10

ELEMENTARY MEROMORPHIC FUNCTIONS

43. Rational Functions

In the preceding chapter, we became acquainted with some of the simplest members of the class of entire functions. A more general class of functions is the class of meromorphic functions. By a *meromorphic function* we mean a function which can be represented as a ratio of two entire functions.[1] Obviously, every entire function $f(z)$ is meromorphic, since it can be written in the form $f(z)/1$, but the converse is not true, as shown by the function $1/z$ which is meromorphic but not entire, since it becomes infinite for $z = 0$.

The simplest members of the class of meromorphic functions in the full sense of the term (i.e., meromorphic functions which do not reduce to entire functions) are the rational functions. By a *rational function* we mean a function which can be written as a ratio

$$f(z) = \frac{P(z)}{Q(z)} = \frac{a_0 + a_1 z + \cdots + a_m z^m}{b_0 + b_1 z + \cdots + b_n z^n} \qquad (a_m \neq 0, \quad b_n \neq 0) \quad (10.1)$$

of two polynomials $P(z)$ and $Q(z)$, where it is assumed that the fraction $P(z)/Q(z)$ is in *lowest terms*, i.e., that the equations $P(z) = 0$ and $Q(z) = 0$ have no common roots. Let $\alpha_1, \ldots, \alpha_p$ denote the distinct zeros of the polynomial $P(z)$, and let α_s be of order k_s ($s = 1, \ldots, p$). Similarly, let

[1] The word *meromorphic* stems from the Greek μερος = *fraction* and μορφή = *form*, and means "like a fraction."

$\beta_1, \ldots, \beta_q$ denote the distinct zeros of the polynomial $Q(z)$, and let β_t be of order l_t ($t = 1, \ldots, q$). Then (10.1) can be written in the form

$$f(z) = \frac{P(z)}{Q(z)} = \frac{a_m(z - \alpha_1)^{k_1} \cdots (z - \alpha_p)^{k_p}}{b_n(z - \beta_1)^{l_1} \cdots (z - \beta_q)^{l_q}}.$$

Obviously

$$k_1 + \cdots + k_p = m, \qquad l_1 + \cdots + l_q = n,$$

and none of the numbers $\alpha_1, \ldots, \alpha_p$ equals any of the numbers $\beta_1, \ldots, \beta_q$, since $f(z)$ is assumed to be in lowest terms. The function $f(z)$ vanishes at each point α_s ($s = 1, \ldots, p$), called a *zero* (or *zero-point*) *of order* k_s [of $f(z)$], and becomes infinite at each point β_t ($t = 1, \ldots, q$), called a *pole of order* l_t. The zero α_s is said to be *simple* if $k_s = 1$ and *multiple* if $k_s > 1$; similarly, the pole β_t is said to be *simple* if $l_t = 1$ and *multiple* if $l_t > 1$. It follows from our definition that the zeros of $f(z)$ are poles of $1/f(z)$, while the poles of $f(z)$ are zeros of $1/f(z)$, and moreover that orders are preserved, i.e., a zero of $f(z)$ of a given order becomes a pole of $1/f(z)$ of the same order, and *vice versa*.

To define $f(z)$ when z is the point at infinity, we use the relation

$$f(\infty) = \lim_{z \to \infty} f(z),$$

obtaining

$$\begin{array}{lll} 1) & f(\infty) = 0 & \text{if } m < n; \\ 2) & f(\infty) = \dfrac{a_m}{b_n} & \text{if } m = n; \\ 3) & f(\infty) = \infty & \text{if } m > n. \end{array} \qquad (10.2)$$

In the first case, we say that $f(z)$ has a *zero at* ∞, and in the third case we say that $f(z)$ has a *pole at* ∞. To assign a definite order to a zero or pole at ∞, we make the preliminary transformation $\zeta = 1/z$ (cf. Sec. 24), which carries the point $z = \infty$ into the point $\zeta = 0$. Then

$$f^*(\zeta) = f(1/\zeta) = \frac{a_0 + a_1 \dfrac{1}{\zeta} + \cdots + a_m \dfrac{1}{\zeta^m}}{b_0 + b_1 \dfrac{1}{\zeta} + \cdots + b_n \dfrac{1}{\zeta^n}} = \frac{\zeta^n}{\zeta^m} \frac{a_m + a_{m-1}\zeta + \cdots + a_0 \zeta^m}{b_n + b_{n-1}\zeta + \cdots + b_0 \zeta^n},$$

and analyzing the three cases (10.2), we obtain the following results:

1. If $m < n$, then

$$f^*(\zeta) = \frac{\zeta^{n-m}(a_m + a_{m-1}\zeta + \cdots + a_0 \zeta^m)}{b_n + b_{n-1}\zeta + \cdots + b_0 \zeta^n}.$$

Since this rational function has a zero of order $n - m$ at $\zeta = 0$, we say that $f(z)$ has a zero of order $n - m$ at $z = \infty$.

2. If $m = n$, then

$$f^*(\zeta) = \frac{a_m + a_{m-1}\zeta + \cdots + a_0\zeta^m}{b_m + b_{m-1}\zeta + \cdots + b_0\zeta^m}.$$

Since this rational function has neither a zero nor a pole at $\zeta = 0$ [in fact, $f^*(0) = f(\infty) = a_m/b_m$], we say that $f(z)$ has neither a pole nor a zero at $z = \infty$.

3. If $m > n$, then

$$f^*(\zeta) = \frac{a_m + a_{m-1}\zeta + \cdots + a_0\zeta^m}{\zeta^{m-n}(b_n + b_{n-1}\zeta + \cdots + b_0\zeta^n)}.$$

Since this rational function has a pole of order $m - n$ at $\zeta = 0$, we say that $f(z)$ has a pole of order $m - n$ at $z = \infty$.

Remark. It is easy to see that $f(z)$ is analytic at $z = \infty$ (see Remark 1, p. 125) if $m < n$ or $m = n$, but not if $m > n$.

By the *order* N of a rational function $f(z)$, we mean the total number of zeros (or poles) of the function in the extended plane, counting each zero (or pole) a number of times equal to its order. It is easy to see that

$$N = \max(m, n),$$

i.e., N is the larger of the two integers m and n (or their common value if $m = n$). In fact, consider the three possibilities in turn:

1. If $m < n$, $f(z)$ has $k_1 + \cdots + k_p = m$ zeros at finite points and a zero of order $n - m$ at infinity, so that $f(z)$ has a total of

$$N = m + (n - m) = n = \max(m, n)$$

zeros. Moreover, $f(z)$ has $l_1 + \cdots + l_q = n$ poles at finite points and no pole at infinity, so that $f(z)$ also has a total of $N = n$ poles.

2. If $m = n$, $f(z)$ has neither a zero nor a pole at infinity, and precisely

$$m = k_1 + \cdots + k_p = l_1 + \cdots + l_q = m$$

zeros and poles at finite points.

3. If $m > n$, $f(z)$ has $l_1 + \cdots + l_q = n$ poles at finite points and a pole of order $m - n$ at infinity, so that $f(z)$ has a total of

$$N = n + (m - n) = m = \max(m, n)$$

poles. Moreover, $f(z)$ has $k_1 + \cdots + k_p = m$ zeros at finite points and no zero at infinity, so that $f(z)$ also has a total of $N = m$ zeros.

Now let A be an arbitrary complex number, and consider the roots of the equation

$$f(z) = \frac{P(z)}{Q(z)} = A, \tag{10.3}$$

or equivalently, the roots of the equation

$$F(z) = \frac{P(z) - AQ(z)}{Q(z)} = 0. \tag{10.4}$$

By the multiplicity of a root z_0 of equation (10.3), we mean the multiplicity of the same root z_0 of equation (10.4).

THEOREM 10.1. *Given a rational function $f(z)$ of order N and an arbitrary complex number A, the total number of roots of equation (10.3) is N (with due regard for multiplicity).*

Proof. If $A = 0$, or $A = \infty$, the result follows directly from the definition of the order of $f(z)$. Let r be the degree of the polynomial $P(z) - AQ(z)$ and s the order of the rational function $F(z)$. If $A \neq 0$, $A \neq \infty, m \neq n$, then $r = \max(m, n) = N$, and hence $s = \max(p, n) = N$. If $A \neq 0$, $A \neq \infty$, $m = n$, then $r \leqslant \max(m, n) = N$, and hence $s = \max(p, n) = N$.

Remark. As we know from Sec. 35, if $P_n(z)$ is a polynomial of degree n, the equation $P_n(z) = 0$ has exactly n roots (with due regard for multiplicity). Thus we see that this same property holds for rational functions, with the concept of *order* replacing that of *degree*.

Example. The rational function

$$f(z) = \frac{z^2 + 1}{z^2 - 1}$$

has two simple zeros $\pm i$ and two simple poles ± 1. If $A = 1$, the equation $f(z) = A$ becomes

$$\frac{z^2 + 1}{z^2 - 1} = 1$$

or

$$\frac{2}{z^2 - 1} = 0, \tag{10.5}$$

which obviously has no finite roots. However, (10.5) has a double root at ∞, since the degree of the denominator is two higher than that of the numerator.

It follows from Theorem 10.1 that *a rational function $w = f(z)$ of order $N > 0$ maps the extended plane onto itself in such a way that every point w has at most N distinct inverse images* (cf. the analogous property of polynomials, given on p. 129). Moreover, only certain exceptional values of w can have fewer than N distinct inverse images [e.g., the point $w = 0$ if $f(z)$ has multiple zeros, or the point $w = \infty$ if $f(z)$ has multiple poles]. The following result is the natural generalization of Theorem 9.1:

THEOREM 10.2. *Let $f(z)$ be a rational function whose numerator $P(z)$ is of degree m and whose denominator $Q(z)$ is of degree n. Then there are at most $m + n$ points in the extended w-plane with fewer than $N = \max(m, n)$ distinct inverse images ($N > 1$).*

Proof. First suppose $A \neq 0$, $A \neq \infty$, and suppose that A has fewer than N distinct inverse images. Then the equation

$$f(z) = \frac{P(z)}{Q(z)} = A,$$

or equivalently

$$\frac{P(z) - AQ(z)}{Q(z)} = 0, \tag{10.6}$$

must have a multiple root. Any finite multiple root of (10.6) is a multiple root of the equation

$$P(z) - AQ(z) = 0, \tag{10.7}$$

and conversely. Moreover, any multiple root of (10.7) satisfies the equation

$$P'(z) - AQ'(z) = 0,$$

and hence also satisfies the equation

$$P'(z)Q(z) - P(z)Q'(z) = 0, \tag{10.8}$$

of degree no higher than $m + n - 1$. Equation (10.8) has no more than $m + n - 1$ distinct roots $\gamma_1, \ldots, \gamma_r$ ($1 \leqslant r \leqslant m + n - 1$). Since any finite multiple zero of $f(z)$ satisfies the equations

$$P(z) = 0, \qquad P'(z) = 0,$$

and since any finite multiple pole of $f(z)$ satisfies the equations

$$Q(z) = 0, \qquad Q'(z) = 0,$$

all the finite multiple roots of (10.6) satisfy (10.8), and are therefore already included among the numbers $\gamma_1, \ldots, \gamma_r$, even if we let $A = 0$ or $A = \infty$. In other words, the numbers

$$f(\gamma_1), \ldots, f(\gamma_r), f(\infty) \qquad (1 \leqslant r \leqslant m + n - 1)$$

are the only values of A for which the equation $f(z) = A$ can have a multiple root, where we finally allow for the possibility that $z = \infty$ is a multiple root. Since at most $m + n$ of these numbers are distinct, the theorem is proved.[2]

[2] If $N = 0$ or $N = 1$, the theorem is meaningless.

Remark. The mapping $w = f(z)$ is conformal at all but a finite number of points, since the condition that z should not equal ∞, any of the numbers $\gamma_1, \ldots, \gamma_r$ or any of the zeros of $Q(z)$ is certainly sufficient to guarantee that the derivative

$$f'(z) = \frac{P'(z)Q(z) - P(z)Q'(z)}{[Q(z)]^2}$$

is finite and nonzero.[3]

44. The Group Property of Möbius Transformations

It follows from Theorems 10.1 and 10.2 that the rational function of order 1, i.e., the *fractional linear transformation* or *Möbius transformation*

$$w = L(z) = \frac{az + b}{cz + d} \qquad (ad - bc \neq 0)$$

is the only rational function which maps the extended plane onto itself in a one-to-one fashion. Here, we assume that the quantity $ad - bc$, called the *determinant* of the function $L(z)$, is nonzero, since otherwise $L(z)$ is identically equal to a constant mapping the extended plane into a single point. As we know from Sec. 34 (see also Problem 8.9), the mapping $w = L(z)$ is conformal at all points of the extended plane. This mapping (or its various special cases) plays an important role in a variety of problems encountered in the theory of functions of a complex variable, and it therefore merits a detailed study (to which we devote most of this chapter).

Let $\mathcal{M}$ denote the set of all Möbius transformations. Two such transformations

$$L_1(z) = \frac{a_1 z + b_1}{c_1 z + d_1}, \qquad L_2(z) = \frac{a_2 z + b_2}{c_2 z + d_2} \tag{10.9}$$

are regarded as *identical* if and only if $L_1(z) = L_2(z)$ for all z.

THEOREM 10.3. *A necessary and sufficient condition for the two Möbius transformations* (10.9) *to be identical is that*

$$a_2 = \lambda a_1, \quad b_2 = \lambda b_1, \quad c_2 = \lambda c_1, \quad d_2 = \lambda d_1 \qquad (\lambda \neq 0). \tag{10.10}$$

Proof. The sufficiency of the condition is obvious. To prove the necessity, suppose $L_1(z) = L_2(z)$. Then, in particular,

$$L_1(0) = L_2(0), \quad L_1(1) = L_2(1), \quad L_1(\infty) = L_2(\infty),$$

which means that[4]

$$\frac{b_1}{d_1} = \frac{b_2}{d_2} = p, \quad \frac{a_1 + b_1}{c_1 + d_1} = \frac{a_2 + b_2}{c_2 + d_2}, \quad \frac{a_1}{c_1} = \frac{a_2}{c_2} = q. \tag{10.11}$$

[3] However, the condition is not necessary. (For further details, see Problems 10.1 and 10.2.)

[4] Here we allow p or q to take the improper value ∞.

Substituting

$$b_1 = pd_1, \quad b_2 = pd_2, \quad a_1 = qc_1, \quad a_2 = qc_2$$

into the second of the equations (10.11), we obtain

$$\frac{qc_1 + pd_1}{c_1 + d_1} = \frac{qc_2 + pd_2}{c_2 + d_2}$$

or

$$(c_1d_2 - c_2d_1)(q - p) = 0.$$

But $q \neq p$, since otherwise

$$\frac{a_1}{c_1} = \frac{b_1}{d_1}, \qquad \text{i.e.,} \quad a_1d_1 - b_1c_1 = 0,$$

contrary to hypothesis, and therefore

$$\frac{c_1}{d_1} = \frac{c_2}{d_2}.$$

Together with (10.11), this implies (10.10), as required.

Remark. A Möbius transformation is not characterized by the value of its determinant, since the determinant is multiplied by λ^2 when the coefficients change as described by (10.10). It can only be asserted that the determinant remains nonzero under any substitution (10.10).

The transformation

$$U(z) = z,$$

which obviously belongs to the set $\mathscr{M}$, is called the *unit transformation* (or the *identity transformation*). By the inverse of a given transformation

$$w = \frac{az + b}{cz + d}, \tag{10.12}$$

we mean the transformation which assigns to each w its inverse image z under the transformation (10.12). Thus the transformation

$$z = \frac{dw - b}{-cw + a}$$

(whose coefficients are unique only to within a multiplicative constant, as in Theorem 10.3) is the inverse of the transformation (10.12). We denote the inverse of the transformation L by L^{-1}.

Given two arbitrary Möbius transformations

$$L_1(z) = \frac{a_1z + b_1}{c_1z + d_1}, \qquad L_2(z) = \frac{a_2z + b_2}{c_2z + d_2},$$

we define the *product* of L_1 and L_2 as the result of first carrying out one transformation and then carrying out the other. There are two possible

products corresponding to the two orders in which the transformations can be carried out. One product, written $L_1L_2(z)$, equals

$$\begin{aligned} L_1L_2(z) &= \frac{a_1[(a_2z + b_2)/(c_2z + d_2)] + b_1}{c_1[(a_2z + b_2)/(c_2z + d_2)] + d_1} \\ &= \frac{a_1(a_2z + b_2) + b_1(c_2z + d_2)}{c_1(a_2z + b_2) + d_1(c_2z + d_2)} \\ &= \frac{(a_1a_2 + b_1c_2)z + (a_1b_2 + b_1d_2)}{(c_1a_2 + d_1c_2)z + (c_1b_2 + d_1d_2)}, \end{aligned} \tag{10.13}$$

and the other, written $L_2L_1(z)$, is obtained from (10.13) by permuting the indices 1 and 2. Moreover, since

$$\begin{aligned} (a_1a_2 + b_1c_2)(c_1b_2 + d_1d_2) - (a_1b_2 + b_1d_2)(c_1a_2 + d_1c_2) \\ = (a_1d_1 - b_1c_1)(a_2d_2 - b_2c_2) \neq 0, \end{aligned}$$

each of the transformations $L_1L_2(z)$ and $L_2L_1(z)$ belongs to the set $\mathscr{M}$. In general, $L_1L_2(z) \neq L_2L_1(z)$, but obviously

$$LL^{-1}(z) = L^{-1}L(z) = U(z),$$

and

$$LU(z) = UL(z) = L(z)$$

for any $L(z) \in \mathscr{M}$.

Example. If

$$L_1(z) = \frac{z}{z + 1}, \qquad L_2(z) = \frac{z + 1}{z - 1},$$

then

$$L_1L_2(z) = \frac{z + 1}{2z}, \qquad L_2L_1(z) = -2z - 1.$$

Multiplication of transformations, as just defined, is an *associative* operation, i.e.,

$$(L_1L_2)L_3(z) = L_1(L_2L_3)(z). \tag{10.14}$$

To see this, we merely write $z_3 = L_3(z)$, and then both sides of (10.14) reduce at once to $L_1L_2(z_3)$. This associative property generalizes immediately to the product of an arbitrary number of transformations, and makes it unnecessary to use parentheses when writing products. For example, we have

$$L_1[L_2(L_3L_4)](z) = L_1L_2(L_3L_4)(z) = L_1(L_2L_3)L_4(z) = \cdots = L_1L_2L_3L_4(z).$$

Thus we have shown that the set $\mathscr{M}$ of Möbius transformations has the following properties:[5]

[5] Henceforth, for simplicity, we shall often omit the argument z, writing L_1 instead of $L_1(z)$, U instead of $U(z)$, etc.

1. $\mathcal{M}$ *is closed under multiplication*, i.e., if $L_1 \in \mathcal{M}$, $L_2 \in \mathcal{M}$, then $L_1L_2 \in \mathcal{M}$, $L_2L_1 \in \mathcal{M}$.
2. Multiplication is associative.
3. There is an element $U \in \mathcal{M}$ such that $LU = UL = L$ for any $L \in \mathcal{M}$.
4. For each $L \in \mathcal{M}$, there is an element $L^{-1} \in \mathcal{M}$ such that $LL^{-1} = L^{-1}L = U$.

In algebraic language, these four properties are summarized by saying that $\mathcal{M}$ *is a group of transformations.*[6]

45. The Circle-Preserving Property of Möbius Transformations

We now prove that any Möbius transformation carries a straight line or a circle into another straight line or circle. We call this the *circle-preserving property*, since a straight line can be regarded as a limiting case of a circle (corresponding to infinite radius). The entire linear transformation $L(z) = \alpha z + \beta$ ($\alpha \neq 0$) is obviously circle-preserving, since the mapping $w = L(z)$ is just a shift (if $\alpha = 1$), or a shift combined with a rotation and a uniform magnification (if $\alpha \neq 1$), as discussed in Sec. 33.

LEMMA. *The transformation*

$$w = \Lambda(z) = \frac{1}{z} \tag{10.15}$$

is circle-preserving.

Proof. The equation of any straight line or circle in the z-plane can be written in the form

$$A(x^2 + y^2) + 2Bx + 2Cy + D = 0, \tag{10.16}$$

where we have a straight line if $A = 0$ and at least one of the numbers B, C is nonzero, and a circle if $A \neq 0$ and $B^2 + C^2 - AD > 0$. Since

$$x^2 + y^2 = z\bar{z}, \quad 2x = z + \bar{z}, \quad 2y = -i(z - \bar{z}),$$

where $\bar{z} = x - iy$ is the complex conjugate of $z = x + iy$, we can rewrite (10.16) as

$$Az\bar{z} + \bar{E}z + E\bar{z} + D = 0, \tag{10.17}$$

where $E = B + iC$. It is easy to see that equation (10.17), where A and D are real and E is complex, is the equation of a straight line if and only if $A = 0$, $E \neq 0$, and the equation of a circle if and only if $A \neq 0$, $E\bar{E} - AD > 0$.

[6] See e.g., G. Birkhoff and S. MacLane, *op. cit.*, Chap, 6, Sec. 2, or V. I. Smirnov, *Linear Algebra and Group Theory* (translated by R. A. Silverman), McGraw-Hill Book Co., New York (1961), Sec. 62.

We now find the image of the curve with equation (10.17) under the transformation (10.15). Replacing z by $1/w$ in (10.17), we obtain

$$A\frac{1}{w\bar{w}} + \bar{E}\frac{1}{w} + E\frac{1}{\bar{w}} + D = 0$$

or

$$Dw\bar{w} + Ew + \bar{E}\bar{w} + A = 0. \tag{10.18}$$

Equation (10.18) has the same form as equation (10.17), with D, $\bar{E}$ and A substituted for A, E and D, respectively. It follows that (10.18) is the equation of a straight line if $D = 0$, since then either $A = 0$ and $E \neq 0$ if (10.17) is the equation of a straight line, or else $A \neq 0$ and $E\bar{E} - AD = E\bar{E} > 0$ (so that $E \neq 0$ again) if (10.17) is the equation of a circle. Moreover, (10.18) is the equation of a circle if $D \neq 0$, since then either $A \neq 0$ and $E\bar{E} - AD > 0$ if (10.17) is the equation of a circle, or else $A = 0$ and $E \neq 0$ (so that $E\bar{E} - AD = E\bar{E} > 0$ again) if (10.17) is the equation of a straight line.

THEOREM 10.4. *Every Möbius transformation*

$$w = L(z) = \frac{az + b}{cz + d} \qquad (ad - bc \neq 0) \tag{10.19}$$

is circle-preserving.

Proof. If $c = 0$, (10.19) reduces to an entire linear transformation and hence is circle-preserving. If $c \neq 0$, (10.19) can be written in the form

$$w = \frac{a}{c} + \frac{bc - ad}{c(cz + d)}.$$

Setting

$$z_1 = L_1(z) = cz + d, \quad z_2 = \Lambda(z_1) = \frac{1}{z_1},$$
$$w = L_2(z_2) = \frac{a}{c} + \frac{bc - ad}{c} z_2,$$

we can write $L(z)$ as a product

$$L = L_2\Lambda L_1$$

of three transformations which are all circle-preserving (use the lemma). It follows that L itself is circle-preserving.

COROLLARY. *Let $\delta = -d/c$ be the pole of the function (10.19). Then (10.19) transforms every straight line or circle which passes through δ into a straight line, and every other straight line or circle into a circle.*

Proof. If the circle or straight line passes through δ, its image under (10.19) contains the point at infinity, and hence must be a straight line, since it cannot be a circle (no circle contains ∞). Similarly, if the circle or straight line does not pass through δ, its image does not contain the point at infinity, and hence must be a circle, since it cannot be a straight line (every straight line contains ∞).

Remark. Let $w = L(z)$ be any Möbius transformation, let γ be a straight line or circle in the z-plane, and let $\Gamma = L(\gamma)$ be the image of γ in the w-plane (Γ is itself a straight line or a circle). The two domains G_1 and G_2 with boundary γ are either two half-planes or the interior and exterior of a circle. Let $L(G_1)$ and $L(G_2)$ be the images of these two domains under the mapping $w = L(z)$. We now show that $L(G_1)$ and $L(G_2)$ are the two domains whose common boundary is the curve Γ.[7]

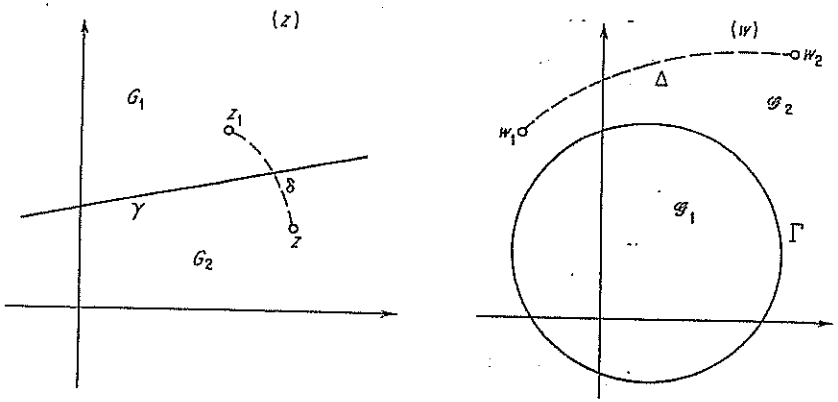

FIGURE 10.1

First suppose $z_1 \in G_1$, $z_2 \in G_2$, and let $w_1 = L(z_2)$, $w_2 = L(z_2)$. Then $w_1 \notin \Gamma$, $w_2 \notin \Gamma$, since $z_1 \notin \gamma$, $z_2 \notin \gamma$, and hence w_1 and w_2 must belong to the union of the two (disjoint) domains into which Γ divides the extended w-plane. If w_1 and w_2 both belong to one of these two domains, we can join w_1 to w_2 by a line segment or circular arc Δ which does not intersect Γ (see Figure 10.1). The inverse image of Δ in the z-plane must be a line segment or circular arc δ, which joins z_1 to z_2 and does not intersect γ. But the existence of δ contradicts the assumption that z_1 and z_2 belong to different domains G_1 and G_2. Therefore, if z_1 and z_2 belong to different domains with boundary γ, their images w_1 and w_2 must belong to different domains with boundary Γ.

We now denote the domains containing w_1 and w_2 by $\mathscr{G}_1$ and $\mathscr{G}_2$, respectively. If z is an arbitrary point in G_1, then, since z and z_2 belong to different domains G_1 and G_2, their images w and w_2 belong to different

[7] Of course, this result follows at once from Theorem 6.3, knowledge of which is not presupposed here.

domains $\mathscr{G}_1$ and $\mathscr{G}_2$. But $w_2 \in \mathscr{G}_2$, and hence $w \in \mathscr{G}_1$, i.e., $L(z) \in \mathscr{G}_1$, if $z \in G_1$. Similarly, $L(z) \in \mathscr{G}_2$ if $z \in G_2$, and hence

$$\mathscr{G}_1 \supset L(G_1), \qquad \mathscr{G}_2 \supset L(G_2). \tag{10.20}$$

Conversely, let w be an arbitrary point in $\mathscr{G}_1$. Then w must be the image of a point z in G_1 or G_2. But $z \in G_2$ implies $w \in \mathscr{G}_2$, contrary to hypothesis, and hence $z \in G_1$, i.e., $\mathscr{G}_1 \subset L(G_1)$. Similarly, we find that $\mathscr{G}_2 \subset L(G_2)$. It follows by comparison with (10.20) that

$$\mathscr{G}_1 = L(G_1), \qquad \mathscr{G}_2 = L(G_2),$$

i.e., the two domains with boundary Γ are just the images of the two domains G_1 and G_2, as asserted. Moreover, to determine which of the two domains with boundary Γ is actually the image of a given domain G_1 with boundary γ, it is sufficient to locate the image w_1 of any point $z_1 \in G_1$, for then the domain $\mathscr{G}_1$ containing w_1 is the image of G_1.

46. Fixed Points of a Möbius Transformation. Invariance of the Cross Ratio

By a *fixed point* of a transformation or mapping $w = f(z)$, we mean a point which is carried into itself by the transformation. Obviously, every such point is a solution of the equation

$$z = f(z).$$

Moreover, every point of the z-plane is trivially a fixed point of the unit transformation $U(z) = z$.

THEOREM 10.5. *Every Möbius transformation different from the unit transformation has two fixed points, which in certain cases coalesce into a single fixed point.*

Proof. First let $c = 0$ ($d \neq 0$), so that $L(z)$ reduces to the entire linear transformation

$$L(z) = \alpha z + \beta \qquad \left(\alpha = \frac{a}{d},\ \beta = \frac{b}{d}\right).$$

Then, since $L(\infty) = \infty$, one fixed point is the point at infinity. If $\alpha \neq 1$, there exists another fixed point determined from the equation

$$z = \alpha z + \beta,$$

i.e., the point $\beta/(1 - \alpha)$, but if $\alpha = 1$, $\beta \neq 0$, there is no finite fixed point. Moreover, if $\alpha \neq 1$, $\beta \neq 0$, the finite fixed point $\beta/(1 - \alpha)$ approaches ∞ as $\alpha \to 1$. Therefore, in the case of the transformation

$$L(z) = z + \beta \qquad (\beta \neq 0),$$

the point at infinity can be regarded as two fixed points which have coalesced.

Now let $c \neq 0$, so that

$$L(\infty) = \frac{a}{c} \neq \infty,$$

i.e., the point at infinity is not a fixed point.[8] Similarly, the pole $\delta = -d/c$ of the transformation is not a fixed point, since

$$L(\delta) = \infty \neq \delta.$$

Assuming that $z \neq \infty$ and $z \neq \delta$, we solve the equation

$$z = \frac{az + b}{cz + d}$$

or

$$cz^2 + (a - d)z - b = 0,$$

obtaining

$$z = \frac{a - d \pm \sqrt{(a - d)^2 + 4bc}}{2c}.$$

If $(a - d)^2 + 4bc \neq 0$, we obtain two different finite fixed points; if $(a - d)^2 + 4bc = 0$, these two points coalesce to form a single finite fixed point $(a - d)/2c$.

COROLLARY. *The only Möbius transformation with more than two fixed points is the unit transformation $U(z) = z$, for which all points are fixed points.*

THEOREM 10.6. *A sufficient condition for two Möbius transformations $L(z)$ and $\Lambda(z)$ to be identical is that the equation*

$$L(z) = \Lambda(z)$$

hold for three distinct points z_1, z_2 and z_3. In particular, there cannot exist two distinct Möbius transformations taking three given values w_1, w_2, w_3 at three given distinct points z_1, z_2, z_3.

Proof. It follows from

$$L(z_k) = \Lambda(z_k) = w_k \qquad (k = 1, 2, 3)$$

that

$$\Lambda^{-1}(w_k) = z_k \qquad (k = 1, 2, 3),$$

and hence

$$\Lambda^{-1}L(z_k) = z_k \qquad (k = 1, 2, 3).$$

[8] Therefore, in the class of Möbius transformations, an entire linear transformation is characterized by the fact that at least one of its fixed points is the point at infinity.

Therefore the transformation $\Lambda^{-1}L$ has three distinct fixed points, which, according to the corollary, implies

$$\Lambda^{-1}L = U,$$

where U is the unit transformation. Multiplying both sides of (10.21) by Λ from the left, we obtain

$$\Lambda(\Lambda^{-1}L) = \Lambda U,$$

which implies

$$L = \Lambda,$$

as asserted.

We now set about determining the (unique) Möbius transformation $w = L(z)$ carrying the points z_1, z_2, z_3 into the points w_1, w_2, w_3. First we consider the problem of finding the special Möbius transformation $w = \Lambda(z)$ carrying three finite points z_1, z_2, z_3 into the points $0, \infty, 1$. Since the function

$$w = \Lambda(z) = \frac{az + b}{cz + d}$$

vanishes for $z = z_1$ and becomes infinite for $z = z_2$ if and only if z_1 is a zero of the numerator and z_2 is a zero of the denominator, it follows that

$$w = \Lambda(z) = \frac{a}{c}\,\frac{z - z_1}{z - z_2}.$$

But w must equal 1 for $z = z_3$, i.e.,

$$1 = \frac{a}{c}\,\frac{z_3 - z_1}{z_3 - z_2},$$

which implies[9]

$$\frac{a}{c} = 1 : \frac{z_3 - z_1}{z_3 - z_2}.$$

Therefore the Möbius transformation carrying z_1, z_2, z_3 into $0, \infty, 1$ is

$$w = \Lambda(z) = \frac{z - z_1}{z - z_2} : \frac{z_3 - z_1}{z_3 - z_2}. \tag{10.21}$$

Next we determine the more general Möbius transformation $w = L(z)$ satisfying

$$L(z_1) = w_1, \quad L(z_2) = w_2, \quad L(z_3) = w_3,$$

where w_1, w_2, w_3 are three arbitrary (but distinct) finite points. As we have just seen, the transformation

$$\Lambda_1(z) = \frac{z - w_1}{z - w_2} : \frac{w_3 - w_1}{w_3 - w_2} \tag{10.22}$$

[9] By $x:y$ is meant the *ratio of x to y*, i.e., the quantity x/y.

carries the points w_1, w_2, w_3 into the points 0, ∞, 1. Therefore the transformation $\Lambda_1 L$ carries the points z_1, z_2, z_3 into the points 0, ∞, 1, so that

$$\Lambda_1 L(z) = \Lambda(z) = \frac{z - z_1}{z - z_2} : \frac{z_3 - z_1}{z_3 - z_2}. \tag{10.23}$$

Multiplying both sides of the equation

$$\Lambda_1 L(z) = \Lambda(z) \tag{10.24}$$

by Λ_1^{-1} from the left, we obtain

$$L(z) = \Lambda_1^{-1}\Lambda(z).$$

This solves the problem, since the functions $\Lambda(z)$ and $\Lambda_1(z)$, and hence $\Lambda_1^{-1}(z)$, are known [cf. (10.23) and (10.22)]. However, it is more convenient to use (10.24) directly, after writing $w = L(z)$. The result is

$$\Lambda_1(w) = \Lambda(z)$$

or

$$\frac{w - w_1}{w - w_2} : \frac{w_3 - w_1}{w_3 - w_2} = \frac{z - z_1}{z - z_2} : \frac{z_3 - z_1}{z_3 - z_2}, \tag{10.25}$$

which expresses the Möbius transformation $w = L(z)$ in implicit form.

Remark. In finding the Möbius transformation carrying the points z_1, z_2, z_3 into the points w_1, w_2, w_3, it was assumed that all six points are finite. However, the case of infinite points is easily handled. For example, the transformation carrying the points ∞, z_2, z_3 into the points 0, ∞, 1 has the form

$$w = \Lambda(z) = \frac{1}{z - z_2} : \frac{1}{z_3 - z_2},$$

which can be found by inspection or by writing (10.21) in the form

$$w = \frac{\dfrac{z}{z_1} - 1}{z - z_2} : \frac{\dfrac{z_3}{z_1} - 1}{z_3 - z_2}$$

and then taking the limit as $z_1 \to \infty$. Therefore (10.25) is replaced by

$$\frac{w - w_1}{w - w_2} : \frac{w_3 - w_1}{w_3 - w_2} = \frac{1}{z_1 - z_2} : \frac{1}{z_3 - z_2},$$

where it is assumed that the points w_1, w_2, w_3 are finite. Similarly, the transformation carrying the points z_1, ∞, z_3 into the points 0, ∞, 1 has the form

$$\Lambda(z) = (z - z_1) : (z_3 - z_1),$$

and hence we have

$$\frac{w - w_1}{w - w_2} : \frac{w_3 - w_1}{w_3 - w_2} = (z - z_1) : (z_3 - z_1),$$

instead of (10.25). Finally, the transformation carrying the points z_1, z_2, ∞ into the points 0, ∞, 1 has the form

$$\Lambda(z) = \frac{z - z_1}{z - z_2},$$

and hence we have

$$\frac{w - w_1}{w - w_2} : \frac{w_3 - w_1}{w_3 - w_2} = \frac{z - z_1}{z - z_2},$$

instead of (10.25).

In just the same way, we have to replace the left-hand side of (10.25) by

$$\frac{1}{w - w_2} : \frac{1}{w_3 - w_2}, \quad (w - w_1) : (w_3 - w_1) \quad \text{or} \quad \frac{w - w_1}{w - w_2},$$

depending on whether $w_1 = \infty$, $w_2 = \infty$ or $w_3 = \infty$. As a result, we arrive at the following mnemonic rule: If $z_k = \infty$ or $w_l = \infty$ ($k, l = 1, 2, 3$), the differences involving z_k or w_l have to be replaced by 1. The reader can easily verify this rule by taking the appropriate limits (as $z_k \to \infty$ or $w_l \to \infty$) in equation (10.25).

Equation (10.25) implies an important general property of Möbius transformations. If a, b, c and d are arbitrary distinct finite complex numbers, then the ratio

$$\frac{c - a}{c - b} : \frac{d - a}{d - b},$$

denoted by (a, b, c, d), is called the *cross ratio* (or *anharmonic ratio*) of the four numbers (or points) a, b, c and d. If one of the four points a, b, c, d is the point at infinity, we define the cross ratio as the limit of the cross ratio of four finite points, three of which coincide with the three given finite points, as the fourth point approaches infinity. Thus, according to this definition, we have

$$(\infty, b, c, d) = \frac{1}{c - b} : \frac{1}{d - b},$$
$$(a, \infty, c, d) = (c - a) : (d - a),$$
$$(a, b, \infty, d) = 1 : \frac{d - a}{d - b},$$
$$(a, b, c, \infty) = \frac{c - a}{c - b}.$$

Now let $w = L(z)$ be an arbitrary Möbius transformation, and let A, B, C, D be the points into which $L(z)$ maps four arbitrary (but distinct) points a, b, c, d. Since the points A, B and D are the images of the points a, b and d, it follows from (10.25) that the relation between z and $w = L(z)$ is given by

$$\frac{w - A}{w - B} : \frac{D - A}{D - B} = \frac{z - a}{z - b} : \frac{d - a}{d - b},$$

where differences involving the point at infinity have to be replaced by 1. Moreover, since the point C is the image of the point c, we have

$$\frac{C-A}{C-B}:\frac{D-A}{D-B}=\frac{c-a}{c-b}:\frac{d-a}{d-b},$$

(where again differences involving the point at infinity have to be replaced by 1), or equivalently

$$(A, B, C, D) = (a, b, c, d).$$

In other words, *the cross ratio of any four distinct points is invariant under a Möbius transformation.*

47. Mapping of a Circle onto a Circle[10]

Using the circle-preserving property of Möbius transformations and the possibility of mapping any given triple of distinct points z_1, z_2, z_3 into any other given triple of distinct points w_1, w_2, w_3, we obtain the following basic result:

THEOREM 10.7. *Let γ and Γ be any two straight lines or circles, and let z_1, z_2, z_3 and w_1, w_2, w_3 be any two triples of distinct points belonging to γ and Γ, respectively. Then there exists a Möbius transformation $w = L(z)$ mapping γ onto Γ in such a way that*

$$L(z_k) = w_k \qquad (k = 1, 2, 3). \tag{10.26}$$

Proof. Construct the Möbius transformation $w = L(z)$ satisfying the conditions (10.26), which according to Theorem 10.6 and the subsequent construction, exists and is unique. According to Theorem 10.4, $w = L(z)$ maps the straight line or circle γ onto another straight line or circle Γ^*. But since γ goes through the points z_1, z_2 and z_3, Γ^* must go through the points w_1, w_2 and w_3. Moreover, since two different straight lines or circles cannot be drawn through the same three points, Γ^* must coincide with Γ, as asserted.

Remark. Again consider two arbitrary straight lines or circles γ and Γ (which may coincide). Let G be one of the two domains with boundary γ, and let $\mathscr{G}$ be one of the two domains with boundary Γ, so that G is either a half-plane, the interior of a circle or the exterior of a circle, and the same is true of $\mathscr{G}$. We now show how to map G onto $\mathscr{G}$. Choose an arbitrary triple of distinct points z_1, z_2, z_3 on γ, and suppose an observer moving along γ in the direction from z_1 to z_3 through z_2 finds the domain G on his *left*, say. Next choose a triple of distinct points w_1, w_2, w_3 on Γ such that an observer moving along Γ in the direction from w_1 to w_3 through w_2 finds the domain $\mathscr{G}$ on his left, but let w_1, w_2, w_3 be otherwise arbitrary. As in Theorem

[10] As usual, a straight line is regarded as a limiting case of a circle (cf. p. 168).

10.7, we form the Möbius transformation $w = L(z)$ which satisfies the conditions (10.26) and hence maps γ onto Γ. Then $w = L(z)$ also maps G onto $\mathscr{G}$, i.e., $\mathscr{G} = L(G)$. In fact, if δ is a segment of the normal to the curve γ drawn from the point z_2 and pointing into the interior of G, so that an observer at z_2 facing in the direction established on γ finds δ on his left, then, since the mapping $w = L(z)$ is conformal, an observer at w_2 facing in the direction established on Γ will also find the image $\Delta = L(\delta)$, which is a line segment or circular arc, on his left (see Figure 10.2 and Remark 2, p. 156). Therefore $\Delta \subset \mathscr{G}$, and hence $\mathscr{G}$ contains images of certain points belonging to G (i.e., the points of the segment δ). But, according to the remark on p. 170, $L(G)$ is one of the two domains with boundary $\Gamma = L(\gamma)$, in fact just the domain containing the image of any point in G. In other words, $\mathscr{G} = L(G)$, as asserted.

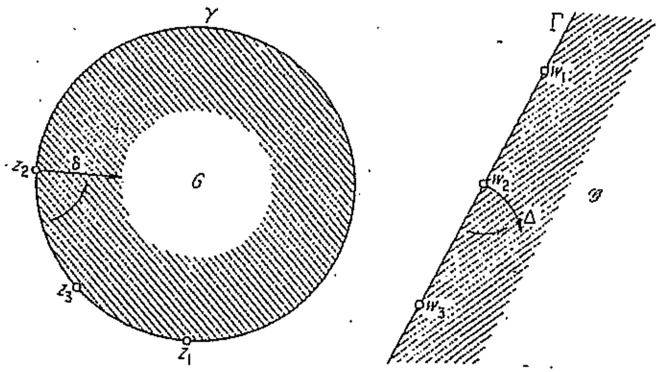

FIGURE 10.2

Example. Find a conformal mapping of the upper half-plane $\text{Im}\, z > 0$ *onto the interior of the unit circle.*

To solve this problem, we choose $z_1 = -1$, $z_2 = 0$, $z_3 = 1$, say, so that the upper half-plane is on the left of an observer moving along the real axis in the direction from z_1 to z_3 through z_2. We also choose three points w_1, w_2, w_3 on the unit circle, such that the interior of the circle is on the left of an observer moving along the circle in the direction from w_1 to w_3 through w_2. For simplicity, we choose $w_1 = 1$, $w_2 = i$, $w_3 = -1$. Then the desired Möbius transformation satisfies the conditions $L(z_k) = w_k$, $k = 1, 2, 3$, and can be represented in the form

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

or

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

where we have used (10.25).

48. Symmetry Transformations

Let z_1 and z_2 be two points which are symmetric with respect to a given straight line γ, i.e., such that γ is the perpendicular bisector of the line segment joining z_1 and z_2. By definition, the straight line passing through z_1 and z_2 is orthogonal to γ. Moreover, the center of any circle δ passing through z_1 and z_2 lies on γ, and hence δ is also orthogonal to γ. It is easy to see that the converse is true as well, i.e., if every straight line or circle passing through a pair of points z_1 and z_2 is orthogonal to a given straight line γ, then z_1 and z_2 are symmetric with respect to γ. Generalizing the concept of symmetry with respect to a straight line, we introduce the following definition: *Two points z_1 and z_2 are symmetric with respect to a given circle γ if and only if every straight line or circle passing through z_1 and z_2 is orthogonal to γ.*

THEOREM 10.8. *Let z_1 and z_2 be any two points symmetric with respect to a given straight line or circle γ, and let $w = L(z)$ be any Möbius transformation. Then the points $w_1 = L(z_1)$ and $w_2 = L(z_2)$ are symmetric with respect to the straight line or circle $\Gamma = L(\gamma)$.*[11]

Proof. We have to show that an arbitrary straight line or circle Δ passing through w_1 and w_2 is orthogonal to Γ. Let $z = L^{-1}(w)$ be the inverse of the transformation $w = L(z)$. Clearly, L^{-1} is also a Möbius transformation, and

$$L^{-1}(w_1) = z_1, \quad L^{-1}(w_2) = z_2, \quad L^{-1}(\Gamma) = \gamma.$$

Moreover, $\delta = L^{-1}(\Delta)$ is a straight line or circle passing through z_1 and z_2. Since z_1 and z_2 are symmetric with respect to γ, by hypothesis, it follows that δ is orthogonal to γ. But then, since the mapping $w = L(z)$ is conformal (see Sec. 33), $\Delta = L(\delta)$ is orthogonal to Γ, and the proof is complete.

COROLLARY. *There is only one point z_2 symmetric to a given point z_1 with respect to a given straight line or circle γ.*

Proof. If γ is a straight line, the statement is obvious. Thus let γ be a circle, and suppose that besides z_2, there is another point $z_2' \neq z_2$ symmetric to z_1 with respect to γ. Choosing a Möbius transformation $w = L(z)$ mapping γ onto a straight line Γ, we find that $w_2 = L(z_2)$ and $w_2' = L(z_2')$ are two distinct points symmetric with respect to Γ, which is impossible.

Remark. Suppose $w = L(z)$ maps a straight line or circle γ onto a circle Γ with center w_1, and let z_1 be the inverse image of w_1. Then the point z_2 symmetric to z_1 with respect to γ must be mapped into the point at infinity. To see this, we note that $w_2 = \infty$ is symmetric to $w_1 = 0$ with respect to the circle Γ, since any straight line or circle passing through 0 and ∞, i.e., any straight line passing through the center of Γ, is orthogonal to Γ. The uniqueness of w_2 follows from the corollary.

[11] It is in this sense that Möbius transformations are said to be *symmetry-preserving*.

Let γ be an arbitrary straight line or circle. A transformation of the extended plane into itself, which carries each point z into the point z^* symmetric to z with respect to γ is called a *symmetry transformation* with respect to γ or a *reflection* in γ. In the case where γ is a circle, the transformation is also called an *inversion* in γ. We now derive analytical expressions for symmetry transformations.

First let γ be a straight line with an assigned direction, and consider reflection in γ. The straight line γ is completely characterized by one of its points a and by the unit vector

$$e^{i\theta} = \cos\theta + i\sin\theta$$

pointing in the direction of γ. Suppose we carry out the entire linear transformation

$$z = L(w) = a + e^{i\theta}w, \tag{10.27}$$

which obviously maps the real axis onto γ, since (10.27) corresponds to a shift by the vector a (carrying the origin of coordinates to the point a), followed by a rotation through the angle θ about the point a. Since the inverse transformation $w = L^{-1}(z)$ maps γ onto the real axis, it maps every pair of points z and z^* symmetric with respect to γ into a pair of points w and w^* symmetric with respect to the real axis. But the points w and w^* are represented by two conjugate complex numbers, i.e.,

$$w = t, \qquad w^* = \bar{t}. \tag{10.28}$$

Therefore $z - a = e^{i\theta}t$, and

$$\overline{z - a} = e^{-i\theta}\bar{t}, \qquad z^* - a = e^{i\theta}w^* = e^{i\theta}\bar{t}. \tag{10.29}$$

Eliminating $\bar{t}$ from (10.29), we obtain

$$z^* - a = e^{2i\theta}(\overline{z - a}). \tag{10.30}$$

According to (10.30), reflection in a straight line γ going through a point a at an angle θ with the real axis can be accomplished by first constructing the vector $\overline{z - a}$ which is the reflection of the vector $z - a$ in the real axis, and then rotating $\overline{z - a}$ through the angle 2θ about the point a.

Next let γ be a circle, and consider inversion in γ. Let R $(0 < R < \infty)$ be the radius and a the center of γ. We begin by finding a Möbius transformation