Automorphisms of the Unit Disc

Richard Chapling[*]

v.1 March 2, 2015

We derive the automorphism group of the unit disc in the complex plane, using the method suggested by questions on the Part IB example sheets.

Let $D = D(0,1)$ be the unit disc, and $\mathbb{H} = \{x + iy \in \mathbb{C} : y > 0\}$. As usual, write $\mathrm{Aut}(D)$ for the automorphism group of D: the set of bijective conformal maps $D \to D$. (Conformality gives us that the inverses are also bijective.)

1 Möbius Transformations Preserving D

Lemma 1. *The Möbius transformations preserving D are precisely those of the form*

$$T(z) = \lambda \frac{z-a}{\bar{a}z-1}, \tag{1.1}$$

where $|\lambda| = 1$ and $|a| < 1$.

Let the set of Möbius transformations of this form be A.

Proof 1: Starting in a sensible form. A Möbius transformation is determined by its action on three points. In particular, the general Möbius transformation can be written in the form

$$S(z) = c\frac{z-a}{bz-1}, \tag{1.2}$$

where $c \neq 0$ and $ab \neq 1$ so that the determinant is nonzero. (Note that $S(a) = 0$, $S(1/b) = \infty$ and $S((ac-1)/(c-b)) = 1$, and (a,b,c) can be chosen so that these points can be any three in $\mathbb{C}$, so this is sufficiently general.) Now, to specialise to the form that we want, we have two conditions: $|S(0)| < 1$, and if $|z| = 1$, then $|S(z)| = 1$. The first implies that $|ca| < 1$. Now the second is more fiddly to apply:

$$\begin{aligned} 1 = |S(z)| &= |c| \frac{|z-a|}{|bz-1|} \\ (bz-1)(\bar{b}\bar{z}-1) &= |c|^2 (z-a)(\bar{z}-\bar{a}) \\ |b|^2 - 2\Re(bz) + 1 &= |c|^2 (1 - 2\Re(\bar{a}z) + |a|^2) \end{aligned}$$

Now, this has to be true for any unit z, so equating coefficients, we find

$$b = |c|^2 \bar{a}, \qquad 1 + |b|^2 = |c|^2 (1 + |a|^2)$$

This two equations imply that either $|c| = 1$ or $|ac| = 1$; since the latter contradicts $|ac| < 1$, the former must hold, and hence we find

$$b = \bar{a}, \qquad |c| = 1, \qquad |a| < 1,$$

which puts S in the form of T. It is easy to check that conversely this form is sufficient to map $D \to D$. □

[*]Trinity College, Cambridge

Proof 2: The group property. First we check that transformations of type T form a group. Since the whole group of Möbius transformations is a group, it suffices to check that if S, T are in the subgroup, so is ST^{-1}. A calculation shows that

$$S(z) = \lambda\frac{z-a}{\bar{a}z-1}, \quad T(z) = \mu\frac{z-b}{\bar{b}z-1} \qquad \Longrightarrow \qquad ST^{-1}(z) = \nu\frac{z-c}{\bar{c}z-1},$$

where

$$\nu = \frac{\lambda}{\mu}\frac{a\bar{b}-1}{1-b\bar{a}}, \qquad c = \frac{(b-a)\mu}{1-a\bar{b}};$$

$|\nu| = 1$ is obvious, and if $|c| \geqslant 1$, then

$$\begin{aligned}\left|1-a\bar{b}\right|^2 &\geqslant |b-a|^2 \\ 1 - 2\Re(a\bar{b}) + |a|^2|b|^2 &\geqslant |b|^2 + |a|^2 - 2\Re(a\bar{b}) \\ (|a|^2-1)(|b|^2-1) &\geqslant 0,\end{aligned}$$

contradicting the assumptions on $|a|$ and $|b|$. Therefore such maps form a group.

Now, we need to show that these Möbius transformations are the most general possible that preserve D. In particular, let $S(z)$ be a Möbius transformation preserving D, with $S(0) = b$ and $S(1) = \mu$. Then with $R(z) = (z-b)/(\bar{b}z-1)$ and $Q(z) = z(\bar{b}\mu-1)/(\mu-b)$, then $s = Q \circ R \circ S$ has $s(0) = 0$ and $s(1) = 1$. A Möbius transformation that fixes 0 and 1 is of the form $(1-c)z/(1-cz)$, and the only one of these that maps -1, say, to a point with unit modulus is with $c = 0$, which is the identity. So $s(z) = z$, and hence $S = Q^{-1} \circ R^{-1}$, which is easily checked to be of the required form. □

2 Schwarz's Lemma

Lemma 2. *Let $f(z) : D \to D$ with $f(0) = 0$. Then for every $z \in D$, $|f(z)| \leqslant |z|$, and $|f'(0)| \leqslant 1$. Moreover, if equality occurs, for some $w \neq 0$ in the first, or in the second, then $f(z) = cz$ for some c with $|c| = 1$.*

Proof. Let $g(z) = f(z)/z$ for $z \neq 0$. As $g(z) \to f'(0)$ as $z \to 0$, g can be analytically continued by setting $g(0) = f'(0)$. Then g is analytic on $D(0, r)$ for every r with $0 < r < 1$, and the maximum modulus principle implies that

$$\sup_{|z|<r} |g(z)| = \sup_{|z|=r} |g(z)| = \frac{1}{r}\sup_{|z|=r} |f(z)| \leqslant \frac{1}{r}$$

for any $r < 1$. This tends to 1 as $r \to 1$, so $|g(z)| \leqslant 1$ for all $z \in D$, and hence the first part of the result. Either of the second statements holding implies that there is $w \in D$ with $|g(w)| = 1$, so $|g(z)|$ has an interior local maximum, and hence is constant, again by the maximum modulus principle. □

3 The Automorphism Group

Proposition 3. $A = \mathrm{Aut}(D)$

Proof. Let $f : D \to D$ be an automorphism, with $f(0) = a$. Then if

$$S(z) = \frac{z-a}{\bar{a}z-1},$$

then $h = S \circ f : D \to D$ is also an automorphism, and $h(0) = 0$. Now, h satisfies the conditions of Schwarz's Lemma, and hence

$$|h(z)| \leqslant |z|. \tag{3.1}$$

On the other hand h^{-1} also satisfies the conditions of Schwarz's Lemma, and so

$$|z| = \left|h^{-1}(h(z))\right| \leqslant |h(z)|, \tag{3.2}$$

and (3.1) and (3.2) together give $|h(z)| = |z|$ for all $z \in D$, and so $h(z)$ must be a rotation, $h(z) = \lambda z$ for some λ with $|\lambda| = 1$. But then $f = S^{-1} \circ h$ is a composition of two elements of A. Since A is a group, $f \in A$. This proves $\mathrm{Aut}(D) \subseteq A$; the reverse inclusion, $A \subseteq \mathrm{Aut}(D)$, was checked in Lemma 1. □

4 Buy One Automorphism Group, Get One Free: The Automorphism Group of the Upper Half-Plane

The Möbius transformation

$$\phi(z) = i\frac{1-z}{1+z} \tag{4.1}$$

is a conformal equivalence $\phi : D \to \mathbb{H}$. Hence if $f \in \operatorname{Aut}(\mathbb{H})$, then $\hat{f} = \phi^{-1} \circ f \circ \phi : D \to D$ is a conformal bijection, so $\hat{f} \in \operatorname{Aut}(D)$. Since $\operatorname{Aut}(D) = A$, $\hat{f}$ is therefore a Möbius transformation. But then $f = \phi \circ \hat{f} \circ \phi^{-1}$ is also a Möbius transformation, so $\operatorname{Aut}(D) \subseteq \mathcal{M}$.

Corollary 4. $\operatorname{Aut}(\mathbb{H})$ *is the set of Möbius transformations of the form*

$$T(z) = \frac{az+b}{cz+d}, \qquad \{a,b,c,d\} \in \mathbb{R}, \quad ad - bc > 0. \tag{4.2}$$

Proof. Let the set of such transformations be B. We just showed that $\operatorname{Aut}(\mathbb{H}) \subseteq \mathcal{M}$. Take an $S \in \mathcal{M}$, $S(z) = (az+b)/(cz+d)$. S is a bijection $\mathbb{H} \to \mathbb{H}$, if and only if it preserves the extended real axis $\mathbb{R} \cup \{\infty\}$; it is easy to see that this can occur if and only if all of $\{a,b,c,d\}$ are real. Now, the next requirement is that the upper half-plane $\mathbb{H}$ is mapped to itself (and not the lower half-plane). The simplest way to check this is that $\Im(S(i)) > 0$:

$$S(i) = \frac{ai+b}{ci+d} = \frac{(ai+b)(-ci+d)}{c^2+d^2} = \frac{(ac+bd)+i(ad-bc)}{c^2+d^2},$$

and hence we require $ad - bc > 0$; therefore if $S : \mathbb{H} \to \mathbb{H}$, $S \in B$, and so $\operatorname{Aut}(\mathbb{H}) \subseteq B$. Continuity of Möbius transformations on the Riemann sphere then implies that $S \in B$ really does map $\mathbb{H} \to \mathbb{H}$ bijectively, and so any $S \in B$ is an automorphism of $\mathbb{H}$; hence $B \subseteq \operatorname{Aut}(\mathbb{H})$, and so $\operatorname{Aut}(\mathbb{H}) = B$. □

5 Aside: Why Are Möbius Transformations Like Matrices?

One of the more puzzling aspects of $\mathcal{M}$, the set of Möbius transformations, is that we are told that we can represent

$$z \mapsto \frac{az+b}{cz+d} \qquad \text{by} \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The purpose of this section is to explain that this is not really mysterious, once you have a gap in your knowledge filled.

5.1 The Complex Projective Line

Definition 5. The *complex projective line*, $\mathbb{CP}^1$, is defined as the set

$$(\mathbb{C}^2 \setminus \{(0,0)\})/\sim\,,$$

where $\sim$ is an equivalence relation given by $(z_0, z_1) \sim (w_0, w_1)$ when $z_0 w_1 - z_1 w_0 = 0$.

There are a number of ways to think about this, including lines through the origin in $\mathbb{C}^2$, a sphere in three dimensions, and the one that we are actually interested in: this is equivalent to the Riemann sphere.

Why? Suppose we look at points of the form $(z, 1) \in \mathbb{CP}^1$. Then z can be any complex number we like. Therefore the set $\{(z,1) : z \in \mathbb{C}$ is isomorphic to $\mathbb{C}$. Also, for any $z_1 \neq 0$, we have $(z_0, z_1) \sim (z_0/z_1, 1)$, so in fact any point in $\mathbb{CP}^1$ with its second coordinate nonzero can be thought of as within the complex plane. On the other hand, suppose $z_1 = 0$. Then $(z_0, 0) \sim (1, 0)$, so there is only one point of this form, and this becomes the point at infinity, ∞. Therefore we have the decomposition

$$\mathbb{CP}^1 = \mathbb{C} \cup \{\infty\}.$$

Okay, that looks like the Riemann sphere, but how do we know that it has the same topology? Recall that the topology of the Riemann sphere is given by taking a basis made out of the balls in the plane, but adjoining sets of the form $\{z \in \mathbb{C} : |z| > r\} \cup \{\infty\}$, for all $r > 0$.

On the other hand, $\mathbb{CP}^1$ inherits the topology of $\mathbb{C}^2$, but with equivalent points identified. The part of this we actually care about is that if we have a ball that does not contain the point at infinity, it is exactly a ball in the $(z, 1)$ complex plane. On the other hand, a ball centred at $(1, 0)$ is of the form $\{(1, z) \in \mathbb{C}^2 : |z| < \varepsilon\}$, which is equivalent to $\{(w, 1) \in \mathbb{C}^2 : |w| > 1/\varepsilon\} \cup \{(1, 0)\}$. It should be clear that these are of the same form as the open sets we add to $\mathbb{C}$ to get the Riemann sphere.

If that wasn't clear, think of it this way: convergence to a limit in $\mathbb{C}$ works as before. On the other hand, convergence to ∞ means that $|z|$ becomes larger and larger. In $\mathbb{CP}^1$, this means that the z in $(z, 1)$ gets larger and larger. But this point is equivalent to $(1, 1/z)$, which we can see approaches $(1, 0)$, which we identified with the point ∞.[1]

5.2 Back to Möbius Transformations

That's all very well, but what does this have to do with Möbius transformations? The linear transformations of $\mathbb{C}^2$ are obviously given by matrices

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

But suppose I consider acting on a point (z_0, z_1) where $z_1 \neq 0$ so $(z_0, z_1) \sim (z, 1)$, and suppose it maps to a point where $cz_0 + dz_1$ is not zero. Then

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} (z, 1) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \left(\frac{z_0}{z_1}, 1 \right) \\ &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} (z_0, z_1) = (az_0 + bz_1, cz_0 + dz_1) = \left(\frac{az_0 + bz_1}{cz_0 + bz_1}, 1 \right) = \left(\frac{az + b}{cz + d}, 1 \right), \end{aligned}$$

which is the same answer as the Möbius transformation acting on z! It follows that, on the complex plane part of $\mathbb{CP}^1$, Möbius transformations and the "projective linear transformations" given by complex 2×2 matrices are identical.

It is left as an exercise to the reader to fiddle through the two other cases: the algebra works in exactly the same way, and you will find that the continuity method we normally use to evaluate at ∞ falls out of this formalism entirely naturally.

[1] Admittedly I have (at least partially deliberately) blurred the distinction between the copy of $\mathbb{C}$ we can identify in $\mathbb{CP}^1$ and $\mathbb{C}$ itself, but hopefully the main idea is clear.