Notes on Complex Hyperbolic Geometry

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

June 16, 2010

Contents

These notes are an updated version of a similar set of notes that were distributed by me at the workshop *Géométrie hyperbolique complexe* held in Luminy in 2003.

1 Introduction

The unit ball in $\mathbb{C}^2$ has a natural metric of constant negative holomorphic sectional curvature (which we normalise to be -1), called the Bergman metric. As such it forms a model for *complex hyperbolic 2-space* $\mathbf{H}^2_{\mathbb{C}}$ analogous to the ball model of (real) hyperbolic space $\mathbf{H}^n_{\mathbb{R}}$. The main difference is that the (real) sectional curvature is no longer constant, but is pinched between -1 and $-1/4$. Another standard model for complex hyperbolic space is a paraboloid in $\mathbb{C}^2$ called the Siegel domain. This is analogous to the the half space model of $\mathbf{H}^n_{\mathbb{R}}$. As complex hyperbolic 1-space is just the unit disc in $\mathbb{C}$ with the Poincaré metric (or the upper half plane), $\mathbf{H}^2_{\mathbb{C}}$ is a natural generalisation of plane hyperbolic geometry which is different from the more familiar generalisation of higher dimensional real hyperbolic space.

An alternative description of $\mathbf{H}^2_{\mathbb{C}}$ is given by the projective model. Here we take a Hermitian form of signature $(2,1)$ on $\mathbb{C}^3$, that is complex Minkowski space. Projectivising the set of complex lines on which this form is negative gives another model for complex hyperbolic space. This is the natural complex generalisation of the projective model of real hyperbolic space. By taking a suitable Hermitian form and making a choice of section we can recover the ball model and the Siegel domain model. The Bergman metric is given by a simple distance formula in terms of the Hermitian form which is closely related to the Cauchy-Schwarz inequality. From this description we can show that all holomorphic isometries of complex hyperbolic space are given by the projectivisation of unitary matrices preserving the Hermitian form. All antiholomorphic isometries are given applying such a matrix followed by complex conjugation. This means that we can use complex linear algebra to study the geometry of complex hyperbolic space.

As well as studying isometries, we want to consider certain special classes of submanifolds of complex hyperbolic space. We will see that the totally geodesic submanifolds have dimension at most 2. (In fact, for n dimensional complex hyperbolic space, totally geodesic subspaces are are either embedded copies of $\mathbf{H}^m_{\mathbb{C}}$ or $\mathbf{H}^m_{\mathbb{R}}$ for $1 \leq m \leq n$. Thus, the real dimension of a totally geodesic submanifold is either at most n, for embedded copies of $\mathbf{H}^m_{\mathbb{R}}$, or else is even, for embedded copies of $\mathbf{H}^m_{\mathbb{C}}$.) In particular, there are no totally geodesic real hypersurfaces in $\mathbf{H}^2_{\mathbb{C}}$. This increases the difficulty of constructing polyhedra (for example fundamental polyhedra for discrete groups of complex hyperbolic isometries). In a later chapter we will describe some classes of real hypersurfaces that can be used to build polyhedra.

The boundary of complex hyperbolic 2-space is the one point compactification of the Heisenberg group in the same way that the boundary of real hyperbolic space is the one point compactification of Euclidean space of one dimension lower. Just as the internal geometry of real hyperbolic space may be studied using conformal geometry on the boundary, so the internal geometry of complex hyperbolic space may be studied using CR-geometry on the Heisenberg group. Moreover, the Heisenberg group is 3 dimensional and so it is easy to illustrate geometrical objects.

In order to make things as concrete as possible, we have chosen restrict our attention to $\mathbf{H}^2_{\mathbb{C}}$. Many of the results we develop will hold for complex hyperbolic space in all dimensions. There will often be analogues for other rank 1 symmetric spaces of non-compact type, quaternionic hyperbolic space $\mathbf{H}^n_{\mathbb{H}}$ and the octonionic hyperbolic plane $\mathbf{H}^2_{\mathbb{O}}$. We will not discuss these here.

2 Complex hyperbolic 2-space

2.1 Hermitian forms on $\mathbb{C}^{2,1}$

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 $\big((A^*)^*\big) = A$. A $k \times k$ complex matrix A is said to be *Hermitian* if it equals its own Hermitian transpose $A = A^*$. Let A be a Hermitian matrix and μ an eigenvalue of A with eigenvector $\mathbf{x}$. We claim that μ is real. In order to see this, observe that

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

Observe that $\mathbf{x}^*\mathbf{x}$ is real and non-zero and so 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

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

From these we see that

$$\begin{aligned}
\langle\mathbf{z},\mathbf{z}\rangle &\in \mathbb{R},\\
\langle\mathbf{z},\lambda\mathbf{w}\rangle &= \overline{\lambda}\,\langle\mathbf{z},\mathbf{w}\rangle\,,\\
\langle\lambda\mathbf{z},\lambda\mathbf{w}\rangle &= |\lambda|^2\,\langle\mathbf{z},\mathbf{w}\rangle\,.
\end{aligned}$$

Let $\mathbb{C}^{2,1}$ be the complex vector space of (complex) dimension 3 equipped with a non-degenerate, indefinite Hermitian form $\langle\cdot,\cdot\rangle$ of signature $(2,1)$. This means that $\langle\cdot,\cdot\rangle$ is given by a non-singular 3×3 Hermitian matrix H with 2 positive eigenvalues and 1 negative eigenvalue. There are two standard matrices H which give different Hermitian forms on $\mathbb{C}^{2,1}$. Following Epstein [7] we call these the first and second Hermitian forms. Let $\mathbf{z}$, $\mathbf{w}$ be the column vectors $(z_1,\, z_2,\, z_3)^t$ and $(w_1,\, w_2,\, w_3)^t$ respectively. The *first Hermitian form* is defined to be:

$$\langle \mathbf{z}, \mathbf{w} \rangle_1 = z_1\overline{w}_1 + z_2\overline{w}_2 - z_3\overline{w}_3. \tag{1}$$

It is given by the Hermitian matrix H_1:

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{2}$$

The *second Hermitian form* is defined to be:

$$\langle \mathbf{z}, \mathbf{w} \rangle_2 = z_1\overline{w}_3 + z_2\overline{w}_2 + z_3\overline{w}_1. \tag{3}$$

It is given by the Hermitian matrix H_2:

$$H_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{4}$$

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 $\mathbb{C}^3$ of signature $(2,1)$).

There are other Hermitian forms which are widely used in the literature. In particular, Chen and Greenberg (page 67 of [3])give a close relative of the second Hermitian form. We will refer to this as the *third Hermitian form*. It is given by

$$\langle \mathbf{z}, \mathbf{w} \rangle_3 = -z_1\overline{w}_2 - z_2\overline{w}_1 + z_3\overline{w}_3.$$

It is given by the Hermitian matrix H_3:

$$H_3 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The third Hermitian form has been used extensively by Kamiya, Hersonsky and Paulin.

2.2 Three models of complex hyperbolic space

If $\mathbf{z} \in \mathbb{C}^{2,1}$ then we know that $\langle \mathbf{z}, \mathbf{z} \rangle$ is real. Thus we may define subsets V_-, V_0 and V_+ of $\mathbb{C}^{2,1}$ by

$$\begin{aligned} V_- &= \left\{ \mathbf{z} \in \mathbb{C}^{2,1} | \, \langle \mathbf{z}, \mathbf{z} \rangle < 0 \right\}, \\ V_0 &= \left\{ \mathbf{z} \in \mathbb{C}^{2,1} - \{0\} | \, \langle \mathbf{z}, \mathbf{z} \rangle = 0 \right\}, \\ V_+ &= \left\{ \mathbf{z} \in \mathbb{C}^{2,1} | \, \langle \mathbf{z}, \mathbf{z} \rangle > 0 \right\}. \end{aligned}$$

We say that $\mathbf{z} \in \mathbb{C}^{2,1}$ 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.

Define an equivalence relation on $\mathbb{C}^{2,1} - \{0\}$ by $\mathbf{z} \sim \mathbf{w}$ if and only if there is a non-zero complex scalar λ so that $\mathbf{w} = \lambda\mathbf{z}$. Let $\mathbb{P} : \mathbb{C}^{2,1} - \{0\} \longmapsto \mathbb{CP}^2$ denote the standard projection map defined by $\mathbb{P}(\mathbf{z}) = [\mathbf{z}]$ where $[\mathbf{z}]$ is the equivalence class of $\mathbf{z}$. 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. On the chart of $\mathbb{C}^{2,1}$ with $z_3 \neq 0$ the projection map $\mathbb{P}$ is given by

$$\mathbb{P} : \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \longmapsto \begin{pmatrix} z_1/z_3 \\ z_2/z_3 \end{pmatrix} \in \mathbb{C}^2.$$

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

We define the other two standard models of complex hyperbolic space by taking the section defined by $z_3 = 1$ for the first and second Hermitian forms. In other words, if we take column vectors

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ 1 \end{bmatrix}$$

in $\mathbb{C}^{2,1}$ then consider what it means for $\langle \mathbf{z}, \mathbf{z}\rangle$ to be negative.

For the first Hermitian form we obtain $\mathbf{z} \in \mathbf{H}^2_{\mathbb{C}}$ provided:

$$\langle \mathbf{z}, \mathbf{z}\rangle_1 = z_1\overline{z}_1 + z_2\overline{z}_2 - 1 < 0.$$

In other words

$$|z_1|^2 + |z_2|^2 < 1.$$

Thus $z = (z_1, z_2)$ is in the unit ball in $\mathbb{C}^2$. This forms the *unit ball model* of complex hyperbolic space. The boundary of the unit ball model is the sphere S^3 given by

$$|z_1|^2 + |z_2|^2 = 1.$$

For the second Hermitian form we obtain $\mathbf{z} \in \mathbf{H}^2_{\mathbb{C}}$ provided:

$$\langle \mathbf{z}, \mathbf{z}\rangle_2 = z_1 + z_2\overline{z}_2 + \overline{z}_1 < 0.$$

In other words

$$2\Re(z_1) + |z_2|^2 < 0.$$

Thus $z = (z_1, z_2)$ is in a domain in $\mathbb{C}^2$ whose boundary is the paraboloid defined by

$$2\Re(z_1) + |z_2|^2 = 0.$$

This domain is called the Siegel domain and forms the *Siegel domain model* of $\mathbf{H}^2_{\mathbb{C}}$.

Given a point $z = (z_1, z_2) \in \mathbb{C}^2$ we define the *standard lift* of z to be the point $\mathbf{z} \in \mathbb{C}^{2,1}$ given by

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ 1 \end{bmatrix}.$$

It is clear that $\mathbb{P}(\mathbf{z}) = z$. Therefore the standard lift enables us to give a well defined inverse of $\mathbb{P}$ whose domain is $\mathbb{C}^2$. We extend this definition to include the point ∞. We define the *standard lift* of ∞ to be

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{C}^{2,1}.$$

We will freely pass between points of $\mathbb{C}^2 \cup \{\infty\}$ and their standard lifts. Most of the time we aim to make our constructions independent of which element of $\mathbb{P}^{-1}(z)$ we choose but, for definiteness, one may always choose the standard lift. There will be some occasions when we need to be careful about lifts, and we will mention this explicitly.

For the projective model the metric on $\mathbf{H}^2_{\mathbb{C}}$, called the *Bergman metric* is given by

$$ds^2 = \frac{-4}{\langle \mathbf{z}, \mathbf{z} \rangle^2} \det \begin{pmatrix} \langle \mathbf{z}, \mathbf{z} \rangle & \langle d\mathbf{z}, \mathbf{z} \rangle \\ \langle \mathbf{z}, d\mathbf{z} \rangle & \langle d\mathbf{z}, d\mathbf{z} \rangle \end{pmatrix}. \tag{5}$$

Alternatively, the Bergman metric is given by the distance function $\rho(\cdot, \cdot)$ defined by the formula

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

For the ball model and Siegel domain model one can find the distance between points z and w by plugging their standard lifts $\mathbf{z}$ and $\mathbf{w}$ into the above formula. However, as may easily be seen, this formula is independent of which lifts $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{C}^{2,1}$ of z and w we choose.

Proposition 2.1 *In the ball model of* $\mathbf{H}^2_{\mathbb{C}}$ *the volume form is given by*

$$d\mathrm{Vol} = \frac{16}{(1 - |z_1|^2 - |z_2|^2)^3} d\,\mathrm{vol}$$

where $d\,\mathrm{vol}$ *is the volume element*

$$(1/2i)^2 d\,z_1 \wedge d\,\overline{z}_1 \wedge d\,z_2 \wedge d\,\overline{z}_2 = d\,x_1 d\,y_1 d\,x_2 d\,y_2.$$

Proof: Substituting for the first Hermitian form in (5) we have

$$\begin{aligned} ds^2 &= \frac{-4}{(|z_1|^2 + |z_2|^2 - 1)^2} \det \begin{pmatrix} |z_1|^2 + |z_2|^2 - 1 & \overline{z}_1 d\,z_1 + \overline{z}_2 d\,z_2 \\ z_1 d\,\overline{z}_1 + z_2 d\,\overline{z}_2 & |d\,z_1|^2 + |d\,z_2|^2 \end{pmatrix} \\ &= \frac{4(1 - |z_1|^2 - |z_2|^2)(|d\,z_1|^2 + |d\,z_2|^2) + 4\left|\overline{z}_1 d\,z_1 + \overline{z}_2 d\,z_2\right|^2}{(|z_1|^2 + |z_2|^2 - 1)^2}. \end{aligned}$$

Converting to real coordinates, $x_1 + iy_1 = z_1$, $x_2 + iy_2 = z_2$, and denoting

$$r^2 = {x_1}^2 + {y_1}^2 + {x_2}^2 + {y_2}^2 = |z_1|^2 + |z_2|^2 < 1,$$

we have

$$ds^2 = \begin{pmatrix} d\,x_1 & d\,y_1 & d\,x_2 & d\,y_2 \end{pmatrix} \mathbf{g}_{\mathbb{R}} \begin{pmatrix} d\,x_1 \\ d\,y_1 \\ d\,x_2 \\ d\,y_2 \end{pmatrix}$$

where

$$\mathbf{g}_{\mathbb{R}} = \frac{4}{(1-r^2)^2} \begin{pmatrix} 1-x_2^2-y_2^2 & 0 & x_1x_2+y_1y_2 & x_1y_2-y_1x_2 \\ 0 & 1-x_2^2-y_2^2 & -x_1y_2+y_1x_2 & x_1x_2+y_1y_2 \\ x_1x_2+y_1y_2 & -x_1y_2+y_1x_2 & 1-x_1^2-y_1^2 & 0 \\ x_1y_2+y_1x_2 & x_1x_2+y_1y_2 & 0 & 1-x_1^2-y_1^2 \end{pmatrix}.$$

Now

$$\det(\mathbf{g}_{\mathbb{R}}) = \frac{256}{(1-r^2)^6}.$$

Thus the volume form is

$$d\mathrm{Vol} = \sqrt{\det(\mathbf{g}_{\mathbb{R}})}\,d\,x_1 d\,y_1 d\,x_2 d\,y_2 = \frac{16}{(1-r^2)^3} d\,x_1 d\,y_1 d\,x_2 d\,y_2$$

as claimed. □

By switching to spherical polar coordinates, we can use this to compute the volume of a hyperbolic ball of radius δ.

Proposition 2.2 *The volume of a ball of (Bergman) radius δ in $\mathbf{H}^2_{\mathbb{C}}$ is*

$$\mathrm{Vol}(B_\delta) = 2\pi^2(\cosh\delta - 1)^2.$$

Proof: We switch to spherical polar coordinates on the ball by writing $z_1 = r\cos\theta e^{i\phi}$, $z_2 = r\sin\theta e^{i\psi}$ where $r \geq 0$, $\theta \in [0, \pi/2]$ and ϕ, $\psi \in [0, 2\pi)$. With these coordinates the volume form is

$$d\,\mathrm{vol} = r^3 \cos\theta \sin\theta\, dr\, d\theta\, d\phi,\ d\psi.$$

Also a point with Euclidean distance r from the origin has Bergman distance ρ from the origin, where $\tanh(\rho/2) = r$. Therefore the volume of the ball of radius δ is

$$\begin{aligned} \mathrm{Vol}(B_\delta) &= \int_{r \leq \tanh(\delta/2)} \frac{16}{(1-r^2)^3}\, d\,\mathrm{vol} \\ &= \int_{r=0}^{\tanh(\delta/2)} \int_{\theta=0}^{\pi/2} \int_{\phi=0}^{2\pi} \int_{\psi=0}^{2\pi} \frac{16}{(1-r^2)^3} r^3 \cos\theta \sin\theta\, dr\, d\theta\, d\phi,\ d\psi \\ &= 4\pi^2 \int_0^{\tanh(\delta/2)} \frac{8r^3}{(1-r^2)^3}\, dr. \end{aligned}$$

We use the substitution $r = \tanh(\rho/2)$ to obtain:

$$\begin{aligned} \mathrm{Vol}(B_\delta) &= 4\pi^2 \int_0^\delta 4\sinh^3(\rho/2)\cosh(\rho/2)\, d\rho \\ &= 8\pi^2 \sinh^4(\delta/2) \\ &= 2\pi^2(\cosh\delta - 1)^2. \end{aligned}$$

□

Proposition 2.3 *In the Siegel domain model of $\mathbf{H}^2_{\mathbb{C}}$ the volume form is given by*

$$d\mathrm{Vol} = \frac{16}{(-z_1 - \overline{z}_1 - |z_2|^2)^3} d\,\mathrm{vol}$$

where $d\,\mathrm{vol}$ is the volume element

$$(1/2i)^2 d\,z_1 \wedge d\,\overline{z}_1 \wedge d\,z_2 \wedge d\,\overline{z}_2 = d\,x_1 d\,y_1 d\,x_2 d\,y_2.$$

Proof: Substituting for the second Hermitian form in (5) we have

$$\begin{aligned} ds^2 &= \frac{-4}{(z_1 + |z_2|^2 + \overline{z}_1)^2} \det \begin{pmatrix} z_1 + |z_2|^2 + \overline{z}_1 & d\,z_1 + \overline{z}_2 d\,z_2 \\ d\,\overline{z}_1 + z_2 d\,\overline{z}_2 & |d\,z_2|^2 \end{pmatrix} \\ &= \frac{-4(z_1 + |z_2|^2 + \overline{z}_1)|d\,z_2|^2 + 4\left|d\,z_1 + \overline{z}_2 d\,z_2\right|^2}{(z_1 + |z_2|^2 + \overline{z}_1)^2}. \end{aligned}$$

Converting to real coordinates, $x_1 + iy_1 = z_1$, $x_2 + iy_2 = z_2$, we have

$$ds^2 = \begin{pmatrix} d\,x_1 & d\,y_1 & d\,x_2 & d\,y_2 \end{pmatrix} \mathbf{g}_{\mathbb{R}} \begin{pmatrix} d\,x_1 \\ d\,y_1 \\ d\,x_2 \\ d\,y_2 \end{pmatrix}$$

where

$$\mathbf{g}_{\mathbb{R}} = \frac{4}{(2x_1 + {x_2}^2 + {y_2}^2)^2} \begin{pmatrix} 1 & 0 & x_2 & y_2 \\ 0 & 1 & -y_2 & x_2 \\ x_2 & -y_2 & -2x_1 & 0 \\ y_2 & x_2 & 0 & -2x_1 \end{pmatrix}.$$

Now

$$\det(\mathbf{g}_{\mathbb{R}}) = \frac{256}{(2x_1 + {x_2}^2 + {y_2}^2)^6}.$$

Thus the volume form is

$$d\mathrm{Vol} = \sqrt{\det(\mathbf{g}_{\mathbb{R}})}\,d\,x_1 d\,y_1 d\,x_2 d\,y_2 = \frac{16}{(-2x_1 - {x_2}^2 - {y_2}^2)^3} d\,x_1 d\,y_1 d\,x_2 d\,y_2$$

as claimed (since $2x_1 + {x_2}^2 + {y_2}^2 < 0$ we take the negative sign in the square root). □

2.3 Cayley transforms

Given two Hermitian forms H and H' of signature $(2,1)$ we can can pass between them using a *Cayley transform* C. That is, we can write

$$H' = C^* H C.$$

The Cayley transform C is not unique for we may precompose and postcompose by any unitary matrix preserving the relevant Hermitian form. The following Cayley transform interchanges the first and second Hermitian forms

$$C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix}. \tag{6}$$

Observe that $C^* = C$. In order to see that C is a Cayley transform, we calculate

$$C^* H_1 C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = H_2.$$

Also, $C^{-1} = C$ and so $C^* H_2 C = H_1$.

When we are dealing with groups of matrices whose entries lie in a ring $\mathcal{O}$ (for example the Picard modular groups) it will be necessary to choose a Cayley transform C so that the entries of C and C^{-1} are all integers. This will show that group of matrices in $\mathrm{GL}(3, \mathcal{O})$ preserving the Hermitian forms $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are conjugate in $\mathrm{GL}(3, \mathcal{O})$. In this case we may choose the Cayley transform sending the ball model to the Siegel domain to be:

$$C_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix} \quad \text{and} \quad C_1^{-1} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 0 & -1 \end{pmatrix}. \tag{7}$$

It is easy to check that $C_1^* H_1 C_1 = H_2$ and so $(C_1^{-1})^* H_2 C_1^{-1} = H_1$.

3 Isometries

3.1 The unitary groups of the first Hermitian form

Let A be a matrix which preserves the first Hermitian form, that is a unitary matrix. In other words for all $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{C}^{2,1}$ we have

$$\mathbf{w}^* A^* H_1 A \mathbf{v} = \langle A\mathbf{v}, A\mathbf{w} \rangle_1 = \langle \mathbf{v}, \mathbf{w} \rangle_1 = \mathbf{w}^* H_1 \mathbf{v}.$$

By letting $\mathbf{v}$ and $\mathbf{w}$ run through a basis of of $\mathbb{C}^{2,1}$ we see that this means $A^* H_1 A = H_1$. In other words, $H_1{}^{-1} A^* H_1 A = I$ and so $A^{-1} = H_1{}^{-1} A^* H_1$. Writing A in terms of its entries gives

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{bmatrix}, \qquad A^{-1} = H_1{}^{-1} A^* H_1 = \begin{bmatrix} \overline{a} & \overline{d} & -\overline{g} \\ \overline{b} & \overline{e} & -\overline{h} \\ -\overline{c} & -\overline{f} & \overline{j} \end{bmatrix}. \tag{8}$$

We now use this expression to find relationships between the entries of A. The resulting identities will be used many times in later sections.

From elementary linear algebra, we know that $A^{-1} = \mathrm{adj}(A)/\det(A)$ where $\mathrm{adj}(A)$ is the adjugate matrix:

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} ej - fh & ch - bj & bf - ce \\ fg - dj & aj - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{bmatrix}.$$

Writing $\Delta = \det(A)$ and comparing these two expressions for A^{-1} gives

$$
\begin{aligned}
\overline{a}\Delta &= ej - fh && (9)\\
\overline{b}\Delta &= fg - dj && (10)\\
\overline{c}\Delta &= eg - dh && (11)\\
\overline{d}\Delta &= ch - bj && (12)\\
\overline{e}\Delta &= aj - cg && (13)\\
\overline{f}\Delta &= ah - bg && (14)\\
\overline{g}\Delta &= ce - bf && (15)\\
\overline{h}\Delta &= af - cd && (16)\\
\overline{j}\Delta &= ae - bd && (17)
\end{aligned}
$$

As A is unitary we have

$$|\Delta|^2 = \det(A^*)\det(A) = \det({H_1}^{-1}A^*H_1A) = \det(A^{-1}A) = 1.$$

From the equations $AA^{-1} = I$ and (8) we have the following identities relating the entries of A:

$$
\begin{aligned}
1 &= |a|^2 + |b|^2 - |c|^2, && (18)\\
1 &= |d|^2 + |e|^2 - |f|^2, && (19)\\
1 &= -|g|^2 - |h|^2 + |j|^2, && (20)\\
0 &= a\overline{d} + b\overline{e} - c\overline{f}, && (21)\\
0 &= a\overline{g} + b\overline{j} - c\overline{j}, && (22)\\
0 &= d\overline{g} + e\overline{h} - f\overline{j}. && (23)
\end{aligned}
$$

Similarly from the relation $A^{-1}A = I$ we have

$$
\begin{aligned}
1 &= |a|^2 + |d|^2 - |g|^2, && (24)\\
1 &= |b|^2 + |e|^2 - |h|^2, && (25)\\
1 &= -|c|^2 - |f|^2 + |j|^2, && (26)\\
0 &= \overline{a}b + \overline{d}e - \overline{g}h, && (27)\\
0 &= \overline{a}c + \overline{d}f - \overline{g}j, && (28)\\
0 &= \overline{b}c + \overline{e}f - \overline{h}j. && (29)
\end{aligned}
$$

3.2 The unitary groups of the second Hermitian form

Let A be a matrix which preserves the second Hermitian form, that is a unitary matrix. In other words for all $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{C}^{2,1}$ we have

$$\mathbf{w}^*A^*H_2A\mathbf{v} = \langle A\mathbf{v},\, A\mathbf{w}\rangle_2 = \langle \mathbf{v},\, \mathbf{w}\rangle_2 = \mathbf{w}^*H_2\mathbf{v}.$$

As before, letting $\mathbf{v}$ and $\mathbf{w}$ run through a basis of of $\mathbb{C}^{2,1}$, we see that this means $A^*H_2A = H_2$. In other words, $A^{-1} = {H_2}^{-1}A^*H_2$.

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & j \end{bmatrix}, \qquad A^{-1} = {H_2}^{-1}A^*H_2 = \begin{bmatrix} \overline{j} & \overline{f} & \overline{c} \\ \overline{h} & \overline{e} & \overline{b} \\ \overline{g} & \overline{d} & \overline{a} \end{bmatrix}. \tag{30}$$

Again, we use this expression to find useful identities between the entries of A.

Using the expression of A^{-1} in terms of the adjugate matrix we obtain

$$\begin{aligned}
\bar{a}\Delta &= ae - bd && (31)\\
\bar{b}\Delta &= cd - af && (32)\\
\bar{c}\Delta &= bf - ce && (33)\\
\bar{d}\Delta &= bg - ah && (34)\\
\bar{e}\Delta &= aj - cg && (35)\\
\bar{f}\Delta &= ch - bj && (36)\\
\bar{g}\Delta &= dh - eg && (37)\\
\bar{h}\Delta &= fg - dj && (38)\\
\bar{j}\Delta &= ej - fh && (39)
\end{aligned}$$

From the equations $AA^{-1} = I$ and (30) we have the following identities relating the entries of A:

$$\begin{aligned}
1 &= a\bar{j} + b\bar{h} + c\bar{g}, && (40)\\
1 &= d\bar{f} + |e|^2 + f\bar{d}, && (41)\\
0 &= a\bar{f} + b\bar{e} + c\bar{d}, && (42)\\
0 &= a\bar{c} + |b|^2 + c\bar{a}, && (43)\\
0 &= d\bar{j} + e\bar{h} + f\bar{g}, && (44)\\
0 &= g\bar{j} + |h|^2 + j\bar{g}. && (45)
\end{aligned}$$

Similarly from the relation $A^{-1}A = I$ we have

$$\begin{aligned}
1 &= \bar{j}a + \bar{f}d + \bar{c}g, && (46)\\
1 &= \bar{h}b + |e|^2 + \bar{b}f, && (47)\\
0 &= \bar{j}b + \bar{f}e + \bar{c}h, && (48)\\
0 &= \bar{j}c + |f|^2 + \bar{c}j, && (49)\\
0 &= \bar{h}a + \bar{e}d + \bar{b}g, && (50)\\
0 &= \bar{g}a + |d|^2 + \bar{a}g. && (51)
\end{aligned}$$

3.3 $\mathrm{PU}(2,1)$ and its action on complex hyperbolic space

We now show unitary matrices act on complex hyperbolic space. Any matrix in $\mathrm{U}(2,1)$ which is a (non-zero) complex scalar multiple of the identity maps each line in $\mathbb{C}^{2,1}$ to itself and so acts trivially on complex hyperbolic space. Since this matrix is unitary with respect to $\langle\cdot,\cdot\rangle$ then the scalar must have unit norm. Because of this, we define the projective unitary group $\mathrm{PU}(2,1) = \mathrm{U}(2,1)/\mathrm{U}(1)$ where $\mathrm{U}(1)$ is canonically identified with $\{e^{i\theta}I | 0 \le \theta < 2\pi\}$, where I is the identity matrix in $\mathrm{U}(2,1)$. Sometimes it will be useful to consider $\mathrm{SU}(2,1)$, the group of matrices with determinant 1 which are unitary with respect to $\langle\cdot,\cdot\rangle$. The group $\mathrm{SU}(2,1)$ is a 3-fold covering of $\mathrm{PU}(2,1)$:

$$\mathrm{PU}(2,1) = \mathrm{SU}(2,1)/\{I, \omega I, \omega^2 I\}$$

where $\omega = (-1 + i\sqrt{3})/2$ is a cube root of unity. This is completely analogous to the fact that $\mathrm{SL}(2,\mathbb{C})$ is a double cover of $\mathrm{PSL}(2,\mathbb{C})$. Cube roots of unity are used because $\mathrm{SU}(2,1)$ comprises 3×3 matrices.

We now use Hermitian linear algebra to show that $\mathrm{PU}(2,1)$ acts transitively on $\mathbf{H}^2_{\mathbb{C}}$ and doubly transitively on $\partial\mathbf{H}^2_{\mathbb{C}}$.

We will begin by working with the first Hermitian form. Let $\mathbf{z} \in \mathbb{C}^{2,1}$ be any negative vector. That is $\langle \mathbf{z}, \mathbf{z} \rangle_1 < 0$. Then $\hat{\mathbf{z}} = \mathbf{z}/\sqrt{-\langle \mathbf{z}, \mathbf{z} \rangle_1}$ is a negative vector with $\langle \hat{\mathbf{z}}, \hat{\mathbf{z}} \rangle_1 = -1$. We can now construct a matrix A in $\mathrm{U}(2,1)$ whose third column is $\hat{\mathbf{z}}$. In order to do this, we take any basis for $\mathbb{C}^{2,1}$ containing $\hat{\mathbf{z}}$. We can then use a version of the Gram-Schmidt process in signature $(2,1)$ to produce vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ so that $\langle \mathbf{e}_1, \mathbf{e}_1 \rangle_1 = \langle \mathbf{e}_2, \mathbf{e}_2 \rangle_1 = 1$ and $\langle \mathbf{e}_1, \mathbf{e}_2 \rangle_1 = \langle \mathbf{e}_j, \hat{\mathbf{z}} \rangle_1 = \langle \mathbf{e}_j, \hat{\mathbf{z}} \rangle_1 = 0$. The vectors $\mathbf{e}_j$ now form the first two columns of A. By construction $A^* H_1 A = H_1$ and so $A \in \mathrm{U}(2,1)$. Moreover, the image of the column vector $(0,\ 0,\ 1)^t$, that is the canonical lift of the origin o, under A is just $\hat{\mathbf{z}}$. This process leads to the following result which shows that $\mathrm{PU}(2,1)$ acts transitively on $\mathbf{H}^2_{\mathbb{C}}$.

Proposition 3.1 *For any point z in $\mathbf{H}^2_{\mathbb{C}}$ (using the ball model) there is an element of* $\mathrm{PU}(2,1)$ *sending the origin o to z.*

Proof: We work with the unit ball model. Let $\mathbf{z}$ be the canonical lift of z to $\mathbb{C}^{2,1}$. As above we can scale $\mathbf{z}$ to form $\hat{\mathbf{z}} = \mathbf{z}/\sqrt{-\langle \mathbf{z}, \mathbf{z} \rangle_1}$ and find a matrix A in $\mathrm{U}(2,1)$ sending the canonical lift of the origin, o, to $\hat{\mathbf{z}}$. Projectivising, we can view A as an element of $\mathrm{PU}(2,1)$ sending o to z as required. □

If we know a vector $\mathbf{e}_1$ Hermitian orthogonal to $\mathbf{z}$, then instead of using the Gram-Schmidt process, we could find $\mathbf{e}_2$ using the Hermitian cross product. That is, if

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}$$

then define $\mathbf{n}$ by

$$\mathbf{n} = \begin{bmatrix} \overline{p}_2\overline{q}_3 - \overline{p}_3\overline{q}_2 \\ \overline{p}_3\overline{q}_1 - \overline{p}_1\overline{q}_3 \\ \overline{p}_2\overline{q}_1 - \overline{p}_1\overline{q}_2 \end{bmatrix}. \tag{52}$$

Then $\mathbf{n}$ is orthogonal to $\mathbf{p}$ and $\mathbf{q}$ with respect to the first Hermitian form and

$$\langle \mathbf{n}, \mathbf{n} \rangle_1 = \langle \mathbf{p}, \mathbf{q} \rangle_1 \langle \mathbf{q}, \mathbf{p} \rangle_1 - \langle \mathbf{p}, \mathbf{p} \rangle_1 \langle \mathbf{q}, \mathbf{q} \rangle_1.$$

Corollary 3.2 *The stabiliser of a point in $\mathbf{H}^2_{\mathbb{C}}$ under* $\mathrm{PU}(2,1)$ *is* $\mathrm{P}\big(\mathrm{U}(2) \times \mathrm{U}(1)\big)$ *which is conjugate to* $\mathrm{U}(2)$. *Moreover, the stabiliser of the origin o in the ball model acts on $\mathbb{B}^2$ with the usual action of* $\mathrm{U}(2)$ *on* $\mathbb{C}^2$.

Proof: We work in the unit ball model. By the above proposition we can conjugate so that the point in question is the origin. Now any matrix in $\mathrm{PU}(2,1)$ fixing the origin is the projectivisation of a block diagonal matrix in $\mathrm{U}(2) \times \mathrm{U}(1)$ in $\mathrm{U}(2,1)$. In other words, it has the form

$$\begin{pmatrix} A & 0 \\ 0 & e^{i\theta} \end{pmatrix}$$

where $A \in \mathrm{U}(2)$ and $e^{i\theta} \in \mathrm{U}(1)$. Projectivising we may assume that $e^{i\theta} = 1$. Clearly all matrices of this form stabilise the origin. This gives the result. □

We now consider the action of $\mathrm{PU}(2,1)$ on the boundary. We choose to work with the second Hermitian form. We first show how to find a vector $\mathbf{n}$ that is Hermitian orthogonal to $\mathbf{p}$ and $\mathbf{q}$. If

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}$$

then

$$\mathbf{n} = \begin{bmatrix} \overline{p}_1\overline{q}_2 - \overline{p}_2\overline{q}_1 \\ \overline{p}_3\overline{q}_1 - \overline{p}_1\overline{q}_3 \\ \overline{p}_2\overline{q}_3 - \overline{p}_3\overline{q}_2 \end{bmatrix}. \tag{53}$$

A short computation shows that we again have

$$\langle \mathbf{n}, \mathbf{n} \rangle_2 = \langle \mathbf{p}, \mathbf{q} \rangle_2 \langle \mathbf{q}, \mathbf{p} \rangle_2 - \langle \mathbf{p}, \mathbf{p} \rangle_2 \langle \mathbf{q}, \mathbf{q} \rangle_2.$$

We now show that $\mathrm{PU}(2,1)$ acts doubly transitively on $\partial\mathbf{H}^2_{\mathbb{C}}$. For this we use the Siegel domain and the group preserving the second Hermitian form.

Proposition 3.3 *For any pair of points p and q in $\partial\mathbf{H}^2_{\mathbb{C}}$ there is an element of* $\mathrm{PU}(2,1)$ *sending the origin o to p and ∞ to q.*

Proof: We use the Siegel domain model.

Choose any lifts $\mathbf{p}$ and $\mathbf{q}$ of p and q to $\mathbb{C}^{2,1}$. Consider $\hat{\mathbf{p}} = \mathbf{p}/\langle \mathbf{p}, \mathbf{q} \rangle_2$. This means that $\langle \hat{\mathbf{p}}, \hat{\mathbf{p}} \rangle_2 = \langle \mathbf{q}, \mathbf{q} \rangle_2 = 0$ and $\langle \hat{\mathbf{p}}, \mathbf{q} \rangle_2 = 1$. Let $\mathbf{n}$ be the Hermitian orthogonal to $\hat{\mathbf{p}}$ and $\mathbf{q}$ given by (53). Then $\langle \mathbf{n}, \mathbf{n} \rangle_2 = 1$ and $\langle \mathbf{n}, \hat{\mathbf{p}} \rangle_2 = \langle \mathbf{n}, \mathbf{q} \rangle_2 = 0$. Let A be the matrix whose columns are $\mathbf{q}$, $\mathbf{n}$, $\hat{\mathbf{p}}$ respectively. Then $A^*H_2A = H_2$, that is A is unitary with respect to the second Hermitian form. Moreover, projectivising to a matrix in $\mathrm{PU}(2,1)$, we see A and sends o to p and ∞ to q as required. □

Lemma 3.4 *Let $\mathbf{p}$ and $\mathbf{q}$ be null vectors with $\langle \mathbf{p}, \mathbf{q} \rangle = -1$. Let $\mathbf{n}$ be normal to $\mathbf{p}$ and $\mathbf{q}$ with $\langle \mathbf{n}, \mathbf{n} \rangle = 1$. Then for any $\mathbf{z} \in \mathbb{C}^{2,1}$*

$$\langle \mathbf{z}, \mathbf{p} \rangle \langle \mathbf{q}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{q} \rangle \langle \mathbf{p}, \mathbf{z} \rangle = \left| \langle \mathbf{z}, \mathbf{n} \rangle \right|^2 - \langle \mathbf{z}, \mathbf{z} \rangle.$$

Proof: Write $\mathbf{z}$ in terms of $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{n}$. Then

$$\mathbf{z} = -\langle \mathbf{z}, \mathbf{q} \rangle \mathbf{p} - \langle \mathbf{z}, \mathbf{p} \rangle \mathbf{q} + \langle \mathbf{z}, \mathbf{n} \rangle \mathbf{n}.$$

Then

$$\langle \mathbf{z}, \mathbf{z} \rangle = -\langle \mathbf{z}, \mathbf{p} \rangle \langle \mathbf{q}, \mathbf{z} \rangle - \langle \mathbf{z}, \mathbf{q} \rangle \langle \mathbf{p}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{n} \rangle \langle \mathbf{n}, \mathbf{z} \rangle.$$

□

3.4 Complex hyperbolic isometries

Since the Bergman metric is given in terms of the Hermitian form $\langle\cdot,\cdot\rangle$ it is clear that if A is unitary with respect to $\langle\cdot,\cdot\rangle$ then A acts isometrically on the projective model of complex hyperbolic space. Thus $\mathrm{PU}(2,1)$ is a subgroup of the complex hyperbolic isometry group.

There are isometries of $\mathbf{H}^2_{\mathbb{C}}$ not in $\mathrm{PU}(2,1)$. For example, consider coordinate-wise complex conjugation $z \longmapsto \overline{z}$. Then

$$\cosh^2\Big(\frac{\rho(\overline{z},\overline{w})}{2}\Big) = \frac{\overline{\langle \mathbf{z},\mathbf{w}\rangle}\ \overline{\langle \mathbf{w},\mathbf{z}\rangle}}{\langle \mathbf{z},\mathbf{z}\rangle\langle \mathbf{w},\mathbf{w}\rangle} = \frac{\langle \mathbf{w},\mathbf{z}\rangle\langle \mathbf{z},\mathbf{w}\rangle}{\langle \mathbf{z},\mathbf{z}\rangle\langle \mathbf{w},\mathbf{w}\rangle} = \cosh^2\Big(\frac{\rho(z,w)}{2}\Big).$$

Therefore complex conjugation is also an isometry of complex hyperbolic space.

We now show that the holomorphic isometry group of $\mathbf{H}^2_{\mathbb{C}}$ is $\mathrm{PU}(2,1)$ and that the full isometry group is generated by $\mathrm{PU}(2,1)$ and complex conjugation.

Theorem 3.5 *Every isometry of $\mathbf{H}^2_{\mathbb{C}}$ is either holomorphic or anti-holomorphic. Moreover, each holomorphic isometry of $\mathbf{H}^2_{\mathbb{C}}$ is given by a matrix in $\mathrm{PU}(2,1)$ and each anti-holomorphic isometry is given by complex conjugation followed by a matrix in $\mathrm{PU}(2,1)$.*

Proof: We use the ball model $\mathbb{B}^2$ of $\mathbf{H}^2_{\mathbb{C}}$ and the first Hermitian form. Let Φ be any isometry of $\mathbf{H}^2_{\mathbb{C}}$. By applying an element of $\mathrm{PU}(n,1)$ and using Proposition 3.1, we may assume that Φ fixes the origin.

Let (z_1,z_2) be any point in $\mathbb{B}^2$ and let $(w_1,w_2)=\Phi(z_1,z_2)$. Then

$$\begin{aligned}\frac{1}{1-|z_1|^2-|z_2|^2} &= \cosh^2\left(\frac{\rho\big((z_1,z_2),(0,0)\big)}{2}\right)\\ &= \cosh^2\left(\frac{\rho\big((w_1,w_2),(0,0)\big)}{2}\right)\\ &= \frac{1}{1-|w_1|^2-|w_2|^2}.\end{aligned}$$

Thus $|z_1|^2+|z_2|^2=|w_1|^2+|w_2|^2$.

Also, using Corollary 3.2, we may assume that Φ maps $(1/2,0)\subset\mathbb{B}^2$ to some point $(x,0)\subset\mathbb{B}^2$ with $0\le x<1$. Using this identity applied to $\Phi(1/2,0)=(x,0)$ we can see that $x=1/2$ and so Φ fixes $(1/2,0)\in\mathbb{B}^2$.

Now consider $\Phi(r,0)=(a+ib,c+id)$ for any $0<r<1$. From the above remark we see that

$$r^2=a^2+b^2+c^2+d^2.$$

In particular, $a\le r<1$ and so $(1-r/2)^2\le(1-a/2)^2$. Also

$$\begin{aligned}\frac{(1-r/2)^2}{(1-r^2)(1-1/4)} &= \cosh^2\left(\frac{\rho\big((r,0),(1/2,0)\big)}{2}\right)\\ &= \cosh^2\left(\frac{\rho\big((a+ib,c+id),(1/2,0)\big)}{2}\right)\\ &= \frac{(1-a/2)^2+(b/2)^2}{(1-a^2-b^2-c^2-d^2)(1-1/4)}.\end{aligned}$$

Thus we have $(1-r/2)^2 = (1-a/2)^2 + b^2$. In other words

$$(1-a/2)^2 \geq (1-r/2)^2 = (1-a/2)^2 + b^2 \geq (1-a/2)^2.$$

Thus $a = r$ and $b = c = d = 0$. Hence Φ fixes $(r, 0)$ for all $0 \leq r < 1$.

Now consider $\Phi(0, 1/2) = (a + ib, c + id)$. As before $1/4 = a^2 + b^2 + c^2 + d^2$. Moreover, for all $0 < r < 1$ we have:

$$\begin{aligned}
\frac{1}{(1-1/4)(1-r^2)} &= \cosh^2\left(\frac{\rho\big((0,1/2),(r,0)\big)}{2}\right) \\
&= \cosh^2\left(\frac{\rho\big((a+ib,c+id),(r,0)\big)}{2}\right) \\
&= \frac{(1-ar)^2+(br)^2}{(1-a^2-b^2-c^2-d^2)(1-r^2)}.
\end{aligned}$$

In other words $(a^2+b^2)r^2 - 2ar + 1 = 1$ for all $0 < r < 1$. Therefore $a = b = 0$ and $\Phi(0, 1/2) = (0, c + id)$. We may apply an element of $\mathrm{PU}(2, 1)$ fixing $(r, 0)$ and sending $(0, c+id)$ to $(0, s)$ for $0 < s < 1$. It is then clear that $s = 1/2$ and, reasoning as above, we can show $\Phi(0, r) = (0, r)$ for all $0 \leq r < 1$.

Finally consider $\Phi(z_1, z_2) = (w_1, w_2)$ for any $(z_1, z_2) \in \mathbb{B}^2$. Then, arguing as above, $|z_1|^2 + |z_2|^2 = |w_1|^2 + |w_2|^2$ and, for all r with $0 < r < 1$, we have

$$\begin{aligned}
\frac{(1-rz_1)(1-r\overline{z}_1)}{(1-|z_1|^2-|z_2|^2)(1-r^2)} &= \cosh^2\left(\frac{\rho\big((z_1,z_2),(r,0)\big)}{2}\right) \\
&= \cosh^2\left(\frac{\rho\big((w_1,w_2),(r,0)\big)}{2}\right) \\
&= \frac{(1-rw_1)(1-r\overline{w_1})}{(1-|w_1|^2-|w_2|^2)(1-r^2)}.
\end{aligned}$$

Thus $|1 - rz_1|^2 = |1 - rw_1|^2$ and equating coefficients of r we see $|z_1|^2 = |w_1|^2$ and $\Re(z_1) = \Re(w_1)$. In other words $z_1 = w_1$ or $z_1 = \overline{w_1}$. Also

$$\begin{aligned}
\frac{(1-rz_2)(1-r\overline{z}_2)}{(1-|z_1|^2-|z_2|^2)(1-r^2)} &= \cosh^2\left(\frac{\rho\big((z_1,z_2),(0,r)\big)}{2}\right) \\
&= \cosh^2\left(\frac{\rho\big((w_1,w_2),(0,r)\big)}{2}\right) \\
&= \frac{(1-rw_2)(1-r\overline{w_2})}{(1-|w_1|^2-|w_2|^2)(1-r^2)}.
\end{aligned}$$

A similar argument gives that $z_2 = w_2$ or $z_2 = \overline{w_2}$.

It is easy to check that $(z_1, z_2) \longmapsto (z_1, \overline{z}_2)$ and $(z_1, z_2) \longmapsto (\overline{z}_1, z_2)$ are not isometries. Thus Φ is either the identity, or complex conjugation.

Therefore any isometry of $\mathbf{H}^2_{\mathbb{C}}$ is either in $\mathrm{PU}(2, 1)$, which means it is holomorphic, or it is an element of $\mathrm{PU}(2, 1)$ followed by complex conjugation, which means it is anti-holomorphic. $\square$

4 The boundary

4.1 Relation to the Heisenberg group

We recall that one model of real hyperbolic n-space $\mathbf{H}^n_{\mathbb{R}}$ is the upper half space in $\mathbb{R}^n$, that is $\big\{(x_1, \dots, x_n)|x_1, \dots, x_n \in \mathbb{R},\, x_n > 0\big\}$. The boundary of this model is the one point compactification of $\mathbb{R}^{n-1}$ thought of as the subspace of $\mathbb{R}^n$ given by $x_n = 0$. This model of real hyperbolic n space is foliated by horospheres H_u for $u > 0$. The horosphere H_u is a copy of $\mathbb{R}^{n-1}$ given by points with $x_n = u$. We want to form the analogous construction for complex hyperbolic space.

In this section we work in the Siegel domain model and we consider $\mathrm{PU}(2,1)$ preserving the second Hermitian form. First we study the boundary a little more carefully. A finite point z is in the boundary of the Siegel domain if its standard lift to $\mathbb{C}^{2,1}$ is $\mathbf{z}$ where

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ 1 \end{bmatrix} \qquad \text{where} \qquad z_1 + \overline{z}_1 + |z_2|^2 = 0.$$

We write $\zeta = z_2/\sqrt{2} \in \mathbb{C}$ and this condition becomes $2\Re(z_1) = -2|\zeta|^2$. Hence we may write $z_1 = -|\zeta|^2 + iv$ for $v \in \mathbb{R}$. That is for $\zeta \in \mathbb{C}$ and $v \in \mathbb{R}$:

$$\mathbf{z} = \begin{bmatrix} -|\zeta|^2 + iv \\ \sqrt{2}\zeta \\ 1 \end{bmatrix}$$

Therefore we may identify the boundary of the Siegel domain with the one point compactification of $\mathbb{C} \times \mathbb{R}$.

We now investigate the effect of isometries in $\mathrm{PU}(2,1)$ on these finite boundary points. We will show that the collection of these points has a group law giving it the structure of the Heisenberg group. Thus the boundary of the Siegel domain is the one point compactification of the Heisenberg group.

Lemma 4.1 *Suppose that $A \in \mathrm{PU}(2,1)$ has the standard form (30). Then the following are equivalent:*

(i) A fixes ∞,

(ii) A is upper triangular,

(iii) $g = 0$.

Proof: Using the notation of (30), we see that A fixes ∞, if and only if $d = g = 0$. Moreover, A fixes ∞ if and only if A^{-1} also fixes ∞. Using the expression for A^{-1} given in (30) we see that A^{-1} fixes ∞ if and only if $h = g = 0$. Thus A fixes ∞ if and only if it is upper triangular.

Clearly if A is upper triangular then $g = 0$. Conversely, assume $g = 0$. Using (45) and (51) we see that this implies $|d|^2 = |h|^2 = 0$. This proves the result. $\square$

Consider the map T from $\mathbb{C}\times\mathbb{R}$ to $\mathrm{GL}(3,\mathbb{C})$ given by

$$T(\zeta,v)=\begin{bmatrix}1 & -\sqrt{2}\,\overline{\zeta} & -|\zeta|^2+iv\\ 0 & 1 & \sqrt{2}\zeta\\ 0 & 0 & 1\end{bmatrix}.$$

It is easy to check that this fixes infinity and sends the origin to the point (ζ,v). It is also easy to check that $T(\zeta,v)$ is in $\mathrm{PU}(2,1)$ since

$$\big(T(\zeta,v)\big)^{-1}=H_2\big(T(\zeta,v)\big)^*H_2=\begin{bmatrix}1 & \sqrt{2}\,\overline{\zeta} & -|\zeta|^2-iv\\ 0 & 1 & -\sqrt{2}\zeta\\ 0 & 0 & 1\end{bmatrix}=T(-\zeta,-v).$$

In order to find the group law we multiply two such matrices

$$\begin{aligned}T(\zeta,v)T(\xi,t)&=\begin{bmatrix}1 & -\sqrt{2}\,\overline{\zeta} & -|\zeta|^2+iv\\ 0 & 1 & \sqrt{2}\zeta\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix}1 & -\sqrt{2}\,\overline{\xi} & -|\xi|^2+it\\ 0 & 1 & \sqrt{2}\xi\\ 0 & 0 & 1\end{bmatrix}\\ &=\begin{bmatrix}1 & -\sqrt{2}(\overline{\zeta}+\overline{\xi}) & -|\zeta+\xi|^2+iv+it+\overline{\xi}\zeta-\overline{\zeta}\xi\\ 0 & 1 & \sqrt{2}(\zeta+\xi)\\ 0 & 0 & 1\end{bmatrix}\\ &=T(\zeta+\xi,v+t+2\Im(\overline{\xi}\zeta)).\end{aligned}$$

This means that T is a group homomorphism to $\mathrm{PU}(2,1)$ from $\mathbb{C}\times\mathbb{R}$ with the group law

$$(\zeta,v)*(\xi,t)=(\zeta+\xi,v+t+2\Im(\overline{\xi}\zeta)).$$

This group law gives $\mathbb{C}\times\mathbb{R}$ the structure of the 3 dimensional Heisenberg group $\mathcal{N}$. We also remark that $\Im(\overline{\xi}\zeta)=\omega(\zeta,\xi)$ where ω is the standard symplectic form on $\mathbb{C}$.

The Heisenberg group is not Abelian but is 2-step nilpotent. In order to see this observe that

$$(\zeta,v)*(\xi,t)*(-\zeta,-v)*(-\xi,-t)=(0,4\Im(\overline{\xi}\zeta)).$$

Therefore any point in $\mathcal{N}$ of the form $(0,t)$ is central and the commutator of any two elements lies in the centre.

Geometrically, we think of the $\mathbb{C}$ factor of $\mathcal{N}$ as being horizontal and the $\mathbb{R}$ factor as being vertical. We refer to $T(\zeta,v)$ as *Heisenberg translation* by (ζ,v). A Heisenberg translation by $(0,t)$ is called *vertical translation* by t. It is easy to see the Heisenberg translations are ordinary translations in the horizontal direction and shears in the vertical direction. The fact that $\mathcal{N}$ is nilpotent means that translating around a horizontal square gives a vertical translation, rather like going up a spiral staircase. We define *vertical projection* $\Pi:\mathcal{N}\longrightarrow\mathbb{C}$ to be the map $\Pi(\zeta,v)=\zeta$.

4.2 Horospherical coordinates

Fix $u\in\mathbb{R}_+$ and consider all those points $z\in\mathbf{H}^2_{\mathbb{C}}$ for which the standard lift $\mathbf{z}$ has $\langle\mathbf{z},\mathbf{z}\rangle=-2u$. This is equivalent to saying

$$\mathbf{z}=\begin{bmatrix}z_1\\ z_2\\ 1\end{bmatrix}\qquad\text{where}\qquad z_1+\overline{z}_1+|z_2|^2=-2u.$$

In other words, $2\Re(z_1) = -|z_2|^2 - 2u$. We again write $z_2 = \sqrt{2}\zeta$ which means that $z_1 = -|\zeta|^2 - u + iv$. Thus $\mathbf{z}$ corresponds to a point $(\zeta, v, u) \in \mathbb{C} \times \mathbb{R} \times \mathbb{R}_+$ via

$$\mathbf{z} = \begin{bmatrix} -|\zeta|^2 - u + iv \\ \sqrt{2}\zeta \\ 1 \end{bmatrix}.$$

Let H_u denote the set of points in $\mathbf{H}^2_{\mathbb{C}}$ with $\langle \mathbf{z}, \mathbf{z} \rangle_2 = -2u$. This is called the *horosphere* of *height* u. It carries the structure of the Heisenberg group. Thus, for example (left) Heisenberg translation by (τ, t) is given by

$$T(\tau, t) \,:\, (\zeta, v, u) \longmapsto (\zeta + \tau, v + t + 2\Im(\bar{\zeta}\tau), u).$$

In this way we canonically identify a point z in the Siegel domain with $(\zeta, v, u) \in \mathcal{N} \times \mathbb{R}_+$ and we call (ζ, v, u) the *horospherical coordinates* of z. Sometimes it is useful to identify the finite boundary points with the horosphere of height zero, that is $H_0 = \partial \mathbf{H}^2_{\mathbb{C}} - \{\infty\}$. This means that $(\zeta, v) = (\zeta, v, 0) \in \partial \mathbf{H}^2_{\mathbb{C}} - \{\infty\}$.

Likewise, we define the *horoball* U_t of height t to be the union of all horospheres of height $u > t$. This is an open (topological) ball of dimension 4. Thus $\mathbf{H}^2_{\mathbb{C}}$ is itself the horoball U_0 of height 0.

With respect to horospherical coordinates the second Hermitian form is given by

$$\langle (\zeta, v, u), (\xi, t, s) \rangle_2 = -|\zeta - \xi|^2 - u - s + i\big(v - t + 2\Im(\bar{\xi}\zeta)\big).$$

With respect to horospherical coordinates $(x + iy, v, u)$ the Bergman metric is given by

$$\begin{aligned} ds^2 &= \frac{-4}{\langle \mathbf{z}, \mathbf{z} \rangle^2} \det \begin{pmatrix} \langle \mathbf{z}, \mathbf{z} \rangle & \langle d\mathbf{z}, \mathbf{z} \rangle \\ \langle \mathbf{z}, d\mathbf{z} \rangle & \langle d\mathbf{z}, d\mathbf{z} \rangle \end{pmatrix} \\ &= \frac{1}{u^2} \left(4u\, dx^2 + 4u\, dy^2 + du^2 + (dv + 2x\, dy - 2y\, dx)^2 \right) \\ &= \begin{pmatrix} dx & dy & du & dv \end{pmatrix} \begin{pmatrix} \frac{4(u+y^2)}{u^2} & \frac{-4xy}{u^2} & 0 & \frac{-2y}{u^2} \\ \frac{-4xy}{u^2} & \frac{4(u+x^2)}{u^2} & 0 & \frac{2x}{u^2} \\ 0 & 0 & \frac{1}{u^2} & 0 \\ \frac{-2y}{u^2} & \frac{2x}{u^2} & 0 & \frac{1}{u^2} \end{pmatrix} \begin{pmatrix} dx \\ dy \\ du \\ dv \end{pmatrix}. \end{aligned}$$

Hence, as a Riemannian metric, the Bergman metric is given by the inner product on the (real) tangent space to the Siegel domain defined, with respect to the basis $(dx,\, dy,\, du,\, dv)$, by the matrix $\mathbf{g}$, where

$$\mathbf{g} = \begin{pmatrix} 4(u + y^2)/u^2 & -4xy/u^2 & 0 & -2y/u^2 \\ -4xy/u^2 & 4(u + x^2)/u^2 & 0 & 2x/u^2 \\ 0 & 0 & 1/u^2 & 0 \\ -2y/u^2 & 2x/u^2 & 0 & 1/u^2 \end{pmatrix}. \tag{54}$$

Therefore the volume form on the Siegel domain is given by

$$d\mathrm{Vol} = \sqrt{\det(\mathbf{g})}\, dx\, dy\, du\, dv = \frac{4}{u^3}\, dx\, dy\, du\, dv. \tag{55}$$

Alternatively, we could have used $-z_1-\overline{z}_1-|z_2|^2=2u$ to derive this from Proposition 2.3 and the Jacobian relating horospherical coordinates and Siegel domain coordinates:

$$dx\,dy\,du\,dv = 2\,dx_1\,dy_1\,dx_2\,dy_2.$$

We now investigate the how the Cayley transform from Section 2.3 changes between horospherical and ball coordinates. Consider (x_1+iy_1, x_2+iy_2) in the unit ball in $\mathbb{C}^2$. Taking the canonical lift to $\mathbb{C}^{2,1}$ and then applying the Cayley transform gives

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0 & 1\\ 0 & \sqrt{2} & 0\\ 1 & 0 & -1\end{bmatrix}\begin{bmatrix}x_1+iy_1\\ x_2+iy_2\\ 1\end{bmatrix} = \begin{bmatrix}(x_1+1+iy_1)/\sqrt{2}\\ x_2+iy_2\\ (x_1-1+iy_1)/\sqrt{2}\end{bmatrix}.$$

Projectivising so that the third coordinate is 1 this becomes

$$\begin{aligned}&\begin{bmatrix}(x_1+1+iy_1)/(x_1-1+iy_1)\\ \sqrt{2}(x_2+iy_2)/(x_1-1+iy_1)\\ 1\end{bmatrix}\\ &= \begin{bmatrix}({x_1}^2+{y_1}^2-1-2iy_1)/\big((x_1-1)^2+{y_1}^2\big)\\ \sqrt{2}(x_1x_2+y_1y_2-x_2+ix_1y_2-iy_1x_2-iy_2)/\big((x_1-1)^2+{y_1}^2\big)\\ 1\end{bmatrix}.\end{aligned}$$

Thus the horospherical coordinates become $(\zeta,\, v\, u)$ where

$$\begin{aligned}\zeta &= \frac{x_1x_2+y_1y_2-x_2+ix_1y_2-iy_1x_2-iy_2}{(x_1-1)^2+{y_1}^2},\\ v &= \frac{-2y_1}{(x_1-1)^2+{y_1}^2},\\ u &= \frac{1-{x_1}^2-{y_1}^2-{x_2}^2-{y_2}^2}{(x_1-1)^2+{y_1}^2}.\end{aligned}$$

A straightforward, though lengthy, computation shows that the Jacobian of this transformation is

$$\mathcal{J} = \frac{4}{\big((x_1-1)^2+{y_1}^2\big)^3}.$$

Hence horospherical coordinates and ball coordinates are related by

$$dx\,dy\,du\,dv = \frac{4}{\big((x_1-1)^2+{y_1}^2\big)^3}dx_1\,dy_1\,dx_2\,dy_2.$$

Using Proposition 2.1, we can see again how the volume form transforms between different sets of coordinates:

$$d\mathrm{Vol} = \frac{16}{(1-{x_1}^2-{y_1}^2-{x_2}^2-{y_2}^2)^3}dx_1\,dy_1\,dx_2\,dy_2. = \frac{4}{u^3}dx\,dy\,du\,dv.$$

We may also define horospherical coordinates based at a point other than infinity. Later we will want to do this for horospherical coordinates based at the origin $o=(0,0)\in\mathcal{N}$.

Because horospherical coordinates are not defined intrinsically but require some normalisation we need to be careful about what the horosphere centred at o of height u means.

We define horospheres and horoballs based at o as the image of those based at ∞ under the inversion ι given by the matrix

$$\iota = \begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix}. \tag{56}$$

We now investigate the effect of ι on horospherical coordinates

$$\begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -|z|^2 - u + iv \\ \sqrt{2}z \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -\sqrt{2}z \\ |z|^2 + u - iv \end{bmatrix} \approx \begin{bmatrix} \frac{-|z|^2-u-iv}{(|z|^2+u)^2+v^2} \\ \frac{-\sqrt{2}z}{|z|^2+u-iv} \\ 1 \end{bmatrix}.$$

Thus the map ι carries the point of $\mathbf{H}^2_{\mathbb{C}}$ with horospherical coordinates (ζ, v, u) to the point with coordinates

$$\iota(\zeta, v, u) = \left(\frac{-\zeta}{|\zeta|^2 + u - iv}, \frac{-v}{\left(|\zeta|^2 + u\right)^2 + v^2}, \frac{u}{\left(|\zeta|^2 + u\right)^2 + v^2} \right). \tag{57}$$

Using this we define horospherical coordinates about o as the image under ι of horospherical coordinates about ∞.

Similarly elements of $\mathrm{PU}(2,1)$ fixing o may be obtained from those fixing ∞ by conjugating by ι. Thus we may speak of Heisenberg translation by (τ, t) fixing o. This is just the conjugate by ι of the Heisenberg translation by (τ, t) fixing ∞. As a matrix in $\mathrm{PU}(2,1)$ it is given by

$$\begin{bmatrix} 1 & 0 & 0 \\ \sqrt{2}\tau & 1 & 0 \\ -|\tau|^2 + it & -\sqrt{2}\,\overline{\tau} & 1 \end{bmatrix}.$$

4.3 The Cygan metric

In this section we define a metric on the Heisenberg group, the Cygan metric. We extend the Cygan metric to an incomplete metric on $\mathbf{H}^2_{\mathbb{C}}$ which agrees with the Cygan metric on each horosphere. This metric should be thought of as the counterpart to the Euclidean metric on the upper half space model of real hyperbolic space.

The Heisenberg norm is given by

$$\left|(\zeta, v)\right| = \left| |\zeta|^2 - iv \right|^{1/2}.$$

This gives rise to a metric, the Cygan metric, on $\mathcal{N}$ by

$$\rho_0\big((\zeta_1, v_1),\, (\zeta_2, v_2)\big) = \left|(\zeta_1, v_1)^{-1} * (\zeta_2, v_2)\right|.$$

In other words

$$\rho_0\big((\zeta_1,v_1),\,(\zeta_2,v_2)\big)=\Big||\zeta_1-\zeta_2|^2-iv_1+iv_2-2i\Im(\zeta_1\overline{\zeta}_2)\Big|^{1/2}. \tag{58}$$

If we take the standard lift of points on $\partial\mathbf{H}^2_{\mathbb{C}}-\{\infty\}$ to $\mathbb{C}^{2,1}$ we can write the Cygan metric in terms of the second Hermitian form:

$$\rho_0\big((\zeta_1,v_1),\,(\zeta_2,v_2)\big)=\left|\left\langle\begin{bmatrix}-|\zeta_1|^2+iv_1\\ \sqrt{2}\zeta_1\\ 1\end{bmatrix},\begin{bmatrix}-|\zeta_2|^2+iv_2\\ \sqrt{2}\zeta_2\\ 1\end{bmatrix}\right\rangle\right|^{1/2}.$$

Exercise 4.2 *Prove that the Cygan metric given by (58) satisfies the triangle inequality. (See Proposition 4.3 below.)*

We remark that the Cygan metric is not a path metric. That is, there exist pairs of points such that the Cygan distance between them is strictly shorter than the Cygan length of any path joining them (see section 3.1 of [8] for more details of the connection between metrics and path metrics). In order to demonstrate this fact it suffices to give a pair of points (ζ_1,v_1) and (ζ_2,v_2) so that for all points (ζ_3,v_3) with $(\zeta_3,v_3)\neq(\zeta_1,v_1),\,(\zeta_2,v_2)$ the triangle inequality is strict. That is

$$\rho_0\big((\zeta_1,v_1),\,(\zeta_2,v_2)\big)<\rho_0\big((\zeta_1,v_1),\,(\zeta_3,v_3)\big)+\rho_0\big((\zeta_3,v_3),\,(\zeta_2,v_2)\big).$$

By the triangle inequality, the Cygan length of any path joining (ζ_1,v_1) and (ζ_2,v_2) is at least as big as the right hand side of this inequality. It is easiest to demonstrate this when $(\zeta_1,v_1)=(0,0)$ and $(\zeta_2,v_2)=(0,1)$. Then we have

$$\begin{aligned}
&\rho_0\big((0,0),\,(\zeta_3,v_3)\big)+\rho_0\big((\zeta_3,v_3),\,(0,1)\big)\\
&\quad=\ \Big||\zeta_3|^2+iv_3\Big|^{1/2}+\Big||\zeta_3|^2+i(v_3-1)\Big|^{1/2}\\
&\quad\geq\ |v_3|^{1/2}+|v_3-1|^{1/2}\\
&\quad\geq\ 1\\
&\quad=\ \rho_0\big((0,0),(0,1)\big).
\end{aligned}$$

In the first inequality we have have equality if and only if $\zeta_3=0$ and in the second inequality we have equality if and only if $v_3=0$ or $v_3=1$. Thus we have strict inequality whenever $(\zeta_3,v_3)\neq(0,0),\,(0,1)$.

We can extend the Cygan metric to an incomplete metric on $\overline{\mathbf{H}^2_{\mathbb{C}}}-\{\infty\}$ as follows

$$\rho_0\big((\zeta_1,v_1,u_1),\,(\zeta_2,v_2,u_2)\big)=\Big||\zeta_1-\zeta_2|^2+|u_1-u_2|-iv_1+iv_2-2i\Im(\zeta_1\overline{\zeta}_2)\Big|^{1/2}. \tag{59}$$

We remark that this agrees with $\big|\langle\mathbf{z}_1,\mathbf{z}_2\rangle\big|^{1/2}$ if and only if one (or both) of $\mathbf{z}_1$ or $\mathbf{z}_2$ is null, that is it corresponds to a point of $\partial\mathbf{H}^2_{\mathbb{C}}$. We now show that the extended Cygan metric on $\overline{\mathbf{H}^2_{\mathbb{C}}}-\{\infty\}$ is indeed a metric. By restricting to points on $\partial\mathbf{H}^2_{\mathbb{C}}-\{\infty\}$, this will also show that the original Cygan metric on $\mathcal{N}$ is a metric.

Proposition 4.3 *The function $\overline{\mathbf{H}^2_{\mathbb{C}}}-\{\infty\}$ given by equation (59) is a metric.*

Proof: It is completely obvious that that both $\rho_0\big((\zeta_1,v_1,u_1),\,(\zeta_2,v_2,u_2)\big)=0$ if and only if $(\zeta_1,v_1,u_1)=(\zeta_2,v_2,u_2)$ and that

$$\rho_0\big((\zeta_1,v_1,u_1),\,(\zeta_2,v_2,u_2)\big)=\rho_0\big((\zeta_2,v_2,u_2),\,(\zeta_1,v_1,u_1)\big)$$

for all points in $\overline{\mathbf{H}^2_{\mathbb{C}}}-\{\infty\}$. Therefore, it suffices to show that

$$\begin{aligned}&\rho_0\big((\zeta_1,v_1,u_1),\,(\zeta_2,v_2,u_2)\big)\\ &\quad\le\ \rho_0\big((\zeta_1,v_1,u_1),\,(\zeta_3,v_3,u_3)\big)+\rho_0\big((\zeta_3,v_3,u_3),\,(\zeta_2,v_2,u_2)\big).\end{aligned}$$

Now we have

$$\begin{aligned}&|\zeta_1-\zeta_2|^2-2i\Im(\zeta_1\overline{\zeta}_2)\\ &\quad=\ |\zeta_1|^2-2\zeta_1\overline{\zeta}_2+|\zeta_2|^2\\ &\quad=\ |\zeta_1-\zeta_3|^2-2(\zeta_1-\zeta_3)(\overline{\zeta}_2-\overline{\zeta}_3)+|\zeta_2-\zeta_3|^2-2i\Im(\zeta_1\overline{\zeta}_3)-2i\Im(\zeta_3\overline{\zeta}_2).\end{aligned}$$

Therefore

$$\begin{aligned}&\rho_0^2\big((\zeta_1,v_1,u_1),\,(\zeta_2,v_2,u_2)\big)\\ &\quad=\ \Big||\zeta_1-\zeta_2|^2+|u_1-u_2|-iv_1+iv_2-2i\Im(\zeta_1\overline{\zeta}_2)\Big|\\ &\quad\le\ \Big||\zeta_1-\zeta_2|^2+|u_1-u_3|+|u_3-u_2|-iv_1+iv_3-iv_3+iv_2-2i\Im(\zeta_1\overline{\zeta}_2)\Big|\\ &\quad\le\ \Big||\zeta_1-\zeta_3|^2+|u_1-u_3|-iv_1+iv_3-2i\Im(\zeta_1\overline{\zeta}_3)\Big|\\ &\qquad\quad+2|\zeta_1-\zeta_3|\,|\zeta_3-\zeta_2|+\Big||\zeta_3-\zeta_2|^2+|u_3-u_2|-iv_3+iv_2-2i\Im(\zeta_3\overline{\zeta}_2)\Big|\\ &\quad\le\ \Big(\rho_0\big((\zeta_1,v_1,u_1),\,(\zeta_3,v_3,u_3)\big)+\rho_0\big((\zeta_3,v_3,u_3),\,(\zeta_2,v_2,u_2)\big)\Big)^2.\end{aligned}$$

$\square$

We conclude this section by considering spheres with respect to the Cygan metric. The *Cygan sphere* of radius $r\in\mathbb{R}_+$ and centre $z_0=(\zeta_0,v_0)=(\zeta_0,v_0,0)\in\partial\mathbf{H}^2_{\mathbb{C}}$ is defined by

$$S_r(z_0)=\{z=(\zeta,v,u)\ :\ \rho_0(z,z_0)=r\}.$$

In terms of coordinates, $S_r(z_0)$ is given by

$$S_r(z_0)=\Big\{z=(\zeta,v,u)\ :\ \Big||\zeta-\zeta_0|^2+u+iv-iv_0-2i\Im(\zeta\overline{\zeta}_0)\Big|=r^2\Big\}.$$

Suppose that $\zeta_0=0$. Cygan spheres centred at $z_0=(0,v_0)$ are ovoids with the property that along the locus $\zeta=0$ they have fourth order contact with their tangent plane. The diameter of their equator, that is the points $(\zeta,0,0)$ with $|\zeta|=r$, grows linearly with r. On the other hand, the diameter of their meridians, that is the points $(0,v,u)$ with $|u+iv-iv_0|=r^2$, grows quadratically with r. Thus, as r tends to zero, Cygan spheres become very short and fat, like a pancake, and, as r tends to infinity, Cygan spheres of radius r become very long and thin, like a cigar.

When $\zeta_0\neq 0$, Cygan spheres are sheared ovoids, the magnitude of the shear being proportional to $|\zeta_0|$. Otherwise they enjoy the same properties outlined above. In particular, Cygan spheres are always convex.

4.4 Complex hyperbolic isometries and the Cygan metric

In this section we consider the action of complex hyperbolic isometries on the boundary of complex hyperbolic space. Of course not every isometry is an isometry of the Cygan metric and it is interesting to see how the Cygan metric changes when we apply an isometry. This is a direct generalisation of the way Möbius transformations in $\mathrm{PSL}(2,\mathbb{R})$ or $\mathrm{PSL}(2,\mathbb{C})$ act on the extended real line or the Riemann sphere respectively and how they distort the Euclidean metric. Readers might find it useful to keep this example in mind.

First consider the subgroup of $\mathrm{PU}(2,1)$ stabilising the point at infinity, whose elements will be called *Heisenberg similarities*. We have already seen the group of Heisenberg translations. This is a normal subgroup of the group of Heisenberg similarities. Using this subgroup, it is sufficient to classify those elements of $\mathrm{PU}(2,1)$ fixing both ∞ and the origin $o = (0,0)$. For example, we have *Heisenberg rotations*. These are given by $(\zeta, v) \longmapsto (e^{i\theta}\zeta, v)$ and are boundary elliptic. Also we have (real) dilations $(\zeta, v) \longmapsto (r\zeta, r^2 v)$ where $r \in \mathbb{R}_+$. A product of a Heisenberg rotation and a real dilation is a *complex dilation*

$$A : (\zeta, v) \longmapsto (re^{i\theta}\zeta, r^2 v) = (\lambda\zeta, |\lambda|^2 v).$$

Here $\lambda = re^{i\theta}$ is the *multiplier* or *complex dilation factor* of A. Complex dilations are isomorphic to $\mathbb{R}_+ \times \mathrm{U}(1) = \mathbb{C}^*$. The group of Heisenberg similarities is the semi-direct product of the complex dilations and the Heisenberg translations, isomorphic to $\big(\mathbb{R}_+ \times \mathrm{U}(1)\big) \ltimes \mathcal{N}$.

Using the exact sequence

$$0 \longrightarrow \mathbb{R} \longrightarrow \mathcal{N} \xrightarrow{\Pi} \mathbb{C} \longrightarrow 0$$

where *vertical projection* $\Pi : \mathcal{N} \longrightarrow \mathbb{C}$ by $\Pi : (\zeta, v) \longmapsto \zeta$. There is an induced map from the group of Heisenberg isometries to the group of Euclidean isometries of the plane. Using this we obtain the following exact sequence

$$0 \longrightarrow \mathbb{R} \longrightarrow Isom(\mathcal{N}) \xrightarrow{\Pi_*} Isom(\mathbb{C}) \longrightarrow 1 \tag{60}$$

Consider $e^{i\theta} \in \mathrm{U}(1)$ and $\zeta_0 \in \mathbb{C}$. This pair corresponds to a Euclidean isometry

$$\zeta \longmapsto e^{i\theta}\zeta + \zeta_0.$$

This isometry can be represented by a matrix in $\mathrm{GL}(2,\mathbb{C})$ as follows:

$$\begin{bmatrix} e^{i\theta} & \zeta_0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \zeta \\ 1 \end{bmatrix} = \begin{bmatrix} e^{i\theta}\zeta + \zeta_0 \\ 1 \end{bmatrix}.$$

Therefore the map Π_* can be explicitly given by

$$\Pi_* : \begin{bmatrix} 1 & -\sqrt{2}\zeta_0 e^{i\theta} & -|\zeta_0|^2 + iv_0 \\ 0 & e^{i\theta} & \sqrt{2}\zeta_0 \\ 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} e^{i\theta} & \zeta_0 \\ 0 & 1 \end{bmatrix} \tag{61}$$

It is clear that

$$\ker(\Pi_*) = \left\{ \begin{bmatrix} 1 & 0 & iv_0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} : v_0 \in \mathbb{R} \right\},$$

the group of vertical translations fixing q_∞.

Lemma 4.4 *As a matrix in* $\mathrm{PU}(2,1)$ *(with respect to the second Hermitian form) the complex dilation* $A : (\zeta, v, u) \longmapsto (\lambda\zeta, |\lambda|^2 v, |\lambda|^2 u)$ *is given by the loxodromic matrix, also denoted* A*:*

$$A = \begin{bmatrix} \overline{\lambda} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda^{-1} \end{bmatrix}.$$

Proof: We may write

$$A = \begin{bmatrix} \overline{\lambda} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda^{-1} \end{bmatrix}$$

Choose any if $z = (\zeta, v, u) \in \mathbf{H}^2_{\mathbb{C}}$. Then taking the canonical lift $\mathbf{z}$ of z to $\mathbb{C}^{2,1}$, we see that $A(\zeta, v, u)$ is

$$\begin{aligned} \begin{bmatrix} \overline{\lambda} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda^{-1} \end{bmatrix} \begin{bmatrix} -|\zeta|^2 - u + iv \\ \sqrt{2}\zeta \\ 1 \end{bmatrix} &= \begin{bmatrix} \overline{\lambda}\big(-|\zeta|^2 - u + iv\big) \\ \sqrt{2}\zeta \\ \lambda^{-1} \end{bmatrix} \\ &= \overline{\lambda}^{-1} \begin{bmatrix} |\lambda|^2\big(-|\zeta|^2 - u + iv\big) \\ \sqrt{2}\lambda\zeta \\ 1 \end{bmatrix}. \end{aligned}$$

Thus the canonical lift $A(\mathbf{z})$ of $A(\zeta, v, u)$ to $\mathbb{C}^{2,1}$ is

$$A(\mathbf{z}) \begin{bmatrix} |\lambda|^2\big(-|\zeta|^2 - u + iv\big) \\ \sqrt{2}\lambda\zeta \\ 1 \end{bmatrix}.$$

□

The following lemma shows how complex dilations distort the Cygan metric and also how their Cygan translation lengths vary. These will be very useful to us when considering the action of complex dilations on $\partial\mathbf{H}^2_{\mathbb{C}}$.

Lemma 4.5 *Suppose that* $A \in \mathrm{PU}(n,1)$ *fixes* o *and* ∞ *and has complex multiplier* λ. *Writing* $M = |\lambda - 1| + |\lambda^{-1} - 1|$, *we have*

(i) $\rho_0\big(A(z), A(w)\big) = |\lambda|\rho_0(z, w)$ *for all* z, $w \in \mathbf{H}^2_{\mathbb{C}}$,

(ii) $\rho_0\big(A(z), z\big) \leq |\lambda|^{1/2} M^{1/2} \rho_0(z, o)$ *for all* $z \in \partial\mathbf{H}^2_{\mathbb{C}} - \{\infty\}$.

Proof: Let $z = (\zeta, v, u)$ and $w = (\xi, t, s)$. The canonical lifts $A(\mathbf{z})$ of $A(\zeta, v, u)$ and $A(\mathbf{w})$ of $A(\xi, t, s)$ to $\mathbb{C}^{2,1}$ are

$$A(\mathbf{z}) = \begin{bmatrix} |\lambda|^2\big(-|\zeta|^2 - u + iv\big) \\ \sqrt{2}\lambda\zeta \\ 1 \end{bmatrix} \quad \text{and} \quad A(\mathbf{w}) = \begin{bmatrix} |\lambda|^2\big(-|\xi|^2 - s + it\big) \\ \sqrt{2}\lambda\xi \\ 1 \end{bmatrix}.$$

From this we see that

$$\begin{aligned}\rho_0\big(A(z),A(w)\big) &= \Big||\lambda|^2\big(|\zeta-\xi|^2+|u-s|-iv+it-2i\Im(\zeta\overline{\xi})\big)\Big|^{1/2}\\ &= |\lambda|\Big||\zeta-\xi|^2+|u-s|-iv+it-2i\Im(\zeta\overline{\xi})\Big|^{1/2}\\ &= |\lambda|\rho_0(z,w).\end{aligned}$$

This proves (i). For part (ii) we argue similarly with $z=(\zeta,v,0)$:

$$\begin{aligned}\rho_0\big(A(z),z\big) &= \Big||\lambda|^2|\zeta|^2-2(\lambda\zeta)\overline{\zeta}+|\zeta|^2-i|\lambda|^2v+iv\Big|^{1/2}\\ &= \Big|\lambda(\overline{\lambda}-1)\big(|\zeta|^2-iv\big)-(\lambda-1)\big(|\zeta|^2+iv\big)\Big|^{1/2}\\ &\le \big(|\lambda|+1\big)^{1/2}|\lambda-1|^{1/2}\Big||\zeta|^2-iv\Big|^{1/2}\\ &= |\lambda|^{1/2}M^{1/2}\rho_0(z,0).\end{aligned}$$

This completes the proof of (ii). □

The next lemma shows how a map in $\mathrm{PU}(2,1)$ not fixing ∞ distorts the Cygan metric on the boundary.

Lemma 4.6 *Let B be any element of* $\mathrm{PU}(2,1)$ *that does not fix ∞. Then there exists a positive real number r_B depending only on B so that for all $z,\,w\in\partial\mathbf{H}^2_{\mathbb{C}}-\{\infty,\,B^{-1}(\infty)\}$ we have*

(i)

$$\rho_0\big(B(z),B(w)\big)=\frac{{r_B}^2\rho_0(z,w)}{\rho_0\big(z,B^{-1}(\infty)\big)\rho_0\big(w,B^{-1}(\infty)\big)},$$

(ii)

$$\rho_0\big(B(z),B(\infty)\big)=\frac{{r_B}^2}{\rho_0\big(z,B^{-1}(\infty)\big)}.$$

Proof: As above let $z=(\zeta,v,0)$ and $w=(\xi,t,0)$ have canonical lifts

$$\mathbf{z}=\begin{bmatrix}-|\zeta|^2+iv\\ \sqrt{2}\zeta\\ 1\end{bmatrix}\quad\text{and}\quad\mathbf{w}=\begin{bmatrix}-|\xi|^2+it\\ \sqrt{2}\xi\\ 1\end{bmatrix}.$$

Let

$$B=\begin{bmatrix}a&b&c\\ d&e&f\\ g&h&j\end{bmatrix}.$$

Define $r_B=1/|g|^{1/2}$. Since $B(\infty)\neq\infty$ we know $g\neq 0$ as so r_B is well defined. Clearly r_B only depends on B.

The canonical lifts of $B^{-1}(\infty)$ and $B(\infty)$ are

$$\begin{bmatrix} \overline{j}/\overline{g} \\ \overline{h}/\overline{g} \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a/d \\ d/g \\ 1 \end{bmatrix}.$$

Hence

$$\begin{aligned} \rho_0\big(z, B^{-1}(\infty)\big) &= \Big|-|\zeta|^2 + iv + \sqrt{2}\zeta h/g + j/g\Big|^{1/2}, \\ \rho_0\big(w, B^{-1}(\infty)\big) &= \Big|-|\xi|^2 + it + \sqrt{2}\xi h/g + j/g\Big|^{1/2}. \end{aligned}$$

Moreover, the canonical lifts $B(\mathbf{z})$ and $B(\mathbf{w})$ of Bz and Bw are

$$\begin{aligned} B(\mathbf{z}) &= \begin{bmatrix} \big(a(-|\zeta|^2 + iv) + b\sqrt{2}\zeta + c\big)/\big(g(-|\zeta|^2 + iv) + h\sqrt{2}\zeta + j\big) \\ \big(d(-|\zeta|^2 + iv) + e\sqrt{2}\zeta + f\big)/\big(g(-|\zeta|^2 + iv) + h\sqrt{2}\zeta + j\big) \\ 1 \end{bmatrix}, \\ B(\mathbf{w}) &= \begin{bmatrix} \big(a(-|\xi|^2 + it) + b\sqrt{2}\xi + c\big)/\big(g(-|\xi|^2 + it) + h\sqrt{2}\xi + j\big) \\ \big(d(-|\xi|^2 + it) + e\sqrt{2}\xi + f\big)/\big(g(-|\xi|^2 + it) + h\sqrt{2}\xi + j\big) \\ 1 \end{bmatrix}. \end{aligned}$$

Hence

$$\begin{aligned} &\rho_0\big(B(z), B(w)\big) \\ &= \frac{\Big|-|\zeta|^2 + iv + 2\zeta\overline{\xi} - |\xi|^2 - it\Big|^{1/2}}{\Big|g(-|\zeta|^2 + iv) + h\sqrt{2}\zeta + j\Big|^{1/2}\Big|g(-|\xi|^2 + it) + h\sqrt{2}\xi + j\Big|^{1/2}} \\ &= \frac{{r_B}^2 \rho_0(z, w)}{\rho_0\big(z, B^{-1}(\infty)\big)\rho_0\big(w, B^{-1}(\infty)\big)}. \end{aligned}$$

This proves (i). Similarly

$$\begin{aligned} \rho_0\big(B(z), B(\infty)\big) &= \frac{1}{\Big|g(-|\zeta|^2 + iv) + h\sqrt{2}\zeta + j\Big|^{1/2}|g|^{1/2}} \\ &= \frac{{r_B}^2}{\rho_0\big(z, B^{-1}(\infty)\big)}. \end{aligned}$$

This proves (ii). □

An important consequence of this proposition is that B sends the Cygan sphere of radius r_B with centre $B^{-1}(\infty)$ to the Cygan sphere of radius r_B with centre $B(\infty)$. Motivated by the analogous Euclidean spheres in real hyperbolic space, we define the *isometric sphere* of B to be the Cygan sphere of radius r_B and centre $B^{-1}(\infty)$.

Lemma 4.7 *Let B be a loxodromic map with multiplier $\lambda \in \mathbb{C}$, attractive fixed point p and repulsive fixed point q and isometric sphere of radius r_B. Suppose that p, $q \neq \infty$, and let $M = |\lambda - 1| + |\lambda^{-1} - 1|$. Then*

(i) $\rho_0\big(p, B^{-1}(\infty)\big) = |\lambda|^{1/2} r_B$ *and* $\rho_0\big(p, B(\infty)\big) = |\lambda|^{-1/2} r_B$,

(i) $\rho_0\big(q, B^{-1}(\infty)\big) = |\lambda|^{-1/2} r_B$ *and* $\rho_0\big(q, B(\infty)\big) = |\lambda|^{1/2} r_B$,

(iii) $\rho_0(p, q) \le M^{1/2} r_B$.

Proof: Let C be any element of $\mathrm{PU}(2,1)$ with $C(o) = p$ and $C(\infty) = q$ as found in Proposition 3.3. Let r_C be the radius of its isometric sphere. Then $A = C^{-1}BC$ is a complex dilation with multiplier λ. Using Lemma 4.6 (ii) with $z = B(z) = q$ we have

$$r_B{}^2 = \rho_0\big(q, B(\infty)\big)\rho_0\big(q, B^{-1}(\infty)\big).$$

Also, substituting for $B = CAC^{-1}$, $q = C(\infty)$ and using Lemma 4.6 (ii) again, but this time with C, we have

$$\begin{aligned}
\rho_0\big(q, B(\infty)\big) &= \rho_0\big(C(\infty), CAC^{-1}(\infty)\big) \\
&= \frac{r_C{}^2}{\rho_0\big(AC^{-1}(\infty), C^{-1}(\infty)\big)} \\
&= \frac{r_C{}^2}{|\lambda|\rho_0\big(C^{-1}(\infty), A^{-1}C^{-1}(\infty)\big)} \\
&= |\lambda|^{-1}\rho_0\big(C(\infty), CAC^{-1}(\infty)\big) \\
&= |\lambda|^{-1}\rho_0\big(q, B^{-1}(\infty)\big).
\end{aligned}$$

Part (ii) follows immediately. Part (i) follows by applying part (ii) to B^{-1}.

For part (iii) we begin with

$$\begin{aligned}
r_B{}^2 &= \rho_0\big(q, B(\infty)\big)\rho_0\big(q, B^{-1}(\infty)\big) \\
&= \rho_0\big(C(\infty), CAC^{-1}(\infty)\big)\rho_0\big(C(\infty), CA^{-1}C^{-1}(\infty)\big) \\
&= \frac{r_C{}^4}{\rho_0\big(C^{-1}(\infty), AC^{-1}(\infty)\big)\rho_0\big(C^{-1}(\infty), A^{-1}C^{-1}(\infty)\big)}.
\end{aligned}$$

Now using Lemma 4.5 (ii) we have

$$\begin{aligned}
\rho_0\big(C^{-1}(\infty), AC^{-1}(\infty)\big) &\le |\lambda|^{1/2}M^{1/2}\rho_0\big(o, C^{-1}(\infty)\big) \\
\rho_0\big(C^{-1}(\infty), A^{-1}C^{-1}(\infty)\big) &\le |\lambda|^{-1/2}M^{1/2}\rho_0\big(o, C^{-1}(\infty)\big).
\end{aligned}$$

Therefore

$$\begin{aligned}
r_B{}^2 &\ge \frac{r_C{}^4}{M\rho_0\big(o, C^{-1}(\infty)\big)^2} \\
&= \frac{\rho_0\big(C(o), C(\infty)\big)^2}{M}.
\end{aligned}$$

where we have used Lemma 4.6 (ii) again. Substituting $p = C(o)$ and $q = C(\infty)$ gives the result. □

We now discuss parabolic maps. As a matrix in $\mathrm{PU}(2,1)$ (left) Heisenberg translation by (τ, t) is given by

$$A = \begin{bmatrix} 1 & -\sqrt{2}\overline{\tau} & -|\tau|^2 + it \\ 0 & 1 & \sqrt{2}\tau \\ 0 & 0 & 1 \end{bmatrix}. \tag{62}$$

At a point $z=(\zeta,v)\in\mathcal{N}$ the Cygan translation length of A is given by

$$t_A(z)=\rho_0\big(A(z),z\big)=\Big||\tau|^2+it+4i\Im(\tau\overline{\zeta})\Big|^{1/2}.$$

We now estimate how this translation length changes with z.

Lemma 4.8 *Let A be Heisenberg translation by (τ,t) and let $t_A(z)$, $t_A(w)$ denote the Cygan translation length of A at the points $z=(\zeta,v)$ and $w=(\xi,s)$ in $\mathcal{N}$. Then*

$$t_A(w)^2\le t_A(z)^2+4|\tau|\,|\xi-\zeta|.$$

Proof: We have

$$\begin{aligned} t_A(w)^2 &= \Big||\tau|^2+it+4i\Im(\tau\overline{\xi})\Big| \\ &= \Big||\tau|^2+it+4i\Im(\tau\overline{\zeta})+4i\Im\big(\tau(\overline{\xi}-\overline{\zeta})\big)\Big| \\ &\le \Big||\tau|^2+it+4i\Im(\tau\overline{\zeta})\Big|+4|\tau|\,|\xi-\zeta|. \end{aligned}$$

$\square$

Let A be the screw parabolic map with fixed point ∞, multiplier $e^{i\theta}$ and axis the complex line $L_A=\big\{(\zeta,v,u)\,:\,\zeta=\zeta_0\in\mathbb{C}\big\}$. Suppose that A acts as Heisenberg translation by $(0,t)$ on L_A. In horospherical coordinates A is given by

$$A:(\zeta,v,u)\longmapsto\Big(\zeta e^{i\theta}+\zeta_0(1-e^{i\theta}),v+t+2\Im\big((\zeta-\zeta_0)\overline{\zeta}_0(1-e^{i\theta})\big),u\Big). \tag{63}$$

As a matrix in $\mathrm{PU}(2,1)$ the map A is given by

$$A=\begin{bmatrix}1 & \sqrt{2}\,\overline{\zeta}_0(1-e^{i\theta}) & -2|\zeta_0|^2(1-e^{i\theta})+it\\ 0 & e^{i\theta} & \sqrt{2}\zeta_0(1-e^{i\theta})\\ 0 & 0 & 1\end{bmatrix}.$$

At a point $z=(\zeta,v)\in\mathcal{N}$ the Cygan translation length of A is

$$t_A(z)=\rho_0\big(A(z),z\big)=\Big|2|\zeta-\zeta_0|^2(e^{i\theta}-1)+it\Big|^{1/2}.$$

We now give a result analogous to Lemma 4.8 for screw parabolic maps. There is a difference between those maps which, when given by (63), have $t\sin(\theta)\ge 0$ and those with $t\sin(\theta)<0$. When $t\sin(\theta)\ge 0$ it is easy to see that $t_A(z)$ is a monotone increasing function of $|\zeta-\zeta_0|$. On the other hand, if $t\sin(\theta)<0$ then $t_A(z)$ has a minimum when $|\zeta-\zeta_0|^2=-t\sin(\theta)/4(1-\cos(\theta))$. The minimum value of $t_A(z)$ is $\sqrt{|e^{i\theta}-1|t/2}$.

Lemma 4.9 *Let A be the screw parabolic map given by (63). Let $t_A(z)$, $t_A(w)$ denote the Cygan translation length of A at the points $z=(\zeta,v)$ and $w=(\xi,s)$ in $\mathcal{N}$.*

(i) If $t\sin(\theta)\ge 0$ then

$$t_A(w)\le t_A(z)+\Big|2(e^{i\theta}-1)\Big|^{1/2}|\xi-\zeta|.$$

(ii) If $t\sin(\theta) < 0$ *then*

$$t_A(w) \leq t_A(z) + 2|\xi - \zeta|.$$

Proof: Suppose that $t\sin(\theta) \geq 0$. It is easy to see that

$$\Big|2(e^{i\theta} - 1)\Big|\,|\zeta - \zeta_0|^2 \leq t_A(z)^2.$$

Using this fact and the triangle inequality we have

$$\begin{aligned}
t_A(w)^2 &= \Big|2|\xi - \zeta_0|^2(e^{i\theta} - 1) + it\Big| \\
&\leq \Big|2|\zeta - \zeta_0|^2(e^{i\theta} - 1) + it\Big| + \Big|2\big(|\zeta - \zeta_0|^2 - |\xi - \zeta_0|^2\big)(e^{i\theta} - 1)\Big| \\
&= \Big|2|\zeta - \zeta_0|^2(e^{i\theta} - 1) + it\Big| \\
&\quad +2\Big||\zeta - \zeta_0| - |\xi - \zeta_0|\Big|\Big(|\zeta - \zeta_0| + |\xi - \zeta_0|\Big)|e^{i\theta} - 1| \\
&\leq \Big|2|\zeta - \zeta_0|^2(e^{i\theta} - 1) + it\Big| \\
&\quad +2|\xi - \zeta|\Big(2|\zeta - \zeta_0| + |\xi - \zeta|\Big)|e^{i\theta} - 1| \\
&\leq t_A(z)^2 + 2\Big|2|\zeta - \zeta_0|^2(e^{i\theta} - 1)\Big|^{1/2}\Big|2|\xi - \zeta|^2(e^{i\theta} - 1)\Big|^{1/2} \\
&\quad +\Big|2|\xi - \zeta|^2(e^{i\theta} - 1)\Big| \\
&\leq t_A(z)^2 + 2t_A(z)\Big|2(e^{i\theta} - 1)\Big|^{1/2}|\xi - \zeta| + \Big|2(e^{i\theta} - 1)\Big|\,|\xi - \zeta|^2 \\
&= \left(t_A(z) + \Big|2(e^{i\theta} - 1)\Big|^{1/2}|\xi - \zeta|\right)^2.
\end{aligned}$$

On the other hand, if $t\sin(\theta) < 0$ then we can find ζ so that $2|\zeta - \zeta_0|^2\sin(\theta) + t = 0$. Thus the initial estimate must be weakened to

$$|e^{i\theta} - 1|^2|\zeta - \zeta_0|^2 = 2(1 - \cos(\theta))\,|\zeta - \zeta_0|^2 \leq t_A(z)^2.$$

We can repeat the previous argument, but we must use the weaker estimate in the last inequality to get

$$\begin{aligned}
t_A(w)^2 &\leq t_A(z)^2 + 4|e^{i\theta} - 1|\,|\zeta - \zeta_0|\,|\xi - \zeta| + 2|e^{i\theta} - 1|\,|\xi - \zeta|^2 \\
&\leq t_A(z)^2 + 4t_A(z)\,|\xi - \zeta| + 2|e^{i\theta} - 1|\,|\xi - \zeta|^2 \\
&\leq \Big(t_A(z) + 2|\xi - \zeta|\Big)^2.
\end{aligned}$$

□

5 Subspaces of complex hyperbolic space

5.1 Geodesics

Consider a pair distinct null vectors $\mathbf{p}, \mathbf{q} \in V_0$. Without loss of generality normalise so that $\langle \mathbf{p}, \mathbf{q} \rangle = -1$. These vectors correspond to a pair of points p and q in $\partial\mathbf{H}^2_{\mathbb{C}}$. We want to describe the geodesic γ with endpoints p and q.

Proposition 5.1 *Let* $\mathbf{p},\,\mathbf{q}\in V_0$ *be null vectors with* $\langle\mathbf{p},\mathbf{q}\rangle=-1$. *For all real* t *let* $\gamma(t)$ *be the point in* $\mathbf{H}^2_{\mathbb{C}}$ *corresponding to the vector* $e^{t/2}\mathbf{p}+e^{-t/2}\mathbf{q}$ *in* $\mathbb{C}^{2,1}$. *Then* $\gamma=\big\{\gamma(t)\,|\,t\in\mathbb{R}\big\}$ *is the geodesic in* $\mathbf{H}^2_{\mathbb{C}}$ *with endpoints* p *and* q *parametrised by arc length* t.

Proof: First observe that $\gamma(t)$ is in $\mathbf{H}^2_{\mathbb{C}}$. This is because

$$\begin{aligned}\langle e^{t/2}\mathbf{p}+e^{-t/2}\mathbf{q},\,e^{t/2}\mathbf{p}+e^{-t/2}\mathbf{q}\rangle &= e^t\langle\mathbf{p},\mathbf{p}\rangle+\langle\mathbf{p},\mathbf{q}\rangle+\langle\mathbf{q},\mathbf{p}\rangle+e^{-t}\langle\mathbf{q},\mathbf{q}\rangle\\ &= -2.\end{aligned}$$

It suffices to show that $\rho\big(\gamma(t),\,\gamma(s)\big)=|t-s|$ for all real s and t.

$$\begin{aligned}&\cosh^2\Big(\frac{\rho\big(\gamma(t),\,\gamma(s)\big)}{2}\Big)\\ &= \frac{\langle e^{t/2}\mathbf{p}+e^{-t/2}\mathbf{q},\,e^{s/2}\mathbf{p}+e^{-s/2}\mathbf{q}\rangle\langle e^{s/2}\mathbf{p}+e^{-s/2}\mathbf{q},\,e^{t/2}\mathbf{p}+e^{-t/2}\mathbf{q}\rangle}{\langle e^{t/2}\mathbf{p}+e^{-t/2}\mathbf{q},\,e^{t/2}\mathbf{p}+e^{-t/2}\mathbf{q}\rangle\langle e^{s/2}\mathbf{p}+e^{-s/2}\mathbf{q},\,e^{s/2}\mathbf{p}+e^{-s/2}\mathbf{q}\rangle}\\ &= \left(\frac{-e^{(t-s)/2}-e^{(-t+s)/2}}{-2}\right)^2.\end{aligned}$$

This proves the result. $\square$

Any pair of points $z,\,w\in\mathbf{H}^2_{\mathbb{C}}$ lie on a unique geodesic. We now use the description of geodesics given above to find an expression for this geodesic.

Proposition 5.2 *Let* $\gamma(t)$ *be a geodesic in* $\mathbf{H}^2_{\mathbb{C}}$ *parametrised by arc length* t. *Suppose that* $\gamma(r)$ *and* $\gamma(s)$ *correspond to the points* $\mathbf{z}$ *and* $\mathbf{w}$ *in* V_- *where* $\langle\mathbf{z},\mathbf{z}\rangle=\langle\mathbf{w},\mathbf{w}\rangle=-2$ *and* $\langle\mathbf{z},\mathbf{w}\rangle$ *is real and negative. Then* $\gamma(t)$ *is given by the vector*

$$\frac{\sinh\big((t-s)/2\big)}{\sinh\big((r-s)/2\big)}\mathbf{z}+\frac{\sinh\big((r-t)/2\big)}{\sinh\big((r-s)/2\big)}\mathbf{w}.$$

Proof: Suppose that the endpoints of γ correspond to the null vectors $\mathbf{p}$ and $\mathbf{q}$ with $\langle\mathbf{p},\mathbf{q}\rangle=-1$. Then

$$\mathbf{z}=e^{r/2}\mathbf{p}+e^{-r/2}\mathbf{q},\quad \mathbf{w}=e^{s/2}\mathbf{p}+e^{-s/2}\mathbf{q}.$$

Then we see that

$$2\sinh\big((r-s)/2\big)\mathbf{p}=e^{-s/2}\mathbf{z}-e^{-r/2}\mathbf{w},\quad 2\sinh\big((r-s)/2\big)\mathbf{q}=-e^{s/2}\mathbf{z}+e^{r/2}\mathbf{w}.$$

The point $\gamma(t)$ then corresponds to

$$e^{t/2}\mathbf{p}+e^{-t/2}\mathbf{q}=\frac{\sinh\big((t-s)/2\big)}{\sinh\big((r-s)/2\big)}\mathbf{z}+\frac{\sinh\big((r-t)/2\big)}{\sinh\big((r-s)/2\big)}\mathbf{w}.$$

$\square$

We now find the height of a point on a geodesic neither of whose endpoints is ∞.

Proposition 5.3 *Let* p *and* q *be points of* $\partial\mathbf{H}^2_{\mathbb{C}}$ *neither of which is* ∞. *Let* u_{pq} *be the maximal height of a point on the geodesic with endpoints* p *and* q. *Then* $u_{pq}\le\rho_0(p,q)^2/2$.

Proof: Without loss of generality, we assume that in Heisenberg coordinates $p=(0,0)$ and $q=(\zeta,v)$. Therefore $\rho_0(p,q)=\left||\zeta|^2-iv\right|^{1/2}=\left(|\zeta|^4+v^2\right)^{1/4}$. We then lift p and q to vectors in $\mathbb{C}^{2,1}$ with $\langle \mathbf{p},\mathbf{q}\rangle=-1$. Thus

$$\mathbf{p}=\begin{bmatrix}0\\0\\1\end{bmatrix},\quad \mathbf{q}=\begin{bmatrix}-1\\ \sqrt{2}\zeta/\left(|\zeta|^2-iv\right)\\ 1/\left(|\zeta|^2-iv\right)\end{bmatrix}.$$

Using Proposition 5.1, an arbitrary point $\gamma(t)$ of the geodesic with endpoints p and q is given by

$$e^{t/2}\mathbf{p}+e^{-t/2}\mathbf{q}=\begin{bmatrix}-e^{-t/2}\\ \sqrt{2}e^{-t/2}\zeta/\left(|\zeta|^2-iv\right)\\ e^{t/2}+e^{-t/2}/\left(|\zeta|^2-iv\right)\end{bmatrix}.$$

In order to be able to use the Cygan metric, we must renormalise this vector so that its bottom entry is 1. This is:

$$\mathbf{z}_t=\begin{bmatrix}-\left(|\zeta|^2-iv\right)/\left(1+e^t\left(|\zeta|^2-iv\right)\right)\\ \sqrt{2}\zeta/\left(1+e^t\left(|\zeta|^2-iv\right)\right)\\ 1\end{bmatrix}.$$

From this we see that $-2u_t=\langle \mathbf{z}_t,\mathbf{z}_t\rangle$. That is

$$\begin{aligned}
-2u_t &= \frac{2|\zeta|^2}{\left|1+e^t\left(|\zeta|^2-iv\right)\right|^2}+\frac{-|\zeta|^2+iv}{1+e^t\left(|\zeta|^2+iv\right)}+\frac{-|\zeta|^2-iv}{1+e^t\left(|\zeta|^2-iv\right)}\\
&= -2e^t\left|\frac{|\zeta|^2-iv}{1+e^t\left(|\zeta|^2-iv\right)}\right|^2\\
&= -2\left(\frac{e^t\left(|\zeta|^4+v^2\right)}{1+2e^t|\zeta|^2+e^{2t}\left(|\zeta|^4+v^2\right)}\right).
\end{aligned}$$

Using elementary calculus, we see that u_t attains its maximum when $e^t=\left(|\zeta|^4+v^2\right)^{-1/2}$. Therefore the maximum height of a point on $\gamma(t)$ is

$$\begin{aligned}
u_{\max} &= \frac{\left(|\zeta|^4+v^2\right)^{1/2}}{1+2|\zeta|^2\left(|\zeta|^4+v^2\right)^{-1/2}+1}\\
&\le \frac{\left(|\zeta|^4+v^2\right)^{1/2}}{2}\\
&= \rho_0(p,q)^2/2.
\end{aligned}$$

□

5.2 Complex lines

Consider a complex line L in $\mathbb{C}^2$ that intersects the unit ball (which we think of as $\mathbf{H}^2_{\mathbb{C}}$). Let z be any point in $L\cap\mathbf{H}^2_{\mathbb{C}}$. We can apply an element of $\mathrm{PU}(2,1)$ to L so that it becomes the last coordinate axis $\{(0,\, z_2)|z_2 \in \mathbb{C}\}$. The intersection of this with the unit ball is the disc $|z_2| < 1$. We claim that the restriction of the Bergman metric to this disc is the Poincaré metric of constant curvature -1.

In order to see this, let $z = (0, z_2)$ and $w = (0, w_2)$ have lifts to $\mathbb{C}^{2,1}$ given by

$$\mathbf{z} = \begin{bmatrix} 0 \\ z_2 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 0 \\ w_2 \\ 1 \end{bmatrix}$$

The distance between these points is given by

$$\begin{aligned} \cosh^2\Big(\frac{\rho(z,w)}{2}\Big) &= \frac{\langle \mathbf{z}, \mathbf{w}\rangle_1 \langle \mathbf{w}, \mathbf{z}\rangle_1}{\langle \mathbf{z}, \mathbf{z}\rangle_1 \langle \mathbf{w}, \mathbf{w}\rangle_1} \\ &= \frac{|z_2\overline{w}_2 - 1|^2}{(|z_2|^2-1)(|w_2|^2-1)}. \end{aligned}$$

This is just the Poincaré metric (see page 132 of [2]).

In other words any complex line L is an embedded copy of $\mathbf{H}^1_{\mathbb{C}}$. This subgroup of $\mathrm{PU}(2,1)$ preserving this disc is the projectivisation of the block diagonal matrices

$$\begin{bmatrix} e^{i\theta} & 0 \\ 0 & A \end{bmatrix}$$

where $e^{i\theta} \in \mathrm{U}(1)$ acts on $L^{\perp}$ rotating $\mathbf{H}^2_{\mathbb{C}}$ around L and $A \in \mathrm{U}(1,1)$ is an isometry of the Poincaré metric on L acting by Möbius transformations. The group of such transformations is then $\mathrm{P}\big(\mathrm{U}(1)\times\mathrm{U}(1,1)\big) < \mathrm{PU}(2,1)$. Clearly this group is isomorphic to $\mathrm{U}(1,1)$. Any other complex line intersecting $\mathbf{H}^2_{\mathbb{C}}$ is the image of V under an element B of $\mathrm{PU}(2,1)$. The stabiliser of this complex line is the conjugate of $\mathrm{P}\big(\mathrm{U}(1)\times\mathrm{U}(1,1)\big)$ by B.

Taking the second Hermitian form and z lying in the subset given by $\{z = (z_1, 0)|z \in \mathbb{C}\}$ we see that z lies in $\mathbf{H}^2_{\mathbb{C}}$ if and only if $\langle \mathbf{z}, \mathbf{z}\rangle_2 = 2\Re(z_1) < 0$. This is a half-plane, in fact the left half plane. In order to get the more familiar upper half plane we write $z = (iz_1, 0)$. This point corresponds to a negative vector if and only if $\Im(z_1) > 0$. Lifting the points z and w to

$$\mathbf{z} = \begin{bmatrix} iz_1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} iw_1 \\ 0 \\ 1 \end{bmatrix}$$

we find that the distance function is given by The metric is given by the distance function

$$\begin{aligned} \cosh^2\Big(\frac{\rho(z,w)}{2}\Big) &= \frac{\langle \mathbf{z}, \mathbf{w}\rangle_2 \langle \mathbf{w}, \mathbf{z}\rangle_2}{\langle \mathbf{z}, \mathbf{z}\rangle_2 \langle \mathbf{w}, \mathbf{w}\rangle_2} \\ &= \frac{|z_1 - \overline{w_1}|^2}{2\Im(z_1)2\Im(w_1)}. \end{aligned}$$

This is just the Poincaré metric on the upper half plane (see Theorem 7.2.1(iv) of [2]).

5.3 Totally real Lagrangian planes

Now consider a totally real Lagrangian plane R. This may be characterised by $\langle \mathbf{v}, \mathbf{w} \rangle \in \mathbb{R}$ for all $\mathbf{v}$, $\mathbf{w} \in R$.

Any totally real Lagrangian plane R is the image under an element of $\mathrm{PU}(2,1)$ of the subspace comprising those points of $\mathbf{H}^2_{\mathbb{C}}$ with real coordinates, that is an embedded copy of real hyperbolic space $\mathbf{H}^2_{\mathbb{R}} = \big\{(x_1,\, x_2)\,|\,x_1,\, x_2 \in \mathbb{R}\big\}$. This subspace intersects the unit ball in the subset consisting of those points with ${x_1}^2 + {x_2}^2 < 1$. We claim that the Bergman metric restricted to this disc is just the Klein-Beltrami metric on the unit ball in $\mathbb{R}^2$ with constant curvature $-1/4$. To see this, write $x = (x_1,\, x_2)$ and $y = (y_1,\, y_2)$ in $\mathbf{H}^2_{\mathbb{R}}$. Lift x and y to column vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{C}^{2,1}$ as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ 1 \end{bmatrix}.$$

The we have

$$\begin{aligned} \cosh^2\Big(\frac{\rho(x,y)}{2}\Big) &= \frac{\langle \mathbf{x}, \mathbf{y} \rangle_1 \, \langle \mathbf{y}, \mathbf{x} \rangle_1}{\langle \mathbf{x}, \mathbf{x} \rangle_1 \, \langle \mathbf{y}, \mathbf{y} \rangle_1} \\ &= \frac{(x_1 y_1 + x_2 y_2 - 1)^2}{({x_1}^2 + {x_2}^2 - 1)({y_1}^2 + {y_2}^2 - 1)}. \end{aligned}$$

This is the Klein-Beltrami metric on the unit ball in $\mathbb{R}^2$ with constant curvature $-1/4$. It is more usual to define the Klein-Beltrami metric with curvature -1. In order to do this, replace $\rho(x,y)/2$ in this formula with $\rho(x,y)$ (see Chapter 3 of [25]).

Thus we obtain an embedded copy of $\mathbf{H}^2_{\mathbb{R}}$. The isometries preserving this space lie in the projectivisation of the natural inclusion $\mathrm{O}(2,1) < \mathrm{U}(2,1)$. It is also preserved by complex conjugation. Any other totally real Lagrangian plane is the image of this one under B in $\mathrm{PU}(2,1)$ and is stabilised by the conjugate by B of the projectivisation of $\mathrm{O}(2,1)$.

5.4 Totally geodesic subspaces

In this section we show that complex lines and totally real Lagrangian planes are totally geodesic. Together with geodesics, these are the only totally geodesic proper subspaces of $\mathbf{H}^2_{\mathbb{C}}$. We will not show the latter fact.

Proposition 5.4 *All complex lines L in $\mathbf{H}^2_{\mathbb{C}}$ are totally geodesic.*

Proof: Let L be a complex line. We need to show that for all choices of z and w in L the geodesic segment joining z to w lies in L. We may represent z and w by negative vectors with $\langle \mathbf{z}, \mathbf{z} \rangle = \langle \mathbf{w}, \mathbf{w} \rangle = -2$. From Proposition 5.2 we see that a general point on this geodesic segment is given by

$$\frac{\sinh\big((t-s)/2\big)}{\sinh\big((r-s)/2\big)}\mathbf{z} + \frac{\sinh\big((t-r)/2\big)}{\sinh\big((s-r)/2\big)}\mathbf{w}.$$

As this is a linear combination of $\mathbf{z}$ and $\mathbf{w}$ it corresponds to a point of L. $\square$

Proposition 5.5 *All totally real Lagrangian planes R in $\mathbf{H}^2_{\mathbb{C}}$ are totally geodesic.*

Proof: Let R be a totally real Lagrangian plane. We must show that for all choices of z and w in R the geodesic segment joining z to w lies in R. Totally real Lagrangian planes are characterised by $\langle \mathbf{v}, \mathbf{w}\rangle \in \mathbb{R}$ for all choices of $\mathbf{v}$ and $\mathbf{w}$ in R. Therefore we must show that, every point on the geodesic segment joining z and w corresponds to a vector whose Hermitian product with every point in R is real.

As before we may lift z and w in R to vectors with $\langle \mathbf{z}, \mathbf{z}\rangle = \langle \mathbf{w}, \mathbf{w}\rangle = -2$ and $\langle \mathbf{z}, \mathbf{w}\rangle \in \mathbb{R}$. Then a general point on the geodesic segment joining z and w is given by

$$\frac{\sinh\big((t-s)/2\big)}{\sinh\big((r-s)/2\big)}\mathbf{z} + \frac{\sinh\big((t-r)/2\big)}{\sinh\big((s-r)/2\big)}\mathbf{w}.$$

Let $\mathbf{v}$ be any vector corresponding to a point of R. Then $\langle \mathbf{v}, \mathbf{z}\rangle$ and $\langle \mathbf{v}, \mathbf{w}\rangle$ are both real. Thus

$$\begin{aligned}&\left\langle \mathbf{v}, \frac{\sinh\big((t-s)/2\big)}{\sinh\big((r-s)/2\big)}\mathbf{z} + \frac{\sinh\big((t-r)/2\big)}{\sinh\big((s-r)/2\big)}\mathbf{w}\right\rangle \\ &= \frac{\sinh\big((t-s)/2\big)}{\sinh\big((r-s)/2\big)}\langle \mathbf{v}, \mathbf{z}\rangle + \frac{\sinh\big((t-r)/2\big)}{\sinh\big((s-r)/2\big)}\langle \mathbf{v}, \mathbf{w}\rangle\end{aligned}$$

is real for all t. Hence the geodesic segment from z to w is in R. □

Alternatively, we could have used the following lemma together with the fact that a complex line is the fixed point set of a boundary elliptic isometry and that any totally real Lagrangian plane is fixed by an anti-holomorphic isometry conjugate to complex conjugation.

Lemma 5.6 *Any subset S of $\mathbf{H}^2_{\mathbb{C}}$ that is precisely the fixed point set of an isometry is totally geodesic.*

Proof: Consider two points z_0, $z_1 \in S$. By negative curvature there is a unique geodesic $\alpha : [0,1] \longrightarrow \mathbf{H}^2_{\mathbb{C}}$ joining $z_0 = \alpha(0)$ to $z_1 = \alpha(1)$. The distance of $z_t = \alpha(t)$ from z_0 is a monotone increasing function of t.

We need to show that the geodesic α lies in S. Suppose it does not. Then there is a point $z_t = \alpha(t)$ for some $t \in (0,1)$ on this geodesic not lying in S. By assumption, there exists an $A \in \mathrm{PU}(2,1)$ so that $A(z) = z$ if and only if $z \in S$. In particular $A(z_0) = z_0$, $A(z_1) = z_1$ but $A(z_t) \neq z_t$. Therefore $A(\alpha)$ is a geodesic joining z_0 to z_1. Now z_t is the unique point of α a distance $\rho(z_0, z_t)$ from z_0 and $A(z_t)$ is the unique point of $A(\alpha)$ a distance $\rho(z_0, z_t)$ from z_0. As these two points are different we see that $A(\alpha) \neq \alpha$. □

The following theorem follows using the general theory of symmetric spaces. Its proof is beyond the scope of these notes A sketch proof is given in Section 3.1.11 of Goldman [12].

Theorem 5.7 *All totally geodesic subspaces of $\mathbf{H}^2_{\mathbb{C}}$ are either complex linear or totally real.*

Corollary 5.8 *Every totally geodesic subspace of $\mathbf{H}^2_{\mathbb{C}}$ has real dimension which is either* 1 *or* 2*. In particular, there are no totally geodesic real hypersurfaces.*

This corollary means that there are no polyhedra in the standard meaning of the term. We therefore have to generalise the notion of polyhedra by defining suitable classes of real hypersurface for their boundaries.

We conclude this section by showing how totally real and complex linear subspaces fit together. The real line $\big\{(0,\, x)|x \in \mathbb{R}\big\}$ is a geodesic. It is contained in a unique complex line, namely $\{(0,\, z)|z \in \mathbb{C}\}$. It is also contained in a one parameter family of totally real subspaces, namely for each $\theta \in [0, \pi)$ there is a totally real subspace $\big\{(re^{i\theta},\, x)|r,\, x \in \mathbb{R}\big\}$.

5.5 Boundaries of totally geodesic subspaces

We now describe the intersection of complex lines and totally real planes with the boundary of the Siegel domain.

First consider a complex line L passing through the point at infinity. By applying a suitable Heisenberg translation, we may suppose that L also passes through the origin $o = (0, 0) \in \mathcal{N}$. In other words, L is the complex line spanned by

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

This complex line L consists of points

$$\begin{bmatrix} -u + iv \\ 0 \\ 1 \end{bmatrix}.$$

These points have horospherical coordinates $(0, v, u)$. Hence, L intersects the finite part of the boundary in the vertical axis $\big\{(0, v)|v \in \mathbb{R}\big\}$ of $\mathcal{N}$. By applying a Heisenberg translation, it is easy to see that any other complex line passing through infinity intersects the finite part of the boundary in the vertical line $\big\{(\zeta_0, v)|v \in \mathbb{R}\big\}$ for some fixed $\zeta_0 \in \mathbb{C}$. This is called an *infinite chain* or *infinite* $\mathbb{C}$*-circle*.

Now consider a complex line not passing through the point at infinity. The simplest example of such a line which intersects complex hyperbolic space is the line L spanned by

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

This complex line L consists of points

$$\begin{bmatrix} -1 \\ \sqrt{2}\zeta \\ 1 \end{bmatrix} = \begin{bmatrix} -|\zeta|^2 - (1 - |\zeta|^2) \\ \sqrt{2}\zeta \\ 1 \end{bmatrix}.$$

These points have horospherical coordinates $(\zeta, 0, 1-|\zeta|^2)$. Hence, L intersects the boundary of the Siegel domain in such points with $|\zeta| = 1$. In other words, ∂L is the circle $\{(e^{i\theta}, 0)|\theta \in [0, 2\pi)\}$. By applying a Heisenberg dilation we see that the circle $\{(r_0 e^{i\theta}, 0)|\theta \in [0, 2\pi)\}$ for any fixed $r_0 \in \mathbb{R}_+$ is also the boundary of a complex line. Now applying Heisenberg translation by (x_0+iy_0, v_0), we see that the most general complex line not passing through infinity intersects the boundary of the Siegel domain in the following ellipse whose vertical projection is a circle

$$\left\{ \left(r_0 e^{i\theta} + x_0 + iy_0, v_0 + 2r_0 y_0 \cos(\theta) - 2r_0 x_0 \sin(\theta) \right) | \theta \in [0, 2\pi) \right\}$$

for fixed $r_0 \in \mathbb{R}_+$ and $(x_0 + iy_0, v_0) \in \mathcal{N}$. Observe that the eccentricity of the ellipse increases with $|x_0 + iy_0|$. This is called a *finite chain* or *finite* $\mathbb{C}$*-circle*.

We do the same for totally real subspaces. First consider the totally real subspace R passing through o and ∞ which is fixed by complex conjugation. Hence R consists of vectors in $\mathbb{C}^{2,1}$ with real entries. Finite points in the Siegel domain with real entries have the form

$$\begin{bmatrix} -x^2 - u \\ \sqrt{2}x \\ 1 \end{bmatrix}$$

where $x \in \mathbb{R}$. These points have horospherical coordinates $(x, 0, u)$. Hence L intersects the boundary at ∞ and in the points

$$\begin{bmatrix} -x^2 \\ \sqrt{2}x \\ 1 \end{bmatrix}$$

where $x \in \mathbb{R}$. In other words ∂R is the x axis of the Heisenberg group, that is the subset of the Heisenberg group given by $\{(x, 0)|x \in \mathbb{R}\}$. By applying Heisenberg rotations we see that for any fixed $\theta_0 \in [0, 2\pi)$ the line $\{(xe^{i\theta_0}, 0)|x \in \mathbb{R}\}$ is also the finite part of the boundary of a totally real plane containing o and ∞. In particular, the y axis of the Heisenberg group is such a line. By applying Heisenberg translation by $(x_0 + iy_0, v_0)$ we find the general form for the boundary of a totally real plane passing through ∞. It is

$$\left\{ \left(xe^{i\theta_0} + x_0 + iy_0, v_0 + 2xy_0 \cos(\theta_0) - 2xx_0 \sin(\theta_0) \right) | x \in \mathbb{R} \right\}$$

for fixed $\theta_0 \in [0, 2\pi)$ and $(x_0 + iy_0, v_0) \in \mathcal{N}$. Observe that the gradient of the line is proportional to its distance from o. Such boundaries of totally real Lagrangian planes containing ∞ are called *infinite* $\mathbb{R}$*-circles*.

We now do the same for the boundaries of totally real subspaces not passing through ∞. This is more complicated. We begin with the boundary of the totally real subspace R fixed by the following involution

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \longmapsto \begin{bmatrix} \overline{z}_3 \\ \overline{z}_2 \\ \overline{z}_1 \end{bmatrix}.$$

Points in the boundary of the Siegel domain fixed by this involution have the form

$$\begin{bmatrix} -e^{2i\theta} \\ i\sqrt{2\cos(2\theta)}e^{i\theta} \\ 1 \end{bmatrix} \qquad \text{where } \theta \in [-\pi/4, \pi/4) \cup (3\pi/4, 5\pi/4].$$

The values of θ are chosen to make $\cos(2\theta)$ non-negative. In other words the subset of the Heisenberg group given by

$$\left\{ \left(i\sqrt{\cos(2\theta)}e^{i\theta}, -\sin(2\theta) \right) | \theta \in [-\pi/4, \pi/4) \cup (3\pi/4, 5\pi/4] \right\}.$$

This is an example of a *finite* $\mathbb{R}$-*circle*, and is called the *standard finite* $\mathbb{R}$-*circle*.

This $\mathbb{R}$-circle R is a non-planar space curve and we now discuss it slightly more carefully. In order to simplify notation, define

$$p(\theta) = \left(i\sqrt{\cos(2\theta)}e^{i\theta}, -\sin(2\theta) \right).$$

Observe that R is connected in spite of the fact that the values of the parameter θ are contained in two disjoint intervals. To see this, observe that $p(-\pi/4) = p(3\pi/4) = (0, 1)$ and $p(\pi/4) = p(5\pi/4) = (0, -1)$. Alternatively, we can see that R is homoeomorphic to a circle using the following re-parametrisation of R:

$$\begin{bmatrix} -(1 + i\cos\phi)/(1 - i\cos\phi) \\ \sqrt{2}i\sin\phi/(1 - i\cos\phi) \\ 1 \end{bmatrix} = \begin{bmatrix} -(\sin^2\phi - 2i\cos\phi)/(1 + \cos^2\phi) \\ \sqrt{2}i\sin\phi(1 + i\cos\phi)/(1 + \cos^2\phi) \\ 1 \end{bmatrix}$$

where $\phi \in [0, 2\pi)$.

The vertical projection of R is a plane curve given by the parametric equation

$$i\sqrt{\cos(2\theta)}e^{i\theta} \qquad \text{where } \theta \in [-\pi/4, \pi/4) \cup (3\pi/4, 5\pi/4].$$

In Cartesian coordinates $x = -\sin(\theta)\sqrt{\cos(2\theta)}$ and $y = \cos(\theta)\sqrt{\cos(2\theta)}$ we have

$$x^2 + y^2 = \cos(2\theta), \qquad x^2 - y^2 = -\cos^2(2\theta).$$

Thus the equation of this curve is

$$(x^2 + y^2)^2 + x^2 - y^2 = 0.$$

This is a lemniscate of Bernoulli, see Chapter 12 of Lockwood [21].

We rewrite $v = -\sin(2\theta)$ as

$$v = -\sin(2\theta) = -\frac{\sin(\theta)}{\cos(\theta)}\left(\cos^2(\theta) - \sin^2(\theta) + 1\right) = \frac{x}{y}(x^2 + y^2 + 1).$$

Observe that as x and y tend to 0 they do so along the lines $x = y$ and $x = -y$. Therefore when $x = y = 0$ we have $v = \pm 1$. Thus the $\mathbb{R}$-circle is a space curve given points $(x + iy, v) \in \mathcal{N}$ satisfying the equations

$$(x^2 + y^2)^2 + x^2 - y^2 = 0, \quad v = \frac{x}{y}(x^2 + y^2 + 1) \quad \text{for } y^2 > x^2 > 0$$

together with the points $(0, \pm 1)$.

From our parametrisation we have

$$v^2 = 1 - \cos^2(2\theta), \quad 2x^2 = \cos(2\theta)\big(1 - \cos(2\theta)\big), \quad 2y^2 = \cos(2\theta)\big(1 + \cos(2\theta)\big).$$

Therefore the projection of this $\mathbb{R}$-circle onto the (x, v) plane is given by all real solutions to the equation

$$(v^2 - 2x^2)^2 = v^2 - 4x^2.$$

Similarly the projection onto the (y, v) plane is given by all real solutions other than $(0, 0)$ to the equation

$$(v^2 + 2y^2)^2 = v^2 + 4y^2.$$

In order to obtain all *finite* $\mathbb{R}$-*circles* we need to apply Heisenberg rotations, dilations and translations to this one.

6 Classification of isometries

6.1 Eigenvalues and eigenspaces

In this section we will classify holomorphic isometries. The familiar trichotomy from real hyperbolic geometry applies in the complex hyperbolic setting as well: A holomorphic complex hyperbolic isometry A is said to be:

(i) *loxodromic* if it fixes exactly two points of $\partial\mathbf{H}^2_{\mathbb{C}}$;

(ii) *parabolic* if it fixes exactly one point of $\partial\mathbf{H}^2_{\mathbb{C}}$;

(iii) *elliptic* if it fixes at least one point of $\mathbf{H}^2_{\mathbb{C}}$.

We will analyse each of these types in more detail. As we do so, this coarse classification will be refined. We will give three different approaches to classifying isometries. First, we will discuss an algebraic approach by considering eigenvectors and eigenvalues of matrices in $\mathrm{SU}(2,1)$. Secondly, we will discuss a dynamical approach by looking at fixed points and subsets that are preserved. Finally, (in Section 6.3) we will discuss a geometrical approach by considering products of involutions in totally real Lagrangian planes.

First we explain the classification in terms of the corresponding matrices in $\mathrm{SU}(2,1)$. It is clear that a fixed point of an isometry A lying in $\mathbf{H}^2_{\mathbb{C}}$ or its boundary corresponds to an eigenvector of the corresponding matrix lying in V_- or V_0 respectively. The goal of the first part of this section is to prove the following theorem, which verifies that the trichotomy above exhausts all possibilities:

Theorem 6.1 *Let A be a matrix in* $\mathrm{SU}(2,1)$. *Then one of the following possibilities occurs:*

(i) A has two null eigenvectors with eigenvalues λ and $\overline{\lambda}^{-1}$ where $|\lambda| \neq 1$. In this case A is loxodromic.

(ii) *A has a repeated eigenvalue of unit modulus whose eigenspace is either spanned by a null vector or it is spanned by vectors a null vector $\mathbf{v}$ and a positive vector $\mathbf{w}$ with the property that $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. In this case A is parabolic.*

(iii) *A has a negative eigenvector. In this case A is elliptic.*

We will prove Theorem 6.1 by way of a series of lemmas. First we investigate some general properties of eigenvalues of matrices in $\mathrm{SU}(2,1)$.

Lemma 6.2 *Let $A \in \mathrm{SU}(2,1)$ 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 J. Hence, $A^*JA = J$ and so $A = J^{-1}(A^*)^{-1}J$. 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. □

Corollary 6.3 *If λ is an eigenvalue of $A \in \mathrm{SU}(2,1)$ with $|\lambda| \neq 1$ then $\overline{\lambda}^{-1}$ is a distinct eigenvalue. In particular, either A has all three eigenvalues of absolute value 1 or else A has a pair of eigenvalues λ and $\overline{\lambda}^{-1}$ with $|\lambda| \neq 1$ and the third eigenvalue $\overline{\lambda}\lambda^{-1}$ of absolute value 1.*

Next we show that any eigenvalue not of unit modulus corresponds to a null eigenvector and that any eigenvectors that are not (Hermitian) orthogonal have eigenvalues λ and $\mu = \overline{\lambda}^{-1}$.

Lemma 6.4 *Let λ, μ be eigenvalues of $A \in \mathrm{SU}(2,1)$ 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 $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. □

The following lemma is an easy consequence of the signature of $\mathbb{C}^{2,1}$.

Lemma 6.5 *If $\mathbf{v}, \mathbf{w} \in \mathbb{C}^{2,1} - \{0\}$ with $\langle \mathbf{v}, \mathbf{v} \rangle \leq 0$ and $\langle \mathbf{w}, \mathbf{w} \rangle \leq 0$ then either $\mathbf{w} = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{C}$ or $\langle \mathbf{v}, \mathbf{w} \rangle \neq 0$.*

Proof: Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ be the standard basis vectors for $\mathbb{C}^{2,1}$ with the first Hermitian form. That is

$$\begin{aligned} 1 &= \langle \mathbf{e}_1, \mathbf{e}_1 \rangle_1 = \langle \mathbf{e}_2, \mathbf{e}_2 \rangle_1 = -\langle \mathbf{e}_3, \mathbf{e}_3 \rangle_1, \\ 0 &= \langle \mathbf{e}_1, \mathbf{e}_2 \rangle_1 = \langle \mathbf{e}_2, \mathbf{e}_3 \rangle_1 = \langle \mathbf{e}_3, \mathbf{e}_1 \rangle_1. \end{aligned}$$

Write

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3, \quad \mathbf{w} = w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + w_3\mathbf{e}_3.$$

Because $\langle \mathbf{v}, \mathbf{v} \rangle_1 \leq 0$ and $\langle \mathbf{w}, \mathbf{w} \rangle_1 \leq 0$ we have

$$|v_1|^2 + |v_2|^2 \leq |v_3|^2, \quad |w_1|^2 + |w_2|^2 \leq |w_3|^2.$$

Therefore v_3 and w_3 are both non-zero. Suppose that $\langle \mathbf{v}, \mathbf{w} \rangle_1 = 0$. This is equivalent to

$$v_1\overline{w}_1 + v_2\overline{w}_2 = v_3\overline{w}_3.$$

Hence, for all $\lambda \in \mathbb{C}$ we have

$$\begin{aligned} |v_3 - \lambda w_3|^2 &= |v_3|^2 - \lambda w_3\overline{v}_3 - \overline{\lambda} v_3\overline{w}_3 + |\lambda|^2|w_3|^2 \\ &\geq |v_1|^2 - \lambda w_1\overline{v}_1 - \overline{\lambda} v_1\overline{w}_1 + |\lambda|^2|w_1|^2 \\ &\quad +|v_2|^2 - \lambda w_2\overline{v}_2 - \overline{\lambda} v_2\overline{w}_2 + |\lambda|^2|w_2|^2 \\ &= |v_1 - \lambda w_1|^2 + |v_2 - \lambda w_2|^2. \end{aligned}$$

Choose $\lambda = -v_3/w_3$. The left hand side of this inequality is then zero. Since the right hand side is non-negative, it too must be zero. This means that $v_1 - w_1v_3/w_3 = v_2 - w_2v_3/w_3 = 0$. In other words, $v_1/w_1 = v_2/w_2 = v_3/w_3 = \lambda$ and so $\mathbf{v} = \lambda\mathbf{w}$. □

Using these lemmas we can analyse the eigenvectors of any A in $\mathrm{SU}(2,1)$ with an eigenvalue not of unit modulus. This will prove Theorem 6.1 (i).

Lemma 6.6 *Suppose that the eigenvalues of $A \in \mathrm{SU}(2,1)$ are $re^{i\theta}$, $r^{-1}e^{i\theta}$, $e^{-2i\theta}$ where $r \neq 1$ and they have eigenvectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ respectively then*

$$\langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle = 0, \quad \langle \mathbf{w}, \mathbf{w} \rangle > 0, \quad \langle \mathbf{u}, \mathbf{v} \rangle \neq 0.$$

Proof: We know that $\langle \mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle = 0$ by Lemma 6.4 (i) as $r \neq 1$. We also know that $\langle \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle = 0$ by part Lemma 6.4 (ii) as $re^{3i\theta} \neq 1$ and $r^{-1}e^{3i\theta} \neq 1$. Using the Lemma 6.5 we see that $\langle \mathbf{u}, \mathbf{v} \rangle \neq 0$. If $\langle \mathbf{w}, \mathbf{w} \rangle \leq 0$ then, by Lemma 6.5 we would have $\langle \mathbf{u}, \mathbf{w} \rangle \neq 0$ which is a contradiction. □

From now on we may assume that all the eigenvalues of A have unit modulus. We begin with the case where they are all distinct. We show that such a matrix satisfies the condition of Theorem 6.1 (iii).

Lemma 6.7 *Suppose that the eigenvalues of $A \in \mathrm{SU}(2,1)$ are distinct with absolute value 1, and suppose they have eigenvectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ respectively then*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{u} \rangle = 0$$

and two of $\langle \mathbf{u}, \mathbf{u} \rangle$, $\langle \mathbf{v}, \mathbf{v} \rangle$, $\langle \mathbf{w}, \mathbf{w} \rangle$ are positive while the other is negative.

Proof: Since the eigenvalues are distinct we have $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{u} \rangle = 0$, from Lemma 6.4 (ii) Now using Lemma 6.5, this implies that at most one of $\langle \mathbf{v}, \mathbf{v} \rangle$, $\langle \mathbf{v}, \mathbf{v} \rangle$, $\langle \mathbf{w}, \mathbf{w} \rangle$ is non-positive. Since $\langle \cdot, \cdot \rangle$ is non-degenerate and indefinite none of $\langle \mathbf{u}, \mathbf{u} \rangle$, $\langle \mathbf{v}, \mathbf{v} \rangle$, $\langle \mathbf{w}, \mathbf{w} \rangle$ are zero and at least one of them is negative. □

We need to consider the case of repeated eigenvalues. We begin by supposing the eigenvalues are $e^{i\theta}$, $e^{i\theta}$ and $e^{-2i\theta}$ where $e^{i\theta} \neq e^{-2i\theta}$.

Lemma 6.8 *Suppose that $A \in \mathrm{SU}(2,1)$ has a repeated eigenvalue $e^{i\theta}$. Suppose that the $e^{i\theta}$-eigenspace is spanned by $\mathbf{v}$, then $\langle \mathbf{v}, \mathbf{v} \rangle = 0$.*

Proof: Since $e^{i\theta}$ is a repeated eigenvalue, there exists a vector $\mathbf{u}$ that is not a multiple of $\mathbf{v}$ and which satisfies $A\mathbf{u} = e^{i\theta}\mathbf{u} + \mathbf{v}$. (To see this, put A into Jordan normal form.) Then

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle A\mathbf{u}, A\mathbf{v} \rangle = \langle e^{i\theta}\mathbf{u} + \mathbf{v}, e^{i\theta}\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + e^{-i\theta}\langle \mathbf{v}, \mathbf{v} \rangle.$$

This implies that $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ as claimed. □

Lemma 6.9 *Suppose that $A \in \mathrm{SU}(2,1)$ has eigenvalues $e^{i\theta}$, $e^{i\theta}$ and $e^{-2i\theta}$. Let $\mathbf{v}$ be an eigenvector corresponding to $e^{i\theta}$ and $\mathbf{w}$ be the eigenvector corresponding to $e^{-2i\theta}$. Suppose that $\langle \mathbf{w}, \mathbf{w} \rangle \leq 0$. Then in fact $\langle \mathbf{w}, \mathbf{w} \rangle < 0$, $\langle \mathbf{v}, \mathbf{w} \rangle = 0$, $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ and the $e^{i\theta}$-eigenspace is two dimensional.*

Proof: As $\mathbf{v}$ and $\mathbf{w}$ correspond to distinct eigenvalues of unit modulus then, by Lemma 6.4 (ii) $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. If $\langle \mathbf{v}, \mathbf{v} \rangle \leq 0$ then by Lemma 6.5 $\langle \mathbf{v}, \mathbf{w} \rangle \neq 0$, a contradiction. Thus $\langle \mathbf{v}, \mathbf{v} \rangle > 0$. Using Lemma 6.8, this means that the $e^{i\theta}$-eigenspace cannot be spanned by $\mathbf{v}$. Therefore there are orthogonal $e^{i\theta}$-eigenvectors $\mathbf{v}$ and $\mathbf{u}$ with $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle > 0$. Since $\langle \cdot, \cdot \rangle$ is indefinite and non-degenerate we have $\langle \mathbf{w}, \mathbf{w} \rangle < 0$. □

Lemma 6.10 *Suppose that $A \in \mathrm{SU}(2,1)$ has eigenvalues $e^{i\theta}$, $e^{i\theta}$ and $e^{-2i\theta}$. Let $\mathbf{v}$ be an eigenvector corresponding to $e^{i\theta}$ and $\mathbf{w}$ be an eigenvector corresponding to $e^{-2i\theta}$. Suppose that $\langle \mathbf{w}, \mathbf{w} \rangle > 0$. Then either*

(i) the $e^{i\theta}$-eigenspace is spanned by $\mathbf{v}$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$, or

(ii) the $e^{i\theta}$-eigenspace is two dimensional and indefinite.

Proof: If the $e^{i\theta}$-eigenspace is one dimensional, say it is spanned by $\mathbf{v}$, then, by Lemma 6.8 we have $\langle \mathbf{v}, \mathbf{v} \rangle = 0$.

On the other hand, suppose that there exist linearly independent $e^{i\theta}$-eigenvectors $\mathbf{v}$ and $\mathbf{u}$. Without loss of generality we assume $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. We know that $\langle \mathbf{u}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle = 0$ as $e^{i\theta}\overline{e^{-2i\theta}} = e^{3i\theta} \neq 1$. Therefore $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are linearly independent and hence form a basis for $\mathbb{C}^{2,1}$. Suppose that $\mathbf{z}$ is a negative vector in $\mathbb{C}^{2,1}$ and write $\mathbf{z} = \alpha\mathbf{u} + \beta\mathbf{v} + \gamma\mathbf{w}$. Then

$$0 > \langle \mathbf{z}, \mathbf{z} \rangle = |\alpha|^2 \langle \mathbf{u}, \mathbf{u} \rangle + |\beta|^2 \langle \mathbf{v}, \mathbf{v} \rangle + |\gamma|^2 \langle \mathbf{w}, \mathbf{w} \rangle .$$

As $\langle \mathbf{w}, \mathbf{w} \rangle > 0$, at least one of $\langle \mathbf{u}, \mathbf{u} \rangle$ or $\langle \mathbf{v}, \mathbf{v} \rangle$ is negative. As the Hermtian form is non-degenerate and has signature $(2,1)$ one of them is negative and the other positive. (This statement relies on the fact that we have chosen $\mathbf{u}$ and $\mathbf{v}$ so that $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.) This gives the result. $\square$

Combining these lemmas we see that if A has two distinct eigenvalues, one of which is repeated, then A either satisfies the condition of Theorem 6.1 (ii) or Theorem 6.1 (iii).

Finally we need to consider the case where all three eigenvalues are the same, necessarily a cube root of unity. We will show that A satisfies the conditions of Theorem 6.1 (ii). So this will complete the proof of Theorem 6.1.

Lemma 6.11 *Suppose that $A \in \mathrm{SU}(2,1)$ has exactly one eigenvalue. Then this eigenvalue has modulus 1. Then $|\lambda| = 1$ and one of the following is true:*

(i) A is a multiple of the identity,

(ii) the eigenspace of A is spanned by a null vector, or

(iii) the eigenspace of A is spanned by vectors a null vector $\mathbf{v}$ and a positive vector $\mathbf{w}$ with the property that $\langle \mathbf{v}, \mathbf{w}\rangle = 0$.

Proof: Let λ be the eigenvalue. Since A is unitary we must have $1 = \big|\det(A)\big| = |\lambda^3|$. So we write $\lambda = e^{i\theta}$.

The eigenspace of A has dimension 1, 2 or 3. If it has dimension 3 then, necessarily, A is a multiple of the identity. If it has dimension 1 then, by Lemma 6.8, this eigenspace is spanned by a null vector.

Thus we only have to consider the case where the eigenspace has dimension 2. By examining the Jordan normal form we can find an eigenvector $\mathbf{v}$ and a vector $\mathbf{u}$ so that

$$A\mathbf{u} = e^{i\theta}\mathbf{u} + \mathbf{v}$$

Let $\mathbf{w}$ be any eigenvector. Then, arguing as in Lemma 6.8, we see that

$$\langle \mathbf{u}, \mathbf{w}\rangle = \langle A\mathbf{u}, A\mathbf{w}\rangle = \langle e^{i\theta}\mathbf{u} + \mathbf{v}, e^{i\theta}\mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + e^{-i\theta}\langle \mathbf{v}, \mathbf{w}\rangle.$$

This implies that $\langle \mathbf{v}, \mathbf{w}\rangle = 0$. In particular, $\langle \mathbf{v}, \mathbf{v}\rangle = 0$. Let $\{\mathbf{v}, \mathbf{w}\}$ be a basis for the eigenspace. Then $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a basis for $\mathbb{C}^{2,1}$. Let H be the matrix of the Hermitian form with respect to this basis. It is not hard to see that since $\langle \mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle = 0$ we have

$$\det(H) = -\langle \mathbf{v}, \mathbf{u}\rangle \langle \mathbf{u}, \mathbf{v}\rangle \langle \mathbf{w}, \mathbf{w}\rangle = -\big|\langle \mathbf{u}, \mathbf{v}\rangle\big|^2 \langle \mathbf{w}, \mathbf{w}\rangle .$$

Since H has signature $(2,1)$ it has negative determinant, and so $\langle \mathbf{w}, \mathbf{w}\rangle > 0$ as claimed. $\square$

6.2 Isometry classification and trace

We now show that we can use the trace of $A \in \mathrm{SU}(2,1)$ to decide the class of A. From Corollary 6.3 we see that, if λ_1, λ_2 and λ_3 are eigenvalues of A, then $\overline{\lambda_1}^{-1}$, $\overline{\lambda_2}^{-1}$ and $\overline{\lambda_3}^{-1}$ form some permutation of λ_1, λ_2 and λ_3. Let $\chi_A(x)$ be the characteristic polynomial of A. Suppose that

$$\chi_A(x) = x^3 - a_2x^2 + a_1x - a_0.$$

Then $a_2 = \lambda_1 + \lambda_2 + \lambda_3 = \mathrm{tr}(A)$ and $a_0 = \lambda_1\lambda_2\lambda_3 = \det(A) = 1$. The other coefficient is

$$\begin{aligned} a_1 &= \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1 \\ &= {\lambda_3}^{-1} + {\lambda_1}^{-1} + {\lambda_2}^{-1} \\ &= \overline{\lambda_1} + \overline{\lambda_2} + \overline{\lambda_3} \\ &= \overline{\mathrm{tr}(A)}. \end{aligned}$$

We denote the trace of A by $\mathrm{tr}(A) = \tau$. Thus we have

$$\chi_A(x) = x^3 - \tau x^2 + \overline{\tau}x - 1.$$

We want to find out when $A \in \mathrm{SU}(2,1)$ has repeated eigenvalues. In other words, we want to find conditions on τ for which $\chi_A(x) = 0$ has repeated solutions. This is true if and only if $\chi_A(x)$ and its derivative $\chi'_A(x)$ have a common root. Two polynomials have a common root if and only if their resultant vanishes (see Kirwan [18]). Now

$$\chi'_A(x) = 3x^2 - 2\tau x + \overline{\tau}.$$

Therefore a brief calculation shows that the resultant of $\chi_A(x)$ and $\chi'_A(x)$ is given by:

$$R(\chi_A, \chi'_A) = \begin{vmatrix} 1 & -\tau & \overline{\tau} & -1 & 0 \\ 0 & 1 & -\tau & \overline{\tau} & -1 \\ 3 & -2\tau & \overline{\tau} & 0 & 0 \\ 0 & 3 & -2\tau & \overline{\tau} & 0 \\ 0 & 0 & 3 & -2\tau & \overline{\tau} \end{vmatrix} = |\tau|^4 - 8\Re(\tau^3) + 18|\tau|^2 - 27.$$

Theorem 6.12 *Let $f(\tau) = |\tau|^4 - 8\Re(\tau^3) + 18|\tau|^2 - 27$. Let $A \in \mathrm{SU}(2,1)$ then:*

(i) A has an eigenvalue λ with $|\lambda| \neq 1$ if and only if $f\big(\mathrm{tr}(A)\big) > 0$,

(ii) A has a repeated eigenvalue if and only if $f\big(\mathrm{tr}(A)\big) = 0$,

(iii) A has distinct eigenvalues of unit modulus if and only if $f\big(\mathrm{tr}(A)\big) < 0$.

It is easy to see that part (ii) of Theorem 6.12 follows from the reasoning given above. We now look at the other cases separately.

Lemma 6.13 *Suppose that $A \in \mathrm{SU}(2,1)$ has an eigenvalue λ with $|\lambda| \neq 1$. Then $f\big(\mathrm{tr}(A)\big) > 0$.*

Proof: Suppose that $re^{i\theta}$ is an eigenvalue of A where r is positive and $r \neq 1$. Then by Corollary 6.3 we see that $\overline{re^{i\theta}}^{-1} = r^{-1}e^{i\theta}$ is also an eigenvalue. Since the determinant of A is 1, we see that the third eigenvalue is $e^{-2i\theta}$. Therefore

$$\tau = re^{i\theta} + r^{-1}e^{i\theta} + e^{-2i\theta}.$$

Hence

$$\begin{aligned} |\tau|^2 &= (r + r^{-1})^2 + 2(r + r^{-1})\cos(3\theta) + 1, \\ \Re(\tau^3) &= (r + r^{-1})^3 + 3(r + r^{-1})^2 + 3(r + r^{-1})\cos(3\theta) + \cos(6\theta). \end{aligned}$$

From this it is east to see that

$$f\big(re^{i\theta}+r^{-1}e^{i\theta}+e^{-2i\theta}\big)=(r-r^{-1})^2\big(r+r^{-1}-2\cos(3\theta)\big)^2>0.$$

□

Lemma 6.14 *Suppose that $A\in\mathrm{SU}(2,1)$ has three distinct eigenvalues, all of unit modulus. Then $f\big(\mathrm{tr}(A)\big)>0$.*

Proof: We write the eigenvalues as $e^{i\theta}$, $e^{i\phi}$, $e^{i\psi}$ where θ, ϕ and ψ are distinct and $e^{i\theta+i\phi+i\psi}=1$. Then

$$\tau=e^{i\theta}+e^{i\phi}+e^{i\psi}$$

and

$$\begin{aligned}|\tau|^2 &= 3+2\cos(\theta-\phi)+2\cos(\phi-\psi)+2\cos(\psi-\theta),\\ \Re(\tau^3) &= \cos(3\theta)+\cos(3\phi)+\cos(3\psi)+6\cos(\theta-\phi)+6\cos(\phi-\psi)+6\cos(\psi-\theta)+6.\end{aligned}$$

But,

$$\cos(3\theta)=\cos(2\theta-\phi-\psi)=\cos(\theta-\phi)\cos(\phi-\psi)+\sin(\theta-\phi)\sin(\phi-\psi).$$

Hence

$$\begin{aligned}\Re(\tau^3) &= \cos(\theta-\phi)\cos(\phi-\psi)+\cos(\phi-\psi)\cos(\psi-\theta)+\cos(\psi-\theta)\cos(\theta-\phi)\\ &\quad+\sin(\theta-\phi)\sin(\phi-\psi)+\sin(\phi-\psi)\sin(\psi-\theta)+\sin(\psi-\theta)\sin(\theta-\phi)\\ &\quad+6\cos(\theta-\phi)+6\cos(\phi-\psi)+6\cos(\psi-\theta)+6.\end{aligned}$$

Using this we calculate

$$f\big(e^{i\theta}+e^{i\phi}+e^{i\psi}\big)=-4\Big(\sin(\theta-\phi)+\sin(\phi-\psi)+\sin(\phi-\psi)\Big)^2<0.$$

□

As these two lemmas exhaust all the possibilities, we have proved Theorem 6.12.

The curve $f(\tau)=0$ is a classical curve called a *deltoid*, see Chapter 8 of Lockwood [21] or page 26 of Kirwan [18] where it is written in terms of $x=\Re(\tau)$ and $y=\Im(\tau)$. The points outside correspond to case (i) in the theorem. This may be seen by considering A with eigenvalues r, r^{-1} and 1 which lie in the interval $(3,\infty)$ and those with eigenvalues $e^{i\theta}$, $e^{-i\theta}$ and 1 which lie in $(-1,3)$. The rest follows by continuity.

Lemma 6.15 *Suppose that A is an elliptic element of $\mathrm{SU}(2,1)$ with real trace, that is $\mathrm{tr}(A)\in[-1,3)$. Then the eigenvalues of A are 1, $e^{i\theta}$ and $e^{-i\theta}$ where $2\cos(\theta)=\mathrm{tr}(A)-1$.*

Proof: If the eigenvalues of A are $e^{i\theta}$, $e^{i\phi}$ and $e^{-i\theta-i\phi}$ then

$$0=\Im\big(\mathrm{tr}(A)\big)=\sin(\theta)+\sin(\phi)-\sin(\theta+\phi)=4\sin(\theta/2)\sin(\phi/2)\sin(\theta/2+\phi/2)$$

and so at least one of θ, ϕ or $\theta+\phi$ is an integer multiple of 2π. Hence A has eigenvalue $+1$. The rest of the lemma follows easily. □

Lemma 6.16 *If $A \in \mathrm{SU}(2,1)$ has trace $\mathrm{tr}(A) = 3$ then A is unipotent. That is, each of the eigenvalues of A is 1.*

Proof: Since
$$f(3) = 81 - 216 + 162 - 27 = 0,$$
using Theorem 6.12 (ii) we see that A has a repeated eigenvalue, say λ. Since $\det(A) = 1$ the eigenvalues of A must by λ, λ and $1/\lambda^2$. Hence $3 = \mathrm{tr}(A) = 2\lambda + 1/\lambda^2$. Solving the resulting cubic we see that $\lambda = 1$ or $\lambda = -1/2$. In the latter case we get a contradiction from Lemma 6.2. □

We now sum up the previous results about eigenvectors and eigenvalues and interpret them in terms of information about fixed points.

First, consider a loxodromic map A. We know that A corresponds to a matrix with eigenvectors $re^{i\theta}$, $r^{-1}e^{i\theta}$, $e^{-2i\theta}$, where $r > 1$, with eigenvectors $\mathbf{p}, \mathbf{q} \in V_0$ and $\mathbf{n} \in V_+$ respectively. These correspond to an attractive fixed point p and a repulsive fixed point q on $\partial\mathbf{H}^2_{\mathbb{C}}$. The complex line L spanned by $\mathbf{p}$ and $\mathbf{q}$ is preserved by A. This line has polar vector $\mathbf{n}$.

Next, consider a parabolic map A. Such a map corresponds to a matrix with a repeated eigenvalue of unit modulus whose eigenspace is spanned by a null eigenvector $\mathbf{p}$. This vector corresponds to a neutral fixed point p on $\partial\mathbf{H}^2_{\mathbb{C}}$. There are two cases to consider, namely when A has a single eigenvalue of multiplicity 3 and when A has two distinct eigenvalues, one of which is repeated. In the first case we say that A is *pure parabolic*. Later we shall see that pure parabolic maps correspond to Heisenberg translations. A pure parabolic map has trace 3 or $(-3 \pm 3i\sqrt{3})/2$, that is it corresponds to one of the three corners of the deltoid. In the second case we say that A is *screw parabolic*. In this case the non-repeated eigenvalue has an eigenvector $\mathbf{n}$ in V_+. The complex line polar to $\mathbf{n}$ is preserved by A, and A acts as a translation there. Moreover, A rotates $\mathbf{H}^2_{\mathbb{C}}$ around this complex line. Screw parabolic maps correspond to smooth points of the deltoid.

Finally, consider an elliptic map A. There are now three cases. First, suppose that A has a repeated eigenvalue with a two dimensional eigenspace containing both positive and negative vectors. This eigenspace corresponds to a complex line L on which A acts as the identity. In particular, there are points of $\partial\mathbf{H}^2_{\mathbb{C}}$ fixed by A and so A is called *boundary elliptic*. As A fixes L and rotates $\mathbf{H}^2_{\mathbb{C}}$ around L, it is *complex reflection in the line L*. If A is not boundary elliptic then it has an eigenspace spanned by a negative vector $\mathbf{w}$. This corresponds to a fixed point $w \in \mathbf{H}^2_{\mathbb{C}}$. In this case A is called *regular elliptic*. There are two possibilities. Either A has a repeated eigenvalue with an eigenspace spanned by two positive vectors. In this case A is *complex reflection in the point w*. Otherwise, A has three distinct eigenvalues. Complex reflections again correspond to smooth points of the deltoid while other elliptic maps correspond to points of the deltoid's interior.

Exercise 6.17 *Let A be any element of $\mathrm{SU}(2,1)$ and let $\tau = \mathrm{tr}(A)$. Let $f(\tau)$ be the function given in Theorem 6.12.*

(i) Show that $\mathrm{tr}(A^2) = \tau^2 - 2\overline{\tau}$ and
$$f(\tau^2 - 2\overline{\tau}) = \left(|\tau|^2 - 1\right)^2 f(\tau).$$

(ii) Show that $\mathrm{tr}(A^3) = \tau^3 - 3|\tau|^2 + 3$ *and*

$$f\big(\tau^3 - 3|\tau|^2 + 3\big) = \big(|\tau|^4 - \tau^3 - \overline{\tau}^3\big)^2 f(\tau).$$

Interpret the above formulae in terms of eigenvalues.

6.3 Isometries as products of involutions

Inversion in a totally real Lagrangian plane is an anti-holomorphic involution. A product of two of these inversions is holomorphic and so is necessarily in $\mathrm{PU}(2,1)$. The theorem below shows, first, that all elements of $\mathrm{PU}(2,1)$ may be written as the product of inversions in two $\mathbb{R}$-circles and, secondly, they may be classified as loxodromic, parabolic, boundary elliptic or regular elliptic by the intersection and linking properties of these $\mathbb{R}$-circles.

Theorem 6.18 *Any* $A \in \mathrm{PU}(2,1)$ *may be decomposed as the product of a pair of inversions in totally real Lagrangian planes. Moreover, if these totally real Lagrangian planes have boundary* $\mathbb{R}$*-circles* R_1 *and* R_2 *then:*

(i) if R_1 *and* R_2 *are disjoint and unlinked then* A *is loxodromic,*

(ii) if R_1 *and* R_2 *are disjoint and linked then* A *is regular elliptic,*

(iii) if R_1 *and* R_2 *intersect in exactly one point then* A *is parabolic,*

(iv) if R_1 *and* R_2 *intersect in two points then* A *is boundary elliptic.*

Every complex hyperbolic isometry is either holomorphic, and so is given by a matrix in $\mathrm{PU}(2,1)$, or anti-holomorphic, so is an inversion in a totally real Lagrangian subspace followed by an element of $\mathrm{PU}(2,1)$. Thus we have the following corollary.

Corollary 6.19 *Any complex hyperbolic isometry may be written as a product of at most three inversions in totally real Lagrangian planes.*

We prove Theorem 6.18 by conjugating A to a normalised form and then showing such an A may be written as a product of inversions in totally real Lagrangian planes. We do this case by case. It will follow from our reasoning that the $\mathbb{R}$-circles have the required properties. As the cases exhaust all possibilities (except for $R_1 = R_2$ which gives $A = I$) this proves the result. For simplicity we work on the boundary $\partial\mathbf{H}^2_{\mathbb{C}} = \mathcal{N} \cup \{\infty\}$.

Lemma 6.20 *Suppose that* $A : (\zeta, v) \longmapsto (\lambda\zeta, |\lambda|^2 v)$ *with* $|\lambda| \neq 1$ *is a loxodromic map fixing* o *and* ∞*. Then* A *may be written as* $A = \iota_2\iota_1$ *where* ι_j *is inversion in the* $\mathbb{R}$*-circle* R_j*. Here* R_1 *is the standard imaginary* $\mathbb{R}$*-circle and* R_2 *is the image of the standard imaginary* $\mathbb{R}$*-circle under* $B : (\zeta, v) \longmapsto (\lambda^{1/2}\zeta, |\lambda| v)$*. These two* $\mathbb{R}$*-circles are disjoint and unlinked.*

Proof: Inversion in the standard imaginary $\mathbb{R}$-circle R_1 and the loxodromic map B are given by

$$\iota_1 \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} \overline{z}_3 \\ \overline{z}_2 \\ \overline{z}_1 \end{bmatrix}, \quad B \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} \overline{\lambda}^{1/2} z_1 \\ z_2 \\ \lambda^{-1/2} z_3 \end{bmatrix}.$$

Inversion in $R_2 = B(R_1)$ is given by

$$\iota_2 \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = B\iota_1 B^{-1} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = B\iota_1 \begin{bmatrix} \overline{\lambda}^{-1/2} z_1 \\ z_2 \\ \lambda^{1/2} z_3 \end{bmatrix} = B \begin{bmatrix} \overline{\lambda}^{1/2} \overline{z}_3 \\ \overline{z}_2 \\ \lambda^{-1/2} \overline{z}_1 \end{bmatrix} = \begin{bmatrix} \overline{\lambda} \overline{z}_3 \\ \overline{z}_2 \\ \lambda^{-1} \overline{z}_1 \end{bmatrix}.$$

Therefore

$$\iota_2\iota_1 \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \iota_2 \begin{bmatrix} \overline{z}_3 \\ \overline{z}_2 \\ \overline{z}_1 \end{bmatrix} = \begin{bmatrix} \overline{\lambda} z_1 \\ z_2 \\ \lambda^{-1} z_3 \end{bmatrix} = A \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix},$$

using Lemma 4.4. Thus $A = \iota_2\iota_1$ as required. Observe R_1 is on the unit Cygan sphere centred at o and R_2 is on the Cygan sphere of radius $|\lambda|^{1/2} \neq 1$ centred at o. These two spheres are disjoint and nested. Hence the two $\mathbb{R}$-circles are disjoint and unlinked. $\square$

Lemma 6.21 *Suppose that $A : (\zeta, v) \longmapsto \big(\zeta + \tau, v + t - 2\Im(\overline{\tau} z)\big)$ is Heisenberg translation by $(\tau, t) \in \mathcal{N} - \{o\}$ fixing ∞. Then A may be written as $\iota_2\iota_1$ where ι_j is inversion in the $\mathbb{R}$-circle R_j. Here R_1 is the infinite $\mathbb{R}$-circle given by $R_1 = \big\{(\zeta, v) = (ki\tau, 0) : k \in \mathbb{R}\big\}$ and R_2 is the image of R_1 under the map $B : (\zeta, v) \longmapsto \big(\zeta + \tau/2, v + t/2 - \Im(\overline{\tau}\zeta)\big)$. These two $\mathbb{R}$-circles only intersect in the point ∞.*

Proof: Without loss of generality, assume that $\tau \in \mathbb{R}$. Inversion in R_1 is given by $\iota_1(\zeta, v) = (-\overline{\zeta}, -v)$. Then $R_2 = B(R_1)$ is given by

$$R_2 = \big\{\big((ki + 1/2)\tau, -k|\tau|^2 + t/2\big) : k \in \mathbb{R}\big\}.$$

If $\tau \neq 0$ then the ζ coordinates of R_1 and R_2 are distinct and if $\tau = 0$ the v coordinates are distinct. Hence they only intersect at ∞. Inversion in R_2 is given by

$$\begin{aligned} \iota_2(\zeta, v) &= B\iota_1 B^{-1}(\zeta, v) \\ &= B\iota_1\big(\zeta - \tau/2, v - t/2 + \tau\Im(\zeta)\big) \\ &= B\big(-\overline{\zeta} + \tau/2, -v + t/2 - \tau\Im(\zeta)\big) \\ &= \big(-\overline{\zeta} + \tau, -v + t - 2\tau\Im(\zeta)\big). \end{aligned}$$

Thus $\iota_2\iota_1(\zeta, v) = \big(\zeta + \tau, v + t - 2\tau\Im(\zeta)\big) = A(\zeta, v)$. Hence $A = \iota_2\iota_1$ as claimed. $\square$

Lemma 6.22 *Suppose that $A : (\zeta, v) \longmapsto \big(e^{i\theta}\zeta, v + t)\big)$ where $\theta \in (0, 2\pi)$ is screw parabolic, for $t \neq 0$, and boundary elliptic, for $t = 0$. Then A may be written as $\iota_2\iota_1$ where ι_j is inversion in the $\mathbb{R}$-circle R_j. Here R_1 is the x-axis in the Heisenberg group, that is $R_1 = \big\{(z, v) = (x, 0) : x \in \mathbb{R}\big\}$, and R_2 is the image of R_1 under the map $B : (\zeta, v) \longmapsto (e^{i\theta/2}\zeta, v + t/2)$. If $t \neq 0$ then these two $\mathbb{R}$-circles only intersect in the point ∞. If $t = 0$ then the two $\mathbb{R}$-circles intersect at o and ∞.*

Proof: This is very similar to the previous lemmas. We have seen earlier that inversion in R_1 is complex conjugation. In horospherical coordinates this inversion is given by

$\iota_1(\zeta,v,u)=(\overline{\zeta},-v,u)$. Hence

$$\begin{aligned}\iota_2(\zeta,v) &= B\iota_1B^{-1}(\zeta,v)\\ &= B\iota_1(e^{-i\theta/2}\zeta,v-t/2)\\ &= B(e^{i\theta/2}\overline{\zeta},-v+t/2)\\ &= (e^{i\theta}\overline{\zeta},-v+t).\end{aligned}$$

As before it is easy to see that $A=\iota_2\iota_1$ and to find the intersections of the two $\mathbb{R}$-circles. □

It remains to consider the case of regular elliptic maps. We begin by using the ball model. We know that a regular elliptic map may be written in the form

$$A=\begin{bmatrix} e^{i\theta} & 0 & 0\\ 0 & e^{i\phi} & 0\\ 0 & 0 & 1\end{bmatrix}$$

where θ, $\phi\in(0,2\pi)$. The map A can be decomposed as $A=\iota_2\iota_1$ where

$$\iota_1:\begin{bmatrix} z_1\\ z_2\\ z_3\end{bmatrix}\longmapsto\begin{bmatrix}\overline{z}_1\\ e^{-i\phi}\overline{z}_2\\ \overline{z}_3\end{bmatrix},\qquad \iota_2:\begin{bmatrix} z_1\\ z_2\\ z_3\end{bmatrix}\longmapsto\begin{bmatrix}e^{i\theta}\overline{z}_1\\ \overline{z}_2\\ \overline{z}_3\end{bmatrix}$$

are inversions in the totally real planes whose boundaries are the $\mathbb{R}$-circles

$$\begin{aligned} R_1 &= \Big\{(z_1,z_2)=\big(\cos(\psi),e^{-i\phi/2}\sin(\psi)\big)\ :\ \psi\in[0,2\pi)\Big\},\\ R_2 &= \Big\{(z_1,z_2)=\big(e^{i\theta/2}\cos(\psi),\sin(\psi)\big)\ :\ \psi\in[0,2\pi)\Big\}.\end{aligned}$$

Since $\theta/2$ and $\phi/2$ lie in the interval $(0,\pi)$, it is clear that these two $\mathbb{R}$-circles are disjoint.

In order to be able to continue our analysis of the boundary using the Siegel domain model we conjugate A, ι_1 and ι_2 be the Cayley transform (6), but still use the same names. These maps become

$$\iota_1:\begin{bmatrix} z_1\\ z_2\\ z_3\end{bmatrix}\longmapsto\begin{bmatrix}\overline{z}_1\\ e^{-i\phi}\overline{z}_2\\ \overline{z}_3\end{bmatrix},\qquad \iota_2:\begin{bmatrix} z_1\\ z_2\\ z_3\end{bmatrix}\longmapsto\begin{bmatrix}\cos(\theta/2)e^{i\theta/2}\overline{z}_1+i\sin(\theta/2)e^{i\theta/2}\overline{z}_3\\ \overline{z}_2\\ i\sin(\theta/2)e^{i\theta/2}\overline{z}_1+\cos(\theta/2)e^{i\theta/2}\overline{z}_3\end{bmatrix}$$

and

$$A=\begin{bmatrix}\cos(\theta/2)e^{i\theta/2} & 0 & i\sin(\theta/2)e^{i\theta/2}\\ 0 & e^{i\phi} & 0\\ i\sin(\theta/2)e^{i\theta/2} & 0 & \cos(\theta/2)e^{i\theta/2}\end{bmatrix}. \tag{64}$$

The $\mathbb{R}$ circles R_j fixed by the involutions ι_j become

$$R_1=\{(x+iy,v)\ :\ \sin(\phi/2)x+\cos(\phi/2)y=v=0\} \tag{65}$$

and

$$R_2=\left\{(x+iy,v)\ :\ \begin{matrix}\sin(\theta/2)\big((x^2+y^2)^2+y^2-x^2\big)-2\cos^2(\theta/2)xy=0\\ vx+y(x^2+y^2+1)=0\end{matrix}\right\} \tag{66}$$

It is clear that R_2 intersects the plane defined by $v = 0$ in the points where $y = 0$ and hence $x = \pm 1$. The $\mathbb{R}$-circle R_1 is a line through the origin and gradient $-\tan(\phi/2) \neq 0$ in this plane. Each of the halfplanes determined by this line contains one of the points $(\pm 1, 0)$. Thus the two $\mathbb{R}$-circles may easily be seen to be disjoint and linked. Thus we have proved:

Lemma 6.23 *Suppose that A be a regular elliptic map of the form (64). Then A may be written as $A = \iota_2\iota_1$ where ι_j is inversion in the $\mathbb{R}$-circles R_j given in (65) and (66). These two $\mathbb{R}$-circles are disjoint and linked.*

We now show that there is a certain amount of flexibility in choosing the anti-holomorphic involutions ι_j.

Lemma 6.24 *Let $A \in \mathrm{PU}(2,1)$ be elliptic and let $\mathbf{z}$ be any point in $\mathbb{C}^{2,1}$. Then there is an anti-holomorphic involution ι so that $A\iota$ is an involution and ι fixes $\mathbf{z}$.*

Proof: Without loss of generality assume that A preserves the first Hermitian form and is given by

$$A = \begin{bmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{i\phi} & 0 \\ 0 & 0 & e^{i\psi} \end{bmatrix}.$$

For any angles α, β, γ define the antiholomrphic involution ι by

$$\iota \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} e^{i\alpha}\overline{z}_1 \\ e^{i\beta}\overline{z}_2 \\ e^{i\gamma}\overline{z}_3 \end{bmatrix}.$$

Then clearly $A\iota$ has order two.

If $\mathbf{z}$ is a given point of $\mathbb{C}^{2,1}$ then choosing $\alpha = 2\arg(z_1)$, $\beta = 2\arg(z_2)$ and $\gamma = 2\arg(z_3)$ we see that ι fixes $\mathbf{z}$. $\square$

7 Distance formulae

7.1 Cross ratios

Cross-ratios were generalised to complex hyperbolic space by Korányi and Reimann [20]. Following their notation, we suppose that z_1, z_2, w_1, w_2 are four distinct points of $\overline{\mathbf{H}^2_{\mathbb{C}}}$, and we define their *complex cross-ratio* to be

$$[z_1, z_2, w_1, w_2] = \frac{\langle \mathbf{w}_1, \mathbf{z}_1 \rangle \langle \mathbf{w}_2, \mathbf{z}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{z}_1 \rangle \langle \mathbf{w}_1, \mathbf{z}_2 \rangle}$$

We will only use the absolute value $\big|[z_1, z_2, w_1, w_2]\big|$ which we call the *cross-ratio*. Observe that if two of the entries are the same then the cross ratio is still defined and equals one of 0, 1 or ∞. If z_1, z_2, w_1 and w_2 all lie on $\partial\mathbf{H}^2_{\mathbb{C}}$ then we can express the cross ratio in terms of the Cygan metric as follows:

$$\left|[z_1, z_2, w_1, w_2]\right| = \frac{\rho_0(w_1, z_1)^2 \rho_0(w_2, z_2)^2}{\rho_0(w_2, z_1)^2 \rho_0(w_1, z_2)^2}$$

provided none of the four points is ∞. If $w_1 = \infty$ then

$$\left|[z_1, z_2, \infty, w_2]\right| = \frac{\rho_0(w_2, z_2)^2}{\rho_0(w_2, z_1)^2}.$$

7.2 Distance between a point and a geodesic

Let p and q be points of $\partial\mathbf{H}^2_{\mathbb{C}}$ and let z be any point of $\mathbf{H}^2_{\mathbb{C}}$. We choose lifts $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{z}$ in $\mathbb{C}^{2,1}$ of of p, q and z.

Following Goldman, we define

$$\eta(p, q, z) = [z, q, p, z] = \frac{\langle \mathbf{p}, \mathbf{z}\rangle \langle \mathbf{z}, \mathbf{q}\rangle}{\langle \mathbf{z}, \mathbf{z}\rangle \langle \mathbf{p}, \mathbf{q}\rangle}.$$

Then we have

Proposition 7.1 *Let p and q be points of $\partial\mathbf{H}^2_{\mathbb{C}}$ and let z be any point of $\mathbf{H}^2_{\mathbb{C}}$. Let γ be the geodesic with endpoints p and q. The distance $\rho(\gamma, z)$ from z to γ is given by*

$$\cosh^2\left(\frac{\rho(\gamma, z)}{2}\right) = \left|\eta(p, q, z)\right| + \Re\big(\eta(p, q, z)\big).$$

Proof: Without loss of generality, normalise so that $\langle \mathbf{p}, \mathbf{q}\rangle = -1$. This means that a general point on the lift of γ to $\mathbb{C}^{2,1}$ is given by $e^{t/2}\mathbf{p} + e^{-t/2}\mathbf{q}$. The corresponding point in $\mathbf{H}^2_{\mathbb{C}}$ will be denoted by $\gamma(t)$. Therefore

$$\begin{aligned}
\cosh^2\left(\frac{\rho(\gamma(t), z)}{2}\right) &= \frac{\left|\langle e^{t/2}\mathbf{p} + e^{-t/2}\mathbf{q}, \mathbf{z}\rangle\right|^2}{\langle e^{t/2}\mathbf{p} + e^{-t/2}\mathbf{q}, e^{t/2}\mathbf{p} + e^{-t/2}\mathbf{q}\rangle \langle \mathbf{z}, \mathbf{z}\rangle} \\
&= \frac{e^t\left|\langle \mathbf{p}, \mathbf{z}\rangle\right|^2 + 2\Re\big(\langle \mathbf{p}, \mathbf{z}\rangle\langle \mathbf{z}, \mathbf{q}\rangle\big) + e^{-t}\left|\langle \mathbf{q}, \mathbf{z}\rangle\right|^2}{2\langle \mathbf{p}, \mathbf{q}\rangle\langle \mathbf{z}, \mathbf{z}\rangle}.
\end{aligned}$$

Using elementary calculus we see that, as t varies over $\mathbb{R}$, the minimum of the right hand side is attained when

$$e^t = \left|\frac{\langle \mathbf{q}, \mathbf{z}\rangle}{\langle \mathbf{p}, \mathbf{z}\rangle}\right|.$$

Substituting this in the above expression we see that

$$\cosh^2\left(\frac{\rho(\gamma, z)}{2}\right) = \frac{\left|\langle \mathbf{p}, \mathbf{z}\rangle\langle \mathbf{z}, \mathbf{q}\rangle\right| + \Re\big(\langle \mathbf{p}, \mathbf{z}\rangle\langle \mathbf{z}, \mathbf{q}\rangle\big)}{\langle \mathbf{p}, \mathbf{q}\rangle\langle \mathbf{z}, \mathbf{z}\rangle}.$$

As the denominator is real and positive this proves the result. □

We observe that if z is on γ, then $\mathbf{z} = e^{s/2}\mathbf{p} + e^{-s/2}\mathbf{q}$ and so $\eta(p, q, z) = 1/2$.

7.3 Distance between pairs of geodesics

For $j = 1, 2$ let p_j and q_j be points of $\partial\mathbf{H}^2_{\mathbb{C}}$ with lifts $\mathbf{p}_j$ and $\mathbf{q}_j$ in $\mathbb{C}^{2,1}$. Normalise these lifts so that $\langle \mathbf{p}_j, \mathbf{q}_j \rangle = -1$. We want to investigate the distance between the geodesics γ_1 and γ_2 where γ_j has endpoints $\mathbf{p_j}$ and $\mathbf{q}_j$. These geodesics are given by

$$\gamma_1 = \left\{ e^{t/2}\mathbf{p}_1 + e^{-t/2}\mathbf{q}_1 : t \in \mathbb{R} \right\}, \quad \gamma_2 = \left\{ e^{s/2}\mathbf{p}_2 + e^{-s/2}\mathbf{q}_2 : t \in \mathbb{R} \right\}.$$

We want to find the distance between γ_1 and γ_2.

First we show that it is possible to find pairs of geodesics so that, as a function of the Hermitian products of $\mathbf{p}_j$ and $\mathbf{q}_j$, the distance between them cannot be expressed using radicals. This method and example is due to Hanna Sandler [26]. Suppose that $\gamma_1(t)$ is the point on γ_1 with lift $e^{t/2}\mathbf{p}_1 + e^{-t/2}\mathbf{q}_1$. Using Proposition 7.1 we see that

$$\cosh^2\left(\frac{\rho\big(\gamma_1(t), \gamma_2\big)}{2}\right) = \Big|\eta\big(p_2, q_2, \gamma_1(t)\big)\Big| + \Re\Big(\eta\big(p_2, q_2, \gamma_1(t)\big)\Big).$$

We now express this as a function of t and the inner products of the $\mathbf{p}_j$ and $\mathbf{q}_j$.

$$\begin{aligned} \eta\big(p_2, q_2, \gamma_1(t)\big) &= \frac{\langle \mathbf{p_2}, e^{t/2}\mathbf{p}_1 + e^{-t/2}\mathbf{q}_1 \rangle \langle e^{t/2}\mathbf{p}_1 + e^{-t/2}\mathbf{q}_1, \mathbf{q}_2 \rangle}{\langle e^{t/2}\mathbf{p}_1 + e^{-t/2}\mathbf{q}_1, e^{t/2}\mathbf{p}_1 + e^{-t/2}\mathbf{q}_1 \rangle \langle \mathbf{p}_2, \mathbf{q}_2 \rangle} \\ &= \frac{1}{2}\Big(\langle \mathbf{p}_2, \mathbf{p}_1 \rangle \langle \mathbf{p}_1, \mathbf{q}_2 \rangle e^t + \langle \mathbf{p}_2, \mathbf{q}_1 \rangle \langle \mathbf{p}_1, \mathbf{q}_2 \rangle \Big) \\ &\quad + \frac{1}{2}\Big(\langle \mathbf{p}_2, \mathbf{p}_1 \rangle \langle \mathbf{q}_1, \mathbf{q}_2 \rangle + \langle \mathbf{p}_2, \mathbf{q}_1 \rangle \langle \mathbf{q}_1, \mathbf{q}_2 \rangle e^{-t} \Big) \\ &= \frac{1}{2}\Big(ae^t + b + ce^{-t} \Big) \end{aligned}$$

where

$$\begin{aligned} a &= \langle \mathbf{p}_2, \mathbf{p}_1 \rangle \langle \mathbf{p}_1, \mathbf{q}_2 \rangle, \\ b &= \langle \mathbf{p}_2, \mathbf{q}_1 \rangle \langle \mathbf{p}_1, \mathbf{q}_2 \rangle + \langle \mathbf{p}_2, \mathbf{p}_1 \rangle \langle \mathbf{q}_1, \mathbf{q}_2 \rangle, \\ c &= \langle \mathbf{p}_2, \mathbf{q}_1 \rangle \langle \mathbf{q}_1, \mathbf{q}_2 \rangle. \end{aligned}$$

Suppose that $x(t)$ and $y(t)$ are the real and imaginary parts of $\eta(p_2, q_2, \gamma_1(t))$. Then we need to find the minimum of the function

$$g(t) = \sqrt{x(t)^2 + y(t)^2} + x(t).$$

Differentiating and setting $g'(t) = 0$, we need to solve

$$0 = \frac{x(t)x'(t) + y(t)y'(t)}{\sqrt{x(t)^2 + y(t)^2}} + x'(t).$$

This is equivalent to

$$x(t)x'(t) + y(t)y'(t) = -x'(t)\sqrt{x(t)^2 + y(t)^2}.$$

Squaring both sides and simplifying we see that

$$0 = y(t)\Big(y(t)x'(t)^2 - y(t)y'(t)^2 - 2x(t)y(t)x'(t)y'(t) \Big).$$

If $y(t)=0$ then $g'(t)=2x'(t)$. Thus a minimum occurs when $y(t)=x'(t)=0$ and we have $y(t)x'(t)^2-y(t)y'(t)^2-2x(t)y(t)x'(t)y'(t)=0$ as well. Thus it suffices to solve

$$\begin{aligned} 0 &= y(t)x'(t)^2-y(t)y'(t)^2-2x(t)y(t)x'(t)y'(t) \\ &= \Im\Big(\eta\big(p_2,q_2,\gamma_1(t)\big)\overline{\eta'\big(p_2,q_2,\gamma_1(t)\big)}^2\Big). \end{aligned}$$

Since $\eta(p_2,q_2,\gamma_1(t))=(ae^t+b+c^{-t})/2$ we have $\eta'(p_2,q_2,\gamma_1(t))=(ae^t-ce^{-t})/2$. Thus,

$$\begin{aligned} & 8\eta\big(p_2,q_2,\gamma_1(t)\big)\overline{\eta'\big(p_2,q_2,\gamma_1(t)\big)}^2 \\ &= \big(ae^t+b+ce^{-t}\big)\big(\overline{a}^2e^{2t}-2\overline{a}\,\overline{c}+\overline{c}^2e^{-2t}\big) \\ &= a|a|^2e^{3t}+\overline{a}^2be^{2t}+\big(\overline{a}^2c-2|a|^2\overline{c}\big)e^t-2\overline{a}b\overline{c} \\ &\qquad +\big(a\overline{c}^2-2\overline{a}|c|^2\big)e^{-t}+b\overline{c}^2e^{-2t}+\overline{c}|c|^2e^{-3t}. \end{aligned}$$

Multiplying by e^{3t} and taking the imaginary part, we see that finding the shortest distance between γ_1 and γ_2 is equivalent to solving a sixth order polynomial in e^t. The coefficients of this polynomial may be expressed in terms of the Hermitian products of $\mathbf{p}_j$ and $\mathbf{q}_j$. The following example shows that there exist geodesics for which this sixth order polynomial is not solvable by radicals in terms of its coefficients.

Consider the following null vectors (with respect to the second Hermitian form):

$$\mathbf{p}_1=\begin{bmatrix}-1\\0\\0\end{bmatrix},\quad \mathbf{q}_1=\begin{bmatrix}-(1+4i)/4\\-(1+i)/2\\1\end{bmatrix},\quad \mathbf{p}_2=\begin{bmatrix}-1/2\\1/\sqrt{2}\\1/2\end{bmatrix},\quad \mathbf{q}_2=\begin{bmatrix}-1/2\\-1/\sqrt{2}\\1/2\end{bmatrix}.$$

These have been normalised so that $\langle\mathbf{p}_1,\mathbf{q}_1\rangle=\langle\mathbf{p}_2,\mathbf{q}_2\rangle=-1$. A short calculation shows that

$$a=\frac{1}{4},\qquad b=\frac{5-2i\sqrt{2}}{8},\qquad c=\frac{25-4i\sqrt{2}}{64}.$$

From this, we see that the minimum distance occurs when t satisfies

$$\begin{aligned} 0 &= 2e^{3t}\Im\Big(\eta\big(p_2,q_2,\gamma_1(t)\big)\overline{\eta'\big(p_2,q_2,\gamma_1(t)\big)}^2\Big) \\ &= \Im\Big(a|a|^2e^{6t}+\overline{a}^2be^{5t}+\big(\overline{a}^2c-2|a|^2\overline{c}\big)e^{4t}-2\overline{a}b\overline{c}e^{3t}\Big) \\ &\quad +\Im\Big(\big(a\overline{c}^2-2\overline{a}|c|^2\big)e^{2t}+b\overline{c}^2e^t+\overline{c}|c|^2\Big) \\ &= \frac{-\sqrt{2}}{16}\Big((4e^t)^5+3(4e^t)^4-30(4e^t)^3-50(4e^t)^2+93(4e^t)-657\Big). \end{aligned}$$

Thus writing $x=4e^t$ we need to find the roots of

$$g(x)=x^5+3x^4-30x^3-50x^2+93x-657.$$

Evaluating g at $x=-6,-5,-4,5,6$ we see that $g(x)$ has three real roots x_1, x_2, x_3 satisfying $-6<x_1<-5<x_2<-4$ and $5<x_3<6$. We claim that $g(x)$ has no more real roots. In order to see this, consider

$$g'(x)=5x^4+12x^3-90x^2-100x+93.$$

Evaluating g' at $x = -6, -5, 0, 1, 4$ we see that $g'(x)$ has roots x_4, x_5, x_6, x_7 with

$$-6 < x_4 < -5 < x_5 < 0 < x_6 < 1 < x_7 < 4.$$

Moreover, x_6 must be a local maximum of $g(x)$. However, when $0 < x < 1$ we have

$$g(x) < 1 + 3 + 93 - 657 < -560.$$

Therefore $g(x_6) < 0$ and so $g(x)$ has a local maximum on which it takes a negative value. Hence, g cannot have the maximum number of real zeros and so has a pair of conjugate complex roots. Finally, we claim that $g(x)$ is irreducible over the integers and hence over the rationals. First we transform $g(x)$ to

$$h(x) = g(2x - 5)/32 = x^5 - 11x^4 + 40x^3 - 50x^2 - 2x + 4.$$

Evaluating $h(x)$ at $-4, -2, -1, 1, 2, 4$ we see that $h(x)$ has no linear factors. Now suppose that

$$h(x) = (x^2 + ax + b)(x^3 + cx^2 + dx + e)$$

for integers a, b, c, d, e. Then b must be one of ± 1, ± 2, ± 4. Expanding and equating coefficients and simplifying in each case we arrive at a contradiction.

(i) Suppose $b = 1$ and so $e = 4$. Evaluating the coefficients of x and x^2 we have $-2 = ae + bd = 4a + d$ and $-50 = e + ad + bc = 4 + a(-2 - 4a) + c$. Moreover, evaluating the coefficient of x^3 and substituting for b, c and d gives

$$40 = d + ac + b = -2 - 4a + a(-50 - 4 + 2a(1 + 2a)) + 1.$$

The right hand side of this is odd but the left is even, a contradiction.

(ii) Suppose $b = 2$ and $e = 2$. Then $-2 = 2a + 2d$ and $-11 = a + c$. Hence $d = -1 - a$ and $c = -11 - a$. Also, $-50 = e + ad + bc = 2 - a - a^2 - 22 - 2a$ and $40 = d + ac + b = -1 - a - 11a - a^2 + 2$. Combining these gives $a^2 = 30 - 3a = -39 - 12a$ which means $23 = 3a$, a contradiction.

(ii) Suppose that $b = 4$ and $e = 1$. Then $-2 = ae + bd = a + 4d$. Substituting for a gives $-50 = e + ad + bc = 1 - 2d - 4d^2 + 4c$. This is a contradiction.

Similar arguments work when $b = -1$, -2 and -4.

Putting this together, we see that $g(x)$ is irreducible over $\mathbb{Z}$, and hence over $\mathbb{Q}$, and has exactly three real roots. Thus, using the same argument as Lemma 14.7 of [28] we see that $g(x)$ is not soluble by radicals. This argument says that the Galois group of $g(x)$ must contain a 5-cycle and an involution (complex conjugation). Hence this Galois group is S_5.

We conclude this section by giving a lower bound on the distance between geodesics. As at the start of this section, for $j = 1, 2$, let p_j, q_j be points on the boundary of $\mathbf{H}^2_{\mathbb{C}}$ with lifts $\mathbf{p}_j$ and $\mathbf{q}_j$ to $\mathbb{C}^{2,1}$ satisfying $\langle \mathbf{p}_j, \mathbf{q}_j \rangle = -1$. Let $\mathbf{n}_j$ be the unit polar vector to the complex line spanned by $\mathbf{p}_j$ and $\mathbf{q}_j$. The following lemma follows from Lemma 3.4:

Lemma 7.2 *With the definitions above we have*

$$\begin{aligned}
0 &= \langle \mathbf{p}_1,\mathbf{p}_1\rangle = \langle \mathbf{p}_1,\mathbf{n}_2\rangle\langle \mathbf{n}_2,\mathbf{p}_1\rangle - \langle \mathbf{p}_1,\mathbf{q}_2\rangle\langle \mathbf{p}_2,\mathbf{p}_1\rangle - \langle \mathbf{p}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_2,\mathbf{p}_1\rangle,\\
-1 &= \langle \mathbf{p}_1,\mathbf{q}_1\rangle = \langle \mathbf{p}_1,\mathbf{n}_2\rangle\langle \mathbf{n}_2,\mathbf{q}_1\rangle - \langle \mathbf{p}_1,\mathbf{q}_2\rangle\langle \mathbf{p}_2,\mathbf{q}_1\rangle - \langle \mathbf{p}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_2,\mathbf{q}_1\rangle,\\
0 &= \langle \mathbf{q}_1,\mathbf{q}_1\rangle = \langle \mathbf{q}_1,\mathbf{n}_2\rangle\langle \mathbf{n}_2,\mathbf{q}_1\rangle - \langle \mathbf{q}_1,\mathbf{q}_2\rangle\langle \mathbf{p}_2,\mathbf{q}_1\rangle - \langle \mathbf{q}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_2,\mathbf{q}_1\rangle,\\
0 &= \langle \mathbf{p}_2,\mathbf{p}_2\rangle = \langle \mathbf{p}_2,\mathbf{n}_1\rangle\langle \mathbf{n}_1,\mathbf{p}_2\rangle - \langle \mathbf{p}_2,\mathbf{q}_1\rangle\langle \mathbf{p}_1,\mathbf{p}_2\rangle - \langle \mathbf{p}_2,\mathbf{p}_1\rangle\langle \mathbf{q}_1,\mathbf{p}_2\rangle,\\
-1 &= \langle \mathbf{p}_2,\mathbf{q}_2\rangle = \langle \mathbf{p}_2,\mathbf{n}_1\rangle\langle \mathbf{n}_1,\mathbf{q}_2\rangle - \langle \mathbf{p}_2,\mathbf{q}_1\rangle\langle \mathbf{p}_1,\mathbf{q}_2\rangle - \langle \mathbf{p}_2,\mathbf{p}_1\rangle\langle \mathbf{q}_1,\mathbf{q}_2\rangle,\\
0 &= \langle \mathbf{q}_2,\mathbf{q}_2\rangle = \langle \mathbf{q}_2,\mathbf{n}_1\rangle\langle \mathbf{n}_1,\mathbf{q}_2\rangle - \langle \mathbf{q}_2,\mathbf{q}_1\rangle\langle \mathbf{p}_1,\mathbf{q}_2\rangle - \langle \mathbf{q}_2,\mathbf{p}_1\rangle\langle \mathbf{q}_1,\mathbf{q}_2\rangle.
\end{aligned}$$

Lemma 7.3 *For $j=1,2$, let γ_j be a geodesic with endpoints p_j and q_j in $\partial\mathbf{H}^2_{\mathbb{C}}$. Let $\mathbf{p}_j$ and $\mathbf{q}_j$ be lifts of p_j and q_j to $\mathbb{C}^{2,1}$ and let $\mathbf{n}_j\in\mathbb{C}^{2,1}$ be the normal vector to $\mathbf{p}_j$ and $\mathbf{q}_j$. Then*

$$\cosh\big(\rho(\gamma_1,\gamma_2)\big) \geq \left|\frac{\langle \mathbf{p}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_2,\mathbf{q}_1\rangle}{\langle \mathbf{p}_1,\mathbf{q}_1\rangle\langle \mathbf{q}_2,\mathbf{p}_2\rangle}\right| + \left|\frac{\langle \mathbf{p}_1,\mathbf{q}_2\rangle\langle \mathbf{p}_2,\mathbf{q}_1\rangle}{\langle \mathbf{p}_1,\mathbf{q}_1\rangle\langle \mathbf{p}_2,\mathbf{q}_2\rangle}\right| + \left|\frac{\langle \mathbf{p}_1,\mathbf{n}_2\rangle\langle \mathbf{n}_2,\mathbf{q}_1\rangle}{\langle \mathbf{p}_1,\mathbf{q}_1\rangle\langle \mathbf{n}_2,\mathbf{n}_2\rangle}\right|.$$

Proof: We normalise so that $\langle \mathbf{p}_j,\mathbf{q}_j\rangle=-1$ and $\langle \mathbf{n}_j,\mathbf{n}_j\rangle=1$. Then

$$\begin{aligned}
&\cosh\big(\rho(\gamma_1(t),\gamma_2(s))\big)\\
&= 2\cosh^2\left(\frac{\rho(\gamma_1(t),\gamma_2(s))}{2}\right)-1\\
&= \frac{2\big|\langle e^{t/2}\mathbf{p}_1+e^{-t/2}\mathbf{q}_1, e^{s/2}\mathbf{p}_2+e^{-s/2}\mathbf{q}_2\rangle\big|^2}{\big|e^{t/2}\mathbf{p}_1+e^{-t/2}\mathbf{q}_1\big|^2\big|e^{s/2}\mathbf{p}_2+e^{-s/2}\mathbf{q}_2\big|^2}-1\\
&= \frac{\big|e^{(t+s)/2}\langle \mathbf{p}_1,\mathbf{p}_2\rangle+e^{(t-s)/2}\langle \mathbf{p}_1,\mathbf{q}_2\rangle+e^{(-t+s)/2}\langle \mathbf{q}_1,\mathbf{p}_2\rangle+e^{(-t-s)/2}\langle \mathbf{q}_1,\mathbf{q}_2\rangle\big|^2}{2}-1\\
&= \frac{1}{2}\left(e^{t+s}\big|\langle \mathbf{p}_1,\mathbf{p}_2\rangle\big|^2+e^{-t-s}\big|\langle \mathbf{q}_1,\mathbf{q}_2\rangle\big|^2+2\Re\big(\langle \mathbf{p}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_2,\mathbf{q}_1\rangle\big)\right)\\
&\quad+\frac{1}{2}\left(+e^{t-s}\big|\langle \mathbf{p}_1,\mathbf{q}_2\rangle\big|^2+e^{-t+s}\big|\langle \mathbf{q}_1,\mathbf{p}_2\rangle\big|^2+2\Re\big(\langle \mathbf{p}_1,\mathbf{q}_2\rangle\langle \mathbf{p}_2,\mathbf{q}_1\rangle\big)\right)\\
&\quad+\frac{1}{2}\left(e^{t}\big|\langle \mathbf{p}_1,\mathbf{n}_2\rangle\big|^2+e^{-t}\big|\langle \mathbf{q}_1,\mathbf{n}_2\rangle\big|^2+e^{s}\big|\langle \mathbf{n}_1,\mathbf{p}_2\rangle\big|^2+e^{-s}\big|\langle \mathbf{n}_1,\mathbf{q}_2\rangle\big|^2\right)-1\\
&\geq \big|\langle \mathbf{p}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_2,\mathbf{q}_1\rangle\big|+\big|\langle \mathbf{p}_1,\mathbf{q}_2\rangle\langle \mathbf{p}_2,\mathbf{q}_1\rangle\big|+\big|\langle \mathbf{p}_1,\mathbf{n}_2\rangle\langle \mathbf{n}_2,\mathbf{q}_1\rangle\big|\\
&\quad+\big|\langle \mathbf{n}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_2,\mathbf{n}_1\rangle\big|+\Re\big(\langle \mathbf{n}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_2,\mathbf{n}_1\rangle\big)\\
&\geq \big|\langle \mathbf{p}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_1,\mathbf{q}_2\rangle\big|+\big|\langle \mathbf{p}_1,\mathbf{q}_2\rangle\langle \mathbf{q}_1,\mathbf{p}_2\rangle\big|+\big|\langle \mathbf{p}_1,\mathbf{n}_2\rangle\langle \mathbf{q}_1,\mathbf{n}_2\rangle\big|.
\end{aligned}$$

Equality in the penultimate line occurs if and only if all the following are true

$$e^{t+s}=\left|\frac{\langle \mathbf{q}_1,\mathbf{q}_2\rangle}{\langle \mathbf{p}_1,\mathbf{p}_2\rangle}\right|,\quad e^{t-s}=\left|\frac{\langle \mathbf{q}_1,\mathbf{p}_2\rangle}{\langle \mathbf{p}_1,\mathbf{q}_2\rangle}\right|,\quad e^{t}=\left|\frac{\langle \mathbf{q}_1,\mathbf{n}_2\rangle}{\langle \mathbf{p}_1,\mathbf{n}_2\rangle}\right|,\quad e^{s}=\left|\frac{\langle \mathbf{n}_1,\mathbf{q}_2\rangle}{\langle \mathbf{n}_1,\mathbf{p}_2\rangle}\right|.$$

Equality in the last line happens if and only if $\langle \mathbf{n}_1,\mathbf{p}_2\rangle\langle \mathbf{q}_2,\mathbf{n}_1\rangle$ is real and negative. $\square$

Corollary 7.4 *The geodesics γ_1 and γ_2 intersect if and only if*

$$[p_2,q_1,p_1,q_2] \quad \textit{and} \quad [q_2,q_1,p_1,p_2]$$

are both real, non-negative and their sum is at most 1.

Proof: Again, we normalise so that $\langle \mathbf{p}_j, \mathbf{q}_j \rangle = -1$ and $\langle \mathbf{n}_j, \mathbf{n}_j \rangle = 1$. Using this, our hypotheses become that

$$\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{q}_1 \rangle \quad \text{and} \quad \langle \mathbf{p}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle$$

are both real, non-negative and their sum is at most 1.

We see that

$$\begin{aligned} -\langle \mathbf{p}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{q}_1 \rangle &= -\langle \mathbf{n}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{n}_1 \rangle \\ &= 1 - \langle \mathbf{p}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle - \langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{q}_1 \rangle > 0. \end{aligned}$$

Define

$$e^{t_0} = -\frac{\langle \mathbf{q}_1, \mathbf{n}_2 \rangle}{\langle \mathbf{p}_1, \mathbf{n}_2 \rangle} > 0 \quad \text{and} \quad e^{s_0} = -\frac{\langle \mathbf{n}_1, \mathbf{q}_2 \rangle}{\langle \mathbf{n}_1, \mathbf{p}_2 \rangle} > 0. \tag{67}$$

Thus

$$\begin{aligned} e^{t_0+s_0} &= \frac{\langle \mathbf{q}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_1, \mathbf{q}_2 \rangle}{\langle \mathbf{p}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_1, \mathbf{p}_2 \rangle} \\ &= \frac{\langle \mathbf{q}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{n}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{n}_1 \rangle} \end{aligned}$$

where we have multiplied top and bottom by $\langle \mathbf{n}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{n}_1 \rangle = \langle \mathbf{p}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{q}_1 \rangle$. Using the identities from Lemma 7.2 to eliminate $\mathbf{n}_1$ and $\mathbf{n}_2$ we obtain:

$$\begin{aligned} e^{t_0+s_0} &= \frac{\langle \mathbf{q}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle + \langle \mathbf{q}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle + \langle \mathbf{q}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{p}_1 \rangle} \\ &= \frac{\langle \mathbf{q}_1, \mathbf{q}_2 \rangle}{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle} \cdot \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle + \langle \mathbf{q}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{p}_1 \rangle} \\ &\quad + \frac{\langle \mathbf{q}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{p}_2, \mathbf{p}_1 \rangle} \cdot \frac{\langle \mathbf{q}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{p}_1 \rangle}{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle + \langle \mathbf{q}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{p}_1 \rangle} \\ &= \frac{\langle \mathbf{q}_1, \mathbf{q}_2 \rangle}{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle} \cdot \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle + \langle \mathbf{q}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{p}_1 \rangle}{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle + \langle \mathbf{q}_1, \mathbf{p}_2 \rangle \langle \mathbf{p}_2, \mathbf{p}_1 \rangle} \\ &= \frac{\langle \mathbf{q}_1, \mathbf{q}_2 \rangle}{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle} \end{aligned}$$

where we have used our hypothesis that

$$\frac{\langle \mathbf{q}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{p}_2, \mathbf{p}_1 \rangle} = \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{q}_1 \rangle}{\left| \langle \mathbf{p}_1, \mathbf{p}_2 \rangle \right|^2}$$

is real.

A similar argument shows that

$$e^{t_0-s_0} = \frac{\langle \mathbf{q}_1, \mathbf{p}_2 \rangle}{\langle \mathbf{p}_1, \mathbf{q}_2 \rangle}.$$

Therefore we have equality at each stage in Lemma 7.3 and so we have

$$\begin{aligned} \cosh\big(\rho(\gamma_1(t_0), \gamma_2(s_0))\big) &= 2\langle \mathbf{p}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle + 2\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{q}_1 \rangle - 2\langle \mathbf{p}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{q}_1 \rangle - 1 \\ &= 1. \end{aligned}$$

Hence the two points are the same.

Conversely, if

$$-\langle \mathbf{p}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{q}_1 \rangle, \quad \langle \mathbf{p}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle, \quad \langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{q}_1 \rangle$$

do not all lie in the interval $[0, 1]$, applying the triangle inequality to

$$-1 = \langle \mathbf{p}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{q}_1 \rangle - \langle \mathbf{p}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle - \langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{q}_1 \rangle$$

we have the strict inequality

$$\left| \langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_1, \mathbf{q}_2 \rangle \right| + \left| \langle \mathbf{p}_1, \mathbf{q}_2 \rangle \langle \mathbf{q}_1, \mathbf{p}_2 \rangle \right| + \left| \langle \mathbf{p}_1, \mathbf{n}_2 \rangle \langle \mathbf{q}_1, \mathbf{n}_2 \rangle \right| > 1.$$

Hence the lower bound in Lemma 7.3 becomes $\cosh\big(\rho(\gamma_1(t), \gamma_2(s))\big) > 1$. Hence the two geodesics are disjoint. □

Proposition 7.5 *For $j = 1, 2$, let γ_j be a geodesic with endpoints p_j and q_j in $\partial \mathbf{H}^2_{\mathbb{C}}$. Then*

$$\cosh\Big(\rho(\gamma_1(t), \gamma_2(s))\Big) \geq \big|[p_2, q_1, p_1, q_2]\big| + \big|[q_2, q_1, p_1, p_2]\big|.$$

Proof: From Lemma 7.3 we have

$$\cosh\big(\rho(\gamma_1, \gamma_2)\big) \geq \left| \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{p}_1, \mathbf{q}_1 \rangle \langle \mathbf{q}_2, \mathbf{p}_2 \rangle} \right| + \left| \frac{\langle \mathbf{p}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{p}_1, \mathbf{q}_1 \rangle \langle \mathbf{p}_2, \mathbf{q}_2 \rangle} \right| + \left| \frac{\langle \mathbf{p}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{p}_1, \mathbf{q}_1 \rangle \langle \mathbf{n}_2, \mathbf{n}_2 \rangle} \right|.$$

Neglecting the third term and using

$$\left| \frac{\langle \mathbf{p}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{q}_2, \mathbf{p}_2 \rangle \langle \mathbf{p}_1, \mathbf{q}_1 \rangle} \right| = \big|[p_2, q_1, p_1, q_2]\big|, \qquad \left| \frac{\langle \mathbf{p}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{q}_1 \rangle}{\langle \mathbf{p}_2, \mathbf{q}_2 \rangle \langle \mathbf{p}_1, \mathbf{q}_1 \rangle} \right| = \big|[q_2, q_1, p_1, p_2]\big|$$

gives the result. □

7.4 Distance to complex lines

We begin by finding the distance from a point to a complex line. This should be compared with Proposition 7.1.

Proposition 7.6 *Let p and q be points of $\partial \mathbf{H}^2_{\mathbb{C}}$ and let $\mathbf{p}$ and $\mathbf{q}$ be lifts of p and q. Let L be the complex line spanned by $\mathbf{p}$ and $\mathbf{q}$. Then for any point z in $\mathbf{H}^2_{\mathbb{C}}$ we have*

$$\cosh^2\left(\frac{\rho(z, L)}{2}\right) = 2\Re\eta(p, q, z).$$

Proof: Without loss of generality suppose that $\langle \mathbf{p}, \mathbf{q} \rangle = -1$. Let $\mathbf{n}$ be the unit polar vector to L. Choose a lift $\mathbf{z}$ of z with $\langle \mathbf{z}, \mathbf{z} \rangle = -1$. Let $\mathbf{w} = \lambda\mathbf{p} + \mu\mathbf{q}$ be a point on L with

$\langle \mathbf{w}, \mathbf{w} \rangle = -\lambda\overline{\mu} - \mu\overline{\lambda} = -1$. Then

$$\begin{aligned}
&\cosh^2\left(\frac{\rho(z,w)}{2}\right) \\
&= \langle \mathbf{z}, \mathbf{w} \rangle \langle \mathbf{w}, \mathbf{z} \rangle \\
&= \left| \langle \mathbf{z}, \mathbf{p} \rangle \overline{\lambda} + \langle \mathbf{z}, \mathbf{q} \rangle \overline{\mu} \right|^2 \\
&= \left| \langle \mathbf{z}, \mathbf{p} \rangle \right| |\lambda|^2 + \langle \mathbf{z}, \mathbf{p} \rangle \langle \mathbf{q}, \mathbf{z} \rangle \overline{\lambda}\mu + \langle \mathbf{z}, \mathbf{q} \rangle \langle \mathbf{p}, \mathbf{z} \rangle \overline{\mu}\lambda + \left| \langle \mathbf{z}, \mathbf{q} \rangle \right| |\mu|^2 \\
&= \left| \langle \mathbf{z}, \mathbf{p} \rangle \right| |\lambda|^2 - \langle \mathbf{z}, \mathbf{p} \rangle \langle \mathbf{q}, \mathbf{z} \rangle \overline{\mu}\lambda - \langle \mathbf{z}, \mathbf{q} \rangle \langle \mathbf{p}, \mathbf{z} \rangle \overline{\lambda}\mu + \left| \langle \mathbf{z}, \mathbf{q} \rangle \right| |\mu|^2 \\
&\qquad + \langle \mathbf{z}, \mathbf{p} \rangle \langle \mathbf{q}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{q} \rangle \langle \mathbf{p}, \mathbf{z} \rangle \\
&= \left| \langle \mathbf{z}, \mathbf{p} \rangle \lambda - \langle \mathbf{z}, \mathbf{q} \rangle \mu \right|^2 + \langle \mathbf{z}, \mathbf{p} \rangle \langle \mathbf{q}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{q} \rangle \langle \mathbf{p}, \mathbf{z} \rangle \\
&\geq \langle \mathbf{z}, \mathbf{p} \rangle \langle \mathbf{q}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{q} \rangle \langle \mathbf{p}, \mathbf{z} \rangle.
\end{aligned}$$

We obtain equality in the last line with the point

$$\mathbf{w} = \frac{\langle \mathbf{z}, \mathbf{q} \rangle}{\sqrt{\langle \mathbf{z}, \mathbf{p} \rangle \langle \mathbf{q}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{q} \rangle \langle \mathbf{p}, \mathbf{z} \rangle}} \mathbf{p} + \frac{\langle \mathbf{z}, \mathbf{p} \rangle}{\sqrt{\langle \mathbf{z}, \mathbf{p} \rangle \langle \mathbf{q}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{q} \rangle \langle \mathbf{p}, \mathbf{z} \rangle}} \mathbf{q}$$

where we have used Lemma 3.4 to check that the denominator is well defined. (This denominator is chosen to ensure $\langle \mathbf{w}, \mathbf{w} \rangle = -1$.) Using $\langle \mathbf{z}, \mathbf{z} \rangle = \langle \mathbf{p}, \mathbf{q} \rangle = -1$, this gives the result. □

Corollary 7.7 *Let L be a complex line with polar vector $\mathbf{n}$. Let z be any point of $\mathbf{H}^2_{\mathbb{C}}$ with lift $\mathbf{z}$. Then*

$$\cosh^2\left(\frac{\rho(z,L)}{2}\right) = 1 - \frac{\langle \mathbf{z}, \mathbf{n} \rangle \langle \mathbf{n}, \mathbf{z} \rangle}{\langle \mathbf{z}, \mathbf{z} \rangle \langle \mathbf{n}, \mathbf{n} \rangle} \geq 1.$$

Proof: This follows from Proposition 7.6 using Lemma 3.4 and $\langle \mathbf{z}, \mathbf{z} \rangle = -1$. □

Proposition 7.8 *Let L_1 and L_2 be complex lines with polar vectors $\mathbf{n}_1$ and $\mathbf{n}_2$. Let*

$$N(L_1, L_2) = \frac{\langle \mathbf{n}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{n}_1 \rangle}{\langle \mathbf{n}_1, \mathbf{n}_1 \rangle \langle \mathbf{n}_2, \mathbf{n}_2 \rangle}.$$

(i) If $N(L_1, L_2) > 1$ then L_1 and L_2 are ultraparallel and

$$\cosh^2\left(\frac{\rho(L_1, L_2)}{2}\right) = N(L_1, L_2).$$

(ii) If $N(L_1, L_2) = 1$ then L_1 and L_2 are asymptotic or coincide.

(iii) If $N(L_1, L_2) < 1$ then L_1 and L_2 intersect.

Proof: First suppose that $N(L_1, L_2) > 1$. Let L_2 be spanned by $\mathbf{p}_2$ and $\mathbf{q}_2$. Suppose that $\langle \mathbf{p}_2, \mathbf{q}_2 \rangle = -1$. Also suppose, without loss of generality, that $\langle \mathbf{n}_1, \mathbf{n}_1 \rangle = \langle \mathbf{n}_2, \mathbf{n}_2 \rangle = 1$.

Then a general point w on L_2 has lift $\mathbf{w} = \lambda \mathbf{p}_2 + \mu \mathbf{q}_2$ in $\mathbb{C}^{2,1}$. Without loss of generality, we suppose that $\langle \mathbf{w}, \mathbf{w} \rangle = -\lambda\overline{\mu} - \mu\overline{\lambda} = -1$. From Corollary 7.7, we know that

$$\begin{aligned}
\cosh^2\left(\frac{\rho(L_1, w)}{2}\right) &= 1 - \frac{\langle \mathbf{w}, \mathbf{n}_1 \rangle \langle \mathbf{n}_1, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle \langle \mathbf{n}_1, \mathbf{n}_1 \rangle} \\
&= 1 + \left| \langle \mathbf{n}_1, \mathbf{p}_2 \rangle \overline{\lambda} + \langle \mathbf{n}_1, \mathbf{q}_2 \rangle \overline{\mu} \right|^2 \\
&= 1 + \left| \langle \mathbf{n}_1, \mathbf{p}_2 \rangle \lambda - \langle \mathbf{n}_1, \mathbf{q}_2 \rangle \mu \right|^2 \\
&\quad + \langle \mathbf{n}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{n}_1 \rangle + \langle \mathbf{n}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{n}_1 \rangle \\
&\geq 1 + \langle \mathbf{n}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{n}_1 \rangle + \langle \mathbf{n}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{n}_1 \rangle
\end{aligned}$$

with equality when $\lambda/\mu = \langle \mathbf{n}_1, \mathbf{q}_2 \rangle / \langle \mathbf{n}_1, \mathbf{p}_2 \rangle$. Thus we choose $\mathbf{w}$ to be the point

$$\mathbf{w} = \nu \langle \mathbf{n}_1, \mathbf{q}_2 \rangle \mathbf{p}_2 + \nu \langle \mathbf{n}_1, \mathbf{p}_2 \rangle \mathbf{q}_2$$

where ν is chosen so that

$$1 = -\langle \mathbf{w}, \mathbf{w} \rangle = |\nu|^2 \Big(\langle \mathbf{n}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{n}_1 \rangle + \langle \mathbf{n}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{n}_1 \rangle \Big).$$

Now writing

$$\mathbf{n}_1 = -\langle \mathbf{n}_1, \mathbf{q}_2 \rangle \mathbf{p}_2 - \langle \mathbf{n}_1, \mathbf{p}_2 \rangle \mathbf{q}_2 + \langle \mathbf{n}_1, \mathbf{n}_2 \rangle \mathbf{n}_2$$

we have

$$1 = \langle \mathbf{n}_1, \mathbf{n}_1 \rangle = -\langle \mathbf{n}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{n}_1 \rangle - \langle \mathbf{n}_1, \mathbf{q}_2 \rangle \langle \mathbf{p}_2, \mathbf{n}_1 \rangle + \langle \mathbf{n}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{n}_1 \rangle.$$

Thus we see that

$$|\nu|^2 = \frac{1}{N(L_1, L_2) - 1}$$

and

$$\begin{aligned}
\cosh^2\left(\frac{\rho(L_1, w)}{2}\right) &= 1 + \langle \mathbf{n}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{n}_1 \rangle + \langle \mathbf{n}_1, \mathbf{p}_2 \rangle \langle \mathbf{q}_2, \mathbf{n}_1 \rangle \\
&= \langle \mathbf{n}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{n}_1 \rangle.
\end{aligned}$$

Using $\langle \mathbf{n}_1, \mathbf{n}_1 \rangle = \langle \mathbf{n}_2, \mathbf{n}_2 \rangle = 1$ this gives part (i).

If $L_1 = L_2$ then $\mathbf{n}_2 = \lambda \mathbf{n}_1$ and so $N(L_1, L_2) = 1$. If L_1 and L_2 are asymptotic, then we can write L_j as the span of null vectors $\mathbf{p}$ and $\mathbf{q}_j$. As usual, suppose that $\langle \mathbf{p}, \mathbf{q}_j \rangle = -1$. Therefore

$$\mathbf{n}_1 = -\langle \mathbf{n}_1, \mathbf{q}_2 \rangle \mathbf{p} - \langle \mathbf{n}_1, \mathbf{p} \rangle \mathbf{q}_2 + \langle \mathbf{n}_1, \mathbf{n}_2 \rangle \mathbf{n}_2 = -\langle \mathbf{n}_1, \mathbf{q}_2 \rangle \mathbf{p} + \langle \mathbf{n}_1, \mathbf{n}_2 \rangle \mathbf{n}_2$$

since $\langle \mathbf{n}_1, \mathbf{p} \rangle = 0$. Hence

$$1 = \langle \mathbf{n}_1, \mathbf{n}_1 \rangle = \langle \mathbf{n}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{n}_1 \rangle.$$

Therefore $N(L_1, L_2) = 1$.

Finally, suppose that L_1 and L_2 intersect in a point w. Then, lift w to the vector $\mathbf{w} = \lambda \mathbf{p}_1 + \mu \mathbf{q}_1$. Since $\mathbf{w}$ is also on L_2 we have

$$0 = \langle \mathbf{w}, \mathbf{n}_2 \rangle = \lambda \langle \mathbf{p}_1, \mathbf{n}_2 \rangle + \mu \langle \mathbf{q}_1, \mathbf{n}_2 \rangle.$$

Therefore, using $\mathbf{w} \in V_-$ we see

$$\begin{aligned} 0 &> \langle \mathbf{w}, \mathbf{w} \rangle \\ &= -\lambda\overline{\mu} - \mu\overline{\lambda} \\ &= |\lambda|^2 \left(\frac{\langle \mathbf{p}_1, \mathbf{n}_2 \rangle}{\langle \mathbf{q}_1, \mathbf{n}_2 \rangle} + \frac{\langle \mathbf{n}_2, \mathbf{p}_1 \rangle}{\langle \mathbf{n}_2, \mathbf{q}_1 \rangle} \right) \\ &= \frac{|\lambda|^2}{\big|\langle \mathbf{q}_1, \mathbf{n}_2 \rangle\big|^2} \big(\langle \mathbf{p}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{q}_1 \rangle + \langle \mathbf{q}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{p}_1 \rangle \big) \\ &= \frac{|\lambda|^2}{\big|\langle \mathbf{q}_1, \mathbf{n}_2 \rangle\big|^2} \big(\langle \mathbf{n}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{n}_1 \rangle - 1 \big). \end{aligned}$$

Therefore $N(L_1, L_2) = \langle \mathbf{n}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{n}_1 \rangle < 1$. □

Suppose that two complex lines L_1 and L_2 intersect in the origin. This means their polar vectors $\mathbf{n}_j$ have the form

$$\mathbf{n}_1 = \begin{bmatrix} n_{11} \\ n_{12} \\ 0 \end{bmatrix}, \quad \mathbf{n}_2 = \begin{bmatrix} n_{21} \\ n_{22} \\ 0 \end{bmatrix}.$$

Therefore, in $\mathbb{C}^2$, the two lines have normal vectors

$$\begin{pmatrix} n_{11} \\ n_{12} \end{pmatrix}, \quad \begin{pmatrix} n_{21} \\ n_{22} \end{pmatrix}.$$

Using the Cauchy-Schwarz inequality, we see that the angle between the two lines is θ where

$$\begin{aligned} \cos^2(\theta) &= \frac{\big|\overline{n}_{21} n_{11} + \overline{n}_{22} n_{12}\big|^2}{\big(|n_{11}|^2 + |n_{12}|^2\big)\big(|n_{21}|^2 + |n_{22}|^2\big)} \\ &= \frac{\langle \mathbf{n}_1, \mathbf{n}_2 \rangle \langle \mathbf{n}_2, \mathbf{n}_1 \rangle}{\langle \mathbf{n}_1, \mathbf{n}_1 \rangle \langle \mathbf{n}_2, \mathbf{n}_2 \rangle} = N(L_1, L_2). \end{aligned}$$

Motivated by this, we define the *angle* $\theta \in [0, \pi/2]$ between any pair of intersecting complex lines by

$$\cos^2(\theta) = N(L_1, L_2). \tag{68}$$

This is clearly invariant under complex hyperbolic isometries.

7.5 Orthogonal projection

Suppose that z_1 and z_2 are points of complex hyperbolic space and let L be the complex line passing through them. In other words, if $\mathbf{z}_1$ and $\mathbf{z}_2$ are lifts of z_1 and z_2 to $\mathbb{C}^{2,1}$ then the complex linear span of $\mathbf{z}_1$ and $\mathbf{z}_2$ maps to L under the canonical projection map $\mathbb{P}$. Since z_1 and z_2 are distinct, we see that $\mathbf{z}_1$ and $\mathbf{z}_2$ are linearly independent.

A vector $\mathbf{n}$ so that $\langle \mathbf{n}, \mathbf{z}_1 \rangle = \langle \mathbf{n}, \mathbf{z}_2 \rangle = 0$ is called a *polar vector* to L. This implies that $\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{n}\}$ is a basis of $\mathbb{C}^{2,1}$. If C is the matrix whose columns are $\mathbf{z}_1$, $\mathbf{z}_2$ and $\mathbf{z}_3$ then

$$C^* H C = \begin{pmatrix} \langle \mathbf{z}_1, \mathbf{z}_1 \rangle & \langle \mathbf{z}_2, \mathbf{z}_1 \rangle & 0 \\ \langle \mathbf{z}_1, \mathbf{z}_2 \rangle & \langle \mathbf{z}_2, \mathbf{z}_2 \rangle & 0 \\ 0 & 0 & \langle \mathbf{n}, \mathbf{n} \rangle \end{pmatrix}$$

has signature $(2, 1)$. Since the upper left hand block has signature $(1, 1)$, we see that $\langle \mathbf{n}, \mathbf{n} \rangle > 0$ and so $\mathbf{n} \in V_+$.

Conversely, given any $\mathbf{n} \in V_+$ we can find linearly independent vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in $\mathbb{C}^{2,1}$ so that $\langle \mathbf{n}, \mathbf{z}_1 \rangle = \langle \mathbf{n}, \mathbf{z}_2 \rangle = 0$. Hence $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{n}\}$ is a basis of $\mathbb{C}^{2,1}$ and a reversing the argument given above, the linear span of $\mathbf{w}_1$ and $\mathbf{w}_2$ must intersect V_-. Therefore $\mathbf{n}$ is the polar vector to some complex line in $\mathbf{H}^2_{\mathbb{C}}$. Of course, all complex scalar multiples of $\mathbf{n}$ are polar vectors of the same complex line.

Let z be a point of $\mathbf{H}^2_{\mathbb{C}}$ and let $\mathbf{z}$ be its standard lift to $V_- \subset \mathbb{C}^{2,1}$. Given a complex line L with polar vector $\mathbf{n}$ we define *orthogonal projection onto* L to be the map

$$\Pi_L(z) = \mathbb{P}\left(\mathbf{z} - \frac{\langle \mathbf{z}, \mathbf{n} \rangle}{\langle \mathbf{n}, \mathbf{n} \rangle} \mathbf{n} \right). \tag{69}$$

It is clear that the image of Π_L is the complex line L.

Lemma 7.9 *Let w be any point of a complex line L then the preimage of w under Π_L, that is*

$$\Pi_L^{-1}(w) = \left\{ z \in \mathbf{H}^2_{\mathbb{C}} \ : \ \Pi_L(z) = w \right\}$$

is a complex line orthogonal to L at w.

Proof: Let $\mathbf{w}$ be the standard lift of w to V_-. Let $\mathbf{v}$ be any point of $\mathbb{C}^{2,1}$ linearly independent from $\mathbf{w}$ and $\mathbf{n}$. Then $\{\mathbf{w}, \mathbf{n}, \mathbf{v}\}$ is a basis for $\mathbb{C}^{2,1}$. Let z be any point of $\mathbf{H}^2_{\mathbb{C}}$, then we can write the standard lift of z as $\mathbf{z} = \alpha\mathbf{w} + \beta\mathbf{n} + \gamma\mathbf{v}$. We have

$$\begin{aligned} \Pi_L(z) &= \mathbb{P}\left(\mathbf{z} - \frac{\langle \mathbf{z}, \mathbf{n} \rangle}{\langle \mathbf{n}, \mathbf{n} \rangle} \mathbf{n} \right) \\ &= \mathbb{P}\left(\alpha\mathbf{w} + \gamma\mathbf{v} - \gamma \frac{\langle \mathbf{v}, \mathbf{n} \rangle}{\langle \mathbf{n}, \mathbf{n} \rangle} \mathbf{n} \right). \end{aligned}$$

This equals w if and only if $\gamma = 0$. Therefore $\Pi_L^{-1}(z)$ is the complex line corresponding to the complex linear span of $\mathbf{w}$ and $\mathbf{n}$. This intersects L in w. Let $\mathbf{m}$ be the polar vector to this complex line, then $\langle \mathbf{m}, \mathbf{n} \rangle = 0$ by definition. Since $\mathbf{n}$ is the polar vector to L we see that these two complex lines are orthogonal using (68). □

We now discuss the analogous formulae for totally real Lagrangian planes R. We discuss orthogonal projection Π_R onto R and its fibres $\Pi^{-1}(z)$. First, we give a formula for the midpoint of two points of complex hyperbolic space.

Proposition 7.10 *Let $\mathbf{z}$, $\mathbf{w}$ be any points of $V_- \subset \mathbb{C}^{2,1}$ and $z = \mathbb{P}\mathbf{z}$, $w = \mathbb{P}\mathbf{w}$ be the corresponding points of $\mathbf{H}^2_{\mathbb{C}}$. Let*

$$\mathbf{m} = \frac{1}{|\mathbf{z}|} \mathbf{z} - \frac{\langle \mathbf{z}, \mathbf{w} \rangle}{\big|\langle \mathbf{z}, \mathbf{w} \rangle\big| \, |\mathbf{w}|} \mathbf{w}. \tag{70}$$

Then $\mathbf{m} \in V_-$ and, writing $m = \mathbb{P}\mathbf{m}$, we have $\rho(m, z) = \rho(m, w) = \rho(z, w)/2$.

If m is as defined in Proposition 7.10 then we call m the *midpoint* of z and w.

Proof: First, observe that

$$\langle \mathbf{m},\, \mathbf{m}\rangle = -2 - \frac{2\big|\langle \mathbf{z},\, \mathbf{w}\rangle\big|}{|\mathbf{z}|\,|\mathbf{w}|} = -2\Big(1+\cosh\big(\rho(z,w)/2\big)\Big) = -4\cosh^2\Big(\rho(z,w)/4\Big).$$

Thus $\mathbf{m}\in V_-$ and $m=\mathbb{P}\mathbf{m}\in \mathbf{H}^2_{\mathbb{C}}$ and we write $|\mathbf{m}| = \sqrt{-\langle \mathbf{m},\, \mathbf{m}\rangle} = 2\cosh\big(\rho(z,w)/4\big)$. Moreover,

$$\begin{aligned}
\langle \mathbf{m},\, \mathbf{z}\rangle &= \frac{\langle \mathbf{z},\, \mathbf{z}\rangle}{|\mathbf{z}|} - \frac{\langle \mathbf{z},\, \mathbf{w}\rangle\langle \mathbf{w}, \mathbf{z}\rangle}{|\langle \mathbf{z},\, \mathbf{w}\rangle|\,|\mathbf{w}|} \\
&= -|\mathbf{z}| - \frac{|\langle \mathbf{z},\, \mathbf{w}\rangle|}{|\mathbf{w}|} \\
&= -|\mathbf{z}|\Big(1+\cosh\big(\rho(z,w)/2\big)\Big) \\
&= -|\mathbf{z}|\, 2\cosh^2\Big(\rho(z,w)/4\Big) \\
&= -|\mathbf{z}|\,|\mathbf{m}|\cosh\Big(\rho(z,w)/4\Big).
\end{aligned}$$

Therefore

$$\cosh\Big(\rho(m,z)/2\Big) = \frac{\big|\langle \mathbf{m},\, \mathbf{z}\rangle\big|}{|\mathbf{z}|\,|\mathbf{m}|} = \cosh\Big(\rho(z,w)/4\Big).$$

Similarly

$$\begin{aligned}
\langle \mathbf{m},\, \mathbf{w}\rangle &= \frac{\langle \mathbf{z},\, \mathbf{w}\rangle}{|\mathbf{z}|} - \frac{\langle \mathbf{z},\, \mathbf{w}\rangle\langle \mathbf{w}, \mathbf{w}\rangle}{|\langle \mathbf{z},\, \mathbf{w}\rangle|\,|\mathbf{w}|} \\
&= \frac{\langle \mathbf{z},\, \mathbf{w}\rangle}{|\langle \mathbf{z},\, \mathbf{w}\rangle|}\left(|\mathbf{w}| + \frac{|\langle \mathbf{z},\, \mathbf{w}\rangle|}{|\mathbf{z}|}\right) \\
&= \frac{\langle \mathbf{z},\, \mathbf{w}\rangle}{|\langle \mathbf{z},\, \mathbf{w}\rangle|}\,|\mathbf{w}|\,|\mathbf{m}|\cosh\Big(\rho(z,w)/4\Big)
\end{aligned}$$

and so

$$\cosh\Big(\rho(m,w)/2\Big) = \frac{\big|\langle \mathbf{m},\, \mathbf{w}\rangle\big|}{|\mathbf{w}|\,|\mathbf{m}|} = \cosh\Big(\rho(z,w)/4\Big).$$

Hence $\rho(m,z)=\rho(m,w)=\rho(z,w)/2$ as required. $\square$

We use Proposition 7.10 to derive a formula for the orthogonal projection onto a Lagrangian plane R. Let ι_R denote the (anti-holomorphic) reflection in R. Then the orthogonal projection $\Pi_R(z)$ of any $z\in\mathbf{H}^2_{\mathbb{C}}$ onto R is defined to be the midpoint m of the points z and $\iota_R(z)$. That is, if $\mathbf{z}\in V_-$ is a lift of z then

$$\Pi_R(z) = \mathbb{P}\left(\frac{1}{|\mathbf{z}|}\,\mathbf{z} - \frac{\langle \mathbf{z},\, \iota_R(\mathbf{z})\rangle}{|\langle \mathbf{z},\, \iota_R(\mathbf{z})\rangle|\,|\iota_R(\mathbf{z})|}\,\iota_R(\mathbf{z})\right).$$

Proposition 7.11 *Let R be a Lagrangian plane stabilised by the subgroup G_R of* $\mathrm{SU}(2,1)$. *Then, for every $A\in G_R$*

$$A\circ\Pi_R = \Pi_R\circ A.$$

Consequently, if $w\in R$,

$$\Pi_R^{-1}\big(A(w)\big) = A\left(\Pi_R^{-1}(w)\right).$$

Proof: Let $z \in \mathbf{H}^2_{\mathbb{C}}$. Then, $\Pi_R(z) = m$ is the midpoint of z and $\iota(z)$. Hence

$$\rho\big(A(z), A(m)\big) = \rho(z, m) = \rho\big(\iota(z), m\big) = \rho\big(A\iota(z), A(m)\big).$$

Also

$$\rho\big(A(z), A\iota(z)\big) = \rho\big(z, \iota(z)\big) = 2\rho(m, z) = 2\rho\big(A(m), A(z)\big).$$

Thus $A(m)$ is the midpoint of $A(z)$ and $A\iota(z)$. But since $A\iota(z) = \iota A(z)$ we see that

$$\Pi_R\big(A(z)\big) = A(m) = A\big(\Pi_R(z)\big).$$

Now suppose that $w \in R$ and choose any z with $\Pi_R(z) = w$. Then

$$A(w) = A\Pi_R(z) = \Pi_R A(z).$$

Thus $A(z) \in {\Pi_R}^{-1}A(w)$ and so $A{\Pi_R}^{-1}(w) \subset {\Pi_R}^{-1}A(w)$. Similarly if z' is chosen so that $\Pi_R(z') = A(w)$ then

$$w = A^{-1}\Pi_R(z') = \Pi_R A^{-1}(z')$$

and so $z' \in A{\Pi_R}^{-1}(w)$. Hence ${\Pi_R}^{-1}A(w) \subset A{\Pi_R}^{-1}(w)$. □

We consider the special case where R is the standard real Lagrangian plane $R_{\mathbb{R}}$ in the ball model of $\mathbf{H}^2_{\mathbb{C}}$, that is

$$R_{\mathbb{R}} = \mathbf{H}^2_{\mathbb{R}} = \big\{(z_1, z_2) \in \mathbf{H}^2_{\mathbb{C}} : \Im(z_1) = \Im(z_2) = 0\big\}$$

and we denote orthogonal projection onto $R_{\mathbb{R}}$ by $\Pi_{\mathbb{R}}$. Consider a point $z = (z_1, z_2) \in \mathbf{H}^2_{\mathbb{C}}$. Then reflection $\iota_{\mathbb{R}}$ in $R_{\mathbb{R}}$ is given by

$$\iota_{\mathbb{R}}(z) = \overline{z} = \big(\overline{z}_1, \overline{z}_2\big).$$

We write

$$\eta(z)^2 = -\langle \mathbf{z},\, \iota_{\mathbb{R}}\mathbf{z}\rangle_1 = 1 - {z_1}^2 - {z_2}^2.$$

Observe that

$$0 < \langle \mathbf{z}, \mathbf{z}\rangle_1 = 1 - |z_1|^2 - |z_2|^2 \leq \Re\big(1 - {z_1}^2 - {z_2}^2\big) = \Re\big(\eta(z)^2\big),$$

and in particular, $\eta(z)^2 \neq 0$.

Applying (70) we find that the midpoint $m = (m_1, m_2)$ of z and $\iota_{\mathbb{R}}(z)$ is given by

$$m_k = \frac{\big|\eta(z)^2\big| z_k + \eta(z)^2 \overline{z}_k}{\big|\eta(z)^2\big| + \eta(z)^2} = 2\big|\eta(z)^2\big| \frac{\Re\Big(\overline{z}_k\big(\big|\eta(z)^2\big| + \eta(z)^2\big)\Big)}{\Big|\big|\eta(z)^2\big| + \eta(z)^2\Big|},$$

for $k = 1, 2$. Clearly, m lies on $R_{\mathbb{R}}$, and if $z \in R_{\mathbb{R}}$, then $\Pi_{\mathbb{R}}(z) = z$.

Corollary 7.12 $\Pi_{\mathbb{R}}$ *is real analytic.*

The subgroup of $\mathrm{SU}(2,1)$ stabilising $R_{\mathbb{R}}$ comprises those matrices with all real entries, that is $\mathrm{SO}(2,1)$ the isometry group of the hyperbolic plane. Proposition 7.11 immediately implies that $\Pi_{\mathbb{R}}$ commutes with all elements of $\mathrm{SO}(2,1)$.

Proposition 7.13 *If $R_{\mathbb{R}}$ is the standard real Lagrangian plane*

$$R_{\mathbb{R}} = \{(z_1, z_2) \in \mathbf{H}^2_{\mathbb{C}} : \Im(z_1) = \Im(z_2) = 0\},$$

then ${\Pi_{\mathbb{R}}}^{-1}(0,0)$ is the purely imaginary Lagrangian plane

$$R_{\mathbb{J}} = \{(z_1, z_2) \in \mathbf{H}^2_{\mathbb{C}} : \Re(z_1) = \Re(z_2) = 0\}.$$

Proof: If z_1 and z_2 are both purely imaginary then $\eta(z)^2 = 1 - z_1^2 - z_2^2$ is a positive real number. It is clear from the above construction that

$$m_1 = \Re(\overline{z}_1) = 0, \quad m_2 = \Re(\overline{z}_2) = 0.$$

Thus the Lagrangian plane $R_{\mathbb{J}}$ is contained in ${\Pi_{\mathbb{R}}}^{-1}(0,0)$.

Conversely, the set $\Pi_{\mathbb{R}}^{-1}(0,0)$ is the collection of points $(z_1, z_2) \in \mathbf{H}^2_{\mathbb{C}}$ satisfying

$$|\eta(z)^2| z_1 + \eta(z)^2 \overline{z}_1 = |\eta(z)^2| z_2 + \eta(z)^2 \overline{z}_2 = 0.$$

When z_1 and z_2 are both non-zero, these two equations are equivalent to

$$\frac{{z_1}^2}{|z_1|^2} = \frac{{z_2}^2}{|z_2|^2} = \frac{-\eta(z)^2}{|\eta(z)|^2}.$$

Writing ${z_1}^2 = |z_1|^2 e^{i\phi}$ and ${z_2}^2 = |z_2|^2 e^{i\phi}$ we obtain

$$|\eta(z)|^2 = -\eta(z)^2 e^{-i\phi} = -(1 - {z_1}^2 - {z_2}^2) e^{-i\phi} = -e^{-i\phi} + |z_1|^2 + |z_2|^2.$$

Therefore $e^{i\phi} \in \mathbb{R}$. Since $|z_1|^2 + |z_2|^2 < 1$ we see that $e^{i\phi} = -1$. Thus z_1 and z_2 are both purely imaginary. When one of z_1 or z_2 is zero, a similar argument shows that the other one is purely imaginary (or zero). Thus ${\Pi_{\mathbb{R}}}^{-1}(0,0)$ is contained in the Lagrangian plane $R_{\mathbb{J}}$. □

Using the fact that $\mathrm{SU}(2,1)$ acts transitively on the set of Lagrangian planes in $\mathbf{H}^2_{\mathbb{C}}$ we immediately have:

Corollary 7.14 *Let w be any point on the Lagrangian plane R. Then ${\Pi_R}^{-1}(w)$ is a Lagrangian plane.*

Corollary 7.15 *For every Lagrangian plane R, the orthogonal projection Π_R is real analytic.*

8 Notes

2.1 The terms first and second Hermitian form were defined by Epstein [7]. The third Hermitian form was defined (but not named) by Chen and Greenberg [3] and has been used extensively by others.

2.2 The connection between Hermitian forms and models of complex hyperbolic space given is completely standard. It is an example of the more general connection between quadratic and Hermitian forms for symmetric spaces, see Chen and Greenberg [3] or Chapter 19 of Mostow [23]. The formula (5) holds for other rank 1 symmetric spaces, see page 135 of [23]. The formula for the Bergman distance using the cross ratio is contained in Giraud [11]. The formula for the volumes of balls Proposition 2.2 may be found on page 104 of Goldman [12], but there are some numerical errors there (compare this to Gray, Lemma 6.18 on page 108 and Corollary A.3 on page 254).

2.3 For other Cayley transforms see Section 4.1.1 of Goldman [12] or page 574 of Kamiya [16], for example.

3.1 Similar formulae to those given in this section were given in Kamiya [15].

3.2 The formulae in this section are analogous to those in the previous section.

3.3 The Hermitian cross product is defined in Section 2.2.7 of Goldman [12]. The fact that the isometry group of $\mathbf{H}^2_{\mathbb{C}}$ acts transitively and acts doubly transitively on the boundary is a special case of similar results for other symmetric spaces.

3.4 Theorem 3.5 is a generalisation of Theorem 7.4.1 of Beardon [2].

4.1 The Heisenberg group is widely studied by analysts, see Korányi [19] or Korányi and Reimann [20]. Its relationship to the boundary of complex hyperbolic space generalises to all rank 1 symmetric spaces of compact type. See for example Section 4.2 of Goldman [12] for more about this.

4.2 Horospherical coordinates were introduced by Goldman and Parker [13].

4.3 The Cygan metric was constructed for the Heisenberg group by Cygan [4], Lemma 2. See also Cygan [5] and Korányi [19], page 227. This metric was extended to the Siegel domain in [24].

4.4 For the exact sequence (60) see Scott [27] page 467. The distortion result, Lemma 4.6, is due to Kamiya, Proposition 2.4 of [17] and is related to Theorem 5.22 of Basmajian and Miner [1].

5.1 Parts (i) and (ii) of Lemma 4.7 are Proposition 2.6 (4) and (5) of Kamiya [17]. The expression of points on a geodesic in terms of their endpoints may be found in Theorem 3.3.3 of Goldman [12]. The treatment we give here follows Sandler, Section 3 of [26] (compare this with page 242 of Goldman [12] for example).

5.2 See Section 3.1.4 of Goldman [12]. For the Poincaré metric on the hyperbolic plane see Section 7.2 of Beardon [2].

5.3 See Section 2.2.1 of Goldman [12]. For the Klein-Beltrami metric on the hyperbolic plane see Section 3.3 of Ratcliffe [25].

5.4 See Sections 3.1.4 and 3.1.9 of Goldman [12].

5.5 There is an extensive treatment of the material in this section in Sections 4.3 and 4.4 of Goldman [12].

6.1 The classification of complex hyperbolic isometries goes back to Giraud's paper of 1921 [11], see also page 52 of Chen and Greenberg [3].

6.2 The use of trace to classify isometries may be found in Theorem 6.2.4 of Goldman [12]. See also Theorem 6.5 of [6].

6.3 The classification of elements of $\mathrm{PU}(2,1)$ by means of products of pairs of real reflections is due to Falbel and Zocca [9]. It may be thought of as a generalisation of the classical idea that holomorphic (orientation preserving) isometries of the hyperbolic plane may be written as products of pairs of reflections in geodesics. Similarly, it generalises Fenchel's idea [10] that all orientation preserving isometries of real hyperbolic 3-space may be decomposed as a product of half-turns (rotation through π) about a pair of geodesics.

7.1 Cross ratios for complex hyperbolic space we introduced by Korányi and Reimann [20].

7.2 The η invariant in Goldman [12], Sandler [26]. It is also related to the A invariant of Kamiya, by $A(p,q;z) = 1/|\eta(p,q,z)|$ of Kamiya [16].

7.3 The first part of this section is mostly taken from Sandler's paper [26]. This includes the example, which is the same as Sandler's but using our conventions. The lower bound for the distance in the last part of this section is taken from Markham and Parker [22].

7.4 See page 100 of Goldman [12] or else Sandler [26].

7.5 For orthogonal projection onto complex lines and Lagrangian planes see Sections 3.3.2 and 3.3.6 of Goldman [12], respectively.

References

[1] A. Basmajian & R. Miner, *Discrete subgroups of complex hyperbolic motions*, Inventiones Math. **131** (1998), 85–136.

[2] A.F. Beardon, The Geometry of Discrete Groups, Springer, 1983.

[3] S.S. Chen & L. Greenberg, *Hyperbolic spaces* in "Contributions to Analysis", ed. L.V. Ahlfors et al., Academic Press, 1974.

[4] J. Cygan, *Wiener's test for the Brownian motion on the Heisenberg group*, Colloquium Math. **39** (1978), 367-373.

[5] J. Cygan, *Subadditivity of homogeneous norms on certain nilpotent Lie groups*, Proc. Amer. Math. Soc. **83** (1981), 69–70.

[6] P. Eberlein & B. O'Niell, *Visibility manifolds*, Pacific J. Maths. **46** (1973) 45–109.

[7] D.B.A. Epstein, *Complex hyperbolic geometry* in "Analytical and Geometric Aspects of Hyperbolic Space", ed. D.B.A. Epstein, London Mathematical Society Lecture Notes Series **111** (1987), 93–111.

[8] D.B.A. Epstein, J.W. Cannon, D.F. Holt, S.V.F. Levy, M.S. Paterson & W.P. Thurston, Word Processing in Groups, Jones and Bartlett, 1992.

[9] E. Falbel & V. Zocca, *A Poincaré's polyhedron theorem for complex hyperbolic geometry*, Journal für die reine und angewandte Mathematik **516** (1999) 133–158.

[10] W. Fenchel, Elementary Geometry in Hyperbolic Space, de Gruyter Studies in Mathematics **11**, 1989.

[11] G. Giraud, *Sur certaines fonctions automorphe de deux variables*, Ann. Ec. Norm. **38** (1921), 43–164.

[12] W.M. Goldman, Complex Hyperbolic Geometry, Oxford University Press, 1999.

[13] W.M. Goldman & J.R. Parker, *Dirichlet polyhedra for dihedral groups acting on complex hyperbolic space*, J. Geometrical Analysis **2** (1992), 517–554.

[14] S. Kamiya, *Notes on non-discrete subgroups of* $\hat{U}(1,n;\mathbb{F})$, Hiroshima Math. J. **13** (1983), 501–506.

[15] S. Kamiya, *Notes on elements of* $U(1,n;\mathbb{C})$, Hiroshima Math. J. **21** (1991), 23–45.

[16] S. Kamiya, *On H-balls and canonical regions of loxodromic elements in complex hyperbolic space*, Math. Proc. Cambridge Phil. Soc. **113**, (1993), 573–582.

[17] S. Kamiya, *On discrete subgroups of* $PU(1,2;\mathbb{C})$ *with Heisenberg translations*, J. London Math. Soc. **62** (2000), 827–842.

[18] F. Kirwan, Complex Algebraic Curves, L.M.S. Student Texts **23**, Cambridge University Press, 1992.

[19] A. Korányi, *Geometric aspects of analysis on the Heisenberg group*, in "Topics in Modern Harmonic Analysis", 209–258 Instituto Nazionale di Alta Mathematica, Rome (1983).

[20] A. Korányi & H.M. Reimann, *The complex cross-ratio on the Heisenberg group*, L'Enseign. Math. **33** (1987), 291–300.

[21] E.H. Lockwood, A Book of Curves, Cambridge University Press, 1961.

[22] S. Markham & J.R. Parker, *Collars in complex and quaternionic hyperbolic manifolds*, Geometriae Dedicata **97** (2003), 199–213.

[23] G.D. Mostow, Strong Rigidity of Locally Symmetric Spaces, Annals of Maths. Studies **78**, Princeton University Press (1973).

[24] J.R. Parker, *Shimizu's lemma for complex hyperbolic space*, International J. Math. **3** (1992), 291–308.

[25] J.G. Ratcliffe, Foundations of Hyperbolic Manifolds, Springer, 1994.

[26] H. Sandler *Distance formulas in complex hyperbolic space*, Forum Mathematicum **8 (1)** (1996), 93–106.

[27] G.P. Scott, *The geometries of* 3*-manifolds*, Bull. London Math. Soc. **15** (1983), 401–487.

[28] I. Stewart, Galois Theory, Chapman and Hall, 1973.