can be evaluated as a double Riemann integral. If $f(z) = u(x,y) + iv(x,y)$ is a bijective differentiable mapping, then by the rule for changing integration variables the area of the image $E' = f(E)$ is given by

$$A(E') = \iint_E |u_x v_y - u_y v_x| \, dx \, dy.$$

But if $f(z)$ is a conformal mapping of an open set containing E, then $u_x v_y - u_y v_x = |f'(z)|^2$ by virtue of the Cauchy-Riemann equations, and we obtain

$$A(E') = \iint_E |f'(z)|^2 \, dx \, dy. \tag{6}$$

The formulas (5) and (6) have important applications in the part of complex analysis that is frequently referred to as geometric function theory.

3. LINEAR TRANSFORMATIONS

Of all analytic functions the first-order rational functions have the simplest mapping properties, for they define mappings of the extended plane onto itself which are at the same time conformal and topological. The linear transformations have also very remarkable geometric properties, and for that reason their importance goes far beyond serving as simple examples of conformal mappings. The reader will do well to pay particular attention to this geometric aspect, for it will equip him with simple but very valuable techniques.

3.1. The Linear Group. We have already remarked in Chap. 2, Sec. 1.4 that a *linear fractional transformation*

$$w = S(z) = \frac{az + b}{cz + d} \tag{7}$$

with $ad - bc \neq 0$ has an inverse

$$z = S^{-1}(w) = \frac{dw - b}{-cw + a}.$$

The special values $S(\infty) = a/c$ and $S(-d/c) = \infty$ can be introduced either by convention or as limits for $z \to \infty$ and $z \to -d/c$. With the latter interpretation it becomes obvious that S is a topological mapping of the extended plane onto itself, the topology being defined by distances on the Riemann sphere.

For linear transformations we shall usually replace the notation $S(z)$

by Sz. The representation (7) is said to be normalized if $ad - bc = 1$. It is clear that every linear transformation has two normalized representations, obtained from each other by changing the signs of the coefficients.

A convenient way to express a linear transformation is by use of homogeneous coordinates. If we write $z = z_1/z_2$, $w = w_1/w_2$ we find that $w = Sz$ if

$$\begin{aligned} w_1 &= az_1 + bz_2 \\ w_2 &= cz_1 + dz_2 \end{aligned} \tag{8}$$

or, in matrix notation,

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

The main advantage of this notation is that it leads to a simple determination of a composite transformation $w = S_1S_2z$. If we use subscripts to distinguish between the matrices that correspond to S_1, S_2 it is immediate that S_1S_2 belongs to the matrix product

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{pmatrix}.$$

All linear transformations form a group. Indeed, the associative law $(S_1S_2)S_3 = S_1(S_2S_3)$ holds for arbitrary transformations, the identity $w = z$ is a linear transformation, and the inverse of a linear transformation is linear. The ratios $z_1 : z_2 \neq 0 : 0$ are the points of the complex projective line, and (8) identifies the group of linear transformations with the one-dimensional projective group over the complex numbers, usually denoted by $P(1,\mathbf{C})$. If we use only normalized representations, we can also identify it with the group of two-by-two matrices with determinant 1 (denoted $SL(2,\mathbf{C})$), except that there are two opposite matrices corresponding to the same linear transformation.

We shall make no further use of the matrix notation, except for remarking that the simplest linear transformations belong to matrices of the form

$$\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The first of these, $w = z + \alpha$, is called a *parallel translation*. The second, $w = kz$, is a *rotation* if $|k| = 1$ and a *homothetic transformation* if $k > 0$. For arbitrary complex $k \neq 0$ we can set $k = |k| \cdot k/|k|$, and hence $w = kz$ can be represented as the result of a homothetic transformation followed by a rotation. The third transformation, $w = 1/z$, is called an *inversion*.

If $c \neq 0$ we can write

$$\frac{az + b}{cz + d} = \frac{bc - ad}{c^2(z + d/c)} + \frac{a}{c},$$

and this decomposition shows that the most general linear transformation is composed by a translation, an inversion, a rotation, and a homothetic transformation followed by another translation. If $c = 0$, the inversion falls out and the last translation is not needed.

EXERCISES

1. Prove that the reflection $z \to \bar{z}$ is not a linear transformation.

2. If

$$T_1z = \frac{z+2}{z+3}, \qquad T_2z = \frac{z}{z+1},$$

find T_1T_2z, T_2T_1z and $T_1^{-1}T_2z$.

3. Prove that the most general transformation which leaves the origin fixed and preserves all distances is either a rotation or a rotation followed by reflexion in the real axis.

4. Show that any linear transformation which transforms the real axis into itself can be written with real coefficients.

3.2. The Cross Ratio. Given three distinct points z_2, z_3, z_4 in the extended plane, there exists a linear transformation S which carries them into $1, 0, \infty$ in this order. If none of the points is ∞, S will be given by

$$Sz = \frac{z - z_3}{z - z_4} : \frac{z_2 - z_3}{z_2 - z_4}. \tag{9}$$

If z_2, z_3 or $z_4 = \infty$ the transformation reduces to

$$\frac{z - z_3}{z - z_4}, \qquad \frac{z_2 - z_4}{z - z_4}, \qquad \frac{z - z_3}{z_2 - z_3}$$

respectively.

If T were another linear transformation with the same property, then ST^{-1} would leave $1, 0, \infty$ invariant. Direct calculation shows that this is true only for the identity transformation, and we would have $S = T$. We conclude that S is uniquely determined.

Definition 12. *The cross ratio* (z_1,z_2,z_3,z_4) *is the image of* z_1 *under the linear transformation which carries* z_2,z_3,z_4 *into* $1, 0, \infty$.

The definition is meaningful only if z_2,z_3,z_4 are distinct. A conventional value can be introduced as soon as any three of the points are distinct, but this is unimportant.

The cross ratio is invariant under linear transformations. In more precise formulation:

Theorem 12. *If z_1, z_2, z_3, z_4 are distinct points in the extended plane and T any linear transformation, then $(Tz_1,Tz_2,Tz_3,Tz_4) = (z_1,z_2,z_3,z_4)$.*

The proof is immediate, for if $Sz = (z,z_2,z_3,z_4)$, then ST^{-1} carries Tz_2, Tz_3, Tz_4 into 1, 0, ∞. By definition we have hence

$$(Tz_1,Tz_2,Tz_3,Tz_4) = ST^{-1}(Tz_1) = Sz_1 = (z_1,z_2,z_3,z_4).$$

With the help of this property we can immediately write down the linear transformation which carries three given points z_1, z_2, z_3 to prescribed positions w_1, w_2, w_3. The correspondence must indeed be given by

$$(w,w_1,w_2,w_3) = (z,z_1,z_2,z_3).$$

In general it is of course necessary to solve this equation with respect to w.

Theorem 13. *The cross ratio (z_1,z_2,z_3,z_4) is real if and only if the four points lie on a circle or on a straight line.*

This is evident by elementary geometry, for we obtain

$$\arg\,(z_1,z_2,z_3,z_4) = \arg\frac{z_1 - z_3}{z_1 - z_4} - \arg\frac{z_2 - z_3}{z_2 - z_4},$$

and if the points lie on a circle this difference of angles is either 0 or $\pm\pi$, depending on the relative location.

For an analytic proof we need only show that the image of the real axis under any linear transformation is either a circle or a straight line. Indeed, $Tz = (z,z_2,z_3,z_4)$ is real on the image of the real axis under the transformation T^{-1} and nowhere else.

The values of $w = T^{-1}z$ for real z satisfy the equation $Tw = \overline{Tw}$. Explicitly, this condition is of the form

$$\frac{aw + b}{cw + d} = \frac{\bar a\bar w + \bar b}{\bar c\bar w + \bar d}.$$

By cross multiplication we obtain

$$(a\bar c - c\bar a)|w^2| + (a\bar d - c\bar b)w + (b\bar c - d\bar a)\bar w + b\bar d - d\bar b = 0.$$

If $a\bar c - c\bar a = 0$ this is the equation of a straight line, for under this condition the coefficient $a\bar d - c\bar b$ cannot also vanish. If $a\bar c - c\bar a \neq 0$ we can divide by this coefficient and complete the square. After a simple computation we obtain

$$\left| w + \frac{\bar{a}d - \bar{c}b}{\bar{a}c - \bar{c}a} \right| = \left| \frac{ad - bc}{\bar{a}c - \bar{c}a} \right|$$

which is the equation of a circle.

The last result makes it clear that we should not, in the theory of linear transformations, distinguish between circles and straight lines. A further justification was found in the fact that both correspond to circles on the Riemann sphere. Accordingly we shall agree to use the word circle in this wider sense.†

The following is an immediate corollary of Theorems 12 and 13:

Theorem 14. *A linear transformation carries circles into circles.*

EXERCISES

1. Find the linear transformation which carries 0, i, $-i$ into 1, -1, 0.

2. Express the cross ratios corresponding to the 24 permutations of four points in terms of $\lambda = (z_1,z_2,z_3,z_4)$.

3. If the consecutive vertices z_1, z_2, z_3, z_4 of a quadrilateral lie on a circle, prove that

$$|z_1 - z_3| \cdot |z_2 - z_4| = |z_1 - z_2| \cdot |z_3 - z_4| + |z_2 - z_3| \cdot |z_1 - z_4|$$

and interpret the result geometrically.

4. Show that any four distinct points can be carried by a linear transformation to positions 1, -1, k, $-k$, where the value of k depends on the points. How many solutions are there, and how are they related?

3.3. Symmetry. The points z and $\bar{z}$ are symmetric with respect to the real axis. A linear transformation with real coefficients carries the real axis into itself and z, $\bar{z}$ into points which are again symmetric. More generally, if a linear transformation T carries the real axis into a circle C, we shall say that the points $w = Tz$ and $w^* = T\bar{z}$ are *symmetric with respect to* C. This is a relation between w, w^* and C which does not depend on T. For if S is another transformation which carries the real axis into C, then $S^{-1}T$ is a real transformation, and hence $S^{-1}w = S^{-1}Tz$ and $S^{-1}w^* = S^{-1}T\bar{z}$ are also conjugate. Symmetry can thus be defined in the following terms:

---

† This agreement will be in force only when dealing with linear transformations.

Definition 13. *The points z and z^* are said to be symmetric with respect to the circle C through z_1, z_2, z_3 if and only if* $(z^*,z_1,z_2,z_3) = \overline{(z,z_1,z_2,z_3)}$.

The points on C, and only those, are symmetric to themselves. The mapping which carries z into z^* is a one-to-one correspondence and is called *reflection* with respect to C. Two reflections will evidently result in a linear transformation.

We wish to investigate the geometric significance of symmetry. Suppose first that C is a straight line. Then we can choose $z_3 = \infty$ and the condition for symmetry becomes

$$\frac{z^* - z_2}{z_1 - z_2} = \frac{\bar z - \bar z_2}{\bar z_1 - \bar z_2} \tag{10}$$

Taking absolute values we obtain $|z^* - z_2| = |z - z_2|$. Here z_2 can be any finite point on C, and we conclude that z and z^* are equidistant from all points on C. By (10) we have further

$$\operatorname{Im} \frac{z^* - z_2}{z_1 - z_2} = -\operatorname{Im} \frac{z - z_2}{z_1 - z_2},$$

and hence z and z^* are in different half planes determined by C.† We leave to the reader to prove that C is the bisecting normal of the segment between z and z^*.

Consider now the case of a finite circle C of center a and radius R. Systematic use of the invariance of the cross ratio allows us to conclude as follows:

$$\begin{aligned}\overline{(z,z_1,z_2,z_3)} &= \overline{(z - a,z_1 - a,z_2 - a,z_3 - a)} \\ &= \left(\bar z - \bar a, \frac{R^2}{z_1 - a}, \frac{R^2}{z_2 - a}, \frac{R^2}{z_3 - a}\right) = \left(\frac{R^2}{\bar z - \bar a}, z_1 - a, z_2 - a, z_3 - a\right) \\ &= \left(\frac{R^2}{\bar z - \bar a} + a, z_1, z_2, z_3\right).\end{aligned}$$

This equation shows that the symmetric point of z is $z^* = R^2/(\bar z - \bar a) + a$ or that z and z^* satisfy the relation

$$(z^* - a)(\bar z - \bar a) = R^2. \tag{11}$$

The product $|z^* - a| \cdot |z - a|$ of the distances to the center is hence R^2. Further, the ratio $(z^* - a)/(z - a)$ is positive, which means that z and z^* are situated on the same half line from a. There is a simple geometric construction for the symmetric point of z (Fig. 3-2). We note that the symmetric point of a is ∞.

---

† Unless they coincide and lie on C.

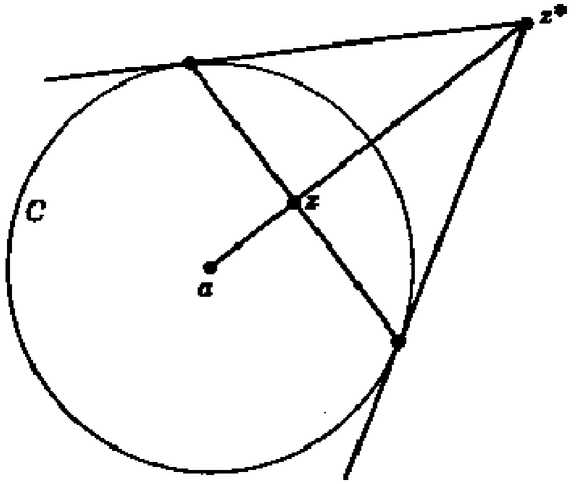

FIG. 3-2. Reflection in a circle.

Theorem 15. (*The symmetry principle.*) *If a linear transformation carries a circle C_1 into a circle C_2, then it transforms any pair of symmetric points with respect to C_1 into a pair of symmetric points with respect to C_2.*

Briefly, linear transformations preserve symmetry. If C_1 or C_2 is the real axis, the principle follows from the definition of symmetry. In the general case the assertion follows by use of an intermediate transformation which carries C_1 into the real axis.

There are two ways in which the principle of symmetry can be used. If the images of z and C under a certain linear transformation are known, then the principle allows us to find the image of z^*. On the other hand, if the images of z and z^* are known, we conclude that the image of C must be a line of symmetry of these images. While this is not enough to determine the image of C, the information we gain is nevertheless valuable.

The principle of symmetry is put to practical use in the problem of finding the linear transformations which carry a circle C into a circle C'. We can always determine the transformation by requiring that three points z_1, z_2, z_3 on C go over into three points w_1, w_2, w_3 on C'; the transformation is then $(w,w_1,w_2,w_3) = (z,z_1,z_2,z_3)$. But the transformation is also determined if we prescribe that a point z_1 on C shall correspond to a point w_1 on C' and that a point z_2 *not* on C shall be carried into a point w_2 *not* on C'. We know then that z_2^* (the symmetric point of z_2 with respect to C) must correspond to w_2^* (the symmetric point of w_2 with respect to C'). Hence the transformation will be obtained from the relation $(w,w_1,w_2,w_2^*) = (z,z_1,z_2,z_2^*)$.

EXERCISES

1. Prove that every reflection carries circles into circles.

2. Reflect the imaginary axis, the line $x = y$, and the circle $|z| = 1$ in the circle $|z - 2| = 1$.

3. Carry out the reflections in the preceding exercise by geometric construction.

4. Find the linear transformation which carries the circle $|z| = 2$ into $|z + 1| = 1$, the point -2 into the origin, and the origin into i.

5. Find the most general linear transformation of the circle $|z| = R$ into itself.

6. Suppose that a linear transformation carries one pair of concentric circles into another pair of concentric circles. Prove that the ratios of the radii must be the same.

7. Find a linear transformation which carries $|z| = 1$ and $|z - \frac{1}{4}| = \frac{1}{4}$ into concentric circles. What is the ratio of the radii?

8. Same problem for $|z| = 1$ and $x = 2$.

3.4. Oriented Circles. Because $S(z)$ is analytic and

$$S'(z) = \frac{ad - bc}{(cz + d)^2} \neq 0$$

the mapping $w = S(z)$ is conformal for $z \neq -d/c$ and ∞. It follows that a pair of intersecting circles are mapped on circles that include the same angle. In addition, the sense of an angle is preserved. From an intuitive point of view this means that right and left are preserved, but a more precise formulation is desirable.

An orientation of a circle C is determined by an ordered triple of points z_1,z_2,z_3 on C. With respect to this orientation a point z not on C is said to lie to the *right* of C if $\operatorname{Im}\,(z,z_1,z_2,z_3) > 0$ and to the *left* of C is $\operatorname{Im}\,(z,z_1,z_2,z_3) < 0$ (this checks with everyday use because $(i,1,0,\infty) = i$). It is essential to show that there are only two different orientations. By this we mean that the distinction between left and right is the same for all triples, while the meaning may be reversed. Since the cross ratio is invariant, it is sufficient to consider the case where C is the real axis. Then

$$(z,z_1,z_2,z_3) = \frac{az + b}{cz + d}$$

can be written with real coefficients, and a simple calculation gives

$$\operatorname{Im}\,(z,z_1,z_2,z_3) = \frac{ad - bc}{|cz + d|^2} \operatorname{Im} z.$$

We recognize that the distinction between right and left is the same as the distinction between the upper and lower half plane. Which is which depends on the sign of the determinant $ad - bc$.

A linear transformation S carries the oriented circle C into a circle which we orient through the triple Sz_1, Sz_2, Sz_3. From the invariance of the cross ratio it follows that the left and right of C will be mapped on the left and right of the image circle.

If two circles are tangent to each other, their orientations can be compared. Indeed, we can use a linear transformation which throws their common point to ∞. The circles become parallel straight lines, and we know how to compare the directions of parallel lines.

In the geometric representation the orientation z_1, z_2, z_3 can be indicated by an arrow which points from z_1 over z_2 to z_3. With the usual choice of the coordinate system left and right will have their customary meaning with respect to this arrow.

When the finite plane is considered as part of the extended plane, the point at infinity is distinguished. We can therefore define an absolute positive orientation of all finite circles by the requirement that ∞ should lie to the right of the oriented circles. The points to the left are said to form the *inside* of the circle and the points to the right form its *outside.*

EXERCISES

1. If z_1, z_2, z_3, z_4 are points on a circle, show that z_1, z_3, z_4 and z_2, z_3, z_4 determine the same orientation if and only if $(z_1,z_2,z_3,z_4) > 0$.

2. Prove that a tangent to a circle is perpendicular to the radius through the point of contact (in this connection a tangent should be defined as a straight line with only one point in common with the circle).

3. Verify that the inside of the circle $|z - a| = R$ is formed by all points z with $|z - a| < R$.

4. The angle between two oriented circles at a point of intersection is defined as the angle between the tangents at that point, equipped with the same orientation. Prove by analytic reasoning, rather than geometric inspection, that the angles at the two points of intersection are opposite to each other.

3.5. Families of Circles.

A great deal can be done toward the visualization of linear transformations by the introduction of certain families of circles which may be thought of as coordinate lines in a circular coordinate system.

Consider a linear transformation of the form

$$w = k \cdot \frac{z - a}{z - b}.$$

Here $z = a$ corresponds to $w = 0$ and $z = b$ to $w = \infty$. It follows that the straight lines through the origin of the w-plane are images of the circles through a and b. On the other hand, the concentric circles about the origin, $|w| = \rho$, correspond to circles with the equation

$$\left|\frac{z-a}{z-b}\right| = \rho/|k|.$$

These are the *circles of Apollonius* with limit points a and b. By their equation they are the loci of points whose distances from a and b have a constant ratio.

Denote by C_1 the circles through a, b and by C_2 the circles of Apollonius with these limit points. The configuration (Fig. 3-3) formed by all the circles C_1 and C_2 will be referred to as the *circular net* or the *Steiner circles* determined by a and b. It has many interesting properties of which we shall list a few:

1. There is exactly one C_1 and one C_2 through each point in the plane with the exception of the limit points.
2. Every C_1 meets every C_2 under right angles.
3. Reflection in a C_1 transforms every C_2 into itself and every C_1 into another C_1. Reflection in a C_2 transforms every C_1 into itself and every C_2 into another C_2.
4. The limit points are symmetric with respect to each C_2, but not with respect to any other circle.

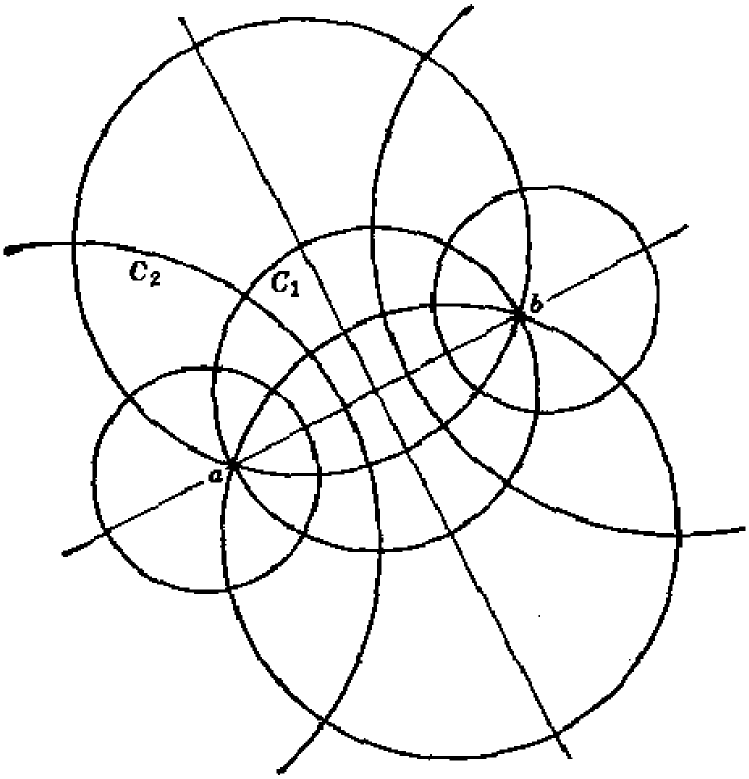

FIG. 3-3. Steiner circles.

These properties are all trivial when the limit points are 0 and ∞, *i.e.*, when the C_1 are lines through the origin and the C_2 concentric circles. Since the properties are invariant under linear transformations, they must continue to hold in the general case.

If a transformation $w = Tz$ carries a, b into a', b' it can be written in the form

$$\frac{w - a'}{w - b'} = k \cdot \frac{z - a}{z - b}. \tag{12}$$

It is clear that T transforms the circles C_1 and C_2 into circles C_1' and C_2' with the limit points a', b'.

The situation is particularly simple if $a' = a$, $b' = b$. Then a, b are said to be *fixed points* of T, and it is convenient to represent z and Tz in the same plane. Under these circumstances the whole circular net will be mapped upon itself. The value of k serves to identify the image circles C_1' and C_2'. Indeed, with appropriate orientations C_1 forms the angle $\arg k$ with its image C_1', and the quotient of the constant ratios $|z - a|/|z - b|$ on C_2' and C_2 is $|k|$.

The special cases in which all C_1 or all C_2 are mapped upon themselves are particularly important. We have $C_1' = C_1$ for all C_1 if $k > 0$ (if $k < 0$ the circles are still the same, but the orientation is reversed). The transformation is then said to be *hyperbolic*. When k increases the points Tz, $z \neq a, b$, will flow along the circles C_1 toward b. The consideration of this flow provides a very clear picture of a hyperbolic transformation.

The case $C_2' = C_2$ occurs when $|k| = 1$. Transformations with this property are called *elliptic*. When $\arg k$ varies, the points Tz move along the circles C_2. The corresponding flow circulates about a and b in different directions.

The general linear transformation with two fixed points is the product of a hyperbolic and an elliptic transformation with the same fixed points.

The fixed points of a linear transformation are found by solving the equation

$$z = \frac{\alpha z + \beta}{\gamma z + \delta}. \tag{13}$$

In general this is a quadratic equation with two roots; if $\gamma = 0$ one of the fixed points is ∞. It may happen, however, that the roots coincide. A linear transformation with coinciding fixed points is said to be *parabolic*. The condition for this is $(\alpha - \delta)^2 = 4\beta\gamma$.

If the equation (13) is found to have two distinct roots a and b, the transformation can be written in the form

$$\frac{w-a}{w-b} = k\frac{z-a}{z-b}.$$

We can then use the Steiner circles determined by a, b to discuss the nature of the transformation. It is important to note, however, that the method is by no means restricted to this case. We can write any linear transformation in the form (12) with arbitrary a, b and use the two circular nets to great advantage.

For the discussion of parabolic transformations it is desirable to introduce still another type of circular net. Consider the transformation

$$w = \frac{\omega}{z-a} + c.$$

It is evident that straight lines in the w-plane correspond to circles through a; moreover, parallel lines correspond to mutually tangent circles. In particular, if $w = u + iv$ the lines u = constant and v = constant correspond to two families of mutually tangent circles which intersect at right angles (Fig. 3-4). This configuration can be considered as a degenerate set of Steiner circles. It is determined by the point a and the tangent to one of the families of circles. We shall denote the images of the lines v = constant by C_1, the circles of the other family by C_2. Clearly, the line $v = \operatorname{Im} c$ corresponds to the tangent of the circles C_1; its direction is given by $\arg \omega$.

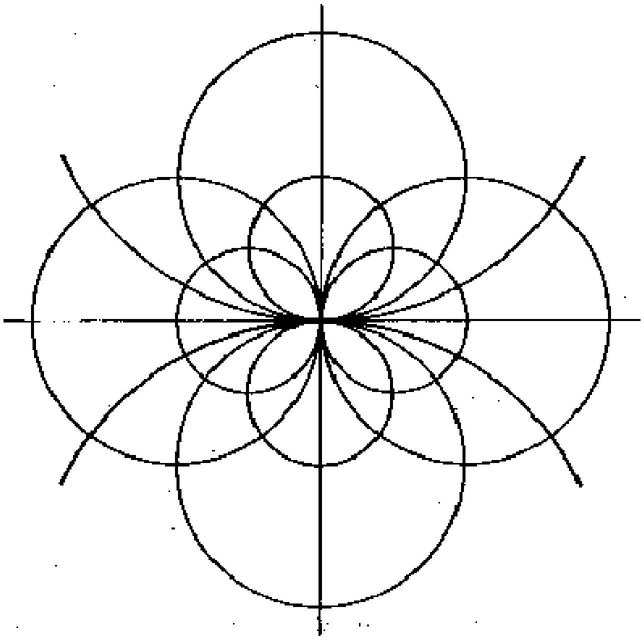

FIG. 3-4. Degenerate Steiner circles.

Any transformation which carries a into a' can be written in the form

$$\frac{\omega'}{w - a'} = \frac{\omega}{z - a} + c.$$

It is clear that the circles C_1 and C_2 are carried into the circles C_1' and C_2' determined by a' and ω'. We suppose now that $a = a'$ is the only fixed point. Then $\omega = \omega'$ and we can write

$$\frac{\omega}{w - a} = \frac{\omega}{z - a} + c. \tag{14}$$

By this transformation the configuration consisting of the circles C_1 and C_2 is mapped upon itself. In (14) a multiplicative factor is arbitrary, and we can hence suppose that c is real. Then every C_1 is mapped upon itself and the parabolic transformation can be considered as a flow along the circles C_1.

A linear transformation that is neither hyperbolic, elliptic, nor parabolic is said to be *loxodromic*.

EXERCISES

1. Find the fixed points of the linear transformations

$$w = \frac{z}{2z - 1}, \qquad w = \frac{2z}{3z - 1}, \qquad w = \frac{3z - 4}{z - 1}, \qquad w = \frac{z}{2 - z}.$$

Is any of these transformations elliptic, hyperbolic, or parabolic?

2. Suppose that the coefficients of the transformation

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

are normalized by $ad - bc = 1$. Show that S is elliptic if and only if $-2 < a + d < 2$, parabolic if $a + d = \pm 2$, hyperbolic if $a + d < -2$ or > 2.

3. Show that a linear transformation which satisfies $S^n z = z$ for some integer n is necessarily elliptic.

4. If S is hyperbolic or loxodromic, show that $S^n z$ converges to a fixed point as $n \to \infty$, the same for all z, except when z coincides with the other fixed point. (The limit is the *attractive*, the other the *repellent* fixed point. What happens when $n \to -\infty$? What happens in the parabolic case?)

5. Find all linear transformations which represent rotations of the Riemann sphere.

6. Find all circles which are orthogonal to $|z| = 1$ and $|z - 1| = 4$.

7. In an obvious way, which we shall not try to make precise, a family of transformations depends on a certain number of real parameters. How many real parameters are there in the family of all linear transformations? How many in the families of hyperbolic, elliptic, parabolic transformations? How many linear transformations leave a given circle C invariant?

4. ELEMENTARY CONFORMAL MAPPINGS

The conformal mapping associated with an analytic function affords an excellent visualization of the properties of the latter; it can well be compared with the visualization of a real function by its graph. It is therefore natural that all questions connected with conformal mapping have received a great deal of attention; progress in this direction has increased our knowledge of analytic functions considerably. In addition, conformal mapping enters naturally in many branches of mathematical physics and in this way accounts for the immediate usefulness of complex-function theory.

One of the most important problems is to determine the conformal mappings of one region onto another. In this section we shall consider those mappings which can be defined by elementary functions.

4.1. The Use of Level Curves. When a conformal mapping is defined by an explicit analytic function $w = f(z)$, we naturally wish to gain information about the specific geometric properties of the mapping. One of the most fruitful ways is to study the correspondence of curves induced by the point transformation. The special properties of the function $f(z)$ may express themselves in the fact that certain simple curves are transformed into curves of a family of well-known character. Any such information will strengthen our visual conception of the mapping.

Such was the case for mappings by linear transformations. We proved in Sec. 3 that a linear transformation carries circles into circles, provided that straight lines are included as a special case. By consideration of the Steiner circles it was possible to obtain a complete picture of the correspondence.

In more general cases it is advisable to begin with a study of the image curves of the lines $x = x_0$ and $y = y_0$. If we write $f(z) = u(x,y) + iv(x,y)$, the image of $x = x_0$ is given by the parametric equations $u = u(x_0,y)$, $v = v(x_0,y)$; y acts as a parameter and can be eliminated or retained according to convenience. The image of $y = y_0$ is determined in the same way. Together, the curves form an orthogonal net in the w-plane. Similarly, we may consider the curves $u(x,y) = u_0$ and $v(x,y) = v_0$ in the z-plane. They are also orthogonal and are called the *level curves* of u and v.

In other cases it may be more convenient to use polar coordinates and study the images of concentric circles and straight lines through the origin.

Among the simplest mappings are those by a power $w = z^\alpha$. We consider only the case of real α, and then we may as well suppose that α is positive. Since

$$|w| = |z|^\alpha$$
$$\arg w = \alpha \arg z$$

concentric circles about the origin are transformed into circles of the same family, and half lines from the origin correspond to other half lines. The mapping is conformal at all points $z \neq 0$, but an angle θ at the origin is transformed into an angle $\alpha\theta$. For $\alpha \neq 1$ the transformation of the whole plane is not one to one, and if α is fractional z^α is not even single-valued. In general we can therefore only consider the mapping of an angular sector onto another.

The sector $S(\varphi_1,\varphi_2)$, where $0 < \varphi_2 - \varphi_1 \leqq 2\pi$, shall be formed by all points $z \neq 0$ such that one value of $\arg z$ satisfies the inequality

$$\varphi_1 < \arg z < \varphi_2. \tag{15}$$

It is easy to show that $S(\varphi_1,\varphi_2)$ is a region. In this region a unique value of $w = z^\alpha$ is defined by the condition

$$\arg w = \alpha \arg z$$

where $\arg z$ stands for the value of the argument singled out by the condition (15). This function is analytic with the nonvanishing derivative

$$De^{\alpha \log z} = \alpha \frac{w}{z}.$$

The mapping is one to one only if $\alpha(\varphi_2 - \varphi_1) \leqq 2\pi$, and in this case $S(\varphi_1,\varphi_2)$ is mapped onto the sector $S(\alpha\varphi_1,\alpha\varphi_2)$ in the w-plane. It should be observed that $S(\varphi_1 + n \cdot 2\pi, \varphi_2 + n \cdot 2\pi)$ is geometrically identical with $S(\varphi_1,\varphi_2)$ but may determine a different branch of z^α.

Let us consider the mapping $w = z^2$ in detail. Since $u = x^2 - y^2$ and $v = 2xy$, we recognize that the level curves $u = u_0$ and $v = v_0$ are equilateral hyperbolas with the diagonals and the coordinate axes for asymptotes. They are of course orthogonal to each other. On the other hand, the image of $x = x_0$ is $v^2 = 4x_0^2(x_0^2 - u)$ and the image of $y = y_0$ is $v^2 = 4y_0^2(y_0^2 + u)$. Both families represent parabolas with the focus at the origin whose axes are pointed in the negative and positive direction of the u-axis. Their orthogonality is well-known from analytic geometry. The families of level curves are shown in Figs. 3-5 and 3-6.

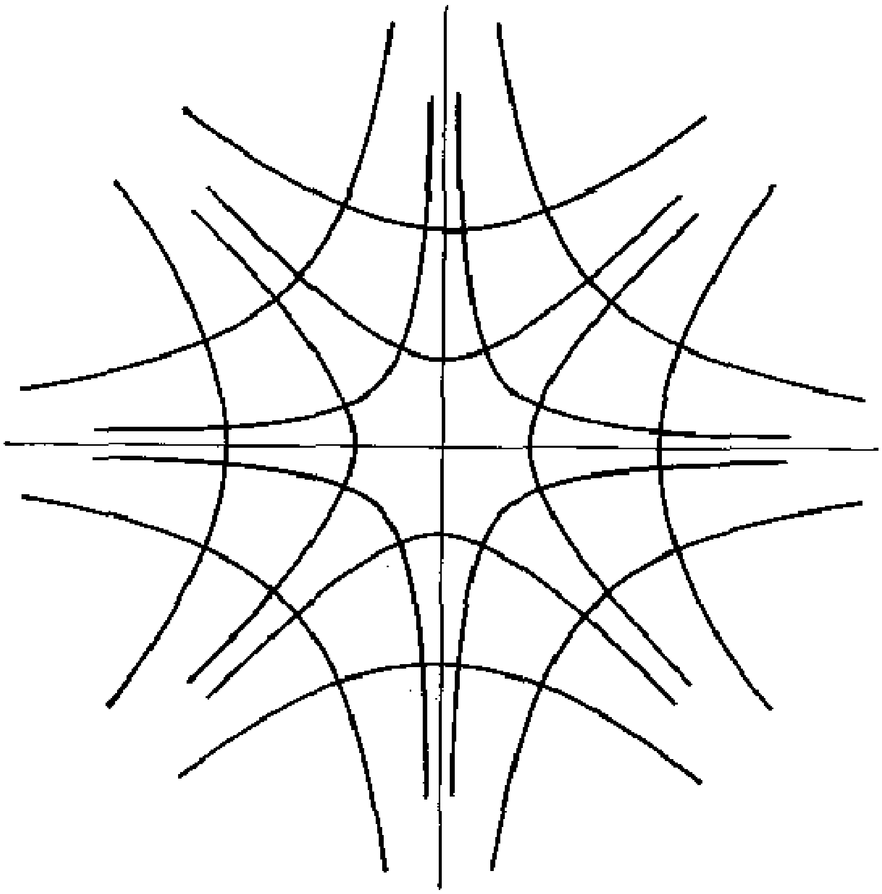

FIG. 3-5. z-plane.

For a different family of image curves consider the circles $|w - 1| = k$ in the w-plane. The equation of the inverse image can be written in the form

$$(x^2 + y^2)^2 = 2(x^2 - y^2) + k^2 - 1$$

and represents a family of lemniscates with the focal points ± 1. The orthogonal family is represented by

$$x^2 - y^2 = 2hxy + 1$$

and consists of all equilateral hyperbolas with center at the origin which pass through the points ± 1.

In the case of the third power $w = z^3$ the level curves in both planes are cubic curves. There is no point in deriving their equations, for their general shape is clear without calculation. For instance, the curves $u = u_0 > 0$ must have the form indicated in Fig. 3-7. Similarly, if we follow the change of arg w when z traces the line $x = x_0 > 0$, we find that the image curve must have a loop (Fig. 3-8). It is a folium of Descartes.

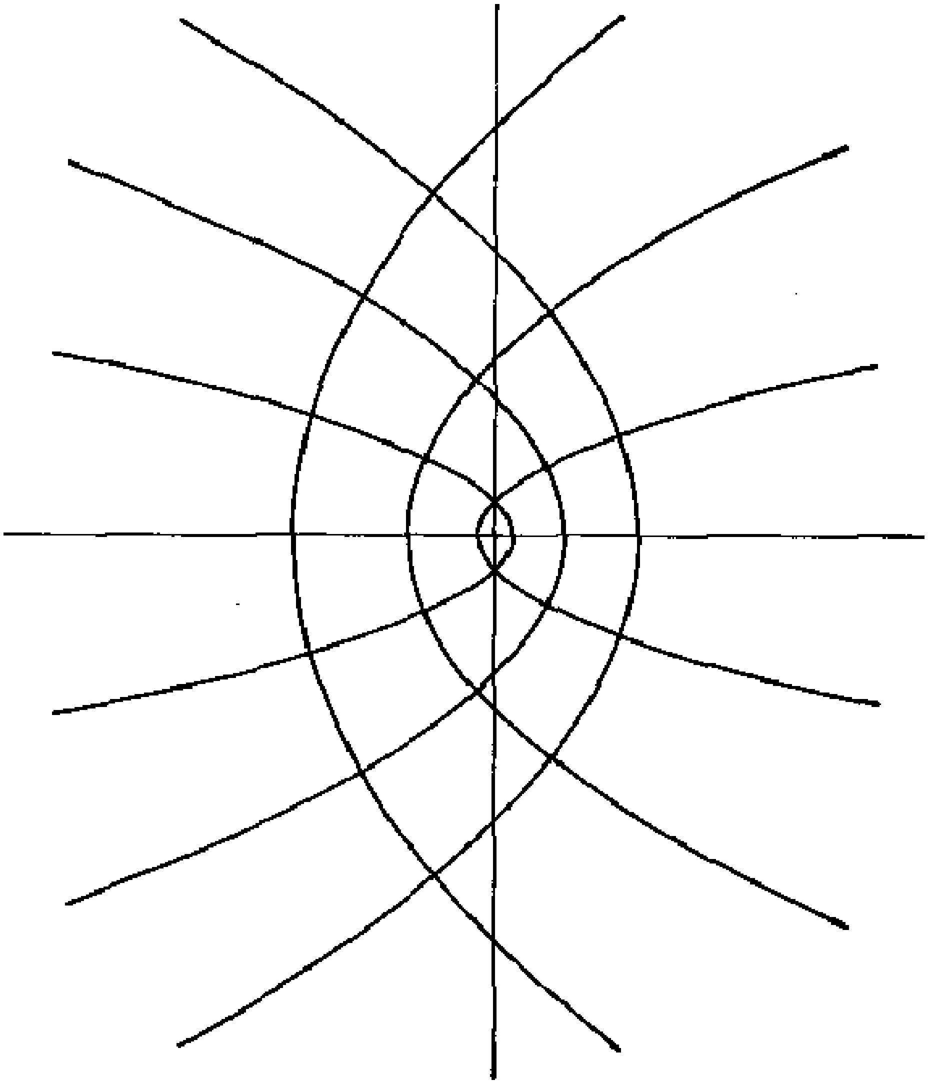

FIG. 3-6. w-plane.

The mapping by $w = e^z$ is very simple. The lines $x = x_0$ and $y = y_0$ are mapped onto circles about the origin and rays of constant argument. Any other straight line in the z-plane is mapped on a logarithmic spiral. The mapping is one to one in any region which does not contain two points whose difference is a multiple of $2\pi i$. In particular, a horizontal strip $y_1 < y < y_2$, $y_2 - y_1 \leq 2\pi$ is mapped onto an angular sector, and if $y_2 - y_1 = \pi$ the image is a half plane. We are thus able to map a parallel strip onto a half plane, and hence onto any circular region. The left half of the strip, cut off by the imaginary axis, corresponds to a half circle.

It is useful to write down some explicit formulas for the mapping. The function $\zeta = \xi + i\eta = e^z$ maps the strip $-\pi/2 < y < \pi/2$ onto the half plane $\xi > 0$. On the other hand,

$$w = \frac{\zeta - 1}{\zeta + 1}$$

maps $\xi > 0$ onto $|w| < 1$. Hence

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

maps the strip $|\text{Im } z| < \pi/2$ on the unit disk $|w| < 1$.

4.2. A Survey of Elementary Mappings. When faced with the problem of mapping a region Ω_1 conformally onto another region Ω_2, it is usually advisable to proceed in two steps. First, we map Ω_1 onto a circular region, and then we map the circular region onto Ω_2. In other words, the general problem of conformal mapping can be reduced to the problem of mapping a region onto a disk or a half plane. We shall prove, in Chap. 6, that this mapping problem has a solution for every region whose boundary consists of a simple closed curve.

The main tools at our disposal are linear transformations and transformations by a power, by the exponential function, and by the logarithm. All these transformations have the characteristic property that they map a family of straight lines or circles on a similar family. For this reason, their use is essentially limited to regions whose boundary is made up of circular arcs and line segments. The power serves the particular purpose of straightening angles, and with the aid of the exponential function we can even transform zero angles into straight angles.

By these means we can for instance find a standard mapping of any

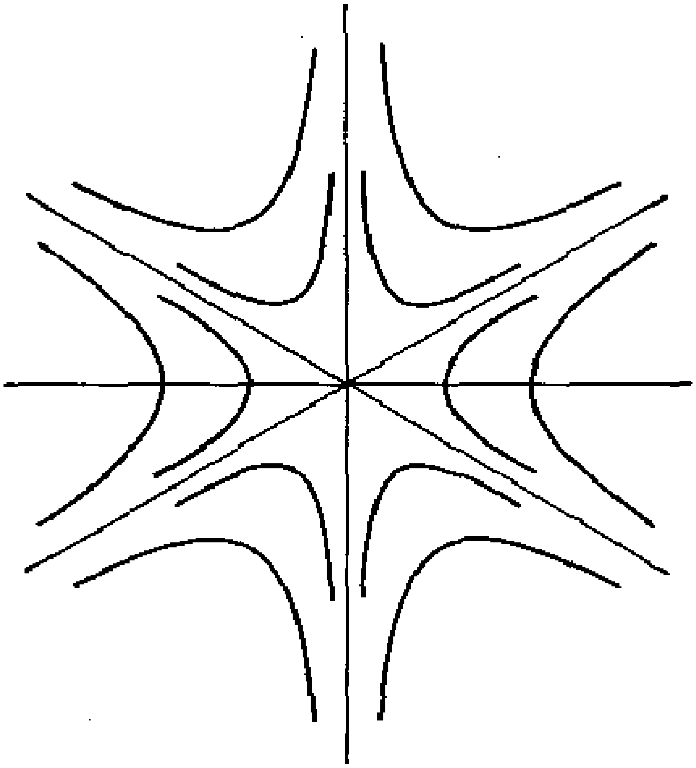

FIG. 3-7

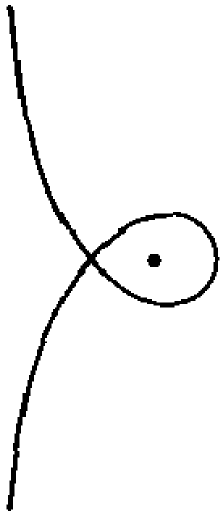

FIG. 3-8

region whose boundary consists of two circular arcs with common end points. Such a region is either a circular wedge, whose angle may be greater than π, or its complement. If the end points of the arcs are a and b, we begin with the preliminary mapping $z_1 = (z - a)/(z - b)$ which transforms the given region into an angular sector. By an appropriate power $w = z_1^\alpha$ this sector can be mapped onto a half plane.

If the circles are tangent to each other at the point a, the transformation $z_1 = 1/(z - a)$ will map the region between them onto a parallel strip, and a suitable exponential transformation maps the strip onto a half plane.

A little more generally, the same method applies to a circular triangle with two right angles. In fact, if the third angle has the vertex a, and if the sides from a meet again at b, the linear transformation $z_1 = (z - a)/(z - b)$ maps the triangle onto a circular sector. By means of a power this sector can be transformed into a half circle; the half circle is a wedge-shaped region which in turn can be mapped onto a half plane.

In this connection we shall treat explicitly a special case which occurs frequently. Let it be required to map the complement of a line segment onto the inside or outside of a circle. The region is a wedge with the angle 2π; without loss of generality we may assume that the end points of the segment are ± 1. The preliminary transformation

$$z_1 = \frac{z+1}{z-1}$$

maps the wedge on the full angle obtained by exclusion of the negative real axis. Next we define

$$z_2 = \sqrt{z_1}$$

as the square root whose real part is positive and obtain a map onto the right half plane. The final transformation

$$w = \frac{z_2 - 1}{z_2 + 1}$$

maps the half plane onto $|w| < 1$.

Elimination of the intermediate variables leads to the correspondence

$$z = \frac{1}{2}\left(w + \frac{1}{w}\right) \tag{16}$$
$$w = z - \sqrt{z^2 - 1}.$$

The sign of the square root is uniquely determined by the condition $|w| < 1$, for $(z - \sqrt{z^2 - 1})(z + \sqrt{z^2 - 1}) = 1$. If the sign is changed, we obtain a mapping onto $|w| > 1$.

For a more detailed study of the mapping (16) we set $w = \rho e^{i\theta}$ and obtain

$$x = \frac{1}{2}\left(\rho + \frac{1}{\rho}\right)\cos\theta$$
$$y = \frac{1}{2}\left(\rho - \frac{1}{\rho}\right)\sin\theta.$$

Elimination of θ yields

$$\frac{x^2}{[\frac{1}{2}(\rho + \rho^{-1})]^2} + \frac{y^2}{[\frac{1}{2}(\rho - \rho^{-1})]^2} = 1 \tag{17}$$

and elimination of ρ

$$\frac{x^2}{\cos^2\theta} - \frac{y^2}{\sin^2\theta} = 1. \tag{18}$$

Hence the image of a circle $|w| = \rho < 1$ is an ellipse with the major axis $\rho + \rho^{-1}$ and the minor axis $\rho^{-1} - \rho$. The image of a radius is half a branch of a hyperbola. The ellipses (17) and the hyperbolas (18) are confocal. The correspondence is illustrated in Fig. 3-9.

Clearly, the transformation (16) allows us to include in our list of elementary conformal mappings the mapping of the outside of an ellipse or the region between the branches of a hyperbola onto a circular region. It does not, however, allow us to map the inside of an ellipse or the inside of a hyperbolic branch.

As a final and less trivial example we shall study the mapping defined by a cubic polynomial $w = a_0z^3 + a_1z^2 + a_2z + a_3$. The familiar transformation $z = z_1 - a_1/3a_0$ allows us to get rid of the quadratic term,

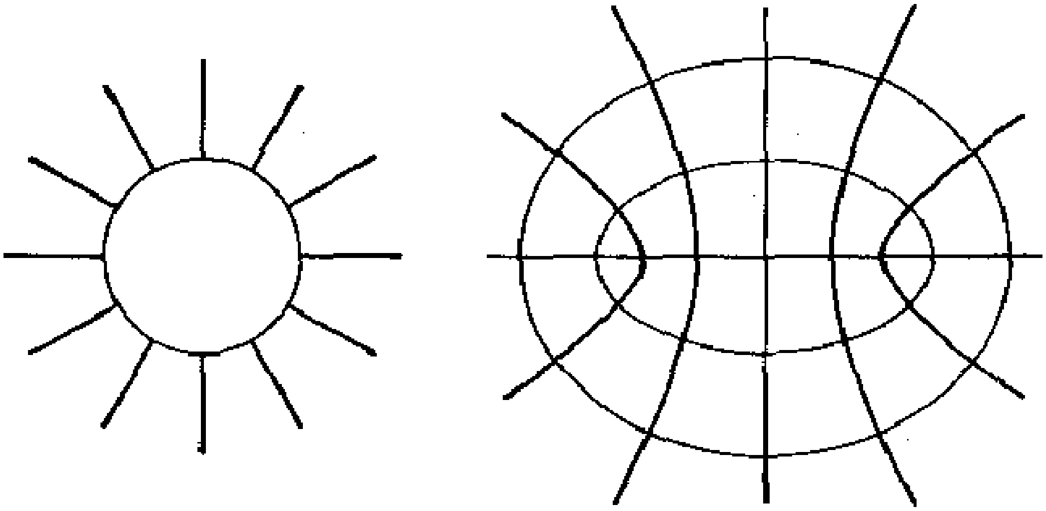

FIG. 3-9. Mapping by $z = \frac{1}{2}(w + w^{-1})$.

and by obvious normalizations we can reduce the polynomial to the form $w = z^3 - 3z$. The coefficient for z is chosen so as to make the derivative vanish for $z = \pm 1$.

Making use of the transformation (16) we introduce an auxiliary variable ζ defined by

$$z = \zeta + \frac{1}{\zeta}.$$

Our cubic polynomial takes then the simple form

$$w = \zeta^3 + \frac{1}{\zeta^3}.$$

We note that each z determines two values ζ, but they are reciprocal and yield the same value of w. In order to obtain a unique ζ we may impose the condition $|\zeta| < 1$, but then the segment $(-2,2)$ must be excluded from the z-plane.

It is now easy to visualize the correspondence between the z- and w-planes. To the circle $|\zeta| = \rho < 1$ corresponds an ellipse with the semiaxes $\rho^{-1} \pm \rho$ in the z-plane, and one with the semiaxes $\rho^{-3} \pm \rho^3$ in the w-plane. Similarly, a radius $\arg \zeta = \theta$ corresponds to hyperbolic branches in the z- and w-planes; the one in the z-plane has an asymptote which makes the angle $-\theta$ with the positive real axis, and in the w-plane the corresponding angle is -3θ. The whole pattern of confocal ellipses and hyperbolas remains invariant, but when z describes an ellipse w will trace the corresponding larger ellipse three times. The situation is thus very similar to the one in the case of the simpler mapping $w = z^3$. For orientation the reader may again lean on Fig. 3-9.

For the region between two hyperbolic branches whose asymptotes make an angle $\leq 2\pi/3$ the mapping is one to one. We note in particular that the six regions into which the hyperbola $3x^2 - y^2 = 3$ and the x-axis divide the z-plane are mapped onto half planes, three of them onto the upper half plane and three onto the lower. The inside of the right-hand branch of the hyperbola corresponds to the whole w-plane with an incision along the negative real axis up to the point -2.

EXERCISES

All mappings are to be conformal.

1. Map the common part of the disks $|z| < 1$ and $|z - 1| < 1$ on the inside of the unit circle. Choose the mapping so that the two symmetries are preserved.

2. Map the region between $|z| = 1$ and $|z - \frac{1}{2}| = \frac{1}{2}$ on a half plane.

3. Map the complement of the arc $|z| = 1, y \geq 0$ on the outside of the unit circle so that the points at ∞ correspond to each other.

4. Map the outside of the parabola $y^2 = 2px$ on the disk $|w| < 1$ so that $z = 0$ and $z = -p/2$ correspond to $w = 1$ and $w = 0$. (Lindelöf.)

5. Map the inside of the right-hand branch of the hyperbola $x^2 - y^2 = a^2$ on the disk $|w| < 1$ so that the focus corresponds to $w = 0$ and the vertex to $w = -1$. (Lindelöf.)

6. Map the inside of the lemniscate $|z^2 - a^2| = \rho^2 (\rho > a)$ on the disk $|w| < 1$ so that symmetries are preserved. (Lindelöf.)

7. Map the outside of the ellipse $(x/a)^2 + (y/b)^2 = 1$ onto $|w| < 1$ with preservation of symmetries.

8. Map the part of the z-plane to the left of the right-hand branch of the hyperbola $x^2 - y^2 = 1$ on a half plane. (Lindelöf.)

Hint: Consider on one side the mapping of the upper half of the region by $w = z^2$, on the other side the mapping of a quadrant by

$$w = z^3 - 3z.$$

4.3. Elementary Riemann Surfaces. The visualization of a function by means of the corresponding mapping is completely clear only when the mapping is one to one. If this is not the case, we can still give our imagination the necessary support by the introduction of generalized regions in which distinct points may have the same coordinates. In order to do this it is necessary to suppose that points which occupy the same place can be distinguished by other characteristics, for instance a tag or a color. Points with the same tag are considered to lie in the same *sheet* or *layer*.

This idea leads to the notion of a *Riemann surface*. It is not our intention to give, in this connection, a rigorous definition of this notion. For our purposes it is sufficient to introduce Riemann surfaces in a purely descriptive manner. We are free to do so as long as we use them merely for purposes of illustration, and never in logical proofs.

The simplest Riemann surface is connected with the mapping by a power $w = z^n$, where $n > 1$ is an integer. We know that there is a one-to-one correspondence between each angle $(k - 1)(2\pi/n) < \arg z < k(2\pi/n)$, $k = 1, \ldots, n$, and the whole w-plane except for the positive real axis. The image of each angle is thus obtained by performing a "cut" along the positive axis; this cut has an upper and a lower "edge." Corresponding to the n angles in the z-plane we consider n identical copies of the w-plane with the cut. They will be the "sheets" of the Riemann surface, and they are distinguished by a tag k which serves to identify the corresponding angle. When z moves in its plane, the corresponding