z' = S(z) = -\frac{z - 1}{z + 1}\cdot \tag{14}$$

The points K and K' project into ∞ and 0, respectively. The corresponding transformations in the plane are the rotations $z' = kz$, where $k = i, -1, -i$. The three rotations are powers of the transformation

$$z' = T(z) = iz. \tag{15}$$

It is evident that, if an axis passes through the midpoints of opposite faces, the corresponding three linear transformations are powers of a single transformation of period *four*. If the axis joins opposite vertices, the corresponding two transformations are powers of a transformation of period *three*. If the axis joins the midpoints of opposite edges the corresponding linear transformation is of period *two*. The fixed points of the elliptic transformations which form the group are thus divided into three sets. All the fixed points of any one of the sets are congruent. Thus, K can be carried into L, K', L', or any point at the end of an axis of the same kind, by means of a rigid motion which carries the cube into itself. Hence, the projections of all these points are congruent. But K cannot be so carried into M or N. Similarly, the points corresponding to the ends of the axes of the kind to which MM' belongs are congruent. The same is true of the points corresponding to N, N' and similar points.

We can form a fundamental region for the group by constructing the isometric circles. We shall, however, proceed in a different manner. Consider the triangle MHG, in the figure, formed by joining the vertices at the ends of an edge to the midpoint of an adjacent face. By a suitable one of the rotations about KK', the triangle can be carried into another such triangle abutting along the side MG, or one abutting along HG. By the rotation about NN', we get a triangle abutting along MH. If we project the triangle MHG on the surface of the sphere, the origin being the center of projection, we have the spherical triangle formed by the arcs of great circles KM, KH, MNH. This spherical triangle is a fundamental region on the sphere for the group of rotations. Its stereographic projection in the z-plane is a fundamental region for the group of linear transformations.

Let z_M, z_N, z_H, z_K $(= \infty)$ be the projections of M, N, H, K. Then the fundamental region has two pairs of congruent sides. The transformation T (Equation (15)) carries $z_K z_H$ into $z_K z_M$; the transformation S (Equation (14)) carries $z_N z_M$ into $z_N z_H$. There are three cycles: z_K constitutes a cycle of angle $\pi/2$; z_N is a cycle of angle π; and z_H and z_M constitute a cycle of angle $2\pi/3$. These cycles give the relations $T^4 = 1$ and $S^2 = 1$, which we know already, and $(ST)^3 = 1$, which is easily verified.

We can construct 23 other triangles on the cube by joining the ends of an edge to the midpoint of an adjacent face. We can carry MHG into any one of these by a suitable rotation. If these 23 triangles be projected on the sphere and then projected stereographically on the z-plane, we have the 23 regions congruent to the fundamental region constructed above.

The transformations S and T connecting congruent sides of the fundamental region are generating transformations for the group. For, by the combinations of S and T, we can construct regions adjacent to the fundamental region, regions adjacent to the new regions, and so on as long as there are any free sides. We thus cover the whole plane; and all the transforms of the fundamental region are accounted for.

The preceding group of 24 linear transformations is known as the "octahedral group." This name is due to the fact that it is the group arising from the regular octahedron. Let the six intercepts of the sphere on the coordinate axes—K, L, K', etc.—be joined by lines to form a regular octahedron (Fig. 37). The octahedron admits the same rotations about KK', LL', etc., as the cube. It is easily seen that an axis, as MM', joining opposite vertices of the cube joins the midpoints of opposite faces of the octahedron; that an axis, as NN', joining the midpoints of opposite edges of the cube also joins the midpoints of opposite edges of the octahedron; and that the octahedron and the cube admit the same rotations about these axes.

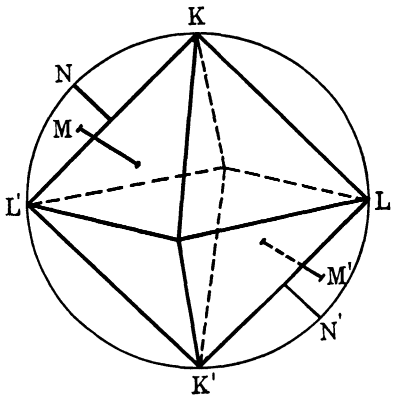


FIG. 37.

56. The General Regular Solid.—Let the number of faces of the solid be F, the number of its vertices be V, and the number of its edges be E. Let μ be the number of edges bounding each face, and let ν faces meet at the vertex.

The number of rigid motions carrying the regular solid into itself is easily found. The solid can be brought into coincidence with itself so that a given edge a_0b_0 is made to coincide with any edge a_ib_i or with b_ia_i. Then *the number of rotations carrying the regular solid into itself is equal to* $2E$.

Let the regular solid be inscribed in Σ_0, and we project stereographically on the plane as in the preceding section. The various fixed points of the linear transformations in the plane are the projections of the ends of the diameters about which the rotations take place. A diameter is an axis about which the solid can be rotated into itself if, and only if, it passes through the midpoint of a face, through a vertex, or through the midpoint of an edge. The number of fixed points of the group is then

$F + V + E$. Using Euler's formula, which applies to any simply connected solid,

$$V + F = E + 2, \tag{16}$$

we have that *the number of fixed points in the group is* $2E + 2$.

The rotations about an axis through the midpoint of a face are all through multiples of an angle $2\pi/\mu$; that is, the rotations are powers of a rotation of period μ. The rotations about an axis through a vertex are powers of a rotation of period ν. The rotation about an axis through the midpoint of a side is of period 2. These statements are true for the corresponding transformations in the plane.

A midpoint of a face can be carried into the midpoint of any other face; a vertex can be carried into any other vertex; a midpoint of an edge can be carried into the midpoint of any other edge by a rotation which carries the solid into itself. Hence, the fixed points of the plane are separated into three sets of congruent points: F fixed points are congruent; V others are congruent; and the remaining E are congruent.

A fundamental region for the group can be got, as in Fig. 36, by joining the vertices at the ends of an edge to the midpoint of an adjacent face to form a triangle; then projecting this triangle on the sphere; and, thence, projecting stereographically on the plane. The transforms of this fundamental region are the like projections of the $2E - 1$ other such triangles which can be drawn on the solid. The congruent sides of this fundamental region are connected by a transformation T of period μ and a transformation S of period 2. S and T are generating transformations for the group. They are connected by the relation $(ST)^\nu = 1$.

In the following table, the various regular solids are listed with the number of their faces, vertices, etc. In the last column N $(=2E)$ is the number of transformations of the group:

TABLE I

	F	V	E	μ	ν	N
Tetrahedron	4	4	6	3	3	12
Cube	6	8	12	4	3	24
Octahedron	8	6	12	3	4	24
Dodecahedron	12	20	30	5	3	60
Icosahedron	20	12	30	3	5	60
Dihedron	2	n	n	n	2	$2n$

The dihedron is an ideal solid of zero volume. Let a regular polygon of n sides be inscribed in Q_0, the equator of Σ_0. We shall look upon this figure as a regular solid with two coincident faces. The line KK' is an axis about which the figure can be rotated into itself through multiples of the angles $2\pi/n$. Also, any diameter through a vertex or through the midpoint of a side of the polygon is an axis about which the figure can be rotated through the angle π into itself. The group of rotations of this figure also carries into itself the double pyramid formed by joining K and K' to the vertices of the polygon inscribed in Q_0. This latter figure, however, is not a regular solid.

We note the same sort of duality between the dodecahedron and the icosahedron that exists between the cube and the octahedron. If we interchange the numbers F and V and the numbers μ and ν in the tabulated values for the dodecahedron, we have the entries for the icosahedron. It is not difficult to verify the fact that the icosahedron can be so placed as to have precisely the same group of rotations as the dodecahedron. The axes joining the midpoints of the opposite faces of one solid are made the axes joining the opposite vertices of the other.

Two further groups remain to be mentioned. The *four group* corresponds to the four rotations of the sphere Σ_0 which carry the real axis into itself and the imaginary axis into itself. It may be regarded as a limiting case of the dihedral group in which $n = 2$.

There is, finally, the simplest of all the finite groups—the elliptic cyclic group. This group has two non-congruent fixed points. We can give it a geometrical origin, if we like, by considering it as arising from the group of rotations which carries into itself a regular pyramid formed by joining the vertices of a regular polygon inscribed in Q_0 to the point K.

We have found five types of finite groups of linear transformations: the *elliptic cyclic group*, the *dihedral group* (including the four group), the *tetrahedral group*, the *octahedral group*, and the *icosahedral group*. Each of these groups can be transformed after the manner of Sec. 15, to get other finite groups. The possibility of further finite groups will be investigated in the next section. We shall find that there are no others.

57. Determination of All the Finite Groups.—We first prove the following theorem:

THEOREM 2.—*If a group contains two non-parabolic transformations which have one fixed point in common and the other fixed* *points different, then the group contains parabolic transformations with the common point as fixed point.*

Let the group be so transformed that ∞ is the common fixed point and 0 and 1 are the remaining fixed points. The two transformations are then of the form

$$S_1 = K_1 z, \quad S_2 = K_2(z - 1) + 1, \quad K_1 \neq 1, K_2 \neq 1.$$

Combining these in the product given below, we find readily

$$U = S_1^{-1}S_2S_1S_2^{-1} = z + \frac{(K_1 - 1)(K_2 - 1)}{K_1}.$$

This is a translation; that is, a parabolic transformation with ∞ as fixed point. Any power of this transformation U^n, $n \neq 0$, is parabolic with ∞ as fixed point.

A finite group has only elliptic transformations. It follows from the preceding theorem that all transformations of the group which have one fixed point in common have the second fixed point in common, also. If ν transformations (including the identical transformation) have a common fixed point, e, we shall call e a fixed point of *order* ν. These transformations form a cyclic subgroup; and the transformations are powers of a transformation with multiplier $e^{2\pi i/\nu}$. Congruent fixed points are of the same order. Thus, if $e' = T(e)$; and S_i, $i = 1, 2, \ldots$ are the transformations with e as fixed point, then the ν transformations TS_iT^{-1}, and no others, have e' as fixed point.

Given a finite group of $N(>1)$ linear transformations, so transformed, if necessary, that ∞ is not a fixed point for any transformation. Let R be the fundamental region exterior to all the isometric circles. We shall investigate the cycles of the region.

Let z_0 be a vertex belonging to a cycle of angle $2\pi/k$, $k > 1$. Then we found in Sec. 26 that z_0 is a fixed point of an elliptic transformation with multiplier $e^{2\pi i/k}$. There is no transformation with z_0 as fixed point and with multiplier $e^{i\theta}$ where $0 < \theta < 2\pi/k$; so z_0 is a fixed point of order k.

Consider now a fixed point e of order ν lying outside R. There is a point e' congruent to e lying within or on the boundary of R. But e' is the fixed point of an elliptic transformation of order ν, and has isometric circles passing through it. Hence, e' is a vertex of R. That is, *each fixed point of order ν is congruent to a vertex of R belonging to a cycle of angle $2\pi/\nu$.*

Let a, b be finite inner points of R. We form the function

$$f(z) = \frac{(z-a)(z_1-a)\cdots(z_{N-1}-a)}{(z-b)(z_1-b)\cdots(z_{N-1}-b)} \tag{17}$$

where $z_1, z_2, \ldots z_{N-1}$ are congruent to z. This function is automorphic with respect to the given group, since if z be transformed into z_i the variables $z, \ldots, z_{N-1}$ are permuted and $f(z)$ is unaltered. It has zeros of the first order at a and the points congruent thereto and poles of the first order at b and congruent points. It has a single pole in R; hence, it takes on every value once in R (Sec. 42, Theorem 11).

The proposition that an automorphic function takes on every value the same number of times in the fundamental region requires that we count the values in a particular way at the vertices. Let $f(z_0) = A$; and let $f(z)$ take on the value A r times. We count the value A as taken on r/k times in the vertices of the cycle (Sec. 42 (2)). In the present case, $r/k = 1$; whence, $f(z)$ takes on the value A exactly k times at z_0. The derivative, $f'(z)$, has $k-1$ zeros at z_0.

Now let us look upon $f(z)$ as a rational function in the whole plane. It has N poles; so it takes on each value N times. It takes on the value A only at the vertex z_0 and the points congruent thereto; and at each such points it takes on the value k times. It follows that the point z_0 and the points congruent to it form a set of N/k congruent points. Since any fixed point is congruent to a vertex of R, we can state this result as follows: *A fixed point of order ν belongs to a set of N/ν congruent fixed points. The order of a fixed point is a submultiple of N.*

Consider $f'(z)$. Since $f(z)$ has N poles of the first order, its derivative has N poles of the second order. Hence, $f'(z)$ has $2N$ zeros. The roots of $f'(z)$ can all be located. At a finite point which is different from a fixed point, $f'(z) \neq 0$. Otherwise, $f(z)$ would take on its value twice at the point; this is contrary to the fact that $f(z)$ takes on no value more than once in R or in any of the regions congruent to R which cover the plane. At infinity we have developments of the form

$$f(z) = c_0 + \frac{c_1}{z} + \cdots; \qquad f'(z) = -\frac{c_1}{z^2} + \cdots$$

Here, $c_1 \neq 0$, since $f(z)$ takes on the value c_0 only once at infinity. Hence, $f'(z)$ has a zero of the second order at infinity. The remaining zeros of $f'(z)$, $2N - 2$ in number, are at the fixed points.

Let the fixed points of the group fall into s sets of congruent points. Let the orders of the points of the sets be $\nu_1, \nu_2, \ldots, \nu_s$. Then the sets consist of $N/\nu_1, N/\nu_2, \ldots N/\nu_s$ points, respectively. Summing the zeros of $f'(z)$ and equating to $2N - 2$, we have

$$\sum_{i=1}^{s} \frac{N}{\nu_i}(\nu_i - 1) = 2N - 2,$$

or

$$\sum_{i=1}^{s}\left(1 - \frac{1}{\nu_i}\right) = 2 - \frac{2}{N}. \tag{18}$$

The integers N and ν_i for any finite group must satisfy (18). Here $N \geqslant 2$ and $\nu_i \geqslant 2$; also N/ν_i is an integer. We have $s > 1$; for is $s = 1$, the first member of (18) is less than 1 and the second is greater than or equal to 1. Also, $s < 4$; for $1 - \frac{1}{\nu_i} \geqslant \frac{1}{2}$ and the first member equals or exceeds $s/2$, which is greater than the second member if $s \geqslant 4$. There are then two cases to consider $s = 2$ and $s = 3$.

If $s = 2$, (18) becomes

$$1 - \frac{1}{\nu_1} + 1 - \frac{1}{\nu_2} = 2 - \frac{2}{N}, \text{ whence, } \frac{N}{\nu_1} + \frac{N}{\nu_2} = 2.$$

We have $N/\nu_1 = N/\nu_2 = 1$. Hence, N may be any integer and $\nu_1 = \nu_2 = N$. The two sets of congruent fixed points contain one point each; and each fixed point is of order N. The groups satisfying these conditions, obviously, consist of the cyclic groups of N transformations.

If $s = 3$, (18) reduces to

$$\frac{1}{\nu_1} + \frac{1}{\nu_2} + \frac{1}{\nu_3} = 1 + \frac{2}{N}. \tag{19}$$

Let the subscripts be so chosen that $\nu_1 \leqslant \nu_2 \leqslant \nu_3$. Then $\nu_1 = 2$, for, otherwise, the first member does not exceed 1. Then,

$$\frac{1}{\nu_2} + \frac{1}{\nu_3} = \frac{1}{2} + \frac{2}{N}. \tag{20}$$

Then, $\nu_2 = 2$ or $\nu_2 = 3$; otherwise, the first member does not exceed $\frac{1}{2}$.

If $\nu_2 = 2$, we have $\nu_3 = N/2$. If N is any even number, $N = 2n$, the equation can be satisfied.

If $\nu_2 = 3$, we have, from (20),

$$\frac{1}{\nu_3} = \frac{1}{6} + \frac{2}{N} \tag{21}$$

whence, $\nu_3 < 6$. We find that each of the possible values $\nu_3 = 3, 4, 5$ gives an integral value of N; namely $N = 12, 24, 60$, respectively.

We have found then that, except for the cyclic group ($s = 2$), all finite groups have three sets of congruent fixed points. The orders of the points, the numbers in each set, and the total number of transformations are only such as appear in the following table:

TABLE II

ν_1	ν_2	ν_3	N/ν_1	N/ν_2	N/ν_3	N
2	2	n	n	n	2	$2n$
2	3	3	6	4	4	12
2	3	4	12	8	6	24
2	3	5	30	20	12	60

That there actually exist finite groups corresponding to the possibilities set forth in Table II is evident on an inspection of Table I. In that table, there are three sets of congruent fixed points containing F, V, and E points. The corresponding orders are μ, ν, and 2, respectively. The first possibility in Table II is realized in the dihedral group (in the four group, if $n = 2$), the second in the tetrahedral group, the third in the octahedral group, and the last in the icosahedral group.

THEOREM 3.—*There are no finite groups of linear transformations other than the elliptic-cyclic groups, the groups of the regular solids (including the four group), and the transforms of these latter groups by means of linear transformations.*

We have found that when $s = 2$ the only groups arising are the cyclic groups. We shall now show that any group with three sets of congruent fixed points is the transform of a group arising from one of the regular solids. We shall prove (which amounts to the same thing) that if two groups have the same values of ν_1, ν_2, ν_3 one is the transform of the other.

Given two finite groups with the same values of ν_1, ν_2, ν_3. Let each be so transformed, if necessary, that ∞ is not a fixed point. If the transformed groups are the transforms of one another, the same is true of the original groups. Let the group S consist of the transformations $S_1, \ldots, S_N$ and have the fixed points $a_1, a_2, \ldots, a_{N/\nu_1}$ of order ν_1, the fixed points $b_1, b_2, \ldots, b_{N/\nu_2}$ of order ν_2, and the fixed points $c_1, c_2, \ldots, c_{N/\nu_3}$ of order ν_3. Let the group S' have the transformations $S_1', \ldots, S_N'$, the fixed points $a_1', \ldots, a'_{N/\nu_1}$ of order ν_1, the fixed points $b_1', \ldots, b'_{N/\nu_2}$ of order ν_2, and the fixed points $c_1', \ldots, c'_{N/\nu_3}$ of order ν_3.

We now use formula (17) to set up an automorphic function $f(z)$ for the group S and an automorphic function $f_1(z)$ for the group S' where the constants a and b in the formula are different from the fixed points. Let $f(a_i) = A$, $f(b_i) = B$, $f(c_i) = C$; $f_1(a_i') = A', f_1(B_i') = B', f_1(c_i') = C'$. Let $t = (at_1 + b)/(ct_1 + d)$ be the transformation carrying the distinct points $t_1 = A'$, B', C' into the distinct points A, B, C, respectively, and form the function

$$F(z) = \frac{af_1(z) + b}{cf_1(z) + d}.$$

Then, $F(a_i') = A$, $F(b_i') = B$, $F(c_i') = C$. The functions $F(z)$ and $f(z)$ have N poles each, and, hence, take on any value N times.

We shall represent the group S in the z-plane and S' in a second or z'-plane. We shall set up a correspondence between the points of the two planes by means of the equation[1]

$$F(z') = f(z). \tag{22}$$

To each value of z there correspond N values of z'; for a given z determines one value of $f(z)$ and this value is taken on by $F(z')$ at N points z'. Likewise, to each z' correspond N values of z.

If $z \neq a_i$, b_i, c_i then $f(z)$ is distinct from A, B, C and the corresponding values of z' are N distinct points in the z'-plane.

If $z = a_i$, then $f(z) = A$ and the corresponding values of z' fall into N/ν_1 sets of ν_1 equal values each; namely, $z' = a_1'$, a_2', . . . Consider the arrangement of the ν_1 values of z' in the neighborhood of a_j' when z is in the neighborhood of a_i. We have the following developments in the neighborhoods of

[1] This correspondence was suggested by Prof. P. Koebe.

the points in question, since each function takes on its value ν_1 times:

$$f(z) = A + k(z - a_i)^{\nu_1} + \cdots, \qquad k \neq 0.$$
$$F(z') = A + k'(z' - a_j')^{\nu_1} + \cdots, \qquad k' \neq 0.$$

Then, (22) becomes

$$(z' - a_j')^{\nu_1}(k' + \cdots) = (z - a_i)^{\nu_1}(k + \cdots).$$

Extracting the ν_1-th root,

$$(z' - a_j')(k'^{1/\nu_1} + \cdots) = \epsilon_\kappa(z - a_i)(k^{1/\nu_1} + \cdots),$$
$$\kappa = 1, 2, \cdots, \nu_1,$$

where $\epsilon_1, \ldots, \epsilon_{\nu_1}$ are the ν_1-th roots of unity. From these we have the ν_1 developments in the neighborhood of a_i,

$$z' = a_j' + \epsilon_\kappa\left(\frac{k}{k'}\right)^{1/\nu_1}(z - a_i) + \cdots$$

Hence, although ν_1 values of z' become equal at a_i, these values belong to distinct branches. Similar reasoning applies to b_i and c_i.

The relation (22), then, gives in the neighborhood of any point of the z-plane N distinct function elements. These elements combine to form N single-valued functions of z. Otherwise, since there are no branch points, the N sheets bearing the values of z' are not connected. Consider one of the functions, $z' = T(z)$. To each value of z there corresponds one and only one value of z'. Interchanging the rôles of z and z', to each z' there corresponds one, and only one, z. Hence, $z' = T(z)$ is a linear transformation (Sec. 1, Theorem 3, Corollary 1).

The transformation T carries the fixed points of the group S into the fixed points of the group S'; since when $z = a_i, b_i, c_i$ then $T(z) = a_i', b_i', c_i'$. Then the group $U_i = TS_iT^{-1}$ is a group with the same fixed points and of the same orders as the group S'. Is the group U the same as the group S'?

Let U_i and S_j' be transformations with the fixed point a_j' and the same multiplier. We shall show that the second fixed points are identical. Suppose they are different. We combine U_i and S_j' in all possible ways to generate a group Γ. A point a_i' $(\neq a_j')$ is carried by each of the transformations into another of the points a_k'; that is, a_i' has a finite number of distinct transforms by the group Γ. But Γ contains parabolic transformations with a_j' as fixed point (Theorem 2); and by repetitions of such a parabolic transformation q_i' has an infinite number of distinct transforms. This contradiction proves that U_i and S_i' have the same fixed points; whence they are identical. We thus identify each transformation of U with a transformation of S'. The group S' is the transform of S by a linear transformation, which was to be proved.

The Polyhedral Functions.—The automorphic functions belonging to the finite groups, which, owing to their connection with the regular solids, are called "polyhedral functions," are readily set up. The function $f(z)$ in (17) takes on each value once in the fundamental region. Hence (Sec. 43, Theorem 14), the most general simple automorphic function belonging to the group is a rational function of $f(z)$.

58. The Extended Groups.—Intimately connected with the rotations which carry a regular solid into itself are the reflections

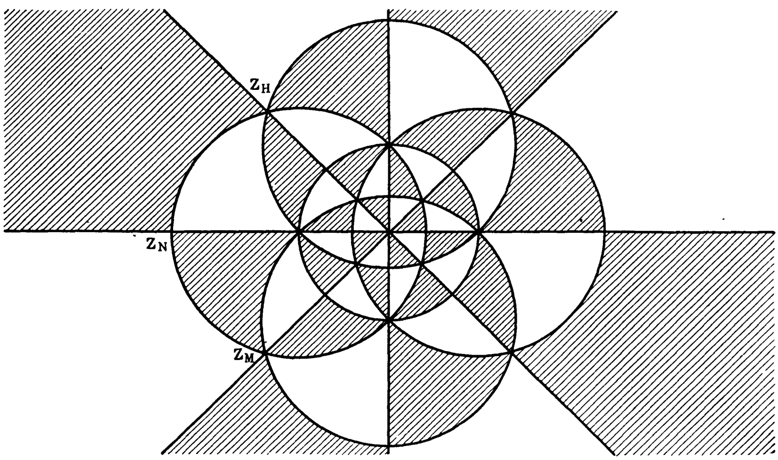


Fig. 38

in planes of symmetry which, likewise, carry the solid into itself. In the cube, Fig. 36, there are nine planes of symmetry: three planes through the midpoints of four faces (the coordinate planes), and six planes containing opposite edges. These planes intersect the sphere Σ_0 in nine great circles. These circles are projected stereographically on the z-plane in Fig. 38. The real and imaginary axes and the unit circle Q_0 arise from the first three planes of symmetry mentioned above; the remaining two straight lines and four circles arise from the latter six planes.

What is the transformation of the z-plane corresponding to a reflection in a diametral plane of the points of Σ_0? On making the inversion which projects Σ_0 stereographically on the z-plane, a diametral plane Π through a great circle C is carried into a sphere Σ through C', the stereographic projection of C. Since Π is orthogonal to Σ_0, Σ is orthogonal to the z-plane and has C' as equator. Two points on Σ_0 inverse with respect to Π are carried into two points of the z-plane inverse with respect to Σ and, hence, inverse with respect to C'. Hence, if the points of Σ_0 be transformed by a reflection in Π the corresponding points of the z-plane undergo an inversion in C'. Reflection in the nine planes of symmetry of the cube, then, correspond to inversions in the nine circles of Fig. 38.

The planes of symmetry divide the surface of the cube into 48 equal triangles. One such is the triangle GHP formed by bisecting the triangle MHG by a line from G to the midpoint of MH. This triangle projects on the triangle KNH on the sphere and the latter projects stereographically on the shaded triangle in the upper left hand corner of Fig. 38. The triangle GPM corresponds to the unshaded region in the lower left-hand corner of Fig. 38; and the remaining triangles on the cube correspond to the remaining triangles of Fig. 38.

Now, let us form a group by combining the reflections in planes of symmetry in all possible ways. Each such reflection carries GHP into one of the 48 triangles on the surface of the cube. Further, by suitable sequences of such reflections, the whole surface of the cube can be covered; that is, GHP can be carried into any other triangle. It is easily seen that two such sequences which carry GHP into the same triangle transform all points in the same way and so are identical. The group thus contains exactly 48 transformations. In the z-plane the corresponding transformations carry a given triangle of Fig. 38 into each of the triangles of the figure.

This group in the z-plane is known as the "extended group." It is found by combining in all ways the inversions in the nine circles in Fig. 38. Each triangle is a fundamental region for the group. The transformations are of two kinds. An even number of inversions is a linear transformation; an odd number is an inversely conformal transformation of the plane into itself, and has the form $z' = (a\bar{z} + b)/(c\bar{z} + d)$ (Sec. 5, Equation (23)). There are 24 transformations of each kind. The former correspond to the rotations of the sphere which carry the cube into itself, since an even number of reflections in diametral planes is equivalent to a rigid motion. Among the inversely conformal transformations are the nine inversions from which we formed the group.

If we shade the alternate triangles in Fig. 38, we note that a transformation which carries a shaded triangle into a shaded triangle or an unshaded triangle into an unshaded triangle is a linear transformation; a transformation which carries a shaded triangle into an unshaded triangle or *vice versa* is a transformation which reverses the sign of the angle. The former 24 transformations constitute the octahedral group. A shaded and an unshaded region together constitute a fundamental region for the group.

It is easy to see that the extended group can be generated by three inversions; namely, in the sides of any one of the triangles of the figure. By repeating these three inversions, any triangle of the figure can be carried into any other.

The groups of the remaining regular solids can be extended in a precisely similar way. We shall not treat the several cases in detail. By repeated reflections in planes of symmetry, a given edge a_0b_0 can be carried into an edge a_ib_i in four ways—so that a_0b_0 coincides with a_ib_i or with b_ia_i, and so that angles are preserved or angles are reversed. The extended group corresponding to the sequence of reflections then consists of $4E$ transformations, where E is the number of edges.

The planes of symmetry divide the surface of the regular solid into $4E$ triangles, each formed by joining the midpoint of a face to the end and the midpoint of an adjacent edge. The projection of these triangles on the sphere and thence on the z-plane gives a system of triangles with properties analogous to those of Fig. 38. From the tetrahedron we get 24 triangles; from the icosahedron 120; from the dihedron $4n$. In each case, half the transformations are linear transformations and half are transformations with reversal of the sign of the angle. The former constitute a subgroup identical with the group of the regular solid.

The extended groups will reappear in a later chapter in the study of the Riemann-Schwarz triangle functions.

II. THE GROUPS WITH ONE LIMIT POINT

59. The Simply and Doubly Periodic Groups.—The groups with one limit point are of particular interest. They comprise the groups whose automorphic functions—the simply and doubly periodic functions—have been studied more than any other automorphic functions. Further, as we shall find later, these groups form one of the three classes of groups whose automorphic functions are employed in the simplest theory of uniformization.

A group with but one limit point can contain no hyperbolic or loxodromic transformation; for each of the two fixed points of such a transformation is a limit point. Each transformation of the group is either parabolic or elliptic. Each transformation has the limit point as fixed point; otherwise, the limit point would have a transform which would be a second limit point. The group cannot consist entirely of elliptic transformations. For, the second fixed points do not all coincide; otherwise, the group would be a finite cyclic group. It follows from Theorem 2 that the group contains parabolic transformations. The group, then, consists either of parabolic transformations alone or of parabolic and elliptic transformations.

Let the group be so transformed that the limit point is carried to infinity. Then all the transformations are of the form

$$z' = Kz + b,$$

where K is the multiplier of the transformation. The parabolic transformations ($K = 1$) are translations; the elliptic transformations are rotations.

If the group contains only translations, it is either a *simply periodic group* with a period strip (Fig. 12) as fundamental region or a *doubly period group* with a period parallelogram (Fig. 13) as fundamental region.

We prove these facts briefly. The translations are of the form $S_i = z + \Omega_i$, Ω_i being called a "period." We have $S_1^{m_1}S_2^{m_2} \cdots S_n^{m_n} = z + m_1\Omega_1 + m_2\Omega_2 + \cdots + m_n\Omega_n$; that is, any linear combination of periods with integral coefficients is a period. Now, let all the periods be plotted in the complex plane. These points have no cluster point. If there were a cluster point, the transformation $S_iS_j^{-1} = z + \Omega_i - \Omega_j$, where Ω_i and Ω_j are taken sufficiently near the cluster point, would have an arbitrarily small period; and the group would be continuous.

Let ω be a nearest period to the origin; that is, one of the finite number at the minimum distance. Let L be the line joining the origin to ω and let Ω be any period lying on L. We can write $\Omega = M\omega$, where M is real. We shall show that M is an integer. Suppose not; and let m be the nearest integer to M. Then $|M - m| \leqslant \frac{1}{2}$; and the period $\Omega' = \Omega - m\omega$ is such that $|\Omega'| = |(M - m)\omega| \leqslant \frac{1}{2}|\omega|$. This is contrary to the assumption that ω is the nearest period to the origin. The transformation $S = z + \Omega$ is then a power of $T = z + \omega$; namely, $S = T^M$. If all the periods lie on L, the group is the simply periodic group generated by T.

Suppose, next, that there are periods not lying on L. Let ω' be at the minimum distance from the origin among these. Then any period in the plane can be written $\Omega = M\omega + M'\omega'$, where M and M' are real. We now prove that M and M' are both integers. Suppose not; and let m and m' be the nearest integers to M, M'. Then $|M - m| \leqslant \frac{1}{2}$, $|M' - m'| \leqslant \frac{1}{2}$. Consider the period

$$\Omega' = \Omega - (m\omega + m'\omega') = (M - m)\omega + (M' - m')\omega' \neq 0.$$

Here we have

$$|\Omega'| < \tfrac{1}{2}|\omega| + \tfrac{1}{2}|\omega'| \leqslant |\omega'|.$$

This is impossible unless Ω' lies on L. But, then, $M' - m' = 0$ and $|\Omega'| \leqslant \frac{1}{2}|\omega|$, which is impossible. It follows that M and M' are integers.

A transformation of the group, $S = z + \Omega$, can then be written in the form $S = T^M T_1^{M'}$, where $T = z + \omega$, $T_1 = z + \omega'$. The group is the doubly periodic group generated by T and T_1.

We can assume, without loss of generality, that a period of smallest absolute value is 1. Let $G = z/\omega$, where ω is a smallest period, and transform the group by G.

$$S = z + \Omega, \quad GSG^{-1} = G(\omega z + \Omega) = z + \frac{\Omega}{\omega}.$$

When $\Omega = \omega$, the period is 1. Otherwise, $|\Omega/\omega| \geqslant 1$, since $|\Omega| \geqslant |\omega|$.

The most general group with one limit point and containing only parabolic transformations is the transform of a simply or doubly periodic group.

60. Groups Allied to the Periodic Groups.[1]—We now consider groups containing parabolic and elliptic transformations. We show first that the multiplier of an elliptic transformation is limited to a small number of possible values. Let $S = z + \Omega$ and $S_1 = Kz + b$ be a parabolic and an elliptic transformation contained in the group. Then the translation

$$S_1SS_1^{-1} = S_1\left(\frac{z - b}{K} + \Omega\right) = z + K\Omega$$

has the period $K\Omega$. Let the minimum period be 1; then, taking $\Omega = 1$, K is a period. Now, the multipliers of the transformations with a common fixed point of order ν are

$$K_\nu(=e^{2\pi i/\nu}),\ K_\nu^2,\ \cdot\cdot\cdot\ ,\ K_\nu^{\nu-1}.$$

[1] For the treatment in this section and the next see KOEBE, P., *Math. Ann.*, vol. 67, pp. 164–168, 1909.

These multipliers—the ν-th roots of unity—lie on the unit circle Q_0.

The possible values of K and the periods ± 1 must be so spaced about Q_0 that the distance between any two distinct periods is not less than 1. Otherwise, there is a period of smaller absolute value than 1, contrary to hypothesis. If $\nu \geqslant 7$, (Fig. 39) we have $|K_\nu - 1| < 1$, which is impossible. If $\nu = 5$, we have $|K_\nu^2 + 1| < 1$, which is impossible. Hence, ν is limited to the values 2, 3, 4, 6. Also $\nu = 4$ and $\nu = 3$, or $\nu = 6$, cannot both appear; since $|K_4 - K_3| < 1$ and $|K_4 - K_6| < 1$.

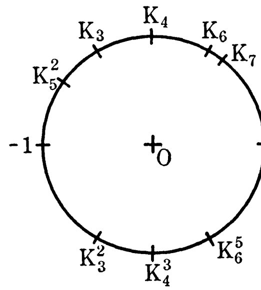


FIG. 39.

Two elliptic transformations with the same multiplier, $S_1 = Kz + b$, $S_2 = Kz + b'$, have constants b, b' which differ by a period. For,

$$S_2S_1^{-1} = K\left(\frac{z - b}{K}\right) + b' = z + b' - b,$$

which has the period $b' - b$. Conversely, if Ω is any period, there is a transformation with the multiplier K and the constant $b + \Omega$; namely, where $S = z + \Omega$,

$$SS_1 = Kz + b + \Omega.$$

Then all transformations with the multiplier K which the group contains, and no others, are comprised in the formula

$$S_i = Kz + b + \Omega_i, \tag{23}$$

where b is the constant for one such transformation and Ω_i is any period or zero. The finite fixed point of (23) we find to be

$$\xi_i = \frac{b + \Omega_i}{1 - K}. \tag{24}$$

In locating the fixed points from (24), it suffices to find those which lie in a period strip or period parallelogram of the subgroup of translations. The remaining fixed points differ from these by periods. For, if U is any transformation of the group, US_iU^{-1} has the same multiplier as S_i and has the fixed point $U(\xi_i)$. In particular, each fixed point has a congruent fixed point lying in the period strip or period parallelogram.

We treat first the case in which $\nu = 2$ for all the fixed points; so $K = -1$. We can transform the group by a translation carrying b to the origin; so we can take $b = 0$. The resulting transformations,

$$S_i = -z + \Omega_i, \qquad T_i = z + \Omega_i,$$

where the T_i form a simply or doubly periodic group, are easily shown to be a group. The fixed points, from (24), are $\xi_i = \Omega_i/2$. The location of the fixed points (marked "2") in a period strip and a period parallelogram are shown in Figs. 40 and 41. The half of the strip or parallelogram whose congruent sides are joined by arrows is a fundamental region for the group. We verify readily that certain transformations carry the region into regions abutting along each of its sides, and that all except the identical transformation carry it outside itself.

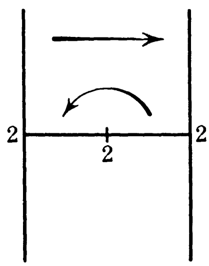


FIG. 40.

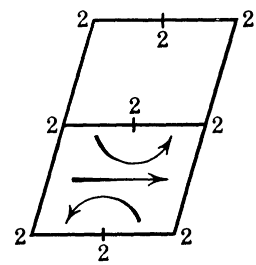


FIG. 41.

If $\nu > 2$, the subgroup of translations cannot be simply periodic. For K_ν is then a period which is not a multiple of 1. Since K_ν is at a distance 1 from the origin and no period can be at a less distance, we may, according to the previous section, take K_ν as the period of the second generating transformation. The doubly periodic subgroup is generated by $T = z + 1$ and $T_1 = z + K_\nu$. The doubly periodic group is, thus, a very special one.

Consider the case $\nu = 4$. We have $K_4 = i$, $K_4^2 = -1 = K_2$, $K_4^3 = -i$. Since $K_4^4 = 1$, we may write the transformations, both elliptic and parabolic, in the form

$$S_i = K_4^n z + \Omega_i = K_4^n z + m + m'i.$$

Both group properties are readily established for these transformations.

For $K = i$ and $K = -1$ we have from (24)

$$\xi_i = \frac{m + m'i}{1 - i} = \frac{(1 + i)(m + m'i)}{2}, \quad \xi_i = \frac{m + m'i}{2}$$

The first formula gives the points marked "4" in the period parallelogram in Fig. 42. The second gives those marked "4" and in addition those marked "2." These latter points are of order 2. The fixed points when $K = -i$ are, of course, the same as those for $K = i$. We can show easily that the square whose congruent sides are connected by arrows is a fundamental region for the group.

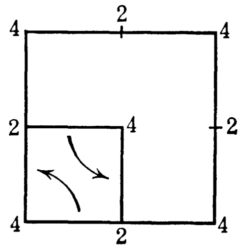


FIG. 42.

The cases $\nu = 3$ and $\nu = 6$ are pictured in Figs. 43 and 44. The group in each case is

$$S_i = K_\nu^n z + m + m'K_\nu, \quad n = 1, 2, \cdots, \nu.$$

The fixed points are located as in the preceding figure. In the latter case there are fixed points of three orders, 2, 3, and 6. In Fig. 43, the parallelogram has been drawn from the periods

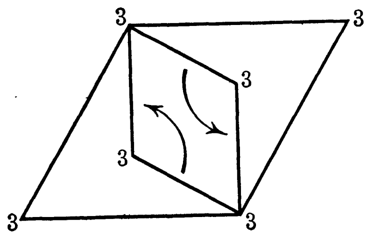


FIG. 43.

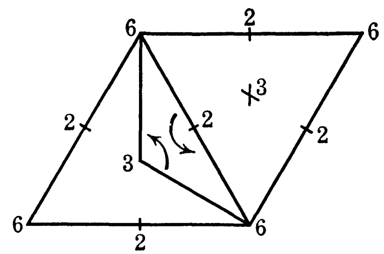


FIG. 44.

1 and $K_3 + 1(=K_6)$; this latter period is at a distance 1 from the origin and may be used, as well as K_3 itself, as the period of the second generating transformation.

We observe that in Figs. 42, 43, and 44, the transformations connecting congruent sides of the fundamental region are elliptic. Each of these groups is generated by rotations.

The groups constructed in this section, together with their transforms by means of linear transformations, include all the groups with a single limit point which contain elliptic transformations.

61. The Automorphic Functions.—We shall now set up all the simple automorphic functions for the groups with one limit point. We suppose the group so transformed that the limit point is at infinity, the smallest period is 1, and an elliptic point of highest order (if elliptic points exist) is at the origin.

For the simply periodic group, the function $e^{2\pi iz}$ is an automorphic function taking on each value once in the period strip. Hence (Sec. 43, Theorem 14), the most general simple automorphic function connected with the group is a rational function of $e^{2\pi iz}$.

In the group of Fig. 40, the function $\cos 2\pi z$ is a simply periodic function which takes on each value twice in the period strip (since it has two zeros, $z = \pm\frac{1}{4}$, in the strip). Now $\cos 2\pi(-z) = \cos 2\pi z$. The function thus admits the elliptic generating transformation; so it is automorphic with respect to the group of the figure. It takes on each value once only in the upper half of the strip, the other value being taken on in the lower half. According to Theorem 14, cited above, the most general simple automorphic function connected with the group is a rational function of $\cos 2\pi z$.

The group of Fig. 41 is treated similarly. The Weierstrassian function

$$\wp(z) = \frac{1}{z^2} + \sum_i{}' \left[\frac{1}{(z - \Omega_i)^2} - \frac{1}{\Omega_i^2} \right], \tag{25}$$

the summation being extended over all non-zero periods Ω_i, is a doubly periodic function with a pole of the second order in the period parallelogram. Hence, $\wp(z)$ takes on each value twice in the period parallelogram. But $\wp(z)$ is an even function, $\wp(-z) = \wp(z)$—as we see on changing z to $-z$ and Ω_i to $-\Omega_i$, which does not alter the set of periods. Then $\wp(z)$ takes on each value once in the fundamental region of the figure. The most general simple automorphic group is a rational function of $\wp(z)$.

For the doubly periodic function itself there is no automorphic function with a single pole in the period parallelogram in terms of which to express all automorphic functions rationally. Let $f(z)$ be a simple automorphic function belonging to the group; that is, an elliptic function. Then $f(z)$ can be written in the form

$$f(z) = \tfrac{1}{2}[f(z) + f(-z)] + \tfrac{1}{2}[f(z) - f(-z)].$$

Here, $f(-z)$ is also an elliptic function. The first term is an even function of z; and so is automorphic with respect to the group of Fig. 41. It is a rational function of $\mathfrak{P}(z)$. The second term is an odd function. Now $\mathfrak{P}'(z)$ is also an odd elliptic function. Hence, $\frac{1}{2}[f(z) - f(-z)]/\mathfrak{P}'(z)$ is an even function and so is a rational function of $\mathfrak{P}(z)$. The most general simple automorphic function is then

$$f(z) = R_1[\mathfrak{P}(z)] + \mathfrak{P}'(z)R_2[\mathfrak{P}(z)],$$

where R_1 and R_2 indicate rational functions.

Consider now the groups of Figs. 42 to 44. The transformations here are of the form

$$z_j = K_\nu{}^n z + \Omega_j, \quad n = 1, 2, \cdots, \nu, \tag{26}$$

where $\nu = 4, 3$, or 6. The function $\mathfrak{P}(z)$ is not automorphic with respect to the groups. We have, from (25),

$$\mathfrak{P}(z_j) = \frac{1}{(K_\nu{}^n z + \Omega_j)^2} + \sum_i{}' \left[\frac{1}{(K_\nu{}^n z + \Omega_j - \Omega_i)^2} - \frac{1}{\Omega_i{}^2}\right]$$
$$= \frac{1}{K_\nu{}^{2n}} \left\{\frac{1}{\left(z + \frac{\Omega_j}{K_\nu{}^n}\right)^2} + \sum_i{}' \left[\frac{1}{\left(z + \frac{\Omega_j - \Omega_i}{K_\nu{}^n}\right)^2} - \frac{1}{\left(\frac{\Omega_i}{K_\nu{}^n}\right)^2}\right]\right\}.$$

Now, the division of a period Ω_i by $K_\nu{}^n$ amounts to rotating it through the angle $-2\pi n/\nu$ about the origin and carries it in each of the three cases into another period Ω_i'. Such a rotation carries the set of all periods into itself. We have, then,

$$\mathfrak{P}(z_j) = \frac{1}{K_\nu{}^{2n}} \left\{\frac{1}{(z + \Omega_j')^2} + \sum_i{}' \left[\frac{1}{(z + \Omega_j' - \Omega_i')^2} - \frac{1}{\Omega_i'^2}\right]\right\}$$
$$= \frac{1}{K_\nu{}^{2n}}\mathfrak{P}(z + \Omega_j') = \frac{1}{K_\nu{}^{2n}}\mathfrak{P}(z). \tag{27}$$

For $\nu = 4$, (Fig. 42) $K = \pm i, \pm 1$ and $K_\nu{}^{2n} = \pm 1$. Then, $\mathfrak{P}(z)^2$ is automorphic with respect to the group. $\mathfrak{P}(z)^2$ has a pole of the fourth order in the period parallelogram and takes on each value four times. The period parallelogram contains four copies of the fundamental region in the figure. Hence, the function takes on each value once in the fundamental region. It follows (Sec. 43, Theorem 14) that the most general simple automorphic function belonging to the group is a rational function of $\mathfrak{P}(z)^2$.

Differentiating (27) and (26), we have

$$\mathfrak{P}'(z_j)\frac{dz_j}{dz} = \frac{1}{K_\nu{}^{2n}}\mathfrak{P}'(z), \quad \frac{dz_j}{dz} = K_\nu{}^n,$$

whence,

$$\mathfrak{P}'(z_j) = \frac{1}{K_\nu^{3n}} \mathfrak{P}'(z). \tag{28}$$

If $\nu = 3$ (Fig. 43), K_ν is a cube root of unity and $K_\nu^{3n} = 1$. Then $\mathfrak{P}'(z)$ is automorphic with respect to the group. $\mathfrak{P}'(z)$ has a pole of the third order and takes on each value three times in the period parallelogram. The period parallelogram contains three copies of the fundamental region of the figure; hence, the function takes on each value once in the fundamental region. The most general simple automorphic function belonging to the group is a rational function of $\mathfrak{P}'(z)$.

If $\nu = 6$ (Fig. 44), K_ν is a sixth root of unity and $K_\nu^{3n} = \pm 1$. Then $\mathfrak{P}'(z)^2$ is automorphic. This function has a pole of the sixth order and takes on each value six times in the period parallelogram. The period parallelogram contains six copies of the fundamental region of the figure; hence, the function takes on every value once in the fundamental region. The most general simple automorphic function belonging to the group is a rational function of $\mathfrak{P}'(z)^2$.

III. THE GROUPS WITH TWO LIMIT POINTS

62. Determination of the Groups.—The remaining elementary groups—those with two limit points—are of less interest. We shall derive them briefly. Given such a group, so transformed that the limit points are zero and infinity. Then each transformation must carry a limit point either into itself or into the other limit point. The transformations (*a*) which have 0 and ∞ as fixed points; and (*b*) which carry 0 to ∞ and ∞ to 0, are readily found to have the following forms:

$$(a)\ z' = K_j z, \quad (b)\ z' = \frac{c_j}{z}. \tag{29}$$

The former are hyperbolic, elliptic, or loxodromic transformations, according to the value of K_j; the latter are elliptic with the fixed points $\pm\sqrt{c_j}$.

Taking logarithms in (29) and setting $\log z = w$, $\log z' = w'$, $\log K_j = \Omega_j$, $\log c_j = \Omega_j'$, we have

$$(a)\ w' = w + \Omega_j + 2m\pi i, \quad (b)\ w' = -w + \Omega_j' + 2n\pi i, \tag{30}$$

the integers m and n depending upon the values used for the logarithms. If the transformations in (29) form a group, so, also, do those in (30), and conversely. The group (29) is continuous or discontinuous according as (30) is continuous or discontinuous. Thus, (30) is continuous if it has arbitrarily small periods, $\Omega_j + 2m\pi i = \epsilon$. Then $\log z' = \log z + \epsilon$ and $z' = e^{\epsilon}z$, which differs arbitrarily little from the identity.

The groups (30) are of the type studied in Secs. 59 and 60 for $\nu = 2$. Suppose, first, that the transformations (*b*) do not appear. In (*a*) we can take one of the primitive periods to be a pure imaginary. Since $2\pi i$ is a period, this primitive period is of the form $\omega = 2\pi i/k$, where k is an integer. If there are no periods other than multiplies of ω, the group (30)(*a*) is simply periodic and is generated by the transformation $w' = w + \omega$. The group (29)(*a*) is then generated by $z' = e^{\omega}z = e^{2\pi i/k}z$. It is a finite cyclic group and can be discarded, since it has no limit points.

If (30)(*a*) is a doubly periodic group, let ω, ω' be a pair of primitive periods, ω being the period found above. Then ω' is not a pure imaginary. The general transformation of the group is $w' = w + m\omega + n\omega'$ and the general transformation of the original group is $z' = e^{m\omega+n\omega'}z$. If $k = 1$, so that $\omega = 2\pi i$, then $e^{m\omega} = 1$ and the group is simply $z' = e^{n\omega'}z$. Putting $K = e^{\omega'}$, $K_1 = e^{\omega}$, we have the following types of groups:

$$(A)\ z' = K^n z; \quad (B)\ z' = K^n K_1{}^m z \qquad |K| \neq 1,\ K_1 = e^{2\pi i/k}.$$

(*A*) is a hyperbolic or loxodromic cyclic group. (*B*) contains elliptic and loxodromic transformations and possibly also hyperbolic transformations.

If the transformations (*b*) appear in (30), they have the form (Equation (23) with $K = -1$)

$$w' = -w + b + m\omega + n\omega'.$$

The group can be transformed by a translation that carries b to the origin (which amounts in (30) to a transformation carrying $\sqrt{c}$ to 1); so we can take $b = 0$. From $\log z' = -\log z + m\omega + n\omega'$, we get $z' = e^{m\omega+n\omega'}/z$. We have then the following further groups:

$$(C)\ z' = K^n z, \quad z' = \frac{K^n}{z}; \quad (D)\ z' = K^n K_1{}^m z, \quad z' = \frac{K^n K_1{}^m}{z}.$$

For any values of K, K_1 such that $|K| \neq 1$ and $K_1 = e^{2\pi i/k}$, the sets of transformations (*A*)-(*D*) form groups of the kind sought.

The most general groups with two limit points are the groups (*A*)-(*D*) *and their transforms by means of linear transformations.*

CHAPTER VII

THE ELLIPTIC MODULAR FUNCTIONS

63. Certain Results from the Theory of Elliptic Functions.—The earliest automorphic functions to be studied were the elementary ones—the circular functions, the elliptic functions, and the rational or polyhedral automorphic functions. Of the non-elementary functions the so-called "elliptic modular functions" were extensively studied before the erection of a general theory of Fuchsian functions. *By an elliptic modular function, or more briefly, a modular function, is meant a simple automorphic function belonging to the modular group or to one of its subgroups.*[1]

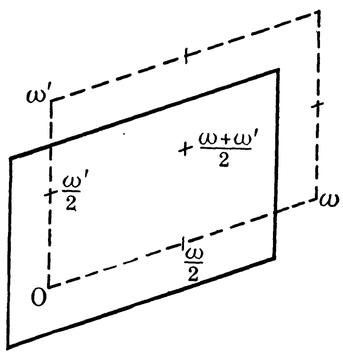


FIG. 45.

In Sec. 37, we constructed the fundamental region for the modular group. By transforming the group so as to reduce infinity to an ordinary point, we could set up automorphic functions by means of the theta series of Poincaré (Chap. V). Instead of following this process, however, we prefer to follow the historical order of development and to derive the modular functions in such a way as to bring out their connections with the elliptic functions. To this end we recall certain of the properties of the Weierstrassian function $\wp(z)$.

Let ω, ω' be a pair of primitive periods so denominated that for the ratio

$$\tau = \frac{\omega'}{\omega} = x + iy, \tag{1}$$

[1] The most complete treatise on these functions is FRICKE-KLEIN, "Vorlesungen über die Theorie der elliptischen Modulfunktionen," two volumes of some 1450 pages. See, also, VIVANTI, "Fonctions polyédriques et modulaires." Brief accounts will be found in HURWITZ-COURANT, "Funktionentheorie," 2nd ed., pp. 220–230, and in BIEBERBACH, *Lehrbuch der Funktionentheorie*, vol. 2, pp. 95–114.

y is positive; that is, which shall be called ω and which ω' is governed by the requirement that the angle $\omega 0 \omega'$ shall be positive and less than π (Fig. 45). We use the symbol $\Omega = m\omega + m'\omega'$ for the general period. The elliptic functions

$$\wp(z) = \frac{1}{z^2} + \sum{}' \left[\frac{1}{(z - \Omega)^2} - \frac{1}{\Omega^2} \right], \tag{2}$$

where the summation extends to all non-zero periods, and its derivative

$$\wp'(z) = -\frac{2}{z^3} - 2\sum{}' \frac{1}{(z - \Omega)^3}, \tag{3}$$

are connected by an algebraic relation (Theorem 12, Sec. 43). This relation is the following

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3, \tag{4}$$

where

$$g_2 = 60\sum{}' \frac{1}{\Omega^4}, \quad g_3 = 140\sum{}' \frac{1}{\Omega^6}. \tag{5}$$

This relation is got most easily by combining $\wp(z)$ and $\wp'(z)$ into a polynomial in such a way that the pole at the origin disappears. Then the polynomial reduces to a constant (Theorem 10, Sec. 42). In the neighborhood of the origin, we have

$$\frac{1}{(z - \Omega)^2} = \frac{1}{\Omega^2} + \frac{2z}{\Omega^3} + \frac{3z^2}{\Omega^4} + \frac{4z^3}{\Omega^5} + \frac{5z^4}{\Omega^6} + \cdots$$

We now insert this value in (2) and sum. We note that $\sum{}' \frac{1}{\Omega^3} = \sum{}' \frac{1}{\Omega^5} = \cdots = 0$ by virtue of the fact that the term containing $\Omega = m\omega + m'\omega'$ is cancelled by the term containing $\Omega = -m\omega - m'\omega'$. We have

$$\wp(z) = \frac{1}{z^2} + 3z^2\sum{}' \frac{1}{\Omega^4} + 5z^4\sum{}' \frac{1}{\Omega^6} + \cdots$$

or

$$\wp(z) = \frac{1}{z^2} + \frac{g_2}{20}z^2 + \frac{g_3}{28}z^4 + \cdots$$

From this series, we have, at once, the following:

$$\wp'(z) = -\frac{2}{z^3} + \frac{g_2}{10}z + \frac{g_3}{7}z^3 + \cdots,$$

$$\wp(z)^3 = \frac{1}{z^6} + \frac{3g_2}{20z^2} + \frac{3g_3}{28} + \cdots,$$

$$\wp'(z)^2 = \frac{4}{z^6} - \frac{2g_2}{5z^2} - \frac{4g_3}{7} + \cdots$$

On combining these equations we find

$$\wp'(z)^2 - 4\wp(z)^3 + g_2\wp(z) = -g_3 + cz^2 + \cdots$$

The first member of this equation is an elliptic function with no singularity in the fundamental region; hence, it is a constant.
Setting $z = 0$, we see that its value is $-g_3$ and (4) is established.

The function $\wp'(z)$ satisfies the following:

$$\wp'\left(-\frac{\Omega}{2}\right) = \wp'\left(\frac{\Omega}{2}\right), \quad \wp'\left(-\frac{\Omega}{2}\right) = -\wp'\left(\frac{\Omega}{2}\right),$$

the first of which follows from the fact that $-\Omega/2$ and $\Omega/2$ differ by a period and the second from the fact that $\wp'(z)$ is an odd function of z. It follows that if $\wp'(z)$ is analytic at $\Omega/2$; that is, if $\Omega/2$ is a *half period,* then $\wp'(\Omega/2) = 0$. Three of these half periods lie in the period parallelogram. The three quantities

$$e_1 = \wp\left(\frac{\omega}{2}\right), \quad e_2 = \wp\left(\frac{\omega + \omega'}{2}\right), \quad e_3 = \wp\left(\frac{\omega'}{2}\right), \tag{6}$$

are then the roots of the equation

$$4t^3 - g_2 t - g_3 = 0. \tag{7}$$

The three quantities e_1, e_2, e_3 are unequal. $\wp(z)$ takes on each value twice in the period parallelogram. Since $\wp'(z)$ vanishes at $\omega/2$, $\wp(z)$ takes on the value e_1 twice there, and so cannot take on the value e_1 elsewhere in the parallelogram. By similar reasoning e_2 and e_3 are unequal. The condition for equal roots in (7) we find readily to be the vanishing of the expression

$$\Delta = g_2^3 - 27g_3^2. \tag{8}$$

Hence,

$$\Delta \neq 0.$$

64. Change of the Primitive Periods.—We consider the problem of finding all pairs of primitive periods; that is, periods in terms of which we can express all other periods by adding integral multiples of the two periods. The periods

$$\omega_1' = a\omega' + b\omega, \quad \omega_1 = c\omega' + d\omega \tag{9}$$

a, b, c, d being integers, are primitive periods if, and only if, ω, ω' can be expressed as sums of integral multiples of ω_1, ω_1'. This condition is necessary, by definition. It is also sufficient. For, any period Ω, being expressible in integral multiples of ω, ω', is then expressible in integral multiples of ω_1, ω_1'. Solving (9) we have

$$\omega' = \alpha\omega_1' + \beta\omega_1, \quad \omega = \gamma\omega_1' + \delta\omega_1,$$

where, putting $D = ad - bc$,

$$\alpha = \frac{d}{D}, \quad \beta = -\frac{b}{D}, \quad \gamma = -\frac{c}{D}, \quad \delta = \frac{a}{D}.$$

Then

$$\alpha\delta - \beta\gamma = \frac{ad - bc}{D^2} = \frac{1}{D}.$$

If α, β, γ, δ are integers, then $1/D$ is an integer; whence, $D = \pm 1$. Obviously, if $D = \pm 1$, the four quantities are integers. Hence, *the necessary and sufficient conditions that* (9) *be primitive periods is that* $ad - bc = \pm 1$.

The ratio of the new periods satisfies the relation

$$\tau_1 = \frac{\omega_1'}{\omega_1} = \frac{a\tau + b}{c\tau + d}. \tag{10}$$

Let us require that the periods be so named that in the expression $\tau_1 = x_1 + iy_1$, y_1 shall be positive. Then τ_1 as well as τ lies in the upper half plane. Now, the transformation (10) is non-loxodromic if $D > 0$ and is loxodromic if $D < 0$. In the former case, the upper half plane is carried into itself; in the latter it is carried into the lower half plane. We have then

$$ad - bc = +1. \tag{11}$$

The set of transformations (10) is the modular group.

65. The Function J(τ).—Expressing the quantities (5) and (8) in terms of the ratio τ, we have

$$\left.\begin{aligned} g_2 = g_2(\omega, \omega') &= 60\sum{}'\frac{1}{(m\omega + m'\omega')^4} \\ &= \frac{60}{\omega^4}\sum{}'\frac{1}{(m + m'\tau)^4} = \frac{1}{\omega^4}g_2(1, \tau), \\ g_3 = g_3(\omega, \omega') &= \frac{40}{\omega^6}\sum{}'\frac{1}{(m + m'\tau)^6} = \frac{1}{\omega^6}\, g_3(1, \tau), \\ \Delta = \Delta(\omega, \omega') &= g_2(\omega, \omega')^3 - 27g_3(\omega, \omega')^2 = \frac{1}{\omega^{12}}\,\Delta(1, \tau). \end{aligned}\right\} \tag{12}$$

It is a well-known result that the series (5) converge absolutely for any pair of primitive periods ω, ω' whose ratio is not real. Hence, the series for $g_2(1, \tau)$ and $g_3(1, \tau)$ converge absolutely for any value of τ which is not real. We show below that these series converge *uniformly* in any closed region not containing real points. From this, it follows that $g_2(1, \tau)$ and $g_3(1, \tau)$ are analytic functions of τ in the whole upper half τ-plane. So also, then, is $\Delta(1, \tau)$.

We may establish the uniform convergence as follows: Let S be a closed region in the τ-plane not containing real points. Let η be the minimum distance from the boundary of S to the real axis; and let $|\tau| < N$ in S. We now show that $\epsilon > 0$ can be chosen sufficiently small that

$$|m + m'\tau| > \epsilon|m + m'i|; \text{ or } |m + m'\tau|^2 - \epsilon^2|m + m'i|^2 > 0,$$

for all values of m, m' in the summation and for all τ in S. Writing $\tau = x + iy$, the first member of the last inequality may be written

$$\begin{aligned} |m + m'x + im'y|^2 - \epsilon^2|m + m'i|^2 \\ &= m^2 + 2mm'x + m'^2x^2 + m'^2y^2 - \epsilon^2m^2 - \epsilon^2m'^2 \\ &= m^2\left[1 - \frac{1}{k^2} - \epsilon^2\right] + \left[\frac{1}{k}m + km'x\right]^2 \\ &\qquad + m'^2[y^2 - (k^2 - 1)x^2 - \epsilon^2]. \end{aligned}$$

Here, since x is bounded, $|x| < N$, and $|y| \geqslant \eta$, we can take $k(>1)$ near enough to 1 and $\epsilon(>0)$ small enough that all the terms in the last member are positive whatever m and m' are; which establishes the desired inequality. The terms of the series for $g_2(1, \tau)$ are less in absolute value than the corresponding terms of the series of positive constant terms

$$\sum{}' \frac{1}{|m + m'i|^4\epsilon^4},$$

which is known to converge. It follows that the series for $g_2(1, \tau)$ converges uniformly in S. The uniform convergence of the series for $g_3(1, \tau)$ is similarly established.

We now combine the quantities in (12) to form a function from which the factor ω cancels and which is, therefore, a function of the ratio τ alone:

$$J(\tau) = \frac{g_2{}^3}{\Delta} = \frac{g_2^3(1, \tau)}{\Delta(1, \tau)}. \tag{13}$$

The numerator and denominator are analytic functions of τ, and the denominator does not vanish, in the upper half plane. Hence, $J(\tau)$ is analytic in the whole upper half τ-plane.

Let ω, ω' of given ratio $\omega'/\omega = \tau$ generate a group. Then, the periods $\omega_1' = a\omega' + b\omega$, $\omega_1 = c\omega' + d\omega$, where a, b, c, d are any integers such that $ad - bc = 1$, are primitive periods for the group. Since the new periods generate the original group, they lead to the same $\wp$-function and to the same constants g_2, g_3, Δ. Hence, the function J is unaltered. Now, the ratio of the new periods $\tau_1 = \omega_1'/\omega_1$ is connected with that of the old by Equation (10); hence, we have, for any transformation of the modular group,

$$J\left(\frac{a\tau + b}{c\tau + d}\right) = J(\tau). \tag{14}$$

66. Behavior of $J(\tau)$ at the Parabolic Points.—$J(\tau)$ is analytic at all finite points of the fundamental region R_0 of Fig. 46. We now consider its behavior at the upper end of the region. The sides which meet at infinity are congruent by the transformation $\tau' = \tau + 1$. In accordance with Sec. 41, we make the change of variable

$$t = e^{2\pi i \tau}. \tag{15}$$

This transformation maps the part of R_0 above the line DEF, with the equation $y = k > 1$, on the interior of the circle K whose equation is $|t| = e^{-2\pi k}$, congruent points on the two sides going into coincident points on a radius (Fig. 47). The function J takes on the same values at a point of this radius other than the

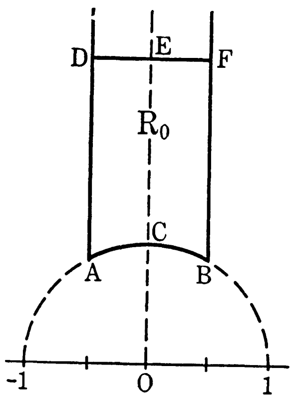


FIG. 46.

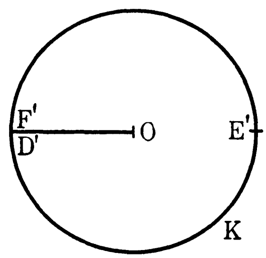


FIG. 47.

origin, when approached from the two sides, and so is single valued and analytic in K, with the possible exception of the origin. This is also true of $g_2(1, \tau)$ and $g_3(1, \tau)$, for each is unaltered by the transformation $\tau' = \tau + 1$. Thus,

$$\begin{aligned} g_2(1, \tau + 1) &= 60 \sum{}' \frac{1}{[m + m'(\tau + 1)]^4} \\ &= 60 \sum{}' \frac{1}{[m + m' + m'\tau]^4} \\ &= g_2(1, \tau), \end{aligned} \tag{16}$$

the terms of the second sum being merely a rearrangement of those of the sum defining $g_2(1, \tau)$. Similarly, $g_3(1, \tau + 1) = g_3(1, \tau)$.

In order to express $g_2(1, \tau)$ and $g_3(1, \tau)$ as functions of t, we shall employ two well-known expressions for the cotangent:[1]

$$\pi \cot \pi u = \frac{1}{u} + \sum_{m=-\infty}^{\infty}{}' \left[\frac{1}{u+m} - \frac{1}{m}\right] \tag{17}$$

and

$$\cot \pi u = i\frac{w+1}{w-1}, \quad w = e^{2\pi i u},$$

whence,

$$\pi \cot \pi u = -\pi i[1 + 2w + 2w^2 + 2w^3 + \cdots]. \tag{18}$$

We now equate the second members of (17) and (18) and differentiate, respectively, three and five times with respect to u, using each time the relation $dw/du = 2\pi i w$. We have the following results:

$$-6\sum_{m=-\infty}^{\infty}\frac{1}{(m+u)^4} = -16\pi^4[w + 8w^2 + \cdots\ ],$$

$$-120\sum_{m=-\infty}^{\infty}\frac{1}{(m+u)^6} = 64\pi^6[w + 32w^2 + \cdots],$$

where the terms $1/u^4$ and $1/u^6$ have been put under the sign of summation. Setting $u = m'\tau$ $(m' > 0)$, so that $w = e^{2\pi i m'\tau} = t^{m'}$ we have

$$\sum_{m=-\infty}^{\infty}\frac{1}{(m+m'\tau)^4} = \frac{8\pi^4}{3}[t^{m'} + 8t^{2m'} + \cdots],$$

$$\sum_{m=-\infty}^{\infty}\frac{1}{(m+m'\tau)^6} = -\frac{8\pi^6}{15}[t^{m'} + 32t^{2m'} + \cdots].$$

We can write[2]

$$g_2(1,\tau) = 60\left[\sum_{m=-\infty}^{\infty}{}'\frac{1}{m^4} + 2\sum_{m'=1}^{\infty}\sum_{m=-\infty}^{\infty}\frac{1}{(m+m'\tau)^4}\right.$$

$$= 60\left[\frac{\pi^4}{45} + \frac{16\pi^4}{3}\sum_{m'=1}^{\infty}(t^{m'} + 8t^{2m'} + \cdots)\right].$$

[1] For the first, see OSGOOD, W. F., "Lehrbuch der Funktionentheorie," 2nd ed., p. 507.

[2] The values of $\sum{}'\frac{1}{m^4}$ and $\sum{}'\frac{1}{m^6}$ which are used here and in (20) are readily got from (17) by expanding $\pi \cot \pi u - \frac{1}{u}$ in powers of u, differentiating three and five times, and setting $u = 0$.

Now, the functions in the last summation are analytic in the whole interior of K, including the origin. Since the series converges uniformly on the circumference of K, the sum is an analytic function throughout K. Also, we can collect the powers of t. We have

$$g_2(1, \tau) = \pi^4[4/3 + 320t + \cdots]. \tag{19}$$

Similarly,

$$\begin{aligned} g_3(1, \tau) &= 140\left[\sum_{m=-\infty}^{\infty}{}' \frac{1}{m^6} + 2\sum_{m'=1}^{\infty}\sum_{m=-\infty}^{\infty} \frac{1}{(m + m'\tau)^6}\right] \\ &= 140\left[\frac{2\pi^6}{945} - \frac{16\pi^6}{15}\sum_{m'=1}^{\infty}(t^{m'} + 32t^{2m'} + \cdots)\right] \\ &= \pi^6[8/27 - 448/3\, t + \cdots]. \end{aligned} \tag{20}$$

From (19) and (20) we find

$$\Delta(1, \tau) = g_2(1, \tau)^3 - 27g_3(1, \tau)^2 = \pi^{12}[4096t + \cdots], \tag{21}$$

and finally

$$\begin{aligned} J(\tau) &= \frac{g_2(1, \tau)^3}{\Delta(1, \tau)} = \frac{\pi^{12}(4/3 + 320t + \cdots)^3}{\pi^{12}(4096t + \cdots)} \\ &= \frac{1}{1728t} + c_0 + c_1 t + \cdots \end{aligned} \tag{22}$$

The function $J(\tau)$, when expressed in terms of t, thus has a pole of the first order at $t = 0$.

67. Further Properties of $\mathbf{J}(\tau)$.—The function $J(\tau)$ has a single pole of the first order in R_0, according to the convention of Sec. 42. It, therefore, takes on each value once in R_0 (Theorem 11, Sec. 42).

THEOREM 1.—*The function $J(\tau)$ takes on each value once, and only once, in the fundamental region.*

Applying Theorem 14, Sec. 43, we determine all functions which are automorphic with respect to the modular group and have no other singularities than poles in the fundamental region.

THEOREM 2.—*The most general simple automorphic function belonging to the modular group is a rational function of $J(\tau)$.*

We next consider the values of $J(\tau)$ at the vertices A, B, C of R_0. At A we have $\tau = \rho = -1/2 + 1/2 i\sqrt{3}$. Here, ρ is a cube root of unity, so $\rho^3 = 1$ and $\rho^2 + \rho + 1 = 0$. If, in

$$g_2(1, \rho) = 60\sum{}' \frac{1}{(m + m'\rho)^4},$$

we multiply numerator and denominator by ρ^8, thus inserting a factor ρ^2 in each factor in the denominator, we have

$$\begin{aligned} g_2(1, \rho) &= 60\rho^8 \sum{}' \frac{1}{(m\rho^2 + m')^4} \\ &= 60\rho^2 \sum{}' \frac{1}{(m' - m - m\rho)^4} = \rho^2 g_2(1, \rho), \end{aligned}$$

for the terms of the last summation are merely those of the original sum arranged in a different order. From this, since $\rho^2 \neq 1$, we have

$$g_2(1, \rho) = 0, \quad J(\rho) = 0. \tag{23}$$

J vanishes also at the congruent point B.

At C we have $\tau = i$. Then,

$$\begin{aligned} g_3(1, i) &= 140 \sum{}' \frac{1}{(m + m'i)^6} \\ &= 140 i^6 \sum{}' \frac{1}{(mi - m')^6} = -g_3(1, i). \end{aligned}$$

From this we have

$$g_3(1, i) = 0, \quad J(i) = 1. \tag{24}$$

We propose now to find all points of R_0 at which $J(\tau)$ is real. We consider the reflection of a point τ in the imaginary axis; namely $\tau' = -\bar{\tau}$, where, as hitherto, the bar indicates the conjugate imaginary. We have

$$g_2(1, \tau') = 60 \sum{}' \frac{1}{(m - m'\bar{\tau})^4}; \; \bar{g}_2(1, \tau) = 60 \sum \frac{1}{(m + m'\bar{\tau})^4};$$

so, $g_2(1, \tau') = g_2(1, \tau)$. Similarly, $g_3(1, \tau') = \bar{g}_3(1, \tau)$; and

$$J(\tau') = \bar{J}(\tau). \tag{25}$$

There are two cases in which $J(\tau') = J(\tau)$. If τ lies on the imaginary axis, it coincides with its reflection, and $\tau' = \tau$. If τ lies on the boundary of R_0, its reflection τ' lies at a *congruent point* on the boundary at which J has the same value as at τ. In both cases we have

$$J(\tau) = \bar{J}(\tau),$$

and $J(\tau)$ is real. That is, $J(\tau)$ is real on the boundary of R_0 and on the imaginary axis.

There are no further points of R_0 at which $J(\tau)$ is real. For, if the function is real at any other point τ_1 of R_0, we have $J(\tau_1) = \bar{J}(\tau_1) = J(\tau_1')$; and the function takes on the same value at the two distinct inner points τ_1, τ_1' of R_0, which is contrary to Theorem 1.

It follows from the preceding that the imaginary part of the function $J(\tau) = u + iv$ has always the same sign in each of the halves into which the imaginary axis divides R_0. If $v_{\tau_1} > 0$ and $v_{\tau_2} < 0$ at two points of R_0 on the same side of the imaginary axis, and if τ trace a curve λ from τ_1 to τ_2, lying in R_0 and not meeting the imaginary axis, then $v = 0$ at some point of λ. At this point, the function is real, which is contrary to what we have just proved.

It remains to find in which half v is greater than zero. As τ moves from A to C along the boundary of R_0, $J(\tau)$ moves from 0 to 1 to the right along the real axis in the J plane. Points in the neighborhood of the first path and to the left of it go into points in the neighborhood of the second path and to the left of it; that is, into points of the upper half J-plane. Then $v > 0$ in the left half of R_0.

We collect our results into the following theorem:

THEOREM 3.—*Writing $J(\tau)$ in terms of its real and imaginary parts, $J(\tau) = u + iv$; then $v > 0$ at the inner points of R_0 to the left of the imaginary axis; $v < 0$ at the inner points of R_0 to the right of the imaginary axis; $v = 0$ on the boundary of R_0 and on the imaginary axis. Also, $v_{\tau'} = -v_\tau$, where τ' is the reflection of τ in the imaginary axis.*

The last statement of the theorem is an immediate consequence of (25).

The function $z = J(\tau)$ maps the left half of R_0 in a one-to-one manner on the upper half z-plane and maps the right half similarly on the lower half-plane. As τ moves upward along the imaginary axis from C to ∞, z moves from 1 to the right to $+\infty$. The two halves of the z-plane are joined to the right of $z = 1$. R_0 is mapped on the z-plane bounded by a slit which extends along the real axis from $z = 1$ to $z = -\infty$.

68. The Function $\lambda(\tau)$.—Of the elliptic modular functions belonging to a subgroup of the modular group, and not automorphic with respect to the whole group, we shall treat here only one. The function which we shall consider is defined as follows,

$$\lambda(\tau) = \frac{e_2 - e_3}{e_1 - e_3}, \tag{26}$$

where e_1, e_2, e_3 have the definitions given in Equation (6).

We note first that λ is, in fact, a function of the ratio τ alone. We write $\wp(z; \omega, \omega')$ to call attention to the dependence of the $\wp$-function on the periods. Since all the terms in the series (2) defining $\wp$ are of the (-2)nd degree in z, ω, ω', we have

$$\wp(z; \omega, \omega') = k^2\wp(kz; k\omega, k\omega'), \ k \neq 0. \tag{27}$$

Applying this to Equations (6), taking $k = 1/\omega$, we have

$$\left.\begin{aligned} e_1 &= \wp\left(\frac{\omega}{2}; \omega, \omega'\right) = \frac{1}{\omega^2}\wp\left(\frac{1}{2}; 1, \tau\right), \\ e_2 &= \wp\left(\frac{\omega + \omega'}{2}; \omega, \omega'\right) = \frac{1}{\omega^2}\wp\left(\frac{1+\tau}{2}; 1, \tau\right), \\ e_3 &= \wp\left(\frac{\omega'}{2}; \omega, \omega'\right) = \frac{1}{\omega^2}\wp\left(\frac{\tau}{2}; 1, \tau\right). \end{aligned}\right\} \tag{28}$$

On forming the quotient in (26). ω cancels, and λ is a function of τ alone.

The uniform convergence of the series that define the functions of τ appearing in (28) can be proved for any closed region in the τ-plane which does not contain points of the real axis. This proof follows the lines of the convergence proof in Sec. 65 and will not be given here. These functions are then analytic in the upper half plane. Since e_1, e_2, e_3 are unequal, *$\lambda(\tau)$ is an analytic function of τ, and nowhere takes on the value 0 or the value 1, in the upper half τ-plane.*

Consider the group generated by a pair of periods ω, ω' whose ratio ω'/ω is τ. Let ω_1, ω_1', Equation (9), be a second pair of primitive periods with the condition that $ad - bc = +1$. The ratio of the new periods is given by (10). For the new periods, we have

$$\begin{gathered} e_1' = \wp\left(\frac{c\omega' + d\omega}{2}\right), \ e_2' = \wp\left(\frac{c\omega' + d\omega + a\omega' + b\omega}{2}\right), \\ e_3' = \wp\left(\frac{a\omega' + b\omega}{2}\right). \end{gathered} \tag{29}$$

Since we are dealing with the same group and, hence, have the same $\wp$-function, we have the same constants as before. They are possibly arranged in a different order. And a rearrangement of the constants in (26) alters the value of λ, in general.

If now $(c\omega' + d\omega)/2$ and $\omega/2$ differ by a period, that is, if $(c\omega' + (d - 1)\omega)/2$ is a period, we have

$$\mathfrak{P}\left(\frac{c\omega' + d\omega}{2}\right) = \mathfrak{P}\left(\frac{\omega}{2}\right), \text{ or } e_1' = e_1.$$

This occurs if, and only if, c is even and d is odd. Similarly, if $(a\omega' + b\omega)/2$ and $\omega'/2$ differ by a period, that is, if $((a - 1)\omega' + b\omega)/2$ is a period, we have $e_3' = e_3$. This occurs if a is odd and b is even. If the preceding conditions hold, the remaining roots are equal, $e_2' = e_2$, and λ is unchanged.

Now if b and c are even, $b = 2b'$, $c = 2c'$, then, a and d are necessarily odd, as a consequence of the relation $ad - bc = 1$; and we have

$$\lambda\left(\frac{a\tau + 2b'}{2c'\tau + d}\right) = \lambda(\tau), \quad ad - 4b'c' = 1. \tag{30}$$

$\lambda(\tau)$ is thus unaltered by all the transformations of the subgroup treated in Sec. 38 and whose fundamental region was found in Fig. 30. (For convenience this region is repeated in Fig. 48.)

69. The Relation between $\lambda(\tau)$ and $J(\tau)$.—If we make a transformation of the modular group which does not belong to the subgroup of (30), there is an interchange of the roots e_1, e_2, e_3. All values of (26) which can result are contained in the following table:

$$\left.\begin{aligned} &\frac{e_2 - e_3}{e_1 - e_3} = \lambda, \quad \frac{e_3 - e_1}{e_2 - e_1} = \frac{1}{1 - \lambda}, \quad \frac{e_1 - e_2}{e_3 - e_2} = \frac{\lambda - 1}{\lambda}, \\ &\frac{e_1 - e_3}{e_2 - e_3} = \frac{1}{\lambda}, \quad \frac{e_3 - e_2}{e_1 - e_2} = \frac{\lambda}{\lambda - 1}, \quad \frac{e_2 - e_1}{e_3 - e_1} = 1 - \lambda. \end{aligned}\right\} \tag{31}$$

These six transformations of λ constitute a group, namely, the group of the anharmonic ratios (Sec. 36).

If, now, we form a rational symmetric function of the six quantities (31), we have a simple automorphic function $F(\tau)$ belonging to the modular group. For, in the first place, any modular transformation merely interchanges the quantities e_1, e_2, e_3 in a certain way and does not alter a symmetric function of these quantities. And, in the second place, such a symmetric function of the roots can be expressed as a rational function of the coefficients $g_2(\omega, \omega')$, $g_3(\omega, \omega')$, and, hence, also of $g_2(1, \tau)$, $g_3(1, \tau)$. The latter two functions have no singularities other than poles in R_0 (including the parabolic point); whence, a rational function of them has no singularities other than poles.

Then $F(\tau)$ is a simple automorphic function. It follows from Theorem 2 that $F(\tau)$ is a rational function of $J(\tau)$.

We shall consider the following symmetric function:

$$F(\tau) = (\lambda + 1)\left(\frac{1}{1 - \lambda} + 1\right)\left(\frac{\lambda - 1}{\lambda} + 1\right)\left(\frac{1}{\lambda} + 1\right)\left(\frac{\lambda}{\lambda - 1} + 1\right)(1 - \lambda + 1)$$

$$= -\frac{(\lambda + 1)^2(2 - \lambda)^2(2\lambda - 1)^2}{\lambda^2(1 - \lambda)^2}. \tag{32}$$

Putting this in terms of e_1, e_2, e_3, we have

$$F(\tau) = -\frac{(e_2 + e_1 - 2e_3)^2(e_2 + e_3 - 2e_1)^2(e_1 + e_3 - 2e_2)^2}{(e_1 - e_2)^2(e_2 - e_3)^2(e_1 - e_3)^2}. \tag{33}$$

From Equation (7) we have $e_1 + e_2 + e_3 = 0$, $e_1e_2e_3 = g_3/4$. Using these relations, we can write the numerator of (33) in the form

$$(-3e_3)^2(-3e_1)^2(-3e_2)^2 = \frac{3^6}{4^2}g_3^2.$$

The denominator of (33) is equal to the discriminant of (7), save for a constant factor. We have, in fact,

$$(e_1 - e_2)^2(e_2 - e_3)^2(e_1 - e_3)^2 = \tfrac{1}{16}(g_2^3 - 27g_2^3).$$

Then,

$$F(\tau) = -\frac{3^6 g_3^2}{g_2^3 - 27g_3^2} = 27[1 - J(\tau)]. \tag{34}$$

From (34) and (32) we have, with a little calculation,

$$J(\tau) = 1 - \frac{F(\tau)}{27} = \frac{4}{27}\,\frac{(1 - \lambda + \lambda^2)^3}{\lambda^2(1 - \lambda)^2}. \tag{35}$$

70. Further Properties of $\lambda(\tau)$.—We first note that for each point of B_0 there exist five other points which are congruent by the modular group. In Fig. 49 we have superposed the modular division of the plane on B_0. Representing the left half of R_0 and the triangles congruent to it by the + sign and the right half and its congruent triangles by the − sign, we observe that there are exactly six copies of each. Then $J(\tau)$ takes on each value *six* times in B_0.

We next show that $\lambda(\tau)$ is a simple automorphic function. The only doubtful matter is its behavior at the parabolic points. Consider the parabolic point at infinity. If, in accordance with Sec. 41, we make the change of variable $t_1 = e^{\pi i \tau}$ (the sides of B_0 being congruent by the transformation $\tau' = \tau + 2$), the upper end of B_0 is mapped on a region similar to Fig. 47, and λ goes over into a function of t_1 which is analytic in the neighborhood of the origin. If λ has an essential singularity at $t_1 = 0$, it takes on certain values an infinite number of times. Then, from (35), J takes on certain values an infinite number of times in B_0, which is impossible. Hence, λ is analytic at $t_1 = 0$ or has a pole there. Similar remarks apply to the other parabolic points. Then $\lambda(\tau)$ is a simple automorphic function. It, therefore, takes on each value the same number of times.

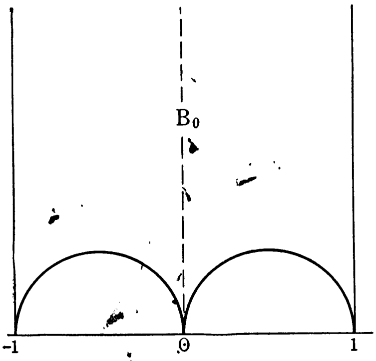


FIG. 48.

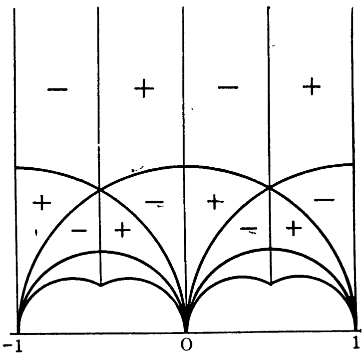


FIG. 49.

The exceptional values 0, 1, ∞, which are taken on nowhere in the upper half plane, are necessarily taken on in the parabolic points $0, \pm 1, \infty$. The points $+1$ and -1 are congruent and bear the same value. We shall now determine in which points these values are taken on. One method of procedure would be to seek the limit approached by $\lambda(\tau)$ as τ approaches the parabolic point from within the region. An easier method is the following:

Let us make the change of periods (see (9) and (10))

$$\omega_1' = \omega' + \omega, \quad \omega_1 = \omega, \quad \tau_1 = \tau + 1.$$

The new constants are

$$e_1' = \wp\left(\frac{\omega_1}{2}\right) = \wp\left(\frac{\omega}{2}\right) = e_1,$$

$$e_3' = \wp\left(\frac{\omega_1'}{2}\right) = \wp\left(\frac{\omega + \omega'}{2}\right) = e_2, \quad e_2' = e_3,$$

and

$$\lambda' = \frac{e_2' - e_3'}{e_1' - e_3'} = \frac{e_3 - e_2}{e_1 - e_2} = \frac{\lambda}{\lambda - 1}.$$

Now, when we make this change of primitive periods, $\tau = 0, -1, \infty$ go into $\tau' = 1, 0, \infty$, respectively; and $\lambda = 0, 1, \infty$ go into $\lambda' = 0, \infty, 1$, respectively. Since $\tau = \infty$ is unaltered, the corresponding value of λ is unaltered. Hence, $\lambda(\infty) = 0$.

Again, make the change of periods

$$\omega_1' = -\omega, \ \omega_1 = \omega', \ \tau_1 = -\frac{1}{\tau}.$$

Then,

$$e_1' = \wp\left(\frac{\omega_1}{2}\right) = \wp\left(\frac{\omega'}{2}\right) = e_3,$$

$$e_3' = \wp\left(\frac{\omega_1'}{2}\right) = \wp\left(-\frac{\omega}{2}\right) = e_1, \ e_2' = e_2,$$

and

$$\lambda' = \frac{e_2 - e_1}{e_3 - e_1} = 1 - \lambda.$$

Here, $\tau = 0, -1, \infty$ go into $\tau_1 = \infty, 1, 0$, respectively; and $\lambda = 0, 1, \infty$ go into $\lambda' = 1, 0, \infty$, respectively; whence, $\lambda(\pm 1) = \infty$, $\lambda(0) = 1$.

We shall now prove the following theorem:

THEOREM 4.—*The function $\lambda(\tau)$ takes on each value once, and only once, in the fundamental region B_0.*

Each simple automorphic function belonging to the subgroup is a rational function of $\lambda(\tau)$.

We find, first, the order of the zero of $\lambda(\tau)$ at the parabolic point at ∞. As $\tau \to \infty$ in B_0, $\lambda \to 0$, and, from (35),

$$\lambda^2 J \to 4/27. \tag{36}$$

Changing the variable to $t_1 = e^{\pi i \tau}$, we have from (22) and (15)

$$J = \frac{1}{1728t} + c_0 + c_1 t + \cdots = \frac{1}{1728t_1^2} + c_0 + c_1 t_1^2 + \cdots$$

J thus has a pole of the second order at $t_1 = 0$. Hence, from (36), λ has a zero of the first order. Since $\lambda(\tau)$ has one, and only one, zero in B_0, it takes on every value exactly once.

The second statement of the theorem is a consequence of Theorem 14, Sec. 43.

The reality of $\lambda(\tau)$ can be treated by the method used in studying the reality of $J(\tau)$. We find here, also, that $\lambda(-\bar{\tau}) = \bar{\lambda}(\tau)$; and by repeating practically verbatim the reasoning in connection with Theorem 3, we show that $\lambda(\tau)$ *is real on the* *boundary of* B_0 *and on the imaginary axis.* The right half of B_0 corresponds to the upper half λ-plane.

We close the chapter with a general theorem concerning modular functions.

THEOREM 5.—*Given a subgroup of the modular group whose fundamental region consists of* k *copies of the fundamental region* R_0 *of the modular group. Then any simple automorphic function* $f(\tau)$ *belonging to the subgroup is connected with* $J(\tau)$ *by a relation of the form*

$$\Phi(f, J) = 0,$$

where Φ *is a polynomial of degree not exceeding* k *in* f.

The fact that an algebraic relation between f and J exists follows from Theorem 12, Sec. 43, since $J(\tau)$ is, obviously, a simple automorphic function with respect to the subgroup. To each value J_0 of J, there correspond not more than k distinct values of f; namely, the values of $f(\tau)$ at the k points of the fundamental region at which $J(\tau) = J_0$. That is, the irreducible relation connecting f and J is of degree k at most in f.

The equation (35) connecting J and λ is an example.

CHAPTER VIII

CONFORMAL MAPPING

71. Conformal Mapping.—The present chapter will be devoted to the problem of mapping one region conformally upon another. This problem, as we shall see subsequently, has an important bearing upon certain of the applications of automorphic functions.

Let $f(z)$ be a function which is single valued and analytic in a plane region S and which takes on no value twice in the region. Then, the relation $z' = f(z)$ sets up a one-to-one correspondence between the inner points of S and the inner points of a plane region S' in the z'-plane. In this correspondence—this transformation of S into S'—angles are preserved. We say that *the function $f(z)$ maps S conformally on S'.* It is well known, also, that, conversely, if one region is transformed into another with preservation of angles, the correspondence between the points is determined by an analytic function $z' = f(z)$. In speaking of conformal mapping we shall understand always that the points of the two regions involved correspond in a one-to-one manner.

We can extend this simple notion in important respects. We can first remove the requirement that $f(z)$ take on each value only once in S'. We require, of course, that $f(z)$ shall not be a constant. The neighborhood of a point a in S at which $f'(a) \neq 0$ is mapped on a plane region in the neighborhood of $b = f(a)$. At a we have the development

$$z' = b + f'(a)(z - a) + \cdot \cdot \cdot ; \qquad (1)$$

and, for a suitably small region about a, the corresponding points in the z'-plane satisfy the relation

$$z = a + \frac{1}{f'(a)}(z' - b) + \cdot \cdot \cdot \qquad (2)$$

That is, the inverse function is single valued, hence, the mapped region is plane. If, however, $f'(a) = 0$, we have a development of the form

$$z' = b + c(z - a)^n + \cdot \cdot \cdot , n > 1, c \neq 0, \qquad (3)$$

and the inverse function

$$z = a + c'(z' - b)^{1/n} + \cdot \cdot \cdot , \qquad (4)$$

is not single valued. To each point near a correspond n points near b. By representing z' on an n-sheeted region with a branch point at b we can secure the desired one-to-one character.

We can combine the various values of z' into branches with suitable branch points (which may of course be absent) into a surface S' with a finite or infinite number of sheets, so that the correspondence between the points of S and those of S' is one-to-one. We say that $f(z)$ maps S conformally on S'. Here S' is merely a part of the Riemann surface—possibly the whole surface—of the function inverse to $f(z)$. The conformal character fails to hold at the branch points, but the correspondence is continuous there. The number of branch points in S' may be infinite, but each is of finite order.

Again, we can replace S by a similar finitely or infinitely sheeted region with interior branch points of finite order on which the function $f(z)$ is analytic, in general, and is single valued. We thus have the mapping of one many-sheeted region upon another.

Finally, we shall admit the point at infinity as an inner point of either S or S'. This presents the existence of poles of $f(z)$ on S. If $f(z)$ has a pole at a, the corresponding point in S' is at infinity.

In the first part of this chapter we shall be concerned primarily with plane regions, although we shall employ certain simple two-sheeted regions in the derivation of some of the results. We turn, first, to the proof of some useful preliminary theorems.

72. Schwarz's Lemma.[1]—In the study of properties of functions, and particularly in connection with the problem to be considered later, of mapping regions on circles, the following proposition is often a powerful tool.

THEOREM 1. SCHWARZ'S LEMMA.—*Let $f(z)$ be analytic in the unit circle Q_0 and vanish at the origin. If $|f(z)| \leqslant 1$ on Q_0, then $|f(z)| \leqslant |z|$ in Q_0 and $|f'(0)| \leqslant 1$. The inequalities $|f(z)| = |z|$, $z \neq 0$, and $|f'(0)| = 1$ hold if, and only if, $f(z) = e^{i\theta}z$.*

The function $f(z)/z$ is analytic in Q_0, if properly defined at the origin. Let Q_r be a circle concentric with Q_0 and of radius r less than 1. The function takes on its maximum absolute value on the boundary; so, if z is in Q_r,

$$\left|\frac{f(z)}{z}\right| \leqslant \frac{1}{r}.$$

[1] SCHWARZ, H. A., *Ges. Abhandl.*, vol. 2, p. 110; C. CARATHÉODORY, *Math. Ann.*, vol. 72, p. 110, 1912.

Since r may be taken as near 1 as we like, z being held fixed, we have

$$\left|\frac{f(z)}{z}\right| \leqslant 1. \tag{5}$$

This is the first of the required inequalities. We get the second by setting $z = 0$:

$$\left|\frac{f(z)}{z}\right|_{z=0} = |f'(0)| \leqslant 1. \tag{6}$$

If, now, $|f(z)/z| = 1$ at an interior point, either there are near-by points at which $|f(z)/z| > 1$, or else $f(z)/z = \text{const.}$ The first alternative is contrary to (5); so, if either of the inequalities of the theorem hold, we have $f(z)/z = e^{i\theta}$, or $f(z) = e^{i\theta}z$. Conversely, if $f(z)$ is so defined, the equalities hold.

As an application of Schwarz's lemma, we shall prove a theorem concerning the mapping of a region and of a subregion on the unit circle. This proposition is the function theoretic equivalent of a well-known property of the Green's function.

By a *subregion of* S is meant a region S_1 such that every interior point of S_1 is an interior point of S but not every interior point of S is an interior point of S_1.

THEOREM 2.—*Let* $w = f(z)$ *map a region* S *conformally on the unit circle* Q_0 *and let* $w_1 = f_1(z)$ *map a subregion* S_1 *of* S *on* Q_0, *the common point* a *being in each case carried to the origin. Then at any point of* S_1 *other than* a

$$|f_1(z)| > |f(z)|. \tag{7}$$

Also, if a *is an ordinary point,*

$$|f_1'(a)| > |f'(a)|. \tag{8}$$

When S is mapped on Q_0 by $w = f(z)$, S_1 is mapped on a subregion S_1' of Q_0. We now make the inverse of the transformation $w_1 = f_1(z)$ and follow it by the transformation $w = f(z)$. The first maps Q_0 on S_1; the second maps S_1 on S_1'. The succession of the two is a function $w = \varphi(w_1)$ which maps Q_0 on S_1'. So $|\varphi(w_1)| \leqslant 1$. Since $w_1 = 0$ goes into $w = 0$, Schwarz's lemma is applicable; and we have

$$|w| = |\varphi(w_1)| < |w_1|,$$

which is the desired result (7). It will be noted that the equality sign of the lemma cannot hold, since Q_0 is not mapped on itself.

At an ordinary point—that is, neither a branch point nor infinity—both functions have non-vanishing derivatives. We have then from Schwarz's lemma

$$|\varphi'(0)| = \left|\frac{dw}{dw_1}\right|_{z=a} = \left|\frac{f'(a)}{f_1'(a)}\right| < 1,$$

which establishes (8).

73. Area Theorems.—Let $f(z)$ be analytic within a circle Q. Then $z' = f(z)$ maps the interior of Q conformally upon a region S of one or many sheets—possibly an infinite number—the mapping being conformal except at the branch points of S at which it is continuous. We shall prove the following theorem relative to the areas of Q and S.[1]

THEOREM 3.—*The function $f(z)$, analytic within a circle of radius R and center a, maps the interior of the circle upon a region whose area A satisfies the inequality*

$$A \geq \pi|f'(a)|^2R^2. \tag{9}$$

In particular A may be infinite. The equality holds if, and only if, $f(z) = a_0 + a_1z$.

We can suppose, without restricting the generality, that $a = 0$. Let Q' be the circle $|z| \leqslant R' < R$. Q' is mapped on a region of finite area A'. The element of area is multiplied by $|f'(z)|^2$ in the mapping, so

$$A' = \iint |f'(z)|^2dxdy,$$

the integral being extended over Q'.

We have

$$f(z) = a_0 + a_1z + a_2z^2 + \cdots,$$

the series converging when $|z| < R$. Then,

$$|f'(z)|^2 = (a_1 + 2a_2z + 3a_3z^2 + \cdots)(\bar{a}_1 + 2\bar{a}_2\bar{z} + 3\bar{a}_3\bar{z}^2 + \cdots).$$

Putting the double integral into polar coordinates, $z = re^{i\theta}$, we have

$$A' = \int_0^{R'} rdr \int_0^{2\pi}(a_1 + 2a_2re^{i\theta} + \cdots)(\bar{a}_1 + 2\bar{a}_2re^{-i\theta} + \cdots)d\theta.$$

We now multiply the two series and integrate term by term. When the integer n is different from zero, we have $\int_0^{2\pi} e^{in\theta}d\theta = 0$, whence,

$$\begin{aligned} A' &= 2\pi\int^{R'} r[a_1\bar{a}_1 + 4a_2\bar{a}_2r^2 + 9a_3\bar{a}_3r^4 + \cdots]dr \\ &= \pi[a_1\bar{a}_1R'^2 + 2a_2\bar{a}_2R'^4 + \cdots + na_n\bar{a}_nR'^{2n} + \cdots]. \end{aligned}$$

[1] BIEBERBACH, L., *Palermo Rend.*, vol. 38, pp. 98–112, 1914.

Letting R' approach R, the sum of this series of positive terms either becomes infinite or approaches the sum $\pi[a_1\bar{a}_1R^2 + 2a_2\bar{a}_2R^4 + \cdots]$, provided this latter series converges. In either case,

$$A \geqslant \pi a_1\bar{a}_1R^2 = \pi|f'(0)|^2R^2.$$

The equality holds if, and only if, $a_2 = a_3 = \cdots = 0$, in which case, $f(z) = a_0 + a_1z$.

As a special case of this theorem, let lengths at the center of Q be unaltered in magnitude, so that $|f'(a)| = 1$. Then, $A \geqslant \pi R^2$; that is, the area of S equals or exceeds the area of Q.

We next consider the mapping of the exterior of the unit circle on a *plane* region where the mapping function is such that the point at infinity remains fixed and where the shift in the position of a sufficiently large z is bounded.

THEOREM 4.—*If the function*

$$w = f(z) = z + c_0 + \frac{c_1}{z} + \frac{c_2}{z^2} + \cdots \tag{10}$$

maps the exterior of the unit circle Q_0 on a plane region, then

$$c_1\bar{c}_1 + 2c_2\bar{c}_2 + 3c_3\bar{c}_3 + \cdots \leqslant 1. \tag{11}$$

Here, $f(z)$ is analytic outside Q_0 except at infinity and the series (10) converges when $|z| > 1$.

Consider the circle Q: $|z| = r > 1$. This is mapped by (10) on a simple closed analytic curve C in the w-plane. Writing $z = re^{i\theta}$, then, as θ increases from 0 to 2π, z moves counter-clockwise around Q and $w = f(re^{i\theta})$ moves counter-clockwise around C.

C encloses an area $A' > 0$ which we now proceed to find. Writing $w = X + iY$, we apply the familiar formula

$$A' = \tfrac{1}{2}\int_C XdY - YdX,$$

the line integral being taken counter-clockwise around C. We have

$$X = \frac{w + \bar{w}}{2}, \quad Y = \frac{w - \bar{w}}{2i}.$$

On substituting these values, we have the following result:

$$A' = \frac{1}{2}\int_0^{2\pi}\left(X\frac{dY}{d\theta} - Y\frac{dX}{d\theta}\right)d\theta = \frac{1}{4i}\int_0^{2\pi}\left(\bar{w}\frac{dw}{d\theta} - w\frac{d\bar{w}}{d\theta}\right)d\theta.$$

The second term in the last integrand is the negative of the conjugate imaginary of the first. We need merely integrate the first term and multiply its imaginary part by 2.

We have, from (10),

$$w = re^{i\theta} + c_0 + \frac{c_1}{r}e^{-i\theta} + \frac{c_2}{r^2}e^{-2i\theta} + \cdots,$$

$$\frac{dw}{d\theta} = i\left[re^{i\theta} - \frac{c_1}{r}e^{-i\theta} - \frac{2c_2}{r^2}e^{-2i\theta} - \cdots\right.$$

$$\bar{w} = re^{-i\theta} + \bar{c}_0 + \frac{\bar{c}_1}{r}e^{i\theta} + \frac{\bar{c}_2}{r^2}e^{2i\theta} + \cdots$$

On multiplying the last two series together and integrating term by term, employing the fact that $\int_0^{2\pi} e^{in\theta}d\theta = 0$ when $n \neq 0$, we have

$$\int_0^{2\pi} \bar{w}\frac{dw}{d\theta}d\theta = 2\pi i\left[r^2 - \frac{c_1\bar{c}_1}{r^2} - \frac{2c_2\bar{c}_2}{r^4} - \cdots\right].$$

Then,

$$A' = \pi\left[r^2 - \frac{c_1\bar{c}_1}{r^2} - \frac{2c_2\bar{c}_2}{r^4} - \frac{3c_3\bar{c}_3}{r^6} - \cdots\right] > 0.$$

Letting r approach 1, we have

$$A = \pi[1 - c_1\bar{c}_1 - 2c_2\bar{c}_2 - 3c_3\bar{c}_3 - \cdots] \geqq 0,$$

from which we have (11). Here A is the (outer) area of the part of the plane not covered by the map of the exterior of Q_0. It may be zero.

We observe that $A < \pi$ unless $f(z) = z + c_0$. Except in the case of a translation, the part of the plane left uncovered by the map has a smaller area than the part not covered by the original region $|z| > 1$.

74. The Mapping of a Circle on a Plane Finite Region.—We are now in a position to prove the following remarkable theorem:

THEOREM 5.—*Let $w = f(z)$ map the interior of the unit circle Q_0 on a plane finite region, subject to the conditions $f(0) = 0, f'(0) = 1$. Then, whatever the mapping function may be, the circle $|w| < \frac{1}{4}$ lies within the mapped region. Further, no other point of the w-plane is interior to all possible maps of the kind stated.*

The mapping function has the series expansion

$$w = f(z) = z + a_2z^2 + a_3z^3 + \cdots, \tag{12}$$

convergent when $|z| < 1$.

We find readily that the function

$$w = F(z) = [f(z^2)]^{\frac{1}{2}} = z + \tfrac{1}{2}a_2z^3 + \cdots \tag{13}$$

also gives a plane finite map of Q_0. For, $t = f(z^2)$ maps Q_0 in a one-to-one manner on a finite region lying on the two-sheeted surface with branch points at 0 and ∞; and $w = \sqrt{t}$ maps this two-sheeted region on a plane finite region.

If we now put $z = 1/Z$, $w = 1/W$, the function just formed gives a plane map of the exterior of Q_0 in the Z-plane on a plane region in the W-plane. We have

$$W = \frac{1}{F(1/Z)} = Z - \frac{a_2}{2Z} + \cdots \tag{14}$$

This is a mapping function to which Theorem 4 applies. We have then from (11) that $\frac{1}{4}a_2\bar{a}_2 \leqslant 1$. That is, if (12) gives a plane finite map of the interior of Q_0 then

$$|a_2| \leqslant 2. \tag{15}$$

Now let c be a finite point of the w-plane not lying in the map of the interior of Q_0 by (12); that is, c is an external or boundary point of the map. Obviously, $c \neq 0$. Then the function

$$w' = \frac{cf(z)}{c - f(z)} = z + \left(a_2 + \frac{1}{c}\right)z^2 + \cdots$$

gives a plane finite map of the interior of Q_0. So we have

$$\left|a_2 + \frac{1}{c}\right| \leqslant 2. \tag{16}$$

From (15) and (16) we have

$$\left|\frac{1}{c}\right| \leqslant 4; \ |c| \geqslant \frac{1}{4}. \tag{17}$$

Then, the points of the circle $|w| < \frac{1}{4}$ are all interior points in the map, which was to be proved.

We now set up a mapping function with a boundary point at the minimum distance from the origin in the map.

The sign of equality in (17) can hold only if $|a_2| = 2$. Taking $a_2 = 2$, then (14) has only two terms, as we see from (11). On working back from (14), we have the function

$$w = \frac{z}{(1 - z)^2} = \frac{1/z}{\left(1 - \frac{1}{z}\right)^2}. \tag{18}$$

We shall show that this function does, in fact, give a plane finite map of Q_0 of the kind mentioned in the theorem and that there is a boundary point of the mapped region on the circle $|w| = \frac{1}{4}$.

The function (18) has the required value and derivative at $z = 0$. Also, it is analytic within Q_0. To each value of w

correspond two values of z. If z_1 is one of them, $1/z_1$ is the other and not both can lie in Q_0. A value of w is then taken on once, at most, in Q_0; hence, we have a plane map. When $z = -1$ we have $w = -\frac{1}{4}$, a point on the circle $|w| = \frac{1}{4}$.

The boundary of the mapped region is got by setting $z = e^{i\theta}$ in (18), and letting θ vary from 0 to 2π. We have

$$w = \frac{e^{i\theta}}{(1 - e^{i\theta})^2} = \frac{1}{\left(e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}}\right)^2} = -\frac{1}{4 \sin^2 \frac{\theta}{2}}. \tag{19}$$

Then (18) maps the interior of Q_0 on the whole plane bounded by a slit along the real axis, from $-\frac{1}{4}$ to $-\infty$.

By combining (18) with suitable rotations about the origin in the z- and w-planes, we can set up a function

$$w = \frac{z}{(1 + e^{-i\alpha}z)^2}, \tag{20}$$

which maps the interior of Q_0 on the w-plane bounded by a slit beginning at any point $w = \frac{1}{4}e^{i\alpha}$ of $|w| = \frac{1}{4}$ and extending radially to infinity. Any point w such that $|w| \geqslant \frac{1}{4}$ lies on the boundary of one of these mapped regions. The circle $|w| < \frac{1}{4}$ is then the *complete locus* of points interior to all possible maps of the kind stated in the theorem.

COROLLARY.—*Let* $w = f(z)$ *map the circle* $|z - a| < \rho$ *conformally on a plane finite region S. Then the circle*

$$|w - f(a)| < \frac{|f'(a)|}{4}\rho \tag{21}$$

lies in S.

By a suitable change of variable this falls under the preceding theorem. Here $f'(a) \neq 0$, since the map is plane. We put

$$W = \frac{w - f(a)}{\rho f'(a)}, \quad Z = \frac{z - a}{\rho}.$$

Then the function

$$W = \frac{f(a + \rho Z) - f(a)}{\rho f'(a)} \tag{22}$$

maps $|Z| < 1$ on a plane finite region and satisfies the conditions of the theorem at $Z = 0$. Then $|W| < \frac{1}{4}$ lies in the map. From this we have (21) at once.

75. The Deformation Theorem for the Circle.—When a mapping is performed by means of a function $f(z)$, infinitesimal lengths in the neighborhood of z are multiplied by $|f'(z)|$. We shall now get limits for this deformation in the mapping considered in the preceding section.

THEOREM 6.—*Let* $w = f(z)$, *where* $f(0) = 0$, $f'(0) = 1$, *map the interior of the unit circle* Q_0 *on a plane finite region. Then at any point* $z = re^{i\theta}$ *within* Q_0

$$\frac{1 - r}{(1 + r)^3} \leqslant |f'(z)| \leqslant \frac{1 + r}{(1 - r)^3}. \tag{23}$$

Further, no closer limits hold for all mapping functions of the type stated.

Holding z fixed, we set up a linear transformation which carries Q_0 into itself and 0 to z; and follow this by a transformation involving $f(z)$ such that z is carried back to the origin, inserting such a factor that the derivative of the final transformation at the origin is 1. Such a sequence with t as the independent variable is the following:[1]

$$t' = \frac{t + z}{\bar{z}t + 1}, \quad w_1 = F(t) = \frac{f(t') - f(z)}{f'(z)(1 - z\bar{z})}. \tag{24}$$

This gives a finite plane map of $|t| < 1$ on the w-plane.

Differentiating in order to get the first terms of the expansion of $F(t)$, we have

$$F'(t) = \frac{f'(t')}{f'(z)(1 - z\bar{z})}\,\frac{1 - z\bar{z}}{(\bar{z}t + 1)^2} = \frac{f'(t')}{f'(z)(\bar{z}t + 1)^2},$$

$$F''(t) = \frac{f''(t')(1 - z\bar{z})}{f'(z)(\bar{z}t + 1)^4} - \frac{2\bar{z}f'(t')}{f'(z)(\bar{z}t + 1)^3}.$$

Setting $t = 0$, whence $t' = z$, we have

$$F(0) = 0, \ F'(0) = 1, \ F''(0) = \frac{f''(z)(1 - r^2)}{f'(z)} - 2\bar{z}.$$

We have, then,

$$w_1 = F(t) = t + \frac{1}{2}\left[\frac{f''(z)(1 - r^2)}{f'(z)} - 2\bar{z}\right]t^2 + \cdots \tag{25}$$

Applying the inequality (15) we have

$$\frac{1}{2}\left|\frac{f''(z)(1 - r^2)}{f'(z)} - 2\bar{z}\right| \leqslant 2,$$

or

$$\left|\frac{f''(z)}{f'(z)} - \frac{2\bar{z}}{1 - r^2}\right| \leqslant \frac{4}{1 - r^2}. \tag{26}$$

[1] BIEBERBACH, L., *Math. Zeit.*, vol. 4, pp. 295–305, 1919; *Glasnik prirod. drustva*, vol. 33, pp. 1–24, 1921; *Lehrbuch der Funktionentheorie*, vol. 2, pp. 87, 88.

Now let z move from the origin along a radius, so that

$$|dz| = |e^{i\theta}dr| = dr; \quad \bar{z}dz = re^{-i\theta}dz = rdr.$$

On integrating (26), using the fact that the absolute value of the integral is less than or equal to the integral of the absolute value, we have

$$\left| \int_0^z \frac{f''(z)dz}{f'(z)} - \int_0^r \frac{2rdr}{1 - r^2} \right| \leqslant \int_0^r \frac{4dr}{1 - r^2};$$

and

$$|\log f'(z) + \log (1 - r^2)| \leqslant 2 \log \frac{1 + r}{1 - r}.$$

Here the logarithms are real except $\log f'(z)$, which we write in terms of its real and imaginary parts,

$$|\log |f'(z)| + \log (1 - r^2) + i \arg f'(z)| \leqslant \log \left(\frac{1 + r}{1 - r}\right)^2. \quad (27)$$

Considering the real part of the expression appearing in the first member, we have

$$|\log [|f'(z)| \cdot (1 - r^2)]| \leqslant \log \left(\frac{1 + r}{1 - r}\right)^2;$$

or

$$-\log \left(\frac{1 + r}{1 - r}\right)^2 \leqslant \log [|f'(z)|(1 - r^2)] \leqslant \log \left(\frac{1 + r}{1 - r}\right)^2;$$

whence,

$$\left(\frac{1 - r}{1 + r}\right)^2 \leqslant |f'(z)|(1 - r^2) \leqslant \left(\frac{1 + r}{1 - r}\right)^2.$$

From this we have (23).

That no better limits are possible appears from a consideration of the mapping function (18)

$$f(z) = \frac{z}{(1 - z)^2}, \quad f'(z) = \frac{1 + z}{(1 - z)^3}.$$

At the real points $\pm r$, both limit values are taken on.

From (27) we have also a limit for the angle through which a lineal element at z is rotated in the mapping; namely,

$$|\arg f'(z)| \leqslant 2 \log \left(\frac{1 + r}{1 - r}\right). \quad (28)$$

COROLLARY.—*Let* $w = f(z)$ *map the circle* $|z - a| < \rho$ *on a plane finite region. Then, at any point within the circle the following inequalities hold*

$$\frac{1 - r}{(1 + r)^3} \leqslant \left|\frac{f'(z)}{f'(a)}\right| \leqslant \frac{1 + r}{(1 - r)^3}, \quad (29)$$

where

$$z = a + \rho re^{i\theta}.$$

This is proved immediately by applying (23) to the function (22).

THEOREM 7.—*Let* $w = f(z)$, *where* $f(0) = 0$, $f'(0) = 1$, *map the interior of the unit circle* Q_0 *on a plane finite region. Then at any point* $z = re^{i\theta}$ *within* Q_0

$$\frac{r}{(1+r)^2} \leqslant |f(z)| \leqslant \frac{r}{(1-r)^2}. \tag{30}$$

No closer limits hold for all mapping functions of the type stated.

On multiplying the second inequality of (23) by $|dz|$ and integrating along a radius from the origin to a point z, we have

$$\left|\int_0^z f'(z)dz\right| \leqslant \int_0^r \frac{1+r}{(1-r)^3}dr,$$

or

$$|f(z)| \leqslant \frac{r}{(1-r)^2}.$$

As to the first inequality, we note that $|f(z)|$ is the length of the line segment L joining $w = 0$ to $w = f(z)$ in the w-plane. If $|f(z)| < \frac{1}{4}$, this line segment lies in the mapped region and is the map of some curve C in Q_0 joining 0 to z. We have for the length of L

$$|f(z)| = \int_C |f'(z)|\ |dz|.$$

Here, the sum of which the integral is the limit is made up of non-negative real terms; so that we can replace each term by a smaller quantity and be assured that, on addition, the inequality persists. We have on C

$$|dz| = |e^{i\theta}dr + ire^{i\theta}d\theta| = |dr + ird\theta| \geqslant |dr|$$

and, using (23),

$$|f(z)| \geqslant \int_C \frac{1-r}{(1+r)^3}|dr| \geqslant \int_0^r \frac{1-r}{(1+r)^3}dr = \frac{r}{(1+r)^2}.$$

If $|f(z)| \geqslant \frac{1}{4}$, the first inequality of (30) holds without further investigation, since $\frac{1}{4} > r/(1+r)^2$.

The function

$$f(z) = \frac{z}{(1-z)^2}$$

attains both limits; whence no closer limits are possible.

It is an immediate consequence of the theorem that $w = f(z)$ maps the circle $|z| < \rho < 1$ on a region in the w-plane whose boundary lies in the ring formed by the two circles $|w| = \rho/(1+\rho)^2$ and $|w| = \rho/(1-\rho)^2$. This ring is independent of the particular mapping function used.

COROLLARY.—*Let $w = f(z)$ map the circle $|z - a| < \rho$ on a plane finite region. Then, at any point within the circle the following inequalities hold:*

$$\frac{r}{(1+r)^2}\rho|f'(a)| \leqslant |f(z) - f(a)| \leqslant \frac{r}{(1-r)^2}\rho|f'(a)|, \tag{31}$$

where

$$z = a + r\rho e^{i\theta}.$$

We find these inequalities on applying the preceding theorem to the function (22).

76. A General Deformation Theorem.—We shall now derive a deformation theorem for more general regions.

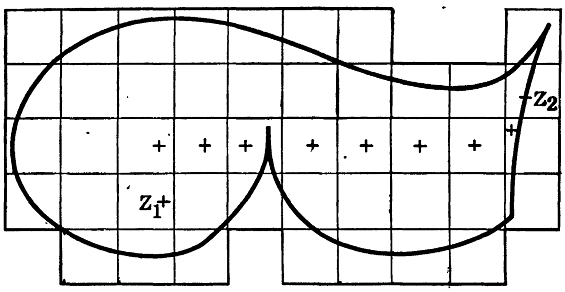


FIG. 50.

THEOREM 8.—*Let Σ' be a plane finite region and let Σ be a subregion whose boundary consists of interior points of Σ'. Let $w = f(z)$ map Σ' on a plane finite region.*

Then there exists a constant M, dependent on Σ and Σ' but independent of $f(z)$, such that if z_1, z_2 are any two interior or boundary points of Σ

$$\frac{1}{M} < \left|\frac{f'(z_1)}{f'(z_2)}\right| < M. \tag{32}$$

Since Σ, together with its boundary, consists of interior points of Σ' there exists a constant $d > 0$ such that a circle C with any point a within or on the boundary of Σ as center and with radius d lies within Σ'. C is then mapped on a plane finite region. For any point z whose distance from a does not exceed rd, $r < 1$, we have, from (29),

$$\left|\frac{f'(z)}{f'(a)}\right| \leqslant \frac{1+r}{(1-r)^3}. \tag{33}$$

We now rule parallels to the real and imaginary axes at a distance $h = \frac{1}{3}d$ apart, thus dividing the z-plane into squares (Fig. 50). Since Σ is a bounded region, only a finite number N

of these squares contain points of Σ. If z lies in the same square as a or in any of the four squares adjacent to it, we have $|z - a| \leqslant \frac{1}{3}\sqrt{5}d < \frac{3}{4}d$; and we have, from (33),

$$\left|\frac{f'(z)}{f'(a)}\right| < \frac{1 + \frac{3}{4}}{(1 - \frac{3}{4})^3} = m. \tag{34}$$

If z_1 and z_2 are two points of Σ, we can since Σ is connected, construct a chain of points of Σ:

$$z_1, \xi_1, \xi_2, \cdots, \xi_{n-1}, z_2, \tag{35}$$

such that adjacent points of the sequence are in adjacent squares (unless z_1 and z_2 are in the same square), the total number of points of the sequence, $n + 1$, not exceeding $N + 1$. We have, then,

$$\left|\frac{f'(z_1)}{f'(z_2)}\right| = \left|\frac{f'(z_1)}{f'(\xi_1)}\right| \cdot \left|\frac{f'(\xi_1)}{f'(\xi_2)}\right| \cdots \left|\frac{f'(\xi_{n-1})}{f'(z_2)}\right| < m^n \leqslant m^N.$$

Taking $M = m^N$, a quantity which depends only on the regions Σ and Σ', we have the second inequality of (32). Interchanging z_1 and z_2, we have the first inequality.

It will be noted that in the preceding proof we have not required that either Σ or Σ' be simply connected.

This theorem is readily extended to the case in which Σ is a closed point set, connected or disconnected, consisting only of interior points of Σ'. For, such a point set can be imbedded in a connected region which satisfies the conditions of the subregion in the theorem. The theorem can also be extended without difficulty to the case in which Σ' has a finite number of sheets, provided no branch point or point at infinity is an interior point of Σ'.

THEOREM 9.—*In the mapping of Theorem* 8 *there exists a constant L, independent of $f(z)$, such that if z_1, z_2, z_3 are any three interior or boundary points of Σ, then,*

$$|f(z_2) - f(z_1)| < L|f'(z_3)|. \tag{36}$$

We use again the chain (35). From (31) and (32), we have

$$|f(\xi_1) - f(z_1)| < \frac{\frac{3}{4}}{(1 - \frac{3}{4})^2}d|f'(z_1)| < \frac{\frac{3}{4}dM}{(1 - \frac{3}{4})^2}|f'(z_3)| = m'|f'(z_3)|,$$

say. Similarly,

$$|f(\xi_2) - f(\xi_1)| < m'|f'(z_3)|, \cdots, |f(z_2) - f(\xi_{n-1})| < m'|f'(z_3)|.$$

Then,

$$|f(z_2) - f(z_1)| \leqslant |f(\xi_1) - f(z_1)| + |f(\xi_2) - f(\xi_1)| + \cdots + |f(z_2) - f(\xi_{n-1}) < nm'|f'(z_3)| \leqslant Nm'|f'(z_3)|.$$

Taking $L = Nm'$, a quantity depending only on the regions Σ and Σ', we have (36).

77. An Application of Poisson's Integral.—Let $f(z) = U_z + iV_z$ be analytic in a circle Q of radius ρ and center at the origin and continuous on the boundary. Putting $z = re^{i\theta}$, and letting $t = \rho e^{i\psi}$ be a point on the circumference, Poisson's integral may be written in either of the following forms:

$$U_z = \frac{1}{2\pi}\int_0^{2\pi} U_t \frac{\left(1 - \frac{z\bar{z}}{\rho^2}\right)d\psi}{\left(1 - \frac{z\bar{t}}{\rho^2}\right)\left(1 - \frac{\bar{z}t}{\rho^2}\right)}, \tag{37}$$

$$U_z = \frac{1}{2\pi}\int_0^{2\pi} U_t \frac{(\rho^2 - r^2)d\psi}{\rho^2 - 2\rho r \cos(\theta - \psi) + r^2}. \tag{38}$$

Poisson's integral may be derived as follows: Consider the integral

$$J = \frac{1}{2\pi i}\int_Q \frac{f(t)dt}{(t - z)\left(1 - \frac{\bar{z}t}{\rho^2}\right)}.$$

The integrand has the single singularity at $t = z$ in Q, so, from the theory of residues,

$$J = \left[\frac{f(t)}{1 - \frac{\bar{z}t}{\rho^2}}\right]_{t=z} = \frac{f(z)}{1 - \frac{z\bar{z}}{\rho^2}} = \frac{U_z + iV_z}{1 - \frac{z\bar{z}}{\rho^2}}.$$

Again, since $dt = i\rho e^{i\psi}d\psi = itd\psi$, the integral can also be written

$$J = \frac{1}{2\pi}\int_0^{2\pi} \frac{f(t)td\psi}{(t - z)\left(1 - \frac{\bar{z}t}{\rho^2}\right)} = \frac{1}{2\pi}\int_0^{2\pi} \frac{U_t + iV_t}{\left(1 - \frac{z}{t}\right)\left(1 - \frac{\bar{z}t}{\rho^2}\right)} d\psi.$$

Now $1 - \frac{z}{t} = 1 - \frac{z\bar{t}}{t\bar{t}} = 1 - \frac{z\bar{t}}{\rho^2}$; the denominator that appears in the last integrand is the product of conjugate imaginaries and is, therefore, real. Equating the real parts in these two expressions for J, we have (37).

If $z = 0$, (37) becomes

$$U_0 = \frac{1}{2\pi}\int_0^{2\pi} U_t d\psi. \tag{39}$$

Suppose, now, that $U_z \geqslant 0$ in Q, and let z lie in or on a circle Q_λ concentric with Q and of smaller radius $\lambda\rho$. We have on Q_λ

$$1 - \frac{z\bar{z}}{\rho^2} = 1 - \lambda^2, \quad \left|\frac{z\bar{t}}{\rho^2}\right| = \frac{\lambda\rho\cdot\rho}{\rho^2} = \lambda, \quad \left|1 - \frac{z\bar{t}}{\rho^2}\right| \geqslant 1 - \lambda.$$

We have then, from (37) and (39),

$$U_z \leqslant \frac{1}{2\pi}\int_0^{2\pi} U_t \frac{1-\lambda^2}{(1-\lambda)^2}\,d\psi,$$

$$U_z \leqslant \frac{1+\lambda}{1-\lambda}U_0. \tag{40}$$

This inequality holds within Q_λ also.

We apply this result as follows: Let $f(z)$ have the further property that $|f(z)| \geqslant 1$ in Q. Then, $\log f(z)$ defined as follows:

$$\log f(z) = \log f(0) + \int_0^z d\log f(z) = \log f(0) + \int^z \frac{f'(z)dz}{f(z)},$$

where $\log f(0)$ is any one of the logarithms of $f(0)$, is analytic in Q and continuous on the boundary. The real part of this function, $\log |f(z)|$, is positive or zero throughout. If z is in Q_λ, we have, from (40),

$$\log |f(z)| \leqslant \frac{1+\lambda}{1-\lambda}\log |f(0)|,$$

whence,

$$|f(z)| \leqslant |f(0)|^{\frac{1+\lambda}{1-\lambda}}. \tag{41}$$

By a simple translation, we can apply formula (41) to a circle Q with center a and radius ρ. If $f(z)$ satisfies the conditions stated above in Q and if Q_λ is concentric with Q and of radius $\lambda\rho$ ($\lambda < 1$), then,

$$|f(z)| \leqslant |f(a)|^{\frac{1+\lambda}{1-\lambda}} \tag{42}$$

when z is in or on Q_λ.

The following theorem will be of frequent use in connection with subsequent convergence proofs.

THEOREM 10.—*Let Σ' be a plane finite region and let Σ be a subregion whose boundary consists of interior points of Σ'. Let $f(z)$ be a function which is analytic in Σ', is bounded, $|f(z)| \leqslant C$, and does not vanish. Then, there exist positive constants K_1, K_2, K such that if z_1, and z_2 are any two points within or on the boundary of Σ*

$$K_1|f(z_2)|^K \leqslant |f(z_1)| \leqslant K_2|f(z_2)|^{1/K}. \tag{43}$$

Here, K_1, K_2, K depend only on Σ, Σ', C, and are independent of $f(z)$.

The function $F(z) = C/f(z)$ is analytic in Σ' and $|F(z)| \geqslant 1$. We employ the chain of points (35) connecting z_1 and z_2. We have, from (42),

$$|F(\xi_1)| \leqslant |F(z_1)|^\mu, \quad \mu = \frac{1 + 3/4}{1 - 3/4}.$$

Similarly,

$$|F(\xi_2)| \leqslant |F(\xi_1)|^\mu \leqslant |F(z_1)|^{2\mu},$$

and so on; so that, finally,

$$|F(z_2)| \leqslant |F(z_1)|^{n\mu} \leqslant |F(z_1)|^{N\mu} = |F(z_1)|^K,$$

where we put $K = N\mu$. On replacing $F(z)$ by its value in terms of $f(z)$, we have, on simplifying,

$$|f(z_1)| \leqslant C^{1-\frac{1}{K}}|f(z_2)|^{\frac{1}{K}}.$$

This is the required second inequality (43). The first inequality is got by interchanging z_1 and z_2 and simplifying. We observe that K is independent of C.

78. The Mapping of a Plane Simply Connected Region on a Circle. The Iterative Process.—A plane simply connected region (p. 222) may consist (*a*) of the whole plane; (*b*) of the plane with the exception of a single boundary point; or (*c*) of a region with a bounding infinite point set, or curve.

In the first two cases it is impossible to map the region on a circle. If $w = f(z)$ maps (*a*) on the unit circle Q_0 with center at the origin (and if it can be mapped on any circle it can be mapped on Q_0), we have $|f(z)| \leqslant 1$ in the whole plane, so, by Liouville's theorem, $f(z) \equiv$ const., which is impossible. Similarly, if (*b*) is mapped on the unit circle we have $|f(z)| \leqslant 1$ in the neighborhood of the bounding point α, hence, the function is analytic at α also, if properly defined there. As before, $f(z) \equiv$ const., which is impossible.

THEOREM 11.—*A plane simply connected region whose boundary consists of more than one point can be mapped conformally on a subregion of the unit circle* Q_0.

Suppose, first, that there is a point z_0 of the plane which is neither an interior nor a boundary point of the given region S. Then, we can construct a circle Q' about z_0 such that no point within or on Q' is an interior or boundary point of S. Then, the linear transformation $z' = T(z)$ which carries Q' into Q_0 and carries the exterior of Q' into the interior of Q_0 maps S on a region S_0 lying within Q_0.

It may be, however, that every point of the plane is an interior or boundary point of S. Such is the case, for example, if S consists of the whole plane bounded by the positive real axis. In this case, let α and β be two boundary points, and consider the mapping of S by the function

$$z' = \sqrt{\frac{z - \alpha}{z - \beta}}, \text{ or } z' = \sqrt{z - \alpha} \text{ (if } \beta = \infty). \tag{44}$$

This function maps the two-sheeted surface Σ with a branch line joining α and β on the whole z'-plane. S is a subregion of Σ and is mapped on a plane region S' which covers only a part of the z'-plane. Thus, if z_1 is an interior point of S, the point

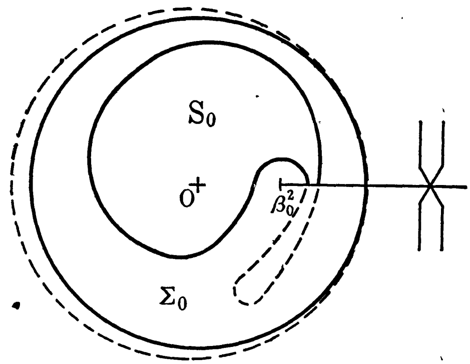


FIG. 51.

P with the same coördinate $z = z_1$ in the second sheet is a point of Σ which is exterior to S. On applying (44), P is carried into a point which is neither an interior nor a boundary point of S'. By the method of the preceding paragraph, S' can be mapped on a region lying within Q_0.

We consider now the problem of mapping a plane simply connected region S_0 which lies within Q_0 upon Q_0 itself. We shall suppose that the origin is an interior point of S_0. This can be secured, when not otherwise true, by making a linear transformation which carries Q_0 into itself and carries an interior point of the region to the origin. Our method of procedure will be to map S_0 upon a region S_1 which fills more of Q_0; to map S_1 upon a still larger region S_2 in Q_0; etc. By a suitable application of the process we shall arrive, in the limit, at a mapping of S_0 on Q_0.

Let β_0^2 (we use the square to avoid radicals) be an interior point of Q_0 and an exterior or boundary point of S_0. We now form a two-sheeted surface (Fig. 51) with a branch point at β_0^2 and with no other branch points within Q_0. Let Σ_0 be the part of this surface within the double circle which $|z| = 1$ cuts from the two sheets. We represent S_0 on this surface so that in the neighborhood of the origin S_0 lies, say, in the upper sheet. S_0 is a subregion of Σ_0.

We now map Σ_0 on the plane interior of Q_0 in the z_1-plane. The following sequence of transformations accomplishes this end:

$$s = \frac{z - \beta_0^2}{-\bar{\beta}_0^2 z + 1}, \quad t = \sqrt{s}, \quad z_1 = \frac{t + c}{\bar{c}t + 1}, \tag{45}$$

where $1 - c\bar{c} > 0$. For, the first transformation maps Σ_0 on a two-sheeted surface bounded by $|s| = 1$ (Equation (47), Sec. 12) and with branch point at the origin; the second maps this latter surface on a single-leaved surface $|t| < 1$; and the third maps this circle on the circle $|z_1| < 1$. The mapping is conformal except at the branch point $z = \beta_0^2$. The mapping of the subregion S_0 is conformal without exception.

We now impose the requirement that *in the mapping of* S_0 *the origin shall remain fixed.* When $z = 0$, $t = \pm i\sqrt{\beta_0^2}$. Let the square root be so chosen that $i\beta_0$ is the value of t when $z = 0$ in the upper sheet. If $z_1 = 0$ when $t = i\beta_0$, we find that $c = -i\beta_0$. Expressing z in terms of z_1, we then have

$$z = z_1 \frac{z_1 + \bar{B}_0}{B_0 z_1 + 1}, \qquad B_0 = -\frac{2i\bar{\beta}_0}{1 + \beta_0\bar{\beta}_0}. \tag{46}$$

We shall further require that *the derivative of the mapping function shall be real and positive at the origin.* This requirement, which prevents the rotation of the region, can be met by inserting a factor $e^{i\theta}$ in one member of (46). This amounts to rotating about the origin in the z-plane before making the sequence of transformations (45), and alters no other requirement of the mapping. From (46), we have

$$\left(\frac{dz}{dz_1}\right)_{z_1=0} = \bar{B}_0 = \frac{2i\beta_0}{1 + \beta_0\bar{\beta}_0}.$$

The derivative will be real and positive if we multiply the second member of (46) by $-i\bar{\beta}_0/\sqrt{\beta_0\bar{\beta}_0}$, which is of absolute value 1. The mapping function is then

$$z = -i\frac{\bar{\beta}_0}{\sqrt{\beta_0\bar{\beta}_0}} z_1 \frac{z_1 + \bar{B}_0}{B_0 z_1 + 1}. \tag{47}$$

The function just derived, and which we shall represent by $z_1 = f_1(z)$ maps S_0 conformally on a plane region S_1 lying on the interior of the circle Q_0 in the z_1-plane.

This map of S_0 has the important property that *the distance of each point (the origin excepted) from the origin is increased in the mapping.* This is an immediate consequence of Schwarz's lemma (Theorem 1). Formula (47) maps Q_0 in the z_1-plane on a two-sheeted region bounded by $|z| = 1$. We have then $|z(z_1)| \leqslant 1$; whence, applying the lemma, $|z| < |z_1|$. It follows, from this, that if h_0 is the shortest distance from the origin to the boundary of S_0 and h_1 is the shortest distance from the origin to the boundary of S_1, then $h_1 > h_0$.

We now select β_1^2, an interior point of Q_0 but not an interior point of S_1, and use a function analogous to (47) to map S_1 on a region S_2. S_2 will be similarly mapped on S_3; and so on. The general mapping function—that which maps S_n in the z_n-plane on S_{n+1} in the z_{n+1}-plane—is

$$z_n = -\frac{i\bar{\beta}_n}{\sqrt{\beta_n\bar{\beta}_n}} z_{n+1}\frac{z_{n+1} + \bar{B}_n}{B_n z_{n+1} + 1}, \qquad B_n = \frac{-2i\bar{\beta}_n}{1 + \beta_n\bar{\beta}_n}. \tag{48}$$

For corresponding points of S_n and S_{n+1}, other than the origin, we have $|z_{n+1}| > |z_n|$. If h_n is the shortest distance from the origin to the boundary of S_n, we have

$$h_{n+1} > h_n. \tag{49}$$

The point β_n^2 is within Q_0 and exterior to or on the boundary of S_n. We shall impose but one condition on its selection. *As n increases β_n^2 shall not approach the boundary of Q_0 unless h_n approaches* 1. This condition can be satisfied by requiring, say, that

$$|\beta_n^2| < \frac{1 + h_n}{2}. \tag{50}$$

(In previous applications of the method β_n^2 has been selected as the nearest point to the origin on the boundary of S_n. Then, $|\beta_n^2| = h_n$; and the condition is satisfied.)

The variable z_n which appears in the n-th step of the process is a function of z, which we shall represent by

$$z_n = f_n(z). \tag{51}$$

This function maps S_0 conformally on S_n.

79. The Convergence of the Process.—We shall now prove:

I. That $\lim h_n = 1$; that is, the boundary of S_n lies in the circular ring bounded by Q_0 and $|z_n| = h_n$, and as n increases the width of this ring approaches zero.

II. That when n becomes infinite, $f_n(z)$ approaches a function $f(z)$ which is analytic in S_0.

III. That $z' = f(z)$ maps the interior of S_0 conformally on the interior of the unit circle Q_0.

I. *The Movement of the Boundary Out to Q_0.*—The distance h_n increases with n and remains less than 1; so h_n approaches a limit $h \leqslant 1$. We shall show that $h < 1$ is impossible.

Consider the derivative at the origin. We have, from (48),

$$\left(\frac{dz_n}{dz_{n+1}}\right)_0 = -\frac{i\bar{\beta}_n}{\sqrt{\beta_n\bar{\beta}_n}}\bar{B}_n' = \frac{2\sqrt{\beta_n\bar{\beta}_n}}{1+\beta_n\bar{\beta}_n} = \frac{2r_n}{1+r_n^2}$$

where $|\beta_n| = r_n$. Now, if $h < 1$, r_n remains always less than a quantity less than 1; $r_n < 1 - \eta$, $\eta > 0$. We have, then,

$$\left(\frac{dz_{n+1}}{dz_n}\right)_0 = \frac{1+r_n^2}{2r_n} = 1 + \frac{(1-r_n)^2}{2r_n} > 1 + \epsilon \qquad (52)$$

where $\epsilon = \eta^2/2$. Then,

$$f_n'(0) = \left(\frac{dz_1}{dz}\right)_0\left(\frac{dz_2}{dz_1}\right)_0 \cdots \left(\frac{dz_n}{dz_{n-1}}\right)_0 > (1+\epsilon)^n. \qquad (53)$$

Now, consider the map of the circle $|z| < h_0$, which lies in S_0. Applying Theorem 3 the area A_n of the map in the z_n-plane satisfies the inequality

$$A_n \geqslant \pi|f_n'(0)|^2h_0^2 > \pi(1+\epsilon)^{2n}h_0^2.$$

By taking n large enough $(1+\epsilon)^{2n}$, and hence A_n, can be made as large as we please. But this is impossible; the map of the circle in question lies in Q_0, whence A_n is less than the area of Q_0. The hypothesis that $h < 1$ thus leads to a contradiction; hence $h = 1$.

II. *The Convergence.*—In order to prove that $f_n(z)$ converges in S_0 and that the limit function $f(z)$ is analytic in S_0 it suffices to prove the following: Given any region Σ which together with its boundary lies in S_0 and which contains the origin, then $f_n(z)$ converges uniformly in Σ. To establish this proposition we shall prove that, given $\epsilon > 0$, then for n sufficiently large the inequality

$$|f_{n+p}(z) - f_n(z)| < \epsilon \qquad (54)$$

holds in Σ for all positive integral values of p.

Again we consider conditions at the origin. We have $(f_n(z)/z)_{z=0} = f_n'(0)$. This derivative is positive by construction —in fact, is greater than 1—and increases with n. since (Equations (52) and (53))

$$f'_{n+1}(0) = f_n'(0)\left(\frac{dz_{n+1}}{dz_n}\right)_0 = \left[1 + \frac{(1 - r_n)^2}{2r_n}\right]f_n'(0) > f_n'(0). \quad (55)$$

Furthermore, we found at the end of (I) that $f_n'(0)$ is bounded. It follows that $f_n'(0)$ approaches a limit as n becomes infinite. Hence, given $\eta > 0$, there exists an n such that

$$|f'_{n+p}(0) - f_n'(0)| < \eta \quad (56)$$

for all positive integral p.

We now apply Theorem 10 to the function $\dfrac{f_{n+p}(z) - f_n(z)}{z}$.

We have just found that this function does not vanish at the orgin. It vanishes nowhere else in S_0 since $|f_{n+p}(z)| > |f_n(z)|$. Let Σ' be a region which with its boundary lies in S_0 and which contains the points of Σ and its boundary as interior points. The function is analytic in and on the boundary of Σ'; so it takes on its maximum absolute value on the boundary. On the boundary $|f_{n+p}(z)| \leqslant 1$, $|f_n(z)| \leqslant 1$, and $|z| \geqslant h'$ where h' is the shortest distance from the origin to the boundary of Σ'. Hence, for any z in Σ' and for any n and p

$$\left|\frac{f_{n+p}(z) - f_n(z)}{z}\right| \leqslant \frac{2}{h'}.$$

The conditions of Theorem 10 are satisfied hence (taking $z_1 = z$ and $z_2 = 0$)

$$\left|\frac{f_{n+p}(z) - f_n(z)}{z}\right| \leqslant K_2\left[f'_{n+p}(0) - f_n'(0)\right]^{1/K}.$$

Here, K_2 and K are independent of n and p, and the inequality holds for z in Σ. When n is so chosen that (56) holds we have in Σ

$$|f_{n+p}(z) - f_n(z)| \leqslant K_2\eta^{1/K}|z| \leqslant K_2\eta^{1/K}. \quad (57)$$

By taking η small enough the last member of this inequality is less than ϵ and (54) holds.

It follows from the uniform convergence of $f_n(z)$ in Σ that the limit function $f(z)$ is analytic in Σ. Since Σ may be chosen large enough to enclose any preassigned interior point of S_0, it follows that $f(z)$ is analytic in the whole interior of S_0.

III. *The Mapping.*—In order to prove that $z' = f(z)$ maps the interior of S_0 upon the interior of Q_0 we shall show that if α is in Q_0 then $f(z)$ takes on the value α once, and only once, in S_0. We shall base the proof upon a theorem of Hurwitz[1] which we shall now establish.

THEOREM 12.—*Let $f_n(z)$, $n = 1, 2, \ldots$, be analytic in a region S and continuous on the boundary, and as n becomes infinite let $f_n(z)$ converge uniformly to a limit function $f(z)$ in the closed region S. Further let $f(z)$ not vanish on the boundary. Then for n sufficiently large, $n > N$, $f_n(z)$ and $f(z)$ have the same number of zeros in S.*

We can draw a regular boundary C within S sufficiently near to the boundary of S that on and outside C in S $|f(z)| > K > 0$. For n sufficiently large, $n > N'$, owing to the uniform convergence, $|f(z) - f_n(z)| < K/2$ on and outside C, so that $|f_n(z)| > K/2$. Then $f(z)$ and $f_n(z)$ have no zeros in S on or outside C. On C, $f_n'(z)$ converges uniformly to $f'(z)$.

The number of zeros of $f(z)$ and $f_n(z)$, $n > N'$, within C are, respectively,

$$m = \frac{1}{2\pi i}\int_C \frac{f'(t)dt}{f(t)}, \quad m_n = \frac{1}{2\pi i}\int_C \frac{f_n'(t)dt}{f_n(t)};$$

so,

$$m - m_n = \frac{1}{2\pi i}\int_C \left[\frac{f'(t)}{f(t)} - \frac{f_n'(t)}{f_n(t)}\right]dt.$$

Owing to the uniform convergence on C we can, by taking n sufficiently large, $n > N \geqslant N'$, make the integrand less than $1/L$, where L is the length of C. We have, then,

$$|m - m_n| < \frac{1}{2\pi L}\int_C |dt| = \frac{L}{2\pi L} < 1.$$

But $|m - m_n|$ is an integer, and can be less than 1 only if it is zero. Hence $m = m_n$, which was to be proved.

Now, consider any interior point α of Q_0. Let $z_m = f_m(z)$ map S_0 on S_m where $h_m > |\alpha|$. Then S_n, $n > m$, contains α; and since the mapping is one-to-one, $f_n(z)$ takes on the value α once and only once. Let Q_m be a circle of radius λ and center at the origin, where $|\alpha| < \lambda < h_m$. Q_m lies in S_m and contains α. Let C_m be the curve in S_0 which $z_m = f_m(z)$ maps on Q_m; and let C be any curve in S_0 enclosing C_m.

[1] HURWITZ, A., *Math. Ann.*, vol. 33, p. 248, 1888.

The functions $f_n(z) - \alpha$ satisfy in C the convergence requirements of Theorem 12. Furthermore, the limit function $f(z) - \alpha$ does not vanish on C. For $z_n = f_n(z)$, $n \geqslant m$, maps C on a curve enclosing Q_m; so that, on C, $|f_n(z) - \alpha| > \lambda - |\alpha|$, and the same inequality holds for the limit function. It follows from Theorem 12 that $f_n(z) - \alpha$ and $f(z) - \alpha$ have, for n sufficiently large, the same number of zeros in C. The former has one zero, since α is taken on once in C; hence, the latter has one zero, and $f(z)$ takes on the value α once in C. Finally, since C may be taken large enough to include any inner point of S_0 it follows that $f(z)$ takes on the value α once and only once in S_0.

It is clear that $f(z)$ takes on, in S_0, no value on or outside Q_0. For we have in $S_0, |f_n(z)| < 1$; hence in the limit $|f(z)| \leqslant 1$. If $|f(z)| = 1$ at an interior point of S_0, we have $|f(z)| > 1$ at a neighboring point, which is impossible. The function $z' = f(z)$ then maps S_0 conformally on Q_0.

We have proved the following important theorem:[1]

THEOREM 13.—*The interior of any plane simply connected region whose boundary consists of more than one point can be mapped conformally upon the interior of the unit circle Q_0.*

Having proved the existence of one function which maps a region on Q_0, there arises the question of the existence of other mapping functions. This question is disposed of by the following general theorem which is easily established:

THEOREM 14.—*If $z' = f(z)$ maps a region S conformally upon the unit circle Q_0, then the most general function mapping S on Q_0 is*

$$Z = \frac{af(z) + \bar{c}}{cf(z) + \bar{a}}, \quad a\bar{a} - c\bar{c} = 1. \tag{58}$$

Let $Z = F(z)$ be any function mapping S conformally on Q_0. Then Z is a function of z', $Z = \varphi(z')$, which maps Q_0 in a one-to-one and conformal manner on itself. For, to each z' in Q_0 corresponds one point in S and to this latter corresponds one Z in Q_0. Conversely, to each Z in Q_0 corresponds one z' in Q_0. Further, the neighborhood of z' is mapped conformally on

[1] This theorem was first stated by Riemann in his Dissertation, 1851. The first proof for the most general simply connected plane region is due to Osgood, *Trans.* Amer. Math. Soc., vol. 1, pp. 310–314, 1900, who proved the existence of the Green's function for such a region.

The method used in the text was outlined by Koebe in *Gött. Nach.* in 1912.

the neighborhood of z and this latter is mapped conformally on the neighborhood of Z. From the form of the most general function mapping the interior of Q_0 conformally on itself, which we found in Theorem 24, Sec. 12, we have $Z = (az' + \bar{c})/(cz' + \bar{a})$, where $a\bar{a} - c\bar{c} = 1$.

It is clear, conversely, that (58) maps S on Q_0; for it is equivalent to the transformation $z' = f(z)$ which maps S on Q_0, followed by a linear transformation which maps Q_0 on itself.

COROLLARY.—*If S can be mapped conformally on Q_0 then there exists one and only one mapping function such that a given inner point z_0 of S is carried into the origin and a given direction at z_0 is carried into a given direction at the origin.*

Let $z' = f(z)$ be one function mapping S on Q_0. Then we can so determine the constants in (58) that $Z = 0$ when $z = z_0$. By a rotation about the origin in the map a lineal element issuing from z_0 can be given the desired direction at the origin.

Let $z_1 = f_1(z)$ and $z_2 = f_2(z)$ be two functions mapping S in the desired way. Then $z_2 = (\alpha z_1 + \bar{\gamma})/(\gamma z_1 + \bar{\alpha})$, $\alpha\bar{\alpha} - \gamma\bar{\gamma} = 1$. When $z = z_0$ we have $z_1 = 0$ and $z_2 = 0$; so $\bar{\gamma} = 0$ and we have $z_2 = \alpha z_1/\bar{\alpha} = e^{i\theta}z_1$. This latter is a rotation about the origin but, since there is a fixed lineal element issuing from the origin, this rotation is through a zero angle. Hence $z_2 \equiv z_1$; and there is but one mapping function of the required kind.

80. The Behavior of the Mapping Function on the Boundary.—The study of the behavior of the mapping function when the variable approaches the boundary has led, in recent years, to a number of brilliant papers. Certain of the results will be derived in the present section.

There are certain elementary cases where the results are immediate. If a plane region S is mapped on a plane region S' by a linear transformation, for example, the points of the two planes correspond in a one-to-one manner, and the correspondence between the boundary points, in particular, is one-to-one. Again, let the mapping function be analytic at a boundary point Z of S; and let $z_1, z_2, \ldots$ be a set of points of S approaching Z. Then the corresponding points $z_1', z_2', \ldots$ of S' approach the unique point $Z' = f(Z)$ on the boundary of S'. This simple fact enables us to determine readily the correspondence between boundary points in the two regions when the mapping is performed by elementary functions analytic throughout the plane except for a few isolated points.

In the study of the general problem we shall make repeated use of the following lemma:[1]

LEMMA.—*Let $f(z)$ be analytic and bounded, $|f(z)| < M$, within and on the boundary of the region ABCD, where AB is an arc of $|z| = 1$, CD is an arc of $|z| = r' < 1$, and BC and DA lie on radii (Fig. 52), except on the arc AB itself. Let $f(z)$ have the further property that in any sub-region $|z| > r$ and for a given $\epsilon > 0$ a cross-cut λ can be drawn from a point of BC to a point of AD along which $|f(z)| < \epsilon$. Then $f(z) \equiv 0$.*

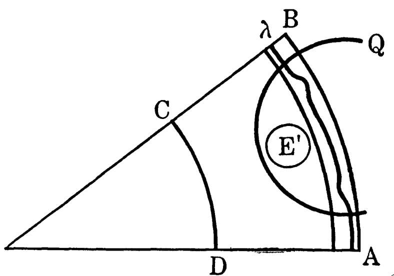


FIG. 52.

Let Q be a circle orthogonal to AB such that the region S common to Q and $|z| < 1$, or Q_0, lies in $ABCD$. By a suitable linear transformation $\tau = T(z)$ we can carry the arc of Q_0 lying in Q into the segment -1, 1 of the real axis and carry S into the upper half of the circle Q_0 in the τ-plane. (Fig. 53). Then, the transformation

$$\frac{t+1}{1-t} = -i\left(\frac{\tau+1}{1-\tau}\right)^2$$

maps this half circle on the whole circle Q_0 in the t-plane (Fig. 54) so that the segment -1, 1 goes into the lower half circumference

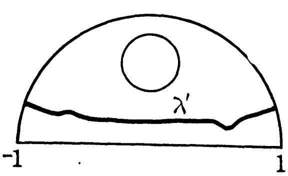


FIG. 53.

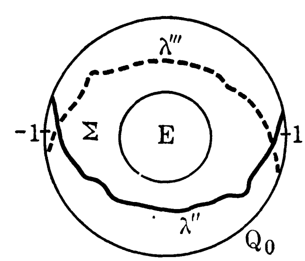


FIG. 54.

in the t-plane and the semicircular boundary goes into the upper half circumference. Further, both these transformations map the boundaries continuously. By these transformations, $f(z)$ goes over into a function $\varphi(t)$ analytic in Q_0 and on the upper half of the boundary.

[1] KOEBE, P., *Jour. für math.*, vol. 145, p. 213, 1915.

Let E be the circle $|t| \leqslant \rho < 1$; and let E' be the map of E in S. We now take r near enough 1 that $|z| = r$ encloses E'; and for a given ϵ we construct the curve λ. There is at least one arc of λ lying in S which has its ends on Q and which separates S into two regions, one of which contains E' and has no point of Q_0 on its boundary. The map of this arc is a curve λ'' in Q_0 in the t-plane whose ends lie on the upper half circumference. This arc divides Q_0 into two regions, of which one contains E and its boundary consists of λ'' and an arc of the upper half circumference not reaching to the points ± 1.

Let λ''' be the curve resulting from rotating λ'' through the angle π about the origin. Then certain arcs of λ'' and λ''' form the boundary of a region Σ lying entirely within Q_0 and containing E. We consider the functions $\varphi(t)$ and $\varphi(-t)$ in this region. If t is on λ''', $-t$ is on λ''. On λ'' we have $|\varphi(t)| < \epsilon$, $|\varphi(-t)| < M$, and on λ''', $|\varphi(t)| < M$ and $|\varphi(-t)| < \epsilon$. So on the boundary of Σ

$$|\varphi(t)\varphi(-t)| < M\epsilon.$$

Since the maximum absolute value of the function is attained at the boundary, this inequality holds also in E. Since ϵ may be made arbitrarily small, we have, in E $\varphi(t)\varphi(-t) \equiv 0$; whence $\varphi(t)$, or $f(z)$, vanishes identically, which was to be proved.

In considering the mapping of the boundary, when a plane region S is mapped on the unit circle Q_0, we may suppose, without loss of generality, that S lies within Q_0 and contains the origin. This is a consequence of Theorem 11. The mapping functions used in the derivation of that theorem are such that when the region in Q_0 is mapped back on the original region the functions employed are analytic on the boundary, except possibly for a few isolated points.

Accessible Boundary Points.—Let the boundary of S be represented by C. A point Z of C will be called an "accessible boundary point" if an interior point of S can be joined to Z by a continuous curve L which lies, except for its end point Z, entirely in S. We may suppose that L begins at the origin and has no multiple points. L is defined by an equation of the form $z = z(t)$, where $z(t)$ is a continuous function of the real variable t in an interval $t_1 \leqslant t \leqslant t_2$; $z(t_1) = 0$, $z(t_2) = Z$, and $z(t') \neq z(t'')$ if t' and t'' are distinct points of the interval. Such a curve is known as a "Jordan curve."

That simply connected regions exist not all of whose points are accessible is easily shown by examples. The region of Fig. 55, formed from a rectangle by making certain cuts from the boundary into the interior, is simply connected. The cuts into the right half of the rectangle are made as follows: Let $AC = h$, $CD = k$. At points of AC whose distance from A are $h/2$, $3h/4$, $5h/6$, $7h/8$, . . . we erect lines perpendicular to AC of length $2k/3$. At points of ED whose distances from E are $2h/3$, $4h/5$, $6h/7$, . . . we erect perpendiculars to ED of length $2k/3$. Then, all points of the line CD are boundary points, but none of them is accessible.

The lines issuing radially from A are all of length l, where $l < h/2$, $l < k$, and make angles $\pi/4$, $\pi/8$, $\pi/16$, . . . with AC. Here the interior points of the line AB are inaccessible boundary points. On the other hand, the point P to which the spiral boundary curve converges is an accessible boundary point.

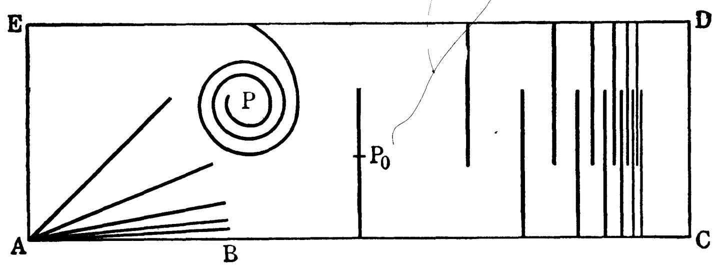


FIG. 55.

In the case of a boundary point, such as P_0 of Fig. 55, to which we can draw continuous curves approaching from one side or the other side of the line on which the point lies, it is desirable, for the simpler statement of our results, to consider P_0 as a *different* accessible boundary point according as it is approached from one side or the other. We shall consider the accessible boundary point as defined not only by its position but by the curve which is drawn to it.

Let Z_1, Z_2 be accessible boundary points defined by the curves L_1, L_2. If $Z_1 \neq Z_2$ they are different accessible points. If $Z_1 = Z_2$ they shall be the same accessible boundary point if, and only if, the following conditions are satisfied: Let a circle C be drawn with Z_1 as center, and of radius sufficiently small that the origin is exterior to C. As L_1 is traced from the origin toward Z_1, let P_1 be the last point of L_1 on the circumference of C. The arc N_1 from P_1 to Z_1 lies entirely in C. Similarly, the arc N_2 of L_2 from the last intersection of L_2 with C lies in C. If, now, for every such circle C, any interior point of the arc N_1 can be joined to any interior point of the arc N_2 by a Jordan curve lying entirely in S and C, then the two accessible points are the same.

With this definition, all curves drawn to P_0 in the figure from the right give the same accessible point; all curves drawn to P_0 from the left give the same accessible point, but different from the former. We thus have two accessible points at P_0. The point A of the figure counts as an infinite number of accessible points.

Let $z' = f(z)$ be a function mapping the interior of S conformally on the interior of the unit circle Q_0. We now prove a series of propositions[1] concerning the behavior of $f(z)$ as z approaches points on the boundary of S.

PROPOSITION 1.—*Let Z be an accessible boundary point of S, defined by the curve L. Then as z approaches Z along L, z' approaches a point Z' on the circumference of Q_0.*

It is clear that if a variable inner point z of S approaches a boundary point P in any way, $|f(z)|$ approaches 1. For an arbitrarily small $\epsilon > 0$, the circle $|z'| < 1 - \epsilon$ is mapped on a region S' of S which, together with its boundary, consists of inner points of S. When z lies in the circle $|z - P| < \eta$ where η is small enough that this circle contains no points of S', then $1 - \epsilon < |z'| < 1$; whence, $|z'|$ or $|f(z)|$ approaches 1.

As z traces L from the origin toward Z, z' traces a continuous curve L' from the origin. L' has the equation $z' = z'(t) = f[z(t)]$, where $z'(t)$ is a continuous function of t in the open interval $t_1 \leqslant t < t_2$. L' is without double points. As z approaches Z, $|z'|$ approaches 1, and there are one or more cluster points of L' on the circumference of Q_0. If there is but one, the proposition is established. Suppose there are two cluster points Z_1', Z_2'. Then, as z approaches Z, L' passes infinitely often from the neighborhood of Z_1' to that of Z_2'. Hence, all the points of one of the two arcs into which Z_1' and Z_2' divide the circumference are cluster points. We now construct a region of the type of Fig. 52 with AB within an arc of cluster points. For $|z - Z| < \epsilon$, there is an arc of L' extending from a point of BC to a point of DA and lying outside any given circle $|z'| = r < 1$. Representing the inverse of the mapping function by $z = \varphi(z')$, we have on this arc $|\varphi(z') - Z| < \epsilon$. Throughout Q_0 we have $|\varphi(z') - Z| \leqslant 2$. Applying the lemma, $\varphi(z') - Z \equiv 0$, or $\varphi(z')$ is a constant, which is impossible. This establishes the proposition. The function $z'(t)$, defining L', is continuous in the closed interval $t_1 \leqslant t \leqslant t_2$, if we put $z'(t_2) = Z'$.

[1] KOEBE, P., *Jour. für Math.*, vol. 145, pp. 215–218, 1915. Osgood, W. F., and Taylor, E. H., *Trans.* Amer. Math. Soc., vol. 14, pp. 277–298, 1913.

PROPOSITION 2.—*Let L_1 and L_2 be curves defining the same accessible point Z on the boundary of S. Then, the corresponding curves L_1', L_2' end at the same point of the circumference of Q_0.*

Suppose L_1' and L_2' to end at different points Z_1', Z_2'. In any circle C, however small, with Z as center, we can draw a curve λ lying in S and in C from a point ξ of L_1 to a point η of L_2. This construction is possible since the two curves define the same accessible point. Let ξ', η', λ' be the maps of ξ, η, λ in the z'-plane. If the radius of C is small enough, ξ' is arbitrarily near Z_1', η' is arbitrarily near Z_2', and λ' lies without a given circle $|z'| = r < 1$. Along λ', if the radius of C is less than ϵ, we have $|z - Z| = |\varphi(z') - Z| < \epsilon$. By taking the arc AB in the lemma on a suitable one of the two arcs into which Z_1', Z_2' separate the circumference, we have, as in the preceding proposition, $\varphi(z') - Z \equiv 0$. This contradiction establishes the proposition.

PROPOSITION 3.—*Let L_1 and L_2 be curves defining different accessible boundary points Z_1 and Z_2 on the boundary of S. Then the corresponding curves L_1', L_2' end at different points on the circumference of Q_0.*

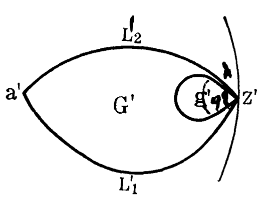


FIG. 56.

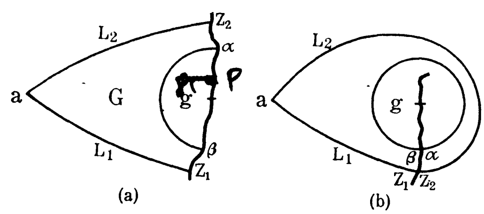


FIG. 57.

The proof here is less simple. Suppose, on the contrary, that L_1' and L_2' end at the same point Z'. As z traces L_1 from the origin toward Z_1, let a (which may be the origin) be the last point of L_1 within S which lies on L_2. Such a point clearly exists, since, otherwise, L_1 and L_2 define the same accessible point. We shall erase the parts of L_1 and L_2 between the origin and a and consider the curves as drawn from a to Z_1 and Z_2. Then L_1' and L_2' extend from the corresponding point a' in Q_0 to the point Z' and have no other common point.

The curves L_1 and L_2 together form a cross-cut in S extending from a boundary point to a boundary point which divides S into two simply connected pieces. The map in Q_0 of one of these parts, which we shall call G' (Fig. 56), is bounded completely by L_1' and L_2' and has the single point Z' of the circumference on its boundary. Let G (Fig. 57(*a*) or (*b*)) be the corresponding region in S.

The boundary of G contains in addition to L_1 and L_2, a piece of the boundary of S. This is obvious if $Z_1 \neq Z_2$. It is true also if $Z_1 = Z_2$; for, otherwise, we could draw in G the curves whose existence suffices to make L_1 and L_2 define the same accessible point.

We next make a cross-cut in G in the form of an arc of a circle joining different accessible boundary points of S. That such a cross-cut is possible we see as follows: Let P be a point on the boundary of G and of S distinct from Z_1 and Z_2. Let P_1 be an inner point of G whose distance from P is less than the distance from P to the points of L_1 and L_2. The arc with P as center and passing through P_1, extended in each direction from P_1 until it meets the boundary at two points α, β, is a cross-cut of the desired kind. (This arc meets the boundary, otherwise S is not simply connected.) One of the two simply connected parts into which G is separated is bounded by the arc $\alpha\beta$ and by a piece of the boundary of S, but by no points of L_1 or L_2. Call this region g.

The reason for getting a cross-cut in the form of a circular arc is that with this particular form of boundary we can map g on a region to which we can apply the lemma. We now show that g can be mapped on a semicircle in such a way that the circular arc $\alpha\beta$ is mapped continuously on a diameter. We first make a linear transformation $\tau = T(z)$ carrying the arc $\alpha\beta$ into a segment $\alpha'\beta'$ of the real axis, and so that g is mapped on a region g_1 in the τ-plane lying, in the neighborhood of $\alpha'\beta'$, in the upper half plane. We may take $\alpha'\beta'$ as a finite segment except in the special case that $\alpha = \beta$. We may suppose that g_1 is entirely in the upper half plane. If not, a second transformation of the form

$$\frac{\tau' - \alpha'}{\beta' - \tau'} = \left(\frac{\tau - \alpha'}{\beta' - \tau}\right)^{1/2}$$

maps the τ-plane bounded by the slit $\alpha'\beta'$ on the upper half τ'-plane. (A second transformation is clearly unnecessary if $\alpha' = \beta'$.) The boundary points of g and g_1 correspond in a one-to-one and continuous manner.

Let g_1' be the reflection of g_1 in the real axis (Fig. 58). We now map the simply connected region $g_1 + g_1'$, formed by erasing the segment $\alpha'\beta'$, on the unit circle Q_0 in the t-plane in such a way that an inner point c of the segment $\alpha'\beta'$ is carried to the origin and the direction of the positive real axis at c is carried into the direction of the positive real axis at the origin (Fig. 59). This determines the mapping function $t = F(\tau)$ uniquely (Theorem 14, Corollary). In this mapping, the interior of the segment $\alpha'\beta'$ is mapped continuously on an analytic curve in Q_0. We shall now show that this curve is precisely the segment $-1, 1$.

If we reflect in the real axis in the τ-plane, thus carrying $g_1 + g_1'$ into itself, perform the mapping with the function F, and then reflect in the real axis in the t-plane, we have a conformal map of $g_1 + g_1'$ on Q_0. The succession of these transformations

$$\tau_1 = \bar{\tau}, \quad t_1 = F(\tau_1), \quad t = \bar{t}_1; \quad \text{or } t = \bar{F}(\bar{\tau}),$$

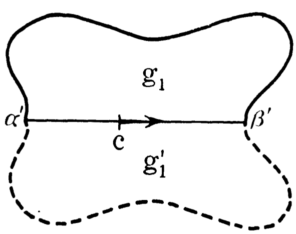


FIG. 58.

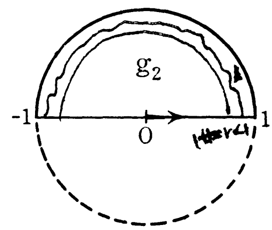


FIG. 59.

is an analytic function of τ which carries c to the origin, $t = 0$, and transforms the direction of the positive real axis at c as before. Hence, $\bar{F}(\bar{\tau}) = F(\tau)$, since there is but one mapping function with these properties. If τ is not real, $\tau \neq \bar{\tau}$, we have in $g_1 + g_1'$, since the mapping is one-to-one, $F(\tau) \neq F(\bar{\tau})$, whence $F(\tau) \neq \bar{F}(\bar{\tau})$. That is, $\bar{t} \neq t$, and t is not real. Hence, the real segment $\alpha'\beta'$ is mapped on the real segment $-1, 1$. Incidentally, from Proposition 1, the one-to-one correspondence may be extended to include the end points themselves. The region g_1 is mapped on the region g_2, the upper half of Q_0.

We shall now use the region g_2 for the application of the lemma. On changing the variable from z to t, we have a function, $z' = f(z) = \psi(t)$, which maps g_2 on a region g' lying in G'. The boundary of g' is a continuous map of $\alpha\beta$ and, hence, of the interval $-1, 1$. If we cut g_2 in two by any circle $|t| = r < 1$, the map of this cut is a curve q cutting g' in two. The part of

g_2 exterior to the circle corresponds to the part of g' which has Z' as boundary point. In this latter region, we can draw a curve λ' within a distance ϵ of Z' cutting off a smaller subregion with Z' on its boundary. The map of λ' is a curve λ in g_2 lying outside $|t| = r < 1$, joining a point of the real axis to the left of the origin to one on the right, and along which $|\psi(t) - Z'| < \epsilon$. Since $|\psi(t) - Z'|$ is bounded in Q_0, it follows from the lemma that $\psi(t) - Z' \equiv 0$; or $f(z)$ is constant. This contradiction establishes Proposition 3.

PROPOSITION 4.—*The points which correspond to the accessible points of the boundary of S are everywhere dense on the circumference of Q_0.*

Suppose, on the contrary, that there is an arc $Z_1'Z_2'$ on the circumference containing no points which correspond to accessible boundary points of S. Let Z' be an inner point of this arc; and let z_1', z_2', . . . be a sequence of interior points of Q_0 approaching Z' as a limit. The corresponding points z_1, z_2, . . . in S have one or more cluster points on the boundary of S. Let Z be one of these cluster points; then, we can select a subsequence ξ_1, ξ_2, . . . of z_1, z_2, . . . approaching Z as a limit. The corresponding sequence ξ_1', ξ_2', . . . approaches the limit Z'. Let L_n be a straight line drawn from ξ_n to the nearest (or one of the nearest) boundary points. L_n defines an accessible boundary point; hence, the corresponding curve L_n' in Q_0 extends from ξ_n' to a point of the circumference outside the interval Z_1' Z_2'. Now, given $\epsilon > 0$ and $r < 1$ we can choose n large enough, $n > N$, that all points of L_n lie within a distance ϵ of Z and outside the map in S of $|z'| \leqslant r$. L_n' is then a curve lying outside the circle $|z'| = r$ along which $|z - Z| = |\varphi(z') - Z| < \epsilon$. By choosing AB of Fig. 52 to lie on a suitable one of the arcs $Z'Z_1'$ or $Z'Z_2'$ (one with an infinite number of the curves L'_n in its neighborhood), and applying the lemma, we have $\varphi(z') - Z \equiv 0$. This identity is impossible; and the proposition is established.

Boundary Elements.[1]—We now interchange the rôles of Q_0 and S and enquire what point or points on the boundary of S correspond to a point Z' on the circumference of Q_0. If a continuous curve L' be drawn from the interior of Q_0 to Z', what point or points on the boundary of S does the corresponding curve L come arbitrarily near to—for all positions of L'? An

[1] These are the *Prime Ends* of Carathéodory's notable paper, *Math. Ann.* vol. 73, pp. 323–370, 1913.

equivalent problem is to consider a sequence of interior points of Q_0, z_1', z_2', . . . approaching Z' as a limit and to seek the cluster points of the corresponding sequence z_1, z_2, . . . in S—for all possible such sequences.

We consider a set of curves C_1, C_2, . . . in S. Let C_1 be a Jordan curve joining different accessible boundary points of S and lying, except at its ends, in S. Let S_1 be one of the simply connected pieces cut off. Let C_2 be a Jordan curve in S_1 joining accessible points on the boundary of S, different from one another and from the ends of C_1. Let S_2 be that part which C_2 cuts off from S_1 which does not have C_1 as part of its boundary. In general, C_{n+1} is a Jordan curve in S_n joining accessible points on the boundary of S, different from one another and from the ends of C_n; and S_{n+1} is that region which C_{n+1} cuts off from S_n which does not have C_n as part of its boundary.

We shall impose one further condition on the curves C_n. Any given interior point of S shall, for n sufficiently large, $n > N$, be exterior to the region S_n. This can be accomplished, for example, by so drawing C_n that S_n contains no points in the map of the circle $|z'| < 1 - \frac{1}{n}$.

Let C_n', S_n' be the maps of C_n, S_n in Q_0.

The closed regions S_1, S_2, . . . have at least one common point on the boundary of S. Likewise the closed regions S_1', S_2', . . . have at least one common point on Q_0. If there are more than one, there is a common arc on Q_0.

DEFINITION.—*If the closed regions* S_1', S_2', . . . *have a single common point* Z', *then the points common to the closed regions* S_1, S_2, . . . *will be said to constitute a boundary element of* S *defined by the curves* C_1, C_2, . . . *and corresponding to* Z'.

Two sets of curves will be said to define the same boundary element if the corresponding points on Q_0 *are the same.*

It is easily seen that two sets of curves defining boundary elements define the same element if, and only if, any region S_m of the first set contains all regions s_n of the second set, for n sufficiently large, and any s_m contains all S_n, for n sufficiently large. For, the corresponding regions S_n' and s_n' in Q_0, which enclose a single point Z', have this property.

To each boundary element corresponds, by definition, a unique point of Q_0. We now prove that to each point Z' on Q_0 corresponds a boundary element.

In the following, we represent by U_n', V_n' points on the boundary of Q_0 which correspond to accessible points U_n, V_n on the boundary of S. Let U_1', V_1' be an arc containing Z'. Let the points of the sequence U_1', U_2', . . . , V_1', V_2', . . . be so chosen that U_{n+1}' lies between U_n' and Z', V'_{n+1} lies between V_n' and Z', and the length of the arc $U_n'V_n'$ approaches zero. Each arc of the sequence $U_1'V_1'$, $U_2'V_2'$, . . . contains the arc which follows it; and Z' is the only point common to all.

We can now draw the curves C_1, C_2, . . . in the manner explained above such that C_n joins U_n and V_n. Then S_1', S_2', . . . have the single common point Z'; whence C_1, C_2, . . . define a boundary element corresponding to Z'.

We have then the following result:

PROPOSITION 5.—*There is a one-to-one correspondence between the points on the circumference Q_0 and the boundary elements of S.*

We now justify our nomenclature by showing that the boundary element constitutes the set of points corresponding to Z' in the sense mentioned originally. Each region of the sequence S_1', S_2', . . . in Q_0 encloses the region which follows it; and a given neighborhood of Z' contains all the regions S_n' from a certain value of n on. Of the points of the sequence z_1', z_2', . . . approaching Z' as a limit, all but a finite number lie in any S_n'. Then, all but a finite number of points of the corresponding sequence z_1, z_2, . . . lie in S_n. The cluster points of the latter sequence all lie within or on the boundary of S_n, for all n. These cluster points must then be points of the boundary element. Conversely, let Z be a point of the boundary element. We can choose z_1 in S_1, z_2 in S_2, etc., in such a way that z_1, z_2, . . . approaches Z as a limit. Then the corresponding sequence z_1' in S_1', z_2' in S_2', . . . approaches Z' as a limit. Thus every point of the boundary element corresponds to Z'.

We now derive certain propositions concerning sets of curves of the kind we have defined.

PROPOSITION 6.—*If S_1, S_2, . . . do not possess more than one common accessible boundary point, then C_1, C_2, . . . define a boundary element.*

It is clear from Proposition 3 that a boundary element cannot possess two accessible boundary points. Suppose, now, that S_1', S_2', . . . have a common bounding arc. Let Z_1', Z_2' be points within this arc corresponding to the accessible points Z_1, Z_2 on the boundary of S. To the curve L_1, L_2 defining Z_1, Z_2

there correspond curves L_1', L_2' terminating in Z_1', Z_2'. In the neighborhood of Z_1', L_1' lies in S_n'; so, in the neighborhood of Z_1, L_1 lies in S_n. Hence, Z_1 and, likewise, Z_2 are accessible boundary points for each region S_n. Since we assume there are not two such points, it follows that S_1', S_2', . . . have no common bounding arc; hence, we have a boundary element.

In Fig. 55, let C_n be drawn vertically across the region from boundary to boundary at a distance ϵ_n from CD, where ϵ_n tends steadily to zero as n increases. The regions S_n have the common boundary line CD, all points of which are inaccessible, and have no other common boundary points. Hence C_1, C_2, . . . define a boundary element comprising the points of CD. When this region is mapped conformally on a circle, any sequence of points z_1, z_2, . . . in the region, such that the distance of z_n from CD approaches zero as n becomes infinite, is mapped on a sequence z_1', z_2', . . . converging to a single point on the boundary of the circle.

PROPOSITION 7.—*If all points on the curve C_n for n sufficiently large lie in an arbitrarily small neighborhood of a point Z, then C_1, C_2, . . . define a boundary element.*

If we suppose the contrary, then S_1', S_2', . . . have a common boundary arc AB. Using this arc, and the function $z - Z = \varphi(Z') - Z$, we apply the lemma at the beginning of this section. For $\epsilon > 0$ given and for n sufficiently great, $|\varphi(z') - Z| < \epsilon$ on C_n', where C_n' is a curve lying near the arc AB. We conclude in the usual way that $\varphi(z')$ is a constant; and this contradiction establishes the proposition.

In Fig. 55, we can draw the curves from the ends of the radial slits to the neighborhood of B in such a manner that the curves converge to B. Then, a boundary element is defined comprising the points of the line AB. This boundary element contains one accessible point; namely, B.

PROPOSITION 8.—*If the diameter of the curve C_n approaches zero as n becomes infinite, then C_1, C_2, . . . define a boundary element.*

By the diameter of C_n is meant the maximum distance between two points of C_n. We can choose a subsequence C_{n_1}, C_{n_2}, . . . converging to a point Z. By Proposition 7 this subsequence defines a boundary element. Each S_{n_m} encloses all S_n for $n > n_m$; hence, C_1, C_2, . . . define a boundary element.

81. Regions Bounded by Jordan Curves.—As a consequence of the preceding propositions, the following theorem is readily established:

THEOREM 15.—*Let $z' = f(z)$ map a plane region S whose boundary C consists of a closed Jordan curve conformally on the unit* *circle Q_0. Then the points on the boundaries of the two regions correspond in a one-to-one and continuous manner.*

If at each point Z of C we set $f(Z) = \lim f(z)$, as z approaches Z from the interior of S, then $f(z)$ is continuous in the closed region S.

C is without double points and each of its points is an accessible boundary point. To each point of C, then, there corresponds a unique point on the circumference of Q_0. To each point of the circumference there corresponds a boundary element consisting of one or more points of C. But, since this boundary element cannot contain two different accessible points, it must consist of a single point. The correspondence is thus one-to-one.

It is a well-known theorem that when a continuous function in a two-dimensional region takes on boundary values in this manner, the sequence of values taken on is continuous on the boundary and forms with the values within the region a continuous function.

Numerous consequences of Theorem 15 will occur to the reader. For example, if we map S on a second plane region Σ bounded by a closed Jordan curve Γ, then the points of C and Γ correspond in a one-to-one and continuous manner; and the mapping function is continuous in the closed region S, if we define the value of the mapping function on C by the limiting process mentioned in the theorem. This we see by mapping S on Q_0 and then mapping Q_0 on Σ.

Many of the facts concerning the mapping on the boundary can be extended to a wide variety of many-sheeted regions. We shall find use for the following theorem and corollary:

THEOREM 16.—*Let S be a plane or many-sheeted region which is mapped conformally on Q_0. Let the boundary of S contain a Jordan curve C_1 with the property that a Jordan curve K_1 can be drawn in S connecting the ends of C_1 and cutting off a plane subregion S_1 whose complete boundary consists of C_1 and K_1. Then the points of C_1 correspond in a one-to-one manner to the points of an arc of Q_0.*

When S is mapped on Q_0, S_1 is mapped on a plane simply connected region S_1' and, so, can be mapped on a circle. S_1' is bounded by a Jordan curve. For, K_1 is mapped on a Jordan curve,—this includes the end points, by the reasoning of Proposition 1—and the rest of the boundary is on the circumference of Q_0. Then, the points on the boundaries of S_1 and S_1', each boundary being a closed Jordan curve, correspond in a one-to-one manner. Hence, the points of C_1 correspond to the points of an arc of Q_0.

It may happen that the boundary of a sheeted region, while not lying in a plane, can be broken up into pieces, each of which is a Jordan curve of the kind mentioned in the theorem. Each piece corresponds to an arc of Q_0, and we have the following result:

COROLLARY.—*If the boundary of S in Theorem* 16 *consists of a finite number of Jordan curves,* A_1A_2, A_2A_3, . . . , A_nA_1, *of the kind specified, then the points of the boundary and the points of the circumference of* Q_0 *correspond in a one-to-one manner.*

When the mapping is continuous on the boundary, there arises the question of performing the mapping so that certain given points on the boundary of one region shall be mapped on given points on the boundary of the other region.

THEOREM 17.—*If two regions S and S′ of the type stated in Theorem* 16, *Corollary, can each be mapped conformally on a circle, then S can be mapped conformally on S′ in one, and only one, way such that three given distinct points on the boundary of S are mapped on three given distinct points arranged in the same order on the boundary of S′.*

Both S and S' can be mapped on Q_0, the mapping being continuous on the boundary. Let m, n, p be three points on the boundary of S, arranged so that, proceeding in a positive sense around the boundary from m, we encounter first n, then p. Let m', n', p' be three points arranged in like manner on the boundary of S'. Let $t = \varphi(z)$ and $t' = \psi(z')$ map S and S', respectively, on Q_0. Then m, n, p and m', n', p' are carried, respectively, into m_1, n_1, p_1 and m_1', n_1', p_1' ordered in the same way on the circumference of Q_0. Now, there is a linear transformation $t' = T(t)$, which carries m_1, n_1, p_1, respectively into m_1', n_1', p_1'. This transformation carries Q_0 into itself, the interior being carried into the interior. Then the transformation $t = \varphi(z)$, followed by the transformation $t' = T(t)$, followed by the inverse of $t' = \psi(z')$, is a sequence of transformations which maps S on S' and carries m, n, p, into m', n', p', respectively.

Suppose, now, that this mapping can be performed by two functions $z' = f(z)$ and $z' = F(z)$. Then, $t' = \psi(z') = \psi[f(z)]$ and $t_1 = \psi[F(z)]$ map S on Q_0. Hence (Theorem 14), t_1 is a linear function of t'. But when $z = m, n, p$, we have $t' = t_1 = m_1, n_1, p_1$; hence (Sec. 3, Theorem 6), $t_1 \equiv t'$, or $\psi[F(z)] \equiv$

$\psi[f(z)]$. It follows that $F(z) \equiv f(z)$, and there is only one function performing the mapping.

82. Analytic Arcs and the Continuation of the Mapping Function across the Boundary.—We return to the mapping of a region S on Q_0. If the mapping function $z' = f(z)$ is analytic at Z on the boundary of S, so that the function can be extended analytically across the boundary, and if $f'(z) \neq 0$, then the inverse function $z = \varphi(z')$ is analytic at $Z' = f(Z)$ on the boundary of Q_0. The neighborhood of Z is mapped conformally on the plane neighborhood of Z'. The boundary near Z' is mapped on an *analytic* arc $z = \varphi(e^{i\theta})$ through Z.[1]

If $f'(Z) = 0$, we have $f'(z) \neq 0$ in the neighborhood of Z except at Z itself; and the boundary of S in the neighborhood of Z consists of analytic arcs. In order that $f(z)$ be capable of continuation across the boundary it is necessary that the boundary contain one or more analytic arcs. If there are no analytic arcs on the boundary C of S, then $f(z)$ has C as a natural boundary.

THEOREM 18.—*If the Jordan curve C_1 in Theorem 16 is an analytic arc, then $f(z)$ is analytic at any interior point Z of C_1 and so can be extended analytically across the boundary in the neighborhood of Z. The neighborhood of Z is mapped conformally by the function.*

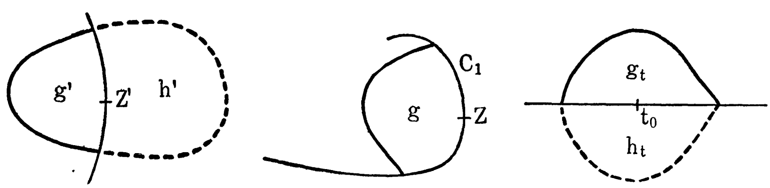


FIG. 60.

Let $z = \psi(t)$ be the equation of C_1, where $\psi(t_0) = Z$, $\psi'(t_0) \neq 0$. The function $z = \psi(t)$ maps the neighborhood of t_0 conformally on the neighborhood of Z so that the real axis near t_0 is mapped on the arc C_1 near Z. A suitably small region g_t (Fig. 60) abutting along the real axis in the neighborhood of t_0—in the upper half plane, let us say—is mapped on a region g of S and abutting along C_1 in the neighborhood of Z. This is mapped by $z' = f(z)$ on g' in Q_0 abutting along an arc through Z', the point

[1] An analytic curve is a Jordan curve, $z = z(t)$, $t_1 \leqslant t \leqslant t_2$, in which $z(t)$ is an analytic function of t in the interval and the derivative $z'(t) \neq 0$. In the present case θ takes the place of t.

corresponding to Z. Then, $z' = f[\psi(t)]$ maps g_t on g', is analytic in g_t, and is continuous on the part of the real axis bounding g_t.

Let h_t, h' be the inversions of g_t and g' in the real axis and the circle Q_0, respectively. If we invert in the real axis carrying h_t into g_t, then map g_t on g', and, finally, invert in Q_0, carrying g' into h', we have a conformal map of h_t on h' by some function $z' = F(t)$, analytic in h_t and continuous on the part of the real axis bounding h_t. Now, this transformation affects the points of the real axis in exactly the same way as the transformation $z' = f[\psi(t)]$. These two functions are analytic in the adjoining regions g_t, h_t and take on the same continuous values along the common boundary. Hence, $F(t)$ is the analytic continuation of $f[\psi(t)]$ throughout h_t. This latter function is analytic in the full neighborhood of t_0; whence, $f(z)$ is analytic in a sufficiently small neighborhood of Z. The mapping of the neighborhood of Z on the neighborhood of t_0 and of this latter on the neighborhood of Z' is in each case conformal; whence, $z' = f(z)$ maps the neighborhood of Z conformally on the neighborhood of Z'.

83. Circular Arc Boundaries.—The following theorem which is of considerable importance in various applications may be derived here:

THEOREM 19.—*Let $z' = f(z)$ map a region S upon a region S' in such a way that a circular arc AB of the boundary of S is mapped continuously on a circular arc $A'B'$ on the boundary of S'. Let S_1 and S_1' be the regions got by inverting S and S' in AB and $A'B'$, respectively. Then $f(z)$ can be extended analytically across AB into S_1 and maps S_1 conformally on S_1'.*

The proof follows the lines of the proof of Theorem 18. Let z_1 be a point of S_1; and let z be the point in S got by inverting in AB; let $z' = f(z)$; and let z_1' be the point in S_1' got by inverting z' in $A'B'$. Then z_1' is a function of z_1, $z_1' = F(z_1)$, which, as a result of one analytic transformation and two inversions, maps S_1 directly conformally on S_1'. Hence, $F(z_1)$ is analytic in S_1. As z_1 and z approach a common point on AB, z_1' and z' approach a common point on $A'B'$. Hence, $F(z_1)$ and $f(z)$ take on the same values on AB. It follows that $F(z_1)$ is the analytic continuation of $f(z)$ across AB into and throughout the region S_1.

COROLLARY.—*If the region S' of Theorem 19 is the unit circle Q_0, and if AB is not the whole boundary of S, then the region formed by joining S and S_1 along AB can be mapped on Q_0.*

In this case, $A'B'$ is an arc of Q_0 but not the whole circumference; and S_1' is the whole exterior of Q_0. The function $z' = f(z)$ maps the region formed by joining S and S_1 along AB upon the region formed by joining the interior and exterior of Q_0 along the arc $A'B'$. The latter region is a plane region whose boundary consists of that part of the circumference of Q_0 which remains after $A'B'$ has been removed. By Theorem 13 it can be mapped on Q_0.

84. The Mapping of Combined Regions.—We shall presently consider the mapping of a many-sheeted region on a circle.

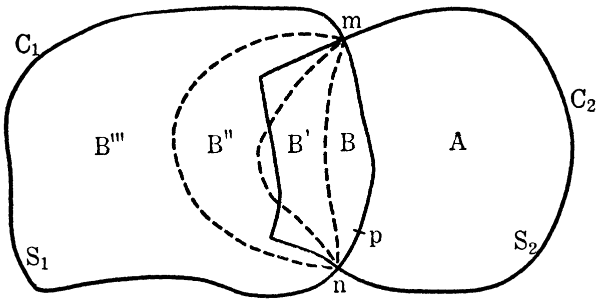


FIG. 61.

As an aid to establishing later a much more general theorem, we shall prove the following proposition:

THEOREM 20.—*Let S_1 and S_2 be regions, with boundaries of the type stated in Theorem* 16, *Corollary, each of which can be mapped conformally on a circle. Let the boundaries C_1, C_2 meet in two points m, n, the region common to S_1 and S_1' being simply connected. Let m, n be interior points of analytic arcs on both boundaries, and let the angle between the bounding arcs of the common region at each point be different from zero. Then, the region formed by combining S_1 and S_2 can be mapped on a circle.*

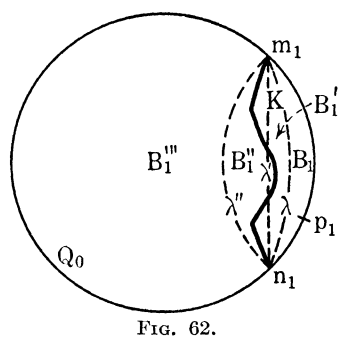


FIG. 62.

The regions S_1 and S_2 are shown schematically in Fig. 61. Let p be a point of C_1 lying in S_2. Let $z_1 = f_1(z)$ map S_1 conformally on Q_0. The points m, n, p are carried into points m_1, n_1, p_1 on the circumference of Q_0 (Fig. 62). The region common to S_1 and S_2 is mapped on a region K in Q_0 bounded by the arc $n_1p_1m_1$. The mapping function is analytic at m and n;

hence, angles at these points are preserved. Hence, the portion of C_2 which lies in S_1 is mapped on a curve in Q_0 (the heavy line in Fig. 61) meeting the arc $n_1p_1m_1$ at m_1 and n_1 at angles greater than zero.

We can draw a circular arc λ joining m_1 and n_1 and lying entirely in K. It will be convenient to take the arc such that it makes with the arc $n_1p_1m_1$ an angle $\theta = \pi/2^s$, where s is an integer. Let λ' be the arc got by inverting $n_1p_1m_1$ in λ; let λ'' be the arc got by inverting $n_1p_1m_1$ in λ'; and so on. We find readily that $\lambda^{(s)}$ lies on the circumference of Q_0. These arcs divide Q_0 into regions B_1, B_1', B_1'', . . . , $B_1^{(s)}$ as in the figure. Let B, B', B'', . . . , $B^{(s)}$ be the corresponding regions in S_1.

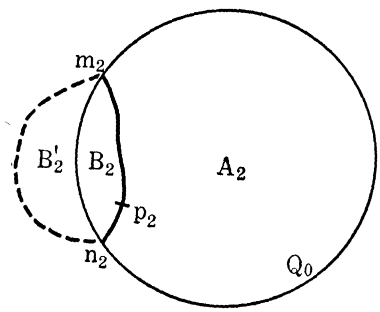


FIG. 63.

Let A be the portion of S_2 exterior to S_1. Let $z_2 = f_2(z)$ map the region formed by combining A and B, which we shall designate by $A + B$, conformally on Q_0. This mapping is clearly possible, for, when S_2 is mapped on a circle, $A + B$ is mapped on a plane region which can be mapped on Q_0 (Theorem 13). Let A_2, B_2, m_2, n_2, p_2 be the regions and points on which A, B, m, n, p are mapped (Fig. 63).

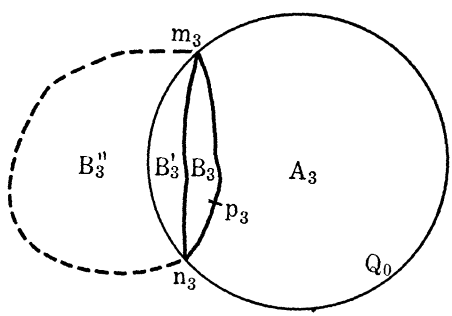


FIG. 64.

The function $f_2(z)$ can be extended analytically across the boundary of B; we shall show that it can be extended throughout the whole of B'. Let $z_2' = \varphi_2(z_1)$ be the unique function mapping B_1 on B_2 so that m_1, n_1, p_1 are carried into m_2, n_2, p_2, respectively.

Then (Theorem 19) this function can be extended analytically across λ and throughout B_1' and maps B_1' on B_2', the inverse of B_2 in Q_0. The function $z' = \varphi_2(z_1) = \varphi_2[f_1(z)]$ exists throughout $B + B'$ and maps $B + B'$ on $B_2 + B_2'$. But, in B, $\varphi_2[f_1(z)] \equiv f_2(z)$, since both functions map the three boundary points m, n, p alike (Theorem 17). Hence, $\varphi_2[f_1(z)]$ is the analytic continuation of $f_2(z)$. The function $f_2(z)$ maps $A + B + B'$ conformally on $A_2 + B_2 + B_2'$. This latter region is a plane region and can be mapped on Q_0.

We now repeat the reasoning, replacing B by $B + B'$ in the argument. Let $z_3 = f_3(z)$ map $A + B + B'$ on Q_0 (Fig. 64). The function $z_3 = \varphi_3(z_1)$ which maps $B_1 + B_1'$ on $B_3 + B_3'$ so that m_1, n_1, p_1 go into m_3, n_3, p_3 can be extended analytically throughout B_1'' and maps B_1'' on B_3'' the inverse of $B_3 + B_3'$ in Q_0. By the identity of $\varphi_3[f_1(z)]$ and $f_3(z)$ in $B + B'$, we prove that $f_3(z)$ can be extended analytically throughout B'' and maps B'' on B_3''. Hence, $z_3 = f_3(z)$ maps $A + B + B' + B''$ on a plane region which can be mapped on Q_0.

Continuing in this manner we arrive in a finite number of steps at a function which maps $A + B + B' + \cdots + B^{(s)}$ on Q_0; and the theorem is proved.

85. The Mapping of Limit Regions.—By the method of the preceding section, we can build up, step by step, by the combination of a finite number of overlapping plane regions, other regions which can be mapped on a circle. These latter regions are finite sheeted and are bounded by Jordan curves. We shall now prove a theorem wherein the final region is constructed by an infinite process. By its use we shall be able to deduce results concerning regions with general boundaries and regions with an infinite number of sheets. This powerful theorem is due to Koebe.

THEOREM 21.—*Let ϕ_1, ϕ_2, . . . be an infinite sequence of regions each of which can be mapped on a circle. Let ϕ_n be a subregion of ϕ_{n+1}. Let ϕ be the region consisting of all points which are interior to any region ϕ_n. Then, ϕ can be mapped conformally either on a circle or on the whole plane exclusive of a single point.*

Let a be an ordinary interior point of ϕ_1. Then, a is within ϕ_n. Let $z_n = f_n(z)$ map ϕ_n on the unit circle Q_0 in the z_n-plane in such a manner that a is carried into the origin and the direction of the positive real axis at a is carried into the direction of the positive real axis at the origin. Then $f_n(a) = 0$; and $f_n'(a) > 0$. This mapping is possible according to the corollary to Theorem 14.

We shall now study the sequence of functions

$$f_1(z), f_2(z), f_3(z), \cdots \tag{59}$$

The regions ϕ_n are shown schematically at the left of Fig. 65. The mapping of the first few regions is shown at the right in the same figure. Let z be a point in ϕ_m and let

$$z_m = f_m(z), \quad z_n = f_n(z), \; n > m \tag{60}$$

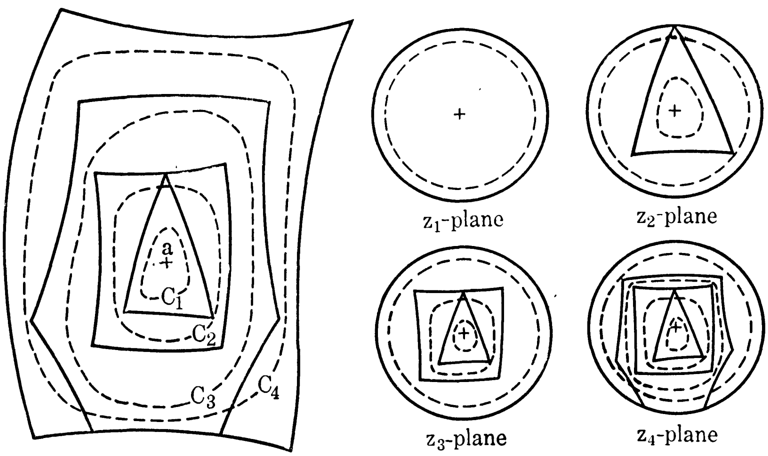


FIG. 65.

be the corresponding points in the z_m- and z_n-planes. The first of these functions maps ϕ_m on Q_0 in the z_m-plane; the second maps ϕ_m on a region lying in Q_0 in the z_n-plane. The inverse of the first function followed by the second,

$$z_n = f_n[f_m^{-1}(z_m)] = \varphi_{n,m}(z_m), \tag{61}$$

maps Q_0 in the z_m-plane on a region lying in Q_0 in the z_n-plane. The function $\varphi_{n,m}(z_m)$ satisfies the conditions of Schwarz's lemma, Theorem 1; hence, we have

$$|z_n| = |\varphi_{n,m}(z_m)| < |z_m|, \; z_m \neq 0; \quad \left|\frac{dz_n}{dz_m}\right|_0 = |\varphi'_{n,m}(0)| < 1. \tag{62}$$

From (60), since

$$\left(\frac{dz_n}{dz_m}\right)_0 = \frac{f_n'(a)}{f_m'(a)},$$

we have

$$|f_n(z)| < |f_m(z)|, \; z \neq a; \quad f_n'(a) < f_m'(a). \tag{63}$$

In particular,

$$|f_{n+1}(z)| < |f_n(z)|, \; z \neq a; \quad f'_{n+1}(a) < f_n'(a). \tag{64}$$

Let a circle Q with center at the origin and radius $r < 1$ be drawn in each plane. Let C_n be the corresponding curve in ϕ_n which the function $z_n = f_n(z)$ maps on Q. The curves C_1, C_2, enclose a set of regions $S_1, S_2, \ldots$ lying in ϕ, such that S_n lies in ϕ_n and contains a, and S_n together with its boundary C_n lies within S_{n+1}. This last statement follows at once from the fact that the interior and boundary of Q in the z_n-plane are, from the first inequality of (62), mapped on a region lying within Q in the z_{n+1}-plane.

Further, r can be taken near enough to 1 that any preassigned point P of ϕ is an inner point of the regions $S_m, S_{m+1}, \ldots$ for m sufficiently large. Let ϕ_m contain P. Then $z_m = f_m(z)$ maps ϕ_m on Q_0, P going into an interior point P'. It then suffices to take r large enough that Q contains P'.

The derivative $f_n'(a)$ is positive and (Inequality (64)) decreases as n increases. Hence, it approaches a limit $h \geqslant 0$ as n becomes infinite. We have two cases to consider:

$$\text{I. } \lim f_n'(a) = h > 0; \quad \text{II. } \lim f_n'(a) = 0$$

The Convergence.—We prove first that *in both cases the sequence $f_n(z)$ converges throughout ϕ and the limit function is analytic in ϕ.* To prove this, it suffices to prove that the sequence converges uniformly in S_m. The limit function is then analytic in S_m. Since m and r can be so chosen that S_m encloses any preasssigned point of ϕ, the limit function is then analytic throughout ϕ.

Instead of considering functions of z in S_m, we shall map S_m on Q in the z_m-plane and consider the corresponding functions of z_m. We have $f_n(z) = \varphi_{n,m}(z_m)$; and we shall establish the uniform convergence of the sequence

$$\varphi_{m,m}(z_m),\ \varphi_{m+1,m}(z_m), \varphi_{m+2,m}(z_m), \cdots \tag{65}$$

in Q. Consider the function

$$\frac{\varphi_{n+p,m}(z_m) - \varphi_{n,m}(z_m)}{z_m}, \quad n \geqslant m, p > 0. \tag{66}$$

This function is nowhere zero in Q_0; for, when $z_m \neq 0$, we have from (63) that $|\varphi_{n+p,m}(z_m)| < |\varphi_{n,m}(z_m)|$, and when $z_m = 0$ we have

$$\left(\frac{\varphi_{n+p,m}(z_m) - \varphi_{n,m}(z_m)}{z_m}\right)_0 = \left(\frac{dz_{n+p}}{dz_m}\right)_0 - \left(\frac{dz_n}{dz_m}\right)_0$$
$$= \frac{f'_{n+p}(a) - f_n'(a)}{f_m'(a)} < 0.$$

The function is also bounded in Q_0, for from (62) $|\varphi_{n,m}(z_m)/z_m| < 1$; so

$$\left|\frac{\varphi_{n+p,m}(z_m) - \varphi_{n,m}(z_m)}{z_m}\right| < 2. \tag{67}$$

The function (66) then satisfies all the conditions of Theorem 10, where the circular regions Q and Q_0 take the place of Σ and Σ', respectively, of that theorem. Hence, we have (setting $z_1 = z_m$ and $z_2 = 0$)

$$\left|\frac{\varphi_{n+p,m}(z_m) - \varphi_{n,m}(z_m)}{z_m}\right| \leqslant K_2\left|\frac{f'_{n+p}(a) - f_n'(a)}{f_m'(a)}\right|^{1/K}, \tag{68}$$

where K_2 and K are independent of m, n, p, and depend only upon r, the radius of Q. The inequality (68) holds for all points in Q.

Since the sequence $f_n'(a)$ converges, we can choose n large enough that the second member of (68) is less than a preassigned positive ϵ. We have, then, within and on Q

$$|\varphi_{n+p,m}(z_m) - \varphi_{n,m}(z_m)| \leqslant \epsilon|z_m| < \epsilon \tag{69}$$

for all $p > 0$. This proves the uniform convergence of the sequence (65) in Q.

In Case II, we find readily that *the limit function is identically zero.* We have but to apply Theorem 10 to the function $\varphi_{n,m}(z_m)/z_m$. We have in Q

$$\left|\frac{\varphi_{n,m}(z_m)}{z_m}\right| \leqslant K_2\left|\frac{\varphi_{n,m}(z_m)}{z_m}\right|_0^{1/K} = K_2\left(\frac{f_n'(a)}{f_m'(a)}\right)^{1/K}. \tag{70}$$

By taking n large enough, we can make $f_n'(a)$ arbitrarily small, whence, $|\varphi_{n,m}(z_m)|$ is arbitrarily small. It follows that

$$\lim_{n=\infty} \varphi_{n,m}(z_m) = 0.$$

Case I.—Let $f(z)$ be the limit of the sequence (59). We shall prove that the $z' = f(z)$ *maps* ϕ *on the circle* Q_0.

The function $f(z)$ is nowhere greater than 1 in absolute value in ϕ, since at each point of ϕ it is the limit of functions which are less than 1 in absolute value. It will suffice to prove that $f(z)$ takes on each value in Q_0 once, and only once. When ϕ_m is mapped on Q_0 by $z_m = f_m(z)$ let $f(z)$ be transformed into a function $\varphi_m(z_m)$. Then when z is in ϕ_m, z_m is in Q_0, and

$$f(z) = \varphi_m(z_m) = \lim_{n=\infty} \varphi_{n,m}(z_m).$$

Let α be a point in Q_0; and let r be taken large enough that Q encloses α. This fixes K_2 and K in (68). Taking $n = m$ we have,

$$\left|\frac{\varphi_{m+p,m}(z_m) - z_m}{z_m}\right| \leqslant K_2 \left|\frac{f'_{m+p}(a) - f_m'(a)}{f_m'(a)}\right|^{1/K}. \tag{71}$$

Let $\epsilon > 0$ be chosen so that

$$\epsilon < r - |\alpha|. \tag{72}$$

Then, since $f_n'(a)$ converges, we can choose m sufficiently large that the numerator in the fraction in the second member of (71) is arbitrarily small, whereas the denominator $f_m'(a)$ is greater than h. We can make the second member less than ϵ, and we have

$$|\varphi_{m+p,m}(z_m) - z_m| \leqslant \epsilon\, |z_m| < \epsilon. \tag{73}$$

Letting p become infinite we have

$$|\varphi_m(z_m) - z_m| \leqslant \epsilon. \tag{74}$$

Now $\varphi_{m+p,m}(z_m)$ maps Q and its interior on a plane region in the z_{m+p}-plane. On Q itself $|z_m| = r$ and we have from (73) and (74)

$$r - \epsilon < |\varphi_{m+p,m}(z_m)| < r + \epsilon;\ r - \epsilon \leqslant |\varphi_m(z_m)| \leqslant r + \epsilon. \tag{75}$$

It follows from the first of these inequalities, together with (72), that Q is mapped on a curve enclosing α. Hence, for all $p > 0$, $\varphi_{m+p,m}(z_m) - \alpha$ has a single zero in Q. It follows from the latter part of (75) that $\varphi_m(z_m) - \alpha$ is not zero on Q. Applying Theorem 12, it results that $\varphi_m(z_m) - \alpha$ has a single zero in Q. That is, $\varphi_m(z_m)$ takes on the value α once, and only once, in S_m. Then $f(z)$ takes on the value α once, and only once, in S_m.

We show finally that $f(z)$ cannot take on the value α at any other point of ϕ. Suppose, on the contrary, that this value is taken on at two points of ϕ. Then, r $(>|\alpha|)$ and ν can be taken large enough that both points are included in S_ν. If in the preceding reasoning we choose $m \geqslant \nu$, we have the contradiction that $f(z)$ takes on the value α more than once in S_m.

Case II.—If $\lim f_n'(a)$ is zero, the limit of the sequence (59) is identically zero; and we get no map of the surface ϕ. In this case, we alter the mapping as follows: Instead of mapping ϕ_n on Q_0, we map it on a circle Q_n with center at the origin and radius $1/f_n'(a)$. We take as the mapping function

$$Z_n = F_n(z) = \frac{f_n(z)}{f_n'(a)} \tag{76}$$

This maps a on the origin, and, since $F'(a) = 1$, lengths at a are unaltered. The radius of Q_n increases with n and approaches infinity as n becomes infinite. The mapped regions are got from the circles shown in Fig. 65, by an expansion from the origin according to the formula $Z_n = z_n/f_n'(a)$. Q is carried into a circle Q_n' of radius $r/f_n'(a)$.

We shall show that the sequence

$$F_1(z),\ F_2(z),\ F_3(z),\ \cdots \tag{77}$$

converges in ϕ, that the limit function $F(z)$ is analytic in ϕ, and that $Z' = F(z)$ maps ϕ on the whole finite plane.

When ϕ_n is mapped on Q_n in the Z_n-plane by (76), each subregion S_m, $m \leqslant n$, is mapped on a subregion of Q_n. Call the boundary of this subregion $C_{n,m}$. The region bounded by $C_{n,m}$ is a plane map of Q_m' in the z_m-plane by the function

$$Z_n = F_n(z) = F_n[F_m^{-1}(Z_m)] = \psi_{n,m}(Z_m). \tag{78}$$

We have

$$\psi'_{n,m}(0) = \left(\frac{dZ_n}{dZ_m}\right)_0 = \frac{F_n'(a)}{F_m'(a)} = 1. \tag{79}$$

We now apply Theorem 7, Corollary, to the map of Q_m'. When Z_m lies on the circumference of Q_m', the corresponding boundary point on $C_{n,m}$ in the Z-plane satisfies the inequalities

$$\frac{r}{(1+r)^2 f_n'(a)} \leqslant |Z_n| = |\psi_{n,m}(Z_m)| \leqslant \frac{r}{(1-r)^2 f_n'(a)}. \tag{80}$$

Fixing our attention on a particular S_m in ϕ and its map in Q_m' in the Z_m plane, we wish to prove that, given $\epsilon > 0$, there exists an n such that

$$|\psi_{n+p,m}(Z_m) - \psi_{n,m}(Z_m)| < \epsilon \tag{81}$$

for all $p > 0$. We arrive at this inequality by considering a larger region S_s, $s > m$, containing S_m, and its map on Q_s' in the Z_s-plane.

Since $f_s'(a)$ approaches zero as s increases without limit, we may choose $s(>m)$ so large that

$$\frac{(1+r)^2 f_s'(a)}{r} < \frac{\epsilon}{2}\left[\frac{(1-r)^2 f_m'(a)}{r}\right]^2. \tag{82}$$

We have from this inequality and from (80), for any n greater than s,

$$\left|\frac{1}{\psi_{n,s}(Z_s)}\right| \leqslant \frac{(1+r)^2 f_s'(a)}{r} < \frac{\epsilon}{2}\left[\frac{(1-r)^2 f_m'(a)}{r}\right]^2. \tag{83}$$

Hence, on Q_s'

$$\left|\frac{1}{\psi_{n+p,s}(Z_s)} - \frac{1}{\psi_{n,s}(Z_s)}\right| < \epsilon\left[\frac{(1-r)^2 f_n'(a)}{r}\right]^2. \tag{84}$$

Now, the quantity in absolute value signs in (84) is analytic throughout Q_s'. Each of the two fractions has a pole at the origin, but is otherwise analytic. In the neighborhood of the origin, we have, since $\psi_{n,s}'(0) = 1$,

$$\frac{1}{\psi_{n,s}(Z_s)} = \frac{1}{Z_s + C_n Z_s^2 + \cdots} = \frac{1}{Z_s} - C_n + \cdots$$

and

$$\frac{1}{\psi_{n+p,s}(Z_s)} - \frac{1}{\psi_{n,s}(Z_s)} = \frac{1}{Z_s} - C_{n+p} + \cdots - \left(\frac{1}{Z_s} - C_n + \cdots\right) = \varphi(Z_s),$$

where $\varphi(Z_s)$ is analytic at the origin. The inequality (84) then holds throughout the whole interior of Q_s', since the maximum absolute value of the function is on the boundary.

Now Q_s' encloses $Q'_{s,m}$, the map of Q_m'. Changing the variable to Z_m, we have on Q_m'

$$\left|\frac{1}{\psi_{n+p,m}(Z_m)} - \frac{1}{\psi_{n,m}(Z_m)}\right| < \epsilon\left[\frac{(1-r)^2 f_m'(a)}{r}\right]^2.$$

Employing the second inequality of (80), we have

$$|\psi_{n+p,m}(Z_m) - \psi_{n,m}(Z_m)| < \epsilon\left[\frac{(1-r)^2 f_m'(a)}{r}\right]^2 |\psi_{n+p,m}(Z_m)\psi_{n,m}(Z_m)| \leqq \epsilon$$

on Q_m'.

Since the expression in the first member is analytic throughout Q_m', this inequality holds throughout Q_m', which was to be proved.

It follows, from the preceding, that $\psi_{n,m}(Z_m)$ approaches a limit function $\psi_m(Z_m)$ which is analytic in Q_m'. The corresponding function of z, namely, $\psi_m(Z_m) \equiv F(z)$, the limit of (77), is analytic in S_m. Since S_m may be taken so large as to enclose any given point of ϕ, it follows that (77) converges in the whole of ϕ and that the limit function is analytic.

We now prove that $Z' = F(z)$ maps ϕ on the whole plane exclusive of the point at infinity. Let α be any finite point; and let m be chosen so that $r/[(1 + r)^2 f'_m(a)] > |\alpha|$. Then on Q_m', from (80),

$$|\psi_m(Z_m)| = \lim_{n=\infty} |\psi_{n,m}(Z_m)| > |\alpha|.$$

Then, $\psi_m(Z_m) - \alpha$ does not vanish on Q_m'; hence, from Theorem 12, this function and $\psi_{n,m}(Z_m) - \alpha$, for n sufficiently large, have the same number of zeros in Q_m'. The latter function has one zero in Q_m', since $Z = \psi_{n,m}(Z_m)$ maps Q_m' on a plane region to which α belongs. It follows that $\psi_m(Z_m)$ takes on the value α once in Q_m'; that is, $F(z)$ takes on the value α once in S_m.

We prove, as in the previous case, that α is not taken on at two different points of ϕ. For, S_m can be so chosen that it contains both points; and we have a contradiction with what we have just proved. Hence, α is taken on only once. This completes the proof of the theorem.

The limit surfaces ϕ fall into the two classes, those which can be mapped on a circle and those which can be mapped on the finite plane. It is conceivable that a surface can belong to both classes; that by forming the approximating regions ϕ_1, ϕ_2, . . . in certain ways we should get one case, and in other ways the other. But this is not possible. For, if ϕ can be mapped both on Q_0 and on the finite plane, then the finite plane can be mapped conformally on Q_0; and we found at the beginning of Sec. 78 that this is impossible.

As a convenient test for distinguishing between the two cases, the following theorem is useful:

THEOREM 22.—*If the limit region ϕ of Theorem 21 is bounded in part by a plane piece of curve, then ϕ is mapped on a circle.*

If we assume the contrary then, making a linear transformation, there exists a function $z' = f(z)$ mapping ϕ on the whole z'-plane exclusive of the origin. We draw a Jordan curve in the form of a cross-cut q cutting off from ϕ a plane finite simply connected region S whose boundary consists of q and a part h of the given boundary of ϕ. We construct S so that the point of ϕ which is carried to infinity in the mapping is exterior to S.

The map of S in the z'-plane is a finite simply connected region S' bounded by a closed Jordan curve. This curve is the map of q. As z traces q, approaching either end, the corresponding point z' approaches the origin. If we map S' conformally on the unit circle Q_0, the boundary points of S' and Q_0 correspond in a one-to-one manner (Theorem 15). The succession of these two mappings is a conformal map of S on Q_0.

Now let L_1, L_2 be curves drawn in S to two different accessible points of h. The corresponding curves L_1', L_2' in S' both terminate at the origin. The corresponding curves L_1'', L_2'' in Q_0

both terminate at the same point on the boundary of Q_0. This contradiction to Proposition 3 of Sec. 80 establishes the theorem.

86. The Mapping of Simply Connected Finite-Sheeted Regions. Before studying the general finite-sheeted region, we consider regions formed by putting square elements together. By a *square element* we shall mean (*a*) the plane interior or exterior of a square together with its boundary; or (*b*) the region formed by n superimposed equal squares hanging together at a single interior branch point of order n.

The square element is simply connected and can be mapped on a circle. The mapping of (*a*) follows directly from Theorem 13. In (*b*), let z_0 be the branch point. Then $z - z_0 = t^n$ maps (*b*) on a plane simply connected region in the t-plane, which can be mapped on a circle.

Suppose, now, that we put square elements together—like the pieces of a patch-work quilt—to form finite-sheeted regions. We shall establish the following proposition:

If a finite number of square elements be adjoined to form a simply connected region with a boundary consisting of more than one point and with no branch points other than those belonging to the individual square elements, the resulting region can be mapped comformally on a circle.

The proof is by induction. We assume that all regions of the kind mentioned formed by adjoining n square elements can be mapped on a circle, and prove that any region formed with $n + 1$ elements can be mapped on a circle. Then, since any region of one square element can be mapped on a circle, the proposition holds.

A region S_{n+1} of the kind mentioned in the proposition formed with $n + 1$ square elements can be constructed by adjoining to a like region S_n a square element S, where a part of the boundary of S belongs to the boundary of S_{n+1}. S and, by hypothesis, S_n can be mapped on circles. We shall apply Theorem 20. Now, S and S_n abut along a common bounding arc mn but there is no common region. But, since the points of mn are not branch points, S can be slightly enlarged to form a square element S' which has with S_n a common simply connected region and where the boundaries of S_n and S' meet as required in the theorem. Then $S_n + S'$ can be mapped on a circle. The subregion S_{n+1} is mapped on a plane simply connected region which can in turn be mapped on a circle, which was to be proved.

Theorem 23.—*A simply connected region with a finite number of sheets and of branch points and whose boundary consists of more than one point can be mapped conformally on a circle.*

We shall show that a given region of the kind stated in the theorem can be got as the limit of a sequence of regions formed from square elements. We first transform the region so that all boundary and branch points lie within the unit circle Q_0. This can be done as follows: Let P be an inner point of the region whose coordinate z_0 is not the z-coordinate of any boundary point or branch point. (Such a point clearly exists, since the number of branch points is finite and the boundary points in any sheet do not fill any area.) We can draw a circle C with z_0 as center such that all boundary and branch points are exterior to C. Then if we make a linear transformation carrying z_0 to ∞ and C into the circle Q_0, we have the branch points and boundary points within Q_0.

We can suppose further that the z-coordinate of each branch point, $z = x + iy$, is such that x and y are both irrational. For, the points whose x-coordinates or whose y-coordinates differ from those of a branch point by a rational quantity lie on a denumerable set of straight lines parallel to the x- and y-axes. Let z_1 be a point in the neighborhood of the origin and not lying on any of these lines. Then, a translation which carries z_1 to the origin carries the branch points into points whose x- and y-coordinates are both irrational. We take z_1 near enough to the origin that the branch points and boundary points lie in Q_0 after the translation.

We now proceed to cut the region, which we shall call ϕ, into square elements by lines parallel to the x- and y-axes. Each cut will be made through all the superimposed sheets. We first cut along the sides of the square K whose vertices are $1 + i$, $1 - i$, $-1 + i$, $-1 - i$. K contains Q_0. The exterior of K in each sheet is a square element lying entirely within or entirely without ϕ. The former we shall call the elements of the first set. By an *element* we mean here and subsequently a *square element all of whose interior and boundary points are interior points of* ϕ. There is at least one element in the first set; namely, that into which the point P was carried.

We next cut up the finite sheeted region within K. Let ν be an integer sufficiently large that the diagonal of a square of side $1/\nu$ is less than the shortest distance between the points of the z-plane at which branch points occur. We cut the square along sides parallel to the y-axis through the points $-1 + \frac{1}{\nu}$, $-1 + \frac{2}{\nu}$, $\cdots$ and along lines parallel to the x-axis through $-i + \frac{i}{\nu}$, $-i + \frac{2i}{\nu}$, $\cdots$. None of the pieces into which the region is cut contains more than one branch point. The square elements resulting from this cutting lying in ϕ are the elements of the second set.

We next cut up what remains of the region. We cut each previous square into quarters along lines joining the midpoints of opposite sides. The elements that result constitute the third set. Similarly, we divide the latter squares into quarters, cut up what remains of the region, and take out the elements of the fourth set; and so on *ad infinitum.*

We now adjoin the elements to form regions. Let ϕ_1 be an element of the first set. Let ϕ_2 consist of ϕ_1 together with all elements of the first and second sets that can be joined to ϕ_1 to form a connected region. In general, let ϕ_n consist of ϕ_{n-1} together with all elements of the first n sets that can be joined to ϕ_{n-1} to form a connected region. In the process of adjunction all cuts separating adjacent squares of ϕ_n are closed up. As the boundary of ϕ_n consists of interior points there are elements adjoining ϕ_n along its boundary. The process is never brought to an end; and we get an infinite number of regions.

The region ϕ_n is made up of a finite number of square elements. All branch points of the region are interior points of elements. For, owing to the manner of making the cuts, a point on the boundary of an element has either a rational x or a rational y, whereas at a branch point both x and y are irrational.

Further, ϕ_n is simply connected. Suppose not; then ϕ_n has two or more bounding curves. Let C_1, C_2 be two of these curves. We can draw a curve C' in ϕ_n separating C_1 and C_2. C' divides ϕ into two parts, one of which contains no boundary points of ϕ. Let this part contain C_1, say. Then C_1 bounds a subregion S_1 of ϕ abutting on ϕ_n, and S_1 has no further boundary. In our process of cutting into elements S_1 is entirely divided into elements when or before we arrive at the n-th set; so, S_1 belongs to ϕ_n, which is a contradiction.

It follows from the proposition at the beginning of this section that ϕ_n can be mapped conformally on a circle.

We next show that for n sufficiently large ϕ_n contains any preassigned point P_1 of ϕ. Let P_2 be a point of ϕ_1. P_1 and P_2 can be joined by a curve L lying entirely within ϕ. There exists a $d > 0$ such that no point of L can be joined to a boundary point of ϕ by a line lying in ϕ and of length less than d. If n be taken so large that the diagonal of the square elements of the n-th set is less than d, L is completely covered by elements of the first n sets. All these elements belong to ϕ_n, whence P_1 is an inner point of ϕ_n.

In the sequence of regions $\phi_1, \phi_2, \ldots$ certain successive regions may be identical. Thus, there may be no elements of the n-th set that can be added to ϕ_{n-1}, so that ϕ_n and ϕ_{n-1} are identical. But, after a finite number of steps, we arrive at elements which can be added to ϕ_{n-1}. We now rewrite the sequence in the same order, but including only one of a succession of identical regions. The resulting sequence,

$$\phi_1, \phi_{n_2}, \phi_{n_3}, \ldots$$

satisfies all the conditions of Theorem **21**. Hence, the limit region, ϕ itself, can be mapped either on a circle or on the whole plane exclusive of a single point.

Suppose the latter case to hold. Then, making a linear transformation of the plane on which we map, ϕ can be mapped on the whole t-plane exclusive of $t = 0$. Let $z = \varphi(t)$ be the mapping function. $\varphi(t)$ is analytic in the t-plane except at the origin and at the points corresponding to $z = \infty$ in ϕ. Let m be large enough that ϕ_m contains all the elements of the first set, and let S_m be the map of ϕ_m in the t-plane. Points outside S_m are mapped on the interior of the square K. Outside S_m then $|z| = |\varphi(t)| < \sqrt{2}$. This inequality holds in the neighborhood of the origin; so, the function is analytic at the origin if properly defined there. The function $z = \varphi(t)$ maps the neighborhood of the origin on a plane sheet or on a finite number of sheets containing a branch point according as $\varphi'(0) \neq 0$ or $\varphi'(0) = 0$. In either case, $t = 0$ goes into a single point; and ϕ has a single boundary point. As this is contrary to the hypothesis that ϕ has more than one boundary point, it follows that ϕ is mapped on a circle.

87. Conformal Mapping and Groups of Linear Transformations. We close this chapter with examples of the connection between conformal mapping and groups of linear transformations.

The ideas here introduced are fundamental in the developments of the following chapter.

Let the simply connected plane strip in Fig. 66 be formed by the adjunction of regions Σ_n each of which is carried into an abutting region by the translation $Z' = Z \pm h$. Let the strip be mapped on the unit circle Q_0 in the z-plane (Fig. 67), S_n being the map of Σ_n.

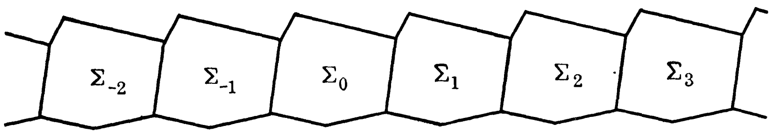


FIG. 66.

The strip—call it ϕ—has a group of conformal transformations into itself; namely, the group $Z_n = Z + nh$. What can be said of the relation between the corresponding points z and z_n in Q_0? If we carry z to Z, Z to Z_n, and Z_n to z_n, we have a sequence of three conformal transformations. Hence, the correspondence between z and z_n sets up a conformal transformation; whence, $z_n = T_n(z)$, where $T_n(z)$ is analytic. Further, in this sequence of mappings Q_0 is mapped on ϕ, ϕ on itself, and ϕ on Q_0; that

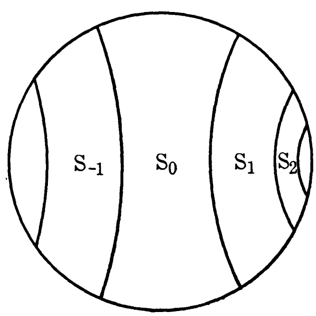


FIG. 67.

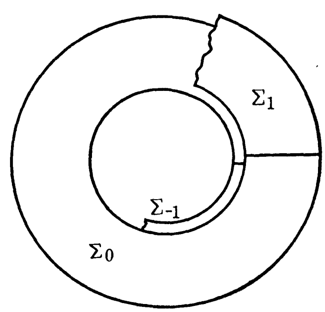


FIG. 68.

is, $z_n = T_n(z)$ maps Q_0 conformally on itself. It follows from Theorem 24, Sec. 12, that T_n is a linear transformation.

We find readily that the transformations T_n satisfy both group properties. Also, the group is properly discontinuous, since each transformation, other than the identical transformation, carries S_0 outside itself. In fact, the group is the cyclic group generated by T_1. A fundamental region for the group is S_0 together with the inverse of S_0 in Q_0.

We now get a similar group in a different way. Consider the ring-shaped region lying between two curves in the Z-plane. Let the region be rendered simply connected by means of a cut connecting the boundaries. Call the severed region Σ_0 (Fig. 68). Now, let us provide ourselves with an infinite number of copies of Σ_0 and use them to build up an infinitely sheeted region. Let one copy Σ_1 be exactly superposed on Σ_0 and joined, as in the figure, along one edge of the cut. Let another copy Σ_{-1} be placed below Σ_0 and joined along the other edge of the cut. Similarly, a copy Σ_2 is superposed and joined along the remaining free cut in Σ_1; Σ_{-2} is joined along the free cut in Σ_{-1}; and so on.

Now, the region $\phi_n = \Sigma_{-n} + \cdot \cdot \cdot + \Sigma_0 + \cdot \cdot \cdot + \Sigma_n$ is simply connected and can be mapped on a circle (Theorem 23). Hence, by Theorem 21, the limit region ϕ can be mapped on a circle or the whole finite plane. In this case (Theorem 22) the mapping is on a circle. Let $Z = f(z)$ map the surface spread over the Z-plane on Q_0 in the z-plane.

The region ϕ admits a group of conformal transformations into itself. If Σ_0 be placed accurately on Σ_n, Σ_1 will coincide with Σ_{n+1}, Σ_{-1} with Σ_{n-1}, and so on; and ϕ will be mapped conformally on itself. The transformations of ϕ for $n = 0, \pm 1, \pm 2, \ldots$ clearly form a group. When ϕ is transformed in this way the corresponding points of Q_0 undergo a transformation $z_n = T_n(z)$. By reasoning identical with that used in the former example, we conclude that T_n is a linear transformation which carries Q_0 into itself. The transformations T_n form a group isomorphic with respect to the group of transformations of ϕ into itself.

Consider the mapping function $Z = f(z)$. Let P be a point of ϕ and let P_n be the point into which it is carried when Σ_0 is carried into Σ_n. The corresponding points z and z_n satisfy the relation $z_n = T_n(z)$. The coordinates of P and P_n are $Z = f(z)$ and $Z_n = f(z_n)$, respectively. But, when Σ_0 is carried into Σ_n, each point of ϕ is superposed on its original position. Hence, the Z-coordinate is unchanged and

$$f(z_n) \equiv f(z).$$

That is, the mapping function is unaltered when a transformation of the group T_n is made.

We can readily generalize the method of Fig. 68 and get much more complicated Fuchsian groups. We can take an initial plane region with n boundaries and cut it into a simply connected region Σ_0. We superpose copies of Σ_0 joining along the edges of the cuts, to form an infinitely sheeted region which can be mapped on a circle and which admits a group of conformal transformations into itself. The details of this generalization offer no difficulty.

CHAPTER IX

UNIFORMIZATION. ELEMENTARY AND FUCHSIAN FUNCTIONS

88. The Concept of Uniformization.—As early in his study as analytical geometry and the calculus, the student of mathematics becomes acquainted with the advantages of the parametric representation of curves. Thus, he finds that the circle

$$W^2 + Z^2 = 1 \tag{1}$$

can be represented in the parametric form

$$Z = \sin z, \quad W = \cos z. \tag{2}$$

Another parametric representation of the same curve is

$$Z = \frac{2z}{1 + z^2}, \quad W = \frac{1 - z^2}{1 + z^2}. \tag{3}$$

In plotting the curve, or in evaluating integrals involving W and Z; that is, containing Z and $\sqrt{1 - Z^2}$, he finds that certain simplifications result from the use of the parametric equations.

Looked at from the standpoint of the theory of functions, the relation (1) defines W as a two-valued function of Z. In (2), and also in (3), both W and Z are expressed in terms of single-valued, or *uniform*, functions. We say that the functions (2), or (3), *uniformize* the function defined by (1).

DEFINITION.—*Let*

$$W = F(Z) \tag{4}$$

be a multiple-valued function of Z; *and let*

$$W = W(z), \quad Z = Z(z) \tag{5}$$

be non-constant single-valued functions of z *such that*

$$W(z) \equiv F(Z(z)); \tag{6}$$

then, the functions (5) *are said to uniformize the function* (4).

The variable z *is called the uniformizing variable.*

If there is one pair of uniformizing functions there are clearly infinitely many others. We have but to put $z = f(t)$, where $f(t)$ is a single-valued function of t, in (5), and we have, if the functions exist, a pair of uniformizing functions with t as the uniformizing variable.

We shall be interested, primarily, in uniformization by means of automorphic functions. In the example given above, we note that the functions (2) are automorphic functions of z.

We have found (Sec. 43) that between two simple automorphic functions $f_1(z)$, $f_2(z)$ belonging to the same group and having the same domain of definition, there exists an algebraic relation $G(f_1, f_2) \equiv 0$. If W is an algebraic function of Z defined by the relation $G(W, Z) = 0$, then the functions $W = f_1(z)$, $Z = f_2(z)$ uniformize the algebraic function. The converse theorem that any algebraic function can be so uniformized will be proved in the present chapter.

Before taking up the study of this problem we turn to a brief discussion of Riemann surfaces and their connectivity.

89. The Connectivity of Regions.—The general algebraic function is defined implicitly by the equation

$$G(W, Z) = P_0(Z)W^m + P_1(Z)W^{m-1} + \cdots + P_m(Z) = 0, \quad (7)$$

where $P_0(Z), \ldots, P_m(Z)$ are polynomials, and $G(W, Z)$ is irreducible. Except for certain isolated values of Z, to each value of Z correspond m distinct values of W. The Riemann surface for W as a function of Z is a closed m-sheeted two-sided surface with a finite number of branch points. The uniformization of the function depends upon the way the Riemann surface is joined together, and brings up the question of the connectivity of the surface.

We consider regions lying on such a surface, and we suppose each region considered to be bounded by curves—Jordan curves, say. As a special case of a bounding curve, we include the point. Thus, the region consisting of the interior of the circle Q_0 exclusive of the origin has two boundaries, Q_0 and the origin.

We shall consider the question of altering a region and of separating it into two or more regions by cutting it along curves. Except possibly at its ends, a cut shall pass only through interior points of the region and shall not pass twice through the same point. As a cut is drawn, it severs the surface; so that two sides or *banks* of the cut belong to the boundary of the new regions or region. These banks are considered a part of the boundary in making subsequent cuts, or in continuing the same cut.

A *cross-cut* is a cut beginning and ending in the boundary. Thus in Fig. 69, where the region is bounded by the rectangle and the circle, q_1 and q_2 are cross-cuts. A *sigma cross-cut* (q_3) is a special kind of cross-cut which begins in the boundary and ends in one of the banks of the cut itself. A *loop-cut* (q_4) begins at an interior point and ends at the initial point. A sigma cross-cut is equivalent to a loop-cut followed by a cross-cut.

A finite-sheeted region with a boundary is called "simply connected" if every cross-cut separates it into two pieces. The elementary properties of such a region—for example, that the two pieces into which a cross-cut separates the region are themselves simply connected; that a loop-cut separates the region into two pieces, one of which is simply connected; that the boundary of the region consists of a single connected set of points—have been considered as being probably sufficiently well known to the reader to be used without proof in the preceding chapter, and they will be taken for granted here. We shall devote our limited space to the derivation of certain propositions which are necessary for the study of regions which are not simply connected, and which are less likely to be already known to the reader.

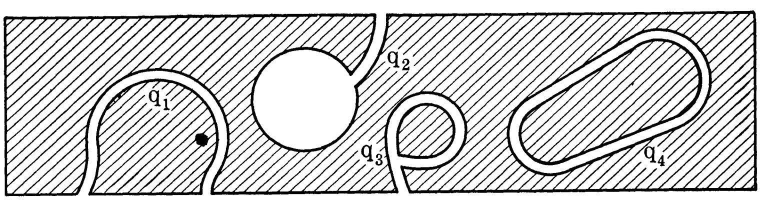


FIG. 69.

We shall deal only with regions with boundaries. A boundaryless region will be given a boundary by marking a point or by cutting out from the surface a small circle at some point.

A region which can be rendered simply connected by one cross-cut is called "doubly connected." In general, a region which can be rendered simply connected by n cross-cuts is $(n + 1)$-ply connected or of connectivity $N = n + 1$. That two different systems of cuts which render the region simply connected contain the same number of cross-cuts is a consequence of the following theorem:

THEOREM 1.—*If a region or set of regions is cut by a system of ν_1 cross-cuts into α_1 simply connected pieces and by a system of ν_2 cross-cuts into α_2 simply connected pieces, then*

$$\nu_1 - \alpha_1 = \nu_2 - \alpha_2. \tag{8}$$

We suppose the cuts slightly deformed, if necessary (this does not affect the number of cuts or of pieces), so that the beginning or end of no cut in one system lies on a cut of the other system, and so that the cuts of the two systems have a finite number s of points of intersection. We now make both systems of cuts and count the number of pieces.

Having made the first system of cuts and got α_1 pieces, we make the second system of cuts. When, in making a cut of the second system, we meet a cut of the first system, the cross-cut ends and a new cross-cut begins; so that each point of intersection increases the number of cross-cuts by one. The second system of cuts then is a set of $\nu_2 + s$ cross-cuts in the regions made by the first system of cuts. Each is a cross-cut in a simply connected region and increases the number of simply connected regions by one. We have, then, $\alpha_1 + \nu_2 + s$ simply connected regions as a result of the two systems of cuts. Making the second system of cuts first, we have, similarly, $\alpha_2 + \nu_1 + s$ simply connected pieces. Equating the two expressions for the number of pieces,

$$\alpha_2 + \nu_1 + s = \alpha_1 + \nu_2 + s,$$

from which (8) follows.

If $\alpha_1 = \alpha_2$ in (8), then $\nu_1 = \nu_2$; that is, if the number of simply connected pieces is the same for two systems, then the number of cross-cuts is the same. In particular, the number of cross-cuts which yield a single simply connected piece is independent of the way the cuts are made.

THEOREM 2.—*If a region is divided by ν cross-cuts into α simply connected pieces, the connectivity of the region is*

$$N = \nu - \alpha + 2. \tag{9}$$

If n suitably made cross-cuts result in a single simply connected piece, we have, from (8),

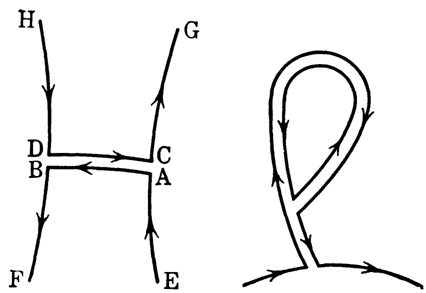
$$\nu - \alpha = n - 1,$$

whence,

$$N = n + 1 = \nu - \alpha + 2. \tag{10}$$

It remains to show that, under the circumstances of the theorem, the region is of finite connectivity; that is, that a finite number of cross-cuts can be made which result in a single simply connected piece. If the region is not simply connected, we can make a cross-cut, by hypothesis, which does not cut it in two. If the resulting region is not simply connected, we can make a second cross-cut not cutting the surface in two; and so on. Let m such cuts be made without cutting the surface in two; and let these cuts have s intersections with the ν cuts of the theorem (the cuts being deformed slightly, if necessary, as in the proof of the preceding theorem). These cuts superimposed on the preceding ν cuts amount to $m + s$ cross-cuts in the set of α pieces; hence we have $\alpha + m + s$ pieces. Now make the m cuts first and then the $\nu + s$ cross-cuts. We start with one piece and each of these $\nu + s$ cuts adds one new piece at most; and we get $\nu + s + 1$ pieces at most. Hence,

$$\alpha + m + s \leqslant \nu + s + 1, \qquad m \leqslant \nu - \alpha + 1.$$



(a) (b)

FIG 70

It follows from this inequality that, in applying the preceding scheme, we cut the region into a simply connected piece in a finite number of steps. The number of necessary cross-cuts is, then, from (10), $n = \nu - \alpha + 1$.

THEOREM 3.—*If m cross-cuts are made in a region of connectivity N without cutting it in two, the resulting region is of connectivity $N - m$.*

$N - 1$ suitable cross-cuts make the original region simply connected. We may use for these the given m cuts together with $N - 1 - m$ other cross-cuts suitably made. These latter cuts render the new region simply connected; hence, its connectivity is $(N - 1 - m) + 1$, or $N - m$.

THEOREM 4.—*A region of finite connectivity is of odd connectivity if it has an odd number of boundaries and of even connectivity if it has an even number of boundaries.*

We show first that a cross-cut increases or decreases the number of boundaries by one. Consider, first, the ordinary cross-cut (Fig. 70 (a)). If the cross-cut joins points on two different boundaries, $CG \ldots EA$ and $BF \ldots HD$ (each traced in a positive sense, that is, with the region on the left), then we have, after the cut, the single boundary $CG \ldots EABF \ldots HDC$; and the number of boundaries is decreased by one. If the cross-cut joins points on the same boundary $CG \ldots HDBF \ldots EA$, then we have, after the cut, two boundaries, $CG \ldots HDC$ and $BF \ldots EAB$; and the number of boundaries is increased by one.

In the case of a sigma cross-cut (Fig. 70 (*b*)), the inner bank of the loop forms a new boundary, whereas the remaining banks of the cut combine with the boundary from which the cut issued to form a single boundary; hence, the number of boundaries is increased by one.

Let the region have k boundaries and let n cross-cuts make it simply connected. The i-th cross-cut increases the number of boundaries by ϵ_i, where $\epsilon_i = \pm 1$. We have as a result of the n cuts $k + \epsilon_i + \cdots + \epsilon_n$ boundaries. But, at the end, there is a simply connected piece with a single boundary; so the connectivity N is

$$N = n + 1 = n + k + \epsilon_1 + \cdots + \epsilon_n,$$

or

$$N = (1 + \epsilon_1) + (1 + \epsilon_2) + \cdots (1 + \epsilon_n) + k.$$

Each term in the second member, excepting the last, is even, being either zero or two; hence N and k are both even or both odd, which was to be proved.

THEOREM 5.—*The Riemann surface of an algebraic function is of finite and odd connectivity.*

The surface is a closed region with a finite number of sheets and of branch points. In determining the connectivity, we first mark a point or cut a small hole in the surface to give it an initial boundary. According to Theorem 4, the connectivity is odd, provided it is finite. We have merely to show that the surface can be divided into simply connected pieces by a finite number of cross-cuts.

It may be remarked that in cutting a region into simply connected pieces, we may put in a finite number of loop-cuts without counting them among the cuts. For, in the process of cutting up the region, the loop-cut is subsequently joined to a boundary point. Now the loop-cut and the cross-cut joining it to the boundary are equivalent to a single sigma cross-cut; so we need only count the cross-cut.

Let the surface have m sheets and s branch points. For convenience, let the surface be slightly deformed, if necessary, so that no two branch points are superposed; that is, no two have the same z coordinate. Let $a_1, a_2, \ldots, a_s$ be the coordinates of the branch points and let a be any other point. We now draw circles $C_1, \ldots, C_s, C$ about $a_1, \ldots, a_s, a$, sufficiently small that they are exterior to one another; and we make cuts along these circles through all m sheets (Fig. 71, where the two-sheeted surface with four branch points is cut up). One of the circles at a is used as an initial boundary, and the piece which it encloses is discarded. The remaining cuts are all loop-cuts, and they enclose simply connected pieces. We now join C to each of the

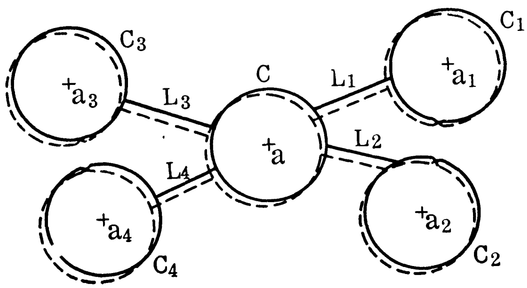


FIG. 71.

circles $C_1, \ldots, C_s$ by lines $L_1, \ldots, L_s$ which nowhere intersect, and we cut through the m sheets along these lines. These cross-cuts divide the m-sheeted region exterior to the preceding loop-cuts into m single-sheeted regions, each of which has a single boundary and is simply connected. The surface is cut into simply connected pieces by this finite set of cuts; hence, its connectivity is finite.

We can get a formula for the connectivity. We made s cross-cuts in each sheet; in all $\nu = ms$ cuts. At the branch point a_i let r_i sheets hang together. The neighborhood of the branch point is simply connected; in addition, there are $m - r_i$ plane circular pieces in C_i. Hence, C_i contains $m - r_i + 1$ simply connected pieces. There are $m - 1$ plane circular pieces in C. Finally, there are m pieces exterior to the circles. All together, then, the number of pieces is

$$\alpha = \sum_{i=1}^{s}(m - r_i + 1) + m - 1 + m.$$

The connectivity, from (9), is

$$N = ms - [\Sigma(m - r_i + 1) + 2m - 1] + 2,$$

or

$$N = \Sigma(r_i - 1) - 2m + 3. \tag{11}$$

The quantity $r_i - 1$ is called the "order of the branch point." Since N and $2m - 3$ are odd numbers, we see that the sum of the orders of the branch points is necessarily even.

As an example of the formula, consider the two-sheeted surface of the function

$$W^2 = A(Z - a_1)(Z - a_2) \cdot \cdot \cdot (Z - a_n), \tag{12}$$

where the constants a_i are distinct. The points a_i are branch points, and in addition ∞ is a branch point if n is odd. Each branch point is of the first order, whence $\Sigma(r_i - 1)$ is equal to the number of branch points, n or $n + 1$. Since $m = 2$, we have, from (11),

$$N = n - 1 \text{ or } N = n$$

according as n is even or odd. If $n = 1$ or $n = 2$, the surface is simply connected. If $n = 3$ or $n = 4$, the connectivity is 3; the surface is called *elliptic*. If $n > 4$, the surface is called *hyperelliptic*.

The Genus of the Algebraic Surface.—Many properties of algebraic surfaces and of the functions to which the surfaces belong are dependent upon the connectivity. These properties are usually stated in terms of the genus, which is defined as follows:

DEFINITION.—*Let $2p + 1$ be the connectivity of the surface; then p is called the genus of the surface.*

We recall that, since the connectivity is odd, it can always be written in the form $2p + 1$, where p is an integer.

There exist surfaces for which p is equal to any given positive integer or zero. If $n = 2p + 1$ or $n = 2p + 2$ in (12), the connectivity is $2p + 1$ and the genus is p.

By the genus of an algebraic function is meant the genus of its Riemann surface.

If $p = 0$, the surface is simply connected. If $p > 0$, it is multiply connected. It requires $2p$ cross-cuts to render the surface simply connected.

From (11), we have the following formula for p:

$$p = \tfrac{1}{2}\Sigma(r_i - 1) - m + 1. \tag{13}$$

If ν cross-cuts result in α simply connected pieces, we have, from (10),

$$p = \tfrac{1}{2}(\nu - \alpha + 1). \tag{14}$$

Euler's Formula.—The formula connecting F, V, and E, the number of faces, vertices, and edges, respectively, of a solid, can now be derived. We first make a loop-cut, around each vertex. After discarding one piece to supply the initial boundary, we have $V - 1$ pieces cut out. What remains is cut into F simply connected pieces by cross-cuts, E in number, along the edges. Formula (14) then gives

$$p = \tfrac{1}{2}[E - (V - 1 + F) + 1];$$

whence,

$$V + F = E + 2(1 - p).$$

This is a generalization of Euler's formula for the case of a solid bounded by a simply connected surface ($p = 0$):

$$V + F = E + 2.$$

On Severing the Surface.—We now describe a method of severing the algebraic surface such that at each stage of the process there are never more than two boundaries. We noted in the proof of Theorem 4 that a cut joining points of the same boundary, also a sigma cross-cut, increases the number of boundaries by one; while a cut joining points of different boundaries decreases it by one. By alternating the cuts, we get alternately one and two boundaries. We note that, in the former case, the opposite banks of the loop of the sigma cross-cut belong to different boundaries.

If a cross-cut joins different boundaries and does not cut the surface into two pieces we can deform the cut—by the simple process of moving its ends along the boundaries—so that the cut joins any given point of one boundary to any given point of the other boundary and the surface is still a single piece. Similarly, we can deform any cut joining two points of the same boundary into a cut joining two prescribed points of that boundary. Or we can deform it into a sigma cross-cut springing from a prescribed point of the boundary; to do this, we move one end of the cut along the boundary to the prescribed point P, then move the other end along the boundary to P, and, thence, along a bank of the cut. Likewise, we can deform a sigma cross-cut into an ordinary cross-cut by reversing the process.

Let P, an ordinary point of the surface, be selected as the initial boundary. If the surface is not simply connected, we can make a cross-cut a_1 beginning and ending at P which does not cut the surface into two pieces. Let b_1 be any cut joining points on opposite banks of a_1. The cut b_1 does not cut the surface in two, the reason being that we now have a single boundary, whereas two pieces would have two boundaries.

We can suppose b_1 to be so drawn that it joins points on opposite banks of a_1 at P.

If $p = 1$, the surface is now simply connected. If not, we can draw a cross-cut a_2 which does not separate the surface. This cut can be deformed so that it begins and ends at a single one of the boundary points at P. A cut b_2 joining points on opposite banks of a_2 does not separate the surface; and we can suppose the ends of b_2 are moved to P. If $p = 2$, the surface is now simply connected. If not, we continue in this manner to get cuts a_3, b_3; a_4, b_4; . . . ; a_p, b_p, all beginning and ending at P which render the surface simply connected.

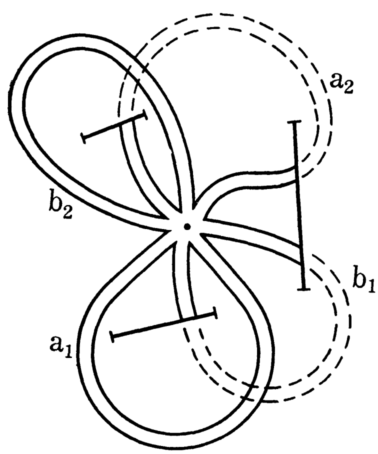


FIG. 72.

The boundary of the severed surface consists of $4p$ curves each beginning and ending at P; to wit, the two banks of each of the p a-cuts and the two banks of each of the p b-cuts.

In Fig. 72, the hyperelliptic surface with six branch points ($p = 2$) is cut into a simply connected piece. The two sheets of the surface are joined along the three rectilinear branch lines of the figure. The cuts drawn with the heavy lines are in the upper sheet, those drawn with broken lines are in the lower sheet.

90. Algebraic Functions of Genus Zero. Uniformization by Means of Rational Functions.—We first prove the following:

THEOREM 6.—*An algebraic surface of genus zero can be mapped conformally on the whole plane.*

Let P be an ordinary point of the surface, and let a circle C_n with P as center and of radius r_n be cut from the sheet in which P lies. Here, r_n is to be sufficiently small that the piece cut out is plane. Then, the part of the surface that remains—call it ϕ_n—is simply connected and can be mapped conformally on a circle (Theorem 23, Sec. 86). We form a sequence of circles with radii $r_1, r_2, \ldots$, where $r_{n+1} < r_n$ and $\lim r_n = 0$. There results a sequence of regions $\phi_1, \phi_2, \ldots$, each of which is a subregion of the following. These regions satisfy the conditions of Theorem 21, Sec. 85; hence, the limit region ϕ, which consists of the surface bounded by the single point P can be mapped conformally either on a circle or on the plane bounded by a single point.

Suppose ϕ can be mapped on the unit circle Q_0 in the z-plane; and let $z = f(Z)$ perform the mapping. At all points of ϕ we have $|f(Z)| < 1$. Since this inequality holds, in particular, in the neighborhood of P the function $f(Z)$ is analytic at P if properly defined there. The function is analytic at all ordinary points of the surface and continuous at the branch points and at infinity. Hence, it is a constant,[1] which is impossible.

Then ϕ can be mapped on the whole z-plane exclusive of a single point, which we may take to be the origin. In the neighborhood of P the mapping function $z = f(Z)$ remains finite as before, and is analytic at P if properly defined there. The plane neighborhood of P is mapped on the plane neighborhood of the origin, P corresponding to the origin. Hence, ϕ, together with its boundary P, is mapped on the whole z-plane.

THEOREM 7.—*Any algebraic function of genus zero can be uniformized by means of rational functions. Conversely, if a function is uniformized by means of rational functions, it is an algebraic function of genus zero.*

Let $z = f(Z)$ map the Riemann surface of the given algebraic function of genus zero on the whole z-plane. Consider the inverse function $Z = Z(z)$. $Z(z)$ is an analytic function of z except at the finite number of points a_i in the z-plane corresponding to the points $Z = \infty$ in the various sheets of the Riemann surface. At a_i, $Z(z)$ becomes infinite. Then $Z(z)$ is analytic except for poles and is, therefore, a rational function of z. It is a rational function of order m, where m is the number of sheets in the Riemann surface. For, an arbitrary value Z_0 is taken on m times; namely, at the points $z_1, z_2, \ldots, z_m$ corresponding to the points $P_1, P_2, \ldots, P_m$ which lie superposed in different sheets of the surface and have the coordinate Z_0.

[1] Briefly, because, if not constant, $|f(Z)|$ takes on its maximum value at some interior point, which is impossible.

Let $W = F(Z)$ be the algebraic function. On the Riemann surface $F(Z)$ is single valued and analytic (continuous at the branch points and at infinity) except at certain points where W becomes infinite. From the equation satisfied by W (Equation (7)), we see that the m values of W are finite except when $P_0(Z) = 0$ and possibly when $Z = \infty$. W becomes infinite only at a finite number of points on the surface. On changing the variable to z, $W = F(Z(z))$ is single valued in the z-plane and analytic except at the points corresponding to the points on the surface at which W becomes infinite; at these points the function has poles. $F[Z(z)]$ is then a rational function of z. The two rational functions, $Z = Z(z)$ and $W = W(z) = F[Z(z)]$, uniformize the algebraic function.

Conversely, let a function $W = H(Z)$ be uniformized by means of rational functions, $Z = R_1(z)$, $W = R_2(z)$. The relation connecting W and Z is got by eliminating z from these latter equations. On clearing of fractions we have two polynomials $G_1(Z, z) = 0$, $G_2(W, z) = 0$, the eliminant of which is a polynomial in Z and W. Or, we can apply the method of Sec. 43. Since $R_1(z)$ and $R_2(z)$ take on each value the same number of times in the z-plane, the reasoning of that section can be applied word for word to show that $R_1(z)$ and $R_2(z)$ are connected by an algebraic relation.

It remains to show that the algebraic function is of genus zero. Let ϕ be the Riemann surface of $H(Z)$. To a point z_0 of the plane corresponds a single point $P_0(Z_0, W_0)$ of the surface, although to P_0 there may correspond other values $z_1, z_2, \ldots z_n$ at which $R_1(z)$ and $R_2(z)$ have the same values as at z_0. If $R_1'(z_0) \neq 0$, the function $Z = R_1(z)$ maps the neighborhood of z_0 on the plane neighborhood of P_0. Then, to a curve on ϕ passing through P_0, there corresponds a single curve passing through z_0.

Now, suppose that ϕ is not simply connected. With P_0 as the initial boundary, we make an ordinary cross-cut a_1 not cutting the surface in two pieces. We suppose a_1 so chosen that it avoids those points on the surface which correspond to the finite number of points in the plane at which $R_1'(z) = 0$ and $z = \infty$. Now, let P start at P_0 and make a circuit of a_1. By virtue of the inverse of the function $Z = R_1(z)$, starting with the branch in which $z = z_0$, z traces a curve λ which begins at z_0 and ends either at z_0 or at one of the points $z_1, \ldots, z_n$.

If the latter is the case, let P make a second circuit around a_1 in the same sense, thus continuing the curve λ to one of the points $z_0, \ldots, z_n$; and so on. In tracing λ, we arrive, eventually, at a point on the part of the curve previously traced; for there is but a finite number of points in the plane corresponding to each point of a_1. Let z' be the first such point encountered; and let P' be the corresponding point on a_1. If z' is different from z_0, we have two different curves through z' which correspond to one curve through P' which, as we noted above, is impossible. Hence $z' = z_0$; and λ is a simple closed curve in the z-plane.

Let b_1 be a cross-cut joining opposite banks of a_1 at P_0 and avoiding the exceptional points mentioned above. Except at P_0, b_1 nowhere meets a_1. Let P trace the cut b_1 repeatedly in the same sense, starting at P_0. Then z traces a curve μ, beginning at z_0 and eventually closing at z_0. Except at z_0, μ nowhere meets λ. At z_0, owing to the conformal mapping of the neighborhood of P_0 on the neighborhood of z_0, the beginning and end of μ are on opposite banks of λ. But, it is impossible to join points on opposite banks of a closed curve in the plane by a curve which does not meet the closed curve. The hypothesis that ϕ is multiply connected is untenable. Hence, the surface is of genus zero.

We can now find all pairs of functions which uniformize the algebraic function of genus zero.

THEOREM 8.—*Let $Z = Z(z)$, $W = W(z)$ be uniformizing functions of an algebraic function of genus zero such that $Z = Z(z)$ maps the whole z-plane in a one-to-one manner on the Riemann surface of the function. Then, the most general uniformizing unctions are got by substituting $z = \varphi(t)$ in this pair of equations, where $\varphi(t)$ is a single-valued function of t.*

Suppose that $Z = Z_1(t)$, $W = W_1(t)$ uniformize the function. The two functions have the same domain of existence S, since the relation $W_1(t) \equiv F[Z_1(t)]$ supplies an analytic continuation of each function into the domain of existence of the other. The equations $Z = Z_1(t)$ and $Z = Z(z)$ define z as a function of t, $z = \varphi(t)$. To each value of t in S corresponds a single point on the Riemann surface and to this point corresponds a single value of z. Further, $\varphi(t)$ does not exist outside S; for, otherwise, $Z_1(t) = Z(z) = Z[\varphi(t)]$ can be continued analytically outside S, contrary to hypothesis. Hence, z is a single-valued function of t. It is obvious that, conversely, if we replace z by a single-valued function of t we have a pair of uniformizing functions.

As an example, we find readily that the functions (3) uniformize the function W defined by (1) in the way mentioned in the theorem. Consider the uniformizing functions $Z = \sin t$, $W = \cos t$. On eliminating Z from the first of these equations and the first equation of (3), we verify that z is a single-valued function of t.

91. Algebraic Functions of Genus Greater than Zero. Uniformization by Means of Automorphic Functions.[1]—If the genus of the surface is greater than zero, the method of the preceding section must be altered; for the uncut surface cannot be mapped on a plane region. In this case, we shall cut the surface and build up an infinitely sheeted region by a method akin to that employed in Sec. 87 (Fig. 68).

Given an algebraic function $W = F(Z)$ of genus $p > 0$. We represent its Riemann surface by F. We shall render the surface simply connected by a system of $2p$ cross-cuts a_1, b_1, . . . , a_p, b_p, beginning and ending at an ordinary point P of the surface (see Fig. 72). The severed surface, which we shall call F_0, is a simply connected region whose single boundary is composed of $4p$ arcs or sides, each of which begins and ends at P. These sides meet at $4p$ vertices—all at P—and the sum of the angles at the vertices is 2π.

We provide ourselves with an infinite number of copies of F_0 and proceed to put them together in such a way that at each step we have a simply connected region. We start with one copy of F_0 which we shall designate by ϕ_0. Let a second copy be superposed on ϕ_0 and joined to it along one of the $4p$ sides—for instance, one bank of a_1 in ϕ_0 is joined to the opposite bank of a_1 in the copy. The combined region is simply connected. Let other copies of F_0 be joined along the $4p - 1$ remaining sides of ϕ_0, a bank of a cut in ϕ_0 being joined to the opposite bank of the same cut in the copy. Call the region formed by ϕ_0 and the copies that have been adjoined along its $4p$ boundaries ϕ_1. ϕ_1 is bounded by a finite number of sides. Along each side of ϕ_1 we join a new copy of F_0, the junction being made between the opposite banks of the same cut; and let ϕ_2 represent the total region.

In general, ϕ_{n+1} is got by joining copies of F_0 along all the free sides of ϕ_n. In applying the process, we shall close up other sides as follows: Whenever in passing around the bound-

[1] P. Koebe, *Math. Ann.*, vol. 67, pp. 145–224, 1909.

ary of a combined region we encounter one of the a- or b-cuts followed by the same cut, we shall join the two banks together and reduce the vertex separating them to an interior point.

In Fig. 73 we show, for the surface of Fig. 72, how the regions fit together at P itself. We start with a particular one of the vertices of ϕ_0. The figure shows how the neighborhood of the vertex is filled out to form a plane sheet. The $4p$ vertices of ϕ_0 lie in $4p$ different plane sheets of the later regions.

By continuing indefinitely the adjunction of copies of F_0 we build up an infinitely sheeted boundaryless region ϕ.[1] Each of the regions ϕ_n is a finitely sheeted simply connected region which can be mapped conformally on a circle. The region ϕ_n

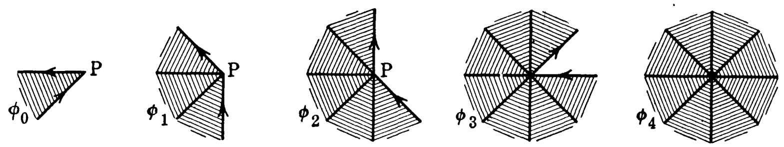


FIG 73.

is a subregion of the region ϕ_{n+1}. The sequence $\phi_0, \phi_1, \ldots$ is one to which Theorem 21, Sec. 85, applies; *hence, ϕ can be mapped conformally either on a circle or on the whole plane exclusive of a single point.*

The function $W = F(Z)$ is single valued on the surface F and, hence, on the severed surface F_0, or ϕ_0. What can be said of the analytic continuation of $F(Z)$ throughout ϕ? To each point of F_0 corresponds a single Z and a single W. When a copy of F_0 is adjoined it is so superposed that each point has the same Z as before. Let it also bear the same value of W as before. Then, the function W so defined is an analytic function (except for poles and branch points) in each of the copies of F_0.

It is a principle of analytic continuation that when two functions are analytic in abutting regions and are equal at the points

[1] The rather obvious fact that there are always free sides along which to adjoin copies of F_0 can be shown in various ways. Assume that ϕ_n has a side l joining vertices P' and P'' each of which belongs to but one of the copies of F_0 comprised in ϕ_n. Then, when copies of F_0 are added to form ϕ_{n+1} that copy which is joined along l is not joined elsewhere. This copy has $4p$ vertices of which P' and P'' belong to other copies. The remaining $4p - 2(\geqslant 2)$ vertices provide at least one side joining vertices not belonging to any other copy in ϕ_{n+1}. Since each side of ϕ_0 joins vertices belonging to but one copy, it follows by induction that ϕ_n has free sides; and the process of constructing the surface proceeds *ad infinitum*.

of the common boundary, each is the analytic continuation of the other. We have joined the copies of F_0 along curves at every point of which W has the same value in the two copies. Hence, the function W of one copy is the analytic continuation of W in the adjacent copy. The function W, or $F(Z)$, borne by the copies F_0 fit together to form a function which is single valued and analytic, except for poles and branch points, over the whole surface ϕ.

It is clear from the manner of constructing the surface that ϕ is infinitely sheeted; but the remarks just made furnish an independent proof of that fact. If a finite number of copies fit together to form a closed surface, we can map ϕ on the whole plane. By the use of the mapping function, we can uniformize $F(Z)$ by means of rational functions after the manner of the preceding section. It follows that $F(Z)$ is of genus zero, contrary to hypothesis.

The Uniformization when ϕ Can Be Mapped on a Circle.—Let $z = f(Z)$ be the function mapping ϕ on the unit circle Q_0 in the z-plane. Let $Z = Z(z)$ be the inverse function. To each point z of Q_0 there corresponds, by the mapping, a single point P of ϕ. To P are attached a unique Z and W. If z traces any closed path in Q_0, P traces a closed curve on ϕ, and the functions $Z(z)$ and $W(z) = F[Z(z)]$ return to their initial values.

It is conceivable that $Z(z)$ or $W(z)$ can be extended analytically around a closed curve not lying entirely within Q_0 such that, on completing the circuit, the function has a value different from its initial value and so is not single valued. To prove that this is impossible, it suffices to show that the circle is a natural boundary for each function; that is, that neither function can be continued analytically across the circumference. Suppose that $Z(z)$ can be continued analytically across the circumference; and let z' be a point on the circumference at which $Z(z)$ is analytic and $Z'(z) \neq 0$. Then $Z = Z(z)$ maps a sufficiently small neighborhood s of z' on a plane region S of the Z-plane. Let z be a point of s lying in Q_0, and let P be the corresponding point of ϕ. The function $z = f(Z)$ maps the region S enclosing P on s. But this is impossible for S is a subregion of ϕ, and is mapped on a region lying entirely within Q_0. Also $W(z)$ cannot be continued analytically across the circumference; otherwise, since Z is an algebraic function of W, $Z(z)$ could be continued analytically across the circumference.

The two functions

$$Z = Z(z), \quad W = W(z) = F[Z(z)] \tag{15}$$

are then single-valued functions of z. We see, from the definition of uniformizing functions, that they uniformize the function $W = F(Z)$. Z and W become infinite at a finite number of points on each copy F_0 of the severed Riemann surface. At the corresponding points in Q_0, $Z(z)$ and $W(z)$ have poles. Since the number of copies of F_0 is infinite, each of the functions has an infinite number of poles in Q_0.

We wish to show next that $Z(z)$ and $W(z)$ are automorphic functions. The surface ϕ admits an infinite set of conformal transformations into itself. If the initial region ϕ_0 is placed in coincidence with any copy F_0' of which ϕ is built up, the copies of F_0 adjacent to ϕ_0 will be carried into coincidence with those adjacent to F_0', and so on, the whole region ϕ being carried into itself. The set of transformations of ϕ into itself which we get by carrying ϕ_0 into each of the copies contained in ϕ constitute a group. For, the succession of two transformations or the inverse of any is equivalent to the carrying of ϕ_0 into a suitable one of the copies.

Let P be a point of ϕ and let P_n be the point into which it is carried by one of these rigid motions of ϕ into itself. Let z and z_n be the corresponding points of the z-plane. We derive z_n from z by the process of mapping Q_0 on ϕ, ϕ on itself, and ϕ on Q_0. The result is that Q_0 is mapped conformally on itself; hence $z_n = T_n(z)$, where $T_n(z)$ is a linear transformation (Sec. 12, Theorem 24). The set of linear transformations of Q_0 into itself form a group isomorphic with the group of transformations of ϕ.

The points P and P_n bear the same values of Z and W. Hence, $Z(z_n) = Z(z)$ and $W(z_n) = W(z)$. In other words, $Z(z)$ and $W(z)$ are automorphic with respect to the group T_n.

The map of any of the copies F_0 of which ϕ is composed is a fundamental region for the group T_n. For instance, let S_0 be the map of ϕ_0 (see Fig. 74). Each of the transformations of ϕ, other than the identical transformation, carries ϕ_0 outside itself; and there exist transformations carrying ϕ_0 into any adjacent copy. Hence, in the z-plane no two points of S_0 are congruent, whereas any region abutting on S_0 contains points congruent to points of S_0. S_0 is, thus, a fundamental region for the group. S_0 is bounded by $4p$ sides—the maps of the $4p$ sides of ϕ_0. These sides are congruent in pairs, each pair being the maps of the opposite banks of the same a- or b-cut.

The boundary of ϕ_0 consists of interior points of ϕ; hence, the boundary of S_0 lies within Q_0. If, now, we form the region R exterior to the isometric circles of the transformations T_n, the part R_0 of R which lies within Q_0 does not extend to the circumference (Sec. 34, Theorem 17). It follows, then, from Theorem 14, Sec. 34, that the group T_n is a Fuchsian group of the first kind.

When ϕ is transformed into itself by carrying ϕ_0 into some copy of F_0 other than itself, no point of ϕ remains fixed. Each inner point of a copy F_0 is carried into an inner point of some other copy. A boundary point of the copy may be carried into another boundary point—if the copy is carried into an adjacent copy—but, if so, it is a different boundary point. It follows that no transformation $T_n(\neq 1)$ has a fixed point lying within Q_0. In other words, *the group T_n contains no elliptic transformations.*

The Uniformization when ϕ Can Be Mapped on the Finite Plane.—Let $z = g(Z)$ be a function mapping ϕ on the finite plane. Let $Z = Z_1(z)$ be the inverse function. To each point z there corresponds, by the mapping, a single point P of ϕ; and P bears a single Z and W. If z traces any closed finite path, P traces a closed curve on ϕ; and $Z_1(z)$ and $W_1(z) \equiv F[Z_1(z)]$ return to their initial values. The functions

$$Z = Z_1(z), \quad W = W_1(z) \equiv F[Z_1(z)] \tag{16}$$

are single-valued functions of z which uniformize the function $W = F(Z)$. Each function is analytic in the finite plane save for an infinite number of poles.

Corresponding to a transformation of the group of transformations of ϕ into itself is a transformation $z_n = T_n(z)$ which maps the finite plane on itself. It follows that T_n is a linear transformation with infinity as fixed point. We can prove, exactly as in the preceding case, that $Z_1(z)$ and $W_1(z)$ are automorphic with respect to the group T_n.

ϕ_0 is mapped on a region S_0 which together with its boundary lies in the finite plane. S_0 is a fundamental region for the group. As before, the group contains no elliptic transformations.

We are now able to identify the group T_n. The transforms of S_0 cluster about no finite point; hence, infinity is the only limit point of the group. All the groups with infinity as the sole limit point were derived in Secs. 59 and 60. The groups of the latter section, however, contain elliptic transformations and so are ruled out. There remain the simply and the doubly periodic groups. But the simply periodic group has no fundamental region lying, together with its boundary, in the finite plane. Hence, the group T_n is a doubly periodic group.

The uniformizing functions $Z_1(z)$ and $W_1(z)$ are elliptic functions. Each is a rational function (Sec. 61) of the Weierstrassian functions $\wp(z)$ and $\wp'(z)$ connected with the doubly periodic group T_n.

92. The Genus of the Fundamental Region of a Group.—S_0 is a map of the severed surface ϕ_0. If we bring the congruent edges of S_0 together (Fig. 74) to form a closed surface, we have

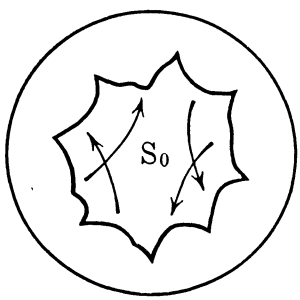


FIG. 74.

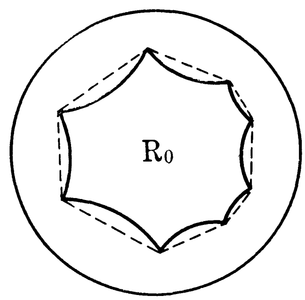


FIG. 75.

a surface whose points correspond in a one-to-one manner to the points of the uncut Riemann surface F. The two surfaces have the same genus, since the connectivity is invariant under continuous deformations; if a set of cuts renders one surface simply connected, the corresponding cuts render the other simply connected.

DEFINITION.—*By the genus of the fundamental region of a group is meant the genus of the closed region formed by bringing congruent sides of the region together so that congruent points coincide.*

We shall derive a formula for the genus in the case in which the fundamental region, before being closed, consists of a single simply connected piece. Let $2n$ be the number of sides, arranged in n congruent pairs, and let k be the number of cycles. When the region is formed into a closed surface, all the vertices of the i-th cycle coincide at a point P_i on the surface. We now cut the surface into simply connected pieces as follows. We first draw k loop cuts about the points $P_1, \ldots, P_k$, cutting out k simply connected pieces. One of these cuts is taken as an initial boundary, and the piece which it encloses is discarded. The surface that remains outside the loop cuts can be rendered simply connected by cutting along the curves where congruent edges have been joined. These cuts give us the original region save for the removal of pieces at the vertices.

Since the loop cuts are not to be counted, we have made $\nu = n$ cross-cuts. In addition to the $k - 1$ pieces at $P_1, \ldots, P_k$, we have a single other piece, so $\alpha = k$. We have, then, from (14),

$$p = \tfrac{1}{2}(n - k + 1). \tag{17}$$

If the unclosed fundamental region is not simply connected, formula (17) does not hold.

93. The Cases $p = 1$ and $p > 1$.—Consider, first, the case in which the infinitely sheeted region ϕ is mapped on Q_0. Let R_0 be the fundamental region of the group T_n, formed by means of isometric circles. R_0 has a finite number of sides. Let Σ_0 be the map of R_0 on ϕ. Σ_0 is a fundamental region for the group of transformations of ϕ into itself. Each point of ϕ_0, or a congruent point in one of the copies of F_0, is contained in Σ_0; and, except on congruent sides, Σ_0 contains no two congruent points. Congruent sides of Σ_0 are superposed; and, if we join them to form a closed surface, we have an exact copy of the closed Riemann surface F. The closed surface formed by joining the congruent sides of R_0 together is a one-to-one continuous transformation of this surface and, hence, has the same genus as the Riemann surface.

We now replace the arcs bounding R_0 by their chords to form a rectilinear polygon (Fig. 75); and we consider the angles in the two polygons. Let R_0 have $2n$ sides and k cycles.

Since there are no elliptic transformations in the group, the sum of the angles at the vertices of each cycle is 2π. Hence, the angles of the circular arc polygon amount to $A = 2\pi k$. The sum of the angles of the rectilinear polygon of $2n$ sides is $B = 2\pi(n - 1)$. Since the latter is the greater, we have $B - A > 0$. Making use of (17), we have

$$B - A = 2\pi(n - k - 1) = 4\pi(p - 1) > 0,$$

whence, $p > 1$.

In the case in which ϕ is mapped on the finite plane, the group T_n is the doubly periodic group. A fundamental region for the group is the period parallelogram (Fig. 13). We note, exactly as before, that this fundamental region has the same genus as the Riemann surface F. The opposite sides of the period parallelogram are congruent and there is one cycle; hence $n = 2$, $k = 1$. We have, then, from (17), $p = 1$.

The surface ϕ can be mapped conformally on the finite plane if, and only if, $p = 1$. It can be mapped conformally on the circle Q_0 if, and only if, $p \geqslant 2$.

The fact that the fundamental region has on its boundary no limit point of the group and no fixed point of an elliptic transformation has the following consequence: Let z_0 lie in the domain of definition of the uniformizing functions and let P_0 be the corresponding point on the Riemann surface F. A sufficiently small neighborhood of z_0 contains no two congruent points; that is, no two points corresponding to the same point of F. The correspondence between the points of the z-plane in the neighborhood of z_0 and the points of the Riemann surface in the neighborhood of P_0 is one-to-one.

We summarize our results in the first part of the following theorem:

THEOREM 9.—*An algebraic functions can be uniformized by means of*

(*a*) *Rational functions, if* $p = 0$,
(*b*) *Elliptic functions, if* $p = 1$,
(*c*) *Fuchsian functions of the first kind, if* $p \geqslant 2$,

in such a manner that in a sufficiently small neighborhood of a point in the domain of existence of the uniformizing functions the correspondence between the points of the plane and the points of the Riemann surface of the algebraic function is one-to-one.

The three cases are mutually exclusive. Further, the most general such pair of uniformizing functions, in each case, is got from any given pair by subjecting the uniformizing variable to a linear transformation.

It is conceivable that by constructing uniformizing functions in some way other than by the use of the sheeted surface ϕ, we should be able to uniformize an algebraic function in two of the ways stated in the theorem. Let $Z = Z(z)$, $W = W(z)$ be one pair of uniformizing functions and let S (the whole plane, the finite plane, or the interior of a circle) be the domain of existence of the functions. Let $Z = Z_1(t)$, $W = W_1(t)$ be another pair with the domain of existence S'. Let t_0, be a point in S';

let $P(Z_0, W_0)$ be the corresponding point on the Riemann surface F; and let $z_0, z_1, \ldots$ be the corresponding points in S. The equations $Z = Z(z)$, $Z = Z_1(t)$ together with the condition that $z = z_0$ when $t = t_0$, define a function $z = \varphi_0(t)$ analytic in the neighborhood of t_0. The equations define other functions $\varphi_1(t), \varphi_2(t), \ldots$ analytic at t_0 and taking on the values $z_1, z_2, \ldots$ at t_0. Each function can be extended analytically throughout S'; but, since for each value of t the corresponding values of z are distinct, each of the functions is single valued in S'.[1] Hence $z = \varphi_0(t)$ is a function single valued in S'. Interchanging the rôles of S and S', the inverse of the function is single valued in S.

The function $z = \varphi_0(t)$ maps S' conformally on S. But of the three domains, the whole plane, the finite plane, and the interior of a circle, no one can be mapped conformally on the other. Hence S and S' are domains of the same kind. An algebraic function which can be uniformized in the way stated in the theorem by one of the three kinds of functions cannot be so uniformized by either of the other two kinds of functions.

Suppose now that S and S' are domains of the same kind. Then, $z = \varphi_0(t)$ maps the whole plane conformally on itself, or the plane bounded by a point on the plane bounded by a point, or the interior of a circle on the interior of a circle. In each of these cases, $z = \varphi_0(t)$ is a linear transformation, which establishes the final statement of the theorem.

94. More General Fuchsian Uniformizing Functions.—In the preceding method, there are no elliptic or parabolic cycles. We now consider the possibility of uniformization by means of Fuchsian functions belonging to groups of the first kind without this restriction. We treat, first, the case $p > 0$.

On the Riemann surface F of the algebraic function, let s points, $P_1, P_2, \ldots, P_s$, be selected at which the mapping from the surface to the plane of the uniformizing variable is not to be one-to-one. With P_i we associate an integer $\nu_i > 1$ and require that the mapping be one-to-ν_i in the neighborhood of P_i. We shall admit $\nu_i = \infty$ as a possible value.

Let the surface be rendered simply connected by a system of $2p$ cuts beginning and ending at an ordinary point O, the cuts being so made that $P_1, \ldots, P_s$ do not lie on the boundary. Let P_i be joined to the boundary at O by a cut C_i. The cuts C_i

[1] OSGOOD, "Lehrbuch der Funktionentheorie," 2nd ed., p. 396.

shall have no common points except the common end point O (Fig. 76). The region F_0 resulting from the $2p$ cross-cuts and the s cuts C_i is simply connected. F_0 has $4p + 2s$ sides; namely, the opposite banks of the $2p + s$ cuts, and an equal number of vertices.

We now provide ourselves with an infinite number of copies of F_0 and proceed to superpose them and join them together in such a manner that (a) the region at each stage is simply connected; (b) in the limit region the neighborhood of O in each sheet is a plane piece; and (c) ν_i copies of F_0 hang together to form a branch point wherever P_i occurs. Take one copy—call it ϕ_0—as the initial region. We superpose copies of F_0 and join

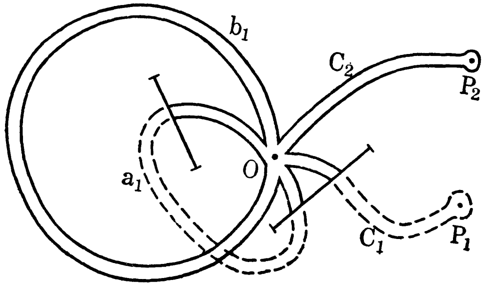


FIG. 76.

along each of the sides of ϕ_0 (opposite banks of the same cut in ϕ_0 and the copy being brought together) to form ϕ_1. We superpose copies of F_0 and adjoin along each of the sides of ϕ_1 to form ϕ_2; and so on. In this process, whenever the neighborhood about a point O is filled out, we join the copy which fills out the neighborhood along two adjacent sides and reduce O to an interior point. Also, when $\nu_i - 1$ copies hang together at P_i, the next copy adjoined along C_i shall have both banks of its cut C_i joined to the two free banks there, thus reducing P_i to an interior branch point. If $\nu_i = \infty$, this latter situation never arises, and P_i is never an interior point.

We can show, easily, as in the footnote of Sec. 91, that there are free sides at every step. The process of construction thus continues indefinitely.

Let ϕ be the infinitely sheeted limit region of the sequence $\phi_0, \phi_1, \ldots$ Let $W = F(Z)$ be the algebraic function to be uniformized. $F(Z)$ is a single valued function of Z on the severed surface F_0. If we let each point of each copy bear the same values of W and Z as were originally associated with the point, the values of W fit together to form a function single valued and analytic, except for poles and branch points, on ϕ. For, we so superpose the copies that each point has the same Z as in its original position; and we join up the copies along lines at which W has the same values in the copies, so that W in the one copy is the analytic continuation of W in the adjacent copy.

The sequence ϕ_0, $\phi_1, \ldots$ satisfies the conditions of Theorem 21, Sec. 85; hence, ϕ can be mapped conformally either on Q_0 or on the finite z-plane. Let $Z = Z(z)$ be a function performing this mapping.

Corresponding to the group of conformal transformations of ϕ into itself, got by carrying one copy of F_0 in ϕ into another or into the same copy, is a group of conformal transformations $z_n = T_n(z)$ which carry Q_0 into itself or the finite plane into itself. The set T_n is a group of linear transformations. The map of any copy F_0 is an fundamental region for the group. Conversely, the map on ϕ of a fundamental region in Q_0 or in the finite plane, of the group T_n is a region Σ_0 which is a fundamental region for the group of transformations of ϕ into itself. If congruent edges of Σ_0 be brought together, we have an exact copy of the unsevered region F. The fundamental region in the plane is of genus $p > 0$.

We now show that the mapping cannot be done on the finite plane. A reference to Secs. 59 and 60 and the figures connected therewith shows that the fundamental regions of all groups with a single limit point are of genus zero, with the exception of the doubly periodic group. The latter group is ruled out by the following considerations. Corresponding to a point P_i (ν_i finite) of ϕ is an elliptic fixed point a_i of order ν_i in the z-plane. For, each point in the neighborhood of a_i has $\nu_i - 1$ congruent points in the neighborhood of a_i, these ν_i points corresponding to ν_i congruent points in different copies of F_0 which are joined together at P_i. The doubly periodic group, however, contains no elliptic transformations. Again, if $\nu_i = \infty$, P_i is a boundary point of ϕ. If the group is doubly periodic, a period parallelogram is mapped on a region not extending to the boundary of ϕ; and not enclosing points in a suitably small neighborhood of the superposed points P_i. Then there are points in the z-plane not having congruent points in the period parallelogram, which is contrary to fact. Hence, in this case also, the group is not the doubly periodic group. It follows that in all cases ϕ can be mapped on Q_0.

The mapping function $Z = Z(z)$ and the function $W = W(z) \equiv F[Z(z)]$, which together uniformize the algebraic function, are automorphic with respect to the group T_n. For, when z is carried into $T_n(z)$, the corresponding point P of ϕ is carried into a point P_n at which Z and W have the same values as at P.

THEOREM 10.—*Let $P_1, P_2, \ldots, P_s$ be points on the Riemann surface F of an algebraic function $W = F(Z)$ of genus greater than zero. With each point P_i let an integer $\nu_i > 1$ be associated ($\nu_i = \infty$ is admitted). Then, the algebraic function can be uniformized by means of Fuchsian functions of the first kind,*

$$Z = Z(z), \quad W = W(z) = F[Z(z)], \tag{18}$$

in such a manner that in a sufficiently small neighborhood of a point a in the principal circle of the group the correspondence between the points of the plane and the points of F is one-to-one, except when a corresponds to P_i, in which case the correspondence is ν_i-to-one.

The most general such uniformizing functions are got from one pair by subjecting the uniformizing variable to a linear transformation.

The latter part of the theorem is proved as in the preceding theorem. Let the functions (18) exist in the principal circle Q_0; and let $Z = Z_1(t)$, $W = W_1(t)$ be uniformizing functions with the principal circle Q and with the properties stated in the theorem. Let t_0 be a point of Q; let $P_0(W_0, Z_0)$ be the corresponding point of F; and let $z_0, z_1, \ldots$ be the corresponding points of Q_0. If P_0 is an ordinary point of F, the equations $Z = Z(z)$ $Z = Z_1(t)$ define distinct function elements $z = \varphi_0(t)$, $z = \varphi_1(t)$, $\ldots$ analytic at t_0 and such that $\varphi_k(t_0) = z_k$. If $P_0 = P_i$ we have

$$Z - Z_0 = c_0(z - z_0)^{\nu_i} + \cdots = d_0(t - t_0)^{\nu_i} + \cdots, c_0 \neq 0, d_0 \neq 0$$

(with suitable changes in the first member if P_i is a branch point of F or a point at ∞). On solving for z, we find distinct function elements in this case also. Any one of these functions, say $z = \varphi_0(t)$, can be extended analytically throughout Q and is a single-valued analytic function of t in Q. Similarly, the inverse is single valued in Q_0; whence, $z = \varphi_0(t)$ maps Q conformally on Q_0. We thus have a linear transformation of the uniformizing variable.

95. The Case $p = 0$.—If the genus of the surface is zero, the initial cross-cuts are absent. From an ordinary point O, we draw cuts $C_1, \ldots, C_s$ to $P_1 \ldots, P_s$, respectively, to form the severed surface F_0 (as in Fig. 77). F_0 has $2s$ sides and vertices, s vertices being at O. We take an initial copy of F_0—call it ϕ_0—and adjoin copies along each of its sides to form ϕ_1, and so on. As before, when ν_i copies lie about P_i, we join up the sides to reduce P_i to an interior point. We join up the sides about O in any sheet when s vertices meet there.

If $s \geqslant 4$, we find readily, by the method of the footnote of Sec. 91, that we can continue to adjoin copies *ad infinitum.* The limit surface ϕ of the sequence $\phi_0, \phi_1, \ldots$ is infinitely many sheeted. This will also appear presently from the fact that the inequality (21) cannot be satisfied.

If $s = 1$, the construction is impossible. It is, likewise, impossible if $s = 2$ and $\nu_1 \neq \nu_2$. For, two copies of F_0 suffice to fill up the region about each verted O. Hence, each added copy is joined along the line P_1OP_2; and when the sides close up about one point they necessarily close up about the other. If $\nu_1 = \nu_2$, we get for ϕ a closed surface containing ν_1 copies of F_0, if ν_1 is finite, and an infinitely sheeted surface, if ν_1 is infinite.

If $s = 3$, we may be able to add copies *ad infinitum* so that ϕ is an infinitely sheeted surface; or the surface may close after a finite number of copies have been added, and ϕ is a finite-sheeted closed surface.

The surface ϕ can be mapped, according to circumstances, on a circle, on the finite plane, or on the whole plane (the last if ϕ is closed). The mapping function $Z = Z(z)$ and the function $W = W(z) \equiv F[Z(z)]$ are uniformizing functions. They are automorphic with respect to the group of transformations in the z-plane corresponding to the group of transformations of ϕ into itself. The group in the z-plane is a Fuchsian group of the first kind, a group with a single limit point at infinity, or a finite group, according to the mapping. We now distinguish between the three cases.

Suppose that ϕ is mapped on Q_0; and let the fundamental region R_0 be formed. R_0 is of genus zero. On its boundary lie points corresponding to each of the points P_i, and each such point belongs to a cycle of angle $2\pi/\nu_i$. Let R_0 have $2n$ sides and k cycles. Of the latter, $k - s$ are of angle 2π. We now compare the sum of the angles A of the fundamental region with the sum of the angles B of the rectilinear polygon formed by joining successive vertices of the region by straight lines (Fig. 75). We have

$$A = \frac{2\pi}{\nu_1} + \cdots + \frac{2\pi}{\nu_s} + 2\pi(k - s), \quad B = 2\pi(n - 1),$$

whence,

$$B - A = 2\pi\left[n - k - 1 + s - \left(\frac{1}{\nu_1} + \cdots + \frac{1}{\nu_s}\right)\right] > 0.$$

But, since $p = 0$, we have, from (17), $n = k - 1$; hence,

$$\frac{1}{\nu_1} + \cdots + \frac{1}{\nu_s} < s - 2. \tag{19}$$

If the mapping is on the finite plane, we can take a rectilinear polygon as fundamental region; for, we have found such fundamental regions for all possible groups with infinity as the only limit point. Then $B = A$, and we have

$$\frac{1}{\nu_1} + \cdots + \frac{1}{\nu_s} = s - 2. \tag{20}$$

We have already found that in the remaining cases, namely those for which

$$\frac{1}{\nu_1} + \cdots + \frac{1}{\nu_s} > s - 2, \tag{21}$$

all possible solutions in integers lead to finite groups (Sec. 57), and, conversely, that for all finite groups the inequality is satisfied.

Those cases in which ϕ is closed may also be determined by Euler's formula. Let ϕ be deformed into the surface of a sphere. Then, the copies with their sides and vertices may be looked upon as the faces, edges, and vertices, respectively, of a solid. Let F be the number of faces. Each face has $2s$ sides; so there are Fs edges altogether. Each face has s O-vertices; but these meet in groups of s, giving F O-vertices in the solid. Each face has one P_i-vertex, but these fall together in sets of ν_i, giving F/ν_i of these vertices in the solid. Euler's formula, $F + V = E + 2$, then becomes

$$F + \left[F + F\left(\frac{1}{\nu_i} + \cdots + \frac{1}{\nu_s}\right)\right] = Fs + 2,$$

or

$$\frac{1}{\nu_i} + \cdots + \frac{1}{\nu_i} = s - 2 + \frac{2}{F} > s - 2.$$

We state our results in the following theorem. We enumerate the possible solutions of (20) and (21) and specify the types of functions resulting. The proof of the last statement of the theorem proceeds as in the previous theorems and will not be repeated here.

THEOREM 11.—*Let* $P_1, \ldots P_s$ $(s > 1)$ *be points on the Riemann surface F of an algebraic function of genus zero. With each point* P_i *let an integer* $\nu_i > 1$ *be associated. Here* ν_i *may be infinite; and* $\nu_1 = \nu_2$ *if* $s = 2$. *Then the algebraic function can be uniformized by means of automorphic functions in such a manner that, in the neighborhood of a point a in the domain of definition of the functions, the correspondence between the points of the plane and the points of F is one-to-one, except when a corresponds to* P_i, *in which case the correspondence is* ν_i*-to-one.*

The uniformizing functions are:

1. *Rational (polyhedral) functions, if*

   (*a*) $s = 2$, ν_1 *finite;*

   (*b*) $s = 3$, $\dfrac{1}{\nu_1} + \dfrac{1}{\nu_2} + \dfrac{1}{\nu_3} > 1.$

2. *Simply periodic functions, if*

   (*a*) $s = 2$, $\nu_1 = \infty$;

   (*b*) $s = 3$, $\nu_1 = \nu_2 = 2$, $\nu_3 = \infty$.

3. *Elliptic functions, if*

   (*a*) $s = 3$, $\dfrac{1}{\nu_1} + \dfrac{1}{\nu_2} + \dfrac{1}{\nu_3} = 1$;

   (*b*) $s = 4$, $\nu_1 = \nu_2 = \nu_3 = \nu_4 = 2$.

4. *Fuchsian functions of the first kind in all other cases.*

The most general such uniformizing functions are got from one pair by subjecting the uniformizing variable to a linear transformation.

The uniformizing variables (2) at the beginning of this chapter fall under 2(*a*). The Riemann surface of $W = \sqrt{1 - Z^2}$ is a two-sheeted surface with branch points at ± 1. The points P_1, P_2 are at $z = \infty$ in the two sheets.

96. Whittaker's Groups.—The following groups for the uniformization of the hyperelliptic algebraic functions are due to Whittaker.[1] They appear as subgroups of groups generated by elliptic transformations of period two.

In Theorem 11 let the Riemann surface of genus zero be the Z-plane; let the relative branch points be

$$e_1, e_2, \cdots, e_{2p+2}, \quad p > 1;$$

[1] *Phil. Trans.*, vol. 192, pp. 1–32, 1899.

and let $\nu_i = 2$ for all the points. The limit surface can be mapped on the circle Q_0.

The initial surface ϕ_0 is formed by cuts extending from a point O to the points $e_1, \ldots, e_{2p+2}$ (Fig. 77, for the case $p = 2$). Any adjacent copy is joined to ϕ_0 along the two banks of the

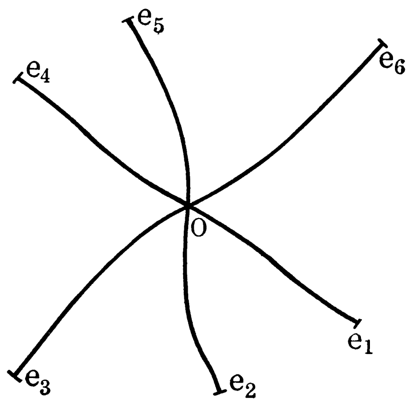


FIG. 77.

cut extending from O to one of the points e_i. When ϕ_0 is placed upon this copy, the copy, falls on ϕ_0, and e_i remains fixed.

Let ϕ_0 and the adjacent copy be mapped on regions S_0, S_i in the z-plane. Either region is a fundamental region for the

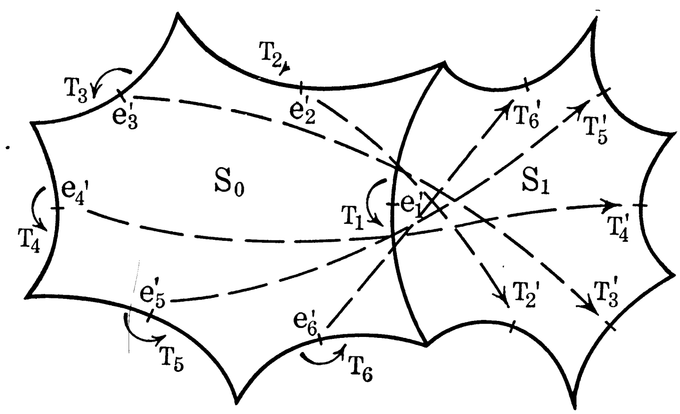


FIG. 78.

group Γ in the z-plane corresponding to the group of conformal transformations of ϕ into itself got by carrying ϕ_0 into the various copies composing ϕ. When S_0 is carried by a transformation T_i into S_i, S_i is carried into S_0 and a point e_i' on the common boundary remains fixed. T_i is thus an elliptic transformation of period two with e_i' as fixed point. The transformations T_1, $T_2, \ldots, T_{2p+2}$, which connect congruent sides of S_0 and so generate the group Γ, are all elliptic of period two (Fig. 78).

The automorphic function $Z = Z(z)$—the inverse of the mapping function—maps S_0 on the whole Z-plane and so takes on each value once in S_0. Hence, all simple automorphic functions belonging to the group are rational functions of $Z(z)$. Any function which can be uniformized by means of these simple automorphic functions, according to Theorem 7, is of genus zero.

Now, let ϕ_0' be the surface formed by adjoining to ϕ_0 one copy of the severed Z-plane, the junction being made along the two banks of Oe_1. The limit surface ϕ may be looked upon as made up of copies of this two-sheeted surface joined together without relative branch points.

To the group of transformations of ϕ into itself got by carrying ϕ_0' into each of the copies composing ϕ, there corresponds a group Γ' in the z-plane. Γ' is a subgroup of Γ. A fundamental region for Γ' is the map S_0' of ϕ_0'. S_0' consists of S_0 together with the adjacent region $S_1 = T_1(S_0)$. The generating transformations, connecting the congruent sides of S_0', are (Fig. 78)

$$T_2' = T_1T_2, \quad T_3' = T_1T_3, \quad \cdots, \quad T'_{2p+2} = T_1T_{2p+2}.$$

Γ' contains no elliptic transformations.

Now, ϕ_0' is a severed two-sheeted surface with branch points at $e_1, \ldots, e_{2p+2}$. This surface is the Riemann surface for the two-valued function

$$W^2 = A(Z - e_1)(Z - e_2) \cdots (Z - e_{2p+2}).$$

The coordinates (Z, W) on ϕ are simple automorphic functions, $Z = Z(z)$, $W = W(z)$, which uniformize this hyperelliptic function.

97. The Transcendental Functions.—Each analytic function

$$W = F(Z) \tag{22}$$

possesses a Riemann surface F spread over the Z-plane on which the function is single valued. The surface is got by analytic continuation from some element of the function. The details will be found in many texts on the Theory of Functions.

The Riemann surfaces of analytic functions exhibit the greatest variety. The number of sheets may be finite or denumerably infinite. There may be branch points of finite or infinite order or branch points may be absent. The function may fail to be analytic at isolated points or along curves, and sheets may contain holes, or lacunary spaces, into which the function cannot be continued analytically and which do not belong to the surface.

We shall consider certain of the isolated singular points of the function as belonging to F. A point in a plane sheet at which the function becomes infinite—a pole—shall belong to F. Also, we include a branch point of finite order at which the function is continuous or becomes infinite. All other singular points shall not belong to F, but lie on its boundary. F consists of interior points.

We shall show that it is possible to make a system of cuts, finite or infinite in number, in the surface F which will render it simply connected.

We may suppose, without loss of generality, that all points of F which lie at infinity are in plane sheets. This can be accomplished by a linear transformation. For, the branch points of F are denumerable, and their Z-coordinates do not include all points of the Z-plane; it suffices to make a linear transformation carrying a point which is not the Z-coordinate of a branch point to infinity. We may suppose, also, that at each branch point, $Z = X + iY$, both X and Y are irrational. For, the points whose abscissas and ordinates differ by rational numbers from those of the branch points lie on a denumerable set of lines parallel to the X- and Y-axes; and we have but to make a translation carrying the origin to a point not lying on one of these lines.

We shall now cut the surface up into square elements (Sec. 86) after which we shall put the pieces together in such a manner that the resulting surface is simply connected. The component square elements shall lie *together with their boundaries* in F. We put aside the case in which F consists of the whole plane.

Consider, first, the points of F at infinity. Each such point lies in an element (the exterior of a square) bounded by lines $X = \pm n$, $Y = \pm n$, where n is a positive integer. We take for n the minimum value such that the element belongs to F. This gives a finite or denumerably infinite number of elements belonging to F.

We next cut what remains of the surface by lines $X = m$. $Y = n$, where m and n take on all integral values. Whatever plane unit squares or superposed squares winding about a *single* branch point as are cut out shall be exempt from further cutting. There is at most a denumerably infinite number.

We divide what remains into quarter squares by lines $X = m + \frac{1}{2}$, $Y = n + \frac{1}{2}$, and take out the square elements. We divide again into quarter squares, taking out the square elements; and so on.

In this process, no branch point lies on the side of a square, for on a side either X or Y is rational. The resulting finite or denumerably infinite set of elements contains all points of F; for the neighborhood of any point of F is either a plane piece or an isolated branch point and is cut up when the squares are small enough.

We propose, now, to put the squares back together in such a way that the resulting surface is simply connected.

When two square elements abut, they have a common piece of straight line boundary equal in length to the side of the smaller square (if unequal). We shall call such a piece an *edge*. Two square elements may have several common edges if each bears a branch point and several sheets. Likewise, an element containing a point at infinity and an adjacent element will have two common edges, if they have a common vertex. Since the number of sides of each square element is finite, it follows that the number of edges is denumerable. We can, therefore, write them in serial order:

$$l_1, l_2, l_3, \cdots \tag{23}$$

Let φ_0 be one of the square elements of F; and l_{n_1} be the first edge of (23) forming part of the boundary of φ_0. We adjoin to φ_0 the square element abutting along the edge l_{n_1} and close up along this edge. If there is a common edge adjacent to l_{n_1} (a possibility if one element is the exterior of a square) we close this edge also. Call the resulting surface φ_1. We cancel from (23) any other common edges of the two square elements.

In general, we form φ_n from φ_{n-1} as follows: Let l_{n_n} be the first edge of the sequence (23) which forms part of the boundary of φ_{n-1}, after cancelling from the sequence each edge common to two elements of φ_{n-1} which have not been joined along that edge. We adjoin to φ_{n-1} the square element abutting along l_{n_n}, closing up the edge l_{n_n}. We shall close up further edges, if possible, in the following way: If in tracing the boundary of the new region we encounter any edge followed by itself, we extend the line of junction by closing this edge. If an edge follows itself in the new boundary, we close it up, and continue the process as long as possible. The resulting surface is φ_n.

The surface φ_n is simply connected. Assume that φ_{n-1} is simply connected. The line of junction made in forming φ_n is a cross-cut (possibly a sigma cross-cut) cutting φ_n into two simply connected pieces—φ_{n-1} and the added element. Hence, φ_n is simply connected. Since φ_0 is simply connected, the result follows by induction; and the limit surface ϕ_0 defined by the sequence

$$\varphi_0,\ \varphi_1,\ \varphi_2,\ \cdot\ \cdot\ \cdot \rightarrow \phi_0 \tag{24}$$

is simply connected.

The surface ϕ_0 contains all the square elements of F. Since F is connected, we can draw a curve in F from φ_0 to any given element s, meeting a finite number of elements φ_0, s_1, s_2, . . . , s_n, s and crossing the edges l_{k_0}, l_{k_1}, . . . , l_{k_n} separating the successive elements. After a finite number of steps, l_{k_0} either will be cancelled or will be the first uncancelled edge of (23) on the boundary of some φ_m. In either case, φ_{m+1} contains s_1. Similarly, s_2, . . . , s_n, s are reached in a finite number of steps.

Consider, now, the possible forms of the surface ϕ_0. It may happen that there is a finite number of square elements and that the surface closes up completely. Then F is of genus zero; and the function $F(Z)$, having no other singularities than poles on F, is algebraic.

It may happen that there is a finite number of square elements but that ϕ_0 has a boundary. The function is again algebraic but of genus greater than zero. The boundary of ϕ_0 constitutes a set of cuts in F which render it simply connected. We put these algebraic cases aside.

If F contains an infinite number of square elements, the function is not algebraic. There are two possibilities. ϕ_0 may have no free edges left, in which case $F(=\phi_0)$ is simply connected. Or, ϕ_0 may have free edges. In the latter case, the free edges constitute a system of cuts in F which render the surface simply connected.

The sequence (24) is one to which Theorem 21, Sec. 85, applies; so, ϕ_0 can be mapped either on the unit circle Q_0 or on the finite plane. If ϕ_0 has free edges, the mapping is certainly on a circle (Theorem 22, Sec. 85).

If ϕ_0 has free edges, we proceed to build up a limit surface. Let the edges on the boundary be arranged serially

$$m_1, m_1'; m_2, m_2'; \cdots, \tag{25}$$

m_i and m_i' being opposite banks of the same edge. We take a copy of ϕ_0 and join to ϕ_0 so that m_1 in ϕ_0 is joined to m_1' in the copy. We continue the junction as far as possible in each direction by closing when an edge follows itself along the boundary. We then take another copy of ϕ_0 and adjoin along the first edge of (25) on the boundary of the original ϕ_0 which has not been closed up, extending the line of junction as before. We continue this process until all edges of (25) are exhausted and the original ϕ_0 is completely embedded. Call the resulting surface ϕ_1.

We next arrange the edges bounding ϕ_1 serially and adjoin copies of ϕ_0 around the boundary in a similar manner to form ϕ_2; and so on. At each step, ϕ_n can be mapped on a circle. Hence, the limit surface ϕ defined by the sequence

$$\phi_0, \phi_1, \phi_2, \cdots \to \phi \tag{26}$$

can be mapped on the circle Q_0 or on the finite plane.

Let ϕ (take $\phi = \phi_0 = F$, if the Riemann surface is simply connected) be mapped on Q_0 or on the finite z-plane. Let $Z = Z(z)$ be the inverse of the mapping function. Then the functions

$$Z = Z(z), \qquad W = W(z) = F[Z(z)] \tag{27}$$

are analytic except possibly for poles in Q_0 or in the finite plane. The poles of $Z(z)$ are the points in the z-plane corresponding to the points at infinity on ϕ; the poles of $W(z)$ correspond to the points of ϕ at which $F(Z)$ becomes infinite.

If the mapping is on the finite plane, at least one of the functions (27) has an essential singularity at infinity. Otherwise, both functions would be rational and $F(Z)$ be algebraic of genus zero.

If the mapping is on Q_0, at least one of the functions has a singularity at each point of the circumference. Suppose, on the contrary, that both functions are analytic at a point a on the circumference. The function $Z = Z(z)$ maps a sufficiently small circle K enclosing a on a region K' in the Z-plane, where K' is a plane piece or is the neighborhood of a branch point of finite order, according as the derivative $Z'(a)$ does not or does vanish. This function maps the part of K lying within Q_0 or a portion of ϕ. But, W can be extended analytically throughout K', so all of K' belongs to ϕ. Then, the corresponding region K lies in Q_0, contrary to hypothesis.

The surface ϕ possesses a group of conformal transformations into itself got by carrying the initial copy ϕ_0 into the various copies of ϕ_0 of which ϕ is composed. There corresponds in the z-plane a group of conformal transformations carrying Q_0 into itself or the finite plane into itself, and which are, therefore, linear transformations. This group consists solely of the identical transformation, if there is but one copy of ϕ_0. The functions (27) are invariant under the transformations of this group.

We have established the first part of the following theorem:

THEOREM 12.—*Any transcendental function*

$$W = F(Z)$$

can be expressed parametrically in terms of two functions

$$Z = Z(z), \qquad W = W(z) \tag{27}$$

which are analytic except for poles in a domain Σ, consisting either of the interior of the unit circle Q_0 or of the finite z-plane, such that each pair of values (Z, W) satisfying the functional relation is given by one or more values of z in Σ, and such that the correspondence between the points in a sufficiently small neighborhood of a point of Σ and the points on the Riemann surface of the function is one-to-one.

The most general functions with these properties may be got from one pair by subjecting z to a linear transformation.

If the Riemann surface of the function is not simply connected, the functions (27) *are invariant under an infinite group of linear transformations.*

The second paragraph of the theorem is proved in the usual way.

As to the final paragraph, it is conceivable that the sequence (26) should break off after a finite number of steps due to the junction of all free edges. The mapping would still be possible, but the group would be finite. The group would contain an elliptic transformation with a fixed point z_0 in Σ. At z_0 the one-to-one character of the correspondence would break down; so this supposition is untenable.

The functions (27) are automorphic, according to the definition of Sec. 39, if they are single valued. This is the case if Σ consists of the finite plane. If, however, Σ consists of the interior of

Q_0, it may be possible to extend one or both functions outside the circle and back again in such a way that new branches appear.

For example, let $F(Z)$ be analytic in a semicircle and have the boundary of the semicircle as a natural boundary. Here, ϕ_0 is the semicircle and it is mapped on Q_0. Q_0 is a natural boundary for $W(z)$. However, $Z(z)$, which maps Q_0 on the semicircle, can be extended analytically across the circumference, and is, in fact, a two-valued function.

As in the case of algebraic functions, it is possible to get quite different pairs of functions by altering the one-to-one character of the correspondence. We may select certain points of the Riemann surface at which there is to be ν-to-one correspondence ($\nu > 1$), the correspondence being one-to-one elsewhere. We shall not work out the details.

CHAPTER X

UNIFORMIZATION. GROUPS OF SCHOTTKY TYPE

98. Regions of Planar Character.—The preceding methods of uniformization have been arrived at by the mapping of simply connected regions. In order to be able to uniformize by means of automorphic functions belonging to groups whose fundamental regions are not simply connected, we shall now consider the mapping of multiply connected regions on plane regions.

A region is said to be *of planar character*[1] if every loop-cut in the region separates it into two parts. This is a property of all plane regions. It is clear that a region which is not of planar character cannot be mapped in a one-to-one manner on, or otherwise continuously deformed into, a plane region; if it could, a loop-cut not cutting the original surface in two would be carried into a loop-cut not cutting the plane region in two, which is impossible.

Let Σ be a finite-sheeted region with a finite number of branch points whose boundary consists of a finite number of closed curves $B_1, B_2, \ldots, B_m$. Further, let Σ be of planar character. We propose to show that Σ can be mapped conformally on a plane region with m bounding curves.

As a first step, we cut Σ into a simply connected region and perform a preliminary mapping. We make a system of m regular cuts $C_1, \ldots, C_m$. The cut C_i joins an ordinary point O of the region to a point of B_i; and the cuts are to have no common points other than O. The severed region Σ_0 has a single boundary. We suppose the cuts so made (or the curves so numbered) that in passing around the boundary in the positive sense we encounter $B_1, B_2, \ldots, B_m$ in order. The region Σ_0 is simply connected. If not, a cross-cut q can be made which does not separate Σ_0 into two parts. If q is not already a sigma cross-cut, it can be deformed into one, since it joins points on the same boundary. If the stem of the sigma cross-cut be erased we have a loop-cut in Σ_0—and hence in Σ—which does not separate Σ, which is contrary to hypothesis.

[1] German, *schlichtartig*.

Figure 79 shows a region of planar character lying on a two-sheeted elliptic surface. In this surface, a loop-cut has been drawn, and the area enclosed by a curve B_1 in the upper sheet has been removed.

We provide ourselves with copies of Σ_0 and proceed to build up an infinitely sheeted surface. We take one copy as the initial region ϕ_0; we adjoin copies along the $2m$ banks of C_i to form ϕ_1; we adjoin copies along the free banks of the cuts C_i of ϕ_1 to form ϕ_2; and so on. In this process, whenever the regions fit together to fill up the plane neighborhood about O,

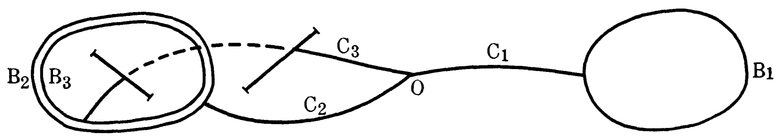


FIG. 79.

we join up two successive banks meeting at O to reduce O to an inner point of the region. The region ϕ_0, ϕ_1, . . . form a sequence to which Theorem 21, Sec. 85, applies. In this case, owing to the presence of the boundaries B_i, the mapping of the infinitely sheeted limit surface ϕ can be made on a circle.

We shall map the surface ϕ, spread over the Z-plane, on the upper half of the t-plane bounded by the real axis, in such a manner that a point of B_m on the boundary of the initial copy ϕ_0 is carried to infinity. Let $Z = Z(t)$ be the mapping function. The inverse function $t = g(Z)$ is single valued on ϕ provided we do not continue the function analytically across the boundaries B_i. To each point of ϕ, there corresponds one, and only one, point of the upper half t-plane. On the unsevered surface Σ, $g(Z)$ is an infinitely many-valued function which takes on each value in the upper half plane once, and only once. In sufficiently small neighborhoods of a point Z_0 on Σ and of one of the corresponding points t_0 in the t-plane, the correspondence between the points of the surface and of the plane is one-to-one.

Corresponding to the group of conformal transformations of ϕ into itself, got by carrying ϕ_0 into any of the copies Σ_0 of which ϕ is built up, we have a group of linear transformations in the t-plane which carry the upper half plane into itself. Let S_0 be the map of ϕ_0 (Fig. 80). A fundamental region for the group is S_0 together with the reflection of S_0 in the real axis. The group is a Fuchsian group of the second kind.

99. Some Accessory Functions.—We propose to set up certain functions which are single valued on ϕ and which take on equal values along opposite banks of each C_i in ϕ_0. The functions will then be single valued on the surface Σ, got by closing up the cuts in ϕ_0. In terms of the variable t we are concerned with

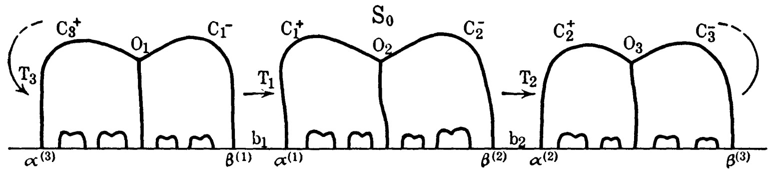


FIG. 80.

functions which take on the same values along congruent boundaries of S_0. The functions which we employ are set up by means of series and products whose convergence has already been established for the Fuchsian group of the second kind in Sec. 50.

We use, as we have previously done, the notation $t_n = T_n(t)$, where $T_0, T_1, \ldots$ are the transformations of the group. Let τ and η be two inner points of S_0, and consider the function

$$\psi(t) = \prod_{n=0}^{\infty} \frac{t - \tau_n}{t - \bar{\tau}_n} \cdot \frac{t - \bar{\eta}_n}{t - \eta_n}, \tag{1}$$

where bars indicate conjugate imaginaries. This function is analytic in the whole t-plane except for poles at the points η_n in the upper half plane and $\bar{\tau}_n$ in the lower half plane, and essential singularities at the limit points of the group on the real axis. It is different from zero except at the points τ_n and $\bar{\eta}_n$. We shall not be interested in its behavior in the lower half plane. At an ordinary point of the real axis we have

$$|t - \tau_n| = |t - \bar{\tau}_n|, \qquad |t - \eta_n| = |t - \bar{\eta}_n|,$$

and

$$|\psi(t)| = 1. \tag{2}$$

Consider the behavior of $\psi(t)$ when a transformation

$$t_k = T_k(t) = \frac{at + b}{ct + d}, \qquad ad - bc = 1, \tag{3}$$

of the group is made. We have

$$\psi(t_k) = \prod \frac{t_k - \tau_n}{t_k - \bar{\tau}_n} \cdot \frac{t_k - \bar{\eta}_n}{t_k - \eta_n}.$$

Let $T_{k'}$ be the inverse of T_k; then, we can write $\tau_n = \tau_{kk'n}$. We have (Sec. 1, Equation 7)

$$t_k - \tau_n = \frac{t - \tau_{k'n}}{(ct + d)(c\tau_{k'n} + d)},$$

and

$$\psi(t_k) = \prod \frac{t - \tau_{k'n}}{t - \bar{\tau}_{k'n}} \cdot \frac{t - \bar{\eta}_{k'n}}{t - \eta_{k'n}} \cdot \frac{c\bar{\tau}_{k'n} + d}{c\tau_{k'n} + d} \cdot \frac{c\eta_{k'n} + d}{c\bar{\eta}_{k'n} + d}.$$

Now, the set of quantities $\tau_{k'n}$ is merely the set τ_n arranged in a different order; so we have

$$\psi(t_k) = \psi(t) \prod \frac{c\bar{\tau}_n + d}{c\tau_n + d} \cdot \frac{c\eta_n + d}{c\bar{\eta}_n + d} = H_k\psi(t), \tag{4}$$

where H_k is a constant different from zero. Since the coefficients in (3) are real, we have $|c\tau_n + d| = |c\bar{\tau}_n + d|$ and

$$|H_k| = 1. \tag{5}$$

Let $T_1, T_2, \ldots, T_m$ be the transformations connecting the sides of S_0—arranged as in the figure. Considering the cycle at O_1, we find the relation $T_m \cdots T_1 = 1$. By a repeated application of (4), we have $\psi(t) = H_m \cdots H_1\psi(t)$; whence,

$$H_1 H_2 \cdots H_m = 1. \tag{6}$$

We shall make repeated use of the elementary formula connecting the number N of zeros and P of poles of a function $f(t)$ analytic except for poles in S_0 and continuous and not vanishing on the boundary.

$$N - P = \frac{1}{2\pi i}\int_C d \log f(t) = \frac{1}{2\pi i} [\log |f(t)| + i \arg f(t)]_C$$
$$= \frac{1}{2\pi}[\arg f(t)]_C \tag{7}$$

where the boundary C of S_0 is traced in the positive sense. As the interval b_s (corresponding to B_s) is traced in this sense, let $\beta^{(s)}$ be its beginning and $\alpha^{(s)}$ its end (Fig. 80). Let h_s be the change in $\arg f(t)$ as t moves from $\beta^{(s)}$ to $\alpha^{(s)}$; and let $\psi(t)$ be substituted in (7). We have $N - P = 0$, since there is one zero and one pole in S_0. Also, $d \log \psi(t)]_{C_s^-} = d \log \psi(t)]_{C_s^+}$ from (4), and the integrals along two congruent sides cancel, being taken in opposite directions. We have then

$$\sum\Big[\arg \psi(t)\Big]_{\beta^{(s)}}^{\alpha^{(s)}} = h_1 + h_2 + \cdots + h_m = 0. \tag{8}$$

We observe that $H_s = e^{ih_s}$; so that (6) is a consequence of (8).

We next set up another type of function

$$\varphi_s(t) = \prod \frac{t - \alpha_n^{(s)}}{t - \beta_n^{(s)}}, \quad s = 1, 2, \cdots, m - 1. \tag{9}$$

The point $\beta_n^{(s)}$ is not a pole; for, since $\alpha^{(s)}$ and $\beta^{(s)}$ are congruent points, there is a factor in the numerator which cancels $t - \beta_n^{(s)}$. The function is analytic in the whole plane exclusive of the limit points of the group. It is nowhere zero.

At the ordinary points of the real axis $\varphi_s(t)$ is real, all factors being real. Each factor in (9) is then positive, with the possible exception of one factor. If t lies in the interval $\beta_j^{(s)}$, $\alpha_j^{(s)}$, $t - \alpha_j^{(s)}$ and $t - \beta_j^{(s)}$ differ in sign and the factor is negative. We have, then,

$$\varphi_s(t)\Big]_{b_s} < 0, \qquad \varphi_s(t)\Big]_{b_n} > 0, n \neq s. \tag{10}$$

It follows from this property that $\varphi_s(t)$ is not a constant.

Let t be subjected to the transformation T_k. We have

$$\varphi_s(t_k) = \prod \frac{t_k - \alpha_n^{(s)}}{t_k - \beta_n^{(s)}} = \prod \frac{t - \alpha_{k'n}^{(s)}}{t - \beta_{k'n}^{(s)}} \cdot \frac{c\beta_{k'n}^{(s)} + d}{c\alpha_{k'n}^{(s)} + d}$$
$$= \varphi_s(t) \prod \frac{c\beta_n^{(s)} + d}{c\alpha_n^{(s)} + d} = K_k^{(s)} \varphi_s(t). \tag{11}$$

Here, each factor in the product for the constant $K_k^{(s)}$ is real. Also, since $-d/c$ is congruent to ∞ and so lies in one of the intervals congruent to b_m, we have, from (10),

$$K_k^{(s)} = \prod \frac{-\frac{d}{c} - \beta_n^{(s)}}{-\frac{d}{c} - \alpha_n^{(s)}} = \frac{1}{\varphi_s\left(-\frac{d}{c}\right)} > 0. \tag{12}$$

In the same way that (6) was established, we show that

$$K_1^{(s)} K_2^{(s)} \cdots K_m^{(s)} = 1. \tag{13}$$

The analogous function $\varphi_m(t)$ requires a slightly different definition. If we form the product in (9) with $\alpha^{(m)}$, $\beta^{(m)}$ at the extremities of b_m (b_3 in the figure) on the boundary of S_0, t lies between $\alpha^{(m)}$ and $\beta^{(m)}$ except when t is on the interval b_m on the boundary of S_0. The result is that the factor $(t - \alpha^{(m)})/(t - \beta^{(m)})$ in (9) has always the opposite sign from that appearing in the previous cases. If we define $\varphi_m(t)$ by the equation

$$\varphi_m(t) = -\prod \frac{t - \alpha_n^{(m)}}{t - \beta_n^{(m)}},$$

the properties (10) to (13) all hold.

Since $\varphi_s(t)$ has no singularities and no zeros in the upper half plane, $\log \varphi_s(t)$ is a single-valued analytic function in that domain, if we restrict ourselves to one branch of the logarithm. We have

$$\log \varphi_s(t) = \log |\varphi_s(t)| + i \arg \varphi_s(t) = \Sigma \log \frac{t - \alpha_n^{(s)}}{t - \beta_n^{(s)}}, \quad (14)$$

where we take the principal value of the logarithm in each term of the summation. We observe from Fig. 81 that the argument of each term in the summation is an angle lying between 0 and π inclusive, and that the sum of these angles does not exceed π; that is,

$$0 \leqslant \arg \varphi_s(t) \leqslant \pi. \quad (15)$$

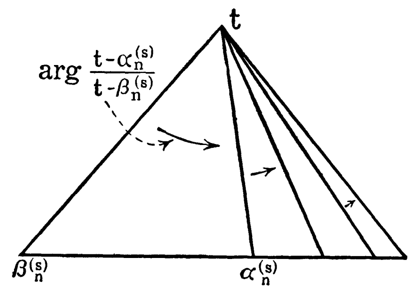


FIG. 81.

The inequalities hold when t is in the upper half plane. When t approaches a point of one of the intervals congruent to b_s, one angle approaches π and the rest approach 0; when t approaches a point of any other interval, all the angles approach zero. Hence,

$$\arg \varphi_s(t)\Big]_{b_s} = \pi, \qquad \arg \varphi_s(t)\Big]_{b_n} = 0,\ n \neq s. \quad (16)$$

(These also hold for $\varphi_m(t)$ if we take $\arg(-1) = \pi$).

At all congruent points of the domain under consideration, $\arg \varphi_s(t)$ has the same value. From (11), we have

$$\arg \varphi_s(t_k) = \arg K_k^{(s)} + \arg \varphi_s(t) = \arg \varphi_s(t), \quad (17)$$

for $K_k^{(s)}$ is real and positive and the assumption that $\arg K_k^{(s)} = 2n\pi$, $n \neq 0$, leads to a contradition of (15). In particular, if t moves from $\beta^{(s)}$ to $\alpha^{(s)}$ along b_s, $\arg \varphi_s(t)$ attains at $\alpha^{(s)}$ the same value as at $\beta^{(s)}$.

Finally, the function

$$G_s(t) = e^{a_s i \log \varphi_s(t)} = e^{a_s [i \log |\varphi_s(t)| - \arg \varphi_s(t)]}, \quad (18)$$

where a_s is a real constant, has the following properties:

$$|G_s(t)|_{b_s} = e^{-\pi a_s}, \qquad |G_s(t)|_{b_n} = 1,\ n \neq s, \quad (19)$$

$$G_s(t_k) = G_s(t) e^{i a_s \omega_k^{(s)}},\ \omega_k^{(s)} = \log K_k^{(s)}, \quad (20)$$

$\omega_k^{(s)}$ being real. The function is analytic and nowhere zero in the upper half plane.

100. The Mapping of a Multiply Connected Region of Planar Character on a Slit Region.[1]—Consider the function

$$G(t) = G_1(t)G_2(t) \cdots G_m(t)\psi(t). \tag{21}$$

From (2) and (19) we have

$$|G(t)|_{b_s} = e^{-\pi a_s}. \tag{22}$$

From (4), (5), and (20)

$$G(t_k) = G(t)e^{-i(a_1\omega_k{}^{(1)} + \cdots + a_m\omega_k{}^{(m)} + h_k)}. \tag{23}$$

The function will take on the same values at congruent points on the boundary of S_0 if $a_1, \ldots, a_m$ be chosen so that the equations

$$\begin{aligned} a_1\omega_1{}^{(1)} + \cdots + a_m\omega_1{}^{(m)} + h_1 &= 0, \\ \cdots\cdots\cdots \\ a_1\omega_m{}^{(1)} + \cdots + a_m\omega_m{}^{(m)} + h_m &= 0, \end{aligned} \tag{24}$$

are satisfied.

Can these equations be satisfied? We observe, first, that they are not independent. We have from (13), taking logarithms,

$$\omega_1{}^{(s)} + \cdots + \omega_m{}^{(s)} = 0.$$

This, together with (8), shows that the sum of the first members of (24) vanishes whatever $a_1, \ldots, a_m$ may be. The equations can be solved, provided the matrix of the coefficients is of rank $m - 1$.

Suppose the matrix of the coefficients is of rank $m - 2$ or less. Then the m equations

$$a_1\omega_k{}^{(1)} + \cdots + a_m\omega_k{}^{(m)} = 0, \quad k = 1, \cdots, m, \tag{25}$$

can be solved by taking at least two of the constants arbitrarily and determining the remainder. Let the arbitrary a's be chosen unequal. Then the function

$$H(t) = G_1(t) \cdots G_m(t) \tag{26}$$

has the same values at congruent points on the boundary of S_0. Also, $|H(t)| = e^{-\pi a_s}$ on b_s. Since not all the a's are equal, $H(t)$ is not a constant. We shall show that a function with these properties is impossible.

Let $H(t_0) = z_0$ be a value taken on by $H(t)$ at an inner point t_0 which is not taken on the boundary; and let $H(t) - z_0$ be substituted in (7). We have $P = 0$ and $N \geqslant 1$. Since $H(t)$ takes on the same values at congruent points of the boundary, the

[1] KOEBE, P., *Acta Math.*, vol. 41, pp. 305–344, 1918.

integrals along the congruent sides cancel. As t moves from $\beta^{(s)}$ to $\alpha^{(s)}$, $z = H(t)$ moves on the circle $|z| = e^{-\pi a_s}$ and returns to its starting point. There is no change in the argument; for

$$\arg H(t) = a_1 \log |\varphi_1(t)| + \cdots + a_m \log |\varphi_m(t)|,$$

and we see from (11) that arg $H(t)$ is increased by

$$a_1 \log K_s{}^{(1)} + \cdots + a_m \log K_s{}^{(m)} = a_1\omega_s{}^{(1)} + \cdots + a_m\omega_s{}^{(m)},$$

which is zero by (25). Thus z moves about the circle and returns to its starting point without completing a revolution. Then $\arg [H(t) - z_0]$, which is the angle between the line segment joining z_0 to z and the positive x-axis, suffers no alteration. We have, then, from (7), that $N = 0$, which is impossible.

The assumption that equations (24) are inconsistent has led to a contradiction. It follows that the equations are consistent. Let $a_1, \ldots, a_m$ be constants satisfying the equations. Then, we have, from (23),

$$G(t_s) = G(t), \quad s = 1, \cdots, m. \tag{27}$$

We have, also,

$$\arg G(t) = a_1 \log |\varphi_1(t)| + \cdots + a_m \log |\varphi_m(t)| + \arg \psi(t);$$

and, as t moves from $\beta^{(s)}$ to $\alpha^{(s)}$, the argument is increased by

$$a_1\omega_s{}^{(1)} + \cdots + a_m\omega_s{}^{(m)} + h_s,$$

which is zero, from (24).

We now consider the mapping of the region S_0 by the function

$$z = G(t). \tag{28}$$

As t traces the boundary of S_0, z traces a path of the character shown in Fig. 82. As t moves from O_1 along C_1^- to $\beta^{(1)}$, z traces a curve C_1' from O', the map of O_1, to a point on the circle $|z| = e^{-\pi a_1}$. As t moves from $\beta^{(1)}$ to $\alpha^{(1)}$, z moves on the circle and returns to its initial position without change of argument. As t moves from $\alpha^{(1)}$ along C_1^+ to O_2, z retraces the curve C_1' (as a consequence of (27)) in the opposite direction to O'. In a similar manner as t continues from O_2 to O_3, z moves along a curve C_2', moves along an arc of $|z| = e^{-\pi a_2}$, and retraces C_2' to O'; and so on.

Let z_0 be a point not lying on the map of the boundary in the z-plane; and let $G(t) - z_0$ be substituted in Formula (7). We have $P = 1$, since the function has a pole of the first order at η. As t moves around the boundary of S_0, $G(t)$, or z, moves around the curve in Fig. 82. Arg $[G(t) - z_0]$, which is the angle between a segment joining the moving point z to z_0 and the positive x-axis, obviously returns to its original value. We have then, from (7), $N - 1 = 0$, or $N = 1$. The value z_0 is taken on once, and only once, in S_0.

If z_0 lies on one of the curves C_s' so that z_0 is taken on at a point t_0 of C_s^{-1}, we deform C_s^{-1} to reduce t_0 to an inner point and make the corresponding deformation in C_s^{+}. Then, z_0 no longer lies on the boundary in the z-plane and the preceding analysis holds. Each value z_0, not lying on an arc b_s', is taken on once, and only once, in S_0, if we count only one of two congruent sides as belonging to the region. Similarly, we count but one of the points $O_1, \ldots, O_m$ as belonging to S_0.

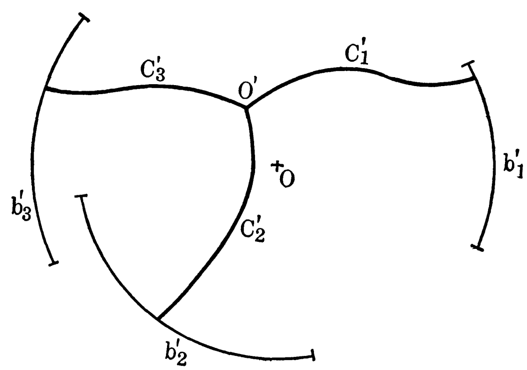


FIG. 82.

Now, let z_0 lie on an arc b_s'; whence, z_0 is taken on at a point t_0 of b_s. Then, $z = \dot{G}(t)$ maps the neighborhood of t_0 in the upper half plane on the neighborhood of z_0 on the interior or exterior of b_s', according as z is moving counter-clockwise or clockwise along b_s' as t moves in a positive sense through t_0. It follows, from this, that z cannot pass twice through z_0 in the same sense; otherwise, a value in the neighborhood of z_0 and not lying on b_s' would be taken on twice in S_0, which is impossible.

If z_0 is at an end of the arc b_s', the neighborhood of t_0 on the upper side of the axis is mapped on the whole neighborhood of z_0, exclusive of the arc b_s'. The arc b_s' cannot consist of the whole circle; otherwise, the two ends would coincide and a value in the neighborhood would be taken on twice in S_0. Finally, if z_0 lies on b_s', the value z_0 is taken on at no point t_1 of S_0 not lying on b_s. For the function maps the neighborhood of t_1 conformally on the neighborhood of z_0; and a nearby value is taken on twice in the region.

It is clear that the arc b_s' cannot consist of a single point. Otherwise, $G(t)$ is constant on b_s, and, hence, is identically constant, which is impossible.

We now carry the function back to the surface ϕ by means of the mapping function $t = g(Z)$:

$$z = G(t) = G[g(Z)] = f(Z). \tag{29}$$

The function $f(Z)$ takes on the same values at points of C_s^- and C_s^+ on the boundary of ϕ_0 on opposite banks of the cut C_s. These banks can be joined and the function remains single valued on the resulting surface; that is, on the original unsevered region Σ. We have a one-to-one correspondence between the points of Σ and the points of the z-plane which do not lie on the arcs b_s'.

We shall speak of a set of arcs, each of which lies on a circle with center at the origin and does not consist of the whole circumference, as *concentric slits.* We have proved the first part of the following theorem:

THEOREM 1.—*A region of planar character which has a finite number of sheets and of branch points and has a finite number, m, of bounding curves can be mapped conformally on a plane region bounded by m concentric slits.*

If it be required that given points P_1, P_2 of the region be carried to 0 and ∞, respectively, in the plane, then the mapping is determined save for a transformation of the plane of the form $z' = cz$.

We now prove the latter part of the theorem. If, for τ and η in (1), we select the points of S_0 which correspond to P_1 and P_2, then the function (29) performs the mapping in the required manner. Let $z' = f_1(Z)$ be any other such mapping function. On mapping in the t-plane, this function is carried into a function of t, $z' = f_1[Z(t)] \equiv G_1(t)$, which has a pole at η and a zero at τ. Elsewhere in S_0 it has no poles or zeros; it takes on the same values at congruent points of the sides of S_0; and its absolute value is constant along each interval b_s. The quotient $G_1(t)/G(t)$ has neither poles nor zeros in S_0. Suppose this quotient is not a constant. Then, it has all the properties found for $H(t)$ in Equation (26); properties which we found to be inconsistent. It follows that $G_1(t)/G(t) \equiv c$, a constant; hence, $z' = cz$. Conversely, it is obvious that if we apply any such transformation to one plane map bounded by concentric slits, the points 0 and ∞ which correspond to P_1 and P_2 remain fixed and the concentric slits are carried into concentric slits.[1]

101. Application to the Uniformization of Algebraic Functions.[2] We turn now to the uniformization of algebraic functions through the use of the mapping theorem just established. Let F be the Riemann surface of an algebraic function $W = F(Z)$ of genus $p > 0$. We show, first, that we can draw p, and not more than p, loop-cuts in F without separating the surface.

In the construction of the cuts explained at the end of Sec. 89, let each of the cuts $a_2, \ldots, a_p$ be deformed into a sigma cross-cut beginning at P before the corresponding b-cut is made. The b-cut joins opposite banks of the loop of the sigma cross-cut. The cuts $b_1, \ldots, b_p$ have no common points; and, if we erase the a-cuts entirely, we have p loop-cuts which do not separate the surface. The surface with the p loop-cuts has $2p$ boundaries, namely, the two banks of each loop-cut. (Fig. 83 shows such loop-cuts for the surface of Fig. 72.)

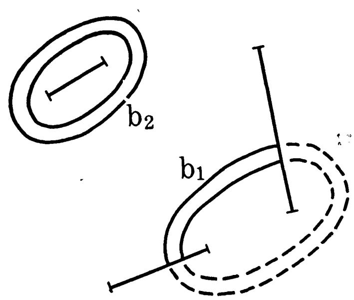


FIG. 83.

Suppose that p' loop-cuts are made in F without separating the surface. Let O be an initial boundary point and let $2p'$ cuts be drawn from O to the banks of the loop-cuts. The surface remains a single piece, for it has a single boundary. Since not more than $2p$ cuts can be made in F without separating the surface—loop-cuts which are subsequently joined to the boundary not being counted—it follows that $p' \leqslant p$. In

[1] In his article in *Acta Math.*, vol. 41, pp. 305–344, KOEBE has considered the possibility and the uniqueness of the mapping of multiply connected regions on a great variety of types of slit regions—regions bounded by radial slits, parallel slits, combinations of radial and concentric slits, slits lying on logarithmic spirals, etc. In all cases the mapping is achieved by the construction of functions based on the group T_n of Fig. 80, the functions being so formed as to behave in particular ways on the boundary of S_0.

The original treatment of the problem of uniformization considered in this chapter was based on the mapping of multiply connected regions on regions bounded by parallel slits.

[2] KOEBE, P., *Math. Ann.*, vol. 69, pp. 1–81, 1910. OSGOOD, W. F., *Annals of Math.* (2), vol. 14, pp. 143*ff*, 1913.

particular, if p loop-cuts have been made without separating the surface, then any additional loop-cut separates it; that is, the surface with the p loop-cuts is of planar character.

Let F_0 be the severed surface after p loop-cuts which do not separate F have been made. We may suppose that the cuts are of elementary character—composed of a finite number of analytic arcs, for instance. We take an infinite number of copies of F_0 and build an infinitely sheeted surface. Let ϕ_0 be the initial copy. We superpose $2p$ copies and join one along each boundary of ϕ_0, opposite banks of the same loop cut in the two copies being joined together. We call the resulting region ϕ_1. We superpose copies and join along each of the free boundaries of ϕ_1 to form ϕ_2, and so on, ϕ_{n+1} being formed from ϕ_n by adjoining copies of F_0 along the free boundaries of ϕ_n. We call the limit surface ϕ.

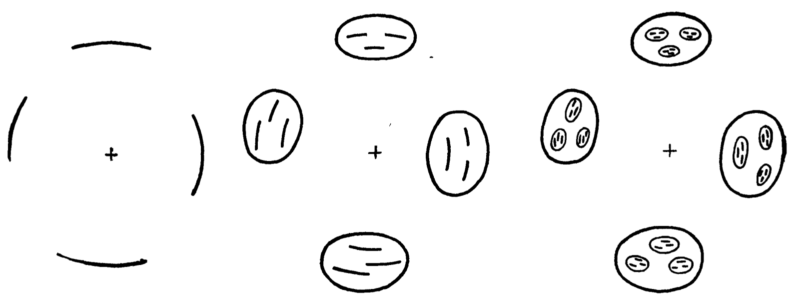


FIG. 84.

Each region ϕ_n is of planar character and satisfies the other conditions of Theorem 1; hence, it can be mapped conformally on a plane region bounded by concentric slits. Let P_1, P_2 be inner points of ϕ_0, with coordinates Z_1, Z_2, and let them be carried to zero and infinity, respectively, in the mapping. If we require, further, that the derivative of the mapping function be unity at P_1, the function is uniquely determined. We have, then,

$$z_n = f_n(Z), \quad f_n(Z_1) = 0, \quad f_n'(Z_1) = 1, \quad f_n(Z_2) = \infty, \qquad (30)$$

the conditions at Z_1 and Z_2 being valid in the sheets in which P_1 and P_2 lie. The successive maps are of the type shown in Fig. 84 (for $p = 2$). The maps of ϕ_0, ϕ_1, and ϕ_2 are shown in the figure.

102. A Convergence Theorem.—The direct proof of the convergence of the sequence $f_n(Z)$ is difficult. We shall make use of a general convergence theorem. This important theorem could have been used to simplify a number of our earlier convergence proofs.[1]

THEOREM 2.—*Given a finitely or infinitely sheeted region ϕ spread over the z-plane, each of whose inner points lies in a plane sheet or is a branch point of finite order; and a sequence of regions $\phi_1, \phi_2, \ldots$ such that ϕ_n is a subregion of ϕ_{n+1} and ϕ is the limit region of the sequence. Let*

$$f_1(z), \quad f_2(z), \quad f_3(z), \quad \cdots \tag{31}$$

be a sequence of functions such that $f_n(z)$ is analytic in ϕ_n (continuous at the branch points and at infinity) and such that in a suitable neighborhood of each inner point p of ϕ the functions of the sequence which exist there are bounded:

$$|f_n(z)| < M_p, \tag{32}$$

where M_p is independent of n.

Then there exists a subsequence of (31)

$$f_{n_1}(z), \quad f_{n_2}(z), \quad \cdots, \quad n_{k+1} > n_k, \tag{33}$$

which converges and whose limit function is analytic in the whole interior of ϕ. Further, the sequence converges uniformly in any finitely sheeted region which, together with its boundary, consists of inner points of ϕ.

Let

$$p_1, p_2, p_3, \cdots \tag{34}$$

be an infinite sequence of inner points of ϕ which are everywhere dense in ϕ. Such a sequence can be formed, for example, from all points whose z-coordinates have rational real and imaginary parts. All rational points in the plane are denumerable; so, likewise, are all the rational points in a region with a finite or a denumerably infinite number of plane sheets. Thus, the rational points of ϕ are denumerable and can be arranged in a sequence.

Consider the values of the functions (31) at p_1. This point lies in one of the regions ϕ_m and in all succeeding regions of the sequence. Then the quantities

$$f_m(p_1), f_{m+1}(p_1), \quad \cdots \tag{35}$$

[1] For a general treatment of theorems of this type see the recent volume of P. Montel, Leçons sur les familles normales de fonctions analytiques, Paris, 1927.

exist, and, from (32) are bounded. They have then at least one cluster point A_1; and we can choose a subsequence of (35)

$$f_{m_{11}}(p_1), f_{m_{12}}(p_1), f_{m_{13}}(p_1), \cdots, \quad m_{1,k+1} > m_{1,k} \tag{36}$$

which converges to the value A_1. We have now found a subsequence of (31)

$$f_{m_{11}}(z), f_{m_{12}}(z), f_{m_{13}}(z), \cdots \tag{37}$$

which converges at p_1. Any subsequence of (37) will, likewise, converge to the value A_1 at p_1.

Considering, next, the sequence (37) and the point p_2, we can repeat the preceding reasoning and get a subsequence of (37)

$$f_{m_{21}}(z), f_{m_{22}}(z), f_{m_{23}}(z), \cdots, \quad m_{2,k+1} > m_{2,k}, \tag{38}$$

whose members exist at p_2 and converge there.

We proceed in this manner *ad infinitum.* From the sequence

$$f_{m_{s1}}(z),\ f_{m_{s2}}(z),\ f_{m_{s3}}(z), \cdots, \quad m_{s,k+1} > m_{s,k}, \tag{39}$$

whose members exist and which converge at each of the points $p_1, p_2, p_3, \cdots, p_s$, we form a subsequence

$$f_{m_{s+1,1}}(z), f_{m_{s+1,2}}(z), \cdots, \quad m_{s+1,k+1} > m_{s+1,k}, \tag{40}$$

whose members exist and which converges at p_{s+1}.

We now take for $f_{n_k}(z)$ in (33) the function $f_{m_{kk}}(z)$; that is, we take for (33) the sequence

$$f_{m_{11}}(z), f_{m_{22}}(z), f_{m_{33}}(z), \cdots \tag{41}$$

formed by taking the first function of (37), the second of (38) $\cdots$, the s-th of (39), etc. Here, owing to the manner of construction, we have $m_{k+1,k+1} > m_{k,k}$. Consider any point p_s of (34). All functions of (41), from the s-th function on, exist at p_s and form a subsequence of (39). Hence, the sequence (41) converges at p_s.

The sequence (41) which we have just constructed converges at all points of the set (34), these points being everywhere dense in ϕ. We next show that it converges at all inner points of ϕ and that the limit function is analytic. Let p be a finite inner point of ϕ other than a branch point. All the functions (41), from a certain point on, are defined at p and satisfy the inequality (32) in a plane region S_p enclosing p. We omit from consideration the finite number of functions of the sequence which are not defined in S_p. Let C_1, C_2 be circles with p as center and lying in S_p, the radii being r_1, r_2 where $r_1 < r_2$. We shall show that the sequence (41) converges uniformly in C_1, which will establish, at the same time, the convergence and the analytic character of the limit function.

Let z_1, z_2 be any two points within or on the boundary of C_1. We have

$$f_n(z_1) = \frac{1}{2\pi i}\int_{C_2}\frac{f_n(t)dt}{t - z_1}, \qquad f_n(z_2) = \frac{1}{2\pi i}\int_{C_2}\frac{f_n(t)dt}{t - z_2},$$

and

$$f_n(z_2) - f_n(z_1) = \frac{z_2 - z_1}{2\pi i}\int_{C_2}\frac{f_n(t)dt}{(t - z_1)(t - z_2)}.$$

Here, since

$$|t - z_1| \geqslant r_2 - r_1 = \delta, \qquad |t - z_2| \geqslant \delta, \qquad |f_n(t)| < M_p,$$

we have

$$|f_n(z_2) - f_n(z_1)| \leqslant \frac{|z_2 - z_1| \cdot M_p \cdot 2\pi r_2}{2\pi\delta^2} = g|z_2 - z_1|, \quad (42)$$

where g is independent of z_1, z_2, and n.

Given $\epsilon > 0$. Let $p_{s_1}, \cdots, p_{s_m}$ be points of (34) in C_1 so chosen that each point in C_1 is within a distance $\epsilon/3g$ of one of the chosen points. This can be accomplished by ruling C_1 into squares of side $\epsilon/6g$ and taking one point in each square. Since the sequence (41) converges at p_{s_k}, there exists an n_k' such that, for the functions of the sequence,

$$|f_{m_{i+\nu,i+\nu}}(p_{s_k}) - f_{m_{ii}}(p_{s_k})| < \frac{\epsilon}{3}, \quad m_{ii} > n_k'. \quad (43)$$

Let n' be the greatest of the numbers n_k'; then, (43) holds for all the points p_{s_k} provided $m_{ii} \geqslant n'$. Let z be any point within or on the boundary of C_1; and let p_{s_k} be one of the m chosen points such that $|z - p_{s_k}| < \epsilon/3g$. Then, we have from (42) and (43), when $m_{ii} \geqslant n'$

$$\begin{aligned}|f_{m_{i+\nu,i+\nu}}(z) - f_{m_{ii}}(z)| &\leqslant |f_{m_{i+\nu,i+\nu}}(z) - f_{m_{i+\nu,i+\nu}}(p_{s_k})| + \\ &\quad |f_{m_{i+\nu,i+\nu}}(p_{s_k}) - f_{m_{ii}}(p_{s_k})| + |f_{m_{ii}}(p_{s_k}) - f_{m_{ii}}(z)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \quad (44)\end{aligned}$$

This inequality, which holds for all points z of C_1 and for all positive values of ν, establishes the uniform convergence of the sequence in C_1.

If p is a finite branch point of order n, the point at infinity in a single sheet, or an infinite branch point, we first map the neighborhood of p on a single-sheeted finite region by means of the functions $z - p = t^n$, $z = 1/t$, or $z = 1/t^n$. The preceding reasoning then applies.

The final statement of the theorem follows from the fact that a finitely sheeted region S lying, together with its boundary, in ϕ can be covered by a finite number of regions, such as C_1, in each of which the inequality of the type (44) holds. We have but to take $m_{ii} > N$, where N is the greatest of the numbers n' for the various regions, and (44) holds throughout S.

103. The Sequence of Mapping Functions.—We now consider the mapping functions (30):

$$f_0(Z), f_1(Z), f_2(Z), \cdots \tag{45}$$

Owing to the pole at P_2 we shall exclude P_2 from each region and from ϕ, considering it as a boundary point. Then, the set of functions (45) and the regions ϕ_0, ϕ_1, $\cdots \rightarrow \phi$, satisfy all the conditions of Theorem 2, with the possible exception of the inequality (32). We show, next, that this condition is satisfied also.

Let P be an inner point of ϕ; and let ϕ_m be the first of the regions ϕ_n containing P. Then, the function $z_m = f_m(Z)$ maps ϕ_m on a slit region S_m, P being mapped on an inner point p of S_m. Let Σ be a region which, together with its boundary, consists of inner points of S_m and which encloses p and the origin. Now, any subsequent mapping function maps ϕ_m on a plane finite region; and, if we change the variable to z_m,

$$z_n = f_n(Z) = \varphi_{n,m}(z_m), \quad n \geqslant m, \tag{46}$$

we have a function which maps S_m on a plane finite region. The regions S_m and Σ are a pair of regions to which Theorem 9, Sec. 76, applies; and, since $\varphi_{n,m}(0) = 0$, $\varphi'_{n,m}(0) = 1$, we have $|\varphi_{n,m}(z)| < L$ in Σ. On returning from the z_m-plane to the surface ϕ, we have

$$|f_n(Z)| < L, \;\; n \geqslant m, \tag{47}$$

in a region Σ_1 enclosing P, Σ_1 being the region on ϕ which was mapped on Σ in the z_m-plane. An inequality of the type (32) thus holds in the neighborhood of any inner point of ϕ.

It follows that there exists a subsequence of (45)

$$f_{n_1}(Z), f_{n_2}(Z), \cdots, n_{k+1} > n_k, \tag{48}$$

which converges throughout ϕ and whose limit function $f(Z)$ is analytic in ϕ. In any finitely sheeted subregion of ϕ which does not reach to the boundary, the convergence is uniform.

What can be said of the limit function $f(Z)$? It is not constant; for, its derivative at P_1, being the limit of the derivative of $f_{n_k}(Z)$ at P_1, is 1. It vanishes at P_1 and at no other point of ϕ. For, consider any other point P. Draw a circle with P_1 as center and in the plane sheet in which P_1 lies of radius r sufficiently small as not to contain P. Then, for each mapping function we have, from Theorem 5, Sec. 74, $|f_{n_k}(P)| \geqslant r/4$. Hence, in the limit, $|f(P)| \geqslant r/4 > 0$.

The function has a pole of the first order at P_2. Let Q be a circle with P_2 as center in the plane sheet in which P_2 lies and of radius sufficiently small as not to include P_1. The function $1/f_n(Z)$ is analytic in and on the boundary of Q and does not vanish except at P_2, where it has a zero of the first order. The sequence $1/f_{n_k}(Z)$ converges uniformly in Q. The limit function $1/f(Z)$ vanishes at P_2, and according to Hurwitz' theorem (Theorem 12, Sec. 79), it has a zero of the first order there. That is, $f(Z)$ has a pole of the first order at P_2.

Finally, $f(Z)$ takes on no value twice in ϕ. Suppose the value α to be taken on at the two points P', P''; and let ϕ_m enclose both points. We map ϕ_m on the slit region S_m by means of the function $z_m = f_m(Z)$, P', P'' being carried to p', p''. Let S_m' be the region S_m with a region about infinity excluded, the part removed being such that p', p'' are interior to S_m'. Then, the sequence $\varphi_{n_k,m}(z_m) - \alpha$ (Equation (46)) converges uniformly in S_m'. We deform the boundary of S_m' slightly, if necessary, so that the limit function $\varphi_m(z_m) - \alpha$ does not vanish on the boundary. Then, applying Hurwitz' theorem, $\varphi_m(z_m) - \alpha$ has the same number of zeros in S_m' as $\varphi_{n_k,m}(z_m) - \alpha$, for k sufficiently large. But, this latter function either has one zero or none in S_m'; hence, $\varphi_m(z_m)$ takes on the value α not more than once in S_m'. This contradicts the hypothesis that $\varphi_m(z_m)$—the function $f(Z)$ with the variable changed—takes on the value α at p' and p''.

It follows from the preceding paragraph that the function

$$z = f(Z) \tag{49}$$

maps the infinitely sheeted region ϕ on a plane region V in the z-plane. If we restore P_2 to the status of an interior point—

the point which is carried to infinity—the boundary of V lies in a finite region.

The function (49) maps ϕ_0 on a region V_0 having zero and infinity as interior points and bounded by $2p$ curves (Fig. 85). The copies of F_0 which adjoin ϕ_0 are mapped on regions adjoining V_0, each having $2p$ boundaries. The infinite number of copies of F_0 which form ϕ are mapped on an infinite number of regions which fit together to make up V.

The surface ϕ possesses an infinite group of conformal transformations into itself; namely, the transformations got by carrying ϕ_0 into each of the copies F_0 of which ϕ is constructed. When ϕ is so transformed, the corresponding points in the z-plane undergo conformal and, hence, analytic transformations $z' = T_n(z)$, which carry V into itself. V_0 is a fundamental region for the group T_n. We shall represent by V_n the transform of V_0 by T_n. Then, the regions V_n, the maps of the copies F_0, fit together to form the region V.

The inverse of (49), $Z = Z(z)$, and the function $W = F[Z(z)] = W(z)$ are unaltered when z is subjected to the transformation T_n. For, the points on ϕ which correspond to z and $T_n(z)$ are similarly situated points on two copies of F_0 and bear the same values of Z and W.

104. The Linearity of T_n.—Hitherto, we have always inferred the linearity of a function from the fact that it maps the whole plane on itself, or the plane bounded by a single point on a plane bounded by a point, or a circle on a circle. Here, the matter is not so simple; and we shall be obliged to make a preliminary study of the regions V_n.

We show first that the series

$$\Sigma|T_n'(0)|^2 \tag{50}$$

converges. Let C be a circle with center at the origin and lying in V_0. Let r be its radius. Consider the map C_n of C by the function $T_n(z)$. It follows from Theorem 5, Corollary, Sec. 74, that C_n contains, on its interior, a circle of radius $\frac{1}{4}r|T_n'(0)|$. Hence, for the area A_n of C_n, we have

$$A_n \geqslant \frac{\pi}{16}r^2|T_n'(0)|^2.$$

But the regions C_n lie in a finite region and are non-overlapping; hence, ΣA_n is finite. It follows that (50) converges.

Consider the boundaries of the regions V_n—the infinite number of closed curves in Fig. 85. Let l_1, l_2, . . . be the lengths of these curves, arranged in some convenient order. We shall prove that

$$\Sigma l_k^2 \tag{51}$$

converges. Let C' be a circle with center at the origin and enclosing all the bounding curves of V_0; and let V_0' be that part of V_0 lying in C'. V_0' can be embedded in a larger region all of whose maps by $T_n(z)$ are plane and finite; for example, the region V with the points congruent to infinity removed. It follows from the deformation theorem (Theorem 8, Sec. 76) that there exists a constant M independent of n such that (taking $z_1 = z$, $z_2 = 0$)

$$|T_n'(z)| < M|T_n'(0)|,$$

where z is any point within or on the boundary of V_0'.

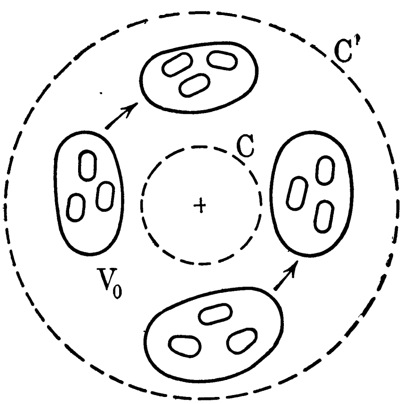


FIG. 85.

Let l' be the length of one of the bounding curves of V_0, and let l_n' be the length of its transform by means of T_n.

We have

$$l_n' = \int_{l'} |T_n'(z)||dz| < M|T_n'(0)|l',$$

and

$$l_n'^2 < M^2 l'^2 |T_n'(0)|^2;$$

whence

$$\Sigma l_n'^2 \tag{52}$$

converges as a consequence of (50). Now, (52) contains the lengths of all curves congruent to l'. V_0 is bounded by p non-congruent curves. The series (51) is the sum of p convergent series of the type (52), and, hence, converges.

We shall base the proof of the linearity of T_n on the following theorem:[1]

THEOREM 3.—*Given a finite region S whose boundary Γ consists of a finite number of regular curves. Let $f(z)$ be a function analytic in S with the exception of a set of points Σ, and continuous on Γ.*

[1] P. Koebe, *Math. Ann.*, vol. 69, p. 29, 1910.

Let it be possible to enclose the points of Σ in a finite number of regular closed curves lying in S, such that the sum of the squares of the lengths of the curves and also the sum of the squares of the oscillations of $f(z)$ on the curves is less than a preassigned positive quantity, however small. Then $f(z)$ is analytic throughout the whole interior of S, if properly defined at the points of Σ.

By the oscillation of $f(z)$ on a curve is meant the maximum value of $|f(z_1) - f(z_2)|$ where z_1 and z_2 are points on the curve.

Let z be an interior point of S at which $f(z)$ is, by hypothesis, analytic. Then there is no point of Σ within a certain distance d of z. Given $\epsilon > 0$. Let the points of Σ be enclosed in curves $C_1, C_2, \ldots, C_m$, lying in S and not containing z, of lengths $\lambda_1, \ldots, \lambda_m$, on which $f(z)$ has the oscillations $\Delta_1, \ldots, \Delta_m$, such that

$$\lambda_1^2 + \cdots + \lambda_m^2 < \pi d\epsilon, \quad \Delta_1^2 + \cdots + \Delta_m^2 < \pi d\epsilon. \tag{53}$$

We suppose the curves taken small enough—which is clearly possible—that C_n contains no part of Γ and contains no point within a distance $d/2$ of z. We have, from Cauchy's formula,

$$f(z) = \frac{1}{2\pi i}\int_\Gamma \frac{f(t)dt}{t - z} + \frac{1}{2\pi i}\sum_1^m \int_{C_k} \frac{f(t)dt}{t - z}, \tag{54}$$

the integral being taken in the positive sense around the boundary. The first member of (54) and the first term of the second member are independent of the ϵ chosen; so, also, then is the remainder of the second member.

Let ξ be a point on C_k. We have

$$\int_{C_k} \frac{f(t)dt}{t - z} = f(\xi)\int_{C_k} \frac{dt}{t - z} + \int_{C_k} \frac{f(t) - f(\xi)}{t - z}dt.$$

The first integral of the second member vanishes, since z is outside C_k; and we have, since $|f(t) - f(\xi)| \leqslant \Delta_k$, and $|t - z| \geqslant d/2$,

$$\left|\int_{C_k} \frac{f(t)dt}{t - z}\right| \leqslant \frac{2\Delta_k}{d}\int_{C_k} |dt| = \frac{2\Delta_k\lambda_k}{d} \leqslant \frac{\Delta_k^2 + \lambda_k^2}{d},$$

the final inequality being got from the algebraic inequality $2AB \leqslant A^2 + B^2$. Then, we have, applying (53),

$$\left|\frac{1}{2\pi i}\sum_1^m \int_{C_k} \frac{f(t)dt}{t - z}\right| \leqslant \frac{\Sigma\lambda_k^2 + \Sigma\Delta_k^2}{2\pi d} < \epsilon.$$

Since the first member of this inequality is less than ϵ, where ϵ may be chosen arbitrarily small, and at the same time is independent of ϵ, it must be zero.

We have, then, from (54)

$$f(z) = \frac{1}{2\pi i}\int_{\Gamma} \frac{f(t)dt}{t - z}. \tag{55}$$

The second member of (55) is analytic throughout the whole interior of S. It furnishes a formula for the analytic continuation of $f(z)$ throughout the interior of S; and the theorem is established.

We now consider $T_n(z)$. This function is analytic in V except at a single point p which is carried to infinity by the transformation; at this point, the function has a pole of the first order. Let S be the region lying within C', previously constructed, and exterior to a circle C'' drawn about p in the region $V_{n'}$ in which p lies. Let Σ be the set of all points of the plane which are not interior points of V. Then $T_n(z)$ is analytic within and on the boundary of S except at the points of Σ.

If we remove a finite number of curves l_n of Fig. 85 (we use l_n indifferently for the curve and for its length) we can enclose the points of Σ in a finite number of the curves that remain. $T_n(z)$ maps l_n on another curve $l_{n'}$ of the set. Hence, the oscillation of $T_n(z)$ on l_n is equal to the maximum distance between two points lying on $l_{n'}$, which is less than the length of $l_{n'}$. Let $\epsilon > 0$ be given. Then, by removing a sufficiently great, but finite, number q of terms at the beginning of the sequence $l_1, l_2, \ldots$, we can accomplish the following: (1) remove all curves containing p; (2) remove a sufficient number that those that remain satisfy the condition

$$\sum_{k=9}^{\infty} l_k{}^2 < \epsilon$$

(This is possible owing to the convergence of (51)); and (3) remove a sufficient number that the transforms by T_n of those that remain satisfy the condition

$$\sum l_{k'}{}^2 < \epsilon.$$

Now, let $l_{k_1}, l_{k_2}, \ldots, l_{k_m}$ be a finite number of the remaining curves which enclose the points of Σ. It follows, from the two inequalities just established, that

$$l_{k_1}{}^2 + \cdots + l_{k_m}{}^2 < \epsilon,$$
$$\Delta_{k_1}{}^2 + \cdots + \Delta_{k_m}{}^2 < l_{k'_1}{}^2 + \cdots + l_{k'_m}{}^2 < \epsilon.$$

The conditions of Theorem 3 are, thus, satisfied. Hence, $T_n(z)$ is analytic throughout the whole interior of S.

The function $T_n(z)$ is analytic in the whole z-plane except at the point p, where it has a pole of the first order. It follows that it is a linear function.

The group T_n is a group of linear transformations. The uniformizing functions $Z(z)$ and $W(z)$ are automorphic. Each has a finite number of poles in the fundamental region V_0; namely, at the points of V_0 which correspond to the points of ϕ_0 where $Z = \infty$ or $W = \infty$.

The group is a group of Schottky type (Sec. 25). We have proved the first part of the following theorem:

THEOREM 4.—*An algebraic function of genus $p > 0$ can be uniformized by means of automorphic functions belonging to a group of Schottky type in such a manner that in a sufficiently small neighborhood of any point a in the domain of definition of the uniformizing functions the correspondence between the points of the plane and the points of the Riemann surface is one-to-one.*

Any other such uniformizing functions can be got by subjecting the uniformizing variable to a linear transformation.

Suppose we have a second pair of uniformizing functions $Z = Z_1(t)$, $W = W_1(t)$ belonging to a group of Schottky type. Let z_0 lie in V, let $Z_0 = Z(z_0)$ and let $t_0, t_1, \ldots$ be the values of t for which $Z_1(t_i) = Z_0$. From the two equations $Z = Z(z)$ and $Z = Z_1(t)$ we can solve for t as a function of z, $t = \varphi_0(z)$, $t = \varphi_1(z)$, $\ldots$, where $t_i = \varphi_i(z_0)$. Any one of these functions, as $t = \varphi_0(z)$, can be extended analytically and is single valued throughout the whole of V. Similarly, its inverse is single valued throughout V', the domain of existence of the function $Z_1(t)$. Then $t = \varphi_0(z)$ maps V on V'.

By considering the new Schottky group we find readily that $\Sigma l_k'^2$ converges, where l_i' is the map in V' of l_i in V. The linearity of $\varphi_0(z)$ is then established exactly as we proved the linearity of $T_n(z)$.

The limit points of the group—the boundary points of V—form a *discrete set.* A closed set of points is called "discrete" if, in any neighborhood of any point of the set, a curve can be drawn which encloses the point and passes through no point of the set. Such a curve lying in a given neighborhood of a boundary point of V can be chosen from among the curves l_n.

We observe that if $p = 1$, the set of limit points of the group reduces to two points.

105. An Extension.—We can get other automorphic uniformizing functions in a simple manner by altering the one-to-one character of the mapping at certain points of the surface. In the severed surface F_0 let us insert a finite number of systems $L_1, \ldots,$

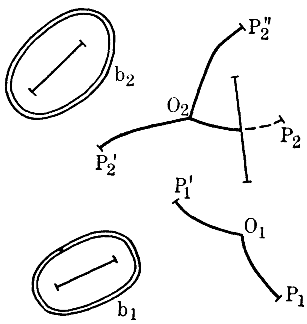


FIG. 86.

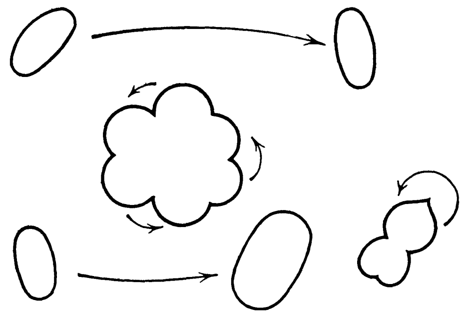


FIG. 87.

L_m of cuts of one of the following types: Either L_i shall consist of a pair of cuts from a point O_i to two points P_i, P_i', or L_i shall consist of three cuts from O_i to P_i, P_i', P_i'', the systems lying within F_0 and not meeting one another (Fig 86). In the former case, we associate with P_i, P_i' two integers ν_i, ν_i', where $\nu_i' = \nu_i$; and, in the latter, we associate with P_i, P_i', P_i'' three integers ν_i, ν_i', ν_i'' where

$$\frac{1}{\nu_i} + \frac{1}{\nu_i'} + \frac{1}{\nu_i''} > 1.$$

Let F_0' designate the resulting surface.

Taking an initial copy of F_0', we add copies of F_0' around its sides and proceed to build up a limit surface ϕ in the usual manner. In the usual way, we close up the surface about P_i, etc., when ν_i copies meet there. We then map ϕ on a region V of the z-plane as in the preceding section. V has a group of analytic transformations $z' = T_n(z)$ into itself.

That this group in the z-plane consists of linear transformations is established without essentially altering the preceding proof. The convergence of

$$\Sigma l_k^2 + \Sigma L_k'^2, \tag{56}$$

where the second summation includes all maps of the cuts L_i, is established in the same way as the convergence of (51). Now, by the addition of copies along a cut L_i in any copy, the boundary closes up after the addition of a finite number of copies, as we found in Sec. 95. We can add copies until the summation (56) extended over the maps of the remaining boundaries is arbitrarily small. We can then establish the linearity of T_n as in the preceding section.

The map of the initial copy F_0', which is a fundamental region for the group T_n, has the character shown in Fig. 87. The group is a combination group whose component groups are the previous group of Schottky type and the m finite groups corresponding to the m systems of cuts $L_1, \ldots, L_m$. In fact, the Schottky group is itself a combination group formed from p cyclic groups.

There arises the general question of the uniformization of algebraic functions by means of functions automorphic with respect to combination groups. Let the Riemann surface be severed by a preassigned finite number of separate systems of cuts into a region of planar character, and let copies be adjoined to form the limit surface ϕ. ϕ can be mapped on a plane region V, which has a group of analytic transformations $z' = T_n(z)$ into itself. This group results from combining the groups arising from each system of cuts. But, in general, the transformations are not linear. There remains the question whether the mapping of ϕ can be done in such a way that the transformations are linear. Such is, in fact, the case. When it is topologically possible to construct the surface ϕ, the mapping can be done in such a way that the transformations of the group are linear and the uniformizing functions are then automorphic. For the treatment of this problem, whose length prevents its inclusion here, the reader is referred to the papers of Koebe listed in the Bibliography.[1]

106. The Mapping of a Multiply Connected Region of Planar Character on a Region Bounded by Complete Circles.—We shall close the chapter with a further theorem on the conformal mapping of multiply connected regions.

THEOREM 5.—*A region of planar character which has a finite number of sheets and of branch points and has a finite number, m,*

[1] In particular, "Über die Uniformisierung der algebraischen Kurven III," *Math. Ann.*, vol. 72, pp. 437–516, 1912.

of bounding curves can be mapped conformally on a plane region bounded by m complete circles.

The latter region is uniquely determined except for a linear transformation of the plane.

We first map the given region on a region ϕ_0 bounded by concentric slits with center at the origin in the Z-plane and then map ϕ_0 on a region of the desired character. We assume that $m > 1$, the theorem having been proved for $m = 1$.

We first build up an infinitely sheeted limit surface containing ϕ_0. Let b_s' be a slit bounding ϕ_0. Let ϕ_0 be inverted in b_s' and let the inverse be superposed on ϕ_0 and joined along the slit b_s', opposite banks being brought together. Let ϕ_1 be the surface resulting from the like addition of sheets along all the slits of ϕ_0.

ϕ_1 is bounded by concentric slits lying in the adjoined sheets. Let b_s'' be one of these slits lying in a sheet F_i. Let F_i be inverted in b_s'' and the inverse be joined to F_i along the slit b_s''. Let ϕ_2 be the surface resulting from the like addition of sheets along all the slits bounding ϕ_1. We continue this process of inversion and junction *ad infinitum;* and we consider the limit surface of the sequence

$$\phi_0, \phi_1, \phi_2, \cdot \cdot \cdot \rightarrow \phi.$$

Each region ϕ_n can be mapped on a slit region in the z-plane. Let $z_n = f_n(Z)$ map ϕ_n so that 0 and ∞ in ϕ_0 go into 0 and ∞, respectively, and $f_n'(0) = 1$ in the first sheet ϕ_0. The condition (32) is fulfilled; for, we can derive the formula (47) without alteration in the reasoning. We have, from Theorem 2, that we can select from $f_n(Z)$ a convergent sequence

$$f_{n_1}(Z), f_{n_2}(Z), \cdot \cdot \cdot \rightarrow f(Z).$$

We find, readily, by the usual application of Hurwitz' theorem, that

$$z = f(Z)$$

maps ϕ on a plane region V in the z-plane.

An inversion in any slit carries ϕ into itself. ϕ has a group of transformations into itself, alternately inversely and directly conformal, got by carrying ϕ_0 by sequences of inversions into the various sheets which were put together to form ϕ. There corresponds in the z-plane a group of inversely and directly conformal transformations of V into itself. These transformations carry V_0, the map of ϕ_0, into the maps V_n of the various sheets of ϕ.

The regions V_n composing V are, alternately, directly conformal and inversely conformal maps of V_0. The former arise from a set of analytic transformations $z_n = T_n(z)$. We can show, repeating word for word the reasoning used in the derivation of (50), that $\Sigma|T_n'(0)|^2$ converges.

Let $l_1, l_2, \ldots$ be the curves bounding the various regions V_n. These curves are the maps of the slits on ϕ. Each separates a directly conformal and an inversely conformal map of V_0. They are, thus, all congruent to the boundaries of V_0 by the subgroup T_n. We can then show, exactly as the convergence of (51) was derived from (50), that the series

$$\Sigma l_k^2 \tag{57}$$

converges.

We now consider the transformation carrying V_0 into an adjacent region V_i abutting along the curve l_i. The transformation may be written

$$z' = U(t), \qquad t = \bar{z},$$

where U is an analytic function of t and $\bar{z}$ is the conjugate imaginary of z. We shall indicate by bars the reflections of regions, lines, etc., in the real axis. Then, U maps $\bar{V}$ in the t-plane on V in the z'-plane. The function is analytic in $\bar{V}$ except at a point p which is carried to infinity. Let S be a region lying within a regular curve so drawn in $\bar{V}_0$ as to enclose all other regions $\bar{V}_n$ and from which a small circle about p has been removed (the piece removed lying entirely in the region $\bar{V}_n$, in which p lies). Let Σ be the set of all points of the plane not interior to V. Then $U(t)$ is analytic within and on the boundary of S except at the points of $\bar{\Sigma}$.

Given $\epsilon > 0$. We can remove from the sequence $l_1, l_2, \ldots$ a finite number of curves so that of those that remain (1) no curve $\bar{l}_k$ encloses p; (2) the condition $\Sigma\bar{l}_k^2 < \epsilon$ is satisfied; and (3) the condition $\Sigma l_{k'}^{2} < \epsilon$ is satisfied, where l_k' is the curve into which U carries $\bar{l}_k$. We can now choose a finite number of these curves $\bar{l}_{k_1}, \ldots, \bar{l}_{k_n}$ enclosing the points of $\bar{\Sigma}$. We have, then,

$$\bar{l}_{k1}^2 + \cdots + \bar{l}_{k_n}^2 < \epsilon,$$

$$\Delta_{k_1}^2 + \cdots + \Delta_{k_n}^2 < l_{k_1}'^2 + \cdots + l_{k_n}'^2 < \epsilon.$$

The conditions of Theorem 3 are satisfied and $U(t)$ is analytic in S, if properly defined at the points of $\bar{\Sigma}$.

The function $U(t)$ is analytic in the whole t-plane except for a pole of the first order at p. It is, therefore, a linear function,

$$z' = \frac{a\bar{z} + b}{c\bar{z} + d}.$$

Consider, now, the curve l_i whose points are unaltered by the inversely conformal transformation $U(\bar{z})$ Let Q be a circle through three points a, b, c of l_i. An inversion in Q followed by U is a linear transformation. This transformation leaves a, b, c unaltered, and so is the identical transformation. The two transformations are inverse transformations; hence, U is equivalent to an inversion in Q. The fixed points are then on the circle Q. Each bounding curve l_i is, thus, a circle, which was to be proved.

The latter part of the theorem is an immediate consequence of the following theorem:

THEOREM 6.—*If a plane region bounded by a finite number of complete circles be mapped conformally on a second such region, the mapping function is linear.*

Let V_0, bounded by the circles $l_1, \ldots, l_m$, be mapped on V_0' bounded by the circles $l_1', \ldots, l_m'$. The mapping is continuous on the boundary and we shall so designate the circles that l_i' corresponds to l_i.

We may suppose—making preliminary linear transformations if necessary—that both V_0 and V_0' have both the origin and the point at infinity as interior points. The mapping function

$$z' = f(z)$$

is analytic in V_0, except for a pole of the first order at the point p which is carried to infinity.

Let V_i, V_i' be the regions got by inverting V_0, V_0' in the circles l_i, l_i', respectively. Then (Theorem 19, Sec. 83) $f(z)$ can be continued analytically across l_i and throughout V_i and maps V_i on V_i'. We extend thus across each circle bounding V_0. The region got by these inversions is bounded by circles and is mapped on a region bounded by circles. We can extend the function analytically across the new circles; and so on *ad infinitum*. The limit region V in the z-plane is mapped by the function on the limit region V' in the z'-plane.

The region V is invariant under the group of linear transformations T_n got by making even numbers of the preceding inversions in the circles of the z-plane. We find here, in the usual way, that $\Sigma|T_n'(0)|^2$ converges and thence that Σl_k^2 converges, where $l_1, l_2, \ldots$ are the circumferences of the infinite number of circles that arise in connection with the repeated inversions.

In a similar manner, for the corresponding circles of the z'-plane, the series $\Sigma l_k'^2$ converges.

Let S be the interior of a regular curve in V_0 which encloses the bounding circles but does not contain p. Then, $f(z)$ is analytic in S except for the points Σ of the plane which do not belong to V. These points can be enclosed in a finite number of circles $l_{k_1}, \cdots, l_{k_n}$ such that

$$l_{k_1}^2 + \cdots + l_{k_n}^2 < \epsilon, \qquad l_{k_1}'^2 + \cdots + l_{k_n}'^2 < \epsilon.$$

It follows, from Theorem 3, that $f(z)$ is analytic throughout S, if properly defined at the points of Σ.

The function $f(z)$ is thus analytic except for a pole of the first order at p. It is, therefore, a linear function, which was to be proved.

CHAPTER XI

DIFFERENTIAL EQUATIONS

107. Connection with Groups of Linear Transformations.—Given a linear homogeneous differential equation of the second order

$$\frac{d^2\eta}{dw^2} + P(w)\frac{d\eta}{dw} + Q(w)\eta = 0, \tag{1}$$

and two linearly independent particular solutions,

$$\eta = \eta_1(w), \qquad \eta = \eta_2(w). \tag{2}$$

Then, the general solution is

$$\eta = A\eta_1 + B\eta_2, \tag{3}$$

where A and B are constants. The connection with the linear transformation arises in the following manner: Let $z = \eta_1/\eta_2$ be the quotient of the two solutions, and let $z' = \eta_1'/\eta_2'$ be the quotient of two other solutions

$$\eta_1' = a\eta_1 + b\eta_2, \qquad \eta_2' = c\eta_1 + d\eta_2; \tag{4}$$

then

$$z' = \frac{a\eta_1 + b\eta_2}{c\eta_1 + d\eta_2} = \frac{az + b}{cz + d}. \tag{5}$$

If the second solutions are linearly independent, that is, if their ratio is not constant, we have $ad - bc \neq 0$.

If the coefficients $P(w)$ and $Q(w)$ are analytic at a point w_0, all solutions of the equation are analytic at w_0. There is, then, one, and only one, solution $\eta(w)$ such that η and $d\eta/dw$ have preassigned values at w_0.[1] A particularly simple pair of linearly independent solutions are those for which

$$\eta_1(w_0) = 0, \frac{d\eta_1(w_0)}{dw} = 1; \qquad \eta_2(w_0) = 1, \frac{d\eta_2(w_0)}{dw} = 0. \tag{6}$$

Then, (3) is a solution which has the value B and whose derivative has the value A at w_0.

[1] Proofs of these well-known properties will be found in WHITTAKER and WATSON's "Modern Analysis," chap. X, or E. L. Ince's "Ordinary Differential Equations," Part II.

If the coefficients are analytic in a region S of the w-plane, all solutions are analytic in S, provided S is simply connected. If, however, S is not simply connected, a solution may fail to be single valued in S. On extending the solution analytically around a closed curve in S which encloses a point not belonging to S, we may return to the starting point with a different value for the solution.

We fix our attention on a pair of linearly independent solutions (2) in the neighborhood of a point w_0 at which $P(w)$ and $Q(w)$ are analytic. Let these solutions be continued analytically from w_0 around a closed curve C which encounters no singularity of $P(w)$ or $Q(w)$ and which is such that these coefficients, when continued analytically along C, return to their original values in the neighborhood of w_0. The functions $\eta_1(w)$ and $\eta_2(w)$ remain solutions of (1) as we continue analytically; and, since the coefficients are unaltered in passing around C, the new values of the solutions—call them $\eta_1'(w)$ and $\eta_2'(w)$—are solutions of (1) with the original coefficients. Hence, $\eta_1'(w)$ and $\eta_2'(w)$ are linear combinations of $\eta_1(w)$ and $\eta_2(w)$ of the form (4). The solution may, of course, remain unaltered; this is certainly the case if C can be shrunk to a point without encountering a singularity of either coefficient. The ratio of the solutions $z = \eta_1(w)/\eta_2(w)$, is subjected to the linear transformation (5), which may, in particular, be the identical transformation. We note that $ad - bc \neq 0$; for $\eta_1'(w)$ and $\eta_2'(w)$ are analytic continuations of linearly independent functions and are, therefore, linearly independent.

THEOREM 1.—*The set of linear transformations of the ratio* $z = \eta_1(w)/\eta_2(w)$, *resulting from extending the solutions analytically from* w_0 *around all possible closed curves* C, *such that* $P(w)$ *and* $Q(w)$ *return to their initial values in the neighborhood of* w_0, *constitute a group.* Let C_i, C_j be two contours of the type considered, and let $z_i = T_i(z)$, $z_j = T_j(z)$ be the transformations of z resulting from passing around these contours. Let C be the contour got by tracing first C_i and then C_j and let $z' = U(z)$ be the resulting transformation. On making the circuit C_i, z is carried into z_i; on continuing around C_j, z_i is carried into z'. Now, by this second circuit, z is carried into z_j; hence, z_i is carried into $T_i(z_j)$. So $z' = T_i[T_j(z)]$, or $U = T_iT_j$. The succession of any two transformations of the set is a transformation of the set; and the second group property is satisfied (Sec. 13). Again, let $z' = V(z)$ be the transformation which z undergoes when C_i is traced in the reverse direction. By this, z_i is carried back to z; that is, $T_i[V(z)] = z$, or $V = T^{-1}$. The inverse of any transformation of the set belongs to the set; and the first group property is satisfied.

The group may consist solely of the identical transformation. For example, if the coefficients are entire functions, the solutions are unchanged on continuing around any closed contour. Otherwise, the group may consist of a finite or an infinite number of transformations. It may or may not be discontinuous. For example, $\frac{d^2\eta}{dw^2} + \frac{1-m}{w}\frac{d\eta}{dw} = 0$ has the solutions $\eta_1(w) = w^m$, $\eta_2(w) = 1$. If the ratio $z = w^m$ be continued around a curve C which passes k times around the origin in the positive sense, we have $z' = e^{m(\log w + 2k\pi i)} = e^{2km\pi i}z$. If m is rational but not an integer, we get an elliptic cyclic group; if irrational, we get a continuous group.

If we take a different pair of independent solutions, we get another group. If z_1 is the ratio of the second pair, we have from (5), $z_1 = G(z)$, where G is linear. On passing around a curve C let $z_1' = S(z)$, $z' = T(z)$. Then z_1 is carried into $G[T(z)]$, that is, $GTG^{-1}(z_1)$; so $S = GTG^{-1}$. The group of transformations S is the transform of the group T, as explained in Sec. 15.

We get no new groups if we use a different initial point w_1. The linear relations between the various values of z_1, the quotient of two solutions at w_1, persist when we extend the functions analytically along a curve from w_1 to w_0. The group to which z is subjected is precisely one of the groups arising from a pair of solutions at w_0.

When $\eta_1(w)$, $\eta_2(w)$ are extended analytically around the curves C they undergo a set of transformations of the form (4). These transformations form a group, according to the definition of Sec. 13. This group is called by Poincaré *the group of the differential equation.*[1] This concept applies equally well to the linear homogeneous equation of any order

$$\frac{d^p\eta}{dw^p} + P_1(w)\frac{d^{p-1}\eta}{dw^{p-1}} + \cdots + P_{p-1}(w)\frac{d\eta}{dw} + P_p(w)\eta = 0.$$

Here, we have to do with p linearly independent solutions which are subjected to a group of transformations each of which is of the form

$$\eta_k' = a_{k1}\eta_1 + \cdots + a_{kp}\eta_p, \qquad k = 1, 2, \cdots, p.$$

Each differential equation gives rise to a unique group, if we do not count as distinct two groups, of which one is the transform of the other.

[1] POINCARÉ, H., "Sur les groupes des équations linéaires," *Acta Math.*, vol. 4, pp. 201–311.

108. The Inverse of the Quotient of Two Solutions.—When w makes the circuit C and the quotient $z = \eta_1(w)/\eta_2(w)$ is subjected to a linear transformation, we have the following result: Although z has been subjected to a linear transformation, w has its original value. That is, the inverse function $w = w(z)$ is unaltered by the linear transformation. The function $w(z)$ has, thus, the first important property of the automorphic function, that of invariance under a group of linear transformations. In general, however, $w(z)$ is not a single-valued function. The requirement that $w(z)$ be single valued can be put in the following geometric form: Starting with the ratio $z(w)$ in the neighborhood of a point w_0, we extend the function in all possible ways and form its Riemann surface Φ spread over the w-plane, making two sheets coincide in the usual manner, when they bear identical functions. The Riemann surface of the inverse function $w(z)$ is a surface spread over the z-plane whose points correspond in a one-to-one manner with the points of Φ. This second surface must be plane. That is, $z(w)$ shall not take on the same value at two distinct points of Φ.

THEOREM 2.—*If the coefficients of* (1) *are analytic at a point* w_0, *then the ratio* $z = \eta_1(w)/\eta_2(w)$ *of two linearly independent solutions maps the neighborhood of* w_0 *on a plane region.*

Consider, first, the two solutions which satisfy (6). We have

$$\eta_1(w) = w - w_0 + a_2(w - w_0)^2 + \cdots,$$
$$\eta_2(w) = 1 + b_2(w - w_0)^2 + \cdots;$$
$$z = w - w_0 + a_2(w - w_0)^2 + (a_3 - b_2)(w - w_0)^3 + \cdots; \quad (7)$$

and, inverting the last series,

$$w - w_0 = z - a_2z^2 + \cdots \quad (8)$$

Then, $w(z)$ is a single-valued function of z in the neighborhood of the point, which establishes the theorem for the two solutions chosen. The ratio of any other pair of solutions (4) is a linear function (5) of z which maps the plane region in the z-plane on a plane region in the z'-plane, which proves the theorem for any other pair.

THEOREM 3.—*Let the coefficients of* (1) *be analytic in the neighborhood of* $w = a$ *except at* a *itself. When the ratio of two linearly independent solutions* $z = \eta_1(w)/\eta_2(w)$ *is continued analytically around a closed curve* C *enclosing* a *but no other singularity of the* *coefficients, let z be altered by the transformation T. Then, in order that the inverse function $w(z)$ be single-valued in the neighborhood of a, it is necessary that T be an elliptic transformation of angle $2\pi/p$, where p is an integer, or a parabolic transformation.*

Suppose, first, that (5) has two fixed points and that its multiplier is K. It is clear that if the inverse function is single-valued for one pair of independent solutions, it is single-valued for any other pair. By taking a suitable pair of solutions we can have the transformation in the form

$$z' = Kz, \qquad K \neq 1. \tag{9}$$

For this purpose, we have merely to replace η_1 and η_2 by $\alpha\eta_1 + \beta\eta_2$ and $\nu\eta_1 + \delta\eta_2$, where $G = (\alpha z + \beta)/(\nu z + \delta)$ is a transformation carrying the fixed points to zero and infinity. We shall suppose that the transformation has the form (9) when we pass counter-clockwise around C.

Let S be the region enclosed by C (Fig. 88); let S' be this region exclusive of the point a; and let S'' be the region bounded by C and a line l extending from a to a point of C. The function $z(w)$ is nowhere 0 or ∞ in S'. For, the Riemann surface Φ of $z(w)$ has a branch point at a; and if $z = 0$ or $z = \infty$ at any point, we see from (9) that when we make a circuit of a we have again $z = 0$ or $z = \infty$. Then $z(w)$ takes on the same value at distinct points of Φ, which is impossible. Then, $\log z$, where we take any branch of the logarithm, is single valued in the simply connected region S''.

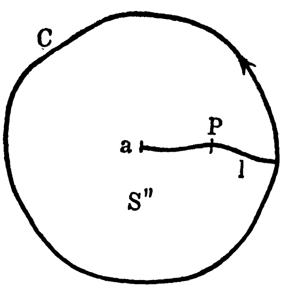


FIG. 88.

Let q be the change in $\log z$ as we pass from a point P of l around a curve in S'', moving counter-clockwise around a, to the point P again. Here, q is one of the logarithms of K and is independent of the position of P. In fact, q is the change in any branch of $\log z$ when we make one counter-clockwise circuit about a in S'.

We now consider the function[1]

$$\tau(w) = e^{\frac{2\pi i}{q}\log z} = z^{\frac{2\pi i}{q}}. \tag{10}$$

[1] KOEBE, P., *Math. Ann.*, vol. 67, p. 157.

This function takes on the same values along the opposite banks of l; and, hence, is a single-valued function of w in S'. Further, $\tau(w)$ takes on no value twice in S'. We have from (10)

$$z(w) = e^{\frac{q}{2\pi i}\log \tau} = \tau^{\frac{q}{2\pi i}}. \tag{11}$$

If $\tau(w_2) = \tau(w_1)$, where w_1 and w_2 are distinct points of S', then $\log \tau(w_2) = \log \tau(w_1) + 2n\pi i$, where n is an integer; and $z(w_2) = e^{nq}z(w_1) = K^n z(w_1)$. But this is impossible, for the value $K^n z(w_1)$ is taken on at w_1 in one of the sheets of Φ, and we have a violation of the requirement that z takes on no value twice on Φ. It follows that (10) maps S' on a plane region Σ in the τ-plane.

The function (10) maps C on a closed curve C' in the τ-plane As z moves once counter-clockwise around C, $\log z$ increases by q and $\arg \tau$ increases by 2π; that is, τ, in tracing C', moves once counter-clockwise around the origin. The region Σ lies on the left as we make this circuit, and, hence, lies in C'. The function $\tau(w)$ is analytic in the neighborhood of a and is bounded, since τ is restricted to lie in the interior of C'; hence, $\tau(w)$ is analytic at a if properly defined there.

Let τ_0 be any point lying in C'. As w passes counterclockwise around C, τ passes around C' and $\arg(\tau - \tau_0)$ increases by exactly 2π. It follows that $\tau(w)$ takes on the value τ_0 exactly once in S. S is mapped in a one-to-one manner on the interior of C'. Now, $\tau(w)$ takes on the value 0 nowhere in S'; hence, $\tau(w)$ has a zero of the first order at a. We have, then,

$$\tau(w) = c_1(w - a) + c_2(w - a)^2 + \cdots, \quad c_1 \neq 0, \tag{12}$$

and

$$w = a + \frac{1}{c_1}\tau + c_2'\tau^2 + \cdots \tag{13}$$

From (10),

$$w = a + \frac{1}{c_1}z^{2\pi i/q} + c_2'z^{4\pi i/q} + \cdots \tag{14}$$

This is a single-valued function of z only if $2\pi i/q = p$, an integer. Then,

$$K = e^q = e^{2\pi i/p}. \tag{15}$$

Hence, if T has two fixed points, it is an elliptic transformation of the type stated in the theorem. If it has only one fixed point, it is, of course, a parabolic transformation. In the former case, we note that Φ has a branch point at a where p leaves hang together. Equation (14) has the form

$$w = a + \frac{1}{c_1}z^p + c_2'z^{2p} + \cdots \tag{16}$$

The relation between w and z in the case of the parabolic transformation can be found in a similar manner. We can choose a pair of solutions whose ratio undergoes the transformation

$$z' = z + b, \quad b \neq 0, \tag{17}$$

when w makes a circuit of C. We then consider the function

$$\tau(w) = e^{\frac{2\pi i}{b}z}. \tag{18}$$

This has all the properties of the function defined in (10) and maps S in a one-to-one manner on the interior of a closed curve C' enclosing the origin in the τ-plane. The relations (12) and (13) then follow. The former may be written

$$\tau = e^{\frac{2\pi i}{b}z} = c_1(w - a) + c_2(w - a)^2 + \cdots, c_1 \neq 0. \tag{19}$$

What can be said of the coefficients of the differential equation at a? We first make a simplification of the equation. If we put

$$\eta = \zeta e^{-\frac{1}{2}\int_{w_0}^{w} P(w)dw}, \tag{20}$$

we find readily that the equation is reduced to the form

$$\frac{d^2\zeta}{dw^2} + Q_1(w)\zeta = 0, \tag{21}$$

where

$$Q_1(w) = Q(w) - \frac{1}{2}\frac{dP(w)}{dw} - \frac{1}{4}P(w)^2. \tag{22}$$

The solutions of (1) are got by multiplying the solutions of (21) by the factor $e^{-\frac{1}{2}\int P dw}$. The quotient of a pair of solutions is unaltered by the multiplication.

We now prove the following theorem:

THEOREM 4.—*Under the conditions of Theorem 3 the function $Q_1(w)$ has a pole of the second order at a.*

If we differentiate the quotient of two linearly independent solutions $z = \zeta_1(w)/\zeta_2(w)$ three times with respect to w and substitute in the expression for the Schwarzian derivative

$D(z)_w$, using the fact that ζ_1 and ζ_2 satisfy (21) to simplify the result, we find

$$\frac{1}{2}D(z)_w = \frac{2\dfrac{dz}{dw}\dfrac{d^3z}{dw^3} - 3\left(\dfrac{d^2z}{dw^2}\right)^2}{4(dz/dw)^2} = Q_1(w). \tag{23}$$

Let primes denote derivatives with respect to w. Eliminating Q_1 from

$$\zeta_1'' + Q_1\zeta_1 = 0, \quad \zeta_2'' + Q_1\zeta_2 = 0,$$

and integrating, we have

$$\zeta_2\zeta_1'' - \zeta_1\zeta_2'' = 0, \quad \zeta_2\zeta_1' - \zeta_1\zeta_2' = \Delta, \text{ a constant.}$$

Here, $\Delta \neq 0$, since the solutions are independent. Then

$$z' = \frac{\Delta}{\zeta_2^2}, \quad z'' = -\frac{2\Delta\zeta_2'}{\zeta_2^3},$$

$$z''' = -\frac{2\Delta\zeta_2''}{\zeta_2^3} + \frac{6\Delta\zeta_2'^2}{\zeta_2^4} = \frac{2\Delta Q_1}{\zeta_2^2} + \frac{6\Delta\zeta_2'^2}{\zeta_2^4}.$$

On inserting these values into $D(z)_w$ we get (23).

Suppose, first, that T is elliptic. We take a pair of solutions whose ratio $z = \zeta_1/\zeta_2$ undergoes the transformation (9), where K has the value (15). We have from (12), using (10),

$$z^p = c_1(w - a) + c_2(w - a)^2 + \cdots, \quad c_1 \neq 0,$$

whence,

$$z = (w - a)^{1/p}[c_1 + c_2(w - a) + \cdots]^{1/p} = (w - a)^{1/p}\varphi_1(w). \tag{24}$$

We use $\varphi_1(w)$, $\varphi_2(w)$, etc., to represent functions analytic at a. Here, $\varphi_1(a) \neq 0$. We now differentiate (24) and use (23) to find the behavior of $Q_1(w)$. We shall use the second expression in Equation (9), Sec. 44, for $D(z)_w$. We have

$$\frac{dz}{dw} = (w - a)^{\frac{1}{p}-1}\left[\frac{1}{p}\varphi_1(w) + (w - a)\frac{d\varphi_1(w)}{dw}\right] = (w - a)^{\frac{1}{p}-1}\varphi_2(w),$$

where $\varphi_2(a) \neq 0$.

$$\log\frac{dz}{dw} = \left(\frac{1}{p} - 1\right)\log(w - a) + \varphi_3(w),$$

$$\frac{d}{dz}\log\frac{dz}{dw} = \frac{\dfrac{1}{p} - 1}{w - a} + \varphi_4(w),$$

$$\frac{d^2}{dz^2}\log\frac{dz}{dw} = \frac{1 - \dfrac{1}{p}}{(w - a)^2} + \varphi_5(w).$$

Whence,

$$\frac{1}{2}D(z)_w = \frac{1}{2}\left[\frac{1-\frac{1}{p}}{(w-a)^2} + \varphi_5(w)\right] - \frac{1}{4}\left[\frac{\frac{1}{p}-1}{w-a} + \varphi_4(w)\right]^2;$$

or

$$Q_1(w) = \frac{1-\frac{1}{p^2}}{4(w-a)^2} + \frac{c'}{w-a} + \varphi_6(w). \tag{25}$$

$Q_1(w)$ thus has a pole of the second order.

If T is parabolic, we have, from, (19)

$$\begin{aligned}\frac{2\pi i}{b}z &= \log(w-a) + \log[c_1 + c_2(w-a) + \cdots]\\ &= \log(w-a) + \varphi_7(w).\\ \frac{2\pi i}{b}\frac{dz}{dw} &= \frac{1}{w-a} + \frac{d\varphi_7(w)}{dw} = \frac{\varphi_8(w)}{w-a}.\\ \log\frac{dz}{dw} &= -\log(w-a) + \varphi_9(w).\end{aligned}$$

On differentiating and substituting as before (or, we note that $\log dz/dw$ above reduces to the form just found if we set $p = \infty$; and so we set $p = \infty$ in (25)), we have

$$Q_1(w) = \frac{1}{4(w-a)^2} + \frac{c'}{w-a} + \varphi_{10}(w). \tag{26}$$

Here, also $Q_1(w)$ has a pole of the second order.

THEOREM 5.—*If $z(w)$ is unaltered when extended analytically around the curve C of Theorem 3 and if $w(z)$ is single-valued, then $Q_1(w)$ is analytic at a.*

The function $z(w)$ is analytic in S', with the possible exception of one point at which it is infinite. Hence, the function has, at most, an isolated singularity at a. It cannot have an essential singularity, for then $z(w)$ takes on values more than once near a, contrary to hypothesis. So $z(w)$ approaches a definite finite or infinite value at a. We can suppose, making a linear transformation on z if necessary, that $z(w)$ approaches the value zero at a; then, if defined to be zero, there the function is analytic. This zero is of the first order, since, otherwise, $w(z)$ is not single-valued; and we have

$$z = c_1(w-a) + c_2(w-a)^2 + \cdots, \quad c_1 \neq 0.$$

On substituting this value in (23), we see that $Q_1(w)$ is analytic at a.

The Point at Infinity.—The character of the solutions at infinity is studied by making the usual transformation $w = 1/t$. Making this change in the independent variable, the differential equation takes the form

$$\frac{d^2\eta}{dt^2} + \left(\frac{2}{t} - \frac{P}{t^2}\right)\frac{d\eta}{dt} + \frac{Q}{t^4}\eta = 0. \tag{27}$$

In order that the inverse of the quotient of two solutions be single-valued at infinity, this differential equation must satisfy the conditions of the previous theorems at $t = 0$.

109. Regular Singular Points of Differential Equations.—If the coefficients of (1) are analytic at a finite point a, a is called an "ordinary point" of the differential equation; otherwise, a is a "singular point." If a is a singular point and if $(w - a)P(w)$ and $(w - a)^2Q(w)$ are analytic at a, then a is called a "regular singular point." That is, $P(w)$, or $Q(w)$, or both, have poles at a, but $P(w)$ cannot have a pole of higher order than the first, nor $Q(w)$ a pole of higher order than the second. The point at infinity is called an "ordinary point" or a "regular singular point" if $t = 0$ is an ordinary point or a regular singular point of the differential equation (27).

If the inverse of the quotient of two solutions is single-valued and if the differential equation be written in the form (21), then, all isolated singular points of the differential equation in the neighborhood of which $Q_1(w)$ is single-valued, are regular. In (1) the singular points may or may not be regular. But the equation can be transformed, without altering the quotients of solutions, into one with regular singular points. If the singular points of (1) happen to be regular, we find readily that the singular points of (21) are regular.

We can find a pair of linearly independent solutions in the form of series valid in the neighborhood of a regular singular point in a fairly simple manner. The equation has the form

$$\frac{d^2\eta}{dw^2} + \frac{p(w)}{w - a}\frac{d\eta}{dw} + \frac{q(w)}{(w - a)^2}\eta = 0, \tag{28}$$

where $p(w)$ and $q(w)$ are analytic at a:

$$\begin{aligned} p(w) &= p_0 + p_1(w - a) + p_2(w - a)^2 + \cdots \\ q(w) &= q_0 + q_1(w - a) + q_2(w - a)^2 + \cdots \end{aligned} \tag{29}$$

If we assume a solution of the form

$$\eta = (w - a)^\alpha[1 + c_1(w - a) + c_2(w - a)^2 + \cdots], \tag{30}$$

differentiate it and substitute into the differential equation, which may be written

$$(w-a)^2\frac{d^2\eta}{dw^2} + (w-a)[p_0 + p_1(w-a) + \cdots]\frac{d\eta}{dw} + [q_0 + q_1(w-a) + \cdots]\eta = 0, \quad (31)$$

we have a result of the following form:

$$\begin{aligned}(w-a)^\alpha\{f(\alpha) &+ [f(\alpha+1)c_1 + \alpha p_1 + q_1](w-a) \\ &+ [f(\alpha+2)c_2 + \varphi_2(c_1)](w-a)^2 + \cdots \\ &+ [f(\alpha+n)c_n + \varphi_n(c_1, \cdots, c_{n-1})](w-a)^n \\ &+ \cdots\cdots\} = 0, \end{aligned} \quad (32)$$

where

$$f(\alpha) = \alpha(\alpha-1) + p_0\alpha + q_0,$$

and $\varphi_n(c_1, \ldots, c_{n-1})$ is a rational integral expression in $c_1, \ldots, c_{n-1}$, which need not be written out.

Equation (42) will be a formal solution if

$$f(\alpha) = \alpha^2 + (p_0 - 1)\alpha + q_0 = 0, \quad (33)$$

$$\left.\begin{aligned} f(\alpha+1)c_1 + \alpha p_1 + q_1 &= 0, \\ \cdots\cdots & \\ f(\alpha+n)c_n + \varphi_n(c_1, \ldots, c_{n-1}) &= 0, \\ \cdots\cdots & \end{aligned}\right\} \quad (34)$$

Equation (33) is known as the "indicial equation." It is satisfied by two values α', α'' which may in particular cases be equal. These roots are called the "exponents" of the differential equation at the singular point.

Taking $\alpha = \alpha'$, the first equation is satisfied. We can then determine $c_1, c_2, \ldots$ in order, unless some $f(\alpha' + n)$ chances to be zero; that is, unless $\alpha' + n$ is equal to the second root α''. This will certainly not occur if we take for α' the root whose real part has the greater value, and this we shall do. We get, in this way, the formal solution

$$\eta_1(w) = (w-a)^{\alpha'}[1 + c_1(w-a) + c_2(w-a)^2 + \cdots]. \quad (35)$$

It can be proved that the series in brackets converges uniformly in the neighborhood of a and that the function $\eta_1(w)$ is a solution of the differential equation.[1]

Proceeding in a similar manner with the second root α'', and supposing $\alpha'' + n \neq \alpha'$ for all positive integers n, we get a solution

$$\eta_2(w) = (w-a)^{\alpha''}[1 + c_1'(w-a) + c_2'(w-a)^2 + \cdots]. \quad (36)$$

We, thus, have two independent solutions except when the roots of the indicial equation are equal or differ by an integer. In the

[1] See WHITTAKER and WATSON, "Modern Analysis," 2nd ed., p. 193.

former case, the two solutions are identical; in the latter the method of forming the series breaks down, except when $\varphi_n(c_1, \ldots, c_{n-1})$ happens to vanish when $f(\alpha'' + n) = 0$.

In any case, if we know one solution, the differential equation can be solved by quadratures, so we have no difficulty in finding a second solution in these exceptional cases. Writing

$$\eta = \eta_1\zeta, \tag{37}$$

we find the following differential equation for ζ:

$$(w - a)^2\frac{d^2\zeta}{dw^2} + \left[2(w - a)^2\frac{\frac{d\eta_1}{dw}}{\eta_1} + (w - a)p(w)\right]\frac{d\zeta}{dw} = 0. \tag{38}$$

Dividing by $(w - a)^2 d\zeta/dw$ and integrating,

$$\log\frac{d\zeta}{dw} = c - 2\log\eta_1 - p_0\log(w - a) - p_1(w - a) - \cdots,$$

$$\frac{d\zeta}{dw} = c'\eta_1^{-2}(w - a)^{-p_0}e^{-p_1(w-a)} - \cdots$$

On substituting for η_1 its value from (35), we have

$$\frac{d\zeta}{dw} = c'(w - a)^{-p_0-2\alpha'}[1 + c_1(w - a) + \cdots]^{-2} \times [1 - p_1(w - a) + \cdots].$$

Let $\alpha' = \alpha'' + \sigma$; then, from (33), $\alpha' + \alpha'' = 1 - p_0$, whence, $-p_0 - 2\alpha' = -1 - \sigma$. Then,

$$\frac{d\zeta}{dw} = c'(w - a)^{-1-\sigma}[1 + k_1(w - a) + k_2(w - a)^2 + \cdots]. \tag{39}$$

On integrating this and multiplying by η_1, we have the general solution of (28).

Suppose the roots of the indicial equation are equal. Then, we have

$$\zeta = c'' + c'[\log(w - a) + k_1(w - a) + \cdots].$$

Taking $c'' = 0$, $c' = 1$ and multiplying by η_1, we have the second solution in the form

$$\eta_3(w) = \eta_1(w)\log(w - a) + (w - a)^\alpha[k_1(w - a) + k_2'(w - a)^2 + \cdots] \tag{40}$$

Suppose the roots of the indicial equation differ by an integer. Then σ is a positive integer, and we have

$$\zeta = c'' + c'\left[\frac{1}{-\sigma(w - a)^\sigma} + \cdots + k_\sigma\log(w - a) + k_{\sigma+1}(w - a) + \cdots\right].$$

Choosing the constants as before and multiplying by η_1, we have a second solution of the form

$$\eta_4(w) = k_\sigma \eta_1(w) \log (w - a) + (w - a)^{\alpha''}\left[-\frac{1}{\sigma} + k_1(w - a) + \cdots \right]. \quad (41)$$

This solution will be devoid of a logarithmic term if $k_\sigma = 0$.

Conversely, it is easily shown that if the solutions of a differential equation are of the types here considered, the equation has a regular singular point at a. The differential equation having the linearly independent solutions $\eta_1(w)$, $\eta_2(w)$ is the following:

$$\begin{vmatrix} \eta'' & \eta' & \eta \\ \eta_1'' & \eta_1' & \eta_1 \\ \eta_2'' & \eta_2' & \eta_2 \end{vmatrix} = 0.$$

On substituting for η_1 and η_2 the expressions (35) and (36) and simplifying, we get an equation of the form (28). We get a like form if we use η_1 and η_3, or η_1 and η_4.

Furthermore, it is not possible to get other solutions of the form here derived with altered values of α', α''. Thus, if the equation has the solutions (35) and (36), it is not possible to choose the constants in the general solution $\eta = A\eta_1 + B\eta_2$ so that

$$\eta = (w - a)^m[1 + l_1(w - a) + \cdots],$$

where m is different from α' and α''; and similar remarks apply to the other types of solutions. Consequently, if we have found in any way a pair of solutions of the kind given here, we know the values of the exponents.

110. The Quotient of Two Solutions at a Regular Singular Point.—The behavior of two solutions at a regular singular point depends upon the roots of the indicial equation (33). If σ $(= \alpha' - \alpha'')$ is not zero or an integer, we have the two solutions (35) and (36), whose ratio is

$$z = \frac{\eta_1}{\eta_2} = (w - a)^\sigma[1 + h_1(w - a) + \cdots].$$

From this, we have

$$z^{1/\sigma} = (w - a)\left[1 + \frac{h_1}{\sigma}(w - a) + \cdots \right];$$

$$w - a = z^{1/\sigma} - \frac{h_1}{\sigma} z^{2/\sigma} + \cdots$$

Then w is a single-valued function of z in the neighborhood of the singular point only if $1/\sigma$ is an integer. We have, then, $\sigma = 1/p$, where p is an integer, which is by hypothesis not 1; and

$$z = \frac{\eta_1}{\eta_2} = (w - a)^{1/p}[1 + h_1(w - a) + \cdots]. \tag{42}$$

We see, from (42), that when w makes a positive circuit about a, z is subjected to the elliptic transformation $z' = e^{2\pi i/p}z$.

If the roots of the indicial equation are equal ($\sigma = 0$), we have

$$z = \frac{\eta_3}{\eta_1} = \log(w - a) + k_1(w - a) + k_2''(w - a)^2 + \cdots \tag{43}$$

$$\tau = e^z = (w - a)e^{k_1(w-a)+\cdots} = (w - a)[1 + k_1(w - a) + \cdots].$$

From this, we can express w as a power series in τ which is single valued in z. When w makes a circuit of a, z, from (43), undergoes a parabolic transformation $z' = z + 2\pi i$.

If σ is a positive integer, we have

$$z = \frac{\eta_4}{\eta_1} = k_\sigma \log(w - a) + \frac{1}{(w - a)^\sigma}\left[-\frac{1}{\sigma} + h_1'(w - a) + \cdots\right]. \tag{44}$$

It is easily seen that w is not a single-valued function of z if $k_\sigma \neq 0$.

We have

$$e^{\frac{z}{k_\sigma}} = (w - a)e^{\frac{1}{(w-a)^\sigma}\left[-\frac{1}{\sigma k_\sigma} + \cdots\right]}.$$

The second member has an essential singularity at a, and so takes on the same value at different points w_1, w_2 near a. Then,

$$\frac{z(w_2)}{k_\sigma} = \frac{z(w_1)}{k_\sigma} + 2n\pi i, \quad z(w_2) = z(w_1) + 2nk_\sigma\pi i,$$

where n is some integer. But from (44), letting w make circuits about a, $z(w)$ takes on all values of the form $z(w_1) + 2nk_\sigma\pi i$ at w_1. We have, then, two values of w for the same value of z, which was to be proved.

If $k_\sigma = 0$, and if we consider the ratio

$$z = \frac{\eta_1}{\eta_4} = (w - a)^\sigma[-\sigma + k_1''(w - a) + \cdots],$$

we can show, exactly as in the case treated at the beginning of this section, that this series gives w as a single-valued function of z if, and only if, σ, which must be a positive integer, has the value 1. Then,

$$z = (w - a)[-1 + k_1'''(w - a) + \cdots]. \tag{45}$$

In this case, z is *unchanged* when w makes a circuit of a.

In this last case, the singularity is not significant. By a suitable transformation of the differential equation, it will disappear. The condition that the difference of the roots of (33) be 1 is that

$$[(p_0 - 1)^2 - 4q_0]^{\frac{1}{2}} = 1; \text{ whence, } q_0 = \tfrac{1}{4}p_0^2 - \tfrac{1}{2}p_0. \qquad (a)$$

We now find the condition $k_1 = 0$. The value k_1 in (39), as we see from the equation which precedes it, is

$$k_1 = -p_1 - 2c_1,$$

where, from the first equation of (34), c_1 is determined from

$$f(\alpha' + 1)c_1 + \alpha' p_1 + q_1 = 0.$$

Now,

$$f(\alpha' + 1) = f(\alpha') + f'(\alpha') + \tfrac{1}{2}f''(\alpha') = 2\alpha' + p_0.$$

On using these values, the equation $k_1 = 0$ leads to

$$q_1 = \tfrac{1}{2}p_0 p_1. \qquad (b)$$

If, now, we eliminate the middle term of (28), the coefficient of the last term becomes, according to (22),

$$Q_1(w) = \left[\frac{q_0}{(w-a)^2} + \frac{q_1}{w-a} + \varphi_1(w)\right] - \frac{1}{2}\left[\frac{-p_0}{(w-a)^2} + \varphi_2(w)\right] - \frac{1}{4}\left[\frac{p_0^2}{(w-a)^2} + \frac{2p_0p_1}{w-a} + \varphi_3(w)\right].$$

Equations (*a*) and (*b*) are precisely the conditions that secure the cancelling of the terms with poles at a; and $Q_1(w)$ is analytic at a.

We may state the preceding results in terms of conformal mapping as follows:

THEOREM 6.—*The necessary and sufficient condition that the ratio of two independent solutions of the differential equation* (28) *map the neighborhood of the regular singular point on a plane region is that the roots of the indicial equation* (33) *be subject to one of the following conditions:*

(*a*) $\alpha' - \alpha'' = \frac{1}{p}$, *where p is an integer greater than* 1*;*

(*b*) $\alpha' = \alpha''$*;*

(*c*) $\alpha' - \alpha'' = 1$, *together with the condition* $q_1 = \frac{1}{2}p_0p_1$.

When w makes a single circuit about the singular point, the quotient of the solutions is subject to (*a*) *an elliptic transformation;* (*b*) *a parabolic transformation;* (*c*) *the identical transformation.*

We have considered in each case a particular pair of solutions. The theorem holds, however, in general; for, the ratio of any pair is merely a linear transformation of the ratio of the particular pair used.

111. Equations with Rational Coefficients.—We shall consider, first, the differential equation

$$\frac{d^2\eta}{dw^2} + P(w)\frac{d\eta}{dw} + Q(w)\eta = 0 \tag{1}$$

whose coefficients are rational functions of w and all of whose singular points are regular. In order that ∞ be not an irregular point, certain conditions are imposed on the coefficients. From (27), P and Q must be of the form

$$\left.\begin{aligned} P &= c_1 t + c_2 t^2 + \cdots = \frac{c_1}{w} + \frac{c_2}{w^2} + \cdots \\ Q &= c_2' t^2 + c_3' t^3 + \cdots = \frac{c_2'}{w^2} + \frac{c_3'}{w^3} + \cdots \end{aligned}\right\}. \tag{46}$$

That is, P must have a zero of the first order, at least, and Q a zero of the second order, at least, at infinity.

Let $a_1, \ldots, a_n$ be the finite singular points, and let the exponents thereat be $\alpha_1', \alpha_1''; \alpha_2', \alpha_2''; \ldots; \alpha_n', \alpha_n''$. Let $p_1, p_2, \ldots, p_n$ be the residues of $P(w)$ at these points. Since the sum of the residues of $P(w)$ is zero, we have

$$p_1 + p_2 + \cdots + p_n - c_1 = 0.$$

But, we have, from (33),

$$\alpha_i' + \alpha_i'' = 1 - p_i;$$

and from (27), representing the exponents at ∞ by $\alpha_\infty', \alpha_\infty''$,

$$\alpha_\infty' + \alpha_\infty'' = 1 - (2 - c_1) = c_1 - 1.$$

From these equations, we have the following relations between the exponents:

$$\alpha_1' + \alpha_1'' + \cdots + \alpha_n' + \alpha_n'' + \alpha_\infty' + \alpha_\infty'' = n - 1. \tag{47}$$

If ∞ is not a singular point, Equation (47) still holds if we take $\alpha_\infty' = 1$, $\alpha_\infty'' = 0$. These are the exponents at an ordinary point, as we see from (33) on setting $p_0 = q_0 = 0$.

The products of $P(w)$ and $Q(w)$ by $(w - a_1) \cdots (w - a_n)$ and $(w - a_1)^2 \cdots (w - a_n)^2$, respectively, have no finite singularities and so are polynomials. Hence, the coefficients have the form

$$\begin{aligned} P(w) &= \frac{P_{n-1}(w)}{(w - a_1) \cdots (w - a_n)}, \\ Q(w) &= \frac{Q_{2n-2}(w)}{(w - a_1)^2 \cdots (w - a_n)^2}, \end{aligned} \tag{48}$$

where $P_{n-1}(w)$ is a polynomial of degree $n - 1$, at most, and $Q_{2n-2}(w)$ is a polynomial of degree $2n - 2$, at most, these degrees being determined by the requirements (46) at infinity.

If we express $P(w)$ in terms of the principal parts at its poles, we have

$$\begin{aligned} P(w) &= \frac{p_1}{w - a_1} + \cdots + \frac{p_n}{w - a_n} \\ &= \frac{1 - \alpha_1' - \alpha_1''}{w - a_1} + \cdots + \frac{1 - \alpha_n' - \alpha_n''}{w - a_n}. \end{aligned} \tag{49}$$

We get an analogous expression for $Q(w)$ by considering the principal parts of $(w - a_1) \cdots (w - a_n)Q(w)$. At a_1 we have, if $n > 1$,

$$\begin{aligned} (w - a_1) \cdots (w - a_n)Q(w) &= (w - a_2) \cdots (w - a_n)\left[\frac{q_0}{w - a_1} + \varphi_1(w)\right] \\ &= \frac{q_0(a_1 - a_2) \cdots (a_1 - a_n)}{w - a_1} + \varphi_2(w). \end{aligned}$$

Here, q_0, the constant term in the indicial equation for a_1, is equal to $\alpha_1'\alpha_1''$. We have like principal parts at $a_2, \ldots, a_n$; whence,

$$Q(w) = \frac{1}{(w - a_1) \cdots (w - a_n)}\left[\frac{\alpha_1'\alpha_1''(a_1 - a_2) \cdots (a_1 - a_n)}{w - a_1} + \cdots + \frac{\alpha_n'\alpha_n''(a_n - a_1) \cdots (a_n - a_{n-1})}{w - a_n} + Q_{n-2}(w)\right]. \tag{50}$$

Here, the conditions at infinity (46) require that $Q_{n-2}(w)$ be a polynomial of degree $n - 2$ at most:

$$\begin{aligned} Q_{n-2}(w) &= c_2'w^{n-2} + \cdots + kw + l \\ &= \alpha_\infty'\alpha_\infty''w^{n-2} + \cdots + l. \end{aligned} \tag{51}$$

We shall introduce two types of transformations of the differential equation which will not alter exponent differences. If we change the dependent variable as follows:

$$\xi = (w - a_i)^{l_i}\eta = \frac{1}{t^{l_i}}(1 - a_it)^{l_i}\eta, \tag{52}$$

$(w = 1/t)$, the solutions of the new equation are got from those of the old by multiplying by $(w - a_i)^{l_i}$. On making this multiplication in the solutions of the form (35), (36), (40), or (41), we see that α_i', α_i'' are both increased by l_i. At $a_j (j \neq i)$, there is no alteration of the exponents, since the factor is analytic and different from zero there. At infinity, each exponent is decreased by l_i. By making n such transformations, we have the exponents

$$\alpha_1' + l_1, \alpha_1'' + l_1; \cdot \cdot \cdot ; \alpha_n' + l_n, \alpha_n'' + l_n;$$
$$\alpha_\infty' - \Sigma l_i, \alpha_\infty'' - \Sigma l_i. \quad (53)$$

We can choose the constants l_i so as to give one of the exponents at each finite point a prescribed value. If one of each pair be made zero, we observe that all the terms in the brackets in (50) disappear with the exception of the polynomial $Q_{n-2}(w)$. Also, by this method we can eliminate the exceptional case (*c*) of Theorem 6. By making $\alpha' = 1$, $\alpha'' = 0$, we have $p_0 = q_0 = q_1 = 0$, and there is no singularity.

Secondly, if we make a linear transformation of the independent variable

$$w' = T(w) = \frac{Aw + B}{Cw + D}, \qquad AD - BC = 1, \qquad (54)$$

the singular points of $P(w)$ and $Q(w)$ are transformed by T. Let a_i be carried into a_i', both points being finite. Then,

$$w - a_i = \frac{-Dw' + B}{Cw' - A} - \frac{-Da_i' + B}{Ca_i' - A} = \frac{w' - a_i'}{(Cw' - A)(Ca_i' - A)},$$

and

$$w - a_i = (w' - a_i')[l_0 + l_1(w' - a_i') + \cdot \cdot \cdot], \; l_0 \neq 0.$$

On substituting in the solutions of the form (35), etc., we see that we have solutions with the same exponents as before. If either the original or the transformed singularity is at infinity, we can show in like manner that there is no change in the exponents.

Let us suppose that finite regular points $a_1, \ldots, a_n$ have been preassigned and that the exponents thereat and at infinity have been given such that in the neighborhood of each singularity the function $w = w(z)$, inverse to the quotient of two solutions $z = \zeta_1/\zeta_2$, is single-valued. There remain at our disposal $n - 2$ of the coefficients of the polynomial $Q_{n-2}(w)$ in (50). There arise the questions whether the polynomial can be chosen so that $w(z)$ is single-valued throughout and so that $w(z)$ is an elementary or Fuchsian function. We shall prove the following theorem:

THEOREM 7.—*Let the regular singular points and the exponent differences satisfying the requirements of Theorem 6 be given.*

Then, there is essentially but one differential equation such that the inverse of the ratio of two solutions is an elementary or Fuchsian function; that is, if the equation be reduced to the form (21), $Q_1(w)$ *is uniquely determined.*

Let curves be drawn from a point O to the singular points to form a severed plane F_0. In order that $w(z)$ be single-valued, it is necessary that $z = \zeta_1/\zeta_2 = z(w)$ take on no value twice in F_0; that is, F_0 is mapped by this function on a plane region S_0. If $z(w)$ be extended analytically across any boundary of F_0 and then extended analytically throughout the severed plane, we must have again a plane map S_i not overlapping S_0. On making a circuit p_i times around a_i, where $\sigma_i = 1/p_i$ is the exponent difference at a_i, $z(w)$ returns to its original value. The p_i maps of the severed plane must meet to fill up the neighborhood of a point of the z-plane.

The topological problem here is that already treated in connection with Theorem 11 of Sec. 95. If we put copies of F_0 together to form a limit surface ϕ, closing up each time about a_i when p_i copies have been put together (never closing when the exponent difference is zero), then, $z(w)$ shall map ϕ on a plane region.

Let $z_1 = z_1(w)$ be the mapping function used in connection with Theorem 11 to map ϕ on a plane region. We have, eliminating w, $z = T(z_1)$, a function which maps the z_1-domain of Theorem 11 on a plane region. If the mapping of Theorem 11 was on the whole plane or the finite plane, it follows, from Theorem 3, Sec. 1, that T is linear. Again, if $w(z)$ is a Fuchsian function, the function $z(w)$ maps ϕ on a circle. So $T(z_1)$ maps Q_0 in the z-plane on a circle, and, hence, is linear.

Now the inverse of $z_1(w)$, $w = \varphi(z_1)$, is a simple automorphic function and appears as the inverse of the quotient of two solutions of a differential equation (Theorem 15, Sec. 44)

$$\frac{d^2\eta}{dw^2} + F(w)\eta = 0.$$

Let (21) be the given equation. Then from (23), together with the fundamental linearity property of the Schwarzian derivative (Equation 11, Sec. 44), we have

$$Q_1(w) = \frac{1}{2}D(z)_w = \frac{1}{2}D[T(z_1)]_w = \frac{1}{2}D(z_1)_w = F(w).$$

$Q_1(w)$ is thus uniquely determined.

Whether the automorphic function arising from a pair of solutions is polyhedral, periodic, or Fuchsian depends upon the number of singularities and the exponent differences (setting $\nu_i = 1/\sigma_i$) as stated in Theorem 11, Sec. 95.

112. The Equation with Two Singular Points.—We first make a linear transformation carrying the singular points to the origin and to infinity. Let α', α'' be the exponents at the origin. Then, from (49) and (48), where $Q_{2n-2}(w)$ is a constant, the equation is

$$\frac{d^2\eta}{dw^2} + \frac{1 - \alpha' - \alpha''}{w}\frac{d\eta}{dw} + \frac{\alpha'\alpha''}{w^2}\eta = 0. \tag{55}$$

This equation can be solved in terms of elementary functions.

If $\alpha' \neq \alpha''$, we have the two solutions

$$\eta_1(w) = w^{\alpha'}, \qquad \eta_2(w) = w^{\alpha''},$$

with the ratio

$$z = \frac{\eta_1(w)}{\eta_2(w)} = w^{\alpha' - \alpha''} = w^{\sigma}.$$

If $\sigma = 1/p$, we have

$$z = w^{1/p}; \qquad w = z^p,$$

the latter function being simply automorphic with respect to the elliptic cyclic group $z' = e^{2n\pi i/p}z$.

If $\alpha' = \alpha''$, we find the solutions

$$\eta_1(w) = w^{\alpha'}, \qquad \eta_2(w) = w^{\alpha'} \log w.$$

Taking the ratio of these,

$$z = \frac{\eta_2(w)}{\eta_1(w)} = \log w; \qquad w = e^z.$$

The function e^z is simply automorphic with respect to the simply periodic group $z' = z + 2n\pi i$.

113. The Hypergeometric Equation.—Given a differential equation with three regular singular points. We first make a linear transformation carrying the three points to 0, 1, ∞. Let the exponent differences be λ, μ, ν, respectively. Then, by transformations of the form (52), we can give the exponents at zero the values $\alpha_0' = \lambda$, $\alpha_0'' = 0$ and at 1 the values $\alpha_1' = \mu$, $\alpha_1'' = 0$. Then, $Q(w)$ in (50) reduces to the form

$$Q(w) = \frac{\alpha_\infty'\alpha_\infty''}{w(w-1)}.$$

But $\alpha_\infty' - \alpha_\infty'' = \nu$, and, from (47), $\alpha_\infty' + \alpha_\infty'' = 1 - \lambda - \mu$; whence,

$$\alpha_\infty'\alpha_\infty'' = \tfrac{1}{4}[(1 - \lambda - \mu)^2 - \nu^2];$$

and the differential equation has the form

$$\frac{d^2\eta}{dw^2} + \left[\frac{1-\lambda}{w} + \frac{1-\mu}{w-1}\right]\frac{d\eta}{dw} + \frac{(1-\lambda-\mu)^2-\nu^2}{4w(w-1)}\eta = 0. \quad (56)$$

This is simply the well-known hypergeometric differential equation.[1]

$$w(1-w)\frac{d^2\eta}{dw^2} + [c-(a+b+c)w]\frac{d\eta}{dw} - ab\eta = 0, \quad (57)$$

where $a = \frac{1}{2}(1-\lambda-\mu+\nu)$, $b = \frac{1}{2}(1-\lambda-\mu-\nu)$, $c = 1-\lambda$. Other familiar equations can be reduced to this form; for example, Legendre's equation

$$(1-w^2)\frac{d^2\eta}{dw^2} - 2w\frac{d\eta}{dw} + n(n+1)\eta = 0,$$

which has the regular singular points 1, -1, ∞ with the exponents 0, 0; 0, 0; $n+1$, $-n$, respectively.

In order that the inverse of the quotient of two solutions of (56) be a single-valued function each of the quantities λ, μ, ν must be (Theorem 6) either zero or the reciprocal of an integer. The coefficients of the equation are then real. We may assume that case (*c*) of Theorem 6 does not occur (for we could then reduce the number of singularities to two) and that λ, μ, ν are not negative. These conditions determine the equation uniquely. If the equation be put in the form (21), $Q_1(w)$ is uniquely determined. There being thus, essentially, but one differential equation with the three singularities and the prescribed exponent differences, it must necessarily be the equation for which the inverse of the quotient of two solutions, $z = \zeta_1/\zeta_2 = z(w)$, is an elementary or Fuchsian function.

Consider the map of the upper half w-plane by any branch of $z(w)$. On the real axis between 0 and 1, since the coefficients are real, we can select a pair of linearly independent real solutions. Their ratio z_1 is real and maps the segment 01 on a segment of the real axis. Since z is a linear function of z_1, it follows that each branch of $z(w)$ maps 01 on the arc of a circle. Similarly, each branch maps the segment between 1 and $+\infty$ on a circular arc and the segment between $-\infty$ and 0 on a circular arc.

[1] The problem considered in this and the following section was treated in a famous paper by SCHWARZ, "Ueber diejenigen Faelle, in welchen die Gaussische hypergeometrische Reihe eine algebraische Funktion ihres vierten Elementes darstellt." *Jour. für Math.*, vol. 75, pp. 292–335; also *Ges. Math. Abhandlungen*, Bd. 2, pp. 211–259.

In the neighborhood of the origin, there exists a pair of solutions whose ratio (Equation (42)) is

$$z_0 = w^\lambda[1 + h_1 w + \cdot \cdot \cdot].$$

This maps the neighborhood of the origin in the upper half plane on the interior of an angle of magnitude $\lambda\pi$. Then z, which is a linear function of z_0, maps this on an angle of magnitude $\lambda\pi$. So, the two circles on which the parts of the real axis to the right and left of the origin are mapped meet at the angle $\lambda\pi$. Similar remarks apply at 1 and at ∞, with the result that *any branch of $z(w)$ in the upper half plane maps the half plane on a circular arc triangle with angles $\lambda\pi$, $\mu\pi$, $\nu\pi$.*

If the branch of $z(w)$ just used be extended analytically across one of the segments, as 01, into and throughout the lower half w-plane, the lower half plane is mapped on a circular arc triangle which is the inverse of the previous triangle in one of its sides (Theorem 19, Sec. 83). Continuing this process, the various branches of $z(w)$ map the two half planes on a finite or infinite system of non-overlapping circular arc triangles fitting together without lacunæ, each with angles $\lambda\pi$, $\mu\pi$, $\nu\pi$, and such that the inverse of any triangle in one of its sides gives another triangle of the system.

It is customary to shade those triangles which are maps of the upper half plane. Then, the triangles are alternately shaded and unshaded. By an inversion in a side of any triangle, the system of triangles is carried into itself, each shaded triangle being carried into an unshaded triangle.

114. The Riemann-Schwarz Triangle Functions.[1]—The inverse $w = w(z)$ of the quotient of the two solutions is automorphic with respect to the group of linear transformations got by carrying any shaded triangle into any other shaded triangle by a sequence of inversions in sides. A shaded and an unshaded triangle together form a fundamental region for the group. Any simple automorphic function belonging to the group is called a "triangle function."

The function $w(z)$ maps a shaded and an unshaded triangle on the whole w-plane. Hence, $w(z)$ takes on each value once in the fundamental region. It follows that any triangle function is a rational function of $w(z)$.

[1] RIEMANN, B., Vorlesungen über die hyperg. Reihe, in his "Werke, Nachträge;" SCHWARZ, H. A., *loc. cit.*

The function $w(z)$, which maps a triangle of the system on a half plane bounded by the real axis, is real on the boundary of each triangle and is not real within the triangle. At points inverse with respect to the side of any triangle, $w(z)$ has conjugate imaginary values.

As appears from Theorem 11, Sec. 95, $w(z)$ is a polyhedral, periodic, or Fuchsian function according as $\lambda + \mu + \nu$ is greater than, equal to, or less than 1. The distinction between these three types can be made, however, by a direct study of the triangle. Let the two sides which issue from a vertex meet again at a point P. If a linear transformation carrying P to ∞ be made, the two sides are carried into straight lines. The character of the

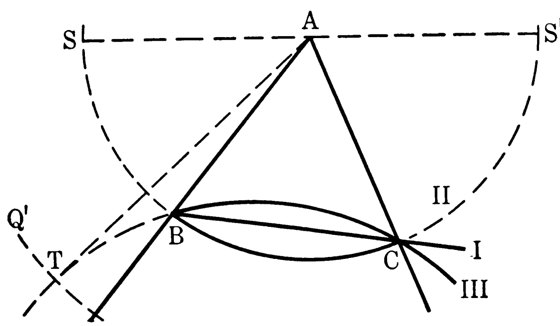


FIG. 89.

third side in each of the three cases is shown in Fig. 89. A is at ∞ and the two sides are parallel if the original two sides were tangent.

Case I. $\lambda + \mu + \nu = 1$.—The sum of the angles is π and the third side is a straight line. There are four possible sets of values:

(*a*) $1/3, 1/3, 1/3$; (*b*) $1/2, 1/4, 1/4$; (*c*) $1/2, 1/3, 1/6$; (*d*) $1/2, 1/2, 0$.
In (*a*), we have an equilateral triangle; in (*b*), an isosceles right triangle; in (*c*), a 30–60-degree right triangle. In each case, by continued reflections in sides, the whole finite plane is covered by an infinite number of triangles. In (*d*), two sides are parallel and the third is a common perpendicular. The repeated reflections of this half strip cover the finite plane with an infinite number of copies.

The groups that appear here are those studied in Sec. 60. Each of these groups contains a periodic subgroup. The triangle function $w(z)$ can be identified, except for possible linear transformations, with the functions (*a*) $\mathfrak{P}'(z)$; (*b*) $\mathfrak{P}(z)^2$; (*c*) $\mathfrak{P}'(z)^2$; (*d*) $\cos z$, of Sec. 61.

Case II. $\lambda + \mu + \nu > 1$.—The sum of the angles is greater than π and the third side is a circle concave toward A. A lies within this circle; otherwise, the triangles would overlap when an inversion in BC is made. Through A draw a chord SS' of

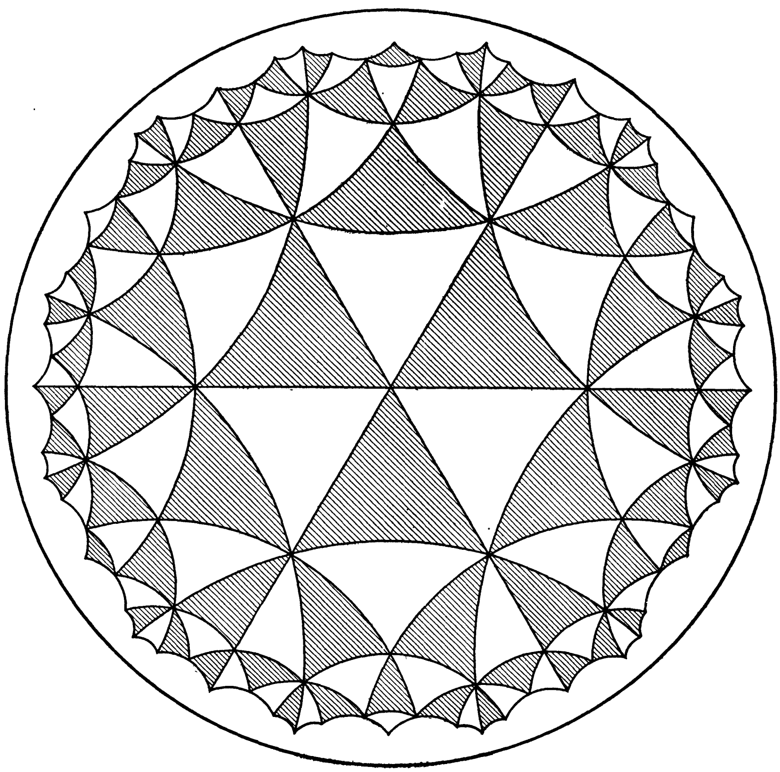


FIG. 90.

the circle of the third side which is bisected at A (SS' is perpendicular to the radius through A). The circle Q with A as center and AS as radius is intersected by each side of the triangle at points which lie at opposite ends of a diameter.

Now, let the plane be projected stereographically on a sphere with Q as equator. The points of Q remain fixed. The sides of the triangle are carried into three circular arcs on the sphere, each of which passes through opposite ends of a diameter. That is, the sides of the corresponding triangle $A'B'C'$ on the sphere are arcs of great circles.

Inversions in the sides of ABC correspond to reflections in the diametral planes containing the sides of $A'B'C'$. The new triangles are bounded by arcs of great circles; and all succeeding inversions in the plane correspond to reflections in diametral planes. The sphere is covered by a finite number of triangles; and the group is finite.

The groups are those of the regular solids treated in Chap. VI. The triangle functions are the polyhedral functions. The set of triangles where λ, μ, ν have the values $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$ is shown in Fig. 38, p. 136.

Case III. $\lambda + \mu + \nu < 1$.—The sum of the angles of the triangle is less than π, and the third side is convex toward A. From A we can draw a tangent AT to the circle of the third side. A circle Q' with center A and passing through the point of tangency T is orthogonal to all three sides. By an inversion in any side of the triangle, Q' is carried into itself. The new triangles will have their sides orthogonal to Q'; and for all succeeding inversions Q' is a fixed circle. The group is Fuchsian.

There are infinitely many values of λ, μ, ν such that their sum is less than 1. Figure 90 shows the system of triangles for the values $\frac{1}{3}$, $\frac{1}{4}$, $\frac{1}{4}$.

The elliptic modular functions $J(\tau)$ and $\lambda(\tau)$ are triangle functions. In the former, λ, μ, ν have the values $\frac{1}{2}$, $\frac{1}{3}$, 0; in the latter 0, 0, 0. The fundamental regions of Figs. 46 and 48 consist of two triangles. The imaginary axis in each figure separates the fundamental region into its two component triangles.

115. Equations with Algebraic Coefficients.—Let the differential equation be

$$\frac{d^2\eta}{dw^2} + P(W, w)\frac{d\eta}{dw} + Q(W, w)\eta = 0, \tag{58}$$

where W and w are connected by a polynomial relation $\Phi(W, w) = 0$, and P and Q are rational functions. We have found that every simple automorphic function is expressible as the inverse of the ratio of two solutions of an equation of this form (Theorem 15, Sec. 44).

The functions P and Q are single valued on the Riemann surface of the algebraic function $W = W(w)$. Their singularities are isolated. In order that the inverse of the quotient of two solutions be single valued at a singularity occurring at an ordinary point on the surface, the requirement of Theorem 4 must be met. At a finite branch point a of order m, the point at infinity in a plane sheet, or an infinite branch point of order m, we make the change of variable $w = a + t^m$, $w = 1/t$, or $w = 1/t^m$. We then apply our previous theorems to the transformed differential equation at $t = 0$.

We shall not go into details, which are lengthy, in this case. With given regular singular points and given suitable exponents, it will be found here, as in the case of rational coefficients, that there are usually certain undefined constants appearing in the coefficients. That is, the singularities and exponents do not determine the coefficients uniquely, except in the simplest cases. A theorem analogous to Theorem 7, and proved in an identical manner, can be established to show that it is always possible to choose these constants, in essentially but one way, so that the inverse of the quotient of two solutions is a polyhedral, elliptic, or Fuchsian function. The kind of function depends upon the number of singularities, the exponent differences, and the genus of the surface, as stated in Theorems 9 to 11 of Chap. IX.

The use of differential equations leads to the same automorphic functions as those given by the method of conformal mapping. We have, however, in the differential equation, in case the free constants can be actually determined, an analytical instrument of great value in the investigation of the properties of the automorphic functions.

A BIBLIOGRAPHY OF AUTOMORPHIC FUNCTIONS

NOTE.—Many particular classes of functions, *e.g.*, the circular functions, elliptic functions, and triangle functions (including the elliptic modular and polyhedral functions), had been studied extensively for many years before the general theory of automorphic functions was created. It seems undesirable to attempt here a bibliography of these earlier researches, and the present list will therefore begin w th the papers in which Poincaré created the general theory. Mention should, however, be made of one or two memoirs of the earlier period in which the general theory was notably foreshadowed; especially—

SCHWARZ, H. A., *Über diejenigen Fälle, in welchen die Gaussische hypergeometrische Reihe eine algebraische Funktion ihres vierten Elementes darstellt,* Crelle's J., **75** (1872), p. 292.

SCHOTTKY, F., *Über die conforme Abbildung mehrfach zusammenhängender ebener Flächen,* Crelle's J., **83** (1877), p. 300.

FUCHS, L., *Über eine Klasse von Funktionen mehrerer Variabeln, welche durch Umkehrung der Integrale von Lösungen der linearen Differentialgleichungen mit rationalen Coeffizienten entstehen,* Crelle's J., **89** (1880), p. 150.

POINCARÉ, H., *Sur les fonctions fuchsiennes,* C.R., **92** (1881), 333, 393, 859, 1198, 1274, 1484; **93** (1881), 301, 581.

POINCARÉ, H., *Sur les fonctions uniformes qui se reproduisent par des substitutions linéaires,* Math. Ann., **19** (1882), 553; **20** (1882), 52.

KLEIN, F., *Über eindeutige Functionen mit linearen Transformationen in sich,* Math. Ann., **19** (1882), 565; **20** (1882), 49.

POINCARÉ, H., *Sur les fonctions fuchsiennes,* C.R., **94** (1882), 163, 1038, 1166; **95** (1882), 626; **97** (1883), 1485.

POINCARÉ, H., *Théorie des groupes fuchsiens,* Acta Math., **1** (1882), 1.

HURWITZ, A., *Ueber eine Reihe neuer Functionen, welche die absoluten Invarianten gewisser Gruppen ganzzahliger linearer Transformationen bilden,* Math. Ann., **20** (1882), 125.

RAUSENBERGER, O., *Ueber eindeutige Functionen mit mehreren nicht vertauschbaren Perioden,* Math. Ann., **20** (1882) 187; **21** (1883), 59.

SCHOTTKY, F., *Ueber eindeutige Functionen mit linearen Transformationen in sich,* Math. Ann., **20** (1882), 293.

FUCHS, L., *Ueber Functionen, welche durch lineare Substitutionen unverändert bleiben,* Gött. Nach. (1882), 81.

POINCARÉ, H., *Mémoire sur les fonctions fuchsiennes,* Acta Math., **1** (1882), 193.

PICARD, E., *Sur une classe de groupes discontinus de substitutions linéaires et sur les fonctions de deux variables indépendantes restant invariables par ces substitutions,* Acta Math., **1** (1882), 297.

Poincaré, H., *Sur un théoreme de la théorie générale des fonctions*, Bull. de la Soc. Math. de France, **11** (1883), 112.

Poincaré, H., *Mémoire sur les groupes kleinéens*, Acta Math., **3** (1883), 49.

Dyck, W., *Ueber die durch Gruppen-linearer Transformationen gegebenen regulären Gebietseinteilungen des Raumes*, Leipz. Ber. (1883), 61.

Klein, F., *Neue Beiträge zur Riemann'schen Functionentheorie*, Math. Ann., **21** (1883), 141.

Picard, E., *Mémoire sur les formes quadratiques binaires indéfinies*, Ann. de l'Éc. Norm. (3), **1** (1884), 9.

Poincaré, H., *Sur les groupes des équations linéaires*, Acta Math., **4** (1884), 201.

Picard, E., *Sur un groupe des transformations des points de l'espace situés du même coté d'un plan*, S.M.F. Bull., **12** (1884), 43.

Poincaré, H., *Mémoire sur les fonctions zétafuchsiennes*, Acta Math., **5** (1884), 209.

Klein, F., *Vorlesungen über das Ikosaeder und die Auflösung der Gleichungen fünften Grades*, Leipzig, 1884; English trans., G. G. Morrice, London, 1888; 2nd. ed., 1913.

Picard, E., *Sur les formes quadratiques ternaires indéfinies à indéterminées conjugées et sur les fonctions hyperfuchsiennes correspondantes*, Acta Math., **5** (1884), 121.

Poincaré, H., *Sur un théorème de M. Fuchs*, Acta Math., **7** (1885), 1.

Biermann, O., *Zur Theorie der Fuchs'schen Functionen*, Wien Sitzungsb., **92** (1885), 1137.

von Mangoldt, H., *Ueber ein Verfahren zur Darstellung elliptischer Modulfunctionen durch unendliche Producte nebst einer Ausdehnung dieses Verfahrens auf allgemeinere Functionen*, Gött. Nach. (1886), 1.

Weber, H., *Ein Beitrag zu Poincaré's Theorie der Fuchs'schen Functionen*, Gött. Nach. (1886), 359.

Poincaré, H., *Sur une transformation des fonctions fuchsiennes et la réduction des intégrales abéliennes*, C.R., **102** (1886), 41.

Poincaré, H., *Sur les fonctiones fuchsiennes et les formes quadratiques ternaires indéfinies*, C. R., **102** (1886), 735.

Humbert, G., *Application de la théorie des fonctions fuchsiennes à l'étude des courbes algébriques*, Liouville's J. (4), **2** (1886) 239.

Schottky, F., *Ueber eine specielle Function, welche bei einer bestimmten linearen Transformation ihres Argumentes unverändert bleibt*, Crelle's J., **101** (1887), 227.

Picard, E., *Démonstration d'un théorème général sur les fonctions uniformes liées par une relation algébrique*, Acta Math., **11** (1887), 1.

Humbert, G., *Sur le théorème d'Abel et quelques-unes de ses applications géométriques*, Liouville's J. (4), **3** (1887), 327.

Poincaré, H., *Les fonctions fuchsiennes et l'arithmétique*, Liouville's J. (4), **3** (1887), 405.

Stouff, X., *Sur la transformation des fonctions fuchsiennes*, Ann. de l'Éc. Norm. (3), **5** (1888), 219.

Stahl, H., *Ueber die Darstellung der eindeutigen Functionen, die sich durch lineare Substitutionen reproduciren, durch unendliche Producte*, Math. Ann., **33** (1888), 291.

BIANCHI, L., *Sulle superficie Fuchsiane*, Lincei Rend. (4), $\mathbf{4}_2$ (1888), 161.

PICARD, E., *Sur les formes quadratiques binaires à indéterminées conjuguées et les fonctions fuchsiennes*, Am. J. Math., **11** (1888), 187.

SCHLESINGER, L., *Zur Theorie der Fuchs'schen Functionen*, Crelle's J., **105** (1889), 181.

DE BRUN, F., *Bevis för nagra teorem af Poincaré*, Öfv. af K. Svenska Vet.-Ak. Förh., **46** (1889), 677.

SCHWARZ, H. A., *Gesammelte mathematische Abhandlungen*, 2 vols., Berlin, 1890.

KLEIN, F., and FRICKE, R., *Vorlesungen über die Theorie der Modulfunktionen*, 2 vols., Leipzig, 1890–1892.

HURWITZ, A., *Ueber die Differentialgleichungen dritter Ordnung, welchen die Formen mit linearen Transformationen in sich genügen*, Math. Ann., **33** (1889), 345.

STOUFF, X., *Sur certains groupes fuchsienes formés avec les racines d'équations binômes*, Toulouse Ann., **4** (1890), P.

FRICKE, R., *Ueber eine besondere Klasse diskontinuierlicher Gruppen reeller linearer Substitutionen*, Math. Ann., **38** (1890), 50, 461.

KLEIN, F., *Zur Theorie der Abel'schen Functionen*, Math. Ann., **36** (1890), 1.

BIERMANN, O., *Ueber die Darstellung der Fuchs'schen Functionen erster Familie durch unendliche Producte*, Monatshefte f. M., **1** (1890), 49; **3** (1892), 143.

KLEIN, F., *Zur Theorie der Laméschen Functionen*, Gött. Nach. (1890), 85.

KLEIN, F., *Ueber Normirung der linearen Differentialgleichungen zweiter Ordnung*, Math. Ann., **38** (1890), 144.

CASSEL, G., *Öfver en afhandling af H. Weber med titel: "Ein Beitrag zu Poincaré's Theorie der Fuchs'schen Functionen,"* Bihang till Sv. Ak., **16** (1891), Afd. I., No. 2.

CASSEL, G., *Sur un problème de représentation conforme*, Acta Math., **15** (1891), 33.

PAINLEVÉ, P., *Mémoire sur les équations différentialles du premier ordre*, Annales de l'Éc. Norm. (3), **8** (1891), 9, 103, 201, 267.

BURNSIDE, W., *On a class of automorphic functions*, Proc. L.M.S., **23** (1892), 49, 281.

RITTER, E., *Die eindeutigen automorphen Formen vom Geschlecht Null*, Gött. Nach. (1892), 283; Math. Ann., **41** (1892), 1.

KLEIN, F., *Ueber den Begriff des functionentheoretischen Fundamentalbereichs*, Math. Ann., **40** (1892), 130.

FRICKE, R., *Weitere Untersuchungen über automorphe Gruppen solcher linearen Substitutionen einer Variabeln, deren Substitutionskoeffizienten Quadratwurzeln ganzer Zahlen enthalten*, Math. Ann., **39** (1891), 62.

FRICKE, R., *Ueber discontinuirliche Gruppen, deren Substitutionscoefficienten ganze Zahlen eines biquadratischen Körpers sind*, Gött. Nach. (1892), 268.

FRICKE, R., *Ueber ein allgemeines arithmetisch-gruppentheoretisches Princip in der Theorie der automorphen Functionen*, Gött. Nach. (1892), 453.

SCHLESINGER, L., *Sur la théorie des fonctions fuchsiennes*, C.R., **114** (1892), 1100, 1409.

SCHLESINGER, L., *Ueber die bei den linearen Differentialgleichungen zweiter Ordnung aufbietenden Primformen*, Crelle's J., **110** (1892), 130.

BURCKHARDT, H., *Ueber die Darstellung einiger Fälle der automorphen Primformen durch specielle Thetareihen,* Math. Ann., **42** (1893), 185.

DEL PEZZO, P., *Sui gruppi Kleineani a due variabili,* Napoli Rend. (2), **7** (1893), 123.

RITTER, E., *Die automorphen Formen von beliebigem Geschlechte,* Gött Nach. (1893), 121.

FRICKE, R., *Zur gruppentheoretischen Grundlegung der automorphen Functionen,* Math. Ann., **42** (1893), 564.

FRICKE, R., *Ueber indefinite quadratische Formen mit drei und vier Veränderlichen,* Gött. Nach. (1893), 705.

PICARD, E., *De l'équation* $\Delta u = ke^u$ *sur une surface de Riemann fermée,* Liouville's J. (4), **9** (1893), 273.

FRICKE, R., *Entwickelungen zur Transformation fünfter und siebenter Ordnung einiger specieller automorpher Functionen,* Acta Math., **17** (1893), 345.

STÄCKEL, P., *Ueber algebraische Gleichungen zwischen eindeutigen Functionen, welche lineare Substitutionen in sich zulassen,* Crelle's J., **112** (1893) 287.

BURNSIDE, W., *Note on linear substitutions,* Mess. Math. (2), **22** (1893), 190.

BURNSIDE, W., *Note on the equation* $y^2 = x(x^4 - 1)$, Proc. L.M.S., **24** (1893), 17.

SCHOENFLIES, A., *Ueber Kreisbogenpolygone,* Math. Ann., **42** (1893) 377.

PICK, G., *Ueber das Formensystem eines Kreisbogenpolygons vom Geschlecht Null,* Math. Ann., **42** (1893), 489.

BIANCHI, L., *Sopra alcune classi di gruppi di sostituzioni lineari a coefficienti complessi,* Math. Ann., **43** (1893), 101.

KAPTEYN, W., *Een paar stellingen van Poincaré,* Handel. v. h. 4de Nederl. Nat. Congres (1893), 149.

FORSYTH, A. R., *Theory of Functions of a complex variable,* Cambridge, (1893), 2nd. ed., 1900.

SCHOENFLIES, A., *Ueber Kreisbogendreiecke und Kreisbogenvierecke,* Math. Ann., **44** (1894), 105.

D'OVIDIO. E., *Sulle funzioni Thetafuchsiane,* Torino Atti, **29** (1894), 741.

RITTER, E., *Die multiplicativen Formen auf algebraischen Gebilde beliebigen Geschlechtes, mit Anwendung auf die Theorie der automorphen Formen,* Math. Ann., **44** (1894), 261.

FRICKE, R., *Ueber die Transformationstheorie der automorphen Functionen,* Math. Ann., **44** (1894), 97.

RITTER, E., *Die Stetigkeit der automorphen Functionen bei stetiger Abänderung des Fundamentalbereichs,* Math. Ann., **45** (1894), 473; **46** (1895), 200.

FRICKE, R., *Eine Anwendung der Idealtheorie auf die Substitutionsgruppen der automorphen Functionen,* Gött. Nach. (1894), 106.

STOUFF, X., *Sur différents points de la théorie des fonctions fuchsiennes,* Toulouse Ann., **8** (1894), D.

CASSEL, G., *Kritiska studier öfver teorin för de automorfa funktionerna.* Diss. Upsala (1894).

FRICKE, R., *Die Kreisbogenvierseite und das Princip der Symmetrie,* Math. Ann., **44** (1894), 565.

KLEIN, F., *Ueber lineare Differentialgleichungen der zweiten Ordnung.* Autograph. Vorlesung. Göttingen, 1894; Math. Ann., **46** (1895), 77.

FRICKE, R., *Ueber die Discontinuitätsbereiche der Gruppen reeller linearen Substitutionen einer complexen Variabeln*, Gött. Nach. (1895), 360.

BAKER, H. F., *On a certain automorphic function*, Cambr. Proc., **8** (1895), 322.

RITTER, E., *Ueber Riemann'sche Formenschaaren auf einem beliebigen algebraischen Gebilde*, Math. Ann., **47** (1895), 157.

GORDAN, P., *Ueber unverzweigte lineare Differentialgleichungen der zweiten Ordnung auf ebenen Curven vierten Grades*, Math. Ann., **46** (1895), 606.

SCHLESINGER, L., *Handbuch der Theorie der linearen Differentialgleichungen*, 2 vols., Leipzig, 1895–1898.

BAGNERA, F., *Sul teorema dell' esistenza delle funzioni Fuchsiane*, Rivista di Mat., **5** (1896), 31.

KEMPINSKI, S., *Ueber Fuchs'sche Functionen zweier Variabeln*, Math. Ann., **47** (1896), 573; Bull. intern. de Cracovie (1895), 288.

FRICKE, R., *Notiz über die Discontinuität gewisser Collineationsgruppen*, Math. Ann., **47** (1896), 557.

FRICKE, R., *Ueber die Theorie der automorphen Modulgruppen*, Gött. Nach. (1896), 91.

FRICKE, R., *Die Theorie der automorphen Functionen und die Arithmetik*, Papers published by the Amer. Math. Soc., I. (1896), 72.

VITERBI, A., *Le equazioni differenziali lineari a coefficienti algebrici integrabili algebricamente, studiate in base alla teoria delle "funzione fuchsiane" del Poincaré*, Giornale di Batt., **35** (1897), 150; **36** (1898), 55.

FRICKE, R., and KLEIN, F., *Vorlesungen über die Theorie der automorphen Functionen*, Leipzig, Teubner, 1897–1912.

FRICKE, R., *Ueber den arithmetischen Charakter gewisser Netze von unendlich vielen congruenten Vierecken*, Braunschw. Festschr. (1897), 85.

BAKER, H. F., *Abel's theorem and the allied theory*, Cambridge, 1897.

FRICKE, R., *Die Discontinuitätsbereiche der Gruppen reeller linearer Substitutionen einer complexen Veränderlichen*, Deutsche Math. Ver., **4** (1897), 151.

KLUYVER, J. C., *A special case of Dirichlet's problem for two dimensions*, Acta Math., **21** (1897), 265.

POINCARÉ, H., *Sur l'intégration algébrique des équations différentielles du premier ordre et du premier degré*, Palermo Rend., **11** (1897), 193.

POINCARÉ, H., *Les fonctions fuchsiennes et l'équation* $\Delta u = e^u$, Liouville's J., (5), **4** (1898), 137; C.R., **126** (1898), 627.

SCHLESINGER, L., *Sur un problème de Riemann*, C. R., **126** (1898), 723.

FRICKE, R., *Ueber die Beziehungen zwischen der Zahlentheorie und der Theorie der automorphen Functionen*, Jahresb. d. Deutsch. Math. Ver., **6** (1898), 94.

WHITTAKER, E. T., *On the connexion of algebraic functions with automorphic functions*, Phil. Trans., **192 A** (1898), 1.

BAGNERA, G., *Un teorema relativo agli invarianti delle sostituzioni di un gruppo Kleinieno*, Rendic. d. Lincei (5^a), **7** (1898), 340.

DIXON, A. C., *Notes on the theory of automorphic functions*, Proc. L.M.S., **31** (1899), 297.

LINDEMANN, F., *Zur Theorie der automorphen Functionen*, Münch. Sitzungsber., **29** (1899), 423; **30** (1900), 493.

WIRTINGER, W., *Zur Theorie der automorphen Functionen von* n *Veränderlichen*, Wien Ber., **108** (1899), 1239.

FRICKE, R., *Zur Theorie der Poincaré'schen Reihen*, Jahresb. d. Deutschen M. Ver., **9** (1900), 78.

FRICKE, R., *Die automorphen Elementarformen*, Gött. Nach. (1900), 303.

FRICKE, R., *Die Ritter'sche Primform auf einer beliebigen Riemann'schen Fläche*, Gött. Nach. (1900), 314.

WELLSTEIN, J., *Ueber Primformen auf Riemann'schen Flächen*, Gött. Nach. (1900), 380.

DIXON, A. C., *Notes on the theory of automorphic functions*, Proc. L.M.S., **32** (1900), 353.

DIXON, A. C., *Prime functions on a Riemann surface*, Proc. L.M.S., **33** (1900), 10.

HILBERT, D., *Mathematische Probleme*, Gött. Nach. (1900), 253; Arch. d. M. u. P. (3), **1** (1901), 44, 213.

FRICKE, R., *Ueber die Poincaré'schen Reihen der* (-1)*-ten Dimension*, Dedekind Festschrift (1901).

HUTCHINSON, J. I., *On a class of automorphic functions*, Trans. Amer. M.S., **3** (1902), 1.

ALEZAIS, R., *Sur une classe de fonctions hyperfuchsiennes et sur certaines substitutions linéaires qui s'y rapportent*, Ann. de l'Éc. Norm , (3), **19** (1902), 261.

WHITTAKER, E. T., *Note on a function analogous to Weierstrass' sigma-function*, Mess. Math., **31** (1902), 145.

FRICKE, R., *Ueber die in der Theorie der automorphen Funktionen auftretenden Polygoncontinua*, Gött. Nach. (1903), 331.

POINCARÉ, H., *Sur l'intégration algébrique des équations linéaires et les périodes des intégrales abéliennes*, Liouville's J. (5), **9** (1903), 139.

BLUMENTHAL, O., *Ueber Modulfunktionen von mehreren Veränderlichen*, Math. Ann., **56** (1903), 509; **58** (1904), 497.

FUBINI, G., *Sulla teoria delle forme quadratiche Hermitiane e dei sistemi di tali forme*, Atti d. Gioenia (4), **17** (1903).

BLUMENTHAL, O., *Ueber Thetafunktionen und Modulfunktionen mehrerer Veränderlichen*, Deutsch. Math. Ver., **13** (1904), 120.

FUBINI, G., *Sulla funzioni automorfe ed iperfuchsiane di più variabili indipendenti*, Annali di Mat. (3), **10** (1904), 1.

YOUNG, J. W., *On the group of the sign* (0, 3; 2, 4, ∞) *and the functions belonging to it*, Trans. Amer. Math. Soc., **5** (1904), 81.

FUBINI, G., *Sulla teoria dei gruppi discontinui*, Ann. di Mat. (3), **11** (1904), 159.

FRICKE, R., *Beiträge zum Continuitätsbeweise der Existenz linear-polymorpher Funktionen auf Riemannschen Flächen*, Math. Ann., **59** (1904), 449.

FUBINI, G., *Una questione fondamentale per la teoria dei gruppi e delle funzioni automorfe*, Rend. Lincei, **13** (1904), 590.

FRICKE, R., *Neue Entwicklungen über den Existenz-beweis der polymorphen Funktionen*, Heidelberg dritt Math. Kongress (1904), 246.

HUTCHINSON, J. I., *On the automorphic functions of the group* (0, 3; 2, 6, 6), Trans. Amer. Math. Soc., **5** (1904), 447.

BLUMENTHAL, O., *Bemerkungen zur Theorie der automorphen Funktionen*, Gött. Nach. (1904), 92.

FUCHS, L., *Gesammelte mathematische Werke*, 3 vols., Berlin, 1904–1909.

POINCARÉ, H., *Sur les invariants arithmétiques*, Crelle's Journ., **129** (1905), 89.

HURWITZ, A., *Zur Theorie der automorphen Funktionen von beliebig vielen Variabeln*, Math. Ann., **61** (1905), 325.

MORRIS, R., *On the automorphic functions of the group* (0, 3; l_1, l_2, l_3), Tr. Am. M.S., **7** (1906), 425.

BRODÉN, T., *Zur Theorie der mehrdeutigen automorphen Funktionen*, Lunds Universitets Års-skrift, **40** (1904), 54.

JOHANSSON, S., *Ein Satz über die konforme Abbildung einfach zusammenhängender Riemann'scher Flächen auf den Einheitskreis*, Math. Ann., **62** (1906), 177.

JOHANSSON, S., *Beweis der Existenz linear-polymorpher Funktionen vom Grenzkreistypus auf Riemann'schen Flächen*, Math. Ann., **62** (1906), 184.

HUTCHINSON, J. I., *On automorphic groups whose coefficients are integers in a quadratic field*, Trans. Am. M.S., **7** (1906), 530.

LEVI, E. E., *Ricerche sulla teoria delle funzioni automorfe*, Atti R. Acc. Lincei (5), **15** (1906), 682.

FUBINI, G., *Sulla teoria delle funzioni automorfe*, Atti R. Inst. Ven., **65** (1905), 1297.

HERGLOTZ, L., *Ueber die Gestalt der auf algebraischen Kurven nirgends singulären linearen Differentialgleichungen zweiter Ordnung*, Math. Ann., **62** (1906), 329.

FUBINI, G., *Sulla costruzione dei compi fondamentali di un gruppo discontinuo*, Ann. d. Mat. (3), **12** (1906), 347.

HUTCHINSON, J. I., *A method for constructing the fundamental region of a discontinuous group of linear transformations*, Tr. Am. M.S., **8** (1907), 261.

KOEBE, P., *Ueber die Uniformisierung reeller algebraischer Kurven*, Gött. Nach. (1907), 177.

KOEBE, P., *Ueber die Uniformisierung beliebiger analytischer Kurven*, Gött. Nach. (1907), 191, 633.

FUBINI, G., *Nuove ricerche interno ad alcune classi di gruppi discontinui*, Palermo Rend., **21** (1906), 177.

FATOU, P., *Sur les solutions uniformes de certaines équations fonctionnelles*, C.R., **143** (1906), 546.

STAHL, H., *Berechtigung einer Arbeit von Herrn E. T. Whittaker*, Arch. d. M. u. P. (3), **10** (1906), 336.

KLEIN, F., *Bemerkungen zur Theorie der linearen Differentialgleichungen zweiter Ordnung*, Math. Ann., **64** (1907), 175.

PICK, G., *Ueber die zu einer ebenen algebraischen Kurve gehörigen transzendenten Formen und Differentialgleichungen*, Monatshefte f. M. u. P., **18** (1907), 219.

KOEBE, P., *Zur Uniformisierung der algebraischen Kurven*, Gött. Nach. (1907), 410.

FUBINI, G., *Sulla teoria delle funzioni automorfe e delle loro transformazioni*, Ann. di Math. (3), **14** (1907), 33.

POINCARÉ, H., *Sur l'uniformisation des fonctions analytiques*, Acta. Math., **31** (1907), 1.

SCHOTTKY, F., *Ueber Beziehungen zwischen veränderlichen Grössen, die auf gegebene Gebiete beschränkt sind;* Berlin Sitzungsb. (1907), 919; and (1908), 119.

KLEIN, F., *Ueber die Zusammenhang zwischen dem sogenannten Oszillationstheorem der linear Differentialgleichungen und dem Fundamentaltheorem der automorphen Funktionen,* Deutsch. Math. Ver., **16** (1907), 537.

YOUNG, J. W., *On a class of discontinuous ζ-groups defined by normal curves of the fourth order,* Palermo Rend., **23** (1907), 97.

YOUNG, J. W., *A fundamental invariant of the discontinuous ζ-groups defined by the normal curves of order* n *in a space of* n *dimensions,* Bull. Am. M.S. (2), **14** (1908), 363.

ROTHE, H., *Ueber das Grundtheorem und die Obertheoreme der automorphen Funktionen im Falle der Hermite-Laméschen Gleichung mit vier singulären Punkten,* Monat. Math. u. Phys., **19** (1908), 258.

KOEBE, P., *Die Uniformisierung der algebraischen Kurven. Mitteilung eines Grenzübergangs durch iterierendes Verfahren,* Gött. Nach. (1908), 112.

HILB, E., *Ueber Kleinsche Theoreme in der Theorie der linearen Differentialgleichungen,* Math. Ann., **66** (1908), 215; **68** (1909), 24.

GARNIER, R., *Sur les équations différentielles du troisieme ordre dont l'intégrale générale est uniforme,* C.R., **147** (1908), 915.

HILB, E., *Neue Entwicklungen über lineare Differentialgleichungen,* Gött. Nach. (1908), 231; (1909), 230.

POINCARÉ, H., *Sur la réduction des intégrales abéliennes et les fonctions fuchsiennes,* Rend. d. Circolo d. Palermo, **27** (1908), 281.

CASPAR, M., *Ueber die Darstellbarkeit der homomorphen Formenscharen durch Poincaré'sche Zetareihen,* Diss. Tübingen (1908).

FUBINI, G., *Sulla discontinuità propria dei gruppi discontinui,* Atti iv. congresso. intern. matem. (Rome), vol. ii. (1908), 169.

FUBINI, G., *Introduzione alla teoria dei gruppi discontinui e delle funzioni automorfe,* Pisa, 1908.

KOEBE, P., *Ueber die Uniformisierung der algebraischen Kurven* I, Math. Ann., **67** (1909), 145.

RICHMOND, H. W., *On automorphic functions in relation to the general theory of algebraic curves,* Quart. Journ. M., **40** (1909), 258.

YOUNG, J. W., *On the discontinuous ζ-groups defined by rational normal curves in a space of* n *dimensions,* Bull. Am. M.S. (2), **16** (1909), 363.

HILB, E., *Ueber Kleinsche Theoreme in der Theorie der linearen Differentialgleichungen,* Math. Ann., **68** (1910), 24.

POINCARÉ, H., *Sechs Vorträge über ausgewählte Gegenstände aus der reinen Mathematik und mathematischen Physik,* Leipzig (1910).

BIERBACH, L., *Zur Theorie der automorphen Funktionen,* Diss. Göttingen (1910).

CRAIG, C. F., *On a class of hyperfuchsian functions,* Trans. Amer. Math. Soc., **11** (1910), 37.

KOEBE, P., *Ueber die Uniformisierung der algebraischen Kurven durch automorphe Funktionen mit imaginärer Substitutionsgruppe,* Gött. Nach. (1909), 68; (1910), 180.

KOEBE, P., *Ueber die Uniformisierung der algebraischen Kurven* II., Math. Ann., **69** (1910), 1.

KOEBE, P., *Ueber die Uniformisierung beliebiger analytischer Kurven*, Gött. Nach. (1908), 337; (1909), 324.

HILBERT, D., *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, Sechste Mitteilung*, Gött. Nach. (1910), 355.

KOEBE, P., *Ueber die Hilbertsche Uniformisierungsmethode*, Gött. Nach. (1910), 59.

HECKE, E., *Ueber die Konstruktion der Klassenkörper reeller quadratischer Körper mit Hilfe von automorphen Funktionen*, Gött. Nach. (1910), 619.

PRYM, F., and ROST, G., *Theorie der Prymschen Funktionen erster Ordnung*, Leipzig, 1911.

COURANT, R., *Ueber die Anwendung des Dirichletschen Prinzipes auf die Probleme der konformen Abbildung*, Math. Ann., **71** (1911), 145.

CHAZY, J., *Sur l'indétermination des fonctions uniformes au voisinage de leurs coupures*, C.R., **152** (1911), 499.

LILJESTRÖM, A., *Sur les fonctions fuchsiennes à multiplicateurs constants*, Arkiv Matem. Astr. Fys., **6** (1910), 18.

KOEBE, P., *Ueber die Uniformisierung beliebiger analytischer Kurven*, Crelle's J., **138** (1910), 192; **139** (1911), 251.

HECKE, E., *Höhere Modulfunktionen und ihre Anwendung auf die Zahlentheorie*, Math. Ann. **71** (1911), 1.

FRICKE, R., *Zur Transformationstheorie der automorphen Funktionen*, Gött. Nach. (1911), 518; (1912), 240.

BIANCHI, L., *Sul gruppo automorfo delle forme ternarie quadratiche suscettibili di rappresentare lo zero*, Atti Accad. Lincei, (5) **21** (1912), 305.

JOHANSSON, S., *Zur theorie der Konvergenz der Poincaré'schen Reihen bei den Hauptkreisgruppen*, Öfversigt af Finska Vetenskaps-Soc. Förhand., **53** (1911), 15, 25.

JOHANSSON, S., *Om konvergensen af de Poincaré'ska θ-serierna i hufvudcirkelfallet.* 2d Congr. math. Scand. (1911), 181.

KÖNIG, R., *Anwendung der Integralgleichungen auf ein Problem der Theorie der automorphen Funktionen*, Math. Ann., **71** (1911), 206.

BIEBERBACH, L., $\Delta u = e^u$ *und die automorphen Funktionen*, Gött. Nach. (1912), 599.

BROUWER, L. E. J., *Ueber die topologischen Schwierigkeiten des Kontinuitäts beweises der Existenztheoreme eindeutig umkehrbarer polymorpher Funktionen auf Riemannschen Flächen*, Gött. Nach. (1912), 603.

OSGOOD, W. F., *Lehrbuch der Funktionentheorie*, Leipzig (2nd edition), 1912.

BROUWER, L. E. J., *Ueber die Singularitätenfreiheit der Modulmannigfaltigkeit*, Gött. Nach. (1912), 803.

KLEIN, F., BROUWER, L. E. J., KOEBE, P., BIEBERBACH, L., und HILB, E., *Zu den Verhandlungen betreffend automorphen Funktionen*, Deut. Math. Ver., **21** (1912), 153.

KOEBE, P., *Ueber die Uniformisierung der algebraischen Kurven* III, Math. Ann., **72** (1912), 437.

KOEBE, P., *Zur Begründung der Kontinuitätsmethode*, Leip. Ber., **64** (1912), 59.

GARNIER, R., *Sur des équations différentielles du troisième ordre dont l'intégrale générale est uniforme, et sur une classe d'équations nouvelles d'ordre* *supérieur dont l'intégrale générale a ses points critiques fixes*, Ann. de l'Éc. Norm. (3), **29** (1912), 1.

LILIJESTRÖM, A., *Études sur la théorie du potentiel logarithmique*, Arkiv Matem. Astro. Fys., **7** (1912), 39, p. 56.

PLEMELJ, J., *Die Grenzkreisuniformisierung analytischer Gebilde*, Monats. f. M. u. P., **23** (1912), 297.

KOEBE, P., *Begründung der Kontinuitätsmethode im Gebiete der konformen Abbildung und Uniformisierung*, Gött. Nach. (1912), 879.

OSGOOD, W. F., *Existenzbeweis betreffend Funktionen, welche zu einer eigentlich discontinuierlichen automorphen Gruppe gehören*, Palermo Rend., **35** (1913), 103.

OSGOOD, W. F., *On the uniformisation of algebraic functions*, Annal. Math. Princeton, (2) **14** (1913), 143.

FRICKE, R., *Automorphe Funktionen mit Einschluss der elliptischen Modulfunktionen*, Encyclopädie der math. Wiss., Band ii. 2, Leipzig (1913).

PICARD, É., *Sur la représentation conforme des aires multiplement connexes*, Ann. de l'Éc. Norm., (3) **30** (1913), 483.

KOEBE, P., *Ueber die Uniformisierung der algebraischen Kurven IV*, Math. Ann., **75** (1914), 42.

EMCH, A., *Two convergency proofs*, Amer. M. S. Bull., (2) **20** (1914), 358.

EMCH, A., *On some geometric properties of circular transformations*, Amer. Math. Monthly, **21** (1914), 139.

BIEBERBACH, L., *Zur Theorie und Praxis der Konformen Abbildung*, Palermo Rend., **38** (1914), 98.

KOEBE, P., *Zur Theorie der Konformen Abbildung und Uniformisierung*, Leipzig Ber., **66** (1914), 67.

EMCH, A., *A certain class of functions connected with Fuchsian groups*, Amer. M. S. Bull., (2) **22** (1915), 33.

CAHEN, E., *Sur les substitutions fondamentales du groupe modulaire*, S. M. F. Bull., **43** (1915), 69.

FORD, L. R., *An introduction to the theory of automorphic functions*, London, Bell, 1915.

BIEBERBACH, L., *Einführung in die konforme Abbildung*, Berlin, Sammlung Göschen, 1915.

REY-PASTOR, J., *Teoria de la representació conforme*, Barcelona, 1915.

BIEBERBACH, L., $\Delta u = e^u$ *und die automorphen Funktionen*, Math. Ann., **77** (1916), 173.

PICARD, É., *Sur certains sous-groupes hyperfuchsiens correspondant aux formes quadratiques ternaires à indéterminées conjuguées*, C.R., **163** (1916), 284.

PICARD, É., *Sur les fonctions de deux variables complexes restant invariables par les substitutions d'un groupe discontinu*, C.R., **163** (1916), 317.

PICARD, É., *Sur certains groupes discontinus correspondant aux formes quadratiques ternaires*, Ann. de l'Éc. Norm., (3) **33** (1916), 363.

POINCARÉ, H., *Oeuvres*, tome II, Paris, 1916.

GIRAUD, G., *Sur une classe de groupes discontinus de transformations birationelles quadratiques et sur les fonctions de trois variables indépendants restant invariables par ces transformations*, Thèse, Paris, 1916.

MYRBERG, P. J., *Zur Theorie der Konvergenz der Poincaréschen Reihen*, Ann. Acad. Sc. Fennicae, (A) **9**, No. 4; (A) **11**, No. 4 (1916).

KOEBE, P., *Begründung der Kontinuitätsmethode in Gebiete der Konformen Abbildung und Uniformisierung*, Gött. Nach., (1916), 226; (1918), 57.

KOEBE, P., *Abhandlungen zur Theorie der konformen Abbildung:* I. J. für Math., **145** (1915), 177; II. Acta Math., **40** (1916), 251; III. J. für Math., **147** (1917), 67; IV. Acta Math., **41** (1918), 305; V. Math. Zs., **2** (1918), 198.

GIRAUD, G., *Sur les groupes des transformations semblables arithmétiques*, Ann. de l'Éc. Norm., (3) **33** (1916), 303.

LICHTENSTEIN, L., *Die Methode des Bogenelementes in der Theorie der Uniformisierungstranszendenten mit Grenz- oder Hauptkreis (Vorläufige Mittenilung)*, Gött. Nach., (1917), 141, 426.

HILB, E., *Zur Topologie der für die linearen Differentialgleichungen zweiter Ordnung geltenden Obertheoreme*, Gött. Nach., (1917), 112.

FALCKENBERG, H., *Zur Theorie der Kreisbogenpolygone*, Math. Ann., **77** (1916), 65; **78** (1917), 234.

KOEBE, P., *Zur Geometrie der automorphen Fundamentalgruppen*, Gött. Nach., (1918), 54.

GIRAUD, G., *Sur les fonctions hyperfuchsiennes et sur les systèmes d'équations aux différentielles totals*, C.R., **164** (1917), 386, 487; **166** (1918), 24.

TMIRNOV, T., *Application du principe de convergence à la théorie d'uniformisation*, Comm. Soc. Math. de Kharkof, **16** (1918), 39.

BIEBERBACH, L., *Über die Einordnung des Hauptsatzes der Uniformisierung in die Weierstrassische Funktionentheorie*, Math. Ann., **78** (1918), 312.

PRICE, H. F., *Fundamental regions for certain finite groups in S_4*, American J., **40** (1918), 108.

FORSYTH, A. R., *Functions of a complex variable*, 3rd Ed., Cambridge Univ. Press., 1918.

LEWICKYJ, W., *Einige Typen der zur Fuchsschen Gruppe gehörigen Funktionen*, Sammelschrift der math.-nat.-ärzt. Lemberg, **18–19** (1919), 103.

VAN VLECK, E. B., *On the combination of non-loxodromic substitutions*, Trans. Amer. Math. Soc., (4) **20** (1919), 299.

GIRAUD, G., *Sur les fonctions automorphes d'un nombre quelconque de variables*, Ann. de l'Éc. Norm., (3) **36** (1919), 187.

GIRAUD, G., *Sur la classification des substitutions de certains groupes automorphes à n variables, et sur les relations algébriques qui existent entre $n+1$ fonctions quelconques correspondant à certains de ces groupes*, C.R., **169** (1919), 131.

MYRBERG, P. J., *Ein Satz über die polymorphen Funktionen auf den eindeutig in sich transformierbaren Riemannschen Flächen*, Ann. Acad. Sc. Fennicae, (A) **12** (1919), No. 8.

MYRBERG, P. J., *Über das identische Verschwinden der Poincaréschen Reihen*, Ann. Acad. Sc. Fennicae, (A) **13** (1919), No. 4.

FUETER, R., *Über die Konstruktion einer speziellen automorphen Funktion*, Zurich Naturf. Ges., **64** (1919), 1.

HUMBERT, G., *Sur la formation du domaine fondamental d'un groupe automorphe*, C.R., **169** (1919), 205.

GIRAUD, G., *Leçons sur les fonctions automorphes. Fonctions automorphes de n variables, fonctions de Poincaré*, Paris, 1920.

GIRAUD, G., *Réponse à une Note de M. Fubini sur les fonctions automorphes*, C.R., **171** (1920), 1365.

BIRKHOFF, G. D., *The work of Poincaré on automorphic functions*, Amer. M. S. Bull., **26** (1920), 164.

SMIRNOFF, V., *Sur quelques points de la théorie des équations differentielles linéaires du second ordre et des fonctions automorphes*, C.R., **171** (1920), 510.

MYRBERG, P. J., *Über die numerische Ausführung der Uniformisierung*, Acta Soc. Sc. Fennicae, **48** (1920), No. **7**.

FUBINI, G., *Sur les fonctions automorphes*, C.R., **171** (1920), 156; **172** (1921), 265.

Poincaré memorial articles: *Analyse de ses travaux scientifiques, lettres*, etc., Acta Math., **38** (1921), 3.

FUHR, H., *Zur Transformationstheorie der Fuchsschen Funktionen*, Mitt. des Math. Sem. d. Giessen, 1 Heft (1921).

UHLER, A., *Sur les séries zétafuchsiennes*, Thèse, Lund, Gleerupska Univ.-Bokh. (1921).

FALCKENBERG, H., *Ableitung der "Ergänzungsrelationen" aus den Formeln von Simon L'Huilier*, Math. Zs., **10** (1921), 17.

GIRAUD, G., *Sur certains fonctions automorphes de deux variables*, Ann. de l'Éc. Norm., **38** (1921), 43.

BIEBERBACH, L., *Neuere Untersuchungen über Funktionen von komplexen Variablen*, Encyk. d. math. Wiss., II C 4, 379, 1921.

GIRAUD, G., *Sur les fonctions automorphes*, C.R., **172** (1921), 354.

SMIRNOFF, V., *Sur les équations différentielles linéaires du second ordre et la théorie des fonctions automorphes*, Bull. des Sci. Math., (2) **45** (1921), 93, 126.

KLEIN, F., *Gesammelte mathematische Abhandlungen*, 3 vols, Berlin, Springer, 1921–1923.

MAURER, L., *Über die Schottkysche Gruppe von linearen Substitutionen*, Tübinger naturwissensch. Abh., Heft 3 (1921), 1.

MYRBERG, P. J., *Sur les fonctions automorphes de plusieurs variables indépendantes*, C.R., **174** (1922), 1402.

MYRBERG, P. J., *Über die automorphen Funktionen zweier Veränderlichen*, Acta Math., **43** (1922), 289.

MYRBERG, P. J., *Sur les singularités des fonctions automorphes*, C.R., **175** (1922), 674, 809.

FALCKENBERG, H., *Zur Theorie der Kleinschen Ergänzungsrelationen*, Math. Ann., **88** (1922), 123.

GOT, TH., *Questions diverses concernant certaines formes quadratiques ternaires indéfinies et les groupes fuchsiens arithmétiques qui s'y rattachent*, Ann. de Toulouse, (3) **5** (1922), 1.

AXEL, H., *Die elliptischen Funktionen als automorphe Funktionen die zu einer aus drei Drehungen der Euklidischen Ebene um* 180° *Komponierten Gruppe gehören*, Mitt. Math. Sem. Giessen, **1** (1922), 1.

POINCARÉ, H., *Sur les fonctions fuchsiennes. Extrait d'un mémoire inédit*, Acta Math., **39** (1923), 58.

Correspondance d'Henri Poincaré et de Felix Klein, Acta Math., **39** (1923), 94.

SANSONE, G., *I sottogruppe del gruppo di Picard e due teoremi sui gruppi finiti analoghi at teorema del Dyck*, Palermo Rend., **47** (1923), 273.

MYRBERG, P. J., *Uber die Singularitäten der automorphen Funktionen mehrerer Veränderlichen*, 5 Kongress Skandinav. Math., (1923), 297.

MYRBERG, P. J., *Über die automorphen Funktionen mehrerer Veränderlichen*, Math. Ann., **93** (1924), 61.

SPIESZ, O., *Über automorphe Funktionen mit rationalem Multikationstheorem*, Math. Ann., **93** (1924), 98.

MYRBERG, P. J., *Ueber die automorphen Funktionen bei einer Klasse Jonquièresscher Gruppen zweier Veränderlichen*, Math. Zeit., **21** (1924), 224.

MYRBERG, P. J., *Ueber die wesentlichen Singularitäten der automorphen Funktionen mehrerer Veränderlichen*, Ann. Acad. Sc. Fennicae, (A) **20** (1924).

DALAKER, H. H., *On the automorphic functions of the group* (0, 3; 2, 4, 6), Ann. of Math., (2) **25** (1924), 241.

SCHLESINGER, L., *Automorphe Funktionen*, Leipzig, W. de Gruyter, 1924.

UHLER, A., *Sur une généralisation des fonctions zêtafuchsiennes*, C.R. 6 Cong des Math. Scan., Copenhagen, (1925), 489.

MYRBERG, P. J., *Über diskontinuierliche Gruppen und automorphe Funktionen von mehreren Variablen*, C.R. 6 Cong. des Math. Scan., Copenhagen, (1925), 191.

HECKE, E., *Ueber einen neuen Zusammenhang zwischen elliptischen Modulfunktionen und indefiniten quadratischen Formen*, Gött. Nach., (1925), 35.

SANSONE, G., *La relazioni fondamentali fra le operazioni generatrici del gruppo modulare finito con coefficienti inleri del campo di Gauss*, Palmero Rend., **49** (1925), 225.

FORD, L. R., *The fundamental region for a Fuchsian group*, Amer. M. S. Bull., **31** (1925), 531.

MYRBERG, P. J., *Sur les fonctions automorphes*, C.R., **180** (1925), 1819.

ABRAMOWICZ, C., *Sur la transformation du p-ième degré d'une fonction automorphe*, Ann. de Soc. Polonaise de Math., **3** (1925), 63.

LEWENT, L., *Conformal Representation*, trans. by R. Jones and D. H. Williams, Methuen, 1925.

SANSONE, G., *I sottogruppi del gruppo modulaire con coefficienti del corpo di Jacobi-Eisenstein e un theorema sui gruppi finiti*, Annali di Math., (4) **3** (1926), 73.

BIEBERBACH, L., *Lehrbuch der Funktionentheorie II*, Berlin and Leipzig, Teubner, 1927.

FUETER, R., *Ueber automorphe Funktionen in Bezug auf Gruppen, die in der Ebene uneigentlich diskontinuerlich sind*, Jour. für Math., **157** (1927), 66.

FORD, L. R., *On the foundations of the theory of discontinuous groups of linear transformations*, Proc. Natl. Acad. Sci., **13** (1927), 286.

FORD, L. R., *On the formation of groups of linear transformations by combination*, Proc. Natl. Acad. Sci., **13** (1927), 289.

KOEBE, P., *Allgemeine Theorie der Riemannschen Mannigfaltigkeiten* (*Konforme Abbildung und Uniformisierung*), Acta Math., **50** (1927), 27.

NEVANLINNA, F., *Über die Anwendung einer Klasse uniformisierenden Tranzendenten zur Untersuchung der Wertverteilung analytischer Funktionen*, Acta Math., **50** (1927), 159.

HECKE, E., *Zur Theorie der elliptischen Modulfunktionen*, Math. Ann., **97** (1927), 210.

FRICKE, R., *Transformation der elliptischen Funktionen bei zusammengesetztem Transformationsgrade*, Math. Ann., **101** (1929), 316.

AUTHOR INDEX

SUBJECT INDEX

G

H

I

J

K

L