UNIVERSITY OF Cincinnati

Department of Mathematical Sciences 839 Old Chemistry Building
PO Box 210025 Phone (513) 556-4075
Cincinnati OH 45221-0025 Fax (513) 556-3417

COMPLEX ANALYSIS NOTES
AUTUMN QUARTER 2009

DAVID A HERRON

Contents

Date: December 3, 2009.

1. Elementary Things

1.A. **The Beginning.** What are complex numbers, and what can we do with them?

What is $\mathbb{C}$? From an informal algebraic point of view, $\mathbb{C}$ is the *field* of complex numbers $z = x + iy$, $w = u + iv$, $\zeta = \xi + i\eta$ (etc.) where the laws of complex arithmetic hold. Here $u, v, x, y, \xi, \eta \in \mathbb{R}$ are real numbers and $i^2 = -1$. Unlike the field $\mathbb{R}$ of real numbers, $\mathbb{C}$ is not an ordered field. Of course, in order to have a field there must exist both additive and multiplicative inverses.

For example, when $z \neq 0$, $z^{-1} = 1/z = \bar{z}/|z|^2$. Here $\bar{z}$ is the *complex conjugate* of z and $|z|$ is the *absolute value* (or *modulus*) of z; when $z = x + iy$, $\bar{z} := x - iy$ and $|z| := \sqrt{x^2 + y^2}$. The map $w = 1/z$ is called *complex inversion*. Notice that $z\bar{z} = |z|^2$.

Here is some more notation: when $z = x + iy$, $\mathfrak{Re}(z) := x$ and $\mathfrak{Im}(z) := y$ are called the *real part* and *imaginary part* of z (respectively). A complex number z is *purely real* if $\mathfrak{Im}(z) = 0$ and is *purely imaginary* when $\mathfrak{Re}(z) = 0$. We identify the real number field $\mathbb{R}$ with $\{z \in \mathbb{C} : \mathfrak{Im}(z) = 0\}$ and write $i\mathbb{R}$ for $\{z \in \mathbb{C} : \mathfrak{Re}(z) = 0\}$. Note the useful inequalities

$$|\,\mathfrak{Re}(z)| \leq |z| \quad \text{and} \quad |\,\mathfrak{Im}(z)| \leq |z|\,.$$

From a geometric point of view, we can picture $\mathbb{C}$ as a Euclidean plane by using cartesian coordinates: the complex number $z = x + iy$ is identified with the point (x, y). In this setting the x-axis is called the *real axis*, which is just $\mathbb{R} \subset \mathbb{C}$, and the y-axis, the so-called *imaginary axis*, is $i\mathbb{R} \subset \mathbb{C}$. Notice that $\bar{z}$ is the reflection of z across the real axis and $|z|$ is the distance from z to the origin. Three other sets which are worthy of special designations are the *unit disk* $\mathbb{D} := \{z : |z| < 1\}$, the *unit circle* $\mathbb{T} := \{z : |z| = 1\}$, and the *right half-plane* $\mathbb{H} := \{z : \mathfrak{Re}(z) > 0\}$. Also, $\mathbb{C}_* := \mathbb{C} \setminus \{0\}$ and $\mathbb{D}_* := \mathbb{D} \setminus \{0\}$ are the punctured plane and punctured disk.

1.1. *Example.* Let $w = T(z) = \dfrac{1 - z}{1 + z} = \dfrac{1 - |z|^2}{|1 + z|^2} - i\dfrac{2\,\mathfrak{Im}(z)}{|1 + z|^2}$.
Notice that T maps $\mathbb{R} \setminus \{-1\}$ bijectively onto itself and $i\mathbb{R}$ bijectively onto $\mathbb{T} \setminus \{-1\}$. Also, $|z| < 1$ if and only if $\mathfrak{Re}(w) > 0$ and $|w| < 1$ if and only if $\mathfrak{Re}(z) > 0$. What is T^{-1}?

We have seen that complex numbers can be viewed algebraically as numbers in a field and geometrically as points in a Euclidean plane. We can also view $z = x + iy$ as a two-dimensional vector in $\mathbb{R}^2$; here we are thinking of $\mathbb{C}$ as a real two-dimensional vector space. In fact $\mathbb{C}$ is a normed linear space where $|z|$ is the norm of the vector/number z. Moreover, this norm arises by way of the (Hermitian) inner product $< z, w >= z\bar{w}$. Notice that we recover the usual Euclidean 'dot product' of z and w by taking the real part of this; i.e., $z \cdot w = \mathfrak{Re}(z\bar{w}) = xu + yv$ (when $z = x + iy$ and $w = u + iv$). Thus, for example, z and w are orthogonal (as vectors in $\mathbb{R}^2$) precisely when $\mathfrak{Re}(z\bar{w}) = 0$.

From a topological viewpoint, $\mathbb{C}$ is a metric space where the distance function is given in the usual way by the underlying norm. Thus, $|z - w|$ is the distance between the complex numbers z and w; of course, this is just the ordinary Euclidean distance between the points z and w in the plane. In this regard we must mention the Triangle Inequalities:

$$||z| - |w|| \leq |z \pm w| \leq |z| + |w|.$$

In conclusion, we can view $\mathbb{C}$ in several different ways: a number field, a metric (or topological) space, a (two-dimensional) real vector space, or the Euclidean plane. Because of these many options, complex numbers have become an indispensable tool employed throughout pure and applied mathematics as well as the physical and engineering sciences.

1.B. **Polar Coordinates.** Here is an especially useful way to describe complex numbers.

We start by defining, for each real number θ, $\exp(i\theta) = e^{i\theta} := \cos\theta + i\sin\theta$; thus we obtain a map $\exp : \mathbb{R} \to \mathbb{T}$. See §2.C for motivation as to why we use this definition.

Recall the notion of polar coordinates (r, θ) used in analytic geometry. For non-zero complex numbers, $r := |z|$ gives the distance from the point z to the origin. The 'θ-coordinate' is the so-called polar angle. We say θ is a *value of the argument of* z, denoted $\theta = \arg(z)$, if $z = |z|e^{i\theta}$; in other words, if $\Re\mathfrak{e}(z) = |z|\cos\theta$ and $\Im\mathfrak{m}(z) = |z|\sin\theta$. Thus the polar coordinate notation $z = re^{i\theta}$ means both that $r = |z|$ and that θ is some value of $\arg(z)$; this notation is only used for $z \in \mathbb{C}_*$.

Clearly, if θ is one of the possible values of the argument of z, then the set of *all* possible values is given by $\{\theta + 2n\pi : n \in \mathbb{Z}\}$. (The reader should note—at least mentally—the set of all arguments of, say, ± 1, $\pm i$, and $1 + i$.) Given $z \in \mathbb{C}_*$, we single out the unique value θ of $\arg(z)$ that satisfies $-\pi < \theta \le \pi$ and call this the *principal value of the argument of* z, or briefly, the *principal value of* $\arg(z)$; we denote this by $\mathrm{Arg}(z)$. Thus we have a mapping $\mathrm{Arg} : \mathbb{C}_* \to (-\pi, \pi]$ with the property that

$$z = |z| \exp[\mathrm{Arg}(z)] \quad \text{for all } z \in \mathbb{C}_* .$$

Note that the set of all possible values of the argument of $z \in \mathbb{C}_*$ can now be described by

$$\arg(z) = \{\mathrm{Arg}(z) + 2n\pi : n \in \mathbb{Z}\} .$$

Finally, we point out the obvious, but oh so important, fact that Arg is not continuous (as a mapping from $\mathbb{C}_*$ to $\mathbb{R}$). (PP=Please Prove)

Recalling the angle addition identities for the sine and cosine functions we readily deduce the following. (PP)

1.2. **Proposition.** *For all* $\alpha, \beta \in \mathbb{R}$, $e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)}$.

We are now ready to take a geometric look at complex multiplication. Writing $z = re^{i\theta}$ and $a = |a|e^{i\alpha}$, say, we find that $az = r|a|e^{i(\theta+\alpha)}$. Thus a geometric description for the mapping $z \mapsto az$ is that the vector z is scaled by $|a|$ and then rotated by $\arg(a)$. We can also provide pictures for the complex arithmetic operations: addition, multiplication, inversion, square root. See my class notes!

As another application of Proposition 1.2 we note that $\arg(ab) = \arg(a) + \arg(b)$. However, this is only true as a set equality. (The reader should be sure to understand the precise meaning of this last statement!) For example, $\mathrm{Arg}(ab) = \mathrm{Arg}(a) + \mathrm{Arg}(b)$ is not always true.

Polar coordinates are especially useful for determining powers z^p and roots $z^{1/p}$ where $p \in \mathbb{Z}$ is an integer. First we remark that

$$\rho e^{i\varphi} = re^{i\theta} \iff \rho = r \quad \text{and} \quad \varphi = \theta + 2k\pi \quad \text{for some integer } k .$$

As an example, the four fourth roots of $4i$ (i.e., the four numbers z with the property that $z^4 = 4i$) are easily found to be $\pm\sqrt{2}e^{i\pi/8}, \pm\sqrt{2}e^{i5\pi/8}$. Similar reasoning provides verification of the following. (PP)

1.3. **Proposition.** *Let $n \in \mathbb{N}$. Then each $z \in \mathbb{C}_*$ has exactly n distinct n^{th} roots; that is, there are n distinct complex numbers w with $w^n = z$. In fact, these roots are given (in polar coordinates) precisely via*

$$w = \rho e^{i\varphi} \quad \textit{where } \rho > 0 \textit{ with } \rho^n = |z|, \textit{ and } \varphi = [\operatorname{Arg}(z) + 2k\pi]/n \textit{ for } k \in [0, n) \cap \mathbb{Z}.$$

We emphasize that the notation $z^{1/n}$ is ambiguous in that it stands for any one of the n distinct roots of z. Typically which root is desired is understood from context, but the reader is urged to use this notation with extreme caution! It is convenient to let

$$\sqrt[n]{z} := \rho \exp\left(\frac{i}{n}\operatorname{Arg}(z)\right), \quad \text{where } \rho > 0 \text{ with } \rho^n = |z|,$$

denote the *principal value of the n^{th} root of z*. Thus, for example, $\sqrt{z} := \sqrt[2]{z} = \rho e^{(i/2)\operatorname{Arg}(z)}$ (where $\rho^2 = |z|$) is the *principal value of the square root of z*; notice that $z^{1/2}$ can be either $+\sqrt{z}$ or $-\sqrt{z}$. One final word: all of the above is for $z \in \mathbb{C}_*$; all n of the n^{th} roots of $z = 0$ are, of course, just 0.

For later use we now define the *oriented angle from z to w* (for $z, w \in \mathbb{C}_*$) by

$$\Theta(z, w) := \operatorname{Arg}(w/z).$$

Notice that w/z is what we must multiply z by to get w, so $\Theta(z, w)$ is the (principal value of the) angle we must rotate z by to get the direction given by w. As simple examples, $\Theta(1, 1+i) = \pi/4$ and $\Theta(i, 1) = -\pi/2$.

Here are some easy to check properties of this concept. (Most of these follow by looking at $\operatorname{Arg}(\zeta)$ for $\zeta \in \mathbb{C}_*$. ◡̈) (PP)

1.4. **Proposition.** *Let z and w be non-zero complex numbers. Then:*

(a) *$\Theta(z, w)$ is a signed angle in $(-\pi, \pi]$ and $|\Theta(z, w)|$ is the measure of the smaller of the two angles formed by the vectors z and w at the origin.*
(b) *$\Theta(z, w) = 0$ if and only if $w = tz$ for some $t > 0$.*
$\Theta(z, w) = \pi$ if and only if $w = tz$ for some $t < 0$.
(c) *When $\Theta(z, w) \neq \pi$, $\Theta(w, z) = -\Theta(z, w) = \Theta(\bar{z}, \bar{w})$.*
When $\Theta(z, w) = \pi$, $\Theta(w, z) = \pi = \Theta(\bar{z}, \bar{w})$.
(d) *$\Theta(-z, w) = \Theta(z, w) - \pi = \Theta(z, -w)$.*
(e) *For all $c \in \mathbb{C}_*$, $\Theta(cz, cw) = \Theta(z, w)$.*
For all positive r and s, $\Theta(rz, sw) = \Theta(z, w)$.

We conclude this subsection by discussing the notion of a branch of the argument function. Let A be a subset of $\mathbb{C}_*$. A function $\vartheta : A \to \mathbb{R}$ is termed *a branch of the argument function* (or simply a *branch of* $\arg(z)$) *in* A provided ϑ is continuous and for all $z \in A$, $z = |z| \exp[\vartheta(z)]$ (i.e., $\vartheta(z)$ is some value of $\arg(z)$). The reader should convince herself that there does not exist a (continuous) branch of the argument in $\mathbb{C}_*$. In particular, $\operatorname{Arg}(z)$ is not a branch of $\arg(z)$ in $\mathbb{C}_*$. Your proof probably can be modified to demonstrate that if A is any set containing the unit circle $\mathbb{T}$, then there does not exists a (continuous) branch of the argument in A. We will see that this seemingly trivial fact is at the heart of much of complex analysis!

Once again: there does not exist a (continuous) branch of the argument in $\mathbb{C}_*$. To remedy this situation, we typically 'cut' the plane from the origin out to infinity and look for continuous branches of $\arg(z)$ in such regions. To provide some concrete examples of branches of

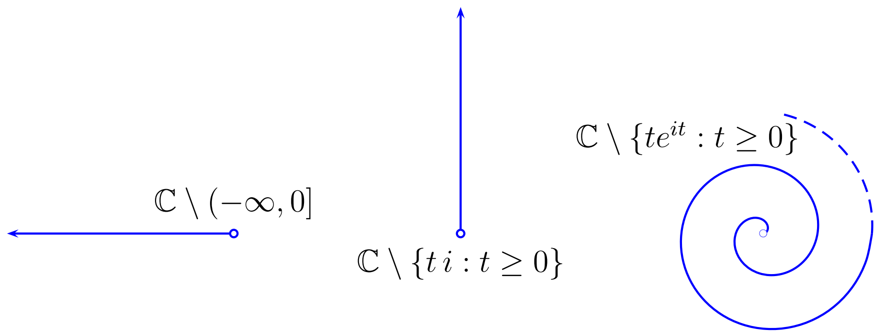

FIGURE 1. Where is there a branch of the argument?

$\arg(z)$ in various sets, let us start with $\mathrm{Arg}(z)$ which is continuous when restricted to the set $\mathbb{C}\setminus(-\infty,0]$ (or any subset thereof) and thus defines a branch of $\arg(z)$ there, the so-called *principal branch of the argument function*. As a second example, the function ϑ given by

$$\vartheta(z) := \begin{cases} \mathrm{Arg}(z) & \text{when } \Im\mathfrak{m}(z)\geq 0 \\ \mathrm{Arg}(z)+2\pi & \text{when } \Im\mathfrak{m}(z)<0 \end{cases}$$

defines a (continuous) branch of the argument in $\mathbb{C}\setminus[0,+\infty)$. (PP) At this time the reader should construct a branch of the argument in $\mathbb{C}\setminus\{te^{i\alpha} : 0\leq t<\infty\}$. We ask: Does there exist a (continuous) branch of $\arg(z)$ defined in $\mathbb{C}\setminus\{te^{it} : t\geq 0\}$? See Figure 1.

We emphasize that whether or not a given set supports a branch of the argument is purely a question of *continuity*!

1.C. **Analytic Geometry.** What are the complex equations for lines and conics?

Let us recall the 'standard' equations for lines and circles. Using our distance function it is easy to express disks and circles: we write

$$D(a;r)=\{z : |z-a|<r\},\ D[a;r]=\{z : |z-a|\leq r\},\ \text{and}\ C(a;r)=\{z : |z-a|=r\}$$

for the *open disk*, *closed disk*, and *circle* (respectively) with *center* a and *radius* r. See below for an alternative description of circles.

There are two standard ways to express an equation for a line. As in calculus, we can parameterize the line through the points a, b and obtain $z=z(t)=a+t(b-a)$ ($t\in\mathbb{R}$). Alternatively, if we know a normal vector, say ν, then the line through the point a with normal ν consists of points z such that the vector $z-a$ is orthogonal to ν. Recalling that $\Re\mathfrak{e}(z\bar{w})$ gives the Euclidean dot product of z and w, we deduce that this line is given by $\Re\mathfrak{e}[(z-a)\bar{\nu}]=0$. As an added bonus, we see that this technique also provides a description for the half-plane determined by this line and ν; namely, $\{z : \Re\mathfrak{e}[(z-a)\bar{\nu}]>0\}$. (Note that the line thru a, b has normal $\nu := i(a-b)$ and since $\Re\mathfrak{e}(-i\,\zeta)=\Im\mathfrak{m}(\zeta)$ we see that this line has equation $\Im\mathfrak{m}((z-a)\overline{(a-b)})$.)

We can also easily express the equations for conic sections. For example, an ellipse is the locus of points z with the property that the sum of the distances from z to two given points, say a and b, is a constant. Thus $\{z : |z-a|+|z-b|=2r\}$ is the ellipse with foci a, b and major axis r. The reader should write down similar expressions for parabolas and hyperbolas (perhaps after reviewing their definitions in terms of foci and directrices).

It is straightforward to verify that the equation

$$r|z|^2+cz+\bar{c}\bar{z}+s=0 \quad \text{where } r,s\in\mathbb{R},\ c\in\mathbb{C} \text{ and } |c|^2>rs\,,$$

represents a line if and only if $r = 0$ and otherwise (i.e., when $r \neq 0$) a circle.

A careful analysis of the complex equation $az + b\bar{z} + c = 0$, with $a, b, c \in \mathbb{C}$, is a wonderful exercise in using complex analysis. Of course this one equation corresponds to two real linear equations (obtained by taking the real and imaginary parts of the left hand side) involving two real variables x, y. From sophomore linear algebra we know that the solution possibilities are precisely: no solutions, a unique solution, or infinitely many solutions. It is a worthy exercise to determine this and explicitly describe the solutions using only complex analysis. [Answer: The trivial case, when $a = b = 0$ has either no solutions (when $c \neq 0$) or every z a solution (when $c = 0$). Assume one of a or b is non-zero. Then we get: a unique solution if and only if $|a| \neq |b|$, in which case $z = (b\bar{c} - \bar{a}c)/(|a|^2 - |b|^2)$; a line if and only if $|a| = |b| \neq 0$ and $b\bar{c} = \bar{a}c$, and when $c \neq 0$ the line is given by $\mathfrak{Re}[(z + c/2a)a\bar{c}] = 0$, which is the line through the point $-c/2a$ with normal vector $\bar{a}c$ (what about when $c = 0$?); otherwise, there are no solutions.]

For later use we set up the following notation: we let $\mathcal{L}$ and $\mathcal{C}$ denote the collections of all lines and all circles (respectively) in $\mathbb{C}$. Thus

$$\mathcal{L} := \{L \subset \mathbb{C} | L \text{ is a straight line}\} \qquad \text{and} \qquad \mathcal{C} := \{C \subset \mathbb{C} | C \text{ is a Euclidean circle}\}\,.$$

2. Elementary Mappings-part I

Here we begin our study of function theory. There are various types of functions including: real-valued functions of a real variable $u = u(x)$, complex-valued functions of a real variable $z = z(t)$, real-valued functions of a complex variable $u = u(z)$, and complex valued functions of a complex variable $w = f(z)$. We are especially interested in visualizing the 'action' of such function via appropriate pictures. In particular, given a set A in the z-plane, what is $f(A)$ in the w-plane?

As in undergraduate calculus, we can (and will) study polynomials, rational functions, trigonometric functions, algebraic functions, and transcendental functions (which include the exponential, logarithm and trigonometric functions). We start with the simplest possible functions.

2.A. **Complex Linear Maps.** What are these?

Well, what are real linear maps? The answer depends on the context. In a linear algebra class, a linear map (often described as a linear transformation) between two vector spaces is simply a function which preserves the (linear) structure. On the other hand, in a freshman calculus class, the so-called linear functions are those of the form $y = ax + b$ whose graphs are straight lines. (In linear algebra, such a function is termed an affine map.) Here we shall adopt the latter point of view, although as a word of caution, we remark that some texts might use the former.

Thus, *complex linear maps* are complex-valued functions of a complex variable of the form $w = az + b$ where $a, b \in \mathbb{C}$ are constants and $z \in \mathbb{C}$ is a complex variable; typically we assume that $a \neq 0$. Special cases include the maps

$$\begin{aligned} w &= z + b && \text{a } \textit{translation by } b \in \mathbb{C}\,, \\ w &= e^{i\theta} z && \text{a } \textit{rotation about the origin}\,, \\ w &= k\,z && \text{a } \textit{dilation wrt the origin} \text{ (here } k > 0)\,. \end{aligned}$$

A geometric description of the map $z \mapsto w = az + b$ is that: z gets rotated by $\arg(a)$, then dilated by $|a|$, and then translated by b. We emphasize this: *every* complex linear map L can be factored as

$$L = T \circ S \circ R \quad \text{where } R \text{ is a rotation, } S \text{ is a scaling (dilation), and } T \text{ is a translation.}$$

As a simple example, look at $w = 2iz + (1 - i)$.

We also have rotations and dilations about a fixed point $c \in \mathbb{C}$ given, respectively, by

$$w = e^{i\theta}(z - c) + c \quad \text{and} \quad w = k(z - c) + c\,.$$

We can consider these as maps $\mathbb{R}^2 \to \mathbb{R}^2$ and from linear algebra we know their matix representations. For example, the dilation $w = kz$ is given by

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and a rotation (about the origin) by the angle θ is

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Again, in general a complex linear map does not give a linear transformation of $\mathbb{R}^2$ to itself but rather an affine transformation.

Suppose we start with a linear transformation $\mathbb{R}^2 \xrightarrow{M} \mathbb{R}^2$. By identifying (x, y) with $x + iy$ we obtain a map $L : \mathbb{C} \to \mathbb{C}$. That is, if $M(x, y) = (u(x, y), v(x, y))$, then for $z = x + iy$ we set $L(z) = u(x, y) + iv(x, y)$. We ask: for which M will L be a complex linear map? (Below we will answer this question, but think about it now!)

It is not difficult to confirm the following fundamental properties of complex linear maps.

2.1. **Proposition.** *A complex linear map $w = L(z) = az + b$ with $a \neq 0$ is a complex polynomial of degree one and is a bijection of $\mathbb{C}$ with inverse $z = L^{-1}(w) = (w - b)/a$ another complex linear map. (In fact, L is a self-homeomorphism of the plane.)*

We also note that if L is a pure rotation or pure dilation or pure translation, then so is L^{-1}.

It turns out that complex linear maps possess many important properties which also hold for more general maps. However, the proofs for the linear maps are nice geometric exercises, so we explain the basic ideas here.

A distance preserving map (between metric spaces) is called an isometry. Thus $\mathbb{C} \xrightarrow{f} \mathbb{C}$ is an *isometry* provided

$$\forall\, z, z' \in \mathbb{C} \;:\quad |f(z) - f(z')| = |z - z'|\,.$$

A map that changes all distances by a fixed multiplicative factor is usually called a similarity. In our setting, a map $f : \mathbb{C} \to \mathbb{C}$ is a *similarity* provided there exists a real constant $\sigma > 0$ such that

$$\forall\, z, z' \in \mathbb{C} \;:\quad |f(z) - f(z')| = \sigma\,|z - z'|\,.$$

It is not too difficult to verify the following.

2.2. **Proposition.** *For $\mathbb{C} \xrightarrow{f} \mathbb{C}$, these are equivalent:*

(a) *f is a similarity.*
(b) *For every triangle $\Delta \subset \mathbb{C}$, $\Delta' = f(\Delta)$ is a triangle that is similar to Δ.*

(c) *For all distinct points* $a, b, c \in \mathbb{C}$, $\left|\dfrac{a-b}{a-c}\right| = \left|\dfrac{f(a)-f(b)}{f(a)-f(c)}\right|$.

(Hints: Show that (a) $\Longrightarrow$ (b) $\Longrightarrow$ (c) $\Longrightarrow$ (a). See me :-)

It is straightforward to demonstrate that complex linear maps are similarities.

2.3. **Proposition.** *Every non-constant complex linear map* $L(z) = az + b$ *is a similarity transformation; when* $|a| = 1$ *it is an isometry.*

Proof. $|L(z) - L(z')| = |(az + b) - (az' + b)| = |a(z - z')| = |a||z - z'|$. □

In fact, complex linear maps are precisely the *orientation preserving similarities* of the Euclidean plane.

Now we want to prove that complex linear maps transform lines and circles to lines and circles respectively. This is geometrically transparent for translations and rotations (isn't it?), but what about dilations? Here is an algebraic argument that should convince you. Let K be a line or a circle. Then K is given by some equation of the form

$$r|z|^2 + cz + \bar{c}\bar{z} + s = 0 \quad \text{where } r, s \in \mathbb{R},\ c \in \mathbb{C} \text{ and } |c|^2 > rs\,,$$

and K is a line if and only if $r = 0$ and otherwise (i.e., when $r \neq 0$) a circle. Now the mapping $w = kz$ transforms K into a set K' which is given by the equation

$$r'|w|^2 + c'w + \bar{c}'\bar{w} + s' = 0\,, \quad \text{with } r' = r/k^2, s' = s \in \mathbb{R} \text{ and } c' = c/k \in \mathbb{C}.$$

Since $|c'|^2 = |c|^2/k^2 > rs/k^2 = r's'$ and $r' = 0$ if and only if $r = 0$, we see that K' is a line or a circle just as K is.

The reader should convince herself that a similar argument works for rotations and translations.

Here is a geometric proof that works in general.

2.4. **Proposition.** *Let* L *be a non-constant complex linear map. If* K *is a line (or circle, respectively), then so is* $L(K)$ *(respectively). Moreover,* L *maps* K *bijectively onto* $L(K)$.

Proof. Assume that $w = L(z) = az + b$. First, suppose $K := C(z_0; r)$ is the circle centered at z_0 and of radius r. Let $w_0 = L(z_0)$. Suppose $w = L(z)$ for some $z \in K$. Recalling the proof of Proposition 2.3 we obtain

$$|w - w_0| = |L(z) - L(z_0)| = |a||z - z_0| = |a|r\,.$$

This means that $L(K) \subset K'$ where $K' = C(w_0; r')$ is the circle centered at w_0 of radius $r' = |a|r$. Exactly the same argument, applied to the inverse map L^{-1}, gives us $L^{-1}(K') \subset K$. Thus $L(K) = K'$.

Next, suppose K is a straight line. Select two points z_1 and z_2 so that K is the perpendicular bisector of the segment $[z_1, z_2]$; i.e., $K = \{z : |z - z_1| = |z - z_2|\}$. Let $w_i = L(z_i)$ for $i = 1, 2$ and define K' to be the perpendicular bisector of the segment $[w_1, w_2]$; so, $K' = \{w : |w - w_1| = |w - w_2|\}$. If $w = L(z)$ for some $z \in K$, then

$$|w - w_1| = |L(z) - L(z_1)| = |a||z - z_1| = |a||z - z_2| = |L(z) - L(z_2)| = |w - w_2|\,.$$

Thus $L(K) \subset K'$. Similarly, $L^{-1}(K') \subset K$, so $L(K) = K'$. □

Here are three noteworthy consequences of the above.

2.5. **Corollary.** *Let* L *be a complex linear map and let* C *be a circle in* $\mathbb{C}$.

(a) *For any points* a, b, $L([a, b]) = [L(a), L(b)]$.

(b) *If $[a,b]$ is a diameter of a circle C, then its image is a diameter of $L(C)$.*
(c) *If T is a line tangent to a circle C, then $L(T)$ is a line tangent to $L(C)$.*

In particular we notice from the above that if T is a line tangent to a circle C, and K is the line through the center of C and the point of intersection of C and T, then T and K are orthogonal and *so are their images* under any complex linear map. What about other angles?

We next verify the angle preservation property: If two lines (or two circles or a line and a circle) intersect, then so do their images and moreover the angles of intersection (for the original sets and their images) are exactly the same. This property just mentioned, angle preservation, is especially important. In fact, complex linear maps preserve *oriented* angles; mappings which possess this property are called *conformal*. The map $z \mapsto \bar{z}$ preserves angles (so it is an *isogonal* map), but it does not preserve orientation and so is not conformal (but rather is *anti-conformal*).

The reader no doubt understands (from calculus or linear algebra) what is meant by the angle between two lines (or two circles or a line and a circle). As a warm-up for later purposes, when we study this in general for two intersecting paths, we now formalize this concept. Recall that for $z, w \in \mathbb{C}_*$, $\Theta(z,w) = \operatorname{Arg}(w/z)$ is the oriented angle from z to w. See Proposition 1.4.

Suppose that A and B are two lines which intersect at some point z_0. Pick direction vectors a, b for A, B (respectively); thus A, B can be described by the parametric equations $\alpha(t) := z_0 + ta$, $\beta(t) := z_0 + tb$ (respectively) where $t \in \mathbb{R}$. We note that here we are actually considering A, B as *oriented* lines: $-a$ is also a direction vector for the line A and it gives the opposite orientation. The *oriented angle from A to B* is

$$\Theta(A,B) := \Theta(a,b) = \operatorname{Arg}(b/a)\,.$$

Note that changing the orientation of A (from that given by a to that given by $-a$) changes this angle by π; see Proposition 1.4(e). Also, $A = B$ as oriented lines if and only if the angle between them is zero; when this angle is π, then the lines are the same but they have opposite orientations.

It is now a simple matter to define the angle between two intersecting oriented circles or between an oriented circle and oriented line that meets it. Let K and K' each be an oriented line or circle (so, two lines or two circles or one of each) that intersect at some point. Let T, T' be the oriented lines that are tangent to K, K' (respectively) at the point of intersection (with the orientations induced by the orientations of K, K'). Then the *oriented angle from K to K'* is

$$\Theta(K,K') := \Theta(T,T')\,.$$

Note that two circles (for example) are tangent at a point exactly when the angle between them is either 0 or π: which value we get depends both on whether the circles have opposite or identical orientation as well as whether they are internally tangent or externally tangent.

The reader is encouraged to not get too tangled up with the above formal definitions!

It is now rather straightforward to demonstrate that complex linear maps are conformal. Notice that if K is an oriented line or circle, then so is its image under any complex linear map. This essentially follows from Proposition 2.4.

2.6. **Proposition.** *Let L be a complex linear map. Suppose A and B are two oriented lines (or two oriented circles or an oriented line and an oriented circle) that intersect. Then so do their images $A' := L(A)$ and $B' := L(B)$, and moreover $\Theta(A',B') = \Theta(A,B)$.*

An outline for a formal proof. First, the reader should convince herself that we need only consider the case when both A, B are oriented lines. Next, the assertion is immediate for translations (right?). Now look at the special case when A, B intersect at the origin and $L(0) = 0$. In this case $L(z) = cz$, so the lines A, B are both rotated through the angle $\text{Arg}(c)$ and hence the oriented angle of intersection is unchanged.

Finally, consider the general case. Let $w_0 := L(z_0)$ where z_0 is the point common to A, B. The translations $T(z) := z - z_0$ and $S(w) := w - w_0$ preserve the oriented angles $\Theta(A, B)$ at z_0 and $\Theta(A', B')$ at w_0 (respectively). Put $F := S \circ L \circ T^{-1}$. Then F is a complex linear map with $F(0) = 0$. Since $T(A)$, $T(B)$ intersect at the origin, F preserves their oriented angle of intersection (by the first part of the argument). It now follows that

$$\begin{aligned}\Theta(A, B) &= \Theta\left(T(A), T(B)\right) = \Theta\left(F[T(A)], F[T(B)]\right)\\ &= \Theta\left(S^{-1}(F[T(A)]), S^{-1}(F[T(B)])\right) = \Theta(A', B')\,.\end{aligned}$$

The penultimate equality above holds because S^{-1} is a translation (by w_0, right?). □

Later we will examine functions of the form $w = (az + b)/(cz + d)$; these are the simplest possible rational functions and are called Möbius transformations. We will see that these maps possess many of the same properties enjoyed by the linear functions.

2.B. **Non-injective Maps.** Here we seek a geometric description for the polynomial maps

$$z \mapsto w = p(z) = a_0 + a_1 z + \cdots + a_n z^n\,.$$

The most elementary yet nonetheless basic representative of these is given by the following.

2.7. *Example.* The "squaring" map $w = z^2$.

See my class notes—what I actually did in the class room—for lotsa pictures including the images and pre-images of straight lines. Note that angles at the origin are doubled whereas away from the origin this map is isogonal. Also: lines,circles are not mapped to lines,circles; much more complicated than linear maps; inverses are $\pm\sqrt{z}$.

Another especially important non-injective map is the following.

2.8. *Example.* The "squash" map $w = z + 1/z$. Here $z \in \mathbb{C}_*$.

Again, see my class notes—what I actually did in the class room—for lotsa pictures.

2.C. **The Complex Exponential Function.** What should e^z be for $z \in \mathbb{C}$?

We assume knowledge about e^x for $x \in \mathbb{R}$. For $z = x + iy$ we define

$$\exp(z) = e^z := e^x e^{iy} = e^x \cos y + i\, e^x \sin y\,.$$

Recall from the beginning of §1.B that for a real number θ we have defined $e^{i\theta} := \cos\theta + i\sin\theta$. To understand why we use this definition of $e^{i\theta}$, we first note that for $n \in \mathbb{N}$, i^n is either ± 1 (when n is even) or $\pm i$ (when n is odd). In particular, if $n = 2k$, then $i^n = (-1)^k$ and if $n = 2k + 1$, then $i^n = (-1)^k i$. Now we use property (1) below and (formally) manipulate power series as follows:

$$e^{i\theta} = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} = \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!} = \cos(\theta) + i\sin(\theta)\,.$$

2.9. **Proposition** (Properties of e^z).

(a) $e^z = \sum_0^\infty z^n/n!$ *absolutely convergent for all z and uniformly on compacta.*

(b) $|e^z| = e^{\mathfrak{Re}(z)}$ *and* $\arg(e^z) = \mathfrak{Im}(z) + 2k\pi$ *for* $k \in \mathbb{Z}$.
(c) $e^z e^w = e^{z+w}$, $e^{-z} = 1/e^z$, *and for all* $n \in \mathbb{N} : (e^z)^n = e^{nz}$.
(d) $e^w = e^z$ *if and only if* $w = z + 2k\pi i$ *for some* $k \in \mathbb{Z}$.

2.10. *Remarks.*

(1) $|e^z| = e^{\mathfrak{Re}(z)} > 0$ implies in particular that $e^z \neq 0$ for all $z \in \mathbb{C}$. We shall see that every non-zero complex number w can be written as e^z; in fact this can be done for infinitely many z. Thus every point $w \in \mathbb{C}$, except $w = 0$, is covered infinitely often by the mapping $w = e^z$.
(2) $\lim_{z\to\infty} e^z$ does not exist. In fact, by choosing an appropriate sequence of points (z_n), with $|z_n| \to \infty$, we can get any value we want for $\lim_{n\to\infty} e^{z_n}$. Another, more sophisticated, way of expressing this is that *in every neighborhood of infinity,* $w = e^z$ *comes arbitrarily close to every complex number infinitely often.*

The **Mapping Properties** of $w = e^z$ are easy to determine. We first look at the images of horizontal and vertical lines. We find that the image of a rectangle $\{x + iy : a \le x \le b, c \le y \le d\}$ with $d - c \le 2\pi$ is the 'circular rectangle' $\{re^{i\theta} : e^a \le r \le e^b, c \le \theta \le d\}$. What is the image of a line $\{x + imx : x \in \mathbb{R}\}$?

2.D. **Trignometric and Hyperbolic Functions.** These are defined as follows:

$$\cos(z) := (e^{iz} + e^{-iz})/2\,, \quad \sin(z) := (e^{iz} - e^{-iz})/2i$$

and

$$\cosh(z) := (e^z + e^{-z})/2\,, \quad \sinh(z) := (e^z - e^{-z})/2\,.$$

The reader is encouraged to express the real and imaginary parts of these functions and to explore the many relations between them. See for example any complex analysis book.

2.E. **The Complex Logarithm Function.** What should $\log(z)$ be for $z \in \mathbb{C}$?

Recall that for real x, the functions $x \mapsto e^x$ and (for $x > 0$) $x \mapsto \ln x$ are inverses of each other; we assume knowledge concerning the natural logarithm function $\ln x$ for $x \in \mathbb{R}$. We want to define an inverse for the complex exponential function, but of course this does not make sense because $z \mapsto e^z$ is not one-to-one (in fact, it is ∞-to-one:-). In spite of this, let us try to do it anyway and see where this leads. We begin with a fixed $z \in \mathbb{C}$ and try to solve the equation $e^w = z$ for w. Obviously this cannot have a solution unless $z \neq 0$, so we assume this. We express z in polar coordinates as $z = re^{i\theta}$ where θ is some value of $\arg(z)$. Now write $w = u + iv$. Then we have

$$re^{i\theta} = z = e^w = e^u e^{iv}$$

which leads to $r = e^u$ and $\theta = v + 2k\pi$ for some $k \in \mathbb{Z}$. From this we conclude that

$$u = \ln(r) = \ln|z| \quad \text{and} \quad v = \theta + 2k\pi \quad \text{for some } k \in \mathbb{Z}.$$

Notice that v above is equal to some value of $\arg(z)$. We record our observations as follows.

2.11. **Lemma.** *For* $z \in \mathbb{C}_*$, $e^w = z \iff w = \ln|z| + i\arg(z)$.

The reader should be certain (s)he understands the precise meaning of the above! It is not hard to see that when w has such a form, then $e^w = z$; the necessity of this is what we verified above. Notice that this Lemma says that for any w with $e^w = z$, $\mathfrak{Re}(w) = \ln|z|$; note that $z \mapsto \ln|z|$ is a well-defined (single-valued) real-valued function of a complex variable.

On the other hand, all we know about $\mathfrak{Im}(w)$ (for such a w) is that it must be some value of $\arg(z)$ (i.e., when $e^w = z$).

Given $z \in \mathbb{C}_*$, we say that $w \in \mathbb{C}$ is a *value of the logarithm of* z (or more simply, a *value of* $\log(z)$) if $e^w = z$. Lemma 2.11 tells us precisely what these values are. As an example, the *principal value of* $\log(z)$ is

$$\mathrm{Log}(z) := \ln|z| + i\,\mathrm{Arg}(z)\,.$$

We cannot use the 'formula' for w provided by Lemma 2.11 to define a function $\log(z)$ because it does not provide a single-value but rather many possible values. To get a true (i.e. single-valued) function we must first choose a branch of the argument. Recall that any continuous real-valued function ϑ satisfying $z = |z|e^{i\vartheta(z)}$ is termed a branch of the argument of z. Here, of course, the domain of ϑ must be a subset of $\mathbb{C}_*$, and this just means that $\vartheta(z)$ is one of the (many) values for $\arg(z)$ (and that ϑ is continuous).

Next, by a *branch of the logarithm function* we mean a continuous (single-valued) complex-valued function λ with the property that for all $z \in \mathbb{C}_*$, $e^{\lambda(z)} = z$. Thus $\lambda(z)$ is a value of $\log(z)$ for each z and so λ serves as an inverse for the exponential function; hence the domain of λ must be a subset of $\mathbb{C}_*$. Again, we also want λ to be continuous, and this is impossible to do in all of $\mathbb{C}_*$, so usually we restrict our attention to the regions described at the end of §1.B (where we took a branch cut from the origin out to infinity).

To make a formal definition, a function $\mathbb{C}_* \supset A \xrightarrow{\lambda} \mathbb{C}$ is a *branch of the logarithm function in* A (or more briefly, a *branch of* $\log(z)$ *in* A) provided λ is continuous and for all $z \in A$, $\exp(\lambda(z)) = z$. Note: Any such λ serves as an inverse for the exponential function, in A.

Let us emphasize the point that, whenever we speak of a branch of the logarithm, there is always an associated domain-set. In fact, whether or not there exists a (continuous) branch of the logarithm in a given set A depends in a very strong way on the topology of A. Moreover, this problem (of determining when there exists a branch of $\log(z)$) is at the very heart of complex analysis.

As an example, the *principal branch of the logarithm* is defined for $z \in \mathbb{C} \setminus (-\infty, 0]$ via

$$\mathrm{Log}(z) := \ln|z| + i\,\mathrm{Arg}(z)\,.$$

The point here is that while $\mathrm{Log}(z)$ is defined for all $z \in \mathbb{C}_*$, it is only continuous in $\mathbb{C} \setminus (-\infty, 0]$.

As one might expect, there is a close tie between branches of the argument and branches of the logarithm. In fact we have the following result, whose proof mimics the argument given for Lemma 2.11.

2.12. **Proposition.** *A function λ is a branch of the logarithm if and only if there exists a branch of the argument, say ϑ, such that for all z, $\lambda(z) = \ln|z| + i\vartheta(z)$.*

We also have the following information about continuous branches of the logarithm.

2.13. **Theorem.** *Let λ_1 and λ_2 be two branches of the logarithm defined (and continuous) in the same open connected set Ω. Then there is a fixed $k \in \mathbb{Z}$ such that for all $z \in \Omega$, $\lambda_1(z) = \lambda_2(z) + 2k\pi i$.*

Proof. □

Note that for any branch λ of the logarithm, defined in some set A, we always have $e^{\lambda(z)} = z$ for all $z \in A$. However, we may not have $\lambda(e^z) = z$. Indeed, even for $z = x > 0$ we need not have $\log(x) = \ln(x)$. For example...

2.F. **Complex Powers and Roots.** Loosely speaking, for $a \in \mathbb{C}^*$ and $b \in \mathbb{C}$ we define

$$a^b := e^{b \log a}.$$

More precisely, given $a \in \mathbb{C}^*$ and $b \in \mathbb{C}$, we say that $c \in \mathbb{C}$ is a *value of the* b^{th}*-power of* a provided there is some value $\lambda(a)$ of $\log(a)$ such that $c = e^{b\lambda(a)}$. For example, $\exp(b \operatorname{Log}(a))$ is always such a value.

Some special cases of this include:

(1) If $b \in \mathbb{Z}$, then a^b is a single value.
(2) If $b = p/q \in \mathbb{Q}$ is rational with $p \in \mathbb{Z}$, $q \in \mathbb{N}$ and p, q having no common factors, then there are q distinct complex values of a^b and these are symmetrically located around the circle $|w| = |a|^b$.
(3) If $b \in \mathbb{R} \setminus \mathbb{Q}$ is irrational, then a^b consists of infinitely many complex values located around the circle $|w| = |a|^b$.

Everywhere above $|a|^b$ stands for the principle value of the b^{th} power of $|a|$. Thus, when $b \in \mathbb{R}$, we find that every value of a^b lies on the circle $|z| = \exp(b \operatorname{Log}(a))$.

Note however that the 'law of exponents' $a^b a^c = a^{b+c}$ fails to hold, even when this is considered as a set equality. The reader should find examples of this phenomenon.

3. Elementary Plane Topology

We consider $\mathbb{C}$ with its usual Euclidean topology given via Euclidean distance.

3.A. **Open and Closed Sets.** What are these? Recall (see §1.C) our notation

$$D(a;r) := \{z : |z-a| < r\}, \ D[a;r] := \{z : |z-a| \le r\}, \text{ and } \ C(a;r) := \{z : |z-a| = r\}$$

(respectively) for the *open disk*, *closed disk*, and *circle* (respectively) with *center* a and *radius* r. Also, $\mathbb{D} := D(0;1) = \{z : |z| < 1\}$ and $\mathbb{T} := C(0;1) = \{z : |z| = 1\}$ are the unit disk and unit circle.

By definition, a set $A \subset \mathbb{C}$ is *open* precisely when each of its points has an open disk about it which lies in A; that is, A is an *open set* provided for each $a \in A$ there exists an $r > 0$ such that $D(a;r) \subset A$. We call A a *closed set* if $A^c := \mathbb{C} \setminus A$ is open.

Examples: an open (closed) disk is an open (closed, resp.) set.

3.1. **Proposition.**

(a) *$\mathbb{C}$ and $\emptyset$ are both open and closed.*
(b) *A union of open sets is open. (An intersection of closed sets is closed.)*
(c) *A finite intersection of open sets is open. (A finite union of closed sets is closed.)*

More notation: *interior*, *closure*, *boundary*, *accumulation points*.

Note that these topological notions are all defined in terms of the Euclidean distance $|z - w|$. A different distance function could (probably would) have different open disks associated with it, but the open sets might be the same or might be different. For example, the 'square' distance

$$|z - w| := \max\{|x-u|, |y-v|\} \qquad (z = x+iy, w = u+iv)$$

has associated open disks which are the insides of squares with horizontal and vertical edges. However, since there is always such a square inside every round disk and conversely a round disk inside every such square, we see that the 'square open' sets are exactly the same as the Euclidean open sets.

3.B. **Compact and Connected Sets.** What are these?

Please consult your favorite text (e.g. Ahlfors has this information in Chapter 3, Section 1). In class we proved the following useful fact.

3.2. **Theorem.** *Every domain (i.e., open connected set) in $\mathbb{C}$ is path-connected. In fact, we can even join two given points by a broken-line-segment path (i.e., a path consisting of finitely many horizontal and vertical segments) that lies in the domain.*

3.C. **Limits and Continuity.** The 'usual' stuff! We assume the reader is familiar with the notion of limits of sequences.

Given $\mathbb{C} \supset A \xrightarrow{f} \mathbb{C}$ and $a \in \bar{A}$, we write

$$b = \lim_{z \to a} f(z)$$

to mean that for all $\varepsilon > 0$ there exists a $\delta > 0$ with the property that

$$\forall z \in A : \qquad 0 < |z - a| < \delta \implies |f(z) - b| < \varepsilon\,.$$

We say that $f(z)$ is *continuous* .... A *homeomorphism* is ...

Example: every non-constant complex linear map is a self-homeomorphism of $\mathbb{C}$.

3.D. **The Riemann Sphere.** Here we introduce and study the extended complex plane.

The *extended complex plane*, also called the *Riemann sphere*, is the set

$$\hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$$

where ∞ is an 'ideal point at infinity'. We obtain a geometric model for $\hat{\mathbb{C}}$ as follows. We identify the complex plane $\mathbb{C}$ with the horizontal x_1x_2-plane in $x_1x_2x_3$-space (i.e., the 2-plane $\mathbb{R}^2 \times \{0\}$ in $\mathbb{R}^3$). Thus we are identifying $z = x + iy$ with the point $(x, y, 0)$ in $\mathbb{R}^3$.

Next, let $\mathbb{S} := \{(x_1, x_2, x_3) : x_1^2 + x_2^2 + x_3^2 = 1\}$ denote the unit sphere in $\mathbb{R}^3$ and consider the 'north and south poles' $N(0, 0, 1)$ and $S(0, 0, -1)$ respectively. Suppose that $P(x_1, x_2, x_3)$ is a point on $\mathbb{S}$ other than N. Then the straight line through P and N meets $\mathbb{C}$ in a unique point $z := \Pi(P)$. Conversely, given a complex number $z = x + iy \in \mathbb{C}$ (thought of as the point $(x, y, 0) \in \mathbb{R}^3$), the line through z and N will meet $\mathbb{S}$ at some point P other than N, and $\Pi(P) = z$.

In this way $\mathbb{S}\setminus\{N\} \xrightarrow{\Pi} \mathbb{C}$ defines a one-to-one correspondence between the points of $\mathbb{S}\setminus\{N\}$ and $\mathbb{C}$. Notice that as P moves closer to N on $\mathbb{S}$, the point z moves further and further away from the origin in $\mathbb{C}$, and conversely. This prompts us to let N correspond to the ideal point ∞, and we obtain a one-to-one correspondence between all points of $\mathbb{S}$ and the extended complex plane $\hat{\mathbb{C}}$. The bijection $\Pi : \mathbb{S} \to \hat{\mathbb{C}}$ is called *stereographic projection*. It is because of this model that $\hat{\mathbb{C}}$ is also called the *Riemann sphere*. For notational convenience we also write $\hat{\mathbb{R}} := \mathbb{R} \cup \{\infty\}$ and $i\hat{\mathbb{R}} := i\mathbb{R} \cup \{\infty\}$; these are called the *extended real line* and the *extended imaginary axis*, respectively.

Let us mention a few simple properties of stereographic projection. Notice that the south pole S corresponds to the origin in $\mathbb{C}$. Also, the unit disk $\mathbb{D}$ corresponds to the lower (southern) hemisphere while the exterior of the unit disk corresponds to the upper (northern) hemisphere.

Next, we consider circles on $\mathbb{S}$; note that these are formed by intersecting $\mathbb{S}$ with some two-dimensional plane in $\mathbb{R}^3$. We see that:

(1) Latitudinal circles on $\mathbb{S}$ correspond to circles in $\mathbb{C}$ centered at the origin.

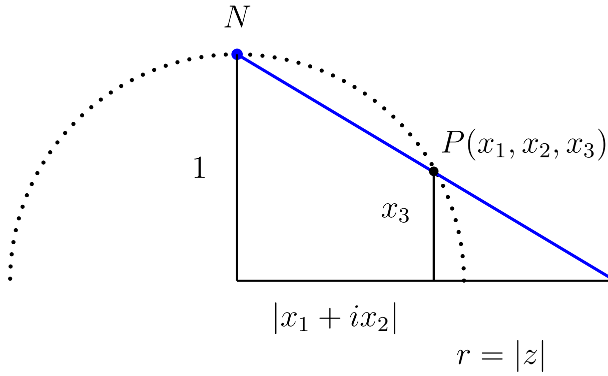

FIGURE 2. Similar triangles

(2) Longitudinal circles on $\mathbb{S}$ correspond to extended lines in $\hat{\mathbb{C}}$ through the origin.
(3) Any circle on $\mathbb{S}$ through N corresponds to an extended line in $\hat{\mathbb{C}}$.

In fact the above are bijective correspondences; the reader should note the precise meaning of this statement.

In order to make our statements about circles through the north pole N precise, we introduced some terminology. Given a line L in $\mathbb{C}$, we set $\hat{L} := L \cup \{\infty\} \subset \hat{\mathbb{C}}$ and call $\hat{L}$ the *extended line* corresponding to L. Even when we forget the adjective 'extended', *every line in* $\hat{\mathbb{C}}$ *contains the point* ∞. As examples we mention that $\hat{\mathbb{R}}$ and $i\hat{\mathbb{R}}$ are extended lines in $\hat{\mathbb{C}}$.

Below we will determine the stereographic image of an arbitrary circle on $\mathbb{S}$. Before doing this, we make a few calculations.

It is useful to have some concrete formulas to work with. Let $P(x_1, x_2, x_3)$ be a point of $\mathbb{S}$ other than N. A quick glance at a diagram (see Figure 2) reveals that P corresponds to the point

$$z = \Pi(P) = r\,\frac{x_1 + ix_2}{|x_1 + ix_2|} \qquad \text{where } r = |z|\,;$$

that is, the vectors z and $x_1 + ix_2$ have the same direction. From the pictured similar triangles we observe that

$$\frac{x_3}{1} = \frac{r - |x_1 + ix_2|}{r} \quad \text{so} \quad r\,x_3 = r - |x_1 + ix_2|$$

and thus

$$z = \Pi(x_1, x_2, x_3) := \frac{x_1 + ix_2}{1 - x_3}\,.$$

To calculate the inverse mapping we parameterize the line through $z = x + iy$ and N by $x_1 = xt$, $x_2 = yt$, $x_3 = 1 - t$ where $t \in \mathbb{R}$, and then determine the parameter value t for the point P where this line meets $\mathbb{S}$. We must have

$$(xt)^2 + (yt)^2 + (1-t)^2 = 1\,, \quad \text{so} \quad t = \frac{2}{|z|^2 + 1}$$

and we find that $(x_1, x_2, x_3) = \Pi^{-1}(z)$ is given by

$$x_1 = \frac{2x}{|z|^2+1} = \frac{z + \bar{z}}{|z|^2+1}\,, \quad x_2 = \frac{2y}{|z|^2+1} = \frac{z - \bar{z}}{i(|z|^2+1)}\,, \quad x_3 = \frac{|z|^2-1}{|z|^2+1}\,.$$

We can use our geometric model to define a distance function on the extended plane. The idea is quite simple: to find the distance between two points in $\hat{\mathbb{C}}$ we just use the Euclidean straight-line distance between the corresponding points of $\mathbb{S} \subset \mathbb{R}^3$; this is not too difficult to calculate using our formulas above. To be explicit, the *chordal distance* on $\hat{\mathbb{C}}$ is defined by

$$\chi(z,w) := \begin{cases} \dfrac{2|z-w|}{\sqrt{(1+|z|^2)(1+|w|^2)}} & \text{if } z \neq \infty \neq w \text{ (i.e., } z,w \in \mathbb{C}\text{)}\,, \\ \dfrac{2}{\sqrt{1+|z|^2}} & \text{if } z \neq \infty = w\,. \end{cases}$$

Note that the chordal distance $\chi(z,w)$ is precisely the length of the chord joining the points $\Pi^{-1}(z), \Pi^{-1}(w) \in \mathbb{S}$. Also, note that as $w \to \infty$, $\chi(z,w) \to \chi(z,\infty)$ Why is it geometrically evident that for all points $z, w \in \hat{\mathbb{C}}$, $\chi(z,w) \leq 2$? When is $\chi(z,w) = 2$? Finally, it is perhaps useful to observe that:

$$\forall\, z,w \in \mathbb{T}\,, \quad \chi(z,w) = |z-w|\,.$$
$$\forall\, z,w \in \mathbb{D}\,, \quad \chi(z,w) \geq |z-w|\,.$$
$$\forall\, z,w \in \hat{\mathbb{C}} \setminus \bar{\mathbb{D}}\,, \quad \chi(z,w) \leq |z-w|\,.$$

Now we return to the problem of determining the stereographic image of an arbitrary circle on $\mathbb{S}$. The definition of chordal distance is such that $\Pi : (\mathbb{S}, |\cdot|) \to (\hat{\mathbb{C}}, \chi)$ is an isometry, where $|\cdot|$ stands for Euclidean distance in $\mathbb{R}^3$. In particular, given a point A on $\mathbb{S}$ and $0 < t < 2$, the set $K = \{P \in \mathbb{S} : |P - A| = t\}$ (a spherical disk on $\mathbb{S}$) is mapped to $K' = \Pi(K) = \{z \in \hat{\mathbb{C}} : \chi(z,a) = t\}$ where $a = \Pi(A)$.

It is convenient to introduce some notation in order to state precisely what happens to circles on $\mathbb{S}$ under stereographic projection. Recall that $\mathcal{C}$ and $\mathcal{L}$ denote the collections of all circles and lines (respectively) in $\mathbb{C}$; see the very end of §1.C. For each $L \in \mathcal{L}$, let $\hat{L} := L \cup \{\infty\}$, put $\hat{\mathcal{L}} := \{\hat{L} : L \in \mathcal{L}\}$, and define $\hat{\mathcal{C}} := \mathcal{C} \cup \hat{\mathcal{L}}$. We also let $\mathcal{K}$ denote the collection of all circles on $\mathbb{S}$.

3.3. **Theorem.** *The stereographic image of each circle on $\mathbb{S}$ is a circle or an extended line in $\hat{\mathbb{C}}$; the image is an extended line precisely when the initial circle passes through the 'north pole' N. In particular, stereographic projection induces a one-to-one correspondence between circles $K \in \mathcal{K}$ and their images $K' = \Pi(K) \in \hat{\mathcal{C}}$, and $K' \in \hat{\mathcal{L}}$ if and only if $N \in K$*

Because of this result, each $C \in \hat{\mathcal{C}}$ is called a circle in $\hat{\mathbb{C}}$; i.e., $\hat{\mathcal{C}}$ is the collection of all circles in $\hat{\mathbb{C}}$. Of course each of these is either a circle in $\mathbb{C}$ or an extended line in $\hat{\mathbb{C}}$.

Ahlfors gives one proof of this result. As shown in class, another proof can be obtained by noticing that any $K \in \mathcal{K}$ can be written as $K = \{P \in \mathbb{S} : |P - A| = t\}$ for some point $A \in \mathbb{S}$ and some $t \in (0,2)$. Letting $a = \Pi(A)$, and using the fact that Π is an isometry, we find that $K' = \Pi(K) = \{z \in \hat{\mathbb{C}} : \chi(z,a) = t\}$. It is then straightforward to show that the equation $\chi(z,a) = t$ has the form $r|z|^2 + cz + \bar{c}\bar{z} + s = 0$ with $r, s \in \mathbb{R}$, $c \in \mathbb{C}$ and $|c|^2 > rs$ (and that $r = 4 - t^2(1+|a|^2) = 0$ if and only if $N \in K$).

Examining this idea more closely, we see that the spherical disk $D = \{P \in \mathbb{S} : \|P - A\| < t\}$ is mapped by Π to the χ-open disk $D' = \Pi(D) = \{z \in \hat{\mathbb{C}} : \chi(z,a) < t\}$, where $a = \Pi(A)$. Upon careful inspection we see that D' is either an open Euclidean disk, an open half-plane, or the exterior of a closed Euclidean disk (together with the point at infinity). More precisely, suppose, as above, that $K = \{P \in \mathbb{S} : \|P - A\| = t\}$ for some $A \in \mathbb{S}$ and $t \in (0,2)$; so

$K' = \Pi(K) = \{z \in \hat{\mathbb{C}} : \chi(z,a) = t\}$ where $a = \Pi(A)$. Then either K' is a Euclidean circle, say $C(c;r)$ (when $N \notin K$) or $K' = \hat{L}$ for some line L (when $N \in K$). (You should be able to calculate c and r, or determine L!) We thus find that

$$D' = \begin{cases} D(c;r) & \text{when } N \notin K \cup D\,, \\ H & \text{when } N \in K\,, \\ \hat{\mathbb{C}} \setminus D[c;r] & \text{when } N \in D\,, \end{cases}$$

where H is one of the open half-planes determined by the line L (which half-plane?). In particular, note that $D' \cap \mathbb{C}$ is an open set in $\mathbb{C}$.

Now we define the chordal topology in $\hat{\mathbb{C}}$ via the chordal distance function. First, we have the χ-open and χ-closed disks

$$D_\chi(a;r) := \{z \in \hat{\mathbb{C}} : \chi(z,a) < r\} \quad \text{and} \quad D_\chi[a;r] := \{z \in \hat{\mathbb{C}} : \chi(z,a) \le r\}\,.$$

A set $A \subset \hat{\mathbb{C}}$ is *χ-open* precisely when each of its points has a χ-open disk about it that lies in A; that is, A is a χ-open provided for each $a \in A$ there exists an $r > 0$ such that $D_\chi(a;r) \subset A$. We call $A \subset \hat{\mathbb{C}}$ *χ-closed* if $\hat{\mathbb{C}} \setminus A$ is χ-open.

Note that this chordal topology on $\hat{\mathbb{C}}$ makes stereographic projection $\mathbb{S} \xrightarrow{\Pi} \hat{\mathbb{C}}$ a homeomorphism. Notice also that each χ-open disk has the property that its intersection with $\mathbb{C}$ is an open set in $\mathbb{C}$; indeed, by the discussion two paragraphs above, this intersection is either an open Euclidean disk, an open half-plane, or the exterior of a closed Euclidean disk.

In addition, each Euclidean open disk is a χ-open set—in fact, it is even a χ-open disk (right?). It follows that each open set in $\mathbb{C}$ is also χ-open. We can summarize these observations by saying that

$$\text{the identity map} \quad (\mathbb{C}, |\cdot|) \xrightarrow{\text{id}} (\mathbb{C}, \chi) \quad \text{is a homeomorphism}\,;$$

there is no typo above—I really do want the set $\mathbb{C}$ in both places. ☺

A good exercise (to check your understanding of the chordal topology) is to validate the following assertion.

> *A set $A \subset \hat{\mathbb{C}}$ is χ-open if and only if either $A \subset \mathbb{C}$ and A is open, or $B := \hat{\mathbb{C}} \setminus A \subset \mathbb{C}$ and B is closed and bounded.*

This latter description for the chordal topology actually shows that $\hat{\mathbb{C}}$, with its chordal topology, is the so-called *one-point compactification* of $\mathbb{C}$.

Here is a natural problem to investigate: Given any continuous map $f : \mathbb{C} \to \mathbb{C}$, we would like to determine whether or not there is a continuous map $F : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ with $F|_{\mathbb{C}} = f$. What can you say about this when F is: a complex linear map? complex inversion? $z \mapsto z + 1/z$? a polynomial? a rational function? the complex exponential?

Let's look at the complex inversion $\mathbb{C}_* \xrightarrow{J} \mathbb{C}$ given by $J(z) := 1/z$. We start by defining

$$J(0) := \infty \quad \text{and} \quad J(\infty) := 0$$

which gives us a bijection $J : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$. We already know that J is continuous as a map of $\mathbb{C}_*$ to itself; here we can use either the Euclidean topology or the chordal one, right? Next, an easy computation shows that

$$\lim_{z \to 0} \chi(J(z), J(0)) = 0 \quad \text{and} \quad \lim_{z \to \infty} \chi(J(z), J(\infty)) = 0\,.$$

(Question: What does it mean to say that the second limit above is zero?) This permits us to assert that J is a continuous map of the metric space $(\hat{\mathbb{C}}, \chi)$ onto itself. Notice too that $J(J(z)) = z$ which means that $J^{-1} = J$. We conclude that J is a self-homeomorphism of $\hat{\mathbb{C}}$ onto itself. (Again, continuity here is with respect to the chordal distance function χ.) In fact it is straightforward to demonstrate that J is an isometry of $(\hat{\mathbb{C}}, \chi)$ onto itself (and then chordal continuity easily follows from this fact :-).

Recall that $1/z$ is a vector with modulus $1/|z|$ and direction given by $\bar{z}$. Thus we see that the map $w = J(z)$ interchanges the interior/exterior of the unit disk and also interchanges the upper/lower half-planes (but not the right/left half-planes). What does this correspond to on the sphere $\mathbb{S}$?

Notice that every self-map of the sphere $\mathbb{S}$ corresponds to a self-map of $\hat{\mathbb{C}}$ and vice versa (the correspondence being given via conjugation by stereographic projection). It is worthwhile to determine the associated maps of $\hat{\mathbb{C}}$ corresponding to the following self-maps of $\mathbb{S}$:

$$P \mapsto -P\,, \quad (x_1, x_2, x_3) \mapsto (x_1, -x_2, -x_3)\,, \quad (x_1, x_2, x_3) \mapsto (x_1, x_2, -x_3);$$

and likewise to determine the associated self-map of $\mathbb{S}$ corresponding to complex inversion $J(z) = 1/z$.

4. Holomorphicity

Here we begin our study of complex differentiability. Henceforth Ω **always** denotes an open subset of $\mathbb{C}$ (or, when so indicated, an open subset of $\hat{\mathbb{C}}$). Recall that $\mathcal{C}^1(\Omega)$ denotes the class of maps $f : \Omega \to \mathbb{C}$ that are $\mathcal{C}^1$, meaning that f has first order partial derivatives f_x and f_y at each point of Ω and these are continuous. (Equivalently, $f = u + iv$ is $\mathcal{C}^1$ provided u_x, u_y, v_x, v_y exist everywhere and are continuous.) We can also talk about real-valued maps that are $\mathcal{C}^1$; this should be clear in context. In addition, we use the notion of a $\mathcal{C}^\infty$ map: these possess partial derivatives of all orders (which are necessarily continuous).

4.A. **Complex Derivatives.** What are these?

We assume the reader is well acquainted with the notion of limit. Our definition of a complex derivative mimics that given in Freshman Calculus. Let $f : \Omega \to \mathbb{C}$ and fix a point $a \in \Omega$. We say that f is (complex) *differentiable* at a provided the limit

$$\lim_{z \to a} \frac{f(z) - f(a)}{z - a}$$

exists as a point in $\mathbb{C}$. When f is differentiable at a, we write $f'(a)$ to denote this limit. Whenever we use the term differentiable we always mean complex differentiable; later we shall also discuss real differentiability and then we shall always add the adjective *real*. We also emphasize that when one looks at the limit for the derivative of some function f at some point $z = a$, one always wants f to be defined in an open neighborhood of a.

A function f is *holomorphic* at a provided there is some disk $D(a; r)$ such that f is differentiable at each point in the disk. We say that f is *holomorphic* in Ω if f is differentiable at each point of Ω. We describe this by writing $f \in \mathcal{H}(\Omega)$.

The term *analytic* is also used to mean holomorphic. The reader should note that many authors also stipulate that holomorphic functions are $\mathcal{C}^1$ smooth. This is not part of our definition. However, we shall eventually prove that holomorphic maps do enjoy this property.

The usual calculus rules (linearity, power rule, product rule, quotient rule, chain rule, implicit and inverse function theorems) all continue to hold in the complex setting. (We shall have more to say about the Inverse Function Theorem later; see §4.D). As a consequence we obtain the following.

4.1. *Examples.* All complex polynomials are holomorphic in the entire plane. Complex rational functions (the quotient of two polynomials) are holomorphic where their denominator is non-zero. E.g., $T(z) = (az+b)/(cz+d)$ is holomorphic, with $T'(z) = (ad-bc)/(cz+d)^2$, for $z \in \mathbb{C} \setminus \{-d/c\}$.

The functions $\mathbb{C} \xrightarrow{f} \mathbb{C}$ that are holomorphic everywhere are called *entire*. Besides the complex polynomials, these include the exponential map e^z and so also the sine and cosine functions too. But at this point we cannot yet prove that e^z is entire. And what about the complex logarithm $\log(z)$, or the square root $\sqrt{z}$, or functions which are polynomials in x and y such as $x^2 + iy^3$: Are these holomorphic?

4.2. *Example.* The map $z \mapsto |z|^2 = z\,\bar{z} = x^2 + y^2$ is differentiable at $z = 0$ but nowhere else (it is real analytic everywhere).

4.3. *Exercise.* Prove that f differentiable at a implies f continuous at a.

4.4. *Exercise.* Prove that for $\mathbb{C} \supset \Omega \xrightarrow{f} \mathbb{C}$, these are equivalent:

(a) f is complex differentiable at a.
(b) There exist $c \in \mathbb{C}$ and $\varepsilon : \Omega \to \mathbb{C}$ such that

$$f(z) = f(a) + c(z-a) + \varepsilon(z) \qquad \text{where} \quad \lim_{z \to a} \frac{\varepsilon(z)}{z-a} = 0\,.$$

(c) There exists a function $\varphi : \Omega \to \mathbb{C}$ which is continuous at $z = a$ and satisfies

$$f(z) = f(a) + (z-a)\varphi(z) \qquad \text{for all } z \in \Omega\,.$$

What is $\varphi(a)$ in part (c) above?

When f is complex differentiable at a, we call $T_{f,a}(z) := f(a) + f'(a)(z-a)$ the complex linear or first-order approximation for f near $z = a$. It gives the 'best' complex linear approximation for f in the sense given in Exercise 4.5. Note that we well understand the geometry of the complex linear map $z \mapsto T_{f,a}(z)$ (right?), and this knowledge is useful in studying the geometry of the map f, at least near a when f is differentiable at a.

4.5. *Exercise.* Suppose that f is differentiable at a. Let $T := T_{f,a}$. Demonstrate that for any complex linear map L, there is a $\delta > 0$ such that for all $z \in D_*(a;\delta)$,

$$|f(z) - L(z)| \geq |f(z) - T(z)|\,.$$

4.B. **Cauchy Riemann Equations.** What are these?

Let us return to a question raised earlier. Suppose that we have a complex valued polynomial in x and y. Evidently such a function is real differentiable; at the very least, it has partial derivatives of all orders. Is there some way that we can decide whether or not such a function is complex differentiable?

Here is a nice 'warm-up' problem to get you thinking in the right direction.

4.6. *Exercise.* Suppose that f is real-valued. If f is differentiable at a, then $f'(a) = 0$.

We begin by deriving an necessary condition for complex differentiability. Suppose f is (complex) *differentiable* at a; then the limit

$$f'(a) = \lim_{z \to a} \frac{f(z) - f(a)}{z - a}$$

exists and is independent of the manner in which z approaches a. Fix θ and consider $z = a + re^{i\theta}$ with $r \searrow 0$. Given that $f'(a)$ exists, we must have

$$f'(a) = \lim_{r \searrow 0} \frac{f(z) - f(a)}{re^{i\theta}} = e^{-i\theta} D_{e^{i\theta}} f(a)$$

where

$$D_{e^{i\theta}} f(a) := \lim_{t \searrow 0} \frac{f(a + te^{i\theta}) - f(a)}{t}$$

is the *directional derivative* of f in the direction $e^{i\theta}$ at a. In particular, looking at the directions given by taking $\theta = 0$ or $\theta = \pi/2$—so $e^{i\theta}$ is 1 (giving the 'x-direction') or i (giving the 'y-direction')—we deduce that

$$\frac{\partial f}{\partial x}(a) = f'(a) = -i \frac{\partial f}{\partial y}(a).$$

CRE: The 'usual' proof—see HW #(59).
TFAE: $f' = f_x = -if_y$ (etc.), $u_x = v_y$ and $u_y = -v_x$

4.7. *Examples.* $z \mapsto |z|$ and $z \mapsto \bar{z}$ are both nowhere differentiable.

4.C. **Consequences of CRE.** Are the CREs NASC?

Here is an especially useful consequence of connectivity.

4.8. **Theorem.** *Let u be any function (real-valued or complex-valued) defined in some domain Ω (in $\mathbb{C}$ or in $\mathbb{R}^n$). Suppose that for each disk (or ball) $D \subset \Omega$, $u\,|_D$ is a constant. Then u is identically constant throughout Ω.*

A consequence of the above is the following generalization of a standard calculus fact (that follows from the Mean Value Theorem).

4.9. **Theorem.** *Suppose that $\mathbb{R}^n \supset \Omega \xrightarrow{u} \mathbb{R}$ has partial derivatives, and Ω is a domain. Then $\nabla u = 0$ if and only if u is constant.*

Here is a complex analog of the above. Compare this with Exercise 4.17.

4.10. **Theorem.** *Suppose f is holomorphic in a domain $\Omega \subset \mathbb{C}$. If any one of the following holds, then f is constant.*

(a) $f' = 0$.
(b) f *is purely real.*
(c) f *is purely imaginary.*
(d) $\bar{f}$ *is holomorphic.*
(e) $|f|$ *is constant.*
(f) $\arg(f)$ *is constant.*

Note that we will, eventually, prove that holomorphic maps are open. The above is an immediate consequence of this non-trivial fact.

NB: CRE holding at $z = a$ do not imply cplx diff at a; exs: (1) CDM's (all drxnl drtvs 0 but not even cts), (2) $f = u + iv$ with $f(0) = 0$ where $u(x, y) = (x^3 - y^3)/|z|^2$ and $v(x, y) = (x^3 + y^3)/|z|^2$ (f cts everywhere and CRE hold at 0, but f not diff'l at 0), (3) $f(z) = e^{-1/z^4}$ for $z \neq 0$ and $f(0) = 0$ (CRE hold everywhere but f not diff'l at 0)

thm: CRE plus cty implies cplx diff. (recall MVT)

cor: For $f \in \mathcal{C}^1(\Omega)$: f is holomorphic if and only if f satisfies the CREs.

exs: (1) $f(z) = \mathfrak{Im}(z^2) + i|z|^2 = 2xy + i(x^2 + y^2)$ is diff'l at each pt of $i\mathbb{R}$—but holomorphic nowhere, (2) e^z is entire with $(e^z)' = e^z$

4.D. **Chain Rule and Inverse Function Theorem.** Palka has an especially nice discussion of the Inverse Function Theorem, although in my opinion his proof is somewhat lacking.

4.E. **Complex Differential Operators.** "Dee Dee Zee & Dee Dee Zee Bar"

The complex differential operators are defined by

$$\frac{\partial}{\partial z} := \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \quad \text{and} \quad \frac{\partial}{\partial \bar{z}} := \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

These guys are specifically defined as above so that that the corresponding Chain Rules work exactly as expected, although these do need to be verified.

4.11. *Examples.* Look at: $|z|^2$, $\bar{f}(\bar{z})$, $|f|$ const implies so is f, $f(z) = |z|^2 + z/\bar{z} = (x^2 + y^2 + (x^2 - y^2)/(x^2 + y^2) + i(2xy/(x^2 + y^2)$ is diff'l exactly at $z = \pm 1$

TFAE: $f' = f_z = f_x = -if_y$, $u_x = v_y$ and $u_y = -v_x$, $f_{\bar{z}} = 0$.

NB: f_z and $f_{\bar{z}}$ may exist without f' existing; these are just abbreviations for certain (complex) linear combinations of f_x and f_y.

4.F. **Real Linear Transformations.** When is such a map holomorphic?

Here we examine real linear transformations $\mathbb{R}^2 \xrightarrow{M} \mathbb{R}^2$, but of course we want to view these as maps $\mathbb{C} \to \mathbb{C}$. Note that $\mathbb{C}$ is a two-dimensional real vector space and is naturally isomorphic to $\mathbb{R}^2$. To make this mathematically rigorous, we define

$$\mathbb{R}^2 \xrightarrow{\iota} \mathbb{C} \quad \text{by } \iota(x, y) := x + iy\,.$$

Then ι is a real linear vector space isomorphism, and $\iota^{-1}(z) = (\mathfrak{Re}(z), \mathfrak{Im}(z))$.

We see that each linear transformation $\mathbb{R}^2 \xrightarrow{M} \mathbb{R}^2$ corresponds to a real linear transformation $\mathbb{C} \xrightarrow{L} \mathbb{C}$ where $L := \iota \circ M \circ \iota^{-1}$, and also $M = \iota^{-1} \circ L \circ \iota$. See the diagram below.

$$\begin{array}{ccc} \mathbb{R}^2 & \xrightarrow{M} & \mathbb{R}^2 \\ \big\downarrow \iota & & \big\downarrow \iota \\ \mathbb{C} & \xrightarrow{L} & \mathbb{C} \end{array}$$

Notice that when $M(x, y) = (\alpha x + \gamma y, \beta x + \delta y)$ we have

$$L(x + iy) = (\alpha x + \gamma y) + i(\beta x + \delta y)\,. \tag{4.12}$$

4.13. **Proposition.** *Let* $\mathbb{R}^2 \xrightarrow{M} \mathbb{R}^2$ *be the linear transformation given by*

$$M(x,y) = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = (\alpha x + \gamma y, \beta x + \delta y)\,.$$

Define $\mathbb{C} \xrightarrow{L} \mathbb{C}$ *by* $L := \iota \circ M \circ \iota^{-1}$. *Then* L *has the form* $L(z) = A\,z + B\,\bar{z}$ *where*

$$A := \frac{1}{2}\left[(\alpha+\delta) + i(\beta-\gamma)\right] \quad \textit{and} \quad B := \frac{1}{2}\left[(\alpha-\delta) + i(\beta+\gamma)\right]\,. \tag{4.14}$$

Proof. Substitute $x = (z+\bar{z})/2$ and $y = (z-\bar{z})/2i$ into the formula (4.12) and simplify. ☐

We note that for a real linear transformation $\mathbb{R}^2 \xrightarrow{M} \mathbb{R}^2$ with associated $L := \iota \circ M \circ \iota^{-1}$, the formulas

$$M(x,y) = (\alpha x + \gamma y, \beta x + \delta y)\,,$$
$$L(x+iy) = (\alpha x + \gamma y) + i(\beta x + \delta y)\,,$$
$$L(z) = \frac{1}{2}\left[(\alpha+\delta) + i(\beta-\gamma)\right] z + \frac{1}{2}\left[(\alpha-\delta) + i(\beta+\gamma)\right] \bar{z}$$

all describe describe essentially the same function; the latter is just a 'complex variables' way of expressing M's action.

Using this information it is now easy to answer a question raised above; see §2.A, a paragraph or two just before 2.1. See also Theorem 5.3.

4.15. **Theorem.** *Define* $\mathbb{C} \xrightarrow{L} \mathbb{C}$ *by* $L := \iota \circ M \circ \iota^{-1}$ *where* $\mathbb{R}^2 \xrightarrow{M} \mathbb{R}^2$ *is the linear transformation with matrix* $\begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}$; *so* $L(z) = A\,z + B\,\bar{z}$ *as described in* (4.14). *The following are equivalent*

(a) $B = 0$.
(b) $\alpha = \delta$ *and* $\beta = -\gamma$.
(c) L *satisfies the CREs everywhere.*
(d) L *is entire (i.e., holomorphic everywhere).*
(e) L *is complex linear (in fact,* $L(z) = (\alpha + i\beta)z$*).*
(f) $L(iz) = iL(z)$

4.G. **Real Differentiability.** As compared to complex differentiability ⌣̈

Recall that a map $\mathbb{R}^n \supset \Omega \xrightarrow{f} \mathbb{R}^m$ is *real differentiable* at $a \in \Omega$ provided there exists a real linear transformation $M : \mathbb{R}^n \to \mathbb{R}^m$ satisfying

$$\lim_{x\to a} \frac{|f(x) - f(a) - M(x-a)|}{|x-a|} = 0\,. \tag{RD}$$

When this holds, we call $Df(a) := M$ the *derivative of* f *at* $x = a$. The matrix representation for $Df(a)$, with respect to the usual Euclidean bases, is given by the $m \times n$ matrix

$$Df(a) = \left(\frac{\partial f_i}{\partial x_j}\right)_{x=a} \qquad (1 \le i \le m, \le j \le n)$$

where $f = (f_1, \dots, f_m)$, $x = (x_1, \dots, x_n)$ and all of the partial derivatives are evaluated at $x = a$. This is called the *Jacobi* matrix for f, and its determinant is the *Jacobian* of f at a, i.e., $Jf(a) = \det Df(a)$.

Let's look at what we get when $n = 2$ and $\mathbb{R}^2 \supset \Omega \xrightarrow{F:=(U,V)} \mathbb{R}^2$. We find that

$$DF = \begin{pmatrix} U_x & U_y \\ V_x & V_y \end{pmatrix} \quad \text{and} \quad JF = \frac{\partial(U,V)}{\partial(x,y)} = \frac{\partial U}{\partial x}\frac{\partial V}{\partial y} - \frac{\partial V}{\partial x}\frac{\partial U}{\partial y}\,.$$

We assume F is real differentiable at the point (x_0, y_0). As in §4.F, we use $\mathbb{R}^2 \xrightarrow{\iota} \mathbb{C}$ to get an associated map $f : \iota(\Omega) \to \mathbb{C}$ and an associated real linear transformation $L := \iota \circ M \circ \iota^{-1}$, where $M = DF(x_0, y_0)$; L is the real derivative of f at $a := \iota(x_0, y_0) = x_0 + iy_0$.

We wish to rewrite the condition (RD) in terms of f and L (in place of F and M). To this end, we note that

$$F(x,y) - F(x_0,y_0) - M(x - x_0, y - y_0) =$$
$$(U(x,y) - U(x_0,y_0), V(x,y) - V(x_0,y_0)) - M(x - x_0, y - y_0)$$

corresponds exactly to

$$f(z) - f(a) - L(z-a) = ([U(x,y) - U(x_0,y_0)] + i[V(x,y) - V(x_0,y_0)]) - L(z-a)\,.$$

Thus we see that $\mathbb{C} \supset \Omega \xrightarrow{f} \mathbb{C}$ is *real differentiable at* $a \in \Omega$ provided

$$\lim_{z\to a} \frac{f(z) - f(a) - L(z-a)}{z-a} = 0$$

for some real linear transformation $L : \mathbb{C} \xrightarrow{L} \mathbb{C}$.

Now recall from Proposition 4.13 that such a linear transformation has the from

$$L(z) = A\,z + B\,\bar{z}\,.$$

Thus $\mathbb{C} \supset \Omega \xrightarrow{f} \mathbb{C}$ is *real differentiable at* $a \in \Omega$ provided there exist $A, B \in \mathbb{C}$ such that

$$\text{(CRD)} \qquad \lim_{z\to a} \frac{f(z) - f(a) - A(z-a) - B\overline{(z-a)}}{z-a} = 0\,.$$

Equivalently, there exist $A, B \in \mathbb{C}$ and $\Omega \xrightarrow{\eta} \mathbb{C}$ with $\lim_{z\to a} \dfrac{\eta(z)}{z-a} = 0$ and such that

$$\forall\, z \in \Omega\,, \quad f(z) = f(a) + A(z-a) + B\overline{(z-a)} + \eta(z)\,.$$

When (CRD) holds, the real linear transformation $Df(a) : \mathbb{C} \to \mathbb{C}$ defined by

$$Df(a)(z) := A\,z + B\,\bar{z}$$

is the real derivative of f at a, and the map $z \mapsto f(a) + Df(a)(z-a)$ is the real linear first order approximation to f near a.

Naturally the above should be contrasted to the case where $\mathbb{C} \supset \Omega \xrightarrow{f} \mathbb{C}$ is complex differentiable at $a \in \Omega$, which holds precisely when there exists $C \in \mathbb{C}$ such that

$$\text{(CD)} \qquad \lim_{z\to a} \frac{f(z) - f(a) - C(z-a)}{z-a} = 0\,.$$

See also Exercise 4.4 and the paragraph immediately thereafter.

In particular, from the above discussion, it should be evident that we have the following.

4.16. **Theorem.** *A map* $\mathbb{C} \supset \Omega \xrightarrow{f} \mathbb{C}$ *is complex differentiable at* $a \in \Omega$ *if and only if it is real differentiable at* a *and* $Df(a) : \mathbb{C} \to \mathbb{C}$ *is a complex linear transformation.*

We close this subsection by writing the real derivative $Df(a)$ in an especially convenient manner. Here we are assuming that $f : \Omega \to \mathbb{C}$ is real differentiable at $a \in \Omega$. Thus there are $A, B \in \mathbb{C}$ with (CRD) holding, and $Df(a)(z) = A\,z + B\,\bar{z}$. Recall (see the beginning of this subsection) that $Df(a)$ comes from the real linear transformation $M = DF(x_0, y_0)$. Here $F = (U, V)$ and $f = u + iv$ where $u = U \circ \iota^{-1}$, $v = V \circ \iota^{-1}$. That is, if $z = x + iy$, then

$$u(z) = U(x,y) \quad \text{and} \quad v(z) = V(x,y)\,.$$

Now

$$DF(x_0,y_0) = \begin{pmatrix} U_x & U_y \\ V_x & V_y \end{pmatrix}_{(x,y)=(x_0,y_0)} = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}_{z=a}$$

and according to 4.13—especially see (4.14)—we thus have

$$A := \frac{1}{2}\left[(u_x(a) + v_y(a)) + i(v_x(a) - u_y(a))\right] = \frac{1}{2}\left[f_x(a) - i\,f_y(a)\right] = f_z(a)$$

and

$$B := \frac{1}{2}\left[(u_x(a) - v_y(a)) + i(v_x(a) + u_y(a))\right] = \frac{1}{2}\left[f_x(a) + i\,f_y(a)\right] = f_{\bar{z}}(a)\,.$$

We conclude that $Df(a) : \mathbb{C} \to \mathbb{C}$ has the form

$$Df(a)(\zeta) = f_z(a)\zeta + f_{\bar{z}}(a)\bar{\zeta}\,.$$

We note that $Df(a)$ is complex linear if and only if $f_{\bar{z}}(a) = 0$.

4.17. *Exercise.* Suppose that $\mathbb{C} \supset \Omega \xrightarrow{f} \mathbb{C}$ is real differentiable at $a \in \Omega$. Prove that the following are equivalent:

(a) $Df(a)$ is complex linear.
(b) f is complex differentiable at $z = a$.
(c) f satisfies the CRE at a.

5. Conformal Mappings-part I

Recall that a map is conformal if it preserves oriented angles.

5.A. **Conformal Linear Transformations.** We are interested in understanding what maps preserve angles. With this in mind, we start by examining linear transformations. In this setting we may as well consider $\mathbb{R}^n$. In fact, we could even look at linear transformations mapping $\mathbb{R}^n$ to $\mathbb{R}^m$ (but we won't :-).

Since angles are defined in terms of inner products, we first examine the linear transformations that preserve the Euclidean inner product. It is not difficult to prove the following.

5.1. **Proposition.** *For a real linear transformation $U : \mathbb{R}^n \to \mathbb{R}^n$, the following are equivalent:*

(a) $U^{\mathrm{tr}}U = I$.
(b) $\forall\, x, y \in \mathbb{R}^n\,,\ U(x) \cdot U(y) = x \cdot y$.
(c) $\forall\, x \in \mathbb{R}^n\,,\ \|U(x)\| = \|x\|$.

Here I is the $n \times n$ identity matrix. The group of linear transformations $U : \mathbb{R}^n \to \mathbb{R}^n$ which satisfy $U^{\mathrm{tr}} U = I$ is usually denoted $\mathcal{O}(n)$ and these are called *orthogonal* transformations. Notice that such transformations have determinants $\det U = \pm 1$.

A real linear transformation $M : \mathbb{R}^n \to \mathbb{R}^n$ is *isogonal* if it preserves angles. Recall that the angle θ between two non-zero vectors $x, y \in \mathbb{R}^n$ is defined by requiring

$$\theta \in [0, \pi] \quad \text{and} \quad \cos(\theta) = \frac{x \cdot y}{\|x\| \|y\|}.$$

The Cauchy-Schwarz Inequality says that $|x \cdot y| \leq \|x\| \|y\|$ and so such an angle θ exists.

An immediate consequence of Proposition 5.1 is that each $U \in \mathcal{O}(n)$ is isogonal. A moments thought reveals that for any $s \in \mathbb{R} \setminus \{0\}$ and any $U \in \mathcal{O}(n)$, sU is also isogonal.

Recall that a real linear transformation $M : \mathbb{R}^n \to \mathbb{R}^n$ is a *similarity* if there is some $s > 0$ such that

$$\forall\, x, y \in \mathbb{R}^n, \quad \|M(x) - M(y)\| = s \|x - y\| .$$

For example, a dilation $x \mapsto s\, x$, is a similarity.

5.2. **Proposition.** *For a real linear transformation $M : \mathbb{R}^n \to \mathbb{R}^n$, the following are equivalent:*

(a) *M is isogonal.*
(b) *There exists $s > 0$ such that $M\, M^{\mathrm{tr}} = s\, I$.*
(c) *There exists $s > 0$ and $U \in \mathcal{O}(n)$ such that $M = s\, U$.*
(d) *There exists $s > 0$ such that $\forall\, x, y \in \mathbb{R}^n$, $M(x) \cdot M(y) = s(x \cdot y)$.*
(e) *There exists $s > 0$ such that $\forall\, x \in \mathbb{R}^n$, $\|M(x)\| = s\, \|x\|$.*
(f) *M is a similarity.*

In class I indicated a geometric proof and gave a detailed algebraic proof.

Let us focus our attention on dimension $n = 2$. Here we know how to talk about *oriented* angles; see the end of §1.B. Notice that $(-\beta, \alpha)$ is the unique vector that is: orthogonal to (α, β), has the same length as (α, β), and is such that the oriented angle from (α, β) to this vector is positive (i.e., is $\pi/2$). Of course in complex notation this readily follows from the observation that $i(\alpha + i\beta) = -\beta + i\alpha$.

A real linear transformation $M : \mathbb{R}^2 \to \mathbb{R}^2$ is *conformal* if it preserves oriented angles. What is the matrix representation for such a map? Well, by what we just observed, if $(\alpha, \beta) = M(1, 0)$, then we must have $(-\beta, \alpha) = M(0, 1)$. Thus the matrix representation for any conformal real linear transformation M has the form

$$M = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} .$$

Glancing back at §4.F we see that the linear transformation $\mathbb{C} \xrightarrow{L} \mathbb{C}$ associated with M is complex linear; see especially Theorem 4.15(b,e).

Let's investigate this further. From Proposition 5.2 we know that there is a constant $r > 0$ such that $r^2 = \det(M) = \alpha^2 + \beta^2$. Thus we can find θ so that

$$\alpha = r \cos\theta \qquad \text{and} \qquad \beta = r \sin\theta \, .$$

It now follows that

$$M = \begin{pmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{pmatrix} = r \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = r\, U$$

so

$$M\begin{bmatrix} x \\ y \end{bmatrix} = r\,U\begin{bmatrix} x \\ y \end{bmatrix} == r\begin{pmatrix}(\cos\theta)x-(\sin\theta)y \\ (\sin\theta)x+(\cos\theta)y\end{pmatrix};$$

here U is the orthogonal matrix

$$U := \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}.$$

Identifying (x,y) with $z = x+iy$ we see that the above becomes $L(z) = re^{i\theta}z$ where $\mathbb{C}\xrightarrow{L}\mathbb{C}$ is the linear transformation associated with M; see (4.14) or Theorem 4.15(e).

We can summarize the above discussion to obtain a long list of equivalent conditions for real linear transformations. See also Theorem 4.15.

5.3. **Theorem.** *Define* $\mathbb{C}\xrightarrow{L}\mathbb{C}$ *by* $L := \iota\circ M\circ\iota^{-1}$ *where* $\mathbb{R}^2\xrightarrow{M}\mathbb{R}^2$ *is the linear transformation with matrix* $\begin{pmatrix}\alpha & \gamma \\ \beta & \delta\end{pmatrix}$; *so* $L(z) = A\,z + B\,\bar{z}$ *as described in* (4.14). *The following are equivalent*

(a) M *is conformal.*
(b) $A \neq 0 = B$.
(c) $\alpha = \delta$ *and* $\beta = -\gamma$ *and* $\alpha\delta - \beta\gamma \neq 0$.
(d) L *satisfies the CREs everywhere and* $L' \neq 0$.
(e) L *is entire (i.e., holomorphic everywhere) and* $L' \neq 0$.
(f) L *is a non-constant complex linear map (in fact,* $L(z) = (\alpha + i\beta)z$*).*
(g) $L(iz) = iL(z)$ *and* L *is non-constant.*

5.B. **Conformal Diffeomorphisms of Plane Domains.** Defns: paths; differentiable and *regular* paths; oriented angle between two regular paths; curvilinear angles; isogonal and conformal $\mathcal{C}^1$-diffeomorphisms.

We call $\mathbb{C}\supset\Omega\xrightarrow{f}\Omega'\subset\mathbb{C}$ a $\mathcal{C}^1$-*diffeomorphism* provided f is a $\mathcal{C}^1$ homeomorphism and $Jf(z)\neq 0$ for all $z\in\Omega$.

5.4. **Theorem.** *Let* $\Omega\xrightarrow{f}\Omega'$ *be a* $\mathcal{C}^1$-*diffeomorphism of domains* $\Omega,\Omega'\subset\mathbb{C}$. *Fix* $a\in\Omega$. *These are equivalent:*

(a) f *is complex differentiable at* a.
(b) f *is conformal at* a.
(c) f *is isogonal at* a *and* $Jf(a) > 0$.
(d) *The infinitesimal linear change of scale induced by* f *at* a *is independent of direction, and* $Jf(a) > 0$.

The first condition in (d) above simply means that there exists $\lambda > 0$ such that for all $\theta\in\mathbb{R}$,

$$\lambda = \lim_{r\searrow 0}\frac{|f(a+re^{i\theta}) - f(a)|}{r};$$

that is, the limit exists and equals λ.

Theorem 5.4 is proven on pages 73–64 in Ahlfors. Another reference is Theorem 1.1 on page 380 in Palka.

Later, next Quarter, we shall prove that complex differentiable maps are always $\mathcal{C}^1$!

6. Möbius transformations-part I

This section could also be called Elementary Mappings-part II or Conformal Mappings-part II.

Here we begin our investigation of mappings of the form

$$z \mapsto w = \frac{az+b}{cz+d} \qquad \text{where} \quad a,b,c,d \in \mathbb{C}\,.$$

These are the nicest possible rational functions being simply a quotient of two complex linear maps. We will see that these maps possess many of the same properties enjoyed by the complex linear functions.

6.A. **The Möbius Group.** A *Möbius transformation* is a mapping of the form

$$w = T(z) := \frac{az+b}{cz+d} \qquad \text{with } a,b,c,d \in \mathbb{C},\ ad-bc \neq 0, \text{ and (for now) } z \in \mathbb{C} \setminus \{-d/c\}\,.$$

Note that we can always assume, if we want, that $ad - bc = 1$.

The reader perhaps is wondering about the stipulation that $ad - bc \neq 0$ in order for the formula $T(z) := (az + b)/(cz + d)$ to define a Möbius transformation. We indicate three reasons why this requirement is included. First, according to the Quotient Rule, such a function T is holomorphic in $\mathbb{C} \setminus \{-d/c\}$ with

$$T'(z) = \frac{ad-bc}{(cz+d)^2}\,.$$

Thus $ad - bc \neq 0$ ensures that T is not a constant map. Another way to see this is to check that

$$T(z) - T(z') = \frac{(ad-bc)(z-z')}{(cz+d)^2}\,;$$

notice that this formula can also be used directly to compute $T'(z)$.

A perhaps more illuminating calculation is to use long-division and write $T(z)$ as follows. When $c \neq 0$ we have

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

Evidently, when $c = 0$, the Möbius transformation T reduces to the complex linear map $w = (a/d)z + (b/d)$. Thus again we see that the requirement $ad - bc \neq 0$ simply ensures that T is not constant map $z \mapsto a/c$.

We record the last observation above as

$$T(z) = \frac{az+b}{cz+d} = \begin{cases} (a/d)z + (b/d) & \text{when } c = 0\,, \\ (a/c) + (b-ad/c)(cz+d) & \text{when } c \neq 0\,. \end{cases}$$

In particular we see that a/c does not belong to the range of T, i.e., $a/c \notin T(\mathbb{C} \setminus \{-d/c\})$.

We emphasize the above: Each Möbius transformation can be expressed as a composition of a complex linear map, followed by complex inversion, followed by another complex linear map. In particular, the above arithmetic permits us to write (when $c \neq 0$)

$$T(z) = \frac{az+b}{cz+d} = M \circ J \circ L(z)$$

where $J(w) = 1/w$ is complex inversion and L, M are the complex linear maps

$$L(z) = cz + d \qquad \text{and} \qquad M(\zeta) = (b - ad/c)\zeta + (a/c)\,.$$

We can deduce many properties of a general Möbius map by knowing the properties of complex linear maps and complex inversion. This is especially helpful and I cannot emphasize it enough!

Many of the basic properties of Möbius transformations can be readily deduced from the analagous properties of complex linear maps, but don't forget about the complex inversion. Because of this, we recall some of the important properties of a complex linear map $w = L(z) := az + b$ where $a, b \in \mathbb{C}$ with $a \neq 0$. (The reader is encouraged to review §2.A.)

(1) L is the composition of: a rotation (about the origin) by the angle $\text{Arg}(a)$, followed by a dilation (with respect to the origin) by $|a|$, followed by a translation by b.
(2) L is a complex polynomial of degree 1.
(3) L is a self-homeomorphism of the plane with inverse $z = L^{-1}(w) = (w - b)/a$.
(4) L maps lines and circles to lines and circles respectively.
(5) L is a conformal entire map.

We continue our study of Möbius transformations $w = T(z) := (az + b)/(cz + d)$. We assume that $c \neq 0$ so that T is a 'true' Möbius transformation. According to the 'quotient rule for limits', we know that T is continuous, with respect to Euclidean distance, at each point $z \in \mathbb{C} \setminus \{-d/c\}$. We would like to consider T as a map of $\hat{\mathbb{C}}$ to itself. Motivated by our earlier work above with complex inversion (see the end of §3.D), we define

$$T(-d/c) := \infty \quad \text{and} \quad T(\infty) := a/c\,.$$

A good exercise is to check that the above definitions are precisely what is needed to ensure that T is a continuous map on $(\hat{\mathbb{C}}, \chi)$. (Alternatively, we can view T as the composition of J with appropriate complex linear maps where we extend a (non-constant) linear map $L : \mathbb{C} \to \mathbb{C}$ to a continuous map $L : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ by defining $L(\infty) := \infty$. Then use the fact that a composition of continuous maps is continuous.) Either way, we obtain a continuous map $T : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$. Moreover, T is a bijection because we can solve $w = T(z)$ for z obtaining

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

Notice that $T^{-1}(\infty) = -d/c$ and $T^{-1}(a/c) = \infty$. In particular, we see that T^{-1} is also a Möbius transformation; hence it too is continuous. Thus we have demonstrated that each Möbius transformation is a self-homeomorphism of the Riemann sphere $\hat{\mathbb{C}}$. (Again, here continuity is with respect to the chordal distance χ.)

We mention that the set of all Möbius transformations forms a group where the group product is given by composition. Indeed, we have already seen that the inverse of a Möbius map is again Möbius and it is straightforward to show that the composition of two such maps is again a Möbius transformation. We can also verify these properties by identifying the Möbius maps with a certain subgroup of 2 by 2 matrices with complex coefficients; of course

$$T(z) := \frac{az + b}{cz + d} \quad \text{is identified with the matrix} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (\text{usually with } ad - bc = 1)\,.$$

One advantage of this viewpoint is the easy calculuation of compositions (products) and inverses of Möbius transformations. But one needs to be extremely careful with this identification! We will not use this in any way.

Here is a list of some of the important (elementary) properties that each Möbius transformation $T(z) := (az + b)/cd + d)$ (with $ad - bc \neq 0$) satisfies:

(1) T is a complex rational function of degree 1.
(2) T a self-homeomorphism of $(\hat{\mathbb{C}}, \chi)$ with inverse $T^{-1}(w) = (dw - b)/(-cw + a)$, provided we put $T(-d/c) = \infty$ and $T(\infty) = a/c$.
(3) T is either itself a complex linear map (when $c = 0$) or the composition of a complex linear map, followed by complex inversion, followed by another complex linear map.
(4) T is a holomorphic conformal map from $\mathbb{C} \setminus \{-d/c\}$ onto $\mathbb{C} \setminus \{-c/a\}$.

6.B. **Möbius Transformations and Circles.** Since each complex linear map transforms lines and circles to lines and circles respectively, we might expect a similar result for Möbius transformations. We must examine is the effect of complex inversion $z \mapsto J(z) := 1/z$ on lines and circles.

6.1. **Proposition.** *If K is a circle or an extended line in $\hat{\mathbb{C}}$, then so is $J(K)$. Also, $J(K)$ is an extended line if and only if $0 \in K$, and $0 \in J(K)$ if and only if K is an extended line.*

Proof. The equation for a line or circle, $r|z|^2 + cz + \bar{c}\bar{z} + s = 0$ with $r, s \in \mathbb{R}$, $c \in \mathbb{C}$ and $|c|^2 > rs$, is easily seen to transform into a similar such equation under complex inversion. □

An alternate proof is available: consider the effect of $\Pi^{-1} \circ J \circ \Pi$ acting on circles on $\mathbb{S}$ and recall that such circles correspond, under stereographic projection Π, precisely to circles and extended lines in $\hat{\mathbb{C}}$.

Recall our notation (see the ends of §1.C and §3.D): $\mathcal{C}$ and $\mathcal{L}$ denote the collections of all circles and lines (respectively) in $\mathbb{C}$, and $\hat{\mathcal{C}} := \mathcal{C} \cup \hat{\mathcal{L}}$ where $\hat{\mathcal{L}} := \{\hat{L} : L \in \mathcal{L}\}$ (with $\hat{L} = L \cup \{\infty\}$ for each $L \in \mathcal{L}$). In these terms the above result says that J maps each $K \in \hat{\mathcal{C}}$ to some $K' \in \hat{\mathcal{C}}$ and also tells us how to determine when $K' \in \hat{\mathcal{L}}$ or when K' passes through the origin.

The following fundamental fact concerning Möbius transformations is an immediate consequence of the related facts concerning complex linear maps and complex inversion. Loosely speaking, each Möbius transformation maps each circle in $\hat{\mathbb{C}}$ to another circle in $\hat{\mathbb{C}}$: the Möbius image of any $K \in \hat{\mathcal{C}}$ is again in $\hat{\mathcal{C}}$.

6.2. **Theorem.** *Let $T(z) := (az + b)/(cz + d)$ be a Möbius transformation; so $a, b, c, d \in \mathbb{C}$ and $ad \neq bc$. If $K \in \hat{\mathcal{C}}$ is a circle or an extended line in $\hat{\mathbb{C}}$, then $T(K) \in \hat{\mathcal{C}}$. Also, $T(K) \in \hat{\mathcal{L}}$ is an extended line if and only if $-d/c \in K$, and $a/c \in T(K)$ if and only if $K \in \hat{\mathcal{L}}$ is an extended line. Moreover, T maps K bijectively onto $T(K)$.*

Proof. Write $T = M \circ J \circ L$ where J is complex inversion and M, L are the appropriate complex linear maps. Let K be a circle or extended line in $\hat{\mathbb{C}}$. Then $L(K)$ is a circle or extended line, so $J(LK)$ is too and therefore so is $K' = T(K) = M(JLK)$. Also, K' will be an extended line exactly when it contains ∞, which is when K goes through $-d/c$. Similarly, K is an extended line precisely when it contains ∞ in which case K' contains the point $T(\infty) = a/c$. □

Our next two examples reveal how we can utilize both the circle-preserving and angle-preserving (i.e., conformality) properties of Möbius transformations to easily determine images of various circles and lines.

6.3. *Example.* Determine the images of $\mathbb{R}$, $i\mathbb{R}$, and $\mathbb{T}$ under the map $w = T(z) = \dfrac{1-z}{1+z}$.

6.4. *Example.* Determine the images of $\mathbb{R}$, $i\mathbb{R}$, and $C(0;2)$ under the map

$$w = T(z) = \frac{z-i}{z+2}\,.$$

We are interested in the following two **Basic Problems** (and their variations).

(a) Given a Möbius transformation T and a circle or extended line K in $\hat{\mathbb{C}}$, find $T(K)$.
(b) Given K and K', each a circle or extended line in $\hat{\mathbb{C}}$, find a Möbius transformation T with $T(K) = K'$.

We obtain variations of each of these problems by considering the 'sides' of K and/or K'; circles have an inside (a disk) and an outside (a χ-disk in $\hat{\mathbb{C}}$) whereas extended lines have two sides (which are both half-planes in $\mathbb{C}$, but chordal disks—i.e., χ-disks—in $\hat{\mathbb{C}}$).

As the examples above illustrate, the circle-preserving and angle-preserving (i.e., conformality) properties of Möbius transformations go a long way towards easy calulations of images of various circles and lines under a given Möbius transformation. Nonetheless, there are other properties of Möbius transformations that will prove especially useful.

Here is an example of the type of problem we would like to be able to answer with a minimal amount of effort.

6.5. *Exercise.* Determine $a \in \mathbb{C}$ and $r > 0$ so that $0,\ 1,\ 2+i$ lie on the circle $C(a;r)$.

Recall that any three distinct points do determine a unique circle in $\hat{\mathcal{C}}$ which passes through the given points. Thus the above problem has a solution. Of course we can resort to high school algebra and just do a bunch of arithmetic to solve this problem. We seek an *elegant* solution :-).

6.C. **Fixed Points-part I.** Recall that z_0 is a *fixed point* of f provided $f(z_0) = z_0$. For many classes of maps, the set of fixed points for a given map plays an important role in understanding the dynamics of the map in question.

6.6. *Example.* Consider a linear map $L(z) = az + b$ with $a \neq 0$. If $a = 1$, then L is just translation by b and so has no fixed points in $\mathbb{C}$; however, recall that we defined $L(\infty) = \infty$, so when $a = 1$ we see that L has a unique fixed point in $\hat{\mathbb{C}}$, namely ∞. When $a \neq 1$ we find (by solving $z = az + b$) that L has a second fixed point given by $z = b/(1-a)$. The interested reader should examine the pictures which illustrate the dynamics of the special linear maps: translations $w = z + b$, dilations $w = kz$ $(k > 0)$, rotations $w = e^{i\theta}z$. (What are the fixed points? The orbits?)

6.7. *Example.* Complex inversion $J(z) := 1/z$ has two fixed points in $\hat{\mathbb{C}}$, namely 1 and -1. What is the dynamical picture on $\hat{\mathbb{C}}$?

Let's determine the fixed points of $T(z) := (az+b)/(cz+d)$. To this end, we solve $z = T(z)$ which leads to the quadratic equation

$$c\,z^2 + (d-a)z - b = 0\,.$$

We note that a general quadratic equation $A\,z^2 + B\,z + c = 0$ has one or two solutions depending on whether or not $B^2 = 4AC$; it can have more than two distinct solutions if and only if $A = B = C = 0$. Now we return to the equation $z = T(z)$. The only way it can have more than two distinct solutions (i.e., T can have more than two distinct fixed points) is if

$a = d$ and $b = c = 0$, which of course means that $T(z) = z$. Assuming that T is not the identity transformation, and also that $ad - bc = 1$, we find solutions

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

these are the only fixed points of T. We record this information.

6.8. **Theorem.** *The only Möbius transformation with more than two fixed points is the identity. More precisely, if*

$$w = T(z) := \frac{az+b}{cz+d} \qquad \text{with } a,b,c,d \in \mathbb{C} \text{ and } ad-bc = 1$$

is not the identity transformation, then T has
(a) *exactly one fixed point if $a + d = \pm 2$,*
(b) *exactly two fixed points if $a + d \neq \pm 2$.*

The above theorem has surprisingly many useful applications. We begin with the following.

6.9. **Corollary.** *Given three distinct points a, b, c in $\hat{\mathbb{C}}$, there exists a unique Möbius transformation T with*

$$T(a) = 0\,, \quad T(b) = \infty\,, \quad T(c) = 1\,.$$

Proof. When $a, b, c \in \mathbb{C}$ (i.e. none of a, b, c are ∞),

$$T(z) = \frac{z-a}{z-b}\frac{b-c}{a-c}$$

has the asserted property. If a or b or c is ∞, we simply replace the above expression for $T(z)$ by

$$\frac{c-b}{z-b} \quad \text{or} \quad \frac{z-a}{c-a} \quad \text{or} \quad \frac{z-a}{z-b}$$

respectively. Thus in all cases we have found a Möbius transformation T with the desired properties. If S is some other Möbius transformation that also has this property, then $T \circ S^{-1}$ is a Möbius transformation which fixes $0, \infty, 1$, so $T \circ S^{-1}$ must be the identity and hence $T = S$. □

6.10. **Theorem.** *Given two pairs of three distinct points a, b, c and a', b', c' in $\hat{\mathbb{C}}$, there exists a unique Möbius transformation T with $T(a) = a'$, $T(b) = b'$, $T(c) = c'$.*

Proof. Let T and S be the Möbius transformations that map a, b, c and a', b', c' to $0, \infty, 1$ respectively. Then $S^{-1} \circ T$ is the desired Möbius map. As above it must be unique. □

In practice we can find T by solving

$$\frac{w-a'}{w-b'}\frac{b'-c'}{a'-c'} = \frac{z-a}{z-b}\frac{b-c}{a-c} \tag{6.11}$$

for $w = T(z)$.

6.12. *Exercise.* The Möbius transformation that maps $i, -2, -i$ to $0, \infty, 1$ is

$$w = \frac{i-2}{2i}\frac{z-i}{z+2}\,.$$

6.13. *Exercise.* The Möbius transformation that maps $i, 2, -2$ to $i, 1, -1$ is

$$w = \frac{3z + 2i}{iz + 6}\,.$$

6.14. *Remark.* We now know how to construct a Möbius transformation that maps any given triple onto another given triple. Recall that any circle or extended line in $\hat{\mathbb{C}}$ is uniquely determined by choosing three distinct points on it. Thus given K and K' (each a circle or extended line in $\hat{\mathbb{C}}$) we can find a Möbius transformation T with $T(K) = K'$; indeed, just choose distinct points $a, b, c \in K$ and distinct $a', b', c' \in K'$, and let T be the Möbius transformation that maps a, b, c to a', b', c'. (Note, however, that T is not unique in the sense that there are *many* Möbius transformations that map K to K'.) In other words, we have now solved the second of our two problems.

Take another look at the problem raised at the end of §6.B; see Exercise 6.5. Clearly, if T is the Möbius transformation that maps $0, 1, 2+i$ to say $0, 1, \infty$, then $T[C(a; r)] = \hat{\mathbb{R}}$. Is there some way that we can use this information to determine a and r?

6.D. **Cross Ratios.** We define the *cross-ratio* $[a, b, c, d]$ of four distinct points $a, b, c, d \in \hat{\mathbb{C}}$ by

$$[a, b, c, d] := \frac{(a-b)(c-d)}{(a-c)(b-d)}\,;$$

of course this must be interpreted appropriately when one of a or b or c or d is the point at infinity. Please note that this is <u>not the same definition</u> as that given by Ahlfors nor that given by Palka. However, there is actually very little difference as should become apparent soon; the only real difference has to do with the ordering of the points a, b, c, d. We believe that the above ordering is the easiest to remember.

6.15. *Remarks.* (a) The cross-ratio $[a, b, c, d]$ is well-defined as soon as we have three distinct points. For example, it is

$$0 \quad \text{if } a = b \text{ or } c = d\,, \qquad \infty \quad \text{if } a = c \text{ or } b = d\,, \qquad 1 \quad \text{if } a = d \text{ or } b = c\,.$$

(b) $T(z) := [z, a, b, c] = \dfrac{z-a}{z-b}\dfrac{b-c}{a-c}$ is the unique Möbius transformation that maps the triple a, b, c to $0, \infty, 1$ respectively. In particular, $z = [z, 0, \infty, 1]$.

(c) The unique Möbius transformation T that maps a, b, c to a', b', c' can be found by solving $[w, a', b', c'] = [z, a, b, c]$ to get $w = T(z)$; cf. (6.11).

(d) Of the twenty-four possible values of $[a, b, c, d]$, $[b, c, d, a]$, ... obtained by looking at all possible permutations of the four points a, b, c, d, there are actually only six different values each of which can be nicely expressed in terms of $z = [a, b, c, d]$. For example, from $z = [z, 0, \infty, 1]$ we find that $[0, \infty, 1, z] = 1 - z$. (See HW#(115).)

Now we demonstrate that cross-ratios are invariant with respect to Möbius transformations.

6.16. **Theorem.** *Möbius transformations preserve cross ratios: If T is any Möbius transformation, then for all distinct $a, b, c, d \in \hat{\mathbb{C}}$,*

$$[Ta, Tb, Tc, Td] = [a, b, c, d]\,.$$

Proof. Let T be a Möbius transformation. Consider any four distinct points $a, b, c, d \in \hat{\mathbb{C}}$. Put $a' := T(a)$, $b' := T(b)$, $c' := T(c)$, $d' := T(d)$. Define Möbius transformations f, g by $f(z) := [z, b, c, d]$ and $g(z) := [z, b', c', d']$. Then f and g are the unique Möbius transformations that map b, c, d and b', c', d' to $0, \infty, 1$ respectively. It follows that $g^{-1} \circ f$ is the unique Möbius transformation that maps b, c, d to b', c', d'. Therefore, $g^{-1} \circ f = T$, so $f = g \circ T$. Thus

$$[a, b, c, d] = f(a) = g(Ta)) = g(a') = [a', b', c', d'] = [Ta, Tb, Tc, Td]$$

as asserted. □

6.E. **Symmetry and Reflections.** When we view complex conjugation as a mapping, its geometric significance is that it gives reflection across the real axis. Writing $\rho_{\mathbb{R}}(z) := \bar{z}$ we see that $\rho_{\mathbb{R}}$, called *reflection across* $\mathbb{R}$, is a self-homeomorphism of $\mathbb{C}$ which is an orientation reversing isogonal isometry with ${\rho_{\mathbb{R}}}^{-1} = \rho_{\mathbb{R}}$, and moreover the set of fixed points of $\rho_{\mathbb{R}}$ is precisely $\mathbb{R}$.

Given any line L, we can define a similar map $\mathbb{C} \xrightarrow{\rho_L} \mathbb{C}$, called *reflection across* L, that is an orientation reversing isogonal (so, anti-conformal) isometric self-homeomorphism of $\mathbb{C}$ with ${\rho_L}^{-1} = \rho_L$, and the set of fixed points of ρ_L is precisely L. This reflection is determined by the requirement that the points z and $z^* = \rho_L(z)$ be *symmetric* with respect to L; that is, L is the perpendicular bisector of the Euclidean line segment $[z, z^*]$. See Figure 3 and HW#'s(30,31). (The reader should derive a formula for ρ_L if L is given in some standard form, e.g. $L = \{z : cz + \overline{cz} + s = 0\}$ where $c \in \mathbb{C}_*$ and $s \in \mathbb{R}$, or alternatively, $L = \{z : \mathfrak{Re}[(z-a)\bar{\nu}] = 0\}$ for some $a \in \mathbb{C}$ and $\nu \in \mathbb{C}_*$.)

We wish to develop a similar notion for circles. We could just give a formula (see Palka). We could do this solely in terms of plane geometry. We follow Ahlfors and use cross ratios. Fix $K \in \hat{\mathcal{C}}$ and select distinct points $a, b, c \in K$. Let T be the unique Möbius transformation that maps a, b, c to $0, \infty, 1$ respectively. We wish to declare points z, z^* to be symmetric with respect to K if $T(z^*) = \overline{T(z)}$. One potential problem with such a definition is that, in principle, it might depend on our choice of the points a, b, c. Does it? Does this notion depend only on the circle K? Two other questions also arise: What does this give if K is an extended line? (Does it give the same thing as discussed above? Certainly we want this to be the case!) And, perhaps most important, what the heck does this mean?

The reader should have no problem verifying the following.

6.17. **Lemma.** *Let T be any Möbius transformation with the property that $T(\hat{\mathbb{R}}) = \hat{\mathbb{R}}$. Then T can be expressed using real coefficients: there exist $a, b, c, d \in \mathbb{R}$ with $ad \neq bc$ and such that for all $z \in \hat{\mathbb{C}}$, $T(z) = (az+b)/(cz+d)$. Consequently, for all $z \in \hat{\mathbb{C}}$, $\overline{T(z)} = T(\bar{z})$.*

Using the above Lemma we can establish the following.

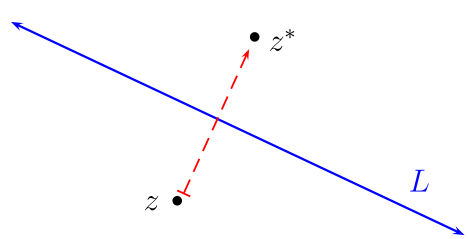

FIGURE 3. Reflection across K

6.18. **Corollary.** *Let $K \in \hat{\mathcal{C}}$. Fix $z, z^* \in \hat{\mathbb{C}} \setminus K$. Then*

$$\exists\, a,b,c \in K \qquad \textit{such that } \; [z^*,a,b,c] = \overline{[z,a,b,c]}$$

if and only if

$$\forall\, a,b,c \in K \qquad \textit{we have } \; [z^*,a,b,c] = \overline{[z,a,b,c]}\,.$$

Thanks to the preceding corollary, we can now make the following definition. Two points z, z^* are said to be *symmetric with respect to* $K \in \hat{\mathcal{C}}$ provided

$$\text{for some (hence for all) distinct } a,b,c \in K\,, \quad [z^*,a,b,c] = \overline{[z,a,b,c]}\,.$$

Notice that if $z \in K$, then $[z,a,b,c] \in \mathbb{R}$ and so the only way that $[z^*,a,b,c] = \overline{[z,a,b,c]}$ can hold is that $[z^*,a,b,c] = [z,a,b,c]$ and thus that $z^* = z$.

We shall see that for each $z \in \hat{\mathbb{C}}$ there is a unique $z^* \in \hat{\mathbb{C}}$ such that z and z^* are symmetric with respect to K. This means that we can define *reflection across* K as the map $\rho_K : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ given by $\rho_K(z) := z^*$. We shall see that ρ_K is an anti-conformal (so, orientation reversing isogonal) self-homeomorphism of $\hat{\mathbb{C}}$ with ${\rho_K}^{-1} = \rho_K$ and K being precisely the set of fixed points of ρ_K.

Here is an especially important and useful property of this notion of symmetry.

6.19. **Theorem.** *Möbius transformations preserve symmetry: If z and z^* are symmetric with respect to some $K \in \hat{\mathcal{C}}$, then for any Möbius transformation T, $T(z)$ and $T(z^*)$ are symmetric with respect to $T(K)$.*

Proof. Assume that z, z^* are symmetric with respect to some $K \in \hat{\mathcal{C}}$. Let T be any Möbius transformation. Fix distinct $a,b,c \in K$. Then, since T preserve cross ratios,

$$[T(z^*),T(a),T(b),T(c)] = [z^*,a,b,c] = \overline{[z,a,b,c]} = \overline{[T(z),T(a),T(b),T(c)]}\,. \quad \square$$

Now we explore the geometric significance of symmetry. First, let us consider the case $K = \hat{L}$ for some straight line L. In this setting we can select the point at infinity as one of our points; since $[z,a,\infty,c] = (z-a)/(c-a)$, we deduce that z, z^* are symmetric with respect to L if and only if for all distinct $a, c \in L$,

$$\frac{z^*-a}{c-a} = \overline{\left(\frac{z-a}{c-a}\right)} \quad \text{or equivalently} \quad z^* - a = \frac{c-a}{\bar{c}-\bar{a}}\,(\bar{z}-\bar{a})\,.$$

In particular, for all $a \in L$, $|z^*-a| = |z-a|$. Of course this trivially holds if $z^* = z$ (which <u>is</u> true if $z \in L$). However, if $z^* \neq z$, then we see that L must be the perpendicular bisector of the line segment $[z, z^*]$ and thus our new notion of symmetry does indeed agree with our original notion for lines.

To see that $z^* \neq z$ when $z \notin L$, choose $a \in L$ to be the point of L which is closest to z. Then $z-a$ is orthogonal to L (why?) and so the point $c = a + i(z-a)$ lies on L. The hypothesis that z, z^* be symmetric with respect to L now tells us that

$$\frac{z^*-a}{c-a} = \overline{\left(\frac{z-a}{c-a}\right)} = \overline{-i} = i\,,$$

so $z^* = a + i(c-a)$ whereas $z = a - i(c-a)$.

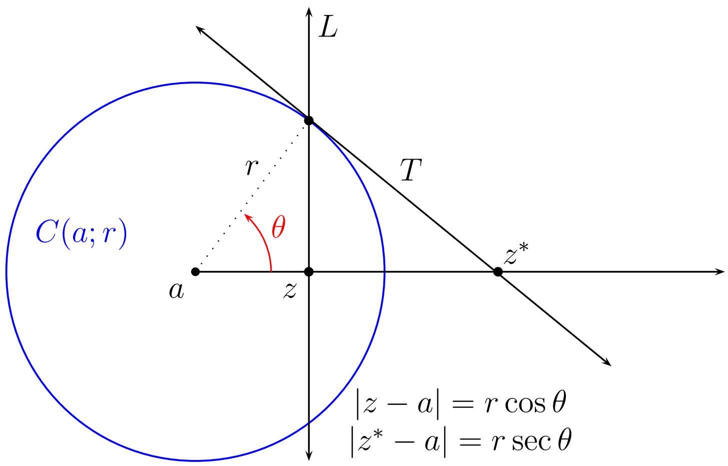

FIGURE 4. Constructing points z, z^* symmetric wrt $C(a;r)$

Next we examine the case when $K = C$ for some circle $C = C(a;r)$. In class we derived the formulas

$$\rho_K(z) := z^* = \frac{a\bar{z} + (r^2 - |a|^2)}{\bar{z} - \bar{a}} = a + r^2 \frac{z-a}{|z-a|^2}\,.$$

This tells us that z, z^* both lie on the same ray from a and that the product of their distances to a is r^2. In Figure 4 we see how to geometrically construct $z^* = \rho_K(z)$, given z and $K = C(a;r)$. We note that the center of a circle and the point ∞ are always symmetric with respect to the given circle.

6.F. **Self-maps of Disks and Half-Planes.** Applications of symmetry

There are many useful ways that symmetry can be employed to help with the Basic Problems (a), (b). Most directly, given $T, z, K, T(z), T(K)$, it is easy to determine $T(z^*)$ just by geometric considerations. To attack Basic Problem (a), we start with T, z, K and then calculate $T(z)$ and $T(z^*)$ which then provides information about $T(K)$. There are two approaches to handle Basic Problem (b). Given K and K', we can choose distinct $a, b, c \in K$ and distinct $a', b', c' \in K'$ and then solve $[Tz, a', b', c'] = [z, a, b, c]$ to find T with $T(K) = K'$. Alternatively, we can choose $a \in K, a' \in K', b \notin K, b' \notin K'$ and solve $[Tz, a', b', (b')^*] = [z, a, b, b^*]$.

Frankly, the best way to understand how to best use symmetry is to work lotsa problems!

Let's solve the problem raised in Exercise 6.5. It is easy to check that

$$T(z) := \frac{(2+i)z}{z + (1+i)}$$

maps $\hat{\mathbb{R}}$ onto a circle $C = C(a;r)$ that goes through the points $0 = T(0), 1 = T(1)$ and $2+i = T(\infty)$. Note that $-1-i, -1+i$ are symmetric with respect to $\hat{\mathbb{R}}$ and $\infty = T(-1-i)$. Thus, since a, ∞ are symmetric with respect to C, it follows that $a = T(-1+i) = (1+3i)/2$. Also, $r = |a| = \sqrt{10}/2$. :-)

It is a routine matter to map any disk or half-plane onto any other disk or half-plane. (NB open Euclidean disks and open Euclidean half-planes are both open chordal disks :-) The standard examples are the self-maps of $\mathbb{D}$ and of $i\mathbb{H}$.

Some missing topics: If we have time, later in the year, we may take yet another look at Möbius transformations. E.g., we may investigate their fixed points, and will examine the so-called Steiner and Apollonian circles.

Department of Mathematics, University of Cincinnati, OH 45221
E-mail address: David.Herron@math.UC.edu