CMFT

Conjugacy Classification
of Quaternionic Möbius Transformations

John R. Parker and Ian Short

(Communicated by Fred Brackx)

Abstract. It is well known that the dynamics and conjugacy class of a complex Möbius transformation can be determined from a simple rational function of the coefficients of the transformation. We study the group of quaternionic Möbius transformations and identify simple rational functions of the coefficients of the transformations that determine dynamics and conjugacy.

Keywords. Möbius transformations, quaternions, conjugacy, dynamics.

2000 MSC. 15A33, 51B10.

1. Introduction

The motivation behind this paper is the classical theory of complex Möbius transformations. A complex Möbius transformation f is a conformal map of the extended complex plane of the form $f(z) = (az + b)(cz + d)^{-1}$, where a, b, c and d are complex numbers such that the quantity $\sigma = ad - bc$ is not zero (to avoid ambiguity, we sometimes denote this quantity by σ_f). There is a homomorphism from $\mathrm{SL}(2, \mathbb{C})$ to the group of complex Möbius transformations that takes the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ to the map f. This homomorphism is surjective because the coefficients of f can always be scaled so that $\sigma = 1$. The collection of all complex Möbius transformations for which σ takes the value 1 forms a group which can be identified with $\mathrm{PSL}(2, \mathbb{C})$. A member f of $\mathrm{PSL}(2, \mathbb{C})$ is *simple* if it is conjugate in $\mathrm{PSL}(2, \mathbb{C})$ to an element of $\mathrm{PSL}(2, \mathbb{R})$. The map f is *k-simple* if it may be expressed as the composite of k simple transformations but no fewer. Let $\tau = a + d$ (we denote this quantity by τ_f where necessary). Conjugacy and dynamics in $\mathrm{PSL}(2, \mathbb{C})$ are determined by the next theorem (see [1, Thms. 4.3.1, 4.3.4]).

Theorem 1.1. *Two non-trivial maps f and g in $\mathrm{PSL}(2, \mathbb{C})$ are conjugate if and only if ${\tau_f}^2 = {\tau_g}^2$. Moreover, given $f \in \mathrm{PSL}(2, \mathbb{C})$,*

(a) *if $\tau^2 \geq 0$ then f is 1-simple, and*
  (i) *if $0 \leq \tau^2 < 4$ then f is elliptic;*
  (ii) *if $\tau^2 = 4$ then f is parabolic;*
  (iii) *if $\tau^2 > 4$ then f is loxodromic.*
(b) *If τ^2 is either not real or is negative then f is 2-simple and loxodromic.*

The purpose of this paper is to generalise Theorem 1.1 to quaternionic Möbius transformations.

We denote the non-commutative field of quaternions by $\mathbb{H}$, and we let $\mathbb{H}_\infty$ denote $\mathbb{H} \cup \{\infty\}$. To each 2×2 matrix

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

with quaternion entries we associate the quantity

$$\alpha = |a|^2|d|^2 + |b|^2|c|^2 - 2\,\mathrm{Re}[a\bar{c}d\bar{b}].$$

According to [14, Thm. 1], the matrix A is invertible if and only if $\alpha \neq 0$, and we have a group

$$\mathrm{SL}(2, \mathbb{H}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{H}, \ \alpha = 1 \right\}.$$

A *quaternionic Möbius transformation* f is an invertible map of $\mathbb{H}_\infty$ of the form

$$f(z) = (az + b)(cz + d)^{-1},$$

for $a, b, c, d \in \mathbb{H}$, subject to the usual conventions regarding the point ∞ (see [2, 8, 14]). The map $A \mapsto f$ from $\mathrm{SL}(2, \mathbb{H})$ to the group of quaternionic Möbius transformations is a surjective homomorphism and its kernel is the group of non-zero real scalar multiples of the identity matrix. Therefore we can identify the group of quaternionic Möbius transformations with $\mathrm{PSL}(2, \mathbb{H}) = \mathrm{SL}(2, \mathbb{H})/\{\pm I\}$. In future, whenever we refer to A, f or a 'quaternionic Möbius transformation', we refer to the quantities described above

For a 2×2 quaternionic matrix A we define the following three quantities:

$$\beta = \mathrm{Re}[(ad - bc)\bar{a} + (da - cb)\bar{d}],$$
$$\gamma = |a + d|^2 + 2\,\mathrm{Re}[ad - bc],$$
$$\delta = \mathrm{Re}[a + d].$$

We sometimes include subscripts in writing these quantities — for example, we write α_A, β_A, γ_A and δ_A — to make it clear to which matrix they correspond. In [6, Thm. 4.1] Foreman proves that the functions β, γ and δ are constant on conjugacy classes in $\mathrm{SL}(2, \mathbb{H})$ (although there is an error in Foreman's definition of β). When we speak of quantities α_f, β_f, γ_f and δ_f, for a map f in $\mathrm{PSL}(2, \mathbb{H})$,

we refer to α_A, β_A, γ_A and δ_A, where A is one of the two matrices in $\mathrm{SL}(2, \mathbb{H})$ associated to f. Since both A and $-A$ are associated to f, it is ambiguous to speak of β_f and δ_f. We can, however, refer to α_f, β_f^2, $\beta_f \delta_f$ and so forth, without ambiguity.

We define two more quantities for each $A \in \mathrm{SL}(2, \mathbb{H})$ as follows:

$$
\sigma = \begin{cases}
cac^{-1}d - cb & \text{when } c \neq 0, \\
bdb^{-1}a & \text{when } c = 0, \ b \neq 0, \\
(d-a)a(d-a)^{-1}d & \text{when } b = c = 0, \ a \neq d, \\
a\bar{a} & \text{when } b = c = 0, \ a = d,
\end{cases}
$$

$$
\tau = \begin{cases}
cac^{-1} + d & \text{when } c \neq 0, \\
bdb^{-1} + a & \text{when } c = 0, \ b \neq 0, \\
(d-a)a(d-a)^{-1} + d & \text{when } b = c = 0, \ a \neq d, \\
a + \bar{a} & \text{when } b = c = 0, \ a = d.
\end{cases}
$$

Observe that in each case $|\sigma|^2 = \alpha = 1$. The quantities σ and τ take the role of the quaternionic determinant and trace in $\mathrm{SL}(2, \mathbb{H})$. They are not conjugation invariant.

A *real Möbius transformation* in $\mathrm{PSL}(2, \mathbb{H})$ is a member of $\mathrm{PSL}(2, \mathbb{H})$ with real coefficients. For such transformations, σ coincides with the usual determinant $ad - bc$. Therefore the group of real Möbius transformations consists of the disjoint union of the subgroup of real Möbius transformations with $\sigma = 1$, namely $\mathrm{PSL}(2, \mathbb{R})$, and the coset of transformations with $\sigma = -1$. A member f of $\mathrm{PSL}(2, \mathbb{H})$ is *simple* if it is conjugate in $\mathrm{PSL}(2, \mathbb{H})$ to an element of $\mathrm{PSL}(2, \mathbb{R})$. The map f is *k-simple* if it may be expressed as the composite of k simple transformations but no fewer.

Theorem 1.2. *A quaternionic Möbius transformation is conjugate to a real Möbius transformation if and only if σ and τ are both real.*

We prove in Proposition 4.3 that if σ and τ are real then they are preserved under conjugation in $\mathrm{SL}(2, \mathbb{H})$. It follows from Theorem 1.2 that, as usual, the quantity τ^2/σ can be used to distinguish conjugacy and dynamics amongst those maps conjugate to real Möbius transformations. In contrast to Theorem 1.2, all quaternionic Möbius transformations are conjugate to Möbius transformations with complex coefficients.

Our second result is about how to determine whether two quaternionic Möbius transformations are conjugate.

Theorem 1.3. *Two non-trivial quaternionic Möbius transformations f and g are conjugate if and only if the following two conditions hold:*

(i) *either both of them or neither of them are conjugate to real Möbius transformations;*

(ii) $\beta_f \delta_f = \beta_g \delta_g$, $\gamma_f = \gamma_g$ and $\delta_f{}^2 = \delta_g{}^2$.

Möbius transformations in any dimension can be described as elliptic, loxodromic or parabolic depending on their dynamics (see [13, p. 142] or Section 3 below). Quaternionic Möbius transformations are either 1-*simple*, 2-*simple* or 3-*simple*. In fact, there are no 3-simple elliptic or parabolic maps, therefore we have partitioned the members of $\mathrm{PSL}(2, \mathbb{H})$, excluding the identity, into seven categories. In Theorem 1.4, below, we describe how one can easily identify the category in which a quaternionic Möbius transformation belongs from its coefficients. Theorem 1.4 should be compared to [4, Thm. 1.3], [3, Thm. 1.3] and [6, Prop. 4.2], although it is stronger and, in some cases, easier to apply than those theorems.

Theorem 1.4. *Let f be a quaternionic Möbius transformation with $\alpha = 1$.*

(a) *If $\sigma = 1$ and $\tau \in \mathbb{R}$ then f is 1-simple, $\beta = \delta$, $\gamma = \delta^2 + 2$ and the following trichotomy holds.*
    (i) *If $0 \le \delta^2 < 4$ then f is elliptic.*
    (ii) *If $\delta^2 = 4$ then f is parabolic.*
    (iii) *If $\delta^2 > 4$ then f is loxodromic.*
(b) *If $\beta = \delta$ and either $\tau \notin \mathbb{R}$ or $\sigma \ne 1$, then f is 2-simple and the following trichotomy holds.*
    (i) *If $\gamma - \delta^2 < 2$ then f is elliptic.*
    (ii) *If $\gamma - \delta^2 = 2$ then f is parabolic.*
    (iii) *If $\gamma - \delta^2 > 2$ then f is loxodromic.*
(c) *If $\beta \ne \delta$ then f is 3-simple loxodromic.*

To determine the dynamics and conjugacy class of a Möbius transformation with quaternion coefficients for which $\alpha \ne 1$, one should replace the coefficients of the transformation with suitably scaled alternatives so that $\alpha = 1$, before applying Theorem 1.3 and Theorem 1.4. We could have stated both theorems in a more general context with α assuming any non-zero value (and the reader can easily derive such theorems), but the exposition is simplified by assuming that α is equal to 1 throughout.

The structure of the remainder of this document is as follows. In §2 we review the work of other authors on topics similar to our own. In §3 we discuss background material in quaternionic linear algebra and the geometry of Möbius transformations. Then in §4, §5 and §6 we prove Theorems 1.2, 1.3 and 1.4.

2. Literature review

Quaternionic Möbius transformations and matrices were originally studied by people such as Brenner [2], Coxeter [5], Gormley [8] and Lee [12]. Later Wilker [14] proved that $\mathrm{SL}(2, \mathbb{H})$ is the double cover of $\mathrm{PSL}(2, \mathbb{H})$ and that $\mathrm{PSL}(2, \mathbb{H})$ is isomorphic to $\mathrm{SO}^+(5, 1)$. He defines the quantity α_f (which he calls Δ^2)

and σ_f (which he calls L_{21}). Wilker does not examine the relationship between conjugacy and the coefficients of quaternionic Möbius transformations. Several papers concerned with dynamics in $PSL(2, \mathbb{H})$ have appeared recently. We now discuss them.

In [3] Cao, Parker and Wang consider the subgroup of quaternionic Möbius transformations preserving the unit ball in $\mathbb{H}$. Using the methods of Huang and So [9] they calculate the fixed points of such quaternionic Möbius transformations. They classify these transformations as elliptic, parabolic or loxodromic and 1-simple or 2-simple using an invariant which, in our notation, is equal to $\gamma - \delta^2 - 2$. Similar work was carried out independently by Kido in [11], who also showed that the quantity $\gamma - \delta^2 - 2$ is conjugation invariant.

The results of [6] come closest to Theorem 1.4. Foreman uses the quantities β, γ and δ (which he calls F_1, F_2 and F_3), and in [6, Thm. 4.1] he proves that they are conjugation invariant. On [6, p. 29], however, he defines a quaternionic Möbius transformation to be loxodromic if and only if it has exactly two fixed points in $\mathbb{H}_\infty$. But 2-simple elliptic transformations have exactly two fixed points, so Foreman's classification disagrees with standard definitions such as [13, p. 141]. In consequence, his main classification result [6, Prop. 4.2] does not distinguish between 2-simple elliptic maps and 2-simple or 3-simple loxodromic maps.

Later Cao, apparently unaware of Foreman's work, provided another classification of dynamics in [4]. Cao did not have the full set of Foreman's conjugacy invariant quantities at his disposal, so his approach is more complicated.

In [7] Gongopadhyay and Kulkarni consider the general problem of giving a dynamical classification of hyperbolic isometries in all dimensions. Quaternionic Möbius transformations can be identified with the group of five dimensional orientation preserving hyperbolic isometries. Gongopadhyay and Kulkarni, like Foreman, embed $SL(2, \mathbb{H})$ in $SL(4, \mathbb{C})$ to obtain a classification of dynamics. Their results are stated in terms of the coefficients of matrices in $SL(4, \mathbb{C})$, and are similar to Foreman's results.

In summary, there have been results similar to, although not as efficient as, Theorem 1.4 (particularly [6, Propo 4.2]). Theorem 1.2 and Theorem 1.3 are, to our knowledge, entirely new.

3. Background material on quaternionic matrices

We review some basic linear algebra in $SL(2, \mathbb{H})$ which can be found in [2, 3, 9, 12, 14], and relate the quantities α, β, γ, δ, σ and τ to eigenvalues of quaternionic matrices. We denote the real part of a quaternion z by $\text{Re}[z]$. Two quaternions z and w are *similar* if there is a third quaternion u such that $uzu^{-1} = w$. It is well known, and easy to prove, that z and w are similar if and only if $|z| = |w|$ and $\text{Re}[z] = \text{Re}[w]$.

A quaternion t is a *right eigenvalue* of a matrix A in $\mathrm{SL}(2, \mathbb{H})$ if and only if there is a non-zero column vector $\mathbf{v} \in \mathbb{H}^2$ such that $A\mathbf{v} = \mathbf{v}t$. For $u \neq 0$, let $\mathbf{w} = \mathbf{v}u^{-1}$. Then we see from the equation $A\mathbf{w} = \mathbf{w}utu^{-1}$ that all quaternions similar to t are also right eigenvalues of A. For $B \in \mathrm{SL}(2, \mathbb{H})$, we see from the equation $BAB^{-1}(B\mathbf{v}) = B\mathbf{v}t$ that conjugate matrices have the same right eigenvalues. According to [2, Thm. 2], each quaternionic matrix is conjugate to an upper triangular matrix. This statement is equivalent to the well known fact that each (quaternionic) Möbius transformation is conjugate to a Möbius transformation that fixes ∞. The diagonal entries t_1 and t_2 of an upper triangular matrix are both representatives from the right eigenvalue similarity classes of the matrix. Recall that the quantities $|t_1|^2$, $|t_2|^2$, $\mathrm{Re}[t_1]$ and $\mathrm{Re}[t_2]$ are independent of the particular representatives t_1 and t_2 of the right eigenvalue similarity classes. Substituting $c = 0$, $a = t_1$ and $d = t_2$ into the equations for α, β, γ and δ from §1, we find that

$$\alpha = |t_1|^2 \, |t_2|^2,$$
$$\beta = |t_1|^2 \, \mathrm{Re}[t_2] + |t_2|^2 \, \mathrm{Re}[t_1],$$
$$\gamma = |t_1|^2 + |t_2|^2 + 4 \, \mathrm{Re}[t_1] \, \mathrm{Re}[t_2],$$
$$\delta = \mathrm{Re}[t_1] + \mathrm{Re}[t_2].$$

It is easy to use these identities to show that all quaternions t similar to either t_1 or t_2 satisfy

$$t^4 - 2\delta t^3 + \gamma t^2 - 2\beta t + \alpha = 0.$$

Thus we have expressions for α, β, γ and δ in terms of the functions $|t_1|^2$, $|t_2|^2$, $\mathrm{Re}[t_1]$ and $\mathrm{Re}[t_2]$. Conversely, given α, β, γ and δ we can determine $|t_1|^2$, $|t_2|^2$, $\mathrm{Re}[t_1]$ and $\mathrm{Re}[t_2]$ using Theorem 3.1, below (compare this to [9, Lem. 2.2]).

Theorem 3.1. *Suppose that f is a quaternionic Möbius transformation with $\alpha = 1$. Let X denote the largest real root of the cubic polynomial*

$$q(x) = 2x^3 - \gamma x^2 + 2(\beta\delta - 1)x + (\gamma - \beta^2 - \delta^2).$$

We let t represent one of the right eigenvalues t_1 or t_2.

  (i) *If $X\delta \geq \beta$ then $|t|^2 = X \pm \sqrt{X^2 - 1}$ and $2\,\mathrm{Re}[t] = \delta \pm \sqrt{2X - \gamma + \delta^2}$.*
  (ii) *If $X\delta < \beta$ then $|t|^2 = X \pm \sqrt{X^2 - 1}$ and $2\,\mathrm{Re}[t] = \delta \mp \sqrt{2X - \gamma + \delta^2}$.*

Notice that $q(1) = -(\beta - \delta)^2 \leq 0$; hence $X^2 \geq 1$. Also,

$$q\left(\frac{\gamma - \delta^2}{2}\right) = -\frac{1}{4}(2\beta - \gamma\delta + \delta^3)^2 \leq 0;$$

hence $2X \geq \gamma - \delta^2$. Therefore the square roots in Theorem 3.1 are real.

Proof. We make use of the expressions relating α, β, γ and δ to t_1 and t_2. One can check that the roots of q are

$$\mathrm{Re}[t_1]\,\mathrm{Re}[t_2] \pm \sqrt{(|t_1|^2 - \mathrm{Re}[t_1]^2)\,(|t_2|^2 - \mathrm{Re}[t_2]^2)}, \qquad \frac{1}{2}(|t_1|^2 + |t_2|^2).$$

Since
$$\frac{1}{2}(|t_1|^2 + |t_2|^2) \geq |t_1|\,|t_2| \geq \mathrm{Re}[t_1]\,\mathrm{Re}[t_2] \pm \sqrt{(|t_1|^2 - \mathrm{Re}[t_1]^2)\,(|t_2|^2 - \mathrm{Re}[t_2]^2)}$$

we see that the largest real root X is the third of these numbers. Using the equation

$$\left(|t_1|^2 - |t_2|^2\right)\left(\mathrm{Re}[t_1] - \mathrm{Re}[t_2]\right) = 2(X\delta - \beta)$$

we can verify the details of Theorem 3.1. $\blacksquare$

Theorem 3.1 alone is insufficient to enable us to determine whether two matrices in $\mathrm{SL}(2, \mathbb{H})$ are conjugate, nor is it sufficient to enable us to find fixed points of quaternionic Möbius transformations. For example, the matrices

$$\begin{pmatrix} i & i \\ 0 & i \end{pmatrix}, \qquad \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$$

have the same right eigenvalues, and each of β, γ and δ takes the same value on each matrix, but they are not conjugate. This is why it is necessary to introduce quantities such as σ and τ described in Section 1. Both σ and τ arise in the search for fixed points of quaternionic Möbius transformations, in a manner described below.

We maintain the notation introduced in Section 1 for a matrix A in $\mathrm{SL}(2, \mathbb{H})$ and a corresponding transformation f in $\mathrm{PSL}(2, \mathbb{H})$. Suppose that $c \neq 0$. Given a fixed point v of f, let $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $t = cv + d$, so that $A\mathbf{v} = \mathbf{v}t$. Conversely, if A has a right eigenvector $\mathbf{v}$ with right eigenvalue t, then after conjugating t suitably we can assume that $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $t = cv + d$ and v is a fixed point of f. Thus fixed points of f correspond to similarity classes of right eigenvalues of A. Possible values of t are solutions to the quadratic equation

$$t^2 - (cac^{-1} + d)t + (cac^{-1}d - cb) = 0.$$

Similarly, when $b \neq 0$, the fixed point v corresponds to the vector $\mathbf{v} = \begin{pmatrix} 1 \\ v^{-1} \end{pmatrix}$ (unless $v = \infty$, in which case v corresponds to $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$). The right eigenvalue associated to $\mathbf{v}$ is $t = a + bv^{-1}$ and it satisfies

$$t^2 - (bdb^{-1} + a)t + (bdb^{-1}a - bc) = 0.$$

When both b and c are non-zero the solutions to these two quadratic equations are conjugate since $a + bv^{-1} = v(cv + d)v^{-1}$. In all cases (including the simple case $b = c = 0$ not considered) the right eigenvalues are solutions to

(3.1) $$t^2 - \tau t + \sigma = 0.$$

We also have $t^2 - 2\,\mathrm{Re}(t)t + |t|^2 = 0$. Following Niven (see [3, 9]), we combine this quadratic equation with (3.1) to obtain, provided $\tau \neq 2\,\mathrm{Re}(t)$,

$$t = (\tau - 2\,\mathrm{Re}(t))^{-1}\left(\sigma - |t|^2\right)$$

Since we know how to find $|t|^2$ and $\mathrm{Re}(t)$ in terms of β, γ and δ (using Theorem 3.1), we can find the similarity class of t. If $\tau = 2\,\mathrm{Re}(t)$, and so $\sigma = |t|^2$, then we can solve (3.1) directly (compare [3, Thm. 1.5]).

Although σ and τ cannot be calculated from β, γ and δ, the converse is true:

$$\alpha = |\sigma|^2, \quad \beta = \mathrm{Re}[\sigma\overline{\tau}], \quad \gamma = |\tau|^2 + 2\,\mathrm{Re}[\sigma], \quad \delta = \mathrm{Re}[\tau].$$

These equations can be verified by expanding out the terms.

We can also describe the dynamics of a quaternionic Möbius transformation f in terms of eigenvalues and fixed points. Let t_1 and t_2 be the eigenvalues of a lift of f in the group $\mathrm{SL}(2, \mathbb{H})$. Then f is *elliptic* if $|t_1| = |t_2|$ and f has at least two fixed points in $\mathbb{H}_\infty$; *parabolic* if $|t_1| = |t_2|$ and f fixes exactly one points in $\mathbb{H}_\infty$; *loxodromic* if $|t_1| \neq |t_2|$. (A geometric description of elliptic, parabolic and loxodromic can be found on [13, p. 141].) If the eigenvalues t_1 and t_2 are not similar, then A may be conjugated to a diagonal matrix with entries t_1 and t_2 on the diagonal. When all eigenvalues of A are similar we may or may not be able to diagonalise A. We can certainly (following Brenner, [2]) conjugate A to a triangular matrix whose diagonal entries will be similar. We now give a criterion which describes when such matrices A can be diagonalised.

Lemma 3.2. *Let $f(z) = (az + b)d^{-1}$ where a and d are similar and not real. Then f has a fixed point in $\mathbb{H}$ if and only if $bd = \overline{a}b$. Moreover, if $bd = \overline{a}b$ then f is conjugate to $g_0(z) = aza^{-1}$ and if $bd \neq \overline{a}b$ then f is conjugate to $g_1(z) = (az + 1)a^{-1}$.*

Proof. Suppose that there exists $v \in \mathbb{H}$ such that $f(v) = v$. That is, $av + b = vd$. Then using $|a|^2 = |d|^2$ and $a + \overline{a} = d + \overline{d}$ we have

$$\begin{aligned}
bd - \overline{a}b &= (vd - av)d - \overline{a}(vd - av) \\
&= vd^2 + |a|^2 v - (a + \overline{a})vd \\
&= vd(d + \overline{d}) - (d + \overline{d})vd \\
&= 0.
\end{aligned}$$

Conversely, assume that $bd = \overline{a}b$. Then set $v = (\overline{a} - a)^{-1}b = (a - \overline{a})b/|a - \overline{a}|^2$, which is defined since we supposed that a is not real. Using $\overline{a}bd^{-1} = b$, we have

$$f(v) = \big(a(\overline{a} - a)^{-1}b + b\big)d^{-1} = (\overline{a} - a)^{-1}(a + \overline{a} - a)bd^{-1} = (\overline{a} - a)^{-1}b = v.$$

For the second part, if f has a fixed point in $\mathbb{H}$ then, by conjugation, we may assume that this fixed point is 0. In other words, we may assume that $b = 0$. Since a and d are similar, we can apply a second conjugation by a map of the form $z \mapsto uz$ to obtain g_0. If f does not have a fixed point in $\mathbb{H}$ then $ba - \overline{a}b \neq 0$. In this case we can check that $(ba - \overline{a}b)$ commutes with a. Conjugate f by $h(z) = \big((a - \overline{a})z + b\big)(ba - \overline{a}b)^{-1}$ to obtain g_1. ∎

We will see later, in Lemma 4.2, that the condition $bd = \overline{a}b$ is equivalent to σ and τ both being real.

4. Proof of Theorem 1.2

Lemma 4.1. *Let A be in $\mathrm{SL}(2, \mathbb{H})$ and suppose that $b \neq 0$ and $c \neq 0$. If $\tau = cac^{-1} + d$ and $\sigma = cac^{-1}d - cb$ are both real then $bdb^{-1} + a = \tau$ and $bdb^{-1}a - bc = \sigma$.*

Proof. Suppose that σ and τ are both real. If $d = 0$ then $\sigma = cb = bc$ and $\tau = cac^{-1} = a$. Suppose $d \neq 0$. Notice that

$$cb = cac^{-1}d - \sigma = \tau d - d^2 - \sigma.$$

Therefore cb commutes with d. Then $bdb^{-1} = c^{-1}(cb)d(cb)^{-1}c = c^{-1}dc$. Hence

$$bdb^{-1} + a = c^{-1}dc + a = c^{-1}\tau c = \tau,$$
$$bdb^{-1}a - bc = c^{-1}dca - c^{-1}(cb)c = c^{-1}d\sigma d^{-1}c = \sigma.$$

$\blacksquare$

Lemma 4.2. *Let A be in $\mathrm{SL}(2, \mathbb{H})$ and suppose that either $b = 0$ or $c = 0$. Then σ and τ are both real if and only if either a and d are both real or else a and d are similar and $\sigma = |a|^2$ and $\tau = 2\,\mathrm{Re}[a]$.*

Proof. Suppose that σ and τ are both real, but a and d are not both real. We claim that each of the following statements holds:

  (i) if $b = 0$ and $c \neq 0$ then $cac^{-1} = \bar{d}$;
 (ii) if $b \neq 0$ and $c = 0$ then $bdb^{-1} = \bar{a}$;
(iii) if $b = c = 0$ and $a \neq d$ then $(d - a)a(d - a)^{-1} = \bar{d}$.

In each case, and in the remaining case in which $b = c = 0$ and $a = d$, a is similar to d, and $\sigma = |a|^2$ and $\tau = 2\mathrm{Re}[a]$.

For case (i), first observe that d is not real; for if d were real then a would also be real. Since $\sigma = cac^{-1}d$ we have $\tau = cac^{-1} + d = \sigma \bar{d}|d|^{-2} + d$. Equating the non-real parts of this equation we see that $|\sigma| = |d|^2$. Hence $cac^{-1} = \sigma d^{-1} = \bar{d}$. Cases (ii) and (iii) can be handled similarly. The proof of the converse implication in the lemma is straightforward. $\blacksquare$

Proposition 4.3. *Given $A \in \mathrm{SL}(2, \mathbb{H})$, if σ_A and τ_A are real then they are preserved under conjugation in $\mathrm{SL}(2, \mathbb{H})$.*

Proof. The group $\mathrm{SL}(2, \mathbb{H})$ is generated by matrices of the form

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

where $|\lambda||\mu| = 1$. Denote one of these matrices by P. Let $B = PAP^{-1}$. It suffices to show that for each choice of P, we have $\sigma_B = \sigma_A$ and $\tau_B = \tau_A$. Denote the coefficients of B by a', b', c' and d'.

In the first case we have

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \lambda^{-1} & 0 \\ 0 & \mu^{-1} \end{pmatrix} = \begin{pmatrix} \lambda a \lambda^{-1} & \lambda b \mu^{-1} \\ \mu c \lambda^{-1} & \mu d \mu^{-1} \end{pmatrix}.$$

If $c \neq 0$ then

$$\sigma_B = \mu(cac^{-1}d - cb)\mu^{-1} = \sigma_A,$$
$$\tau_B = \mu(cac^{-1} + d)\mu^{-1} = \tau_A.$$

Suppose that $c = 0$. We apply Lemma 4.2. If a and d are real then $a' = a$ and $d' = d$. Otherwise, $\sigma_B = |a'|^2 = |a|^2 = \sigma_A$ and $\tau_B = 2\,\mathrm{Re}[d] = 2\,\mathrm{Re}[a] = \tau_A$.

In the second case we have

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a + c & b - a + d - c \\ c & d - c \end{pmatrix}.$$

If $c \neq 0$ then

$$\sigma_B = c(a + c)c^{-1}(d - c) - c(b - a + d - c) = \sigma_A,$$
$$\tau_B = c(a + c)c^{-1} + d - c = \tau_A.$$

If $c = 0$ then $a' = a$ and $d' = d$ and the result follows from Lemma 4.2.

In the third case we have

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} d & c \\ b & a \end{pmatrix}.$$

If $b \neq 0$ and $c \neq 0$ then, using Lemma 4.1, we have

$$\sigma_B = bdb^{-1}a - bc = \sigma_A,$$
$$\tau_B = bdb^{-1} + a = \tau_A.$$

Suppose that $b = 0$ or $c = 0$. If a and d are both real then $\sigma_B = da = \sigma_A$ and $\tau_B = d + a = \tau_A$. Otherwise we get the relations $\sigma_B = |a|^2 = |d|^2 = \sigma_A$ and $\tau_B = 2\,\mathrm{Re}[d] = 2\,\mathrm{Re}[a] = \tau_A$. ∎

We are now in a position to prove Theorem 1.2.

Proof of Theorem 1.2. If a quaternionic Möbius transformation is conjugate to a real Möbius transformation then σ and τ are real, by Proposition 4.3. Conversely, given a quaternionic Möbius transformation f, suppose that σ and τ are both real. We may replace f by a conjugate transformation that fixes ∞ (meaning that $c = 0$). Proposition 4.3 ensures that σ and τ remain real. Now we apply Lemma 4.2. If a and d are real but b is not real then we conjugate f by the map g, given by the equation $g(z) = b^{-1}z$, to obtain the real Möbius transformation $gfg^{-1}(z) = (az + 1)d^{-1}$.

It remains to consider the case when a and d are similar, not real and $\sigma = |a|^2$ and $\tau = 2\,\mathrm{Re}[a]$. If $b \neq 0$ then $bdb^{-1} = \sigma a^{-1} = \bar{a}$ so that, by Lemma 3.2, f has a fixed point v in $\mathbb{H}$. After replacing f by another conjugate transformation we

may assume that $v = 0$ (meaning that $b = 0$). Apply one final conjugation to f by a map of the form $z \mapsto uz$, for $u \in \mathbb{H}$, to ensure that $a \neq d$. This means that $a = x + \mu y$ and $d = x + \nu y$ for real numbers x and y and distinct purely imaginary unit quaternions μ and ν. From the matrix equation

$$\begin{pmatrix} \mu & 1 \\ 1 & -\nu \end{pmatrix} \begin{pmatrix} x + \mu y & 0 \\ 0 & x + \nu y \end{pmatrix} \begin{pmatrix} \mu & 1 \\ 1 & -\nu \end{pmatrix}^{-1} = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

we see that f is conjugate to a real Möbius transformation. ∎

5. Proof of Theorem 1.3

Proof of Theorem 1.3. Suppose that f and g are conjugate. Condition (i) is obvious and condition (ii) follows from the conjugacy invariance of the quantities β, γ and δ in $\mathrm{SL}(2, \mathbb{H})$. Conversely, assume that conditions (i) and (ii) hold. If f and g are both conjugate to real Möbius transformations then σ and τ are real (by Theorem 1.2) and preserved under conjugation (by Proposition 4.3). For real Möbius transformations, $\beta\delta = \sigma\tau^2$. Therefore, because σ is either -1 or 1, we have $\tau_f^2/\sigma_f = \tau_g^2/\sigma_g$. The quantity τ^2/σ determines conjugacy amongst real Möbius transformations, therefore f and g are conjugate. Henceforth we assume that neither f nor g are conjugate to real Möbius transformations.

Using condition (ii), we lift f and g to two matrices A and B in $\mathrm{SL}(2, \mathbb{H})$ for which $\beta_A = \beta_B$, $\gamma_A = \gamma_B$ and $\delta_A = \delta_B$. From Theorem 3.1 we see that A and B have the same pair of right eigenvalue similarity classes. If the two similarity classes in the pair are distinct then A and B are diagonalisable, and we can conjugate one matrix to the other using a diagonal conjugating matrix. In the remaining case, both classes consist of quaternions similar to a particular (non-real) quaternion t. Since A and B are not conjugate to real matrices, we see from Lemma 3.2 that they are both conjugate to $\left(\begin{smallmatrix} t & 1 \\ 0 & t \end{smallmatrix} \right)$. ∎

6. Proof of Theorem 1.4

It is straightforward to establish using either algebraic or geometric methods that each map in $\mathrm{PSL}(2, \mathbb{H})$ can be expressed as a composite of three simple maps (and that the only 3-simple maps are loxodromic). The following proposition shows how to determine whether a map can be expressed as a composite of fewer than three simple maps.

Proposition 6.1. *A map in* $\mathrm{PSL}(2, \mathbb{H})$ *is 3-simple if and only if* $\beta \neq \delta$.

Proof. Suppose that f_1 and f_2 are simple maps; possibly one of them is the identity. Let $f = f_1 f_2$. We use the obvious notation $f_1(z) = (a_1 z + b_1)(c_1 z + d_1)^{-1}$, σ, σ_1, σ_2, and so forth. We have to show that $\beta = \delta$. Conjugating if necessary, we suppose that $c = c_1 a_2 + d_1 c_2 = 0$. If either (and hence both) of c_1 and c_2 are zero then the result is straightforward. So suppose that for $j = 1, 2$ we have

$c_j \neq 0$ and $\sigma_j = c_j a_j c_j^{-1} d_j - c_j b_j = 1$. Then we can write $b_j = a_j c_j^{-1} d_j - c_j^{-1}$. We have

$$a = a_1 a_2 + b_1 c_2 = a_1 c_1^{-1}(c_1 a_2 + d_1 c_2) - c_1^{-1} c_2 = -c_1^{-1} c_2,$$
$$d = c_1 b_2 + d_1 d_2 = (c_1 a_2 + d_1 c_2) c_2^{-1} d_2 - c_1 c_2^{-1} = -c_1 c_2^{-1}.$$

Then we have

$$\begin{aligned}
\beta - \delta &= \mathrm{Re}\left[|a|^2 \,\overline{d} + |d|^2 \,\overline{a} \right] - \mathrm{Re}[a + d] \\
&= \mathrm{Re}\left[-|c_1|^{-2}|c_2|^2 \,\overline{c}_2^{\,-1} \,\overline{c}_1 - |c_1|^2|c_2|^{-2} \,\overline{c}_2 \,\overline{c}_1^{\,-1} \right] - \mathrm{Re}\left[-c_2 c_1^{-1} - c_2^{-1} c_1 \right] \\
&= 0.
\end{aligned}$$

Conversely, if f is 3-simple then it is necessarily loxodromic and we can conjugate f so that it assumes the form $f(z) = (\lambda u)z(\lambda^{-1}v)d^{-1}$, where $\lambda > 1$ and u and v are unit quaternions that are not similar. One can check that $\beta \neq \delta$. ∎

The 3-simple maps are precisely those quaternionic Möbius transformations that do not fix a four dimensional ball, hence we have the following corollary.

Corollary 6.2. *A map in* $\mathrm{PSL}(2, \mathbb{H})$ *fixes a four dimensional ball setwise if and only if* $\beta = \delta$.

Proof of Theorem 1.4. The map f is 1-simple if and only if $\sigma = 1$ and $\tau \in \mathbb{R}$. Otherwise, either $\beta = \delta$, in which case f is 2-simple, or $\beta \neq \delta$, in which case f is 3-simple (by Proposition 6.1). This completes the classification into (a), (b) and (c).

In case (a), we have $\delta^2 = \tau^2$ and the classification into (i), (ii) and (iii) corresponds to the usual classification for real Möbius transformations. In case (b), if f is elliptic then we conjugate f so that it is of the form $f(z) = azd^{-1}$, for unit quaternions a and d. This map satisfies $\gamma - \delta^2 < 2$. If f is parabolic then we conjugate f so that it is of the form $f(z) = (az + 1)a^{-1}$. This map satisfies $\gamma - \delta^2 = 2$. Finally, if f is loxodromic then we conjugate f so that it is of the form $f(z) = (\lambda u)z(\lambda^{-1}v)d^{-1}$, where $\lambda > 1$ and u and v are unit quaternions. Since $\beta = \delta$ we find that u and v are similar. This means that $\gamma - \delta^2 > 2$. ∎

References

1. A. F. Beardon, *The Geometry of Discrete Groups*, Graduate Texts in Mathematics, 91, Springer Verlag, 1983.
2. J. L. Brenner, Matrices of quaternions, *Pacific J. Math.* **1** (1951), 329–335.
3. W. Cao, J. R. Parker and X. Wang, On the classification of quaternionic Möbius transformations, *Math. Proc. Camb. Phil. Soc.* **137** (2004), 349–361.
4. W. Cao, On the classification of four-dimensional Möbius transformations, *Proc. Edinb. Math. Soc.* **50** (2007), 49–62.
5. H. S. M. Coxeter, Quaternions and reflections, *Amer. Math. Monthly* **53** (1946), 136–146.
6. B. Foreman, Conjugacy invariants of $Sl(2, \mathbb{H})$, *Linear Algebra Appl.* **381** (2004), 24–35.
7. K. Gongopadhyay and R. S. Kulkarni, Dynamical types of isometries of hyperbolic space, preprint 2006.

8. P. G. Gormley, Stereographic projection and the linear fractional group transformations of quaternions, *Proc. Roy. Irish Acad.* **57** (1947), 67–85.

9. L. Huang and W. So, Quadratic formulas for quaternions, *Appl. Math. Lett.* **15** (2002), 533–540.

10. R. Kellerhals, Quaternions and some global properties of hyperbolic 5-manifolds, *Canad. J. Math.* **55** (2003), 1080–1099.

11. T. Kido, Möbius transformations on quaternions, Osaka University preprint, 2005.

12. H. C. Lee, Eigenvalues and canonical forms of matrices with quaternion coefficients, *Proc. Roy. Irish Acad. Sect. A.* **15** (1949), 253–260.

13. J. G. Ratcliffe, *Foundations of Hyperbolic Manifolds*, Graduate Texts in Mathematics 149, Springer-Verlag, 1994.

14. J. B. Wilker, The quaternion formalism for Möbius groups in four or fewer dimensions, *Linear Algebra Appl.* **190** (1993), 99–136.

John R. Parker   E-MAIL: `j.r.parker@durham.ac.uk`
ADDRESS: *University of Durham, Department of Mathematical Sciences, Science Laboratories, South Road, Durham DH1 3LE, U.K.*

Ian Short   E-MAIL: `ian.short@nuim.ie`
ADDRESS: *National University of Ireland Maynooth, Mathematics Department, Logic House, Maynooth, County Kildare, Ireland.*