NATIONAL UNIVERSITY OF SINGAPORE

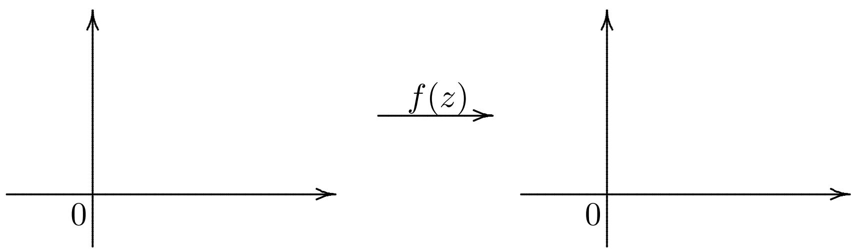
Department of Mathematics

MA4247 Complex Analysis II

Lecture Notes Part V

Chapter 3. Conformal Mappings, Linear Fractional Transformations

In this chapter, we study the geometric properties of analytic functions. Given an analytic function $f(z)$ defined on a domain D, we get a mapping $w = f(z)$ from the subset D of the z plane to the w plane. We will also study some mappings by elementary functions and an important class of mappings called the Möbius transformations or linear fractional transformations.



§3.1. Conformal properties of analytic functions

Let γ be a smooth (oriented) curve parametrized by

$$z = \gamma(t), \ a \leq t \leq b$$

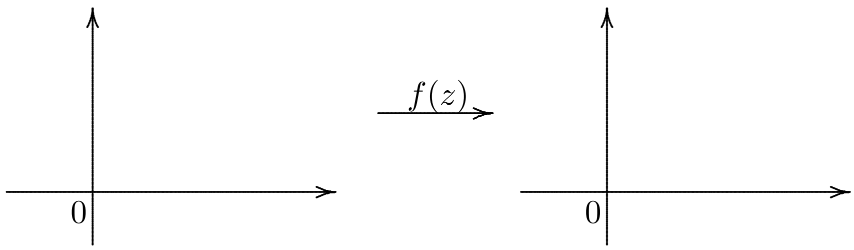
and let $f(z)$ be an analytic function on a domain D containing γ. Then the equation

$$w = f(\gamma(t)), \ a \leq t \leq b$$

is a parametric representation of the curve $f(\gamma)$. Suppose that $z_0 = \gamma(t_o)$, $a < t_o < b$ is a point on γ such that $f'(z_0) \neq 0$. By the chain rule,

$$w'(t_o) = f'(\gamma(t_o))\gamma'(t_o) \tag{1}$$

The assumption that γ is smooth implies that $\gamma'(t_o) \neq 0$, and thus $\gamma'(t_o)$ represents a tangent vector to γ at z_0. Since $f'(z_0) \neq 0$ and $\gamma'(t_o) \neq 0$, one sees from (1) that $w'(t_o) \neq 0$, and thus $w'(t_o)$ also represents a tangent vector to $f(\gamma)$ at $f(z_0)$ (we can think of each complex number as representating a vector in $\mathbb{R}^2$).

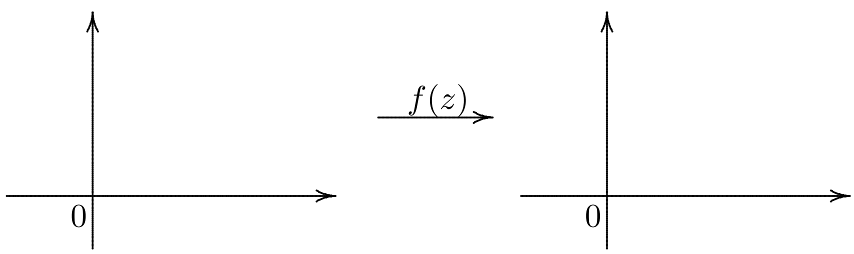


Taking the arguments in (1), we get

$$\arg[w'(t_o)] = \arg[f'(z_0)] + \arg[\gamma'(t_o)] \tag{2}$$

This means that the the direction of the directed tangent of $f(\gamma)$ at $f(z_0)$ is obtained by rotating the directed tangent of the curve γ at z_0 by the angle $\theta_o = \arg[f'(z_0)]$.
Crucial point: The amount of rotation only depends on $f'(z_0)$ and not on the direction of the original curve γ.
In particular, if we have two oriented smooth curves γ_1 and γ_2 intersecting at z_0 with (signed) angle ψ and $f(z)$ is analytic at z_0 with $f'(z_0) \neq 0$, then the image curves $f(\gamma_1)$ and $f(\gamma_2)$ also intersect at $f(z_0)$ with the (signed) angle ψ. Hence the map f preserves the angles (both the magnitude and the sense of the angle) at all points z_0 where $f'(z_0) \neq 0$. To summarize, we have

Proposition 3.1.1. (Angle Preserving Property) Let $f(z)$ be an analytic function on a domain D such that $f'(z_0) \neq 0$ at some point $z_0 \in D$. The the (signed) angle between two smooth curves γ_1, γ_2 in D at z_0 is equal to the angle between the curves $f(\gamma_1)$ and $f(\gamma_2)$ at $f(z_0)$.



In view of Proposition 3.1.1, we have

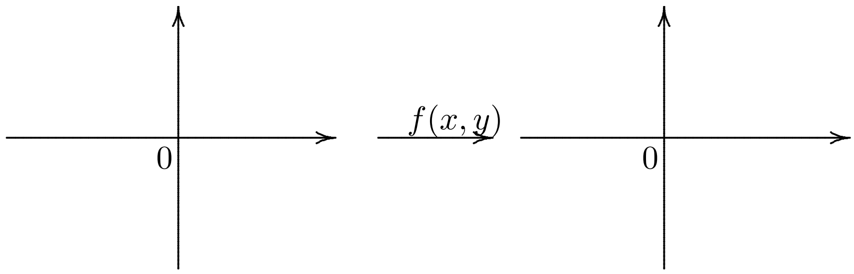
Definition 3.1.2. A function f on a domain D is said to be **conformal** at a point z_0 if f is analytic at z_0 and $f'(z_0) \neq 0$. A function f on a domain D is said to be a **conformal mapping** or **conformal transformation** if f is conformal at each point of D, i.e., if $f(z)$ is analytic on D and $f'(z) \neq 0$ at each $z \in D$.

Remark. (i) Observe that if f is conformal at a point $z_0 \in D$, then f is a conformal transformation on some open ball $B(z_0, r) \subset D$ with $r > 0$.

(ii) In some books (and sometimes in more general contexts), a function f is said to be conformal if it satisfies the angle preserving property in Proposition 3.1.1. Here we follow the definition of Churchill's book.

(iii) In the real case, Proposition 3.1.1 does not hold for general functions $f : \mathbb{R}^2 \to \mathbb{R}^2$. For example, consider the function

$$f(x, y) = (x, 2y) \quad \text{for } (x, y) \in \mathbb{R}^2.$$



Example 3.1.3.
(i)The mapping $w = f(z) = e^z$ is conformal on the entire complex plane.
(ii) The mapping $w = g(z) = \bar{z}$ is not conformal, it preserves angles but not the sense, it is called isogonal.
(iii) The mapping $w = f(z) = z^2$ is conformal everywhere except at $z = 0$ (since $f'(0) = 0$). The points where $f'(z) = 0$ are called the critical points of the mapping.

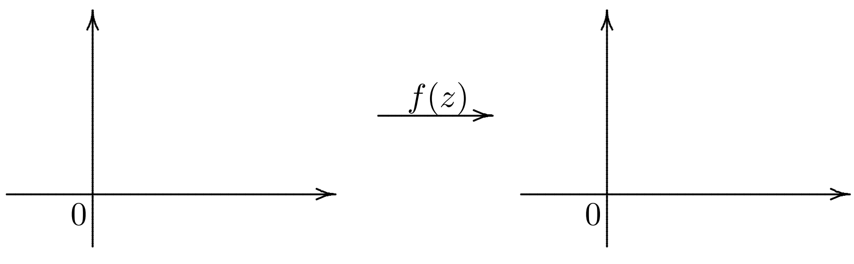
3.1.4. Infinitesimal circle preserving property of conformal mappings

Let D be a domain in $\mathbb{C}$. Suppose a function f is conformal at a point $z_0 \in D$. From the definition of the derivative, we have

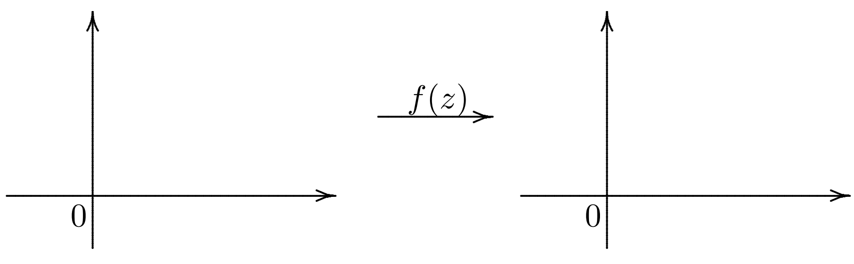
$$|f'(z_0)| = \left| \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} \right| = \lim_{z \to z_0} \frac{|f(z) - f(z_0)|}{|z - z_0|} \tag{3}$$

Now $|z - z_0|$ is the length of the line segment joining z to z_0 and $|f(z) - f(z_0)|$ is the length of the line segment joining $f(z)$ to $f(z_0)$. This means that any small line segment with one endpoint at z_0 is contracted or expanded by a ratio approximately equal to $|f'(z_0)|$. Again, it is important to note that the contraction or expansion is independent of the original direction of the line segment from z_0 to z. Thus for a very small circle C centered at z_0 and of radius r, $f(C)$ will roughly look like a circle centered at $f(z_0)$ and of radius $|f'(z_0)| \cdot r$. This property can also be described as "infinitesimal circles are mapped to infinitesimal circles".
The above tells us that for a small region R surrounding z_0 (say, R is a small circle centered at z_0), $f(R)$ has approximately the same shape as R. However, large regions may be distorted completely by f.

It is also important to note that while the angle of rotation and the dilation factor depend only on $f'(z_0)$, it varies from point to point in the domain D.



Remark 3.1.5.
1. If $f'(z)$ has a zero of order n at z_0, it can be shown that if γ_1 and γ_2 intersect at the angle α at z_0, then $f(\gamma_1)$ and $f(\gamma_2)$ will intersect at angle $(n+1)\alpha$ at $f(z_0)$ so that f does not have the angle preserving property at z_0.



2. Let D be a domain in $\mathbb{C}$, and let f be conformal at a point $z_0 \in D$ (so that $f'(z_0) \neq 0$). Recall that in the proof of the open mapping theorem, we have shown that if $z_0 \in D$ and $f(z_0) = w_0$, then since $f' \neq 0$, there exist $r > 0$ and $\delta > 0$ such that every $w \in B(w_0, r)$ is the image (under f) of a unique $z \in B(z_0, \delta)$. By choosing δ' such that $0 < \delta' < \delta$ and $f(B(z_0, \delta')) \subset B(w_0, r)$, then one sees that (the restriction of) f is a one-to-one map from $B(z_0, \delta')$ to its image. Note that f may not be one-to-one on D.

Definition 3.1.6. We say that a map $f : D \to \mathbb{C}$ is **locally one-to-one** if for every point $z_0 \in D$, we can find an open ball $B(z_0, \delta) \subset D$ with $\delta > 0$ such that the restriction of f to $B(z_0, \delta)$ is a one-to-one map from $B(z_0, \delta)$ to $\mathbb{C}$.

The above discussion gives the following

Proposition 3.1.7. (Locally one-to-one property)
A conformal mapping f on a domain D is locally one-to-one.

It can be shown that the angle preserving property together with some regularity conditions actually implies conformality, i.e., analyticity with non-zero derivatives. More precisely, we have

Theorem 3.1.8 . Let D be a domain in $\mathbb{C}$. Then a function $f : D \to \mathbb{C}$ is a conformal mapping if and only if f has the following properties:
(i) The partial derivatives u_x, u_y, v_x, y_y are continuous on D; and
(ii) f has the angle preserving property on D, i.e., for any point $z_0 \in D$ and any two smooth curves γ_1 and γ_2 in D passing through z_0, the (signed) angle between γ_1, γ_2 at z_0 is equal to the angle between the curves $f(\gamma_1)$ and $f(\gamma_2)$ at $f(z_0)$.

[In short, **conformal $\Longleftrightarrow$ continuous partial derivatives + angle preserving property**.]

Proof: 'only if' part. If f is conformal on D, then f is analytic on D, and it is known that (i) holds for any analytic functions on D (see e.g. [Churchill, p.161]). (ii) follows from Proposition 3.1.1.
Now we proceed to prove the 'if' part. Suppose f has properties (i) and (ii). We need to show that f is analytic on D and $f'(z) \neq 0$ for any $z \in D$. To prove the analyticity of f, in view of (i) and a theorem on sufficient condition for differentiability (see Lecture notes Part I, page 5), we only need to show that f satisfies the CR equations on D. To show that f satisfies the CR equations on D, we take any point $z_0 \in D$ and consider a smooth curve γ passing through z_0 and parametrized by a function $\gamma(t)$, $a \leq t \leq b$, with $\gamma(t_o) = z_0$ for some $a < t_o < b$.

Then $w(t) = f(\gamma(t))$, $a \le t \le b$, is a parametrization of $f(\gamma)$. Regard $f(z) = u(x,y) + iv(x,y)$ as a function in the variables x, y. Then the partial derivatives $f_x = u_x + iv_x$, etc, makes sense. By the Chain Rule,

$$w'(t_o) = \frac{\partial f}{\partial x}x'(t_o) + \frac{\partial f}{\partial y}y'(t_o)$$

where the partial derivatives of f are taken at z_0. Recall that $\gamma'(t_o) = x'(t_o) + iy'(t_o)$, so that we have

$$x'(t_o) = \frac{\gamma'(t_o) + \overline{\gamma'(t_o)}}{2}, \qquad y'(t_o) = \frac{\gamma'(t_o) - \overline{\gamma'(t_o)}}{2i}.$$

Using the above identities, we may rewrite

$$w'(t_o) = \frac{1}{2}(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y})\gamma'(t_o) + \frac{1}{2}(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y})\overline{\gamma'(t_o)}. \tag{4}$$

(Exercise.) If f has the angle preserving property, then

$$\arg[\frac{w'(t_o)}{\gamma'(t_o)}] = \arg w'(t_o) - \arg\gamma'(t_o)$$

must be independent of $\theta = \arg\gamma'(t_o)$. Hence

$$\begin{aligned}\frac{w'(t_o)}{\gamma'(t_o)} &= \frac{1}{2}(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}) + \frac{1}{2}(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y})\frac{\overline{\gamma'(t_o)}}{\gamma'(t_o)} \\ &= \frac{1}{2}(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}) + \frac{1}{2}(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y})e^{-2i\theta} \end{aligned} \tag{5}$$

must have a constant argument. As we vary the curve γ passing through z_0, $\theta = \arg\gamma'(t_o)$ will also vary, and the points represented by (5) will describe a circle centered at $\frac{1}{2}(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y})$ and of radius

$$\frac{1}{2}\Big|(\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y})\Big|.$$

The argument of $w'(t_o)/\gamma'(t_o)$ cannot be constant unless the radius is zero, so we have

$$\begin{aligned} &\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y} = 0 \\ \Longrightarrow\ & u_x + iv_x + i(u_y + iv_y) = (u_x - v_y) + i(v_x + u_y) = 0 \\ \Longrightarrow\ & u_x = v_y, \quad \& \quad u_y = -v_x. \end{aligned}$$

Hence f satisfies the CR equations on D. Together with (i), it follows that f is differentiable and hence analytic on D. Finally the fact that

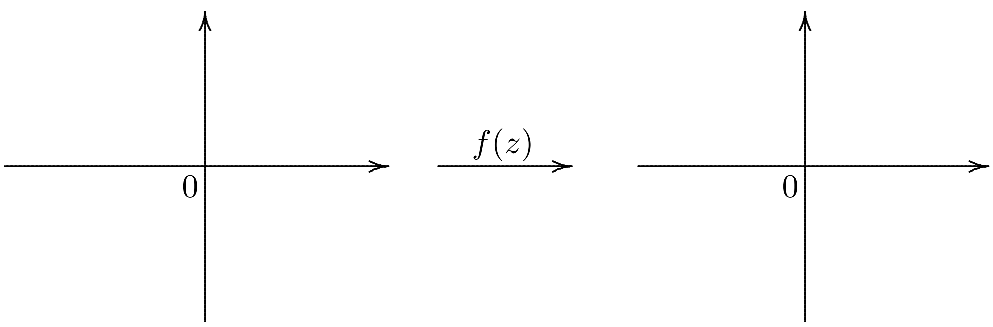
$f'(z)$ never vanishes on D follows from (i) and Remark 3.1.5. This finishes the proof of Theorem 3.1.8.

Remark-Definition 3.1.9. Let $f : U \to V$ be an analytic isomorphism from a domain U to another domain V, i.e., f is a bijective analytic map from U to V. Then by Corollary 2.4.6, $f'(z) \neq 0$ everywhere on U. Thus, f is also called a **conformal isomorphism** from U to V. Note that in this case, it follows from the Inverse Function Theorem (Theorem 2.4.9) that the inverse map f^{-1} is also analytic on V and $f^{-1}(w) \neq 0$ for all $w \in V$. Hence f^{-1} is also a conformal mapping.
Likewise, if U is analytically isomorphic to V, then we sometimes also say U is **conformally isomorphic** to V. Also, analytic isomorphisms are also called **conformal isomorphisms**.

3.1.10. Orthogonal families of curves, level curves

Two families of curves in the complex plane are said to be **orthogonal** if any two curves from the two families of curves always intersect at right angles whenever they intersect. Such two orthogonal families of curves are also called an **orthogonal net**.

For example: the horizontal lines (i.e. those lines of the form $y = y_o$) and the vertical lines (i.e. those lines of the form $x = x_o$) form an orthogonal families of curves (i.e. an orthogonal net).
Another example is the family of straight lines through the origin (i.e. those lines of the form $\arg z = \alpha$) and the family of circles centered at the origin (i.e. those curves of the form $|z| = r$). They also form an orthogonal net.

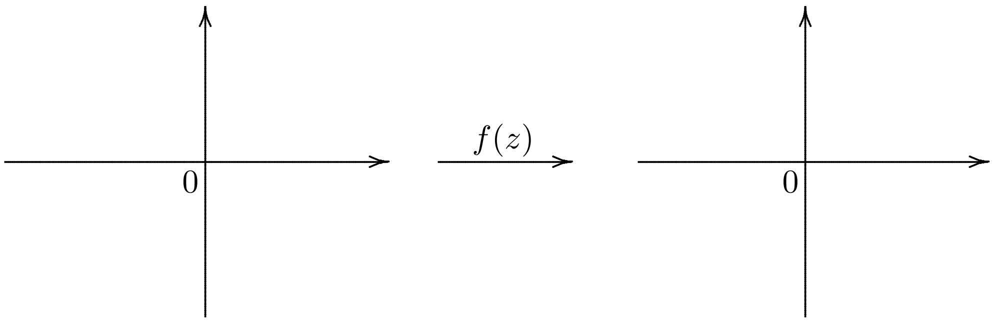


Let $w = f(z)$ be a conformal map. Write $z = x + iy$ and $w = u(x, y) + iv(x, y)$. From the angle-preserving property of f, one easily sees that the conformal map f transforms two orthogonal families of curves in the z-plane to two corresponding orthogonal family of curves in the w-plane (i.e. send an orthogonal net to an orthogonal net). For example we may use the horizontal and vertical lines in the z-plane and look at the image of these lines under the conformal map f. The image will form an orthogonal net in the w-plane. Similarly, we may wish to take the set of straight lines from the origin and the set of circles centered at the origin in the z-plane, and look at their images which again will form an orthogonal net in the w-plane.

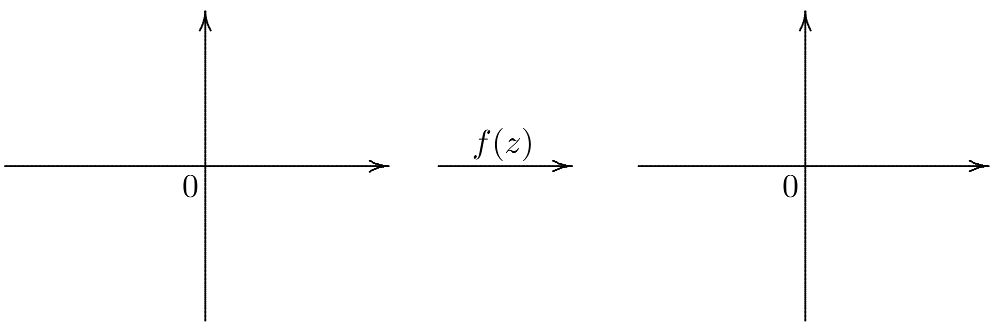
Alternatively, we may wish to take the horizontal and vertical lines in the w-plane and consider their inverse images in the z-plane, i.e., the set of points in the z-plane for which $u(x, y) = u_o$ (for a fixed u_o) as well as the set of points in the z-plane for which $v(x, y) = v_o$ (for a fixed v_o). These are called the **level curves** of u and v.

Example. Consider the mapping $w = z^2$. As usual, write $z = x + iy$, $w = u + iv$.

Consider the concentric circles and straight lines emanating from the origin in the z-plane. These are mapped to concentric circles and straight lines emanating from the origin in the w plane. Note that angles at the origin are multiplied by 2.

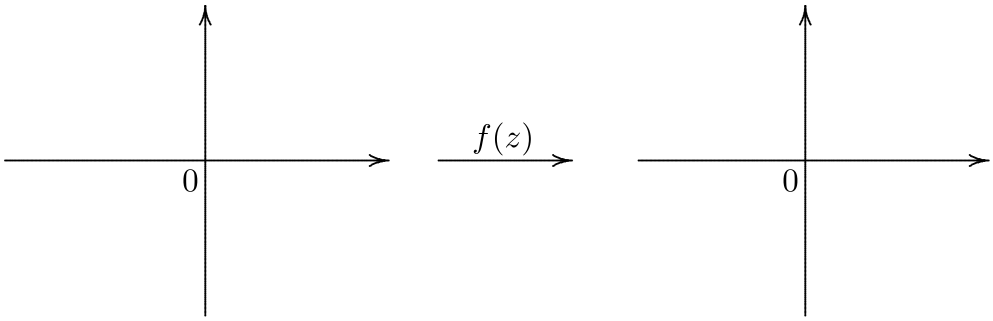


Now consider the image of the horizontal lines $x = a$ and vertical lines $y = b$. The lines $x = a$ are mapped to the curves $v^2 = 4a^2(a^2 - u)$ and the lines $y = b$ are mapped to the curves $v^2 = 4b^2(b^2 + u)$. Both families represent parabolas with focus at the origin and with axes in the negative and positive direction of the u-axis. They are orthogonal.



Now consider the level curves $u = a$ and $v = b$. Since $u = x^2 - y^2$ and $v = 2xy$, the level curves are two othogonal families of hyperbolas

$x^2 - y^2 = a$ and $2xy = b$.

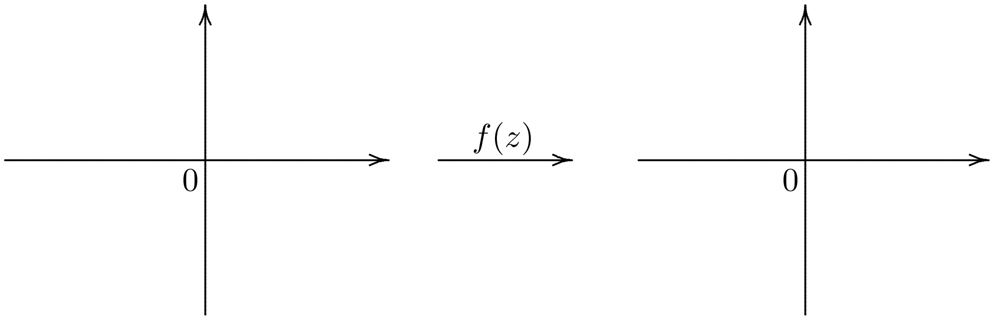


3.2 Mappings by elementary functions

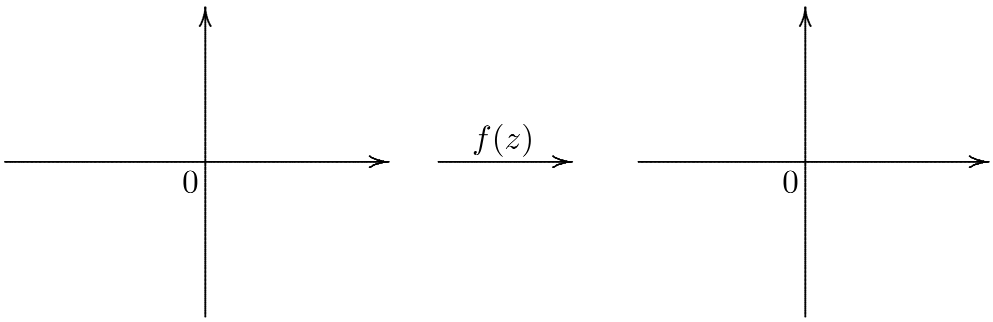
The exponential map $w = e^z$.
If $z = x + iy$, then $w = e^z = e^{x+iy}$ so

$$|w| = e^x \quad \text{and} \quad \arg\ w = y.$$

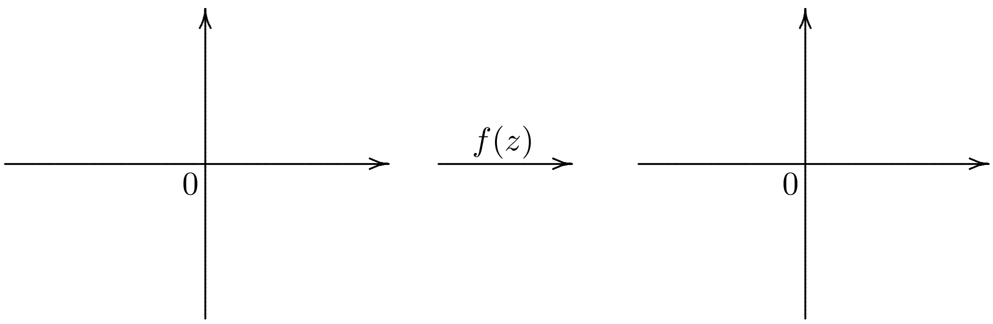
The infinite horizontal strip $\{x + iy \in \mathbb{C} : -\pi < y \leq \pi\}$ is mapped one-to-one to $\mathbb{C} - \{0\}$.



The (open) infinite horizontal strip $\{x + iy \in \mathbb{C} : 0 < y < \pi\}$ is mapped one-to-one to the (open) upper half plane.



Note that e^z is periodic with period $2\pi i$, that is, $e^{z+2\pi i} = e^z$ for all $z \in \mathbb{C}$.
Vertical lines $x = a$ are mapped to circles $|w| = e^a$, and horizontal lines $y = b$ are mapped to semi-infinite rays from the origin $\arg\ w = b$.



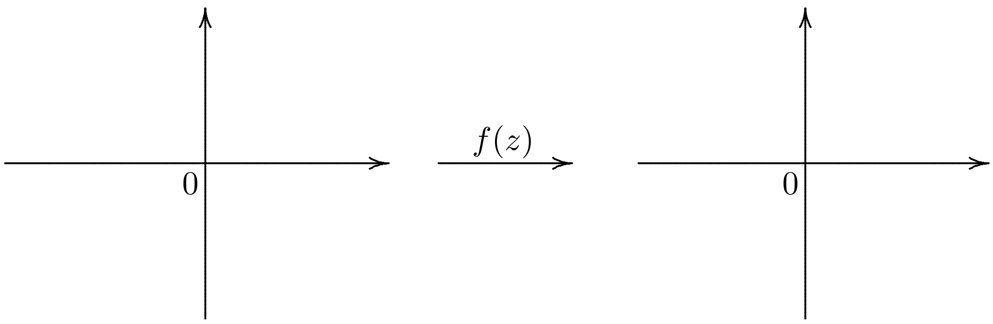
Power maps $w = z^a$.

For a positive real number a, and a suitable domain $D \subset \mathbb{C}$, we can define $w = z^a$ by $w = e^{a \log z}$, where $\log z$ is a branch of the log function (defined on D). Then

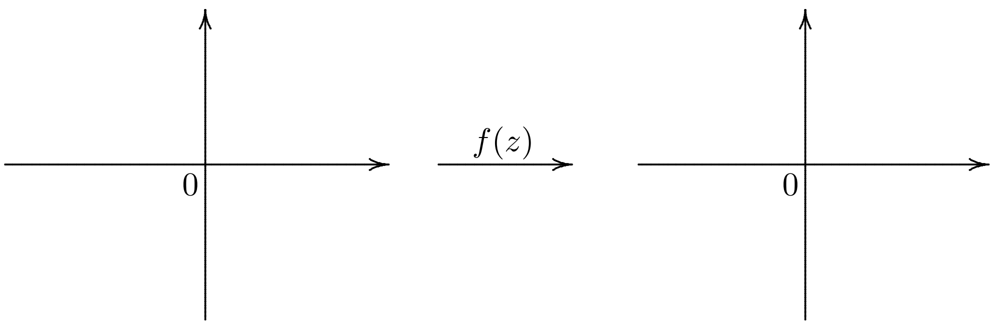
$$|w| = |z|^a, \qquad \arg\ w = a \cdot \arg\ z$$

Note that concentric circles about the origin are mapped to concentric circles about the origin, and rays from the origin are mapped to rays from the origin. A sector of angle β subtended at the origin is mapped to a sector of angle $a \cdot \beta$ subtended at the origin. The map is one to one on the sector if $a \cdot \beta < 2\pi$.

e.g. Consider the principal branch of $w = z^{\frac{1}{2}}$, i.e., let $w = e^{\frac{1}{2} \operatorname{Log} z}$.



e.g. Consider $w = z^3$.

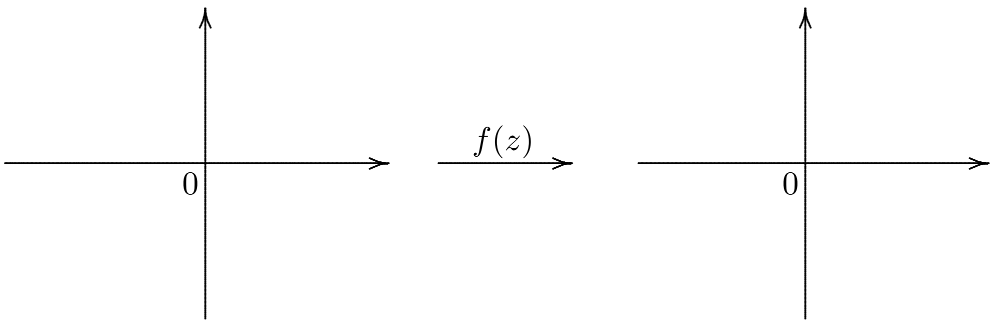


Log map

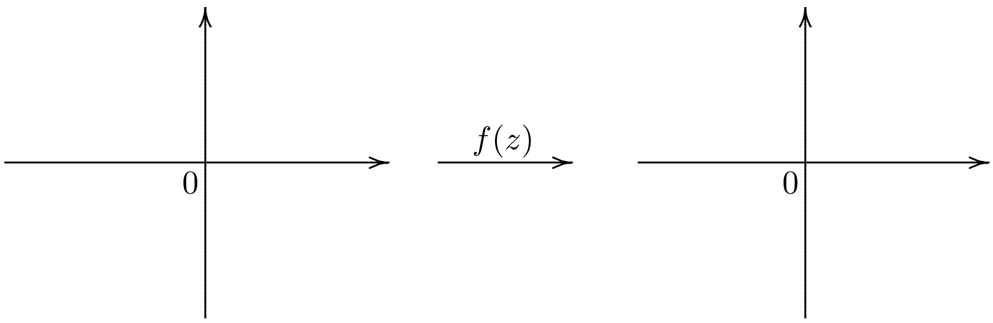
Branches of the logarithmic function $w = \log z$ can be used to map sectors to infinite strips. For example, the principal log map

$$w = \operatorname{Log} z = \ln |z| + i \operatorname{Arg} z$$

maps $\mathbb{C} \setminus \{\text{negative real axis}\}$ to the strip $-\pi < v < \pi$.



Likewise, $w = \operatorname{Log} z$ maps the upper half plane $\operatorname{Im} z > 0$ to the strip $0 < v < \pi$.



3.3. The Extended Complex plane $\hat{\mathbb{C}}$.

Recall that there is only one infinity ∞ for the complex plane $\mathbb{C}$. In studying certain type of conformal mappings, it is convenient to think of ∞ as a point, and we form the **extended complex plane** given by $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. A basis for the open neighborhoods of ∞ is given by $\{z \in \mathbb{C} \;:\; |z| > r\} \cup \{\infty\}$, where $r > 0$.

Let S^2 denote the sphere in $\mathbb{R}^3$ of radius one and with centre at the origin, i.e.,

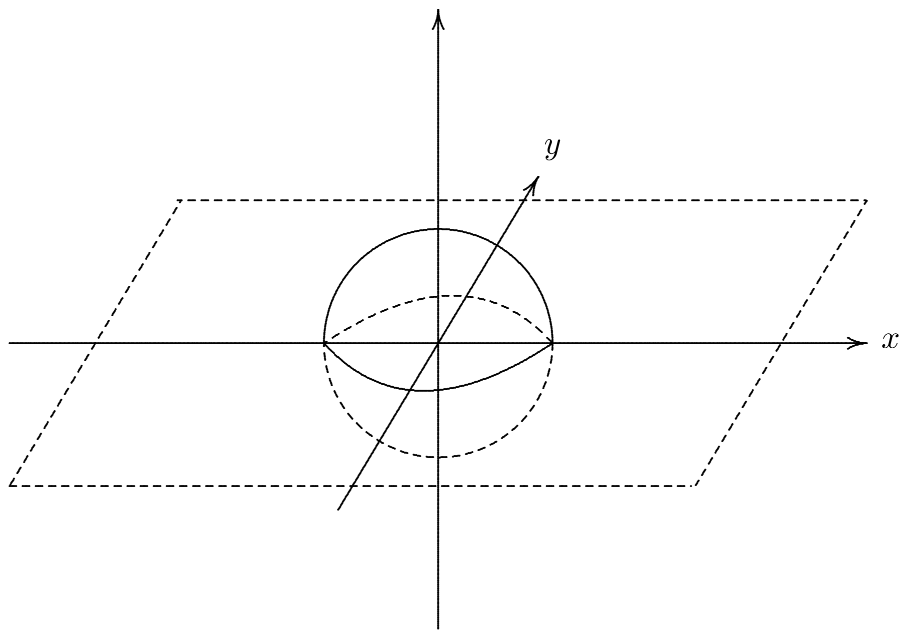
$$S^2 = \{(x, y, t) \in \mathbb{R}^3 : x^2 + y^2 + t^2 = 1\}.$$

It is common to identify $\hat{\mathbb{C}}$ with the sphere S^2 via the **stereographic projection** map as follows:

Identify $\mathbb{C}$ with the xy-plane in $\mathbb{R}^3$ given by

$$z = x + iy \in \mathbb{C} \quad \longleftrightarrow \quad (x, y, 0) \in \mathbb{R}^3.$$

Take any point $z = x + iy \leftrightarrow (x, y, 0)$. Then we can form the straight line in $\mathbb{R}^3$ passing through $(x, y, 0)$ and the north pole $N = (0, 0, 1)$ of S^2. This line meets the sphere S^2 at exactly one other point u. Using this construction, we get a map St from $\mathbb{C}$ to $S^2 - \{(0, 0, 1)\}$ given by $St(z) = u$, called the **stereographic projection** of $\mathbb{C}$ to the sphere S^2. We can extend this to $\hat{\mathbb{C}}$ by mapping ∞ to the north pole $(0, 0, 1)$ thus giving a one-to-one onto map of $\hat{\mathbb{C}}$ to S^2. S^2 is called the **Riemann sphere**.



Exercise: Show that the stereographic projection map is given by

$$St(x+iy) = \Big(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}, \frac{x^2+y^2-1}{1+x^2+y^2}\Big). \tag{1}$$

A straight line in $\mathbb{C}$ together with ∞ will be called an **extended line** of $\mathbb{C}$.

Proposition 3.3.1. The stereographic projection map St maps circles and extended lines of $\mathbb{C}$ bijectively to circles on S^2, and vice versa. Moreover, St maps extended lines of $\mathbb{C}$ to circles on S^2 passing through the north pole.

Sketch of proof. Using the geometric fact that the intersection of S^2 with a plane in $\mathbb{R}^3$ is always a circle in $\mathbb{R}^3$, one easily sees that St will map an extended line to a circle containing the north pole. More generally, a circle on S^2 lies on a plane

$$aX + bY + cT = d, \tag{2}$$

where we may assume that $a^2+b^2+c^2 = 1$ and $0 \le a < 1$. Substitute (1) into (2), it follows that

$$\begin{aligned} a\cdot\frac{2x}{1+x^2+y^2} + b\cdot\frac{2y}{1+x^2+y^2} + c\cdot\frac{x^2+y^2-1}{1+x^2+y^2} = d \\ \iff \quad (d-c)(x^2+y^2) - 2ax - 2ay + (d+c) = 0, \end{aligned} \tag{3}$$

which corresponds to a circle in the z-plane if $d \neq c$ and an extended line in the z-plane if $d = c$. By reversing the argument, one easily sees that extended lines and circles are mapped to circle on S^2. $\square$

With the extended complex plane $\hat{\mathbb{C}}$, it is useful to define algebraic operations involving ∞ which would agree with the limiting operations. We have

Definition.
(i) If $z \in \mathbb{C}$, then $z \pm \infty = \infty \pm z = \infty$, $\dfrac{z}{\infty} = 0$, $\dfrac{\infty}{z} = \infty$;
(ii) If $z \in \mathbb{C}$, $z \neq 0$, then $z \cdot \infty = \infty \cdot z = \infty$.

3.4. Möbius/Linear fractional transformations.

Definition 3.4.1. A **Möbius transformation/map** (also called **linear fractional transformation** (LFT)) is any function of the form

$$w = f(z) = \frac{az+b}{cz+d}, \quad \text{where } ad - bc \neq 0. \tag{1}$$

Examples:

1. $w = f(z) = \dfrac{1}{z}$ (reciprocal map or inversion).

Here $w = \dfrac{0z+1}{1z+0}$ corresponds to $a = 0$, $b = 1$, $c = 1$, $d = 0$, so that $ad - bc = 0 - 1 = -1 \neq 0$.

2. $w = f(z) = az$, $a \neq 0$ (rotation by $\theta = \arg\ a$ followed by dilation by $|a|$ about 0).

Here $w = \dfrac{az+0}{0z+1}$ corresponds to $a = a \neq 0$, $b = 0$, $c = 1$, $d = 0$, so that $ad - bc = a \neq 0$.

3. $w = f(z) = z + c$ (translation by c).

Here $w = \dfrac{1z+c}{0z+1}$, and $ad - bc = 1 \neq 0$.

4. $w = f(z) = az + c$ with $a \neq 0$ (rotation and dilation about 0 followed by translation by c).

Here $w = \dfrac{az+c}{0z+1}$, and $ad - bc = a \neq 0$.

3.4.2. Basic properties.

1. If $w = f(z)$, is a Möbius transformation, then w and z satisfy a second degree equation of the form

$$Awz + Bw + Cz + D = 0, \quad \text{with } AD - BC \neq 0.$$

Proof: Exercise.

2. The transformation

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

is not changed if we multiply the coefficients a, b, c and d by the same non-zero constant k.

3. The Möbius transformation has a pole at $-d/c$, if $c \neq 0$. We can extend the transformation to a transformation from $\hat{\mathbb{C}}$ to $\hat{\mathbb{C}}$ by defining $f(-d/c) = \infty$ and $f(\infty) = a/c$ (if $-d/c \neq \infty$). This is motivated by defining the functions to be equal to the limits of the functions as z approaches these points, i.e.,

$$f(-\frac{d}{c}) = \lim_{z\to -d/c} f(z) = \lim_{z\to -d/c} \frac{az+b}{cz+d} = \infty \quad (\mathit{Exercise}) \qquad (2)$$

and

$$f(\infty) = \lim_{z\to\infty} f(z) = \lim_{z\to\infty} \frac{az+b}{cz+d} = \frac{a}{c}. \qquad (3)$$

(If $c = 0$, then we just define $f(\infty) = \infty$.)

Proposition 3.4.3. With this extension, $w = f(z)$ is then a bijective map from $\hat{\mathbb{C}}$ to $\hat{\mathbb{C}}$ itself.

Proof. We need to show that $w = f(z) : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is one-to-one and onto. We first consider the case when $c \neq 0$. To prove that f is one-to one, we need to show that if $f(z) = f(z')$, then $z = z'$ for $z, z' \in \hat{\mathbb{C}}$.
Case (i): If $f(z) = f(z') = \infty$, then from (1), (2) and (3), we see that $z = z' = -d/c$.
Case (ii): Suppose $f(z) = f(z') = \frac{a}{c}$. By (3), we have $f(\infty) = \frac{a}{c}$. For any finite $z \in \mathbb{C}$, we have, from (1),

$$\begin{aligned} f(z) = \frac{a}{c} &\Longrightarrow \frac{az+b}{cz+d} = \frac{a}{c} \\ &\Longrightarrow az + bc = a(cz+d) \\ &\Longrightarrow bc - ad = 0, \end{aligned}$$

which contradicts the assumption that $ad - bc \neq 0$. Hence we must have $z = z' = \infty$.

Case (iii): Suppose $f(z) = f(z')$ (is finite and $\neq \frac{a}{c}$). Then from equations (2) and (3), one knows that neither z nor z' is ∞ or $-\frac{d}{c}$, so that $f(z)$ and $f(z')$ are both given by (1). Then we have

$$\begin{aligned} f(z) = f(z') &\Longrightarrow \frac{az+b}{cz+d} = \frac{az'+b}{cz'+d} \\ &\Longrightarrow (az+b)(cz'+d) = (az'+b)(cz+d) \\ &\Longrightarrow (ad-bc)(z-z') = 0 \\ &\Longrightarrow z = z' \quad (\text{since } ad - bc \neq 0). \end{aligned}$$

Hence f is injective on $\hat{\mathbb{C}}$. To prove the surjectivity, we consider the three cases:
Case (α): $w = \infty$. In this case, by (2), we have $f(-\frac{d}{c}) = \infty$.
Case (β): $w = \frac{a}{c}$. In this case, by (3), we have $f(\infty) = \frac{a}{c}$.
Case (γ): $w \in \mathbb{C}$ and $w \neq \frac{a}{c}$. In this case, by (1), we solve:

$$\begin{aligned} f(z) = w &\Longleftrightarrow \frac{az+b}{cz+d} = w \\ &\Longleftrightarrow az + b - (cz+d)w = 0 \\ &\Longleftrightarrow (-cw+a)z = dw - b \\ &\Longleftrightarrow z = \frac{dw-b}{-cw+a}, \end{aligned} \tag{4}$$

which is finite, if $w \neq \frac{a}{c}$. Hence, we have $f\left(\frac{dw-b}{-cw+a}\right) = w$. Combining the 3 cases, we conclude that $f(z)$ is surjective on $\hat{\mathbb{C}}$.
The case when $c = 0$ is similar, and is left as an exercise (Try it!)

4. $w = f(z) = \dfrac{az+b}{cz+d}$ has an inverse which is also a Möbius transformation, namely,

$$z = f^{-1}(w) = \frac{dw-b}{-cw+a} \tag{5}$$

This follows from the calculation in equation (4) above. Exercise: Check that the Möbius transformation in equation (5) is the inverse of $f(z)$ as a map on $\hat{\mathbb{C}}$, i.e., check that

$$f \circ f^{-1}(z) = f^{-1} \circ f(z) = z \quad \text{for all } z \in \hat{\mathbb{C}}.$$

5. $f(z)$ is conformal map on $\mathbb{C} \setminus \{-d/c\}$ if $c \neq 0$, since

$$f'(z) = \frac{ad-bc}{(cz+d)^2} \neq 0, \quad \text{if } z \neq -\frac{d}{c}.$$

If $c = 0$, then $f(z)$ is a conformal map on $\mathbb{C}$ (Exercise).

6. The composition of two Möbius transformations is again a Möbius transformation.

Proof. If

$$f(z) = \frac{a_1 z + b_1}{c_1 z + d_1} \quad \text{and} \quad g(z) = \frac{a_2 z + b_2}{c_2 z + d_2},$$

with $a_1 d_1 - b_1 c_1 \neq 0$ and $a_2 d_2 - b_2 c_2 \neq 0$, then

$$\begin{aligned}(f \circ g)(z) &= \frac{a_1 g(z) + b_1}{c_1 g(z) + d_1} \\ &= \frac{a_1 \frac{a_2 z + b_2}{c_2 z + d_2} + b_1}{c_1 \frac{a_2 z + b_2}{c_2 z + d_2} + d_1} \\ &= \frac{(a_1 a_2 + b_1 c_2) z + (a_1 b_2 + b_1 d_2)}{(c_1 a_2 + d_1 c_2) z + (c_1 b_2 + d_1 d_2)} \\ &= \frac{Az + B}{Cz + D}, \qquad (*)\end{aligned}$$

where

$$\begin{aligned}AD - BC &= (a_1 a_2 + b_1 c_2)(c_1 b_2 + d_1 d_2) - (a_1 b_2 + b_1 d_2)(c_1 a_2 + d_1 c_2) \\ &= (a_1 d_1 - b_1 c_1)(a_2 d_2 - b_2 c_2) \neq 0.\end{aligned}$$

A more conceptual approach:
Any element $z \in \hat{\mathbb{C}}$ can be represented by a (non-zero) column vector $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2$ with $z = \dfrac{z_1}{z_2}$. For example, any $z \in \mathbb{C}$ can be represented by $\begin{pmatrix} z \\ 1 \end{pmatrix}$. The ideal point ∞ can be represented by $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Obviously two column vectors represent the same element in $\hat{\mathbb{C}}$ if the two column vectors are non-zero multiples of each other, i.e., $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim \begin{pmatrix} kz_1 \\ kz_2 \end{pmatrix}$, $k \neq 0$. $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ is called homogeneous coordinates of z in $\hat{\mathbb{C}}$. Now if $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ are homogeneous coordinates of w and $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ are homogeneous coordinates of z, then in terms of homogeneous coordinates, $w = f(z)$ can be written in the the form of the matrix equation

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

Using this notation, the Möbius transformation is represented by a non-singular 2×2 matrix (determinant $\neq 0$) and it is easy to see that composition of Möbius mappings correspond to taking the product of matrices (and taking inverse of a Möbius mapping corresponds to taking the inverse of the corresponding matrix). More precisely, if a Möbius transformation f corresponds to the matrix $\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ and a Möbius transformation g corresponds to the matrix $\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$, then $f \circ g$ will correspond to the matrix

$$\begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1 a_2 + b_1 c_2 & a_1 b_2 + b_1 d_2 \\ c_1 a_2 + d_1 c_2 & c_1 b_2 + d_1 d_2 \end{pmatrix}.$$

Since product of two non-singular 2×2 matrice is a non-singular 2×2 matrix, it follows that composition of two Möbius transformations is a Möbius transformation.

7. **Theorem 3.4.4.** The set of Möbius transformations (regarded as bijective mappings from $\hat{\mathbb{C}}$ to itself) form a group under composition (i.e., under the binary operation in (*) above).

Remark. Recall that a set G together with a binary operation $\cdot$ is said to form a group if the following conditions are satisfied:
(G-i) (associativity) $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3 \quad \forall\, g_1, g_2, g_3 \in G$;
(G-ii) (existence of identity) There exists an element $e \in G$ such that

$$g \cdot e = e \cdot g = g \quad \forall g \in G; \quad \text{and}$$

(G-iii) (existence of inverse) For each $g \in G$, there exists $g^{-1} \in G$ such that

$$g \cdot g^{-1} = g^{-1} \cdot g = e.$$

Proof of Theorem 3.4.4. We have seen in (6) that the composition $f \circ g$ of two Möbius transformations f and g is again a Möbius transformation, i.e., it is a binary operation on the set of Möbius transformation. Then the associativity (G-i) follows from the associativity of the composition of maps (it also follows from the associativity of matrix multiplication). It is easy to see that the identity element in the set of Möbius transformations is given by the identity mapping $f(z) = z$, which gives (G-ii). We also have checked that the inverse of a Möbius transformation exists a Möbius transformation given by equation (5), which gives (G-iii).

8. **Proposition 3.4.5.** (i) Two non-singular 2×2 matrices give rise to the same Möbius transformation (i.e. the same map on $\hat{\mathbb{C}}$) if and only if they are non-zero multiples of each other.
(ii) Each Möbius transformation is represented by exactly two matrices whose determinants are equal to 1. Moreover, the two matrices are minuses of each other (i.e., they are of the form A and $-A$).

Proof. (i) 'If' part. It is easy to see that if two non-singular 2×2 matrices are multiples of each other, then their corresponding Möbius transformations are the same map on $\hat{\mathbb{C}}$ (Exercise, see (2)).
'Only if' part. Let A and B be two non-singular 2×2 matrices whose corresponding Möbius transformations are the same map on $\hat{\mathbb{C}}$. It is easy to see that their inverses A^{-1} and B^{-1} represent the inverse of the given Möbius transformation (as a map on $\hat{\mathbb{C}}$). Thus, the matrix $C = A \circ B^{-1}$ represent the identity map on $\hat{\mathbb{C}}$. Write $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then we have

$$\begin{aligned} \frac{az+b}{cz+d} &= z \quad \text{for all} z \in \mathbb{C} \setminus \{-\frac{d}{c}\} \\ \implies cz^2 + (d-a)z + b = 0 &\quad \text{for all} z \in \mathbb{C} \setminus \{-\frac{d}{c}\} \\ \implies c = &0, \ d = a, \ \text{and } b = 0, \end{aligned}$$

since a non-zero quadratic polynomial can have at most 2 distinct roots. Hence

$$C = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

with $\det(C) = a^2 \neq 0 \implies a \neq 0$. Denote by I the 2×2 identity matrix. Thus, $AB^{-1} = aI \implies A = aIB = aB$ with $a \neq 0$. This finishes the proof of (i). To prove (ii), we take any non-singular 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ which represents a Möbius transformation on $\hat{\mathbb{C}}$. Let s be a square root of $\det(A)$, so that $s^2 = ad - bc \neq 0$. Then one easily sees that both $\frac{1}{s}A = \begin{pmatrix} a/s & b/s \\ c/s & d/s \end{pmatrix}$ and $-\frac{1}{s}A = \begin{pmatrix} -a/s & -b/s \\ -c/s & -d/s \end{pmatrix}$ are of determinant 1 and by (i), one knows that they lead to the same map on $\hat{\mathbb{C}}$ as A. Finally suppose that A and B are non-singular 2×2 matrices such that $\det(A) = \det(B) = 1$ and they lead to the same map on $\hat{\mathbb{C}}$. By (i), one knows that $A = kB$ for some complex number $k \neq 0$. But then $1 = \det(A) = k^2 \det B = k^2 \cdot 1 \implies k = \pm 1$. Hence $A = B$ or $A = -B$. This finishes the proof of (i).