Hyperbolic Spaces

The Jyväskylä Notes

John R. Parker
Department of Mathematical Sciences
University of Durham
Durham DH1 3LE, England
J.R.Parker@durham.ac.uk

November 21, 2007

Contents

Chapter 1

Preface

My starting point in these notes is the use of techniques from linear algebra to describe the geometry of the hyperbolic plane $\mathbf{H}^2$. Specifically, we consider geometric objects contained in $\mathbf{H}^2$, such as geodesics, circles and horocycles. We also use linear algebra to analyse the action of hyperbolic isometries, including classifying them and finding their fixed points. This viewpoint differs somewhat from traditional treatments, but follows along very similar lines.

The reason to use this technique is that the methods generalise naturally to higher dimensions and to different algebraic settings. My principal aim is to reach the description of high dimensional hyperbolic isometries using Möbius transformations whose entries lie in a Clifford algebra. Although this idea has a long history, current interest is due to work done by Ahlfors to popularise this field. I feel it is particularly appropriate to include this material in lectures given in Finland in the hundredth year since his birth in Helsinki. Some of the formulae that appear in Ahlfors's work make more sense when viewed in terms of Hermitian forms, and the use of linear algebra streamlines some of the calculations. In the final two sections I indicate further directions in which the main material may be generalised.

These notes are based on material I distributed to those attending the lecture course *Hyperbolic Spaces* which I gave as part of the 17th Jyväskylä Summer School between 13th and 17th August 2007. I have made some changes to the notes I distributed then. Many of these changes have been made at the suggestion of those taking the course and I am very grateful to all those who have pointed out errors to me. It became clear that some of the exercises were not appropriate and I have taken this opportunity to make them clearer. My main thanks go to Jouni Parkkonen who was the main organiser of the mathematics courses in the summer school. I would like to thank him for his invitation and also his support and encouragement. I would also like to thank Henna Koivusalo for her assistance.

Chapter 2

Hermitian linear algebra

2.1 Hermitian forms

Let $A = (a_{ij})$ be a $k \times l$ complex matrix. The *Hermitian transpose* of A is the $l \times k$ complex matrix $A^* = (\overline{a}_{ji})$ formed by complex conjugating each entry of A and then taking the transpose. As with ordinary transpose, the Hermitian transpose of a product is the product of the Hermitian transposes in the reverse order. That is $(AB)^* = B^*A^*$. Clearly $\big((A^*)^*\big) = A$. If $\mathbf{x}$ is a column vector in $\mathbb{C}^k$ then $\mathbf{x}^*\mathbf{x} = |\mathbf{x}|^2$. A $k \times k$ complex matrix H is said to be *Hermitian* if it equals its own Hermitian transpose $H = H^*$. Let H be a Hermitian matrix and μ an eigenvalue of H with eigenvector $\mathbf{x}$. We claim that μ is real. In order to see this, observe that

$$\mu\mathbf{x}^*\mathbf{x} = \mathbf{x}^*(\mu\mathbf{x}) = \mathbf{x}^*H\mathbf{x} = \mathbf{x}^*H^*\mathbf{x} = (H\mathbf{x})^*\mathbf{x} = (\mu\mathbf{x})^*\mathbf{x} = \overline{\mu}\mathbf{x}^*\mathbf{x}.$$

Since $\mathbf{x}^*\mathbf{x} = |\mathbf{x}|^2$ is real and non-zero we see that μ is real.

To each $k \times k$ Hermitian matrix H we can naturally associate an *Hermitian form* $\langle \cdot, \cdot \rangle : \mathbb{C}^k \times \mathbb{C}^k \longrightarrow \mathbb{C}$ given by $\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^*H\mathbf{z}$ (note that we change the order) where $\mathbf{w}$ and $\mathbf{z}$ are column vectors in $\mathbb{C}^k$. Hermitian forms are sesquilinear, that is they are linear in the first factor and conjugate linear in the second factor. In other words, for $\mathbf{z}$, $\mathbf{z}_1$, $\mathbf{z}_2$, $\mathbf{w}$ column vectors in $\mathbb{C}^k$ and λ a complex scalar, we have

$$\begin{aligned}
\langle \mathbf{z}_1 + \mathbf{z}_2, \mathbf{w} \rangle &= \mathbf{w}^*H(\mathbf{z}_1 + \mathbf{z}_2) = \mathbf{w}^*H\mathbf{z}_1 + \mathbf{w}^*H\mathbf{z}_2 = \langle \mathbf{z}_1, \mathbf{w} \rangle + \langle \mathbf{z}_2, \mathbf{w} \rangle, \\
\langle \lambda\mathbf{z}, \mathbf{w} \rangle &= \mathbf{w}^*H(\lambda\mathbf{z}) = \lambda\mathbf{w}^*H\mathbf{z} = \lambda \langle \mathbf{z}, \mathbf{w} \rangle, \\
\langle \mathbf{w}, \mathbf{z} \rangle &= \mathbf{z}^*H\mathbf{w} = \mathbf{z}^*H^*\mathbf{w} = (\mathbf{w}^*H\mathbf{z})^* = \overline{\langle \mathbf{z}, \mathbf{w} \rangle}.
\end{aligned}$$

Exercise 2.1.1 *For all* $\mathbf{z}$, $\mathbf{w} \in \mathbb{C}^k$ *and* $\lambda \in \mathbb{C}$ *show that*

1. $\langle \mathbf{z}, \mathbf{z} \rangle \in \mathbb{R}$,
2. $\langle \mathbf{z}, \lambda\mathbf{w} \rangle = \overline{\lambda} \langle \mathbf{z}, \mathbf{w} \rangle$,
3. $\langle \lambda\mathbf{z}, \lambda\mathbf{w} \rangle = |\lambda|^2 \langle \mathbf{z}, \mathbf{w} \rangle$.

Let $\langle\cdot,\cdot\rangle$ be a Hermitian form associated to the Hermitian matrix H. We know that the eigenvalues of H are real. We say that

(i) $\langle\cdot,\cdot\rangle$ is *non-degenerate* if all the eigenvalues of H are non-zero;

(ii) $\langle\cdot,\cdot\rangle$ is *positive definite* if all the eigenvalues of H are positive;

(iii) $\langle\cdot,\cdot\rangle$ is *negative definite* if all the eigenvalues of H are negative;

(iv) $\langle\cdot,\cdot\rangle$ is *indefinite* if some eigenvalues of H are positive and some are negative.

Suppose that $\langle\cdot,\cdot\rangle$ is a non-degenerate Hermitian form associated to the $k\times k$ Hermitian matrix H. We say that $\langle\cdot,\cdot\rangle$ has *signature* (p,q) where $p+q=k$ if H has p positive eigenvalues and q negative eigenvalues. Thus positive definite Hermitian forms have signature $(k,0)$ and negative definite forms have signature $(0,k)$. We often write $\mathbb{C}^{p,q}$ for $\mathbb{C}^{p+q}$ equipped with a non-degenerate Hermitian form of signature (p,q). This generalises the idea of $\mathbb{C}^p$ with the implied Hermitian form of signature $(p,0)$.

For real matrices the Hermitian transpose coincides with the ordinary transpose. A real matrix that equals its own transpose is called *symmetric*. Symmetric matrices define bilinear forms on real vector spaces, usually called *quadratic forms*.

Exercise 2.1.2 *Show that all the above definitions (non-degenerate, positive definite and so on) may be carried through for quadratic forms on real vector spaces. Hence make sense of the signature of a quadratic form and the real vector space* $\mathbb{R}^{p,q}$.

Theorem 2.1.1 (Sylvester's principle of inertia) *The signature of a Hermitian matrix is independent of the means of finding it. In particular, if* H_1 *and* H_2 *are two* $k\times k$ *Hermitian matrices with the same signature then there exists a* $k\times k$ *matrix* C *so that* $H_2=C^*H_1C$.

2.2 Unitary matrices

Let $\langle\cdot,\cdot\rangle$ be a Hermitian form associated to the $k\times k$ Hermitian matrix H. A $k\times k$ matrix A is called *unitary* if for all $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{C}^k$ we have

$$\langle A\mathbf{z},A\mathbf{w}\rangle=\langle\mathbf{z},\mathbf{w}\rangle.$$

If the Hermitian form is non-degenerate then unitary matrices form a group. The group of matrices preserving this Hermitian form will be denoted $\mathrm{U}(H)$. Sometimes it is only necessary to determine the signature. If $\langle\cdot,\cdot\rangle$ has signature (p,q) then we write $\mathrm{U}(p,q)$.

Since A preserves the form we have

$$\mathbf{w}^*A^*HA\mathbf{z}=(A\mathbf{w})^*H(A\mathbf{z})=\langle A\mathbf{z},A\mathbf{w}\rangle=\langle\mathbf{z},\mathbf{w}\rangle=\mathbf{w}^*H\mathbf{z}.$$

Therefore letting $\mathbf{z}$ and $\mathbf{w}$ run through a basis for $\mathbb{C}^k$ we have $A^*HA = H$. If H is non-degenerate then it is invertible and this translates to an easy formula for the inverse of A:

$$A^{-1} = H^{-1}A^*H.$$

Most of the Hermitian forms we consider will have eigenvalues ± 1 and so will be their own inverse.

One consequence of this formula is that

$$\det(H) = \det(A^*HA) = \det(A^*)\det(H)\det(A).$$

If $\det(H) \neq 0$ (so the form is non-degenerate) then

$$1 = \det(A^*)\det(A) = \overline{\det(A)}\det(A) = \big|\det(A)\big|^2.$$

Thus unitary matrices have unit modulus determinant. The group of those unitary matrices whose determinant is $+1$ is denoted $\mathrm{SU}(H)$.

Exercise 2.2.1 *Let H_1 and H_2 be two $k \times k$ Hermitian matrices with the same signature. Show that $\mathrm{U}(H_1)$ is conjugate to $\mathrm{U}(H_2)$. [Use Sylvester's principle of inertia.]*

2.3 Eigenvalues and eigenvectors

Lemma 2.3.1 *Let $A \in \mathrm{SU}(H)$ and let λ be an eigenvalue of A. Then $\overline{\lambda}^{-1}$ is an eigenvalue of A.*

Proof: We know that A preserves the Hermitian form defined by H. Hence, $A^*HA = H$ and so $A = H^{-1}(A^*)^{-1}H$. Thus A has the same set of eigenvalues as $(A^*)^{-1}$ (they are conjugate). Since the characteristic polynomial of A^* is the complex conjugate of the characteristic polynomial of A, we see that if λ is an eigenvalue of A then $\overline{\lambda}$ is an eigenvalue of A^*. Therefore $\overline{\lambda}^{-1}$ is an eigenvalue of $(A^*)^{-1}$ and hence of A. $\square$

Corollary 2.3.2 *If λ is an eigenvalue of $A \in \mathrm{SU}(H)$ with $|\lambda| \neq 1$ then $\overline{\lambda}^{-1}$ is a distinct eigenvalue.*

Next we show that any eigenvalue not of unit modulus corresponds to a null eigenvector (that is a non-zero vector $\mathbf{v}$ with $\langle \mathbf{v}, \mathbf{v} \rangle = 0$). Likewise, we show that any eigenvectors that are not (Hermitian) orthogonal have eigenvalues λ and $\mu = \overline{\lambda}^{-1}$.

Lemma 2.3.3 *Let λ, μ be eigenvalues of $A \in \mathrm{SU}(H)$ and let $\mathbf{v}$, $\mathbf{w}$ be any eigenvectors with eigenvalues λ, μ respectively.*

(i) Either $|\lambda| = 1$ *or* $\langle \mathbf{v}, \mathbf{v} \rangle = 0$.

(ii) Either $\lambda\overline{\mu} = 1$ *or* $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

Proof: (i)

$$\langle \mathbf{v}, \mathbf{v} \rangle = \langle A\mathbf{v}, A\mathbf{v} \rangle = \langle \lambda\mathbf{v}, \lambda\mathbf{v} \rangle = |\lambda|^2 \langle \mathbf{v}, \mathbf{v} \rangle .$$

Thus either $|\lambda| = 1$ or $\langle \mathbf{v}, \mathbf{v} \rangle = 0$.

(ii)

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle A\mathbf{v}, A\mathbf{w} \rangle = \langle \lambda\mathbf{v}, \mu\mathbf{w} \rangle = \lambda\overline{\mu} \langle \mathbf{v}, \mathbf{w} \rangle .$$

Thus either $\lambda\overline{\mu} = 1$ or $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. $\square$

Chapter 3

The Poincaré models of the hyperbolic plane

3.1 Hermitian forms of signature $(1, 1)$

Let $\mathbb{C}^{1,1}$ be the complex vector space of (complex) dimension 2 equipped with a non-degenerate, indefinite Hermitian form $\langle \cdot, \cdot \rangle$ of signature $(1, 1)$. This means that $\langle \cdot, \cdot \rangle$ is given by a non-singular 2×2 Hermitian matrix H with 1 positive eigenvalue and 1 negative eigenvalue. There are two standard matrices H which give different Hermitian forms on $\mathbb{C}^{1,1}$. We call these the first and second Hermitian forms. Let $\mathbf{z}$, $\mathbf{w}$ be the column vectors $(z_1,\, z_2)^t$ and $(w_1,\, w_2)^t$ respectively. The *first Hermitian form* is defined to be:

$$\langle \mathbf{z}, \mathbf{w} \rangle_1 = z_1\overline{w}_1 - z_2\overline{w}_2. \tag{3.1}$$

It is given by the Hermitian matrix H_1:

$$H_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.2}$$

The *second Hermitian form* is defined to be:

$$\langle \mathbf{z}, \mathbf{w} \rangle_2 = iz_1\overline{w}_2 - iz_2\overline{w}_1. \tag{3.3}$$

It is given by the Hermitian matrix H_2:

$$H_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{3.4}$$

We can pass between these Hermitian forms via a *Cayley transform* C. For example, writing

$$C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \tag{3.5}$$

then

$$H_2 = C^* H_1 C.$$

Sometimes we want to specify which of these two Hermitian forms to use. When there is no subscript then you can use either of these (or your favourite Hermitian form on $\mathbb{C}^2$ of signature $(1,1)$).

If $\mathbf{z} \in \mathbb{C}^{1,1}$ then we know that $\langle \mathbf{z}, \mathbf{z} \rangle$ is real. Thus we may define subsets V_-, V_0 and V_+ of $\mathbb{C}^{1,1} - \{0\}$ by

$$\begin{aligned} V_- &= \left\{ \mathbf{z} \in \mathbb{C}^{1,1} \middle| \langle \mathbf{z}, \mathbf{z} \rangle < 0 \right\}, \\ V_0 &= \left\{ \mathbf{z} \in \mathbb{C}^{1,1} - \{0\} \middle| \langle \mathbf{z}, \mathbf{z} \rangle = 0 \right\}, \\ V_+ &= \left\{ \mathbf{z} \in \mathbb{C}^{1,1} \middle| \langle \mathbf{z}, \mathbf{z} \rangle > 0 \right\}. \end{aligned}$$

We say that $\mathbf{z} \in \mathbb{C}^{1,1}$ is *negative*, *null* or *positive* if $\mathbf{z}$ is in V_-, V_0 or V_+ respectively. Motivated by special relativity, these are sometimes called time-like, light-like and space-like. Because $\langle \lambda\mathbf{z}, \lambda\mathbf{z} \rangle = |\lambda|^2 \langle \mathbf{z}, \mathbf{z} \rangle$ we see that for any non-zero complex scalar λ the point $\lambda\mathbf{z}$ is negative, null or positive if and only if $\mathbf{z}$ is. Therefore we define a projection map $\mathbb{P}$ on those points of $\mathbb{C}^{1,1} - \{0\}$ to $\mathbb{CP}^1 = \widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, the Riemann sphere. This map is given by

$$\mathbb{P} : \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \longmapsto \begin{cases} z_1/z_2 \in \mathbb{C} & \text{if } z_2 \neq 0, \\ \infty & \text{if } z_2 = 0. \end{cases}$$

The *complex projective model* of the hyperbolic plane $\mathbf{H}^2$ is defined to be the collection of negative lines in $\mathbb{C}^{1,1}$ and its ideal boundary is defined to be the collection of null lines. In other words $\mathbf{H}^2$ is $\mathbb{P}V_-$ and $\partial\mathbf{H}^2$ is $\mathbb{P}V_0$.

We define the other two standard models of complex hyperbolic space by taking the section defined by $z_2 = 1$ for the first and second Hermitian forms. In other words, if we take column vectors $\mathbf{z} = (z,\, 1)^t$ in $\mathbb{C}^{1,1}$ then consider what it means for $\langle \mathbf{z}, \mathbf{z} \rangle$ to be negative. Conversely, we can pass from points of $\widehat{\mathbb{C}}$ to $\mathbb{C}^{1,1}$ be taking the *standard lift*. This is defined by

$$z \longmapsto \mathbf{z} = \begin{pmatrix} z \\ 1 \end{pmatrix}, \quad \text{for } z \in \mathbb{C}; \qquad \infty \longmapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

For the first Hermitian form we obtain that if $\mathbf{z}$ and $\mathbf{w}$ are the standard lifts of points z and w we have $\langle \mathbf{z}, \mathbf{w} \rangle_1 = z\overline{w} - 1$. Thus a point $z \in \widehat{\mathbb{C}}$ is in $\mathbf{H}^2 = \mathbb{P}V_-$ provided its standard lift $\mathbf{z}$ satisfies:

$$\langle \mathbf{z}, \mathbf{z} \rangle_1 = z\overline{z} - 1 = |z|^2 - 1 < 0.$$

In other words $|z| < 1$, the unit disc in $\mathbb{C}$. This is called the *Poincaré disc model* of the hyperbolic plane. The ideal boundary of the Poincaré disc is $\mathbb{P}V_0$. A similar argument shows that $\partial\mathbf{H}^2$ is the unit circle given by $|z| = 1$.

For the second Hermitian form we obtain that $\langle \mathbf{z}, \mathbf{w} \rangle_2 = iz - i\overline{w}$. Therefore a point $z \in \widehat{\mathbb{C}}$ lies in in $\mathbf{H}^2$ provided:

$$\langle \mathbf{z}, \mathbf{z} \rangle_2 = iz - i\overline{z} = -2\mathrm{Im}(z) < 0.$$

In other words $\mathrm{Im}(z) > 0$ and we obtain the *upper half plane* model of the hyperbolic plane. The ideal boundary $\mathbb{P}V_0$ is the extended real line $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$.

The *Poincaré metric* on either Poincaré disc or the upper half plane $\mathbf{H}^2$ is given by

$$ds^2 = \frac{-4}{\langle \mathbf{z}, \mathbf{z} \rangle^2} \det \begin{pmatrix} \langle \mathbf{z}, \mathbf{z} \rangle & \langle d\mathbf{z}, \mathbf{z} \rangle \\ \langle \mathbf{z}, d\mathbf{z} \rangle & \langle d\mathbf{z}, d\mathbf{z} \rangle \end{pmatrix} \tag{3.6}$$

where $\mathbf{z}$ is any lift of $z \in \mathbf{H}^2$ and $d\mathbf{z}$ is its differential. Alternatively, the Poincaré metric is given by the distance function $\rho(\cdot, \cdot)$ defined by the formula

$$\cosh^2\left(\frac{\rho(z, w)}{2}\right) = \frac{\langle \mathbf{z}, \mathbf{w} \rangle \langle \mathbf{w}, \mathbf{z} \rangle}{\langle \mathbf{z}, \mathbf{z} \rangle \langle \mathbf{w}, \mathbf{w} \rangle} \tag{3.7}$$

where $\mathbf{z}$ and $\mathbf{w}$ are lifts of z and w to $\mathbb{C}^{1,1}$. However, as may easily be seen, this formula is independent of which lifts $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{C}^{1,1}$ of z and w we choose.

Exercise 3.1.1 *1. For the Poincaré disc, use the first Hermitian form to show that*

$$ds^2 = \frac{4dz\, d\overline{z}}{\left(1 - |z|^2\right)^2}, \qquad \cosh^2\left(\frac{\rho(z, w)}{2}\right) = \frac{|z\overline{w} - 1|^2}{\left(1 - |z|^2\right)\left(1 - |w|^2\right)}. \tag{3.8}$$

2. For the upper half plane, use the second Hermitian form to show that

$$ds^2 = \frac{dz\, d\overline{z}}{\mathrm{Im}(z)^2}, \qquad \cosh^2\left(\frac{\rho(z, w)}{2}\right) = \frac{|z - \overline{w}|^2}{4\mathrm{Im}(z)\mathrm{Im}(w)}. \tag{3.9}$$

After applying the projection map $\mathbb{P}$, the Cayley transform C given by (3.5) becomes a Möbius transformation $C(z)$ sending the upper half plane to the Poincaré disc. Using composition of functions, we can pass between the different expressions for the metric.

3.2 Isometries

We let $\mathrm{U}(1,1)$ denote the collection of all unitary matrices preserving a given Hermitian form of signature $(1,1)$. When the form is H_1 or H_2 this group will be denoted $\mathrm{U}(H_1)$ or $\mathrm{U}(H_2)$ respectively. Sometimes it is convenient to assume that $\det(A) = 1$. The collection of all unitary matrices with determinant 1 will be written as $\mathrm{SU}(1,1)$.

For any Hermitian form of signature $(1,1)$, the corresponding unitary matrices in $\mathrm{SU}(1,1)$ naturally act on $\mathbb{C}^{1,1}$. Since they preserve the Hermitian form they also preserve V_+, V_0 and V_-. Let $\mathbf{z}$ be the standard lift of $z \in \widehat{\mathbb{C}} = \mathbb{CP}^1$ and let $A \in \mathrm{SU}(1,1)$. What does $\mathbb{P}(A\mathbf{z})$ look like as a function of z?

$$\begin{aligned}
\mathbb{P}A\mathbf{z} &= \mathbb{P}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} z \\ 1 \end{pmatrix} \\
&= \mathbb{P}\begin{pmatrix} az+b \\ cz+d \end{pmatrix} \\
&= \begin{cases} (az+b)/(cz+d) & \text{if } z \neq -d/c \\ \infty & \text{if } z = -d/c. \end{cases}
\end{aligned}$$

Therefore, the answer is that A acts on $\mathbb{CP}^1 = \widehat{\mathbb{C}}$ via the Möbius transformation $A(z) = (az+b)/(cz+d)$. We will frequently pass between matrices and Möbius transformations without comment. (We abuse notation by using the same letter for the matrix and the Möbius transformation.) The collection of Möbius transformations preserving the unit disc will be denoted $\mathrm{PSU}(1,1)$ and those preserving the upper half plane by $\mathrm{PSL}(2,\mathbb{R})$.

We now examine what unitary matrices look like in terms of the first Hermitian form. Let

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

Then

$$\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1} = H_1^{-1}A^*H_1 = \begin{pmatrix} \overline{a} & -\overline{c} \\ -\overline{b} & \overline{a} \end{pmatrix}.$$

In particular, if $A \in \mathrm{SU}(H_1)$ then $b = \overline{c}$ and $d = \overline{a}$. Hence $1 = |a|^2 - |c|^2$. Thus

$$\mathrm{SU}(H_1) = \left\{ \begin{pmatrix} a & \overline{c} \\ c & \overline{a} \end{pmatrix} \;:\; a,\, c \in \mathbb{C};\ |a|^2 - |c|^2 = 1 \right\}. \tag{3.10}$$

We now repeat the above construction for the second Hermitian form. Writing A as above, we have

$$\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1} = H_2^{-1}A^*H_2 = \begin{pmatrix} \overline{d} & -\overline{b} \\ -\overline{c} & \overline{a} \end{pmatrix}.$$

Thus, if $\det(A) = 1$ we see that its entries are all real. Hence the special unitary group preserving the second Hermitian form is

$$\mathrm{SU}(H_2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;:\; a,\, b,\, c,\, d \in \mathbb{R};\ ad-bc = 1 \right\}. \tag{3.11}$$

This group is more commonly denoted $\mathrm{SL}(2,\mathbb{R})$ and we will call it this from now on. It is also usual to write $\mathrm{SU}(H_1)$ simply as $\mathrm{SU}(1,1)$.

The Cayley transform C defined by (3.5) acts on $\widehat{C}$ by the Möbius transformation

$$C(z) = (z-i)/(-iz+1). \tag{3.12}$$

Exercise 3.2.1 *Show that the Cayley transformation $C(z)$ given by (3.12) maps the extended real line $\mathbb{R}\cup\{\infty\}$ to the unit circle and sends the point i to 0. Deduce that it maps the upper half plane to the unit disc.*

Lemma 3.2.1 *The group $\mathrm{SU}(H_1)$ is generated by matrices of the form*

$$D = \begin{pmatrix} \cosh(\lambda) & \sinh(\lambda) \\ \sinh(\lambda) & \cosh(\lambda) \end{pmatrix}, \qquad S = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

where $\lambda \in \mathbb{R}_+$ and $\theta \in [0, 2\pi)$.

Proof: Suppose that

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

where a and c are complex numbers with $|a|^2 - |c|^2 = 1$. Let $\lambda \in \mathbb{R}_+$ be defined by $\sinh(\lambda) = |c|$. Then we have $|a|^2 = 1+|c|^2 = 1+\sinh^2(\lambda) = \cosh^2(\lambda)$ so $\cosh(\lambda) = |a|$. Define θ_1 and θ_2 mod 2π by $2\theta_1 = \arg(a) - \arg(c)$ and $2\theta_2 = \arg(a) + \arg(c)$. Then we have

$$\begin{pmatrix} e^{i\theta_1} & 0 \\ 0 & e^{-i\theta_1} \end{pmatrix} \begin{pmatrix} \cosh(\lambda) & \sinh(\lambda) \\ \sinh(\lambda) & \cosh(\lambda) \end{pmatrix} \begin{pmatrix} e^{i\theta_2} & 0 \\ 0 & e^{-i\theta_2} \end{pmatrix} = \begin{pmatrix} a & \overline{c} \\ c & \overline{a} \end{pmatrix}.$$

$\square$

Lemma 3.2.2 *The group $\mathrm{SL}(2,\mathbb{R})$ is generated by matrices of the form*

$$D = \begin{pmatrix} e^{\lambda} & 0 \\ 0 & e^{-\lambda} \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

where t, $\lambda \in \mathbb{R}$.

Proof: If $c \neq 0$ and $ad - bc = 1$ then

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

If $c = 0$ and $ad = 1$ (so $a \neq 0$) then

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

$\square$

Lemma 3.2.3 *Let $A(z)$ be the Möbius transformation in* $\mathrm{PSU}(1,1)$ *corresponding to the matrix* $A \in \mathrm{SU}(1,1)$. *Let* $v \in \mathbb{CP}^1$ *be a fixed point of* $A(z)$ *and let* $\mathbf{v}$ *be the standard lift of* v *to* $\mathbb{C}^{1,1}$. *Then* $\mathbf{v}$ *is an eigenvector of* A *with eigenvalue* $cv + \overline{a}$.

Proof: We have $A(v) = (av + \overline{c})/(cv + \overline{a}) = v$ and so $av + \overline{c} = cv^2 + \overline{a}v$. Hence

$$A\mathbf{v} = \begin{pmatrix} a & \overline{c} \\ c & \overline{a} \end{pmatrix} \begin{pmatrix} v \\ 1 \end{pmatrix} = \begin{pmatrix} av + \overline{c} \\ cv + \overline{a} \end{pmatrix} = \begin{pmatrix} cv^2 + \overline{a}v \\ cv + \overline{a} \end{pmatrix} = (cv + \overline{a}) \begin{pmatrix} v \\ 1 \end{pmatrix}.$$

□

Proposition 3.2.4 *Let* $A \in \mathrm{SU}(1,1)$ *or* $\mathrm{SL}(2,\mathbb{R})$ *and let* $\tau = \mathrm{tr}(A)$. *Then the characteristic polynomial of* A *is*

$$\mathrm{ch}_A(t) = t^2 - \tau t + 1$$

and the eigenvalues of A *are*

$$t = \frac{\tau \pm \sqrt{(\tau - 2)(\tau + 2)}}{2}.$$

In particular,

(i) if $|\tau| > 2$ *the eigenvalues of* A *are both real and reciprocals of one another;*

(ii) if $|\tau| = 2$ *then* A *has a repeated eigenvalue* ± 1;

(iii) if $|\tau| < 2$ *then the eigenvalues of* A *are complex conjugate complex numbers of modulus 1.*

Exercise 3.2.2 *Prove Proposition 3.2.4.*

Corollary 3.2.5 *(i) In case (i) of Proposition 3.2.4* $A(z)$ *has two distinct fixed points which both lie on* $\partial \mathbf{H}^2$.

(ii) In case (ii) either $A(z) = z$ *or* $A(z)$ *has a unique fixed point which lies on* $\partial \mathbf{H}^2$.

(iii) In case (iii) $A(z)$ *has a unique fixed point in* $\mathbf{H}^2$ *(and also one in the exterior).*

Proof: In (i) both eigenvalues are not equal to 1. By Lemma 2.3.3 (i) the eigenvectors must lie in V_0. Hence their image under $\mathbb{P}$ correspond to points of $\partial \mathbf{H}^2$.

In (ii) either $A = \pm I$ or else A has Jordan normal form

$$A = \begin{pmatrix} \pm 1 & 1 \\ 0 & \pm 1 \end{pmatrix}.$$

Then A has an eigenvector $\mathbf{v}$ and there is a non-zero vector $\mathbf{u}$ so that $A\mathbf{u} = \pm\mathbf{u} + \mathbf{v}$. Therefore

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle A\mathbf{u}, A\mathbf{v} \rangle = \langle \pm\mathbf{u} + \mathbf{v}, \pm\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle \pm \langle \mathbf{v}, \mathbf{v} \rangle.$$

Hence $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ and so the fixed point lies on $\partial\mathbf{H}^2$ as before.

Finally, in (iii) the eigenvalues are a pair of complex conjugate complex numbers of unit modulus, say $t = e^{\pm i\theta}$ for some $\theta \in (0, \pi)$. Suppose that they have eigenvectors $\mathbf{v}$ and $\mathbf{w}$ respectively. Since $e^{i\theta}\overline{e^{-i\theta}} = e^{2i\theta} \neq 1$, by using Lemma 2.3.3 (ii) we see that $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. From the signature of the form we see that one must be in V_- and the other in V_+. Thus the eigenvector in V_- corresponds to a unique fixed point of A in $\mathbf{H}^2$. □

This leads to an important classification result for isometries. Let $A(z)$ be a non-trivial Möbius transformation in either $\mathrm{PSU}(1,1)$ or $\mathrm{PSL}(2,\mathbb{R})$ and we think of its action on the hyperbolic plane $\mathbf{H}^2$ and its ideal boundary $\partial\mathbf{H}^2$.

(i) $A(z)$ is *loxodromic* (or *hyperbolic*) if it has exactly two fixed points, each of which lies on $\partial\mathbf{H}^2$.

(ii) $A(z)$ is *parabolic* if it has a unique fixed point on $\partial\mathbf{H}^2$.

(iii) $A(z)$ is *elliptic* if it fixes a point of $\mathbf{H}^2$ (and so exactly one such point).

Combining the previous results, we have a classification of elements of $\mathrm{SU}(1,1)$ or $\mathrm{SL}(2,\mathbb{R})$ by trace.

Proposition 3.2.6 *Let $A \in \mathrm{SU}(1,1)$ or $\mathrm{SL}(2,\mathbb{R})$ and let $A(z)$ be the corresponding Möbius transformation in $\mathrm{PSU}(1,1)$ or $\mathrm{PSL}(2,\mathbb{R})$. Then*

(i) $A(z)$ is loxodromic if $\mathrm{tr}(A) > 2$ or $\mathrm{tr}(A) < -2$;

(ii) $A(z)$ is parabolic or the identity if $\mathrm{tr}(A) = \pm 2$;

(iii) $A(z)$ is elliptic if $-2 < \mathrm{tr}(A) < 2$.

Exercise 3.2.3 *Use theorems about diagonalisation of matrices and the Jordan normal form to show that two loxodromic or parabolic elements of $\mathrm{SU}(1,1)$ or $\mathrm{SL}(2,\mathbb{R})$ are conjugate (within this group) if and only if their traces are equal, but that there are two conjugacy classes of elliptic map with the same trace (and these are inverses of each other).*

The Möbius transformations in $\mathrm{PSU}(1,1)$ or $\mathrm{PSL}(2,\mathbb{R})$ are orientation preserving isometries of the hyperbolic plane. The map $z \longrightarrow -\overline{z}$ is also a hyperbolic isometry and preserves both the Poincaré disc and the upper halfplane. In order to see this, observe that

$$\left\langle \begin{pmatrix} -\overline{z} \\ 1 \end{pmatrix}, \begin{pmatrix} -\overline{w} \\ 1 \end{pmatrix} \right\rangle = \overline{\left\langle \begin{pmatrix} z \\ 1 \end{pmatrix}, \begin{pmatrix} w \\ 1 \end{pmatrix} \right\rangle} = \left\langle \begin{pmatrix} w \\ 1 \end{pmatrix}, \begin{pmatrix} z \\ 1 \end{pmatrix} \right\rangle.$$

We can then use the formula for the hyperbolic cosine of the hyperbolic distance function to show that this map is an isometry. We now show that these are the only isometries of the hyperbolic plane.

Proposition 3.2.7 *Any hyperbolic isometry of the Poincaré disc or the upper half plane has the form $A(z)$ or $A(-\overline{z})$ where $A(z)$ is a Möbius transformation in* $\mathrm{PSU}(1,1)$ *or* $\mathrm{PSL}(2,\mathbb{R})$ *respectively.*

Proof: We work with the upper half plane. Let $\phi : \mathbf{H}^2 \longrightarrow \mathbf{H}^2$ be any isometry of the upper half plane. By applying $A(z) \in \mathrm{PSL}(2,\mathbb{R})$ we may suppose that $\phi(i) = i$ and $\phi(2i) = it$ for some $t > 1$. Using the formula (3.9) and the fact that ϕ is an isometry, we see that

$$\frac{9}{8} = \cosh^2\left(\frac{\rho(2i,i)}{2}\right) = \cosh^2\left(\frac{\rho(\phi(2i),\phi(i))}{2}\right) = \cosh^2\left(\frac{\rho(it,i)}{2}\right) = \frac{(t+1)^2}{4t}.$$

Simplifying, we see that $2t^2 - 5t + 2 = 0$ so $t = 2$ or $t = 1/2$. Since we assumed that $t > 1$ we see that ϕ fixes $2i$.

We now claim that ϕ fixes it for all $t > 0$. Suppose that $\phi(it) = x + iy$ with $y > 0$. Then

$$\begin{aligned} \frac{(t+1)^2}{4t} &= \cosh^2\left(\frac{\rho(it,i)}{2}\right) = \cosh^2\left(\frac{\rho(x+iy,i)}{2}\right) = \frac{x^2+(y+1)^2}{4y}, \\ \frac{(t+2)^2}{8t} &= \cosh^2\left(\frac{\rho(it,2i)}{2}\right) = \cosh^2\left(\frac{\rho(x+iy,2i)}{2}\right) = \frac{x^2+(y+2)^2}{8y}. \end{aligned}$$

Solving these equations we obtain $x = 0$ and $y = t$.

Finally, suppose that $\phi(z) = w$ then for all $y > 0$ we have

$$\frac{|z+iy|^2}{4y\mathrm{Im}(z)} = \cosh^2\left(\frac{\rho(z,iy)}{2}\right) = \cosh^2\left(\frac{\rho(w,iy)}{2}\right) = \frac{|w+iy|^2}{4y\mathrm{Im}(w)}.$$

Expanding and comparing coefficients of y we see that $\mathrm{Im}(w) = \mathrm{Im}(z)$. Using this and comparing the coefficients of y^{-1} leads to $\big(\mathrm{Re}(z)\big)^2 = \big(\mathrm{Re}(w)\big)^2$ and so $\mathrm{Re}(w) = \pm\mathrm{Re}(z)$. Therefore either $\phi(z) = z$ or $\phi(z) = -\overline{z}$. The result follows. $\square$

Proposition 3.2.8 *Let* $A(z) \in \mathrm{PSL}(2,\mathbb{R})$*. Then*

(i) if $a = d$ then $A(-\overline{z})$ fixes a semicircle centred on the real axis or a line orthogonal to the real axis;

(ii) if $a \neq d$ then $A(-\overline{z})$ has two fixed points on $\mathbb{R} \cup \{\infty\}$.

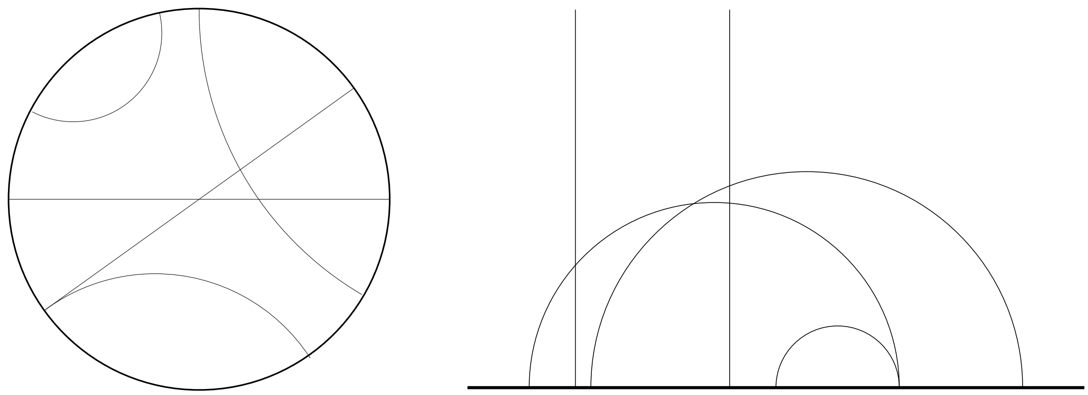

Figure 3.1: *Geodesics in the Poincaré disc and upper half plane.*

Proof: Suppose that v is a fixed point of $A(-\overline{z})$. Then $v = (-a\overline{v}+b)/(-c\overline{v}+d)$. That is

$$c|v|^2 - a\overline{v} - dv + b = 0.$$

If $a = d$ and $c \neq 0$ this is $c|v-2a/c|^2 = 4a^2/c - b = (3a^2+1)/c$. This is the equation of a circle centred on the real axis. If $a = d$ and $c = 0$ then $2a\mathrm{Re}(v) = b$ which is a line orthogonal to the real axis.

If $a \neq d$ then, since a, b, c, d are all real, the imaginary part of this equation is $(a-d)\mathrm{Im}(v) = 0$ and so v is real. It is easy to see that when $c \neq 0$

$$v = \frac{a+d \pm \sqrt{(a-d)^2+4}}{2c}$$

and when $c = 0$ then $v = \infty$ or $v = b/(a+d)$. $\square$

Let $A(z)$ be a Möbius transformation in either $\mathrm{PSU}(1,1)$ or $\mathrm{PSL}(2,\mathbb{R})$ and we think of the action of $A(-\overline{z})$ on the hyperbolic plane $\mathbf{H}^2$ and its ideal boundary $\partial\mathbf{H}^2$.

(i) $A(-\overline{z})$ is a *glide reflection* if it has exactly two fixed points lying on $\partial\mathbf{H}^2$.

(ii) $A(-\overline{z})$ is a *reflection* if it fixes a line or semicircle in $\mathbf{H}^2$.

3.3 The geometry of the hyperbolic plane

A curve is called a *geodesic* if its length is the distance between its endpoints.

Proposition 3.3.1 *Geodesics in the Poincaré disc and upper half plane are arcs of circles orthogonal to the ideal boundary; see Figure 3.1.*

Proof: We show that the imaginary axis in the upper half plane is a geodesic. The result for the upper half plane will then follow by applying isometries and using that, since they are Möbius transformations, they preserve circles (or lines) and angles.

Let iy_0 and iy_1 be two points on the imaginary axis and let $\gamma : [0,1] \longrightarrow \mathbf{H}^2$ be a path between them. That is $\gamma(0) = iy_0$ and $\gamma(1) = iy_1$. We write $\gamma(t) = x(t) + iy(t)$. On $\gamma(t)$ we have

$$ds^2 = \frac{dz\, d\overline{z}}{\mathrm{Im}(z)^2} = \frac{\left|\gamma'(t)\right|^2}{y(t)^2} dt^2 = \frac{x'(t)^2 + y'(t)^2}{y(t)^2} dt^2.$$

Then the length of γ is

$$\begin{aligned} \ell(\gamma) &= \int_{t=0}^{1} \frac{\sqrt{x'(t)^2 + y'(t)^2}}{y(t)}\, dt \\ &\geq \int_{t=0}^{1} \left| \frac{y'(t)}{y(t)} \right| dt \\ &= \left| \Big[\log\big(y(t)\big) \Big]_{t=0}^{1} \right| \\ &= \left| \log(y_1/y_0) \right| \end{aligned}$$

with equality if and only if $x'(t) = 0$ for all t, which implies that γ sends $[0,1]$ to the imaginary axis. Thus the imaginary axis is a geodesic in the upper half plane model, as claimed. □

We remark that we have shown that the length of the geodesic arc from iy_0 to iy_1 is $\rho(iy_0, iy_1) = \left|\log(y_1/y_0)\right|$. Thus

$$\begin{aligned} \cosh^2\left(\frac{\rho(iy_0, iy_1)}{2}\right) &= \left((y_1/y_0)^{1/2} + (y_0/y_1)^{1/2}\right)^2 / 4 \\ &= \frac{(y_0 + y_1)^2}{4y_0y_1}. \end{aligned}$$

This is the second formula of Exercise 3.1.1(1) in the case where $z = iy_0$ and $w = iy_1$.

Lemma 3.3.2 *Hyperbolic circles in the Poincaré disc or the upper half plane are Euclidean circles, but the centre is not necessarily the Euclidean centre. In particular, the hyperbolic circle in the Poincaré disc with centre* 0 *and hyperbolic radius* r *is the Euclidean circle with centre* 0 *and radius* $\tanh(r/2)$.

Proof: Suppose that z lies in the Poincaré disc and is a distance ρ from the origin. From (3.8) we have

$$\cosh^2(r/2) = \cosh^2\left(\frac{\rho(0,z)}{2}\right) = \frac{1}{1 - |z|^2}.$$

Rearranging, we find that $|z| = \tanh(r/2)$. Therefore the locus of such points is a Euclidean circle of radius $\tanh(r/2)$.

Since Möbius transformations map circles to circles we find that all other hyperbolic circles are also Euclidean circles. However, when a Möbius transformation maps one circle to another, it does not map the centre to the centre unless it fixes ∞. $\square$

Cartesian coordinates on the upper half plane have the following interpretation in terms of hyperbolic geometry. The vertical lines where x is constant are geodesics. The horizontal lines where y is constant are horocycles. We define the *horocycle* H_t based at ∞ of *height* t and the *horodisc* B_t based at ∞ of height t by

$$H_t = \left\{ z = x + iy \in \mathbf{H}^2 \ : \ y = t \right\}, \qquad B_t = \left\{ z = x + iy \in \mathbf{H}^2 \ : \ y > t \right\}.$$

Horocycles naturally carry the Euclidean metric, but this is not intrinsically defined.

We can use Möbius transformations to define horocycles and horodiscs based at any point of either the upper half plane or the Poincaré disc. These are Euclidean circles tangent to the boundary. Their height is not intrinsically defined.

Möbius transformations in $\mathrm{PSL}(2,\mathbb{R})$ map $\widehat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ to itself. Möbius transformations fixing ∞ are either Euclidean isometries (necessarily translations) or dilations fixing a finite point of the ideal boundary. In either case, these maps preserve the Cartesian grid comprising geodesics with one endpoint at ∞ and horocycles based at ∞. Furthermore, if $A(x) \in \mathrm{Isom}(\mathbb{R})$ then $A(x) = \pm x + t$ for some $t \in \mathbb{R}$ and if $A(x) \in \mathrm{Isom}^+(\mathbb{R})$ then $A(x) = x + t$ for some $t \in \mathbb{R}$. Such maps clearly satisfy $\bigl|A(x) - A(y)\bigr| = |x - y|$. Moreover, they preserve each horocycle.

Likewise the (loxodromic) dilation $D(x) = d^2 x$ where $d \in \mathbb{R} - \{0, \pm 1\}$ satisfies $\bigl|D(x) - D(y)\bigr| = d^2|x - y|$. Any loxodromic map fixing ∞ is the conjugate of this one by a translation. That is $D(x) = d^2 x + t$ for $t \in \mathbb{R}$. Once again we have $\bigl|D(x) - D(y)\bigr| = d^2|x - y|$. Moreover, $D(x)$ maps the horocycle H_t of height $t > 0$ based at ∞ to the horocycle $H_{d^2 t}$ of height $d^2 t$ based at ∞.

Finally we have $R(x) = -1/x$ which satisfies

$$\bigl|R(x) - R(y)\bigr| = \left| \frac{-1}{x} - \frac{-1}{y} \right| = \frac{|x - y|}{|x|\,|y|}.$$

This enables us to give the dynamical action of hyperbolic isometries.

Proposition 3.3.3 *Let* $A(z) \in \mathrm{PSU}(1,1)$ *or* $\mathrm{PSL}(2,\mathbb{R})$.

(i) If $A(z)$ *is loxodromic then it preserves the geodesic whose endpoints are the fixed points. Moreover, it translates points of this geodesic by a constant hyperbolic distance. Thus one fixed point is attractive and the other is repulsive.*

(ii) If $A(z)$ *is parabolic then it preserves each horocycle based at the fixed point. It translates points of this horocycle by a constant Euclidean distance.*

(iii) If $A(z)$ *is elliptic then it preserves hyperbolic circles centred at the fixed point and rotates them by a constant angle.*

Chapter 4

The Klein model of the hyperbolic plane

4.1 Quadratic forms of signature $(2,1)$

Let $\mathbb{R}^{2,1}$ be the real vector space $\mathbb{R}^3$ equipped with a non-degenerate quadratic form of signature $(2,1)$. We consider the following symmetric matrices of signature $(2,1)$.

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad H_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{4.1}$$

We define quadratic forms by

$$(\mathbf{x},\mathbf{y})_1 = \mathbf{y}^t H_1 \mathbf{x} = x_1y_1 + x_2y_2 - x_3y_3, \tag{4.2}$$
$$(\mathbf{x},\mathbf{y})_2 = \mathbf{y}^t H_2 \mathbf{x} = x_1y_3 + x_2y_2 + x_3y_1. \tag{4.3}$$

In what follows we think of $\mathbb{R}^{2,1}$ to be $\mathbb{R}^3$ equipped with one of these two forms. In fact, we mainly work with H_1. It is an interesting (non-assessed) exercise to work out the details for H_2. As before, we define

$$\begin{aligned} V_- &= \big\{\mathbf{x} \in \mathbb{R}^{2,1} \ : \ (\mathbf{x},\mathbf{x}) < 0\big\}, \\ V_0 &= \big\{\mathbf{x} \in \mathbb{R}^{2,1} - \{0\} \ : \ (\mathbf{x},\mathbf{x}) = 0\big\}, \\ V_+ &= \big\{\mathbf{x} \in \mathbb{R}^{2,1} \ : \ (\mathbf{x},\mathbf{x}) > 0\big\}. \end{aligned}$$

Again $\mathbf{x} \in \mathbb{R}^{2,1}$ is said to be *negative*, *null* or *positive* according to whether it is in V_-, V_0 or V_+ respectively.

We define a projection map $\mathbb{P} : \mathbb{R}^{2,1} - \{0\} \longrightarrow \mathbb{RP}^2$ in the usual way. If $x_3 \neq 0$ then

$$\mathbb{P} : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} x_1/x_3 \\ x_2/x_3 \end{pmatrix} \in \mathbb{R}^2 \subset \mathbb{RP}^2.$$

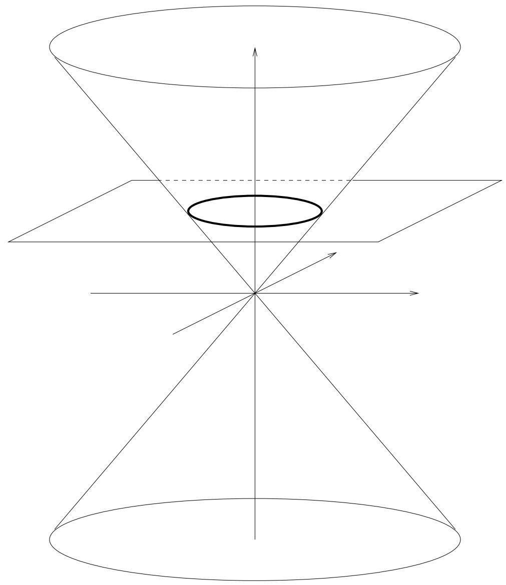

Figure 4.1: *The light cone in* $\mathbb{R}^{2,1}$. *Points of* V_- *lie inside the cone. The Klein disc, as shown, is the intersection of* V_- *with the plane with* $x_3 = 1$.

We define a partial inverse to this map called the *standard lift* by choosing the section where $x_3 = 1$. That is, given $x \in \mathbb{R}^2$ we define a point $\mathbf{x}$ in $\mathbb{R}^{2,1}$ by setting the first two entries of $\mathbf{x}$ to be the entries of x and the third to be 1.

We define the *Klein disc* model of the hyperbolic plane to be $\mathbb{P}V_-$ for the first quadratic form. It is easy to see that if $x \in \mathbb{R}^2$ and has standard lift $\mathbf{x}$, then x is in the Klein disc $\mathbf{H}^2$ if and only if

$$(\mathbf{x}, \mathbf{x})_1 = |x|^2 - 1 < 0.$$

In other words $|x|^2 = x_1^2 + x_2^2 < 1$, see Figure 4.1.

Then the metric is given by

$$ds^2 = \frac{-1}{(\mathbf{x}, \mathbf{x})_1^2} \det \begin{pmatrix} (\mathbf{x}, \mathbf{x})_1 & (d\mathbf{x}, \mathbf{x})_1 \\ (\mathbf{x}, d\mathbf{x})_1 & (d\mathbf{x}, d\mathbf{x})_1 \end{pmatrix} = \frac{dx_1^2 + dx_2^2 - (x_1\, dx_2 - x_2\, dx_1)^2}{(1 - x_1^2 - x_2^2)^2}. \tag{4.4}$$

Note that we omit the constant 4 that appears in (3.6). This ensures that the curvature is the same as for the Poincaré models of the hyperbolic plane. Also

$$\cosh^2\Big(\rho\big((x_1, x_2), (y_1, y_2)\big)\Big) = \frac{(\mathbf{x}, \mathbf{y})_1 (\mathbf{y}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})_1 (\mathbf{y}, \mathbf{y})} = \frac{(x_1 y_1 + x_2 y_2 - 1)^2}{(1 - x_1^2 - x_2^2)(1 - y_1^2 - y_2^2)}.$$

For the second quadratic form, we obtain a region bounded by a parabola:

$$(\mathbf{x}, \mathbf{x})_2 = 2x_1 + x_2^2 < 0.$$

In order to discuss the boundary of this parabola we must extend the definition of $\mathbb{P}$ so that

$$\mathbb{P} : \mathbf{x} \longmapsto \infty$$

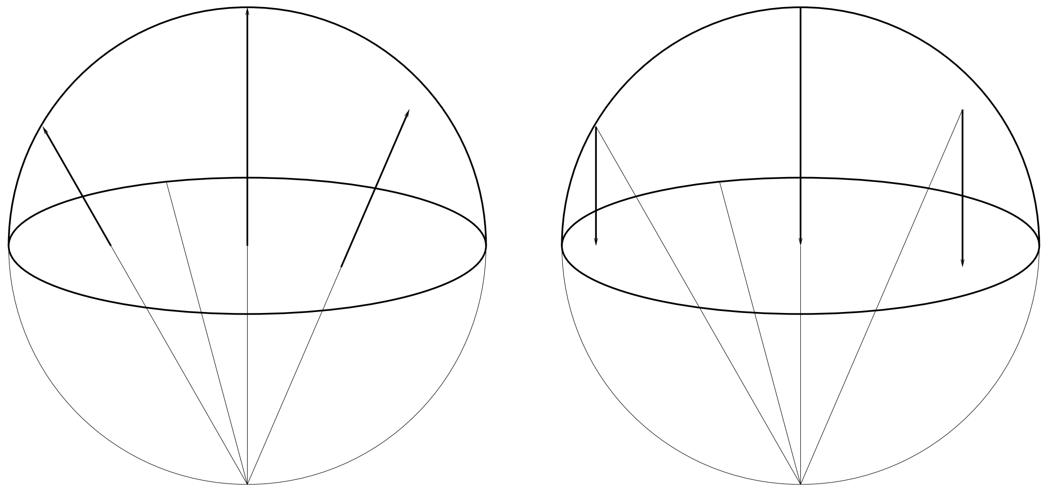

Figure 4.2: *Stereographic projection from the Poincaré disc to the upper hemisphere and vertical projection from the hemisphere to the Klein disc.*

when $x_2 = x_3 = 0$.

The map from the Poincaré disc to the Klein disc is given by first stereographically projecting to the unit sphere in $\mathbb{R}^3$ and then orthogonally projecting onto the first two coordinates, see Figure 4.2. Stereographic projection of the Poincaré disc onto the upper hemisphere is given by:

$$z = x + iy \longmapsto \left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}, \frac{1-x^2-y^2}{1+x^2+y^2} \right).$$

Vertical projection from the upper hemisphere to the Klein disc is given by:

$$\left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2}, \frac{1-x^2-y^2}{1+x^2+y^2} \right) \longmapsto \left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2} \right)$$

The composition of these maps is:

$$z = x + iy \longmapsto \left(\frac{2x}{1+x^2+y^2}, \frac{2y}{1+x^2+y^2} \right). \tag{4.5}$$

We remark that under both of these maps the points on the unit circle are fixed. The map from the Klein disc to the Poincaré disc is given by

$$(x_1, x_2) \longmapsto \frac{x_1 + ix_2}{1 + \sqrt{1 - x_1^2 - x_2^2}}. \tag{4.6}$$

Lemma 4.1.1 *The maps (4.5) and (4.6) are inverses of each other.*

Proof: If

$$x_1 = \frac{2x}{1+x^2+y^2}, \quad x_2 = \frac{2y}{1+x^2+y^2}$$

then
$$\sqrt{1-x_1^2-x_2^2}=\frac{1-x^2-y^2}{1+x^2+y^2}.$$
Thus
$$\frac{x_1+ix_2}{1+\sqrt{1-x_1^2-x_2^2}}=\frac{\frac{2x}{1+x^2+y^2}+i\frac{2y}{1+x^2+y^2}}{1+\frac{1-x^2-y^2}{1+x^2+y^2}}=x+iy.$$
□

Proposition 4.1.2 *The map (4.5) induces an isometry from the Poincaré disc to the Klein disc, that is between the metrics (3.8) and (4.4).*

Proof: If
$$z=\frac{x_1+ix_2}{1+\sqrt{1-x_1^2-x_2^2}}$$
then
$$dz=\frac{\sqrt{1-x_1^2-x_2^2}\left(1+\sqrt{1-x_1^2-x_2^2}\right)(dx_1+idx_2)+(x_1+ix_2)(x_1dx_1+x_2dx_2)}{\sqrt{1-x_2^2-x_2^2}\left(1+\sqrt{1-x_1^2-x_2^2}\right)^2}$$
Therefore
$$dzd\overline{z}=\frac{(1-x_1^2-x_2^2)(dx_1^2+dx_2^2)+(x_1dx_1+x_2dx_2)^2}{(1-x_1^2-x_2^2)\left(1+\sqrt{1-x_1^2-x_2^2}\right)^2}.$$
Also
$$1-|z|^2=\frac{2\sqrt{1-x_1^2-x_2^2}}{1+\sqrt{1-x_1^2-x_2^2}}.$$
Substituting these expressions in (3.8) gives (4.4). □

Exercise 4.1.1 *Using the map (4.5) show that geodesics in the Klein disc are Euclidean line segments. Similarly, show that hyperbolic circles and horocycles in the Klein disc are ellipses. [Do not calculate: argue using geometry. Note that stereographic projection maps arcs of circles to arcs of circles and also preserves angles.]*

4.2 Isometries

Matrices preserving a quadratic form are called *orthogonal*. In particular if H is a real Hermitian form then unitary matrices preserving H that have real entries are orthogonal. The inverse of an orthogonal matrix A is $H^{-1}A^tH$.

We let $\mathrm{O}(2,1)$ be the group of orthogonal matrices preserving a quadratic form of signature $(2,1)$. Let $\mathrm{SO}(2,1)$ be the subgroup with determinant $+1$. If this form is H_1 or H_2 then we write $SO(H_1)$ or $\mathrm{SO}(H_2)$ respectively.

A matrix A in $\mathrm{SO}(H_1)$ acts on the Klein disc as follows. First take the standard lift $\mathbf{x}$ of $x \in \mathbf{H}^2$. Then A acts on $\mathbf{x}$ by left multiplication. Finally we projectivise. That is

$$\begin{aligned} A(x) &= \mathbb{P}A\mathbf{x} \\ &= \mathbb{P}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} \\ &= \mathbb{P}\begin{pmatrix} ax_1 + bx_2 + c \\ dx_1 + ex_2 + f \\ gx_1 + hx_2 + j \end{pmatrix} \\ &= \begin{pmatrix} (ax_1 + bx_2 + c)/(gx_1 + hx_2 + j) \\ (dx_1 + ex_2 + f)/(gx_1 + hx_2 + j) \end{pmatrix}. \end{aligned}$$

This is a generalisation of a Möbius transformation.

This map $A \longrightarrow A(x)$ is a surjective homomorphism. Its kernel is the group of non-zero real multiples of the identity with unit determinant. Since there is a unique real cube root of unity, this means that this map is an isomorphism.

One advantage of the Klein disc model is that $\mathrm{SO}(2,1)$, unlike $\mathrm{PSL}(2,\mathbb{R})$, contains both orientation preserving and orientation reversing isometries.

Lemma 4.2.1 *The isometry of the Poincaré disc given by $z \longmapsto -\overline{z}$ corresponds to the following map in* $\mathrm{SO}(2,1)$*:*

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Proof: Since $-(\overline{x+iy}) = -x + iy$ we see using (4.5) and (4.6) that the map $z \longmapsto -\overline{z}$ corresponds to the map $(x_1, x_2) \longmapsto (-x_1, x_2)$ of the Klein disc. The action of the above matrix is

$$\mathbb{P}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \mathbb{P}\begin{pmatrix} x_1 \\ -x_2 \\ -1 \end{pmatrix} = \begin{pmatrix} -x_1 \\ x_2 \\ 1 \end{pmatrix}.$$

$\square$

Exercise 4.2.1 *Show that the map (4.5) induces the following homomorphism from* $\mathrm{SU}(1,1)$ *to* $\mathrm{SO}(2,1)$

$$\begin{pmatrix} \cosh(\lambda) & \sinh(\lambda) \\ \sinh(\lambda) & \cosh(\lambda) \end{pmatrix} \longmapsto \begin{pmatrix} \cosh(2\lambda) & 0 & \sinh(2\lambda) \\ 0 & 1 & 0 \\ \sinh(2\lambda) & 0 & \cosh(2\lambda) \end{pmatrix},$$

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \longmapsto \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) & 0 \\ \sin(2\theta) & \cos(2\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

In other words, if $\Phi : z \longmapsto (x_1, x_2)$ is the map (4.5) and $A \in \mathrm{SL}(2,\mathbb{R})$ is the matrix on the left, show that $B\Phi(z) = \Phi\big(A(z)\big)$ where $B \in \mathrm{SO}(2,1)$ is the map on the right.

Using the proof of Lemma 3.2.1, deduce that if $A \in \mathrm{SU}(1,1)$ is mapped to $B \in \mathrm{SO}(2,1)$ by the homomorphism you have just constructed, then $\mathrm{tr}(B) = \mathrm{tr}^2(A) - 1$.

Proposition 4.2.2 *Let $A \in \mathrm{SO}(2,1)$ and let $\tau = \mathrm{tr}(A)$. Then the characteristic polynomial of A is*

$$\mathrm{ch}_A(t) = t^3 - \tau t^2 + \tau t - 1$$

and the eigenvalues of A are $t = 1$ and

$$t = \frac{\tau - 1 \pm \sqrt{(\tau-3)(\tau+1)}}{2}.$$

Proof: Suppose that A has eigenvalues t_1, t_2 and t_3. Since $\det(A) = 1$ we have $t_1t_2t_3 = 1$. Also, by Lemma 2.3.1, $\overline{t}_j^{-1}$ is an eigenvalue of A and so $\overline{t}_1^{-1}$, $\overline{t}_2^{-1}$, $\overline{t}_3^{-1}$ is a permutation of t_1, t_2, t_3. Hence

$$t_1t_2 + t_2t_3 + t_3t_1 = t_3^{-1} + t_1^{-1} + t_2^{-1} = \overline{t}_1 + \overline{t}_2 + \overline{t}_3 = \overline{\tau} = \tau.$$

We have used the fact that τ is real in the last step. Hence the characteristic polynomial of A is

$$\begin{aligned}
\mathrm{ch}_A(t) &= (t-t_1)(t-t_2)(t-t_3) \\
&= t^3 - (t_1+t_2+t_3)t^2 + (t_1t_2+t_2t_3+t_3t_1)t - t_1t_2t_3 \\
&= t^3 - \tau t^2 + \tau t - 1 \\
&= (t-1)\big(t^2 - (\tau-1)t + 1\big).
\end{aligned}$$

Notice that $t = 1$ is always a root of this polynomial. The rest of the result follows simply. □

The following result is the analogue of Proposition 3.2.6 but there are now more categories since $\mathrm{SO}(2,1)$ contains the orientation reversing isometries of $\mathbf{H}^2$.

Proposition 4.2.3 *Let $A \in \mathrm{SO}(2,1)$ and write $\tau = \mathrm{tr}(A)$. Then*

(i) $A(z)$ is loxodromic if $\tau > 3$;

(ii) $A(z)$ is parabolic or the identity if $\tau = 3$;

(iii) $A(z)$ is elliptic if $-1 < \tau < 3$;

(iv) $A(z)$ is elliptic or a reflection if $\tau = -1$;

(v) $A(z)$ is a glide reflection if $\tau < -1$.

Chapter 5

Hyperbolic 3-space

5.1 Isometries

We begin by discussing the matrix group $\mathrm{SL}(2,\mathbb{C})$ and the corresponding group of Möbius transformations $\mathrm{PSL}(2,\mathbb{C})$. The group $\mathrm{SL}(2,\mathbb{C})$ is defined by

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

Such matrices naturally act by left multiplication on vectors in $\mathbb{C}^2$. Arguing as in Section 3.2 the matrix $A \in \mathrm{SL}(2,\mathbb{C})$ acts on $\widehat{\mathbb{C}} = \mathbb{CP}^1$ via the Möbius transformation $A(z) = (az+b)/(cz+d)$. The collection of such Möbius transformations will be denoted $\mathrm{PSL}(2,\mathbb{C})$. These Möbius transformations act naturally on the Riemann sphere. The group $\mathrm{PSL}(2,\mathbb{C})$ contains $\mathrm{PSU}(1,1)$ and $\mathrm{PSL}(2,\mathbb{R})$ as subgroups.

Proposition 5.1.1 *Let $A \in \mathrm{SL}(2,\mathbb{C})$ have real trace. Then A is conjugate to an element of* $\mathrm{SL}(2,\mathbb{R})$.

Proof: Suppose first that A has distinct eigenvalues. Then we can conjugate A to diagonal form and the eigenvalues are the diagonal entries. There are two cases. If $\mathrm{tr}(A) > 2$ or $\mathrm{tr}(A) < -2$ then A has eigenvalues t and $1/t$ for some real number with $t > 1$ or $t < -1$ respectively. This is clearly in $\mathrm{SL}(2,\mathbb{R})$.

If $-2 < \mathrm{tr}(A) < 2$ then A has eigenvalues $e^{i\theta}$ and $e^{-i\theta}$ for some $\theta \in (0,\pi)$. Hence A lies in $\mathrm{SU}(1,1)$. In fact we can conjugate A to $\mathrm{SL}(2,\mathbb{R})$ as follows:

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Finally suppose that A has a repeated eigenvalue, necessarily ± 1. If A is diagonalisable it is necessarily $\pm I$. Otherwise, we can conjugate A to its Jordan normal form

$$\begin{pmatrix} \pm 1 & 1 \\ 0 & \pm 1 \end{pmatrix}$$

which clearly lies in $\mathrm{SL}(2,\mathbb{R})$. □

Clearly elements of $\mathrm{SL}(2,\mathbb{C})$ with non-real trace cannot be conjugated to elements of $\mathrm{SL}(2,\mathbb{R})$. However, we can write them as a product of two such maps.

Proposition 5.1.2 *Let $A \in \mathrm{SL}(2,\mathbb{C})$ have non-real trace. Then A can be be written as a product of two maps each of which is conjugate to an element of* $\mathrm{SL}(2,\mathbb{R})$.

Proof: Let t be an eigenvalue of A. Then $1/t$ is necessarily the other eigenvalue. Since $t+1/t$ is not real, we can write t as $t = re^{i\theta}$ where $r > 1$ and $\theta \in (0,\pi)$. Clearly A has distinct eigenvalues. Therefore A can be diagonalised. Thus A is conjugate to

$$\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} = \begin{pmatrix} re^{i\theta} & 0 \\ 0 & r^{-1}e^{-i\theta} \end{pmatrix} = \begin{pmatrix} r & 0 \\ 0 & r^{-1} \end{pmatrix} \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

This proves the result. □

We say that A $\in \mathrm{SL}(2,\mathbb{C})$ is *simple* if it is conjugate to an element of $\mathrm{SL}(2,\mathbb{R})$. Also A $\in \mathrm{SL}(2,\mathbb{C})$ is said to be *k-simple* if it can be written as a product of k simple matrices and no fewer.

Theorem 5.1.3 *Two non-trivial maps A and B in $\mathrm{SL}(2,\mathbb{C})$ are conjugate if and only if their traces are equal. Moreover, given $A \in \mathrm{SL}(2,\mathbb{C})$ not $\pm I$ then write $\tau = \mathrm{tr}(A)$. We have*

(a) if τ is real then A is 1–simple, and

- *(i) if $-2 < \tau < 2$ then A is elliptic;*
- *(ii) if $\tau = \pm 2$ then A is parabolic;*
- *(iii) if $\tau > 2$ or $\tau < -2$ then A is loxodromic.*

(b) If τ is not real then A is 2–simple and loxodromic.

5.2 The Poincaré extension

We want to describe the action of complex Möbius transformations on upper half 3-space. In order to do so we introduce the imaginary unit j, satisfying $ij = -ji$ and $j^2 = -1$. (This is the first hint about how we will use quaternions and, later, Clifford algebras. We will say more about this in later sections.) Points in the upper half space will be written $z + tj$ where $z \in \mathbb{C}$ and $t \in \mathbb{R}_+$:

$$\mathbf{H}^3 = \left\{ z + tj \ : \ z \in \mathbb{C},\, t \in \mathbb{R}_+ \right\}.$$

In fact, these are *horospherical coordinates*: we have foliated $\mathbf{H}^3$ by horospheres H_t centred at ∞:

$$H_t = \left\{ z + tj \ : \ z \in \mathbb{C} \right\}. \tag{5.1}$$

Each of these horospheres is parametrised by $z \in \mathbb{C}$, and the parameter t indexes the *height* above $\widehat{\mathbb{C}} = \partial \mathbf{H}^3$ of each horosphere.

We are now interested in how a Möbius transformation acts on $z + tj$. Because $ij \neq ji$ we have to be careful in which order we multiply numbers involving both i and j. Therefore we choose to write our Möbius transformation as

$$A(z+tj) = \big(a(z+tj)+b\big)\big(c(z+tj)+d\big)^{-1}$$

where a, b, c and d are all complex numbers satisfying $ad - bc = 1$. We now have to interpret what $\big(c(z+tj)+d\big)^{-1}$ means. It should be that number which, when multiplied by $\big(c(z+tj)+d\big)$ gives 1. We claim that it is a real multiple of $(\overline{z}-tj)\overline{c}+\overline{d}$. In order to see this, observe first that, if $z = x + iy$ then

$$j\overline{z} = j(x-yi) = xj - yji = xj + yij = (x+yi)j = zj.$$

Therefore

$$\begin{aligned}
&\big(c(z+tj)+d\big)\big((\overline{z}-tj)\overline{c}+\overline{d}\big) \\
&\quad = c(z+tj)(\overline{z}-tj)\overline{c} + c(z+tj)\overline{d} + d(\overline{z}-tj)\overline{c} + d\overline{d} \\
&\quad = c(|z|^2 - ztj + ztj + t^2)\overline{c} + cz\overline{d} + cdtj + d\overline{zc} - dctj + |d|^2 \\
&\quad = |cz+d|^2 + |c|^2 t^2.
\end{aligned}$$

Therefore

$$\big(c(z+tj)+d\big)^{-1} = \frac{(\overline{z}-tj)\overline{c}+\overline{d}}{|cz+d|^2+|c|^2t^2}.$$

(We are allowed to divide by non-zero real numbers since they commute with i and j and hence the left inverse and right inverse of a real number are the same.)

Proposition 5.2.1 *The group* $\mathrm{PSL}(2,\mathbb{C})$ *maps* $\mathbf{H}^3 = \{z + tj \ : \ z \in \mathbb{C},\ t \in \mathbb{R}_+\}$ *to itself.*

Proof: We have

$$\begin{aligned}
A(z+tj) &= \big(a(z+tj)+b\big)\big(c(z+tj)+d\big)^{-1} \\
&= \frac{\big(a(z+tj)+b\big)\big((\overline{z}-tj)\overline{c}+\overline{d}\big)}{|cz+d|^2+|c|^2t^2} \\
&= \frac{a(z+tj)(\overline{z}-tj)\overline{c} + a(z+tj)\overline{d} + b(\overline{z}-tj)\overline{c} + b\overline{d}}{|cz+d|^2+|c|^2t^2} \\
&= \frac{a(|z|^2 - ztj + ztj + t^2)\overline{c} + az\overline{d} + adtj + b\overline{zc} - bctj + b\overline{d}}{|cz+d|^2+|c|^2t^2} \\
&= \frac{(az+b)\overline{(cz+d)} + a\overline{c}t^2 + tj}{|cz+d|^2+|c|^2t^2}
\end{aligned}$$

which has the required form. □

Consider the matrix

$$H_2 = \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix}. \tag{5.2}$$

It is clear that H_2^2 is the identity so H_2 is its own inverse. If we were to extend the definition of complex conjugation so that $\overline{z+tj} = \overline{z} - tj$ (so $\overline{j} = -j$) we could then extend the definition of the Hermitian conjugate of a matrix whose entries lie in $\mathbb{C} \oplus j\mathbb{R}$ accordingly. By this definition, H_2 is a Hermitian matrix.

Proposition 5.2.2 *Let H_2 be given by (5.2). For any $A \in \mathrm{SL}(2,\mathbb{C})$ we have $A^*H_2A = H_2$.*

Proof: Using $\overline{z}j = jz$ we have

$$\begin{aligned}
A^*H_2A &= \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix} \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\
&= \begin{pmatrix} \overline{c}j & -\overline{a}j \\ \overline{d}j & -\overline{b}j \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\
&= \begin{pmatrix} jc & -ja \\ jd & -jb \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\
&= \begin{pmatrix} 0 & -j \\ j & 0 \end{pmatrix} = H_2.
\end{aligned}$$

as required. □

If we allow the definition of unitary in Section 2.2 to be extended to Hermitian forms involving $z + tj$ then we have:

Corollary 5.2.3 *Each $A \in \mathrm{SL}(2,\mathbb{C})$ is unitary with respect to H_2.*

The Hermitian matrix H_2 given by (5.2) defines a Hermitian form $\langle \cdot, \cdot \rangle_2$. This is done in an analogous manner to the construction for complex Hermitian forms, as explained in Section 2.1. We give more details in the next section. Using this form the definition of the hyperbolic metric on $\mathbf{H}^3$ is given by the formulae (3.6) and (3.7).

Exercise 5.2.1 *Let*

$$\mathbf{z} + \mathbf{tj} = \begin{pmatrix} z + tj \\ 1 \end{pmatrix}$$

be the standard lift of $z + tj \in \mathbf{H}^3$. Let $\langle \cdot, \cdot \rangle$ be the Hermitian form defined by H_2. Show that

$$\langle \mathbf{z} + \mathbf{tj}, \mathbf{w} + \mathbf{sj} \rangle = j(z - w) - t - s.$$

Use this form to show that on $\mathbf{H}^3$ *we have*

$$ds^2 = \frac{dz\,d\overline{z} + dt^2}{t^2}$$

and

$$\cosh^2\Big(\frac{\rho(z+tj, w+sj)}{2}\Big) = \frac{|z+tj-w+sj|^2}{4ts} = \frac{|z-w|^2+(t+s)^2}{4ts}.$$

5.3 The geometry of hyperbolic 3-space

In this section we consider geometric objects in $\mathbf{H}^3$ analogous to those we found in the hyperbolic plane. The properties of these objects and their construction will generalise directly from $\mathbf{H}^2$ to $\mathbf{H}^3$.

Proposition 5.3.1 *Geodesics in* $\mathbf{H}^3$ *are lines or semicircles orthogonal to the ideal boundary* $\widehat{\mathbb{C}}$*.*

Proof: We can show that the j-axis is a geodesic as in Proposition 3.3.1. This is a line orthogonal to $\widehat{\mathbb{C}}$, as is its image under any $A(z) \in \mathrm{PSL}(2,\mathbb{C})$ with $c = 0$. Likewise, if $A(z) \in \mathrm{PSL}(2,\mathbb{C})$ has $d = 0$ then $A(0) = \infty$ and so $A(z)$ sends the imaginary axis to a line orthogonal to $\widehat{\mathbb{C}}$.

Thus we assume that $A(z)$ has c, $d \neq 0$ and we want to show that $\{A(tj) \,:\, t \in \mathbb{R}\}$ is a semicircle orthogonal to $\widehat{\mathbb{C}}$. Now $A(0) = bd^{-1}$ and $A(\infty) = ac^{-1}$. These points will be where our semicircle intersects $\widehat{\mathbb{C}}$. So if our hypothesis is correct, the semicircle will have centre $(ac^{-1}+bd^{-1})/2$ and radius $|ac^{-1}-bd^{-1}|/2 = 1/2|cd|$. Thus

$$\begin{aligned}
A(tj) - (ac^{-1}+bd^{-1})/2 &= \frac{b\overline{d} + a\overline{c}t^2 + tj}{|d|^2+|c|^2t^2} - \frac{ac^{-1}+bd^{-1}}{2} \\
&= \frac{2b\overline{d} + 2a\overline{c}t^2 + 2tj - ac^{-1}|d|^2 - a\overline{c}t^2 - b\overline{d} - bd^{-1}|c|^2t^2}{2\big(|d|^2+|c|^2t^2\big)} \\
&= \frac{(-ac^{-1}+bd^{-1})\big(|d|^2 - 2cdtj - |c|^2t^2\big)}{2\big(|d|^2+|c|^2t^2\big)} \\
&= \frac{-ac^{-1}+bd^{-1}}{2} \cdot \frac{(d-ctj)(\overline{d}-tj\overline{c})}{(d+ctj)(\overline{d}-tj\overline{c})}.
\end{aligned}$$

Therefore

$$\big|A(tj) - (ac^{-1}+bd^{-1})/2\big| = |ac^{-1}-bd^{-1}|/2$$

as claimed. □

Lemma 5.3.2 *The hyperbolic sphere centred at* (z_0+t_0j) *with hyperbolic radius* r_0 *is a Euclidean sphere with centre* $z_0 + t_0\cosh(r_0)j$ *and radius* $t_0\sinh(r_0)$*.*

Proof: A point $z+tj$ lies on this sphere if and only if

$$\cosh^2(r_0/2) = \frac{|z-z_0|^2 + (t+t_0)^2}{4tt_0}.$$

Rearranging, we obtain

$$\begin{aligned} 0 &= |z-z_0|^2 + t^2 - 2\big(2\cosh^2(r_0/2) - 1\big)tt_0 + t_0^2 \\ &= |z-z_0|^2 + t^2 - 2\cosh(r_0)tt_0 + t_0^2 \\ &= |z-z_0|^2 + \big(t - t_0\cosh(r_0)\big)^2 - t_0^2\sinh^2(r_0). \end{aligned}$$

$\square$

We can generalise horocycles and horodiscs to the Poincaré extension as in (5.1). These are called horospheres and horoballs respectively. The *horosphere* H_u based at ∞ of *height* u and the *horoball* B_u based at ∞ of *height* u are defined by

$$H_u = \Big\{ z+tj \in \mathbf{H}^3 \ : \ t = u \Big\}, \qquad B_u = \Big\{ z+tj \in \mathbf{H}^3 \ : \ t > u \Big\}.$$

As before, horospheres naturally carry the Euclidean metric. Horospheres and horoballs based at other points of $\partial\mathbf{H}^3 = \widehat{\mathbb{C}}$ are defined as the images of these ones under Möbius transformations.

Exercise 5.3.1 *Show that elliptic and parabolic elements of* $\mathrm{PSL}(2,\mathbb{C})$ *that fix* ∞ *map each horosphere* H_u *for* $u \geq 0$ *to itself and act there as Euclidean isometries. Show that loxodromic maps fixing* ∞ *map horocycles based at* ∞ *to distinct horocycles based at* ∞.

5.4 The Klein model of hyperbolic 3-space.

Let $\mathbb{R}^{3,1}$ be the real vector space $\mathbb{R}^4$ equipped with a non-degenerate quadratic form of signature $(3,1)$. The forms we choose to work with are given by the following symmetric matrices:

$$H_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad H_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

We define quadratic forms by

$$(\mathbf{x},\mathbf{y})_1 = \mathbf{y}^t H_1 \mathbf{x} = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4, \tag{5.3}$$
$$(\mathbf{x},\mathbf{y})_2 = \mathbf{y}^t H_2 \mathbf{x} = x_1y_4 + x_2y_2 + x_3y_3 + x_4y_1. \tag{5.4}$$

Then defining V_+, V_0, V_- and $\mathbb{P}$ as before, we define the *Klein ball* model of the hyperbolic plane to be $\mathbb{P}V_-$ for the first quadratic form. Then the metric is given by

$$\begin{aligned} ds^2 &= \frac{-1}{(\mathbf{x},\mathbf{x})_1^2}\det\begin{pmatrix} (\mathbf{x},\mathbf{x})_1 & (d\mathbf{x},\mathbf{x})_1 \\ (\mathbf{x},d\mathbf{x})_1 & (d\mathbf{x},d\mathbf{x})_1 \end{pmatrix} \\ &= \frac{(1-x_1^2-x_2^2-x_3^2)(dx_1^2+dx_2^2+dx_3^3)+(x_1\,dx_1+x_2\,dx_2+x_3\,dx_3)^2}{(1-x_1^2-x_2^2-x_3^2)^2}. \end{aligned}$$

We can then mimic the calculations we did in two dimensions to obtain an isometry between the upper half space and the Klein ball. We can define the special orthogonal group

$$\mathrm{SO}(3,1)=\left\{A\in\mathrm{SL}(4,\mathbb{R})\ :\ (A\mathbf{x},A\mathbf{y})=(\mathbf{x},\mathbf{y})\text{ for all }\mathbf{x},\,\mathbf{y}\in\mathbb{R}^{3,1}\right\}.$$

We can then show that isometries of the Klein ball lie in

$$\mathrm{PSO}(3,1)=\mathrm{SO}(3,1)/\{\pm I\}.$$

The details are left as a exercise.

Chapter 6

Quaternionic Möbius transformations

6.1 The quaternions

Let $\mathbb{H}$ denote the division ring of real quaternions. Elements of $\mathbb{H}$ have the form $z = z_0 + z_1 i + z_2 j + z_3 k$ where $z_i \in \mathbb{R}$ and

$$i^2 = j^2 = k^2 = ijk = -1.$$

Let $\overline{z} = z_0 - z_1 i - z_2 j - z_3 k$ be the *conjugate* of z, and define the *modulus* of z to be:

$$|z| = \sqrt{\overline{z}z} = \sqrt{z_0^2 + z_1^2 + z_2^2 + z_3^2}.$$

We define the *real part* of z to be $\mathrm{Re}(z) = (z + \overline{z})/2 = z_0$ and the *imaginary part* of z to be $\mathrm{Im}(z) = (z - \overline{z})/2 = z_1 i + z_2 j + z_3 k$. Note that the imaginary part of a quaternion is *not* a real number (unlike with complex numbers). Also $z^{-1} = \overline{z}|z|^{-2}$ is the *inverse* of z.

Exercise 6.1.1 *Let z and w be quaternions. Show that $\overline{(zw)} = \overline{w}\,\overline{z}$. Deduce that $|zw| = |w|\,|z|$ and that $(zw)^{-1} = w^{-1}z^{-1}$.*

Lemma 6.1.1 *For $z \in \mathbb{H}$, we let $\mathbb{R}(z)$ be the smallest sub-ring of $\mathbb{H}$ containing $\mathbb{R}$ and z. Then we have:*

(i) If $\mathrm{Im}(z) = 0$ then $\mathbb{R}(z) = \mathbb{R}$.

(ii) If $\mathrm{Im}(z) \neq 0$ then $\mathbb{R}(z) = \mathbb{R} \oplus \mathbb{R}z$.

Moreover, in either case $\mathbb{R}(z)$ is a field (that is every non-zero element has an inverse in $\mathbb{R}(z)$ and $\mathbb{R}(z)$ is abelian).

Proof: In the first case, it is clear that $z \in \mathbb{R}$ and so $\mathbb{R}$ is the smallest sub-ring of $\mathbb{H}$ containing z. This is a field.

In the second case, it is clear that $\mathbb{R}(z)$ must contain

$$\mathbb{R} \oplus \mathbb{R}z = \Big\{ a + bz \ : \ a,\, b \in \mathbb{R} \Big\}.$$

It is clear that $\mathbb{R} \oplus \mathbb{R}z$ is an abelian group under addition. We must show that $\mathbb{R} \oplus \mathbb{R}z$ is closed under multiplication. The other ring axioms will then be inherited from $\mathbb{H}$. First observe that

$$z^2 = -|z|^2 + 2\mathrm{Re}(z)z \in \mathbb{R} \oplus \mathbb{R}z$$

and so

$$(a+bz)(c+dz) = ac + (ad+bc)z + bdz^2 = ac - bd|z|^2 + \big(ad + bc + 2bd\mathrm{Re}(z)\big)z$$

which is in $\mathbb{R} \oplus \mathbb{R}z$.

It is clear that this multiplication is commutative. Moreover, every non-zero element has a multiplicative inverse:

$$(a+bz)^{-1} = (a+b\overline{z})\big(a^2 + 2ab\mathrm{Re}(z) + b^2|z|^2\big)^{-1} \in \mathbb{R} \oplus \mathbb{R}z.$$

Therefore again $\mathbb{R}(z)$ is a field. □

For each non-zero quaternion q we define a map $A_q : \mathbb{H} \longrightarrow \mathbb{H}$ by $A_q(z) = qzq^{-1}$. It is clear that $\big|A_q(z)\big| = |q|\,|z|\,|q|^{-1} = |z|$. Moreover,

$$\overline{A_q(z)} = \overline{q}^{-1}\,\overline{z}\,\overline{q} = q\overline{z}q^{-1} = A_q(\overline{z}).$$

Therefore $\mathrm{Re}\big(A_q(z)\big) = \mathrm{Re}(z)$. If $q \in \mathbb{R}$ then $A_q(z) = z$ for all $z \in \mathbb{H}$. Otherwise, $A_q(z) = z$ if and only if $z \in \mathbb{R}(q)$, where $\mathbb{R}(q)$ is the field considered in Lemma 6.1.1.

Exercise 6.1.2 *Consider the canonical identification between quaternions and column vectors in $\mathbb{R}^4$:*

$$z = z_0 + z_1 i + z_2 j + z_3 k \longleftrightarrow \begin{pmatrix} z_0 \\ z_1 \\ z_2 \\ z_3 \end{pmatrix}. \tag{6.1}$$

We write $q = a + bi + cj + dk$ and then consider the action of $A_q(z)$ in coordinates. If we write $z = z_0 + z_1 i + z_2 j + z_3 k$ and $A_q(z) = w = w_0 + w_1 i + w_2 j + w_3 k$, then we have

$$\begin{aligned} & w_0 + w_1 i + w_2 j + w_3 k \\ = \ & (a+bi+cj+dk)(z_0 + z_1 i + z_2 j + z_3 k)(a - bi - cj - dk)/(a^2+b^2+c^2+d^2). \end{aligned}$$

Show that under the identification between $\mathbb{H}$ and $\mathbb{R}^4$ given in (6.1) the map $A_q(z)$ may be identified with the matrix A_q where

$$A_q = \frac{1}{|q|^2}\begin{pmatrix} |q|^2 & 0 & 0 & 0 \\ 0 & a^2+b^2-c^2-d^2 & 2(bc-ad) & 2(bd+ac) \\ 0 & 2(bc+ad) & a^2-b^2+c^2-d^2 & 2(cd-ab) \\ 0 & 2(bd-ac) & 2(cd+ab) & a^2-b^2-c^2+d^2 \end{pmatrix}.$$

Show that $A_q \in \mathrm{SO}(4)$ and that the 3×3 matrix in the bottom right hand corner of A_q is in $\mathrm{SO}(3)$. Deduce that the map $q \longmapsto A_q$ gives a homomorphism from $\mathbb{H}-\{0\}$ to $\mathrm{SO}(3)$.

Two quaternions z and w are *similar* if there exists non-zero $q \in \mathbb{H}$ such that $z = qwq^{-1}$. The similarity class of z is the set $\{qzq^{-1} : q \in \mathbb{H}-\{0\}\}$.

Lemma 6.1.2 *The quaternions z and w are similar if and only if $|z| = |w|$ and $\mathrm{Re}(z) = \mathrm{Re}(w)$.*

Proof: If z and w are similar then there is a non-zero $q \in \mathbb{H}$ so that $w = A_q(z)$. We have already seen that $|z| = |w|$ and $\mathrm{Re}(z) = \mathrm{Re}(w)$.

Conversely, suppose that $z \in \mathbb{H}$. Consider

$$w = \mathrm{Re}(z) + \big|\mathrm{Im}(z)\big| i = \mathrm{Re}(z) + \sqrt{|z|^2 - \mathrm{Re}(z)^2}\, i.$$

It is clear that $\mathrm{Re}(w) = \mathrm{Re}(z)$ and $|w| = |z|$. We claim that we can find a non-zero $q \in \mathbb{H}$ so that $A_q(w) = qwq^{-1} = z$. The result for general quaternions similar to z will follow by composition.

We take

$$q = \mathrm{Im}(z) + \big|\mathrm{Im}(z)\big| i.$$

That is, if we write $z = z_0 + z_1 i + z_2 j + z_3 k$ then let $y = \mathrm{Im}(z) = z_1 i + z_2 j + z_3 k$. This means that $q = y + |y| i$. We have

$$|q|^2 = \big(y + |y|i\big)\big(-y - |y|i\big) = 2|y|^2 + 2|y|z_1.$$

Hence

$$\begin{aligned} q\big(|y|i\big)q^{-1} &= \big(y+|y|i\big)\big(|y|i\big)\big(-y-|y|i\big)/\big(2|y|^2+2|y|z_1\big) \\ &= \big(-|y|yiy + 2|y|^2 y + |y|^3 i\big)/\big(2|y|^2+2|y|z_1\big) \\ &= \big(2|y|^2 y + 2|y|z_1 y\big)/\big(2|y|^2+2|y|z_1\big) \\ &= y. \end{aligned}$$

Since $q\mathrm{Re}(z)q^{-1} = \mathrm{Re}(z)$ we then have $q\big(\mathrm{Re}(z) + \big|\mathrm{Im}(z)\big|i\big)q^{-1} = z$ as claimed. $\square$

6.2 Quaternionic matrices and Möbius transformations

We can define a *quaternionic right vector space* in an analogous way to the way we define real and complex vector spaces. The main difference is that (quaternionic) scalars act by right multiplication. Thus a quaternionic right vector space is a set V with operations addition and right scalar multiplication so that for $\mathbf{v}$, $\mathbf{w} \in V$ and $t \in \mathbb{H}$ we have $\mathbf{v} + \mathbf{w}$ and $\mathbf{v}t$ in V. Then V should be an abelian group under addition, scalar multiplication is associative and distributive and finally $\mathbf{v}1 = \mathbf{v}$ and $\mathbf{v}0 = \mathbf{0}$ (the additive identity) for all $\mathbf{v} \in V$. The quaternionic right vector space we have in mind will be $\mathbb{H}^2$ defined by

$$\mathbb{H}^2 = \left\{ \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \ : \ z_1,\ z_2 \in \mathbb{H} \right\}$$

with the operations

$$\mathbf{z} + \mathbf{w} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} z_1 + w_1 \\ z_2 + w_2 \end{pmatrix}, \qquad \mathbf{z}t = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} t = \begin{pmatrix} z_1 t \\ z_2 t \end{pmatrix}.$$

The vector space axioms then follow from the fact that $\mathbb{H}$ is a division ring.

Linear maps act on $\mathbb{H}^2$ (with the standard basis) as left multiplication by 2×2 matrices with quaternion entries in the usual way:

$$A\mathbf{z} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix}.$$

To each 2×2 matrix

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

with quaternion entries we associate the quantities σ and τ as follows:

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

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

The quantities σ and τ take the role of the *quaternionic determinant* and *quaternionic trace* of a quaternionic 2×2 matrix. They are not conjugation invariant. Below we will construct some real invariants which are conjugation invariant.

When $\sigma \neq 0$ the matrix A is invertible. Furthermore, A^{-1} is given by:

$$\begin{aligned} A^{-1} &= \begin{pmatrix} c^{-1}d\sigma^{-1}c & -a^{-1}b\sigma^{-1}cac^{-1} \\ -\sigma^{-1}c & \sigma^{-1}cac^{-1} \end{pmatrix}, \qquad \text{when } c \neq 0, \\ A^{-1} &= \begin{pmatrix} a^{-1} & -\sigma^{-1}b \\ 0 & d^{-1} \end{pmatrix}, \qquad \text{when } c = 0. \end{aligned}$$

(Note that when $c = 0$ the hypothesis $\sigma \neq 0$ implies that $a \neq 0$ and $d \neq 0$.) If $\sigma = 0$ then either one column of A is the zero vector or else one column is a left (or right) scalar multiple of the other. In either case A is not invertible.

Note that $a^{-1}b\sigma^{-1}cac^{-1} = c^{-1}d\sigma^{-1}cbd^{-1}$ when a, c, d and σ are all non-zero. In order to see this, observe that

$$b^{-1}c^{-1}\sigma = b^{-1}ac^{-1}d - 1.$$

Therefore $b^{-1}c^{-1}\sigma \in \mathbb{R}(b^{-1}ac^{-1}d)$ and so they commute. Therefore

$$\begin{aligned} a^{-1}b\sigma^{-1}cac^{-1} &= c^{-1}d(d^{-1}ca^{-1}b)(\sigma^{-1}cb)b^{-1}ac^{-1} \\ &= c^{-1}d(\sigma^{-1}cb)(d^{-1}ca^{-1}b)b^{-1}ac^{-1} \\ &= c^{-1}d\sigma^{-1}cbd^{-1}. \end{aligned}$$

We define the *real determinant* (or Dieudonné determinant) to be $\sqrt{\alpha}$ where:

$$\alpha = \alpha_A = |a|^2|d|^2 + |b|^2|c|^2 - 2\mathrm{Re}[a\bar{c}d\bar{b}]. \tag{6.2}$$

Observe that in each case $\alpha = |\sigma|^2$. This enables us to define the group $\mathrm{SL}(2, \mathbb{H})$:

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

Exercise 6.2.1 *Show that the group $SL(2, \mathbb{H})$ is generated by matrices of the form*

$$D = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

where $|\lambda||\mu| = 1$. [This is similar to Lemma 3.2.2.]

An element

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{H})$$

acts on $\mathbb{H}^2$ by left matrix multiplication.

Define a right projection map $\mathbb{P} : \mathbb{H}^2 \longmapsto \widehat{\mathbb{H}} = \mathbb{H} \cup \{\infty\}$ by

$$\mathbb{P} : \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \longmapsto \begin{cases} z_1 z_2^{-1} & \text{if } z_2 \neq 0 \\ \infty & \text{if } z_2 = 0. \end{cases}$$

We define the *standard lift* of $z \in \widehat{\mathbb{H}}$ to $\mathbb{H}^2$ by

$$z \longmapsto \mathbf{z} = \begin{pmatrix} z \\ 1 \end{pmatrix} \quad \text{for } z \in \mathbb{H}, \qquad \infty \longmapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Matrices in $\mathrm{SL}(2, \mathbb{H})$ act of $\widehat{\mathbb{H}}$ by left multiplication on the standard lift followed by right projection. For points $z \in \mathbb{H}$ this is

$$\begin{aligned} A(z) &= \mathbb{P}A\mathbf{z} \\ &= \mathbb{P}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} z \\ 1 \end{pmatrix} \\ &= \mathbb{P}\begin{pmatrix} az+b \\ cz+d \end{pmatrix} \\ &= (az+b)(cz+d)^{-1}. \\ A(\infty) &= \mathbb{P}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= ac^{-1}. \end{aligned}$$

This is a *quaternionic Möbius transformation*. Note that $\frac{az+b}{cz+d}$ does not make sense for the quaternions. The map

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto A(z) = (az+b)(cz+d)^{-1}$$

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. In particular, when $\alpha = 1$ the kernel is $\{\pm I\}$. 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, $A(z)$ or a 'quaternionic Möbius transformation', we refer to the quantities described above with coefficients a, b, c and d such that α takes the value 1.

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

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

Since both A and $-A$ are associated to the same quaternionic Möbius transformation $A(z)$, it is ambiguous to speak of β and δ for a Möbius transformation. We can, however, refer to α, β^2, $\beta\delta$ and so forth, without ambiguity.

Exercise 6.2.2 *Show that*

$$\begin{aligned} \alpha &= |\sigma|^2, \\ \beta &= \mathrm{Re}\big(\sigma\overline{\tau}\big), \\ \gamma &= |\tau|^2 + 2\mathrm{Re}(\sigma), \\ \delta &= \mathrm{Re}(\tau) \end{aligned}$$

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 6.2.1 *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 6.2.4 that if σ and τ are real then they are preserved under conjugation in $\mathrm{SL}(2,\mathbb{H})$. It follows from Theorem 6.2.1 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 6.2.1, all quaternionic Möbius transformations are conjugate to Möbius transformations with complex coefficients.

Lemma 6.2.2 *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 $\sigma=-cb$ and $\tau=cac^{-1}$. Then $-bc=c^{-1}\sigma c=\sigma$ and $a=c^{-1}\tau c=\tau$ since σ and τ are real and so commute with c.

Suppose $d\neq 0$. Notice that

$$cb=cac^{-1}d-\sigma=\tau d-d^2-\sigma\in\mathbb{R}(d).$$

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

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

□

Lemma 6.2.3 *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: If a and d are both real then clearly σ and τ are both real.

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 $cac^{-1}=\overline{d}$;

(ii) if $b\neq 0$ and $c=0$ then $bdb^{-1}=\overline{a}$;

(iii) if $b = c = 0$ and $a \neq d$ then $(d-a)a(d-a)^{-1} = \overline{d}$.

In each case, and in the remaining case in which $b = c = 0$ and $a = d$, a is similar to $\overline{d}$ and hence to d, and also $\sigma = |a|^2$ 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. $\square$

Proposition 6.2.4 *Given $A \in \mathrm{SL}(2,\mathbb{H})$, if $\sigma = \sigma_A$ and $\tau = \tau_A$ 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

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

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

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

Suppose that $c = 0$. We apply Lemma 6.2.3. If a and d are both 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

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

If $c = 0$ then $a' = a$ and $d' = d$ and the result follows from Lemma 6.2.3.

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 6.2.2, we have

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

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, $\sigma_B = |a|^2 = |d|^2 = \sigma_A$ and $\tau_B = 2\mathrm{Re}[d] = 2\mathrm{Re}[a] = \tau_A$. $\square$

Lemma 6.2.5 *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, conjugating by a diagonal map if necessary, we may always suppose $d = a$. When $ba - \overline{a}b = 0$, conjugating so that 0 is a fixed point gives the result. When $ba - \overline{a}b \neq 0$ it is easy to check that $(ba - \overline{a}b)$ commutes with a. Conjugating by $h(z) = \big((a - \overline{a})z + b\big)(ba - \overline{a}b)^{-1}$ gives the result. $\square$

Using Lemma 6.2.3, we see that the condition $bd = \overline{a}b$ is equivalent to σ and τ both being real. We are now in a position to prove Theorem 6.2.1.

Proof: (Theorem 6.2.1) If a quaternionic Möbius transformation is conjugate to a real Möbius transformation then σ and τ are real, by Proposition 6.2.4. 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 6.2.4 ensures that σ and τ remain real. Now we apply Lemma 6.2.3. 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 6.2.5, 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. □

6.3 Quaternionic Hermitian forms

Let $A = (a_{ij})$ be a $k \times l$ quaternionic matrix. Then we define $A^* = (\overline{a}_{ji})$ to be the $l \times k$ matrix given by the conjugate transpose of A. This is completely analogous to the complex case. A $k \times k$ quaternionic matrix H is *Hermitian* if and only if $H^* = H$. To each $k \times k$ quaternionic Hermitian matrix we can define a *quaternionic Hermitian form* on the quaternionic right vector space $\mathbb{H}^k$ as a map $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{H}$ by

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^* H \mathbf{z}.$$

It is clear that for all $\mathbf{z}$, $\mathbf{z}_1$, $\mathbf{z}_2$ and $\mathbf{w}$ column vectors in $\mathbb{H}^k$ and all $\lambda \in \mathbb{H}$, this satisfies:

$$\begin{aligned} \langle \mathbf{z}_1 + \mathbf{z}_2, \mathbf{w} \rangle &= \mathbf{w}^* H(\mathbf{z}_1 + \mathbf{z}_2) = \mathbf{w}^* H \mathbf{z}_1 + \mathbf{w} H \mathbf{z}_2 = \langle \mathbf{z}_1, \mathbf{w} \rangle + \langle \mathbf{z}_2, \mathbf{w} \rangle, \\ \langle \mathbf{z}\lambda, \mathbf{w} \rangle &= \mathbf{w}^* H(\mathbf{z}\lambda) = (\mathbf{w}^* H \mathbf{z})\lambda = \langle \mathbf{z}, \mathbf{w} \rangle \lambda, \\ \langle \mathbf{w}, \mathbf{z} \rangle &= \mathbf{z}^* H \mathbf{w} = \mathbf{z}^* H^* \mathbf{w} = (\mathbf{w}^* H \mathbf{z})^* = \overline{\langle \mathbf{z}, \mathbf{w} \rangle}. \end{aligned}$$

Moreover, using the last property listed above

$$\langle \mathbf{z}, \mathbf{z} \rangle = \overline{\langle \mathbf{z}, \mathbf{z} \rangle}$$

and so for all $\mathbf{z} \in \mathbb{H}^k$ we have $\langle \mathbf{z}, \mathbf{z} \rangle \in \mathbb{R}$. Therefore, we can define

$$\begin{aligned} V_+ &= \Big\{ \mathbf{z} \in \mathbb{H}^k : \langle \mathbf{z}, \mathbf{z} \rangle > 0 \Big\}, \\ V_0 &= \Big\{ \mathbf{z} \in \mathbb{H}^k \setminus \{0\} : \langle \mathbf{z}, \mathbf{z} \rangle = 0 \Big\} \\ V_- &= \Big\{ \mathbf{z} \in \mathbb{H}^k : \langle \mathbf{z}, \mathbf{z} \rangle < 0 \Big\}. \end{aligned}$$

We consider $\mathbb{H}^2$ with a quaternionic Hermitian form of signature $(1, 1)$. There are two standard quaternionic Hermitian forms. The first is simply given by the first Hermitian form associated to the Hermitian matrix H_1 :

$$\langle \mathbf{z}, \mathbf{w} \rangle_1 = \mathbf{w}^* H_1 \mathbf{z} = \begin{pmatrix} \overline{w}_1 & \overline{w}_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \overline{w}_1 z_1 - \overline{w}_2 z_2. \tag{6.3}$$

For the second, we introduce an involution $\cdot^*$ on $\mathbb{H}$. This involution is defined by

$$z^* = (z_0 + z_1 i + z_2 j + z_3 k)^* = z_0 + z_1 i + z_2 j - z_3 k.$$

Observe that $\overline{z}k = kz^*$ for all quaternions z. The second form is a direct generalisation of the second Hermitian form on $\mathbb{C}^{1,1}$. It is

$$\langle \mathbf{z}, \mathbf{w} \rangle_2 = \mathbf{w}^* H_2 \mathbf{z} = \begin{pmatrix} \overline{w}_1 & \overline{w}_2 \end{pmatrix} \begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \overline{w}_2 k z_1 - \overline{w}_1 k z_2 = k(w_2^* z_1 - w_1^* z_2). \tag{6.4}$$

We also have

$$\langle \mathbf{z}\lambda, \mathbf{z}\lambda \rangle = \overline{\lambda}\mathbf{z}^* H \mathbf{z}\lambda = \overline{\lambda}\langle \mathbf{z}, \mathbf{z} \rangle \lambda = |\lambda|^2 \langle \mathbf{z}, \mathbf{z} \rangle$$

since $\langle \mathbf{z}, \mathbf{z} \rangle$ is real and so commutes with λ. Therefore the map $\mathbb{P} : \mathbb{H}^2 \longrightarrow \widehat{\mathbb{H}}$ respects the division of $\mathbb{H}^2$ into V_+, V_0 and V_-. We define the hyperbolic 4-space $\mathbf{H}^4$ to be $\mathbf{H}^4 = \mathbb{P}V_-$ and its ideal boundary $\partial\mathbf{H}^4$ to be $\mathbb{P}V_0$.

We define the quaternionic unit ball to be $\mathbb{B} = \{z \in \mathbb{H} \ : \ |z| < 1\}$. It is easy to see that, for the first Hermitian form $\mathbb{P}(V_-) = \mathbb{B}$. Also the unit sphere in $\mathbb{H}$ is $\partial\mathbb{B} = \mathbb{P}(V_0)$. For H_2 we have $\mathbb{P}(V_-) = \{z = z_0 + z_1 i + z_2 j + z_3 k \in \mathbb{H} \ : \ z_3 > 0\}$ and $\mathbb{P}(V_0) = \{z \in \mathbb{H} \ : \ z_3 = 0\} \cup \{\infty\}$.

We define the hyperbolic metric on $\mathbf{H}^4 = \mathbb{P}(V_-)$ by

$$ds^2 = \frac{-4}{\langle \mathbf{z}, \mathbf{z} \rangle^2} \det \begin{pmatrix} \langle \mathbf{z}, \mathbf{z} \rangle & \langle d\mathbf{z}, \mathbf{z} \rangle \\ \langle \mathbf{z}, d\mathbf{z} \rangle & \langle d\mathbf{z}, d\mathbf{z} \rangle \end{pmatrix}, \qquad \cosh^2\left(\frac{\rho(z,w)}{2}\right) = \frac{\langle \mathbf{z}, \mathbf{w} \rangle \langle \mathbf{w}, \mathbf{z} \rangle}{\langle \mathbf{z}, \mathbf{z} \rangle \langle \mathbf{w}, \mathbf{w} \rangle}.$$

Exercise 6.3.1 *1. Show that for the first Hermitian form H_1*

$$ds^2 = \frac{4dz\, d\overline{z}}{\left(1 - |z|^2\right)^2}, \qquad \cosh^2\Big(\frac{\rho(z,w)}{2}\Big) = \frac{|\overline{w}z - 1|^2}{\left(1 - |z|^2\right)\left(1 - |w|^2\right)}.$$

2. For the second Hermitian form H_2 show that

$$ds^2 = \frac{dz\, d\overline{z}}{z_3^2}, \qquad \cosh^2\Big(\frac{\rho(z,w)}{2}\Big) = \frac{|z - w^*|^2}{4 z_3 w_3}.$$

We define a *symplectic transformation* A to be an automorphism of $\mathbb{H}^{1,1}$, that is, a linear bijection such that $\langle A\mathbf{z}, A\mathbf{w} \rangle = \langle \mathbf{z}, \mathbf{w} \rangle$ for all $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{H}^{1,1}$. In other words, "symplectic" is the natural generalisation to quaternionic matrices of "orthogonal" for real matrices and "unitary" for complex matrices. We denote the group of all unitary transformations by $\mathrm{Sp}(1,1)$. Those symplectic transformations that preserve a given Hermitian form H are denoted $\mathrm{Sp}(H)$.

Proposition 6.3.1 *Let A be a 2×2 quaternionic matrix in* $\mathrm{Sp}(H_1)$*. Then*

$$|a| = |d|, \quad |b| = |c|, \quad |a|^2 - |c|^2 = 1, \quad \overline{a}b = \overline{c}d, \quad a\overline{c} = b\overline{d}. \tag{6.5}$$

Proof:

$$\begin{aligned} A^{-1} &= H_1^{-1}A^*H_1 \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} \overline{a} & -\overline{c} \\ -\overline{b} & \overline{d} \end{pmatrix}. \end{aligned}$$

Therefore

$$\begin{aligned} I &= A^{-1}A = \begin{pmatrix} |a|^2-|c|^2 & \overline{a}b-\overline{c}d \\ \overline{d}c-\overline{b}a & |d|^2-|b|^2 \end{pmatrix}, \\ I &= AA^{-1} = \begin{pmatrix} |a|^2-|b|^2 & b\overline{d}-a\overline{c} \\ c\overline{a}-d\overline{b} & |d|^2-|c|^2 \end{pmatrix}. \end{aligned}$$

Therefore $\overline{a}b = \overline{c}d$ and $a\overline{c} = b\overline{d}$. Also

$$1 = |a|^2-|b|^2 = |a|^2-|c|^2 = |d|^2-|b|^2 = |d|^2-|c|^2.$$

The result follows. $\square$

Proposition 6.3.2 *Let A be a 2×2 quaternionic matrix in* $\mathrm{Sp}(H_2)$. *Then*

$$ad^*-bc^* = d^*a-b^*c = 1, \quad ab^*-ba^* = cd^*-dc^* = c^*a-a^*c = d^*b-b^*d = 0. \tag{6.6}$$

Proof:

$$\begin{aligned} A^{-1} &= H_2^{-1}A^*H_2 \\ &= \begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix} \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix} \begin{pmatrix} 0 & -k \\ k & 0 \end{pmatrix} \\ &= \begin{pmatrix} d^* & -b^* \\ -c^* & a^* \end{pmatrix}. \end{aligned}$$

Therefore

$$\begin{aligned} I &= A^{-1}A = \begin{pmatrix} d^*a-b^*c & d^*b-b^*d \\ a^*c-c^*a & a^*d-c^*b \end{pmatrix}, \\ I &= AA^{-1} = \begin{pmatrix} ad^*-bc^* & ba^*-ab* \\ cd^*-dc^* & da^*-cb^* \end{pmatrix}. \end{aligned}$$

The result follows. $\square$

6.4 Fixed points and eigenvalues

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 $A\mathbf{v} = \mathbf{v}t$. Then $\mathbf{v}$ is called a *right eigenvector* of A. 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.

The relationship between fixed points of quaternionic Möbius transformations and right eigenvectors of quaternionic matrices is similar to the real and complex cases. Suppose that v is fixed by $A(z) = (az+b)(cz+d)^{-1}$ and so $\mathbf{v}$, the standard lift of v, is an eigenvector of A. Then we have

$$vcv + vd - av - b = 0. \tag{6.7}$$

Then it is easy to see that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} v \\ 1 \end{pmatrix} = \begin{pmatrix} v \\ 1 \end{pmatrix}(cv+d).$$

We now show that every quaternionic Möbius transformation has a right eigenvalue. In order to do this, we show that $A \in \mathrm{SL}(2,\mathbb{H})$ may be conjugated to triangular form. This method is called *Niven's trick*.

We suppose that $c \neq 0$. Then the right eigenvalue t corresponding to the fixed point v is $t = cv + d$ where $vcv + vd - av - b = 0$. Substituting $v = c^{-1}(t-d)$ in this equation gives

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

Hence t satisfies the *quaternionic characteristic polynomial*

$$t^2 - \tau t + \sigma = 0. \tag{6.8}$$

However, we also know that

$$t^2 - 2\mathrm{Re}(t)t + |t|^2 = t^2 - (t + \bar{t})t + \bar{t}t = 0.$$

Subtracting these two equations gives

$$\big(\tau - 2\mathrm{Re}(t)\big)t = \big(\sigma - |t|^2\big).$$

When $\tau = 2\mathrm{Re}(t)$ then we must have $\sigma = |t|^2$. Hence σ and τ are both real. We have already dealt with this case.

We now suppose that $\tau \neq 2\mathrm{Re}(t)$. Thus $\tau - 2\mathrm{Re}(t)$ is non-zero and so we can left multiply by its inverse t obtain

$$t = \big(\tau - 2\mathrm{Re}(t)\big)^{-1}(\sigma - |t|^2).$$

We substitute this into the equations $t+\overline{t}=2\mathrm{Re}(t)$ and $t\overline{t}=|t|^2$ to obtain

$$\begin{aligned}\big(\tau-2\mathrm{Re}(t)\big)^{-1}\big(\sigma-|t|^2\big)+\big(\overline{\sigma}-|t|^2\big)\big(\overline{\tau}-2\mathrm{Re}(t)\big)^{-1} &= 2\mathrm{Re}(t),\\ \big(\tau-2\mathrm{Re}(t)\big)^{-1}\big(\sigma-|t|^2\big)(\overline{\sigma}-|t|^2\big)\big(\overline{\tau}-2\mathrm{Re}(t)\big)^{-1} &= |t|^2.\end{aligned}$$

Multiplying on the left by $\big(\tau-2\mathrm{Re}(t)\big)$ and on the right by $\big(\overline{\tau}-2\mathrm{Re}(\tau)\big)$ gives

$$\begin{aligned}\big(\sigma-|t|^2\big)\big(\overline{\tau}-2\mathrm{Re}(t)\big)+\big(\tau-2\mathrm{Re}(t)\big)(\overline{\sigma}-|t|^2\big) &= 2\mathrm{Re}(t)\big(\tau-2\mathrm{Re}(t)\big)\big(\overline{\tau}-2\mathrm{Re}(t)\big),\\ \big(\sigma-|t|^2\big)\big(\overline{\sigma}-|t|^2\big) &= |t|^2\big(\tau-2\mathrm{Re}(t)\big)\big(\overline{\tau}-2\mathrm{Re}(t)\big).\end{aligned}$$

Expanding and using the formulae from Exercise 6.2.2 we obtain

$$\big(\gamma-4\delta\mathrm{Re}(t)+4\mathrm{Re}(t)^2\big)2\mathrm{Re}(t) = 2\beta-2\delta|t|^2+4|t|^2\mathrm{Re}(t), \tag{6.9}$$
$$\big(\gamma-4\delta\mathrm{Re}(t)+4\mathrm{Re}(t)^2\big)|t|^2 = 1+|t|^4. \tag{6.10}$$

We have thus converted the quaternionic quadratic polynomial (6.8) into a pair of real simultaneous polynomials in $2\mathrm{Re}(t)$ and $|t|^2$. The coefficients only involve $\alpha=1$, β, γ and δ. Our task is to solve these equations. We begin by eliminating $2\mathrm{Re}(t)$.

Equations (6.9) and (6.10) lead to:

$$\big(2\beta-2\delta|t|^2\big)|t|^2=2\mathrm{Re}(t)\big(1-|t|^4\big)$$

Using this, we can eliminate $2\mathrm{Re}(t)$ to obtain

$$\gamma|t|^2\big(1-|t|^4\big)^2+4|t|^4\big(\beta-\delta|t|^2\big)\big(\beta|t|^2-\delta\big)=\big(1+|t|^4\big)\big(1-|t|^4\big)^2.$$

Dividing by $4|t|^6$ gives

$$\gamma\big(|t|^2-|t|^{-2}\big)^2/4+\beta^2+\delta^2-\beta\delta\big(|t|^2+|t|^{-2}\big)=\big(|t|^2+|t|^{-2}\big)\big(|t|^2-|t|^{-2}\big)^2/4.$$

Define $X=\big(|t|^2+|t|^{-2}\big)/2$. Note that $X\geq 1$ with equality if and only if $|t|^2=1$. Using $\big(|t|^2-|t|^{-2}\big)^2/4=X-1$, this leads to

$$2X^3-\gamma X^2+2(\beta\delta-1)X+(\gamma-\beta^2-\delta^2)=0.$$

We define this cubic to be $q(x)$:

$$q(x)=2x^3-\gamma x^2+2(\beta\delta-1)x+(\gamma-\beta^2-\delta^2). \tag{6.11}$$

Observe that $q(1)=-(\beta-\delta)^2\leq 0$ and so $q(x)=0$ has at least one root in the interval $[1,\infty)$. Moreover, $|t|=1$ implies that $q(1)=-(\beta-\delta)^2=0$ and so $\beta=\delta$.

Let X be a root of $q(x)$ at least 1, then $|t|^4-2X|t|^2+1=0$ and so

$$|t|^2=X\pm\sqrt{X^2-1}. \tag{6.12}$$

Also dividing (6.10) by $|t|^2$ and substituting for X yields

$$2X = |t|^2 + |t|^{-2} = \gamma - 4\delta\mathrm{Re}(t) + 4\mathrm{Re}(t)^2.$$

Hence

$$2\mathrm{Re}(t) = \delta \pm \sqrt{\delta^2 - \gamma + 2X}. \tag{6.13}$$

Note that $q\big((\gamma - \delta^2/2\big) = -(2\beta - \gamma\delta + \delta^3)^2/4 < 0$ and so we can choose $X > (\gamma - \delta^2)/2$ and so $2\mathrm{Re}(t)$ is defined.

Proposition 6.4.1 *Let t be a root of (6.8).*

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

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

Proof: If t is a root of (6.8) then we have seen that $|t|^2$ and $2\mathrm{Re}(t)$ must satisfy (6.12) and (6.13). In each of these equations there is a choice of sign. We need to know how to make these choices consistently. We write

$$|t|^2 = X + \varepsilon_1\sqrt{X^2 - 1}, \quad 2\mathrm{Re}(t) = \delta + \varepsilon_2\sqrt{\delta^2 - \gamma + 2X}.$$

Note that

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

Using $\delta = \mathrm{Re}(\tau)$, we have

$$\begin{aligned}
t &= \big(\tau - 2\mathrm{Re}(t)\big)^{-1}\big(\sigma - |t|^2\big) \\
&= \Big(\tau - \delta - \varepsilon_2\sqrt{\delta^2 - \gamma + 2X}\Big)^{-1}\Big(\sigma - X - \varepsilon_1\sqrt{X^2 - 1}\Big) \\
&= \Big(\mathrm{Im}(\tau) - \varepsilon_2\sqrt{\delta^2 - \gamma + 2X}\Big)^{-1}\Big(\sigma - X - \varepsilon_1\sqrt{X^2 - 1}\Big) \\
&= \frac{\big(-\mathrm{Im}(\tau) - \varepsilon_2\sqrt{\delta^2 - \gamma + 2X}\big)\big(\sigma - X - \varepsilon_1\sqrt{X^2 - 1}\big)}{2\big(X - \mathrm{Re}(\sigma)\big)}.
\end{aligned}$$

From this we can substitute back for $2\mathrm{Re}(t)$ to obtain

$$\begin{aligned}
4\mathrm{Re}(t)\big(X - \mathrm{Re}(\sigma)\big) &= \big(-\mathrm{Im}(\tau) - \varepsilon_2\sqrt{\delta^2 - \gamma + 2X}\big)\big(\sigma - X - \varepsilon_1\sqrt{x^2 - 1}\big) \\
&\quad + \big(\overline{\sigma} - X - \varepsilon_1\sqrt{x^2 - 1}\big)\big(\mathrm{Im}(\tau) - \varepsilon_2\sqrt{\delta^2 - \gamma + 2X}\big) \\
&= -\mathrm{Im}(\tau)\sigma + \overline{\sigma}\mathrm{Im}(\tau) + 2\varepsilon_2\big(X - \mathrm{Re}(\sigma)\big)\sqrt{\delta^2 - \gamma + 2X} \\
&\quad + 2\varepsilon_1\varepsilon_2\big|\delta X - \beta\big|.
\end{aligned}$$

Now

$$\begin{aligned}
\mathrm{Im}(\tau)\sigma - \overline{\sigma}\mathrm{Im}(\tau) - 2\mathrm{Re}(\sigma)\delta &= \mathrm{Im}(\tau)\sigma - \overline{\sigma}\mathrm{Im}(\tau) - 2\mathrm{Re}(\sigma)\mathrm{Re}(\tau) \\
&= -2\mathrm{Re}(\sigma\overline{\tau}) \\
&= -2\beta.
\end{aligned}$$

Therefore

$$\begin{aligned} 2\varepsilon_1\varepsilon_2\big|\delta X-\beta\big| &= 2\big(\delta+\varepsilon_2\sqrt{\delta^2-\gamma+2X}\big)\big(X-\mathrm{Re}(\sigma)\big) \\ &\qquad +\mathrm{Im}(\tau)\sigma-\overline{\sigma}\mathrm{Im}(\tau)-2\varepsilon_2\big(X-\mathrm{Re}(\sigma)\big)\sqrt{\delta^2-\gamma+2X} \\ &= 2(\delta X-\beta). \end{aligned}$$

Thus if $\delta X>\beta$ we have $\varepsilon_1\varepsilon_2=+1$ and the two choices of sign must be the same. Likewise, $\delta X<\beta$ we have $\varepsilon_1\varepsilon_2=-1$ and we have opposite signs. □

Therefore we have shown that every $A\in\mathrm{SL}(2,\mathbb{H})$ with $c\neq 0$ has a right eigenvalue t. Then $v=c^{-1}(t-d)$ is a fixed point. This means we can conjugate A to upper triangular form:

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

Proposition 6.4.2 *Let*

$$A=\begin{pmatrix} cv+d & -c \\ 0 & a-vc\end{pmatrix}.$$

Then $t_1=cv+d$ and $t_2=a-vc$ are right eigenvalues of A.

Proof: We have already seen that $t_1=cv+d$ is a right eigenvalue of A and that $|t_1|^2=X+\varepsilon_1\sqrt{X^2-1}$ and $2\mathrm{Re}(t_1)=\delta+\varepsilon_2\sqrt{\delta^2-\gamma+2X}$ for a given $X\geq 1$.

We substitute $v=c^{-1}(t_1-d)$ in $t_2=a-vc$ to obtain

$$t_2=a-c^{-1}(t_1-d)c=c^{-1}(cac^{-1}+d-t_1)c=c^{-1}(\tau-t_1)c.$$

Therefore

$$2\mathrm{Re}(t_2)=2\mathrm{Re}(\tau-t_1)=2\delta-2\mathrm{Re}(t_1)=\delta-\varepsilon_2\sqrt{\delta^2-\gamma+2X}.$$

Also, $0=t_1^2-\tau t_1+\sigma$ and so $\sigma=t_1(\tau-t_1)$. Therefore

$$|t_2|^2=|\tau-t_1|^2=|\sigma|^2/|t_1|^2=X-\varepsilon_1\sqrt{X^2-1}.$$

We may reverse the steps above to see that t_2 is an eigenvalue. □

We now reinterpret the above calculations in terms of the diagonal entries of an upper triangular matrix, which we denote by t_1 and t_2. As we have seen, they are both representatives for the similarity classes of eigenvalues. 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 Section 6.2, we find that

$$\begin{aligned} &\alpha=|t_1|^2\,|t_2|^2 \;=\; 1, \\ &\beta=|t_1|^2\mathrm{Re}[t_2]+|t_2|^2\mathrm{Re}[t_1], \\ &\gamma=|t_1|^2+|t_2|^2+4\mathrm{Re}[t_1]\mathrm{Re}[t_2], \\ &\delta=\mathrm{Re}[t_1]+\mathrm{Re}[t_2]. \end{aligned}$$

It is easy to use these identities to show that all quaternions t similar to either t_1 or t_2 satisfy the *real characteristic polynomial of A*

$$t^4 - 2\delta t^3 + \gamma t^2 - 2\beta t + \alpha = 0. \tag{6.14}$$

Corollary 6.4.3 *The quantities α, β, γ, δ are invariant under conjugation.*

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 $\mathrm{Re}[t_2]$. Using these expressions we can solve the cubic polynomial (6.11) and hence determine which root corresponds to X.

Theorem 6.4.4 *Suppose that f is a quaternionic Möbius transformation with $\alpha = 1$. Let X denote the largest real root of the cubic polynomial (6.11)*

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

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 $2\mathrm{Re}[t] = \delta \mp \sqrt{2X - \gamma + \delta^2}$.

For large values of x we have $q(x) > 0$. Notice that $q(1) = -(\beta - \delta)^2 \leq 0$ and so we have $X \geq 1$. Also, $q\big((\gamma - \delta^2)/2\big) = -\frac{1}{4}(2\beta - \gamma\delta + \delta^3)^2 \leq 0$ and similarly we have $2X \geq \gamma - \delta^2$. Therefore the square roots in Theorem 6.4.4 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

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

Since

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

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

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, although as we have seen, we can conjugate it to an upper triangular matrix. Then Lemma 6.2.5 gives a criterion which describes when such matrices A can be diagonalised.

6.5 Conjugacy

Theorem 6.4.4 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 6.2.

Theorem 6.5.1 *Two non-trivial quaternionic Möbius transformations $A(z)$ and $B(z)$ 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) $\beta_A\delta_A = \beta_B\delta_B$, $\gamma_A = \gamma_B$ and $\delta_A^2 = \delta_B^2$.

Note that we can use σ and τ to decide whether a quaternionic Möbius transformation is conjugate to a real Möbius transformation.

Proof: Suppose that $A(z)$ and $B(z)$ 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 $A(z)$ and $B(z)$ are both conjugate to real Möbius transformations then σ and τ are real (by Theorem 6.2.1) and preserved under conjugation (by Proposition 6.2.4). For real Möbius transformations, $\beta\delta = \sigma\tau^2$. Therefore, because σ is either -1 or 1, we have $\tau_A^2/\sigma_A = \tau_B^2/\sigma_B$. The quantity τ^2/σ determines conjugacy amongst real Möbius transformations, therefore $A(z)$ and $B(z)$ are conjugate. Henceforth we assume that neither $A(z)$ nor $B(z)$ is conjugate to real Möbius transformations.

Using condition (ii), we lift $A(z)$ and $B(z)$ 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 6.4.4 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 is not conjugate to a real matrix, using Lemma 6.2.5, we see they are both conjugate to $\left(\begin{smallmatrix} t & 1 \\ 0 & t \end{smallmatrix}\right)$ □

6.6 Classification of quaternionic Möbius transformations

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 6.6.1 *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 map. 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

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

Using these equations, we find:

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

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 $\beta \neq \delta$. □

Lemma 6.6.2 *Let* $A \in \mathrm{Sp}(H)$ *for some quaternionic Hermitian form* H *and suppose that* t *is a right eigenvalue of* A*. Then* $\overline{t}^{-1}$ *is also an eigenvalue of* A*.*

Proof: Since A^{-1} is conjugate (via H) to A^* they have the same right eigenvalues. If t is a right eigenvalue of A then $\overline{t}$ is a right eigenvalue of A^*, and so of A^{-1}. Hence $\overline{t}^{-1}$ is a right eigenvalue of A as claimed. □

Corollary 6.6.3 *If* $A \in \mathrm{Sp}(H)$ *then* $\beta = \delta$*.*

Proof: Either $|t_1| = |t_2| = 1$ and $\beta = \mathrm{Re}(t_2) + \mathrm{Re}(t_1) = \delta$ or else $t_2 = \overline{t}_1^{-1}$. In the latter case

$$\beta = |t_1|^2 \mathrm{Re}(\overline{t}_1^{-1}) + |t_1|^{-2} \mathrm{Re}(t_1) = \mathrm{Re}(t_1) + \mathrm{Re}(\overline{t}_1^{-1}) = \delta.$$

□

Thus for elements of $\mathrm{Sp}(H)$ the cubic $q(x)$ given by (6.11) has a root $x = 1$. Thus we can factor it as

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

Therefore the roots of q are

$$x = 1, \quad x = \frac{\gamma - 2 \pm \sqrt{(\gamma+2)^2 - 16\delta^2}}{4}.$$

A brief calculation shows that $x = 1$ is the largest of these roots whenever

$$2 \geq \gamma - \delta^2 = \left|\mathrm{Im}(a+d)\right|^2 + 2\mathrm{Re}(ad - bc).$$

We have already seen the classification into k-simple transformations. A second method of classification is the standard trichotomy based on fixed points.

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 $|t_1| \neq |t_2|$.

For Möbius transformations preserving the unit ball or upper half space this specialises to the usual definitions: Let $A(z) \in \mathrm{PSp}(H_1)$ be a quaternionic Möbius transformation preserving $\mathbb{B}$. We say that

(i) $A(z)$ is *elliptic* if it has at least one fixed point in $\mathbb{B}$;

(ii) $A(z)$ is *parabolic* if it has exactly one fixed point and this point lies in $\partial\mathbb{B}$;

(iii) $A(z)$ is *loxodromic* if it has exactly two fixed points and these points lie in $\partial\mathbb{B}$.

Theorem 6.6.4 *Let $A(z)$ be a quaternionic Möbius transformation with $\alpha = 1$.*

(a) If $\sigma = 1$ and $\tau \in \mathbb{R}$ then $A(z)$ is 1–simple, $\beta = \delta$, $\gamma = \delta^2 + 2$ and the following trichotomy holds.

(i) If $0 \leq \delta^2 < 4$ then $A(z)$ is elliptic.

(ii) If $\delta^2 = 4$ then $A(z)$ is parabolic.

(iii) If $\delta^2 > 4$ then $A(z)$ is loxodromic.

(b) If $\beta = \delta$ and either $\tau \notin \mathbb{R}$ or $\sigma \neq 1$, then $A(z)$ 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 $\beta \neq \delta$ then $A(z)$ is 3–simple loxodromic.

We remark that if $A(z)$ is conjugate to a real Möbius transformation with determinant σ and trace τ then the following possibilities occur:

(i) $\sigma > 0$ and $A(z)$ is 1–simple. In this case τ^2/σ determines whether $A(z)$ is elliptic, parabolic or loxodromic in the usual way;

(ii) $\sigma < 0$ and $\tau = 0$ in which case $A(z)$ is 2–simple and elliptic;

(iii) $\sigma < 0$ and $\tau \neq 0$ in which case $A(z)$ is 3–simple and loxodromic.

To determine the dynamics and conjugacy class of a Möbius transformation with quaternion coefficients for which $\alpha \neq 1$, one should replace the coefficients of the transformation with suitably scaled alternatives so that $\alpha = 1$, before applying Theorem 6.5.1 and Theorem 6.6.4. 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.

Proof: (Theorem 6.6.4) 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 6.6.1). 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 $\gamma - \delta^2 > 2$. □

6.7 The geometry of hyperbolic 4-space

We have two quaternionic models of hyperbolic 4-space $\mathbf{H}^4$, namely the unit ball $\mathbb{B} = \{z \in \mathbb{H} \ : \ |z| < 1\}$ (which corresponds to H_1) and the upper half space $\mathbb{U} = \{z = z_0 + z_1 i + z_2 j + z_3 k \in \mathbb{H} \ : \ z_3 > 0\}$ (which corresponds to H_2). We can carry over the geometric notions from lower dimensions to these models. In all cases the proofs are the same, except that we must be careful about commutativity.

We can pass between the models using a Cayley transform C and the corresponding Möbius transformation $C(z)$ given by

$$C = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -k \\ -k & 1 \end{pmatrix}, \qquad C(z) = (z-k)(-kz+1)^{-1}.$$

Geodesics are arcs of circles (or lines) orthogonal to the boundary of $\mathbf{H}^4$. Loxodromic maps fix two points on the boundary $\partial\mathbf{H}^4$ and preserve the geodesic between them. For example consider $A(z) = azd^{-1} \in \mathrm{PSp}(H_2)$ where $ad^* = 1$ and $|a| \neq 1$. This map fixes 0 and ∞ and preserves the geodesic between them, namely the positive z_3-axis.

We can extend the idea of geodesics to include (totally geodesic) copies of the hyperbolic plane and hyperbolic 3-space embedded in $\mathbf{H}^4$. For the hyperbolic plane these are discs or half-planes orthogonal to the boundary. For example in the unit ball model one such half plane is

$$\left\{z = z_0 + z_1 i + z_2 j + z_3 k \in \mathbb{B} \ : \ z_2 = z_3 = 0\right\}$$

and one in the upper half space is

$$\left\{z = z_0 + z_1 i + z_2 j + z_3 k \in \mathbb{U} \ : \ z_1 = z_2 = 0\right\}$$

Each 1-simple Möbius transformation preserves one of these totally geodesically embedded hyperbolic planes and its action there is the same as the action on the corresponding Poincaré model of the hyperbolic plane. In this case, δ and τ are equal and they are just the usual trace.

We can define spheres as before. These are just Euclidean 3-spheres but with a different hyperbolic centre and radius. For example in the ball model, the hyperbolic sphere of radius r centred at 0 is the Euclidean sphere centred at 0 with radius $\tanh(r/2)$. The proof is the same as for Lemma 3.3.2. Likewise, in the upper half space, the sphere with centre $z = z_0 + z_1 i + z_2 j + z_j k$ and radius r is the Euclidean sphere with centre $z_0 + z_1 i + z_2 j + z_3 \cosh(r) k$ and radius $z_3 \sinh(r)$. The proof of this is the same as for Lemma 5.3.2. Elliptic maps preserve hyperbolic spheres centred at the fixed point. For example, $A(z) = azd^{-1} \in \mathrm{PSp}(H_1)$ where $|a| = |d| = 1$ fixes every hyperbolic sphere centred at 0. If in addition $d = a$ then $A(z)$ fixes every point of $\mathbb{R}(a)$ that lies in the unit ball. This is a copy of the hyperbolic plane. Any sphere centred at a point of this plane is preserved by $A(z)$.

For example, let

$$\begin{aligned} A &= \begin{pmatrix} \cos\theta + i\sin\theta & 0 \\ 0 & \cos\theta + i\sin\theta \end{pmatrix}, \\ A(z) &= (\cos\theta + i\sin\theta) z (\cos\theta + i\sin\theta)^{-1} \\ &= (\cos\theta + i\sin\theta) z (\cos\theta - i\sin\theta). \end{aligned}$$

Then

$$A(z_0 + z_1 i + z_2 j + z_3 k) = z_0 + z_1 i + (\cos\theta + i\sin\theta)^2 (z_2 j + z_3 k).$$

Then $A(z)$ fixes each point in $\mathbb{R}(i) = \mathbb{C} \subset \mathbb{H}$ and rotates around this point by an angle 2θ. Likewise if

$$\begin{aligned} B &= \begin{pmatrix} \cos\theta + i\sin\theta & 0 \\ 0 & \cos\theta - i\sin\theta \end{pmatrix}, \\ B(z) &= (\cos\theta + i\sin\theta) z (\cos\theta - i\sin\theta)^{-1} \\ &= (\cos\theta + i\sin\theta) z (\cos\theta + i\sin\theta). \end{aligned}$$

Then

$$B(z_0 + z_1 i + z_2 j + z_3 k) = (\cos\theta + i\sin\theta)^2(z_0 + z_1 i) + z_2 j + z_3 k$$

This fixes every point of the plane spanned by j and k and again rotates through an angle 2θ. Both $A(z)$ and $B(z)$ are 1-simple.

Finally, horospheres and horoballs are defined as before. For example those based at ∞ of height t are

$$\begin{aligned} H_t &= \Big\{z = z_0 + z_1 i + z_2 j + z_3 k \in \mathbb{U} \ : \ z_3 = t\Big\}, \\ B_t &= \Big\{z = z_0 + z_1 i + z_2 j + z_3 k \in \mathbb{U} \ : \ z_3 > t\Big\}. \end{aligned}$$

Each horosphere is a copy of Euclidean 3-space. Parabolic and 1-simple elliptic maps fixing ∞ preserve each horoball and act as Euclidean isometries. For example, consider $A(z) = (az + b)d^{-1} \in \mathrm{PSp}(H_2)$ with $|a| = |d| = 1$ (and so $ad^* = 1$, $ab^* - ba^* = d^*b - b^*d = 0$). Then

$$0 = kd^*b - kb^*d = \overline{d}kb - \overline{b}kd$$

and so $kbd^{-1} = \overline{d}^{-1}\overline{b}k = \overline{bd^{-1}}k$. This means that the k coordinate of bd^{-1} is zero. Then

$$\begin{aligned} A(z_0 + z_1 i + z_2 j + z_3 k) &= a(z_0 + z_1 i + z_2 j + z_3 k)d^{-1} + bd^{-1} \\ &= a(z_0 + z_1 i + z_2 j)d^{-1} + ad^* z_3 k + bd^{-1} \\ &= a(z_0 + z_1 i + z_2 j)d^{-1} + bd^{-1} + z_3 k. \end{aligned}$$

Thus $A(z)$ preserves each horosphere. Moreover,

$$\Big|A(z) - A(w)\Big| = \Big|azd^{-1} + bd^{-1} - awd^{-1} - bd^{-1}\Big| = \Big|a(z-w)d^{-1}\Big| = \Big|z - w\Big|$$

and so $A(z)$ acts as a Euclidean isometry.

Notice that parabolic maps are no longer necessarily translations. In fact 1-simple parabolic maps are translations, but 2-simple parabolic maps are screw motions. For example

$$A = \begin{pmatrix} \cos\theta + i\sin\theta & \cos\theta + i\sin\theta \\ 0 & \cos\theta + i\sin\theta \end{pmatrix}$$

acts as $A(z) = (\cos\theta + i\sin\theta)z(\cos\theta - i\sin\theta) + 1$. That is

$$A(z_0 + z_1 i + z_2 j + z_3 k) = z_0 + z_1 i + (\cos\theta + i\sin\theta)^2(z_2 j + z_3 k) + 1.$$

In $\mathrm{PSL}(2, \mathbb{H})$ the 3-simple Möbius transformations do not preserve a copy of hyperbolic 4-space (just as 2-simple elements of $\mathrm{SL}(2, \mathbb{C})$ don't preserve a hyperbolic plane). We could develop a Poincaré extension, but the algebra is more complicated. Instead, in the next section we pass to all higher dimensions at once.

Chapter 7

Clifford Möbius transformations

7.1 Clifford algebras

The *Clifford algebra* $\mathcal{C}_n$ is the associative algebra over the real numbers generated by n elements $i_1, \ldots, i_n$ that anti-commute and square to -1. In other words the i_j are subject to the relations $i_j i_k = -i_k i_j$ and $i_j^2 = -1$ for each $j \neq k$ in $\{1, \ldots, n\}$. Each element of $\mathcal{C}_n$ may be written as a real linear combination of products $i_{k_1} \ldots i_{k_m}$ where $1 \leq k_1 < \cdots < k_m \leq n$. We include the empty product, thought of as the real number 1. Hence we can identify $\mathcal{C}_n$ with a real vector space of dimension 2^n. If $x \in \mathcal{C}_n$ is written out in this way, the coefficient x_0 of 1 will be called the *real part* of x and is denoted $\mathrm{Re}(x)$. Also, $x - x_0 = x - \mathrm{Re}(x)$ will be called the *imaginary part* of x and denoted $\mathrm{Im}(x)$.

The Clifford algebras $\mathcal{C}_0$ and $\mathcal{C}_1$ are simply $\mathbb{R}$ and $\mathbb{C}$ respectively. Writing $i_1 = i$ and $i_2 = j$ we may identify $\mathcal{C}_2$ with the quaternions $\mathbb{H}$. In order to see this, observe that any quaternion may be written as a linear combination of 1, $i = i_1$, $j = i_2$ and $k = i_1 i_2$. The relations $i_1^2 = i_2^2 = -1$ and $i_1 i_2 = -i_2 i_1$ imply that $k^2 = (i_1 i_2)^2 = -i_1^2 i_2^2 = -1$ and $jk = i_2 i_1 i_2 = -i_1 i_2^2 = i_1 = i$ and so on.

Exercise 7.1.1 *Show that* $(i_1 i_2 i_3)(i_1 i_2 i_3) = +1$. *Deduce that* $(1 + i_1 i_2 i_3)(1 - i_1 i_2 i_3) = 0$ *and* $(1 + i_1 i_2 i_3)^2 = 2(1 + i_1 i_2 i_3)$.

A consequence of Exercise 7.1.1 is that $\mathcal{C}_n$ contains zero divisors when $n \geq 3$ and so is not a division ring. This means that the constructions we have made for $n = 1$ and $n = 2$ (that is for $\mathbb{C}$ and $\mathbb{H}$) do not generalise directly. To get around the problem of zero divisors we introduce two subsets of $\mathcal{C}_n$ as follows. A *Clifford vector* is a real linear combination of 1, $i_1, \ldots, i_n$ and the set of Clifford vectors, denoted $\mathbb{V}^{n+1}$, may be identified with $\mathbb{R}^{n+1}$ in a canonical way. The *Clifford group* Γ_n is the collection of all finite products of non-zero Clifford vectors. We will see below that this really is a group!

There are three involutions on $\mathcal{C}_n$ that generalise complex conjugation. In what follows, assume that $x \in \mathcal{C}_n$ is written as a linear combination of products of the i_j as above. Then

(i) x' is obtained from x by sending i_j to $-i_j$ in each product. Hence the product $i_{k_1} \dots i_{k_m}$ is sent to $(-1)^m i_{k_1} \dots i_{k_m}$.

(ii) x^* is obtained from x by reversing the order of each product of i_j. Hence the product $i_{k_1} \dots i_{k_m}$ is sent to $i_{k_m} \dots i_{k_1}$ which may be rearranged using the relations to $\pm i_{k_1} \dots i_{k_m}$.

(iii) $\overline{x}$ is obtained from x by sending i_j to $-i_j$ and reversing the order of each product of i_j. Hence the product $i_{k_1} \dots i_{k_m}$ is sent to $(-1)^m i_{k_m} \dots i_{k_1}$ which may be rearranged using the relations to $\pm i_{k_1} \dots i_{k_m}$.

Therefore when $n = 1$ we have $(x+iy)^* = x+iy$ and $(x+iy)' = \overline{(x+iy)} = x-iy$, which is the usual complex conjugation in $\mathbb{C}$. Similarly, when $n = 2$ we have $(x_0 + x_1 i + x_2 j + x_3 ij)^* = x_0 + x_1 i + x_2 j - x_3 ij$ which is the involution x^* defined in the previous chapter. Also, $\overline{x_0 + x_1 i + x_2 j + x_3 ij} = x_0 - x_1 i - x_2 j - x_3 ij$ which is the usual quaternionic conjugate $\overline{x}$ of x. Finally $x' = \overline{x}^*$.

Exercise 7.1.2 *1. Show that*

$$(i_{k_1} \dots i_{k_m})^* = \begin{cases} (-1)^{m/2} i_{k_1} \dots i_{k_m} & \text{if } m \text{ is even,} \\ (-1)^{(m-1)/2} i_{k_1} \dots i_{k_m} & \text{if } m \text{ is odd.} \end{cases}$$

2. Show that

$$\overline{i_{k_1} \dots i_{k_m}} = \begin{cases} (-1)^{m/2} i_{k_1} \dots i_{k_m} & \text{if } m \text{ is even,} \\ (-1)^{(m+1)/2} i_{k_1} \dots i_{k_m} & \text{if } m \text{ is odd.} \end{cases}$$

The following lemma is easy to prove.

Lemma 7.1.1 *For all $a, b \in \mathcal{C}_n$ we have $\overline{a} = (a')^* = (a^*)'$ and*

$$\begin{array}{lll} (a+b)' = a' + b', & (ab)' = a'b', & (a')' = a, \\ (a+b)^* = a^* + b^*, & (ab)^* = b^* a^*, & (a^*)^* = a, \\ \overline{a+b} = \overline{a} + \overline{b}, & \overline{(ab)} = \overline{b}\,\overline{a}, & \overline{(\overline{a})} = a. \end{array}$$

We now consider products of vectors. Suppose that $x = x_0 + x_1 i_1 + \dots + x_n i_n$ is in $\mathbb{V}^{n+1}$. Then applying the definitions, we see that $x' = \overline{x} = x_0 - x_1 i_1 - \dots - x_n i_n$ and $x^* = x$. For vectors we generally write $\overline{x}$ and x and do not use x' and x^*. For a vector x the real part is given by $\mathrm{Re}(x) = x_0 = (x + \overline{x})/2$.

Moreover, $x\overline{x} = \overline{x}x = x_1^2 + x_1^2 + \dots + x_n^2 \in \mathbb{R}$ and we define $|x|$ by $|x|^2 = x\overline{x} = \overline{x}x$. This notion may be extended to products of vectors. If x and y are in $\mathbb{V}^{n+1}$ then

$$xy\overline{(xy)} = xy\overline{y}\,\overline{x} = x|y|^2\overline{x} = x\overline{x}|y|^2 = |x|^2|y|^2$$

where we have used the fact that, since $|y|^2$ it real, it commutes with $\overline{x}$. By induction, we may extend this definition so that if $a \in \Gamma_n$ then $a\overline{a} = \overline{a}a = |a|^2$. Therefore, we can define the *inverse* of $a \in \Gamma_n$ to be $a^{-1} = \overline{a}|a|^{-2}$. Therefore the Clifford group really is a group.

It is *not* true in general that $a + \overline{a}$ is real for all $a \in \Gamma_n$. For example,

$$\overline{i_1 i_2 i_3} = (-i_3)(-i_2)(-i_1) = -i_3 i_2 i_1 = i_2 i_3 i_1 = -i_2 i_1 i_3 = i_1 i_2 i_3.$$

Therefore $i_1 i_2 i_3 + \overline{i_1 i_2 i_3} = 2 i_1 i_2 i_3$ which is not real. However, for products of two vectors it is still true:

Lemma 7.1.2 *Let $x = x_0 + x_1 i_1 + \cdots + x_n i_n$ and $y = y_0 + y_1 i_1 + \cdots + y_n i_n$ lie in $\mathbb{V}^{n+1}$. Then*

$$x\overline{y} + y\overline{x} = x\overline{y} + \overline{x\overline{y}} = 2(x_0 y_0 + x_1 y_1 + \cdots x_n y_n) = 2x \cdot y$$

where $\cdot$ is the usual dot product.

This enables us to write down reflections in hyperplanes in $\mathbb{V}^{n+1}$ in terms of Clifford numbers.

Lemma 7.1.3 *If $y \in \mathbb{V}^{n+1}$ then the reflection in the hyperplane orthogonal to y is given by $R_y : x \longmapsto -y\overline{x}\,\overline{y}^{-1}$.*

Proof: We know that the reflection in the hyperplane orthogonal to y is given by

$$x \longmapsto x - \frac{2(x \cdot y)y}{|y|^2} = x - \frac{(x\overline{y} + y\overline{x})y}{|y|^2} = x - \frac{x|y|^2 + y\overline{x}y}{|y|^2} = -y\overline{x}\,\overline{y}^{-1}.$$

□

Proposition 7.1.4 *For each $a \in \Gamma_n$ consider the map $A_a : x \longmapsto axa'^{-1} = axa^*/|a|^2$. The map $a \longmapsto A_a(x)$ is a surjective homomorphism from Γ_n to $\mathrm{SO}(n+1)$ with kernel $\mathbb{R} - \{0\}$.*

Proof: If $a,\, b \in \Gamma_n$ then

$$A_{ab} = (ab)x(ab)'^{-1} = a(bxb'^{-1})a'^{-1} = a\big(A_b(x)\big)a'^{-1} = A_a A_b(x)$$

and so this map is a homomorphism.

Let $y \in \mathbb{V}^{n+1}$ then, using Lemma 7.1.3, we see that $R_y(x) = -y\overline{x}\,\overline{y}^{-1}$. Also, when $y = 1$ we have $R_1(x) = -\overline{x}$. Therefore the map

$$A_y(x) = yx\overline{y}^{-1} = R_y R_1(x)$$

is the product of reflection in the hyperplane orthogonal to 1 followed by reflection in the hyperplane orthogonal to y. (Note that when $y \in \mathbb{V}^{n+1}$ then $y' = \overline{y}$ and by convention we use the latter.) Therefore this map lies in $\mathrm{SO}(n+1)$. If $a = y_1 y_2 \cdots y_m$

then, since $a \longmapsto A_a(x)$ is a homomorphism, we have $A_a(x) = A_{y_1}A_{y_2}\cdots A_{y_m}(x)$. This is a product of elements of $\mathrm{SO}(n+1)$ and so lies in $\mathrm{SO}(n+1)$.

Finally, any element of $\mathrm{SO}(n+1)$ can be written as a product of an even number of reflections and so the map $a \longmapsto A_a$ is surjective. Clearly if $t \in \mathbb{R} - \{0\}$ then $A_t(x) = txt'^{-1} = txt^{-1} = x$. In fact,this is the only way to get the identity. In order to see this, suppose that $A_a(x) = x$. Then $ax = xa'$ for all $x \in \mathbb{V}^{n+1}$. In particular this is the case when x is real and so $a = a'$. Suppose that we write a as a linear combination of products of the i_j. The condition $a = a'$ means that the only products which appear with non-zero coefficients comprise an even number (possibly zero) of the i_j. If $i_{k_1}\ldots i_{k_{2m}}$ is such a product then for each $j = 1, \ldots, 2m$

$$(i_{k_1}\ldots i_{k_{2m}})i_{k_j} = -i_{k_j}(i_{k_1}\ldots i_{k_{2m}})'$$

and so this term cannot arise in a when $ai_{k_j} = i_{k_j}a'$. Hence the only terms in a are empty products, which is the same as saying that a is real. □

Corollary 7.1.5 *Suppose a and b are in Γ_n. Then $a^{-1}b \in \mathbb{V}^{n+1}$ if and only if $ba^* \in \mathbb{V}^{n+1}$.*

Proof: If $a^{-1}b \in \mathbb{V}^{n+1}$ then so is $A_a(a^{-1}b) = a(a^{-1}b)a^*/|a|^2 = ba^*/|a|^2$. □

Proposition 7.1.6 *Let $y \in \mathbb{V}^{n+1} - \{0\}$ then there exists $a \in \Gamma_n$ so that $y = aa^*$.*

Proof: Write $y = ru$ where r is a positive real number and $u \in \mathbb{V}^{n+1}$ has $|u| = 1$. Then, by construction, u lies on the unit sphere S^n in $\mathbb{V}^{n+1}$. Since $\mathrm{SO}(n+1)$ acts transitively on S^n there is an element of $\mathrm{SO}(n+1)$ sending 1 to u. Using Proposition 7.1.4 we can write this map as $A_a(x) = axa^*/|a|^2$ for some $a \in \Gamma_n$. Moreover, any non-zero real multiple of a yields the same map. Thus we have

$$u = A_a(1) = aa^*/|a|^2.$$

Multiplying a by a positive real number if necessary, we suppose that $|a|^2 = r$. Therefore

$$aa^* = |a|^2A_a(1) = |a|^2u = ru = y$$

as claimed. □

Proposition 7.1.7 *Any Euclidean similarity of $\mathbb{V}^{n+1}$ can be written in the form $S(x) = (ax+b)a^*$ or $S(x) = \big(a(-\overline{x})+b\big)a^*$ for some a in Γ_n and $b \in \Gamma_n \cup \{0\}$ with $ba^* \in \mathbb{V}^{n+1}$.*

Proof: We can write any orientation preserving Euclidean similarity as a rotation followed by a dilation and then a translation. These maps can all be written in terms of Clifford algebras to give the map indicated. For orientation reversing similarities, first apply the reflection $x \longmapsto -\overline{x}$ and then an orientation preserving similarity. □

7.2 Hermitian forms and unitary matrices

We would like to mimic the construction of complex and quaternionic vector spaces equipped with a Hermitian form and the corresponding unitary or symplectic matrices that preserve this form. Unfortunately, we have an additive structure on $\mathbb{V}^{n+1}$ but not a multiplicative structure and we have a multiplicative structure on Γ_n but not an additive structure. This means that we have problems defining a vector space. However, we can still define Hermitian and unitary matrices.

Let $A = (a_{ij})$ be a $k \times l$ matrix with entries in $\Gamma_n \cup \{0\}$. Then we define $A^* = (\overline{a}_{ji})$ to be the matrix obtained by taking the transpose and applying the involution $a \longmapsto \overline{a}$ to each entry – there is an unfortunate problem with the star notation here! A $k \times k$ matrix H with entries in $\Gamma_n \cup \{0\}$ is said to be *Clifford Hermitian* if $H^* = H$. When $k = 2$ we consider two particular Clifford Hermitian matrices that generalise those we considered for the complex numbers and quaternions. They are

$$H_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad H_2 = \begin{pmatrix} 0 & -i_n \\ i_n & 0 \end{pmatrix}.$$

Given a Clifford Hermitian matrix H, we can define *Clifford unitary* matrices to be matrices A with entries in $\Gamma_n \cup \{0\}$ so that $A^*HA = H$. Provided H is invertible we also have $A^{-1} = H^{-1}A^*H$ as before. We now characterise Clifford unitary matrices with respect to H_1 and H_2.

Proposition 7.2.1 *Let A be a 2×2 matrix with entries a, b, c, $d \in \Gamma_n \cup \{0\}$. If $A^*H_1A = H_1$ then*

$$\begin{aligned} &|a| = |d|, \quad |b| = |c|, \quad |a|^2 - |c|^2 = 1, \\ &\overline{a}b = \overline{c}d, \quad a\overline{c} = b\overline{d}. \end{aligned}$$

Proof: We have

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} |a|^2 - |c|^2 & \overline{a}b - \overline{c}d \\ \overline{b}a - \overline{d}c & |b|^2 - |d|^2 \end{pmatrix}.$$

Therefore $|a|^2 - |c|^2 = |d|^2 - |b|^2 = 1$ and $\overline{a}b = \overline{c}d$.

From this we see that $|a|\,|b| = |c|\,|d|$ and so

$$|a|^2 = |a|^2|d|^2 - |a|^2|b|^2 = |a|^2|d|^2 - |c|^2|d|^2 = |d|^2.$$

Therefore $|a| = |d|$ and hence $|b| = |c|$. Finally

$$a\overline{c}d = a\overline{a}b = |a|^2 b = b|d|^2 = b\overline{d}d$$

and so $a\overline{c} = b\overline{d}$. (This uses $|d| \geq 1$ and so $d \neq 0$.) $\square$

Lemma 7.2.2 *If $a \in \Gamma_{n-1}$ then $i_n a^* = \overline{a} i_n$.*

Proposition 7.2.3 *Let A be a 2×2 matrix with entries a, b, c, $d \in \Gamma_{n-1} \cup \{0\}$. If $A^* H_2 A = H_2$ then*

$$\begin{aligned} &ad^* - bc^* = d^*a - b^*c = 1, \\ &ab^* - ba^* = cd^* - dc^* = c^*a - a^*c = d^*b - b^*d = 0. \end{aligned}$$

Proof: We have

$$\begin{pmatrix} 0 & -i_n \\ i_n & 0 \end{pmatrix} = \begin{pmatrix} \overline{a} & \overline{c} \\ \overline{b} & \overline{d} \end{pmatrix} \begin{pmatrix} 0 & -i_n \\ i_n & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} i_n(c^*a - a^*c) & i_n(c^*b - a^*d) \\ i_n(d^*a - b^*c) & i_n(d^*b - b^*d) \end{pmatrix}.$$

Therefore $d^*a - b^*c = 1$ and $c^*a - a^*c = d^*b - b^*d = 0$. From these equations we have $1 = \overline{(d^*a - b^*c)} = \overline{a}d' - \overline{c}b'$. Also $(a'c^* - c\overline{a})|a|^2 = a'(c^*a - a^*c)\overline{a} = 0$ and so $a'c^* - c\overline{a} = 0$. Similarly, $b'd^* - d\overline{b} = 0$.

Therefore

$$\begin{aligned} ab^* - ba^* &= a(\overline{a}d' - \overline{c}b')b^* - b(\overline{d}a' - \overline{b}c')a^* \\ &= |a|^2(d'b^* - b\overline{d}) - |b|^2(a\overline{c} - c'a^*) \;=\; 0, \\ cd^* - dc^* &= c(\overline{a}d' - \overline{c}b')d^* - d(\overline{d}a' - \overline{b}c')c^* \\ &= |d|^2(c\overline{a} - a'c^*) - |c|^2(b'd^* - d\overline{b}) \;=\; 0. \end{aligned}$$

Finally, when $d \neq 0$

$$|d|^2 = d'(d^*a - b^*c)d^* = |d|^2 ad^* - d'b^*cd^* = |d|^2 ad^* - b\overline{d}dc^* = |d|^2(ad^* - bc^*)$$

and so $ad^* - bc^* = 1$. When $d = 0$ we have $-c^*b = 1$ and so

$$-bc^* = b(-c^*b)B^{-1} = bb^{-1} = 1.$$

□

7.3 Clifford Hermitian forms

We now show how to use Hermitian matrices to construct Clifford Hermitian forms. We only do this for the two standard Clifford Hermitian matrices constructed above. As indicated above, $\left(\mathbb{V}^{n+1}\right)^2$ is not a vector space. So we confine our interest to all right multiples of the standard lift of a point in $\widehat{\mathbb{V}}^{n+1}$. Let $z \in \widehat{\mathbb{V}}^{n+1} = \mathbb{V}^{n+1} \cup \{\infty\}$ and define the standard lift of z to be

$$z \longmapsto \mathbf{z} = \begin{pmatrix} z \\ 1 \end{pmatrix} \qquad \text{for } z \in \mathbb{V}^{n+1}, \qquad \infty \longmapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

We can let Γ_n act by right multiplication on such vectors. That is, for all $\lambda \in \Gamma_n$ we define

$$\mathbf{z}\lambda = \begin{pmatrix} z \\ 1 \end{pmatrix} \lambda = \begin{pmatrix} z\lambda \\ \lambda \end{pmatrix}.$$

We can then use the Hermitian matrices H_1 and H_2 to define Hermitian forms on this collection of vectors as follows:

$$\begin{aligned}\langle \mathbf{z}\lambda, \mathbf{w}\mu\rangle_1 &= \overline{\mu}\mathbf{w}^* H_1 \mathbf{z}\lambda = \overline{\mu}(\overline{w}z - 1)\lambda, \\ \langle \mathbf{z}\lambda, \mathbf{w}\mu\rangle_2 &= \overline{\mu}\mathbf{w}^* H_2 \mathbf{z}\lambda = \overline{\mu}(i_n z - \overline{w} i_n)\lambda.\end{aligned}$$

In particular

$$\begin{aligned}\langle \mathbf{z}\lambda, \mathbf{z}\lambda\rangle_1 &= \overline{\lambda}\mathbf{z}^* H_1 \mathbf{z}\lambda = |\lambda|^2\big(|z|^2 - 1\big), \\ \langle \mathbf{z}\lambda, \mathbf{z}\lambda\rangle_2 &= \overline{\lambda}\mathbf{z}^* H_2 \mathbf{z}\lambda = -2|\lambda|^2 z_n.\end{aligned}$$

Hence $\langle \mathbf{z}\lambda, \mathbf{z}\lambda\rangle$ is real and we can define V_+, V_0 and V_- as before.

We may define projection to be the map which identifies all right multiples of $\mathbf{z}$, namely

$$\mathbb{P} : \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \longmapsto z_1 z_2^{-1} \quad \text{when } z_2 \neq 0, \qquad \mathbb{P} : \begin{pmatrix} z_1 \\ 0 \end{pmatrix} \longmapsto \infty.$$

For the first Hermitian form $\mathbb{P}V_-$ is the unit ball in $\mathbb{V}^{n+1}$ given by $|z| < 1$ and for the second Hermitian form $\mathbb{P}V_-$ is the upper half space in $\mathbb{V}^{n+1}$ given by $z_n > 0$. These are both models for hyperbolic $(n+1)$-space $\mathbf{H}^{n+1}$.

7.4 Clifford Möbius transformations

We want to let Clifford unitary matrices act on $\widehat{\mathbb{V}}^{n+1}$ via Möbius transformations. As in previous chapters, this will be done by writing $A(z) = \mathbb{P}A\mathbf{z}$. The main problem is that $\mathbb{P}$ is only defined for right multiples of the standard lift of vector. This is done to ensure that the image lies in $\widehat{\mathbb{V}}^{n+1}$. We need to impose further conditions on A so that $A\mathbf{z}$ has the required form.

We have a map from matrices to Möbius transformations as before

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto A(x) = (ax+b)(cx+d)^{-1} \tag{7.1}$$

whenever the right hand side is well defined. The kernel of this map is $\pm I$.

A necessary condition for the map $A(x) = (ax+b)(cx+d)^{-1}$ to preserve $\widehat{\mathbb{V}}^{n+1}$ is that $A(0)$, $A(\infty)$, $A^{-1}(0)$ and $A^{-1}(\infty)$ are all in $\widehat{\mathbb{V}}^{n+1}$. It turns out that this condition is sufficient. Now

$$A(0) = bd^{-1}, \quad A(\infty) = ac^{-1}, \quad A^{-1}(0) = -a^{-1}b, \quad A^{-1}(\infty) = -c^{-1}d$$

or else is ∞ (when $d = 0$, $c = 0$, $a = 0$ or $c = 0$ respectively). In particular, multiplying by $|d|^2$, $|c|^2$, $|a|^2$ or $|c|^2$ respectively, we see that we must have

$$b\overline{d},\ a\overline{c},\ \overline{a}b,\ \overline{c}d \in \mathbb{V}^{n+1}.$$

Proposition 7.4.1 *Let A be a 2×2 matrix with entries a, b, c, $d \in \Gamma_n \cup \{0\}$ with $ad^* - bc^* = 1$. Suppose that $b\overline{d}$, $a\overline{c}$, $\overline{a}b$, $\overline{c}d \in \mathbb{V}^{n+1}$. Then for all $z \in \widehat{\mathbb{V}}^{n+1}$ we have $A(z) = (az+b)(cz+d)^{-1} \in \widehat{\mathbb{V}}^{n+1}$.*

Proof: Since $a^{-1}b \in \mathbb{V}^{n+1}$ and $c^{-1}d \in \mathbb{V}^{n+1}$ (when $a \neq 0$ and $c \neq 0$ respectively) for all $z \in \mathbb{V}^{n+1}$ we have

$$az+b = a(z+a^{-1}b) \in \Gamma_n, \qquad cz+d = c(z+c^{-1}d) \in \Gamma_n.$$

Then, using $ac^{-1}d - b = (ad^* - bc^*)c^{*-1} = c^{*-1}$, we have

$$\begin{aligned}
(az+b)(cz+d)^{-1} &= a(z+a^{-1}b)(z+c^{-1}d)^{-1}c^{-1} \\
&= a\big((z+c^{-1}d)+(a^{-1}b-c^{-1}d)\big)(z+c^{-1}d)^{-1}c^{-1} \\
&= ac^{-1} + (b-ac^{-1}d)(z+c^{-1}d)^{-1}c^{-1} \\
&= ac^{-1} - (ad^*-bc^*)c^{*-1}(z+c^{-1}d)^{-1}c^{-1} \\
&= ac^{-1} - \big(c(z+c^{-1}d)c^*\big)^{-1}.
\end{aligned}$$

This lies in $\mathbb{V}^{n+1}$ since $c(z+c^{-1}d)c^* \in \mathbb{V}^{n+1}$ and so is its inverse. □

Proposition 7.4.2 *Suppose that A_1 and A_2 are two matrices that satisfy the hypotheses of Proposition 7.4.1. Then their product A_1A_2 also satisfies these hypotheses.*

Proof: When all the entries of A_1 and A_2 are non-zero, we have

$$A_1A_2 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}\begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1(a_2c_2^{-1}+a_1^{-1}b_1)c_2 & a_1(b_2d_2^{-1}+a_1^{-1}b_1)d_2 \\ c_1(a_2c_2^{-1}+c_1^{-1}d_1)c_2 & c_1(b_2d_2^{-1}+c_1^{-1}d_1)d_2 \end{pmatrix}.$$

So the entries of A_1A_2 lie in $\Gamma_n \cup \{0\}$. Similar, but simpler, formulae hold when some of the entries are zero.

Moreover,

$$(c_1b_2+d_1d_2)^* = \big(c_1(b_2d_2^{-1}+c_1^{-1}d_1)d_2\big)^* = d_2^*(d_2^{*-1}b_2^*+d_1^*c_1^{*-1})c_1^* = b_2^*c_1^*+d_2^*d_1^*$$

and so

$$\begin{aligned}
&(a_1a_2+b_1c_2)(c_1b_2+d_1d_2)^* - (a_1b_2+b_1d_2)(c_1a_2+d_1c_2)^* \\
&\quad= (a_1a_2+b_1c_2)(b_2^*c_1^*+d_2^*d_1^*) - (a_1b_2+b_1d_2)(a_2^*c_1^*+c_2^{''}d_1^*) \\
&\quad= a_1a_2d_2^*d_1^* + b_1c_2b_2^*c_1^* - a_1b_2c_2^*d_1^* - b_1d_2a_2^*c_1^* \\
&\quad= (a_1d_1^*-b_1c_1^*)(a_2d_2^*-b_2c_2^*) \;=\; 1.
\end{aligned}$$

Finally,

$$(a_1a_2+b_1c_2)(c_1a_2+d_1c_2)^{-1} = (a_1a_2c_2^{-1}+b_1)(c_1a_2c_2^{-1}+d_1) = A_1(a_2c_2^{-1}) \in \widehat{\mathbb{V}}^{n+1}$$

by Proposition 7.4.1. Similarly

$$(a_1b_2+b_1d_2)(c_1b_2+d_1d_2)^{-1} = A_1(b_2d_2^{-1}) \in \widehat{\mathbb{V}}^{n+1}.$$

The result follows. □

Corollary 7.4.3 *The set of* 2×2 *matrices with entries* a, b, c, $d \in \Gamma\cup\{0\}$ *satisfying the conditions of Proposition 7.4.1 form a group under matrix multiplication.*

Motivated by this, we define

$$\mathrm{SU}(H_1,\Gamma_n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \begin{array}{c} |a| = |d|,\ |b| = |c|,\ |a|^2 - |c|^2 = 1, \\ \overline{a}b = \overline{c}d \in \mathbb{V}^{n+1},\ a\overline{c} = b\overline{d} \in \mathbb{V}^{n+1} \end{array} \right\}.$$

Similarly, using Corollary 7.1.5, we must have

$$d^*b,\ c^*a,\ ba^*,\ dc^* \in \mathbb{V}^{n+1}.$$

Since $x^* = x$ for elements of $\mathbb{V}^n$ we must have $d^*b = (d^*b)^* = b^*d$ and so on. Motivated by this we define

$$\mathrm{SU}(H_2,\Gamma_{n-1}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : ad^* - bc^* = d^*a - b^*c = 1,\ ab^*,\ cd^*,\ c^*a,\ d^*b \in \mathbb{V}^n \right\}.$$

We define $\mathrm{PSU}(H_1,\Gamma_n)$ to be the image of $\mathrm{SU}(H_1,\Gamma_n)$ under the map (7.1) and $\mathrm{PSU}(H_2,\Gamma_{n-1})$ to be the image of $\mathrm{SU}(H_2,\Gamma_{n-1})$ under the map (7.1). These maps are *Clifford Möbius transformations*.

Proposition 7.4.4 *The set* $\mathrm{SU}(H_1,\Gamma_n)$ *is a group and is generated by*

$$D = \begin{pmatrix} \cosh(\lambda) & \sinh(\lambda) \\ \sinh(\lambda) & \cosh(\lambda) \end{pmatrix}, \qquad S = \begin{pmatrix} u & 0 \\ 0 & u' \end{pmatrix}$$

where $\lambda \in \mathbb{R}_+$ *and* $u \in \Gamma_n$ *with* $|u| = 1$.

Proof: Define λ by $\sinh(\lambda) = |b| = |c|$. Therefore $|a| = |d| = \cosh(\lambda)$. Using Corollary 7.1.6, we find u, $v \in \Gamma_n$ with $|u| = |v| = 1$ so that $a\overline{c} = |a|\,|c| uu^* \in \mathbb{V}^{n+1}$ and $\overline{b}a = |a|\,|b| v^*v \in \mathbb{V}^{n+1}$. Thus

$$\begin{aligned}
\frac{\overline{b}a}{|a|\,|b|} &= v^*v = v^*\overline{u}uv = \overline{(uv')}(uv), \\
\frac{\overline{d}c}{|d|\,|c|} &= v^*v = v^*u^*u'v = \overline{(u'v')}(u'v), \\
\frac{a\overline{c}}{|a|\,|c|} &= uu^* = uv\overline{v}u^* = (uv)\overline{(u'v)}, \\
\frac{b\overline{d}}{|b|\,|d|} &= uu^* = uv'v^*u^* = (uv')\overline{(u'v')}.
\end{aligned}$$

Then

$$a = \cosh(\lambda)uv, \quad b = \sinh(\lambda)uv', \quad c = \sinh(\lambda)u'v, \quad d = \cosh(\lambda)u'v'.$$

Therefore $ad^* - bc^* = 1$ and the hypotheses of Proposition 7.4.1 are satisfied. Hence by Corollary 7.4.3 the members of $\mathrm{SU}(H_1, \Gamma_n)$ form a group under matrix multiplication.

Moreover,

$$\begin{pmatrix} u & 0 \\ 0 & u' \end{pmatrix} \begin{pmatrix} \cosh(\lambda) & \sinh(\lambda) \\ \sinh(\lambda) & \cosh(\lambda) \end{pmatrix} \begin{pmatrix} v & 0 \\ 0 & v' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

□

Proposition 7.4.5 *The set* $\mathrm{SU}(H_2, \Gamma_{n-1})$ *is a group and is generated by*

$$D = \begin{pmatrix} a & 0 \\ 0 & a^{*-1} \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

where $a \in \Gamma_{n-1}$ *and* $t \in \mathbb{V}^n$.

Proof: If $c \neq 0$ then

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

We have used

$$ac^{-1}d - c^{*-1} = ad^*c^{*-1} - c^{*-1} = bc^*c^{*-1} = b.$$

□

Once again, we have the fact that if t is an eigenvalue of a Clifford matrix then so is $\overline{t}^{-1}$. If $|t| \neq 1$ then these eigenvalues are distinct and the corresponding eigenvectors are null. In this way we can define loxodromic maps. However it is not straightforward to find eigenvalues from the entries of a Clifford matrix.

7.5 The geometry of hyperbolic $(n+1)$-space

We define the hyperbolic metric by

$$ds^2 = \frac{-4}{\langle \mathbf{z}, \mathbf{z} \rangle^2} \det \begin{pmatrix} \langle \mathbf{z}, \mathbf{z} \rangle & \langle d\mathbf{z}, \mathbf{z} \rangle \\ \langle \mathbf{z}, d\mathbf{z} \rangle & \langle d\mathbf{z}, d\mathbf{z} \rangle \end{pmatrix}, \qquad \cosh^2\left(\frac{\rho(z,w)}{2}\right) = \frac{\langle \mathbf{z}, \mathbf{w} \rangle \langle \mathbf{w}, \mathbf{z} \rangle}{\langle \mathbf{z}, \mathbf{z} \rangle \langle \mathbf{w}, \mathbf{w} \rangle}.$$

For H_1 that is

$$ds^2 = \frac{dz\, d\overline{z}}{\left(1 - |z|^2\right)^2}, \quad \cosh^2\left(\frac{\rho(z,w)}{2}\right) = \frac{(\overline{w}z - 1)(\overline{z}w - 1)}{\left(|z|^2 - 1\right)\left(|w|^2 - 1\right)}.$$

We remark that the numerator is real, using Lemma 7.1.2. For H_2

$$ds^2 = \frac{dz\, d\overline{z}}{4z_n^2}, \qquad \cosh^2\left(\frac{\rho(z,w)}{2}\right) = \frac{(i_n z - \overline{w} i_n)(i_n w - \overline{z} i_n)}{(i_n z - \overline{z} i_n)(i_n w - \overline{w} i_n)} = \frac{|z - w|^2 + 4z_n w_n}{4z_n w_n}.$$

Theorem 7.5.1 *The collection of all orientation preserving isometries of hyperbolic $(n+1)$-space is* $\mathrm{PSU}(H_i, \Gamma_n)$ *where* $i = 1, 2$. *Furthermore, any orientation reversing isometry of hyperbolic $(n+1)$-space is* $A(-\overline{z})$ *where* $A(z) \in \mathrm{PSU}(H_i, \Gamma_n)$ *where* $i = 1, 2$.

Proof: We prove this in the same way that we proved Proposition 3.2.7. We work with the upper half space model. Let ϕ be any hyperbolic isometry. By applying a Clifford Möbius transformation, we may suppose that $\phi(i_n) = i_n$ and $\phi(2i_n) = yi_n$ for some $y > 1$. Arguing as in Proposition 3.2.7 we see that this implies that ϕ fixes the whole i_n axis.

Now suppose that $\phi(z) = w$. Then for all $y > 0$ we have

$$\begin{aligned} \frac{|z - yi_n|^2 + 4z_n y_n}{4z_n y_n} &= \cosh^2\left(\frac{\rho(z, yi_n)}{2}\right) \\ &= \cosh^2\left(\frac{\rho(w, yi_n)}{2}\right) \\ &= \frac{|w - yi_n|^2 + 4w_n y_n}{4w_n y_n}. \end{aligned}$$

From this we see that $z_n = w_n$ and $|z| = |w|$. Therefore $z - z_n i_n$ and $w - w_n i_n$ are in the same sphere in $\mathbb{V}^n$ and so there is $a \in \Gamma_{n-1}$ for which $w - w_n i_n = a(z - z_n i_n)a'^{-1}$ or $w - w_n i_n = a(-\overline{z} - z_n i_n)a'^{-1}$. Hence $w = aza'^{-1}$ or $w = a(-\overline{z})a'^{-1}$ which is in $\mathrm{PSU}(H_2, \Gamma_{n-1})$. $\square$

Exercise 7.5.1 *Show that in the upper half space model of* $\mathbf{H}^{n+1}$ *the* i_n *axis is a geodesic.*

In the unit ball model of $\mathbf{H}^{n+1}$, *describe the hyperbolic ball of radius r centred at the origin.*

Chapter 8

p-adic Möbius transformations

8.1 The p-adic numbers

Let X be a non-empty set. A distance or *metric* on X is a function ρ_p from pairs of elements (x, y) to the real numbers satisfying:

(i) $\rho(x, y) \geq 0$ with equality if and only if $x = y$;

(ii) $\rho(x, y) = \rho(y, x)$;

(iii) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ for all $z \in X$.

The inequality in (iii) is called the triangle inequality. A metric is said to be *non-Archimedean* if the triangle inequality is replaced with the following stronger inequality, called the *ultrametric inequality*:

(iv) $\rho(x, y) \leq \max\big\{\rho(x, z),\ \rho(z, y)\big\}$ for all $z \in X$.

A simple consequence of the ultrametric inequality is the fact that every triangle in a non-Archimedean metric space is isosceles:

Lemma 8.1.1 *Suppose that ρ is a non-Archimedean metric on a space X. If x, y and z are points of X so that $\rho(x, y) < \rho(x, z)$ then $\rho(x, z) = \rho(y, z)$.*

Proof: We have

$$\rho(y, z) \leq \max\Big\{\rho(x, y),\ \rho(x, z)\Big\} = \rho(x, z)$$

by hypothesis. Likewise,

$$\rho(x, z) \leq \max\Big\{\rho(x, y),\ \rho(y, z)\Big\} = \rho(y, z)$$

since otherwise we would have $\rho(x, z) \leq \rho(x, y)$ which would be a contradiction. Therefore, we have

$$\rho(y, z) \leq \rho(x, z) \leq \rho(y, z)$$

and hence these quantities are equal. $\square$

Many metrics arise from valuations on a ring. Let R denote a non-trivial ring. An *absolute value* (or *valuation* or *norm*) on R is a real valued function $x \longmapsto |x|$ on R satisfying:

(i) $|x| \geq 0$ with equality if and only if $x = 0$;

(ii) $|xy| = |x|\,|y|$;

(iii) $|x + y| \leq |x| + |y|$.

Once again, a valuation is said to be *non-Archimedean* if the inequality in (iii) is replaced with the stronger inequality:

(iv) $|x + y| \leq \max\big\{|x|,\ |y|\big\}$.

Given a valuation $|\ |$ on a ring R we may define a metric on R by:

$$\rho(x, y) = |x - y|.$$

For example, the standard absolute value on $\mathbb{R}$, $\mathbb{C}$ or or $\mathbb{H}$ gives rise to the Euclidean metric.

Fix a prime number p and let $r \in \mathbb{Q}$ be non-zero. Write $r = p^f u/v$ where $f \in \mathbb{Z}$ and u, v are coprime integers both of which are also coprime to p. Then define a valuation $|\ |_p$ on $\mathbb{Q}$ by:

$$|r|_p = p^{-f}, \qquad |0|_p = 0. \tag{8.1}$$

One can then show that $|r+s|_p \leq \max\big\{|r|_p,\ |s|_p\big\}$. This valuation is called the *p-adic valuation*. A rational number is p-adically small if it is divisible by a large power of p. We use the p-adic valuation $|\cdot|_p$ to define a metric ρ_p on $\mathbb{Q}$ by $\rho_p(x, y) = |x - y|_p$. This is called the *p-adic metric*. Two rational numbers are p-adically close if their difference is divisible by a large power of p.

The set of *p-adic numbers* $\mathbb{Q}_p$ is the completion of $\mathbb{Q}$ with respect to the p-adic valuation (8.1). The p-adic numbers form a field. Any p-adic number can be written as a semi-infinite power series in p, which resembles a Laurent series:

$$x = \sum_{n=k}^{\infty} x_n p^n. \tag{8.2}$$

Here k is any integer and $x_n \in \{0,\, 1,\, \ldots,\, p-1\}$ for all n with $x_k \neq 0$. Since we are summing over p^n for $n \geq k$ we see that p^k divides each term (and no other power of p does since $x_k \neq 0$). Therefore if x has the form (8.2) then $|x|_p = p^{-k}$. We can extend the p-adic metric to $\mathbb{Q}_p$. Let

$$x = \sum_{n=k}^{\infty} x_n p^n, \qquad y = \sum_{n=l}^{\infty} y_n p^n.$$

Then $\rho_p(x,y) = p^{-j}$ where j is the smallest index so that $x_j \neq y_j$ (where we put $x_n = 0$ for $n < k$ and $y_n = 0$ for $n < l$). In other words x and y are p-adically close if their series agree up to the term corresponding to a large power of p.

A *p-adic integer* is any p-adic number m with $|m|_p \leq 1$. Each p-adic integer m has an expansion

$$m = \sum_{n=0}^{\infty} m_n p^n \tag{8.3}$$

where $m_n \in \{0, 1, \ldots, p-1\}$. The p-adic integers form a ring, denoted $\mathbb{Z}_p$.

These p-adic expansions can be slightly surprising. For example, take $p = 5$. Consider $x = 1+5+5^2+5^3+\cdots$. It is clear that $5x = 5+5^2+5^3+5^4+\cdots = x-1$. Therefore $4x = -1$ and $x = -1/4$.

We can show that $\mathbb{Z}_p$ is compact (with respect to the topology induced by the p-adic valuation). Likewise, a *p-adic unit* is any non-zero p-adic number u so that $u \in \mathbb{Z}_p$ and $u^{-1} \in \mathbb{Z}_p$. That is, u has the form (8.3) with $m_0 \neq 0$. Since the set of units is the intersection of two compact subsets of $\mathbb{Q}_p$, we see that it is compact.

By definition, $\mathbb{Z}_p$ is the p-adic unit ball in $\mathbb{Q}_p$, that is the ball $B_1(0)$ of radius 1 centred at 0. That is, it is the set of p-adic numbers whose p-adic norm is at most 1. If y is any p-adic number, then the unit ball centred at y is

$$B_1(y) = \Big\{ x \in \mathbb{Q}_p \ : \ \rho_p(x,y) = |x-y|_p \leq 1 \Big\}.$$

Note that this means that $x = y + m$ where m is a p-adic integer. Therefore we denote this ball by $y + \mathbb{Z}_p$. Note that if $y \in \mathbb{Z}_p$ then $y + \mathbb{Z}_p = \mathbb{Z}_p$. This means that any point inside a p-adic unit ball is its centre!

It is clear that $p^k\mathbb{Z}_p$ is the p-adic ball centred at the origin (or any other point of its interior) of radius p^{-k}. Likewise $y + p^k\mathbb{Z}_p$ is the p-adic ball centred at y of radius p^{-k}. Again any point in the ball $y + p^k\mathbb{Z}_p$ may be taken to be its centre.

8.2 p-adic matrices and Möbius transformations

In this section we consider

$$\mathrm{SL}(2,\mathbb{Q}_p) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ : \ a,\, b,\, c,\, d \in \mathbb{Q}_p,\ ad - bc = 1 \right\}.$$

There is a natural projection to p-adic Möbius transformations, namely

$$\mathrm{PSL}(2,\mathbb{Q}_p) = \{ A(x) = (ax+b)/(cx+d) \ : \ a,\, b,\, c,\, d \in \mathbb{Q}_p,\ ad-bc = 1. \}.$$

These act on $\mathbb{Q}_p \cup \{\infty\} = \widehat{\mathbb{Q}}_p$.

For example, consider

$$A = \begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix}, \quad D = \begin{pmatrix} p^m & 0 \\ 0 & p^{-m} \end{pmatrix}, \quad R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{8.4}$$

where a is a unit in $\mathbb{Q}_p$ and b is any element of $\mathbb{Q}_p$.

Proposition 8.2.1 *The group* $\mathrm{SL}(2,\mathbb{Q}_p)$ *is generated by* A, D, R *as given in (8.4) where* a *is a unit,* b *is any element of* $\mathbb{Q}_p$ *and* m *is an integer.*

Proof: If $c = 0$ then write $a = p^k u$ where u is a unit. Then $d = p^{-k}u^{-1}$.

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} p^k & 0 \\ 0 & p^{-k} \end{pmatrix} \begin{pmatrix} u & p^{-k}b \\ 0 & u^{-1} \end{pmatrix}$$

Suppose $c \neq 0$. Write $c = p^k u$ where u is a unit. Then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & ac^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} p^k & 0 \\ 0 & p^{-k} \end{pmatrix} \begin{pmatrix} u & p^{-k}d \\ 0 & u^{-1} \end{pmatrix}.$$

□

As Möbius transformations these act as

$$A(x) = (ax+b)a, \qquad D(x) = p^{2m}x, \qquad R(x) = -1/x.$$

Now it is not hard to show that

$$\rho_p\big(A(x), A(y)\big) = \big|A(x) - A(y)\big|_p = \big|a(x-y)a\big|_p = \big|x-y\big|_p = \rho_p(x,y)$$

so that $A(x)$ is an isometry of the metric ρ_p. Similarly,

$$\rho_p\big(D(x), D(y)\big) = \big|D(x) - D(y)\big|_p = |p^{2m}x - p^{2m}y|_p = |p^{2m}|_p|x-y|_p = p^{-2m}\rho_p(x,y).$$

Thus $D(x)$ is a dilation. Likewise,

$$\begin{aligned} \rho_p\big(R(x), 0\big) &= \big|R(x)\big|_p = \left|\frac{-1}{x}\right|_p = \frac{1}{|x|_p} = \frac{1}{\rho_p(x,0)}, \\ \rho_p\big(R(x), R(y)\big) &= \big|R(x) - R(y)\big|_p = \left|\frac{-1}{x} - \frac{-1}{y}\right|_p = \frac{|x-y|_p}{|x|_p|y|_p} \\ &= \frac{\rho_p(x,y)}{\rho_p(x,0), \rho_p(y,0)}. \end{aligned}$$

8.3 The tree T_p

We now show how to define a tree T_p whose boundary is $\mathbb{Q}_p \cup \{\infty\}$. The closed balls in $\mathbb{Q}_p$ are the vertices of T_p, that is

$$V = \Big\{ x + p^k\mathbb{Z}_p \ : x \in \mathbb{Q}_p, \ k \in \mathbb{Z} \Big\}.$$

Two vertices $x + p^k\mathbb{Z}_p$ and $y + p^j\mathbb{Z}_p$ are joined by an edge of T_p if and only if one is a maximal ball properly contained in the other. In other words, either $k = j+1$ and $x - y \in p^j\mathbb{Z}_p$ or else $j = k+1$ and $x - y \in p^k\mathbb{Z}_p$.

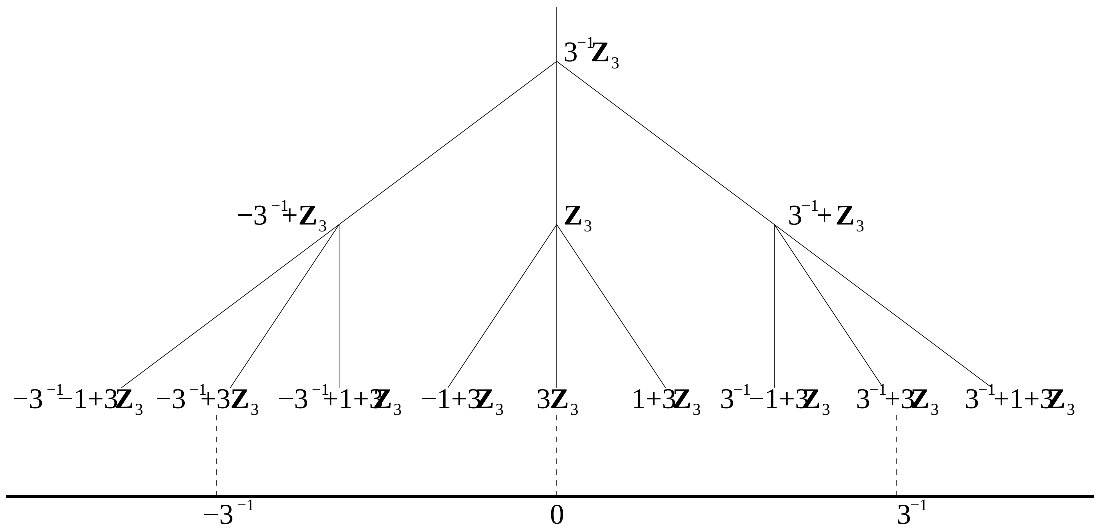

Figure 8.1: *Part of the tree T_3 for $\mathbb{Z}_3$.*

Notice that each ball $x + p^k\mathbb{Z}_p$ of radius p^{-k} is contained in exactly one ball of radius p^{-k+1}, namely $x + p^{k-1}\mathbb{Z}_p$. Similarly, $x + p^k\mathbb{Z}_p$ contains exactly p balls of radius p^{-k-1}, namely $x+yp^k+p^{k+1}\mathbb{Z}_p$ where $y = 0, 1, \ldots, p-1$. Hence each vertex has exactly $p + 1$ edges emanating from it. Therefore the graph T_p we have just constructed is an infinite, regular $p + 1$ tree. The tree T_p has a natural metric, namely the graph metric where each edge has length 1.

We divide the vertices into two classes, namely those for which k is even denoted V^+ and those for which k is odd, denoted V^-:

$$V^+ = \Big\{x + p^{2k}\mathbb{Z}_p \ : x \in \mathbb{Q}_p,\ k \in \mathbb{Z}\Big\}, \quad V^- = \Big\{x + p^{2k+1}\mathbb{Z}_p \ : x \in \mathbb{Q}_p,\ k \in \mathbb{Z}\Big\}.$$

By construction, each edge has one end in V^+ and the other in V^-.

Example. For example, in Figure 8.1 we illustrate the tree T_3 in the case where $p = 3$. (We choose to take the coefficients in (8.2) from $\{-1, 0, 1\}$ rather than $\{0, 1\,2\}$.) In this case, each vertex has the form $x + 3^k\mathbb{Z}_3$. Such a vertex is linked to the vertices $x+3^{k-1}\mathbb{Z}_3$, $x-3^k+3^{k+1}\mathbb{Z}_3$, $x+3^{k+1}\mathbb{Z}_3$ and $x+3^k+3^{k+1}\mathbb{Z}_3$ by edges. Thus T_3 is a regular 4-valent tree. We also indicate the geodesic from 0 to ∞ which passes through the vertices $\ldots, 3\mathbb{Z}_3, \mathbb{Z}_3, 3^{-1}\mathbb{Z}_3, \ldots$ and the geodesic from -3^{-1} to 3^{-1} passing through the vertices $\ldots, -3^{-1}+3\mathbb{Z}_3, -3^{-1}+\mathbb{Z}_3, 3^{-1}\mathbb{Z}_3, 3^{-1}+\mathbb{Z}_3, 3^{-1}+3\mathbb{Z}_3, \ldots$.

We now find the boundary of T_p. We consider geodesic paths through T_p (with respect to the graph metric). In other words, such a path is a (possibly infinite) sequence of vertices v_j so that for all j the vertices v_j, v_{j+1} are joined by an edge and $v_{j-1} \neq v_{j+1}$, that is there is no back tracking. The semi-infinite geodesic path $p^{-k}\mathbb{Z}_p$ for $k = 0, 1, 2, \ldots$ identifies a point of the boundary denoted by ∞. Every other semi-infinite geodesic path starting at the vertex $\mathbb{Z}_p$ eventually consists of a sequence of nested, decreasing balls $x + p^k\mathbb{Z}_p$ for $k = K, K + 1, K + 2, \ldots$. The limit of this sequence is the point x of $\mathbb{Q}_p$. Choosing a starting point other than $\mathbb{Z}_p$ makes only finitely many changes to these paths. Hence the boundary of T_p is $\mathbb{Q}_p \cup \{\infty\}$. From this construction ∂T_p is totally disconnected.

Any two distinct points z, w in $\mathbb{Q}_p \cup \{\infty\}$ are the end points of a unique doubly infinite geodesic path through T_p. We denote this path by $\gamma(z, w)$.

We may also define horocycles. The *horocycle H_k centred at ∞ of height p^k* is defined to be the collection of all vertices in T_p corresponding to balls of radius p^k. In other words

$$H_k = \Big\{x + p^{-k}\mathbb{Z}_p \ : \ x \in \mathbb{Q}_p\Big\}.$$

Likewise, the *horoball B_k centred at ∞ of height p^k* is the union of all horocycles of height greater than k. That is

$$B_k = \Big\{x + p^j\mathbb{Z}_p \ : \ x \in \mathbb{Q}_p, \ j < -k\Big\}.$$

This is suggested in Figure 8.1 where points on the same horocycle are drawn on the same horizontal line.

We can define the p-adic topology on $\mathbb{Q}_p \subset \partial T_p$ by measuring the maximum height of any point on the geodesic joining two boundary points.

8.4 Automorphisms of the tree T_p

An automorphism of T_p is a bijection from T_p to itself that preserves adjacency. Tits classified tree automorphisms into three types. Let $A : T \longrightarrow T$ be a tree automorphism. Then

(i) A is called a *rotation* if there is a vertex v of T_p so that $A(v) = v$.

(ii) A is called an *inversion* if there are two adjacent vertices v and w of T_p so that A interchanges v and w. This means that A preserves the edge of T_p with endpoints v and w but reverses its orientation.

(iii) A is called a *translation* if there is a geodesic $\{v_n \ : \ n \in \mathbb{Z}\}$ and a non-zero integer k so that $A(v_n) = v_{n+k}$ for all $n \in \mathbb{Z}$.

The first two classes of automorphism correspond to elliptic maps and reflections respectively. The third class corresponds to loxodromic maps when k is even and to glide reflections when k is odd.

We now describe the action of A, D and R on T_p. Consider the vertex $v = x + p^j\mathbb{Z}_p$. This is the p-adic ball centred at x with radius p^{-j}. The image of v under A or D should be a p-adic ball (that is another vertex of T_p). We require that a point $z \in \mathbb{Q}_p$ lies in v if and only if $A(z)$ lies in $A(v)$ (or $D(z)$ lies in $D(v)$ respectively). We must have $z = x + p^j w$ where $w \in \mathbb{Z}_p$. Then

$$\begin{aligned} A(z) &= (a(x + p^j w) + b)a \ = \ (ax + b)a + p^j a^2 w \ = \ A(x) + p^j a^2 w, \\ D(z) &= p^{2m}(x + p^j w) \ = \ p^{2m}x + p^{j+2m}w \ = \ D(x) + p^{j+2m}w. \end{aligned}$$

As w ranges over $\mathbb{Z}_p$ then, since a is a unit, a^2w also ranges over the whole of $\mathbb{Z}_p$. Therefore

$$\begin{aligned} A(v) &= A(x+p^j\mathbb{Z}_p) = (ax+b)a+p^j\mathbb{Z}_p = A(x)+p^j\mathbb{Z}_p, && (8.5)\\ D(v) &= D(x+p^j\mathbb{Z}_p) = p^{2m}x+p^{j+2m}\mathbb{Z}_p = D(x)+p^{j+2m}\mathbb{Z}_p. && (8.6) \end{aligned}$$

Thus the image of each vertex of T_p under A and D is also a vertex of T_p.

The action of R on T_p is slightly more complicated. Consider the vertex $v = x+p^j\mathbb{Z}_p$ and suppose for the moment that 0 is not in v. In other words, suppose that x is not in $p^j\mathbb{Z}_p$. We may write $x = p^k u$ where u is a unit. Our hypothesis is that $k < j$. Let $y \in \mathbb{Q}_p$ be given by $y = R(x) = -x^{-1} = p^{-k}u'$ where u' is a unit so that $uu' = -1$. We can write a general point of v as $x+p^jw = p^ku+p^jw$. We claim that $R(p^ku+p^jw) - p^{-k}u' \in p^{j-2k}\mathbb{Z}_p$. In order to see this, observe that

$$\begin{aligned} (p^ku+p^jw)\Big(R(p^ku+p^jw)-p^{-k}u'\Big) &= -1-(p^ku+p^jw)p^{-k}u' \\ &= -1+1-p^{j-k}wu' \\ &= -p^{j-k}wu'. \end{aligned}$$

Because $wv \in \mathbb{Z}_p$ all powers of p in the expansion of the right hand side, and hence also of the left hand side, must be at least p^{j-k}. Moreover, since $k < j$ there is a factor of p^k in (p^ku+p^jw). Thus the smallest power of p in the expansion of $R(p^ku+p^jw)-p^{-k}u'$ must be at least p^{j-2k}. Hence $R(p^ku+p^jw)-p^{-k}u' \in p^{j-2k}\mathbb{Z}_p$ as claimed. Therefore $R(v) = R(p^ku+p^j\mathbb{Z}_p) \subset p^{-k}u'+p^{j-2k}\mathbb{Z}_p$. A similar argument shows that $p^{-k}+p^{j-2k}\mathbb{Z}_p \subset R(p^ku+p^j\mathbb{Z}_p)$ and so these two balls are equal. That is, for $k < j$

$$R(v) = R(p^ku+p^j\mathbb{Z}_p) = p^{-k}u'+p^{j-2k}\mathbb{Z}_p = R(p^ku)+p^{j-2k}\mathbb{Z}_p. \qquad (8.7)$$

Thus when 0 is not contained in v the image of v under R is also a vertex of T_p.

In principle, the same argument works when $0 \in v$. If a ball contains 0 then we may write it as $p^j\mathbb{Z}_p$ for some integer j. The image under R of a ball centred at 0 should be a ball centred at $R(0) = \infty$. Therefore we must make sense of p-adic balls centred at ∞. Just as for the Riemann sphere, a neighbourhood of ∞ is the exterior of a bounded set. Thus balls centred at ∞ are the exteriors of finite balls. Consider $v = p^j\mathbb{Z}_p$. Then certainly v contains points of the form p^ju where u is a unit. Then $R(p^ju) = -p^{-j}u^{-1} \in p^{-j}\mathbb{Z}_p$. However $p^j\mathbb{Z}_p$ also contains points p^ku where $k > j$ and u is a unit. For such points $R(p^ku) = -p^{-k}u^{-1}$ which lies in the exterior of $p^{-j}\mathbb{Z}_p$. Therefore we define

$$R(v) = R(p^j\mathbb{Z}_p) = p^{-j}\mathbb{Z}_p. \qquad (8.8)$$

Proposition 8.4.1 *Suppose that A, D and R act on the vertices of T_p via (8.5), (8.6), (8.7) and (8.8). Then A, D and R act on the edges of T_p.*

Proof: Let $x + p^j\mathbb{Z}_p$ and $y + p^{j+1}\mathbb{Z}_p$ where $x - y \in p^j\mathbb{Z}_p$ be adjacent vertices of T_p.

Then $A(x) - A(y) = (ax + b)a - (ay + b)a = a^2(x - y) \in p^j\mathbb{Z}_p$ as a is a unit. Therefore $A(x+p^j\mathbb{Z}_p) = A(x)+p^j\mathbb{Z}_p$ and $A(y+p^{j+1}\mathbb{Z}_p) = A(y)+p^{j+1}\mathbb{Z}_p$ are adjacent vertices of T_p.

Similarly, we have $D(x) - D(y) = p^{2m}x - p^{2m}y = p^{2m}(x - y) \in p^{j+2m}\mathbb{Z}_p$. Thus $D(x+p^j\mathbb{Z}_p) = D(x)+p^{j+2m}\mathbb{Z}_p$ and $D(y+p^{j+1}\mathbb{Z}_p) = D(y)+p^{j+2m+1}\mathbb{Z}_p$ are adjacent vertices of T_p.

Suppose that $x+p^j\mathbb{Z}_p$ and $y+p^{j+1}\mathbb{Z}_p$ neither contain 0. In other words $x \notin p^j\mathbb{Z}_p$ and $y \notin p^{j+1}\mathbb{Z}_p$. We write $x = p^k u$ and $y = p^l v$ where $k < j$, $l < j + 1$ and u, v are units. If x and y are adjacent then $x - y = p^k u - p^l v \in p^j\mathbb{Z}_p$ and so, in fact, $k = l$ and $u - v \in p^{j-k}\mathbb{Z}_p$ Then $R(x) = -p^{-k}u^{-1}$ and $R(y) = -p^{-k}v^{-1}$. Thus

$$R(x) - R(y) = -p^{-k}u^{-1} + p^{-k}v^{-1} = p^{-k}u^{-1}v^{-1}(u - v) \in p^{j-2k}\mathbb{Z}_p$$

since u^{-1} and v^{-1} are both units. Therefore $R(x + p^j\mathbb{Z}_p) = R(x) + p^{j-2k}\mathbb{Z}_p$ and $R(y + p^{j+1}\mathbb{Z}_p) = R(y) + p^{j-2k+1}\mathbb{Z}_p$ are adjacent vertices.

Suppose now that $x+p^j\mathbb{Z}_p$ contains 0 but $y+p^{j+1}\mathbb{Z}_p$ does not. Therefore we may take $x = 0$. This means that $y \in p^j\mathbb{Z}_p$. Because we assumed that $y + p^{j+1}\mathbb{Z}_p$ does not contain 0 we must have $y = p^j u$ where u is a unit. We have $R(p^j\mathbb{Z}_p) = p^{-j}\mathbb{Z}_p$ and $R(p^j u + p^{j+1}\mathbb{Z}_p) = -p^{-j}u^{-1} + p^{1-j}\mathbb{Z}_p$. Then $-p^{-j}u^{-1} \in p^{-j}\mathbb{Z}_p$ and so $R(p^j\mathbb{Z}_p)$ and $R(y + p^{j+1}\mathbb{Z}_p)$ are adjacent vertices.

Finally, suppose that $x + p^j\mathbb{Z}_p$ and $y + p^{j+1}\mathbb{Z}_p$ both contain 0. Then we may suppose that x and y are both 0. Then $R(p^j\mathbb{Z}_p) = p^{-j}\mathbb{Z}_p$ and $R(p^{j+1}\mathbb{Z}_p) = p^{-j-1}\mathbb{Z}_p$ which are adjacent vertices. □

Theorem 8.4.2 *There is an injective group homomorphism from* $\mathrm{PSL}(2, \mathbb{Q}_p)$ *to the automorphism group of* T_p*. The action of* $\mathrm{PSL}(2, \mathbb{Q}_p)$ *preserves the sets* V^+ *and* V^- *and acts transitively on each of these sets. Non-trivial maps in the image are either rotations or even translations.*

Proof: In Proposition 8.4.1 we showed (8.5), (8.6), (8.7) and (8.8) define actions of $A(x)$, $D(x)$ and $R(x)$ as automorphisms of T_p. Since $A(x)$, $D(x)$ and $R(x)$ generate $\mathrm{PSL}(2, \mathbb{Q}_p)$, this action may be extended to a homomorphism from $\mathrm{PSL}(2, \mathbb{Q}_p)$ to the automorphism group of T_p. If an element B of $\mathrm{PSL}(2, \mathbb{Q}_p)$ is mapped to the identity it must fix all points of T_p. That is it fixes all balls in $\mathbb{Q}_p$. By considering families of nested balls, we see that B fixes all points of the boundary of T_p, that is all points of $\mathbb{Q}_p$. Hence B was in fact the identity. Hence this homomorphism is injective.

It is also obvious that $A(x)$, $D(x)$ and $R(x)$ preserve V^+ and V^-. Therefore the associated tree automorphisms can only be rotations or even translations.

In order to show that the action is transitive on V^+ we show how to map $\mathbb{Z}_p$ to $x + p^{2k}\mathbb{Z}_p$. This is achieved by

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p^k & 0 \\ 0 & p^{-k} \end{pmatrix} = \begin{pmatrix} p^k & p^{-k}x \\ 0 & p^{-k} \end{pmatrix}.$$

Thus $\mathbb{Z}_p \longmapsto p^{2k}\mathbb{Z}_p \longmapsto x + p^{2k}\mathbb{Z}_p$. $\square$

We now investigate the action a little more closely. Consider

$$A = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

where $t = p^j u \in \mathbb{Q}_p$ where u is a unit. It is clear that if $j \geq i + k$ then

$$A\big(p^i(x + p^k\mathbb{Z}_p)\big) = p^i(x + p^k\mathbb{Z}_p) + p^j u = p^i(x + p^k\mathbb{Z}_p).$$

Hence the vertex $p^i(x + p^k\mathbb{Z}_p)$ is fixed by A.

Example. Let $p = 3$. Let $x_0 \in \mathbb{Z}_3$ have 3-adic expansion $x_0 = 1 + 3 + 3^2 + 3^3 + \cdots$. Then

$$2x_0 = 3x_0 - x_0 = (3 + 3^2 + 3^3 + 3^4 + \cdots) - (1 + 3 + 3^2 + 3^3 + \cdots) = -1$$

so $x_0 = -1/2$. Consider $A(x) = x + 1$. Then

$$\begin{array}{rcccccccccl} A(x_0) & = & A(-1/2) & = & -1/2+1 & = & 1/2 & = & -x_0 & = & -1-3-3^2-3^3-\cdots, \\ A^2(x_0) & = & A^2(-1/2) & = & -1/2+2 & = & 3/2 & = & -3x_0 & = & -3-3^2-3^3-3^4-\cdots. \end{array}$$

Thus $x_0 \in 1+3\mathbb{Z}_3$, $A(x_0) \in -1+3\mathbb{Z}_3$, $A^2(x_0) \in 3\mathbb{Z}_3$ and $A^3(x_0) \in 1+3\mathbb{Z}_3$. Therefore $A(x_0)$ fixes $\mathbb{Z}_3$ and cyclically permutes $1 + 3\mathbb{Z}_3$, $-1 + 3\mathbb{Z}_3$ and $3\mathbb{Z}_3$.

On the next level, similar arguments give:

$$\begin{array}{lll} x \in 1+3+3^2\mathbb{Z}_3, & A(x) \in -1-3+3^2\mathbb{Z}_3, & A^2(x) \in -3+3^2\mathbb{Z}_3, \\ A^3(x) \in 1-3+3^2\mathbb{Z}_3, & A^4(x) \in -1+3^2\mathbb{Z}_3, & A^5(x) \in 3^2\mathbb{Z}_3, \\ A^6(x) \in 1+3^2\mathbb{Z}_3, & A^7(x) \in -1+3+3^2\mathbb{Z}_3, & A^8(x) \in 3+3^2\mathbb{Z}_3. \end{array}$$

Proposition 8.4.3 $\mathrm{SL}(2,\mathbb{Z}_p)$ *fixes the vertex* $\mathbb{Z}_p$.

Proof: Let $A \in \mathrm{SL}(2,\mathbb{Z}_p)$ have entries a, b, c, $d \in \mathbb{Z}_p$ with $ad - bc = 1$. If $c = 0$ then we have $ad = 1$ and so a and d are units. It is clear that

$$\mathbb{Z}_p \longmapsto ad^{-1}\mathbb{Z}_p + bd^{-1} = \mathbb{Z}_p.$$

For example suppose that a, b, c, d are all in $\mathbb{Z}_p$ with $ad - bc = 1$. If c is a unit then ac^{-1} and dc^{-1} are in $\mathbb{Z}_p$. Writing

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

we see that each of these four matrices preserve $\mathbb{Z}_p$.

Now suppose that a, b, c, d are all in $\mathbb{Z}_p$ with $ad-bc=1$ and $c=up^j$ where $j>0$ and u is a unit. Note that this implies a and d are units. Consider

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

(Note that these matrices are not all in $\mathrm{SL}(2,\mathbb{Z}_p)$.) Then applying these matrices in turn, we obtain

$$\mathbb{Z}_p \longmapsto p^{-j}(du^{-1}+p^j\mathbb{Z}_p) \longmapsto p^j(du+p^j\mathbb{Z}_p) \longmapsto p^{-j}(-au^{-1}+p^j\mathbb{Z}_p) \longmapsto \mathbb{Z}_p.$$

This proves the result. $\square$

Chapter 9

Rank 1 symmetric spaces of non-compact type

9.1 Symmetric spaces

Let X be a simply connected, complete, geodesic metric space. Consider a point $x \in X$ and a local geodesic $\gamma : \mathbb{R} \longrightarrow X$ parametrised by arc length and with $\gamma(0) = x$. In other words, for all points s and t of $\mathbb{R}$ that are sufficiently close the arc $\gamma : [s, t] \longrightarrow X$ is the geodesic arc between $\gamma(s)$ and $\gamma(t)$. However for long arcs this arc may not be the shortest path. This is the case for great circles on a sphere.

We say that a map $\Phi : \gamma \longrightarrow \gamma$ is a reflection of γ in x if $\Phi\big(\gamma(t)\big) = \gamma(-t)$. We say that X is a *symmetric space* if for each point x the map given by reflection in x of each geodesic through x maps X to itself and is an isometry. Note that this implies that X is homogeneous. For if we take any two points y, $z \in X$ there is a geodesic arc between them. Let x be the midpoint of this arc. Reflection in x maps y to z.

Simple examples of symmetric spaces are spheres of all dimensions, Euclidean space of all dimensions and hyperbolic space of all dimensions.

By construction the map $\gamma : \mathbb{R} \longrightarrow X$ sending $\mathbb{R}$ to a geodesic parametrised by arc length is a locally isometric embedding of the Euclidean line to X. Consider $\mathbb{R}^n$ with its standard Euclidean metric. Suppose we can find a map $\gamma : \mathbb{R}^n \longrightarrow X$ that maps $\mathbb{R}^n$ locally isometrically to X. Then we call the image of $\mathbb{R}^n$ a *flat* subspace of X. The *rank* of X is the largest dimension of a flat subspace of X.

For example, spheres and hyperbolic spaces each have rank 1 whereas Euclidean n-space has rank n.

A symmetric space is of *compact type* if it is compact and of *non-compact* type if it is non-compact. Therefore S^2 is a rank 1 symmetric space of compact type and $\mathbf{H}^2$ is a rank 1 symmetric space of non-compact type. In fact these two examples are indicative of all rank one symmetric spaces.

Theorem 9.1.1 *Let X be a rank 1 symmetric space. Then*

(i) *If X has compact type then X is one of the following: a sphere S^n, a complex projective space $\mathbb{CP}^n$, a quaternionic projective space $\mathbb{HP}^n$ or the octonionic projective plane $\mathbb{OP}^2$.*

(ii) *If X has compact type then X is one of the following: (real) hyperbolic space $\mathbf{H}^n_{\mathbb{R}}$, complex hyperbolic space $\mathbf{H}^n_{\mathbb{C}}$, quaternionic hyperbolic space $\mathbf{H}^n_{\mathbb{H}}$ or the octonionic hyperbolic plane $\mathbf{H}^2_{\mathbb{O}}$.*

The spaces in (ii) are called *the hyperbolic spaces*.

9.2 Real, complex and quaternionic hyperbolic spaces

The construction of the hyperbolic spaces and their isometries follows along the same lines that we have been discussing earlier. In the case of the real, complex and quaternionic hyperbolic spaces the general construction is the following. Consider $\mathbb{R}^{n,1}$, $\mathbb{C}^{n,1}$ or $\mathbb{H}^{n,1}$ the vector space over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ with dimension $n+1$ equipped with a Hermitian (or quadratic) form H of signature $(n,1)$. Write $\langle \mathbf{z}, \mathbf{w}\rangle = \mathbf{w}^* H \mathbf{z}$. There are two standard choices of Hermitian form, which we have seen when dealing with the Klein model of $\mathbf{H}^n$; compare (4.1):

$$H_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & -1 \end{pmatrix},$$

$$H_2 = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Let V_+, V_0 and V_- be the positive, null and negative vectors respectively. Define a right projection from $\mathbb{R}^{n,1}$, $\mathbb{C}^{n,1}$ or $\mathbb{H}^{n,1}$ to $\mathbb{RP}^n$, $\mathbb{CP}^n$ or $\mathbb{HP}^n$ respectively. Then the corresponding hyperbolic space is $\mathbb{P}V_-$ and its ideal boundary is $\mathbb{P}V_0$. In each case we define the metric by

$$ds^2 = \frac{-\kappa^2}{\langle \mathbf{z}, \mathbf{z}\rangle^2} \det \begin{pmatrix} \langle \mathbf{z}, \mathbf{z}\rangle & \langle \mathbf{z}, d\mathbf{z}\rangle \\ \langle d\mathbf{z}, \mathbf{z}\rangle & \langle d\mathbf{z}, d\mathbf{z}\rangle \end{pmatrix}$$

where κ is a positive real number. Once again, this leads to a distance formula

$$\cosh^2\left(\frac{\rho(z,w)}{\kappa}\right) = \frac{\langle \mathbf{z}, \mathbf{w}\rangle \langle \mathbf{w}, \mathbf{z}\rangle}{\langle \mathbf{z}, \mathbf{z}\rangle \langle \mathbf{w}, \mathbf{w}\rangle}$$

We have seen that there is an isometry between complex hyperbolic 1-space (the Poincaré models) and real hyperbolic 2-space (the Klein model). Likewise, there is an isometry between quaternionic hyperbolic 1-space (as discussed in Chapter 6) and real hyperbolic 4-space. Similar identifications do not arise in other dimensions.

Let $\mathrm{O}(n,1)$, $\mathrm{U}(n,1)$ and $\mathrm{Sp}(n,1)$ be the groups of matrices preserving the form H in each case. Then entries lie in $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ respectively. Since they preserve the form it is clear that they preserve the metric on the respective hyperbolic space.

We can define geodesics, totally geodesic subspaces, spheres, balls, horospheres and horoballs as before, although there are some subtle differences. For example, for $\mathbb{C}$ and $\mathbb{H}$ and $n \geq 2$ there are no totally geodesic real hypersurfaces and horospheres naturally carry the structure of a Heisenberg group (that is nil geometry) rather than Euclidean geometry.

9.3 The octonionic hyperbolic plane

There is still one remaining case, namely the octonionic hyperbolic plane. The octonions $\mathbb{O}$ comprise the real vector space spanned by 1 and i_j for $j = 1, \ldots, 7$ together with a non-associative multiplication defined on the basis vectors as follows and then extended to the whole of $\mathbb{O}$ by linearity. First for j, $k = 1, \ldots, 7$ and $j \neq k$ we have

$$1i_j = i_j 1 = i_j, \quad i_j^2 = -1, \quad i_j i_k = -i_k i_j.$$

Finally

$$i_j i_k = i_l$$

precisely when (j, k, l) is a cyclic permutation of one of the triples:

$$(1,2,4), \quad (1,3,7), \quad (1,5,6), \quad (2,3,5), \quad (2,6,7), \quad (3,4,6), \quad (4,5,7).$$

We write an octonion z as

$$z = z_0 + \sum_{j=1}^{7} z_j i_j.$$

Define the *conjugate* $\overline{z}$ of z to be

$$\overline{z} = z_0 - \sum_{j=1}^{7} z_j i_j.$$

Conjugation is an anti-automorphism, that is for all octonions z and w

$$\overline{(zw)} = \overline{w}\,\overline{z}.$$

The *real part* of z is $\mathrm{Re}(z) = \frac{1}{2}(z+\overline{z})$ and the *imaginary part* of z is $\mathrm{Im}(z) = \frac{1}{2}(z-\overline{z})$. The *modulus* $|z|$ of an octonion is the non-negative real number defined by

$$|z|^2 = \overline{z}z = z\overline{z} = {z_0}^2 + {z_1}^2 + \cdots + {z_7}^2.$$

The modulus is multiplicative, that is $|zw| = |z|\,|w|$. Clearly $|z| > 0$ unless $z = 0$ (that is $z_0 = z_1 = \cdots = z_7 = 0$). An octonion z is a *unit* if $|z| = 1$.

The following result is due to Artin

Proposition 9.3.1 *For any octonions x and y the subalgebra with a unit generated by x and y is associative. In particular, any product of octonions that may be written in terms of just two octonions is associative.*

The octonions are not associative and so we loose the notion of a vector space. The basic idea here is to use the fact that two generator subalgebras of $\mathbb{O}$ are associative (and are isomorphic to $\mathbb{H}$).

Consider $z = (z_1, z_2)$ where $z_1, z_2 \in \mathbb{O}$. Then the standard lift of z is the "octonionic vector"

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ 1 \end{pmatrix}.$$

Suppose that λ is an octonion in the same associative subalgebra of $\mathbb{O}$ as the entries of $\mathbf{z}$. Then we can let λ act on $\mathbf{z}$ by right multiplication.

$$\mathbf{z}\lambda = \begin{pmatrix} z_1 \\ z_2 \\ 1 \end{pmatrix} \lambda = \begin{pmatrix} z_1\lambda \\ z_2\lambda \\ \lambda \end{pmatrix}.$$

Motivated by this, we define

$$\mathbb{O}_0^3 = \left\{ \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \ : \ z_1,\, z_2,\, z_3 \text{ all lie in some associative subalgebra of } \mathbb{O} \right\}.$$

We define an equivalence relation on $\mathbb{O}_0^3$ by $\mathbf{z} \sim \mathbf{w}$ if $\mathbf{w} = \mathbf{z}\lambda$ for some λ in an associative subalgebra of $\mathbb{O}$ containing the entries z_1, z_2, z_3 of $\mathbf{z}$. The map from $\mathbb{O}_0^3$ to the set of equivalence classes is the analogue of right projection and so we let $\mathbb{P}\mathbb{O}_0^3$ denote the set of right equivalence classes.

Let H be a real Hermitian matrix of signature $(2,1)$, for example one of those considered in (4.1):

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad H_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Given $\mathbf{z} \in \mathbb{O}_0^3$, define $Z = \mathbf{z}\mathbf{z}^* H$. This is a 3×3 matrix whose entries all lie in an associative subalgebra of $\mathbb{O}$. Right multiplication of $\mathbf{z}$ by λ lying in the same associative subalgebra as the entries of $\mathbf{z}$ leads to multiplication of Z by $|\lambda|^2$

$$(\mathbf{z}\lambda)(\mathbf{z}\lambda)^* H = \mathbf{z}\lambda\overline{\lambda}\mathbf{z}^* H = |\lambda|^2 \mathbf{z}\mathbf{z}^* H.$$

In particular, if $z = (z_1, z_2)$ where $z_1,\ z_2 \in \mathbb{O}$ and if $\mathbf{z}$ is the standard lift of z then $\mathbf{z}$ and $Z = \mathbf{z}\mathbf{z}^* H_2$ are given by

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ 1 \end{pmatrix}, \qquad Z = \begin{pmatrix} z_1 & z_1\overline{z}_2 & |z_1|^2 \\ z_2 & |z_2|^2 & z_2\overline{z}_1 \\ 1 & \overline{z}_2 & \overline{z}_1 \end{pmatrix}.$$

We consider $\mathrm{M}(3,\mathbb{O})$ the real vector space of 3×3 octonionic matrices. Let X^* denote the conjugate transpose of a matrix X in $\mathrm{M}(3,\mathbb{O})$. Given a real Hermitian matrix H, as above, define

$$J = \Big\{ X \in \mathrm{M}(3,\mathbb{O}) \ : \ HX = X^* H \Big\}.$$

Then J is closed under the Jordan multiplication

$$X * Y = \frac{1}{2}(XY + YX),$$

and so we call J the *Jordan algebra* associated to H. Real numbers act on $\mathrm{M}(3,\mathbb{O})$ by multiplication of each entry of X. We define an equivalence relation on J by $X \sim Y$ if and only if $Y = kX$ for some non-zero real number k. Then $\mathbb{P}J$ is defined to be the set of equivalence classes $[X]$.

If $\mathbf{z} \in \mathbb{O}_0^3$ then

$$HZ = H\mathbf{z}\mathbf{z}^* H = (\mathbf{z}\mathbf{z}^* H)^* H = Z^* H$$

and so $Z = \mathbf{z}\mathbf{z}^* H$ lies in the Jordan algebra J. Hence this map defines an embedding $\mathbb{O}_0^3 \longrightarrow J$. Moreover, the two projection maps are compatible and so there is a well defined map $\mathbb{P}\mathbb{O}_0^3 \longrightarrow \mathbb{P}J$.

Moreover, $\mathrm{tr}(Z) = \mathrm{tr}(\mathbf{z}\mathbf{z}^* H) = \mathbf{z}^* H \mathbf{z}$, which is real. This is the analogue of our Hermitian form. Therefore define

$$\begin{aligned} V_- &= \big\{ \mathbf{z} \in \mathbb{O}_0^3 \ : \ \mathrm{tr}(\mathbf{z}\mathbf{z}^* H) < 0 \big\}, \\ V_0 &= \big\{ \mathbf{z} \in \mathbb{O}_0^3 - \{0\} \ : \ \mathrm{tr}(\mathbf{z}\mathbf{z}^* H) = 0 \big\}, \\ V_+ &= \big\{ \mathbf{z} \in \mathbb{O}_0^3 \ : \ \mathrm{tr}(\mathbf{z}\mathbf{z}^* H) > 0 \big\}. \end{aligned}$$

Then we define $\mathbf{H}^2_{\mathbb{O}} = \mathbb{P}V_-$ and $\partial\mathbf{H}^2_{\mathbb{O}} = \mathbb{P}V_0$. To each $z \in \mathbf{H}^2_{\mathbb{O}}$ we can take the standard lift $\mathbf{z}$ of z and the corresponding element Z of the Jordan algebra.

We can define a bilinear form on $\mathrm{M}(3,\mathbb{O})$ by

$$\langle X, Y \rangle = \mathrm{Re}\big(\mathrm{tr}(X * Y)\big) = \frac{1}{2}\mathrm{Re}\big(\mathrm{tr}(XY + YX)\big).$$

Then if $Z = \mathbf{z}\mathbf{z}^* H$ and $W = \mathbf{w}\mathbf{w}^* H$ then

$$\langle Z, W \rangle = |\mathbf{w}^* H \mathbf{z}|^2.$$

Note that $\langle Z, Z \rangle = (\mathbf{z}^* H \mathbf{z})^2 = \mathrm{tr}(Z)^2$

Then the metric on $\mathbf{H}^2_{\mathbb{O}}$ is given by

$$ds^2 = \frac{-4\big(\mathrm{tr}(Z)\mathrm{tr}(dZ) - \langle Z, dZ\rangle\big)}{\mathrm{tr}(Z)^2} = \frac{-4\big((\mathbf{z}^* H\mathbf{z})(d\mathbf{z}^* H d\mathbf{z}) - |\mathbf{z}^* H d\mathbf{z}|^2\big)}{|\mathbf{z}^* H\mathbf{z}|^2}$$

where $\mathbf{z} \in V_-$ and $Z = \mathbf{z}\mathbf{z}^* H \in J$. The distance formula is then given by

$$\cosh^2\left(\frac{\rho(z,w)}{2}\right) = \frac{\langle Z, W\rangle}{\mathrm{tr}(Z)\mathrm{tr}(W)} = \frac{\big|\mathbf{w}^* H\mathbf{z}|^2}{(\mathbf{z}^* H\mathbf{z})(\mathbf{w}^* H\mathbf{w})}.$$

The isometries of $\mathbf{H}^2_{\mathbb{O}}$ form an exceptional group $F_{4(-20)}$. This group may be defined via a set of generators analogous to those we have seen before. First, the following real matrices act on $\mathbb{O}^3_0$ by left multiplication.

$$T = \begin{pmatrix} 1 & -1 & -1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} t & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/t \end{pmatrix}, \quad R = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

where $t \neq 0$ is real. Since each of them satisfies $A^{-1} = H_2^{-1} A^* H_2$ where H_2 is given by (4.1), they also act on J by conjugation:

$$(A\mathbf{z})(A\mathbf{z})^* H_2 = A\mathbf{z}\mathbf{z}^* A^* H_2 = A(\mathbf{z}\mathbf{z}^* H_2)A^{-1}.$$

Finally, let u be a unit imaginary octonion, that is u is an octonion with $|u| = 1$ and $\mathrm{Re}(u) = 0$. Let S_u be the map defined by

$$S_u : \begin{pmatrix} z_1 \\ z_1 \\ 1 \end{pmatrix} = \begin{pmatrix} uz_1\overline{u} \\ z_2\overline{u} \\ 1 \end{pmatrix}.$$

This sends the corresponding $Z = \mathbf{z}\mathbf{z}^* H_2 \in J$ by

$$S_u : \begin{pmatrix} z_1 & z_1\overline{z}_2 & |z_1|^2 \\ z_2 & |z_2|^2 & z_2\overline{z}_1 \\ 1 & \overline{z}_2 & \overline{z}_1 \end{pmatrix} \longmapsto \begin{pmatrix} uz_1\overline{u} & u(z_1\overline{z}_2) & |z_1|^2 \\ z_2\overline{u} & |z_2|^2 & (z_2\overline{z}_1)\overline{u} \\ 1 & u\overline{z}_2 & u\overline{z}_1\overline{u} \end{pmatrix}.$$

This uses the identity $(uz_1\overline{u})(u\overline{z}_2) = u(z_1\overline{z}_2)$ which is valid provided u is a unit imaginary octonion. Therefore S_u acts on J via conjugation by the matrix

$$\begin{pmatrix} u & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & u \end{pmatrix}.$$

Chapter 10

Further reading

The material in Chapters 2, 3, 4 and 5 is completely standard. It would be impossible to give a complete bibliography of text books in this area, not to mention research papers. Therefore we have given a few standard references for further reading. We have taken a mildly non-standard approach in order to justify the material in later chapters. The material in Chapters 6, 7 and 8 is less well known. Therefore we give a more detailed bibliography for this material. No doubt there are many omissions from this list. Finally, we give a brief list of references for the survey in Chapter 9 of how this material may be generalised.

Chapter 2. The material in this section is fairly standard linear algebra and can be found in many text books.

Chapter 3. This material may be found in most text books on hyperbolic geometry, for example Anderson [5] or Beardon [8]. It is not standard to base the treatment so closely on linear algebra, however.

Chapter 4. Ratcliffe [33] and Beardon [8] discuss the relationship between the Poincaré and Klein models of the hyperbolic plane. In [1] Ahlfors gives an account of high dimensional Möbius transformations by generalising the construction in this section.

Chapter 5. This material is again standard and may be found in Beardon [8] or Ratcliffe [33]. In particular, Beardon discusses the Poincaré extension and Ratcliffe discusses the Klein model.

Chapter 6. The foundations of quaternionic linear algebra were laid down by Brenner [9], Coxeter [14], Gormley [19], Lee [26]. The quadratic equation was solved by Niven [30] and this was used to find eigenvalues by Huang and So [20], [21]; see also Zhang [40]. Other more recent work is by Wilker [39].

The classification problem has been discussed by various people, such as Cao [12] Parker and Wang [11] Foreman, [16] Gongopadhyay and Kulkarni, [18] Kido [24] and Parker and Short [31]. See the discussion in [31] for more details.

Applications of quaternionic matrices to hyperbolic geometry have been given by Kellerhals [22], [23].

Chapter 7. Clifford Möbius transformations were first introduced over a century ago by Vahlen [35] and subsequently discussed by Maass [27]. However, it was Ahlfors [2], [3] who really popularised their use. Other authors who have studies Clifford Möbius transformations are Cao [10], Waterman [38] and Wada [36], [37] among others.

Chapter 8. The general theory of automorphisms of trees and their relationship to p-adic numbers is due to Serre [34]. The approach we use here is based on Figà-Talamanca [15], which was extended by Armitage and Parker [6].

Chapter 9. The generalisation of classical hyperbolic geometry to other rank one symmetric spaces of non-compact type may be found in Chen and Greenberg [13] and Mostow [29].

Complex hyperbolic geometry is currently an active topic of research; see [17]. More details of the approach we take here may be found in Parker [32]. There has been some work on quaternionic hyperbolic geometry; see [25]. The account we give here of octonionic hyperbolic geometry is based on Allcock, [4]. See also Markham and Parker [28] for further work in this area. See Baez [7] for background material.

Bibliography

[1] L.V. Ahlfors. *Möbius transformations in several dimensions* Ordway lecture notes, University of Minnesota, 1981.

[2] L.V. Ahlfors. Möbius transformations and Clifford numbers. *Differential Geometry and Complex Analysis*, Springer 1985, 65–73.

[3] L.V. Ahlfors. On the fixed points of Möbius transformation in $\mathbf{R}^{\mathbf{n}}$. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **10** (1985), 15–27.

[4] D.J. Allcock. Reflection groups on the octave hyperbolic plane, *Journal of Algebra* **213** (1998), 467–498.

[5] J.W. Anderson *Hyperbolic Geometry* Springer 1999.

[6] J.V. Armitage & J.R. Parker Jørgensen's inequality for non-Archimedean metric spaces, *Geometry and Dynamics of Groups and Spaces* Progress in Mathematics **265**, Birkhäuser, to appear, 1–15.

[7] J.C. Baez. The octonions, *Bulletin of the American Math. Soc.* **39** (2002), 145–205.

[8] A. F. Beardon. *The Geometry of Discrete Groups*, Graduate Texts in Mathematics, **91**. Springer Verlag (1983).

[9] J. L. Brenner. Matrices of quaternions. *Pacific J. Maths.* **1** (1951) 329–335.

[10] C. Cao & P.L. Waterman. Conjugacy invariants of Möbius groups. *Quasiconformal Mappings and Analysis*, Springer 1998, 109–139.

[11] W. Cao, J. R. Parker & X. Wang. On the classification of quaternionic Möbius transformations. *Math. Proc. Camb. Phil. Soc.* **137** (2004), 349–361.

[12] W. Cao. On the classification of four-dimensional Möbius transformations. *Proc. Edinb. Math. Soc.* **50** (2007), 49–62.

[13] S.S. Chen & L. Greenberg. Hyperbolic spaces. *Contributions to Analysis*, Academic Press, New York, 1974, 49–87.

[14] H. S. M. Coxeter. Quaternions and reflections. *Amer. Math. Monthly* **53** (1946), 136–146.

[15] A. Figà-Talamanca Local fields and trees. *Harmonic Functions on Trees and Buildings*, ed A. Korányi, A.M.S. Contemporary Mathematics **206** (1997), 3–16.

[16] B. Foreman. Conjugacy invariants of $Sl(2,\mathbb{H})$. *Linear Algebra Appl.* **381** (2004), 24–35.

[17] W.M. Goldman. *Complex Hyperbolic Geometry*, Oxford University Press, 1999.

[18] K. Gongopadhyay & R. S. Kulkarni. Dynamical types of isometries of hyperbolic space. Preprint (2006).

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

[20] L. Huang & W. So. On left eigenvalues of a quaternionic matrix. *Linear Algebra Appl.* **323** (2001), 105–116.

[21] L. Huang & W. So. Quadratic formulas for quaternions. *Applied Mathematics Letters* **15** (2002), 533–540.

[22] R. Kellerhals. Collars in $\mathrm{PSL}(2,\mathbb{H})$. *Annales Acad. Sci. Fenn.* **26** (2001), 51–72.

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

[24] T. Kido. Möbius transformations on quaternions. *Osaka University preprint* (2005).

[25] I. Kim & J.R. Parker. Geometry of quaternionic hyperbolic manifolds. *Math. Proc. Cambridge Phil. Soc.* **135** (2003), 291–320.

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

[27] H. Maass Automorphe Funktionen von mehreren Veränderlichen und Dirichletsche Reihen, *Abh. Math. Sem. Univ. Hamburg* **16** (1949), 53–104.

[28] S. Markham & J.R. Parker, Jørgensen's inequality for metric spaces with applications to the octonions, *Advances in Geometry* **7** (2007), 19–38.

[29] G.D. Mostow *Strong rigidity of locally symmetric spaces* Annals Of Maths. Studies **78**, Princeton University Press, 1973.

[30] I. Niven. Equations in quaternions. *American Math. Monthly* **48** (1941), 654–661.

[31] J.R. Parker & I. Short. Conjugacy classification of quaternionic Möbius transformations. Preprint, 2007.

[32] J.R. Parker *Complex Hyperbolic Kleinian Groups* Cambridge University Press, to appear.

[33] J. G. Ratcliffe *Foundations of hyperbolic manifolds*. Graduate Texts in Mathematics **149**. Springer-Verlag (1994).

[34] J.-P. Serre, *Trees*, Springer 1980.

[35] K.Th. Vahlen. Über Bewegungen und complexe Zahlen *Math Annalen* **55** (1902), 585–593.

[36] M. Wada *Conjugacy Invariants and Normal Forms of Isometries of Hyperbolic Space*, Ph.D. thesis, Columbia University 1986.

[37] M. Wada Conjugacy invariants of Möbius transformations *Complex Variables Thy. Appl.* **15** (1990) 125–133.

[38] P.L. Waterman. Möbius groups in several dimensions. *Advances in Maths.* **101** (1993),87–113.

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

[40] F. Zhang. Quaternions and matrices of quaternions. *Linear Algebra Appl.* **251** (1997), 21–57.