Complex Hyperbolic Geometry

William M. Goldman

September 25, 1998

PREFACE

This book attempts a fairly comprehensive treatment of the geometry of complex hyperbolic space and its boundary. This subject's richness is enhanced by the confluence of many fields of mathematics: Riemannian geometry, complex analysis, symplectic and contact geometry, Lie theory, harmonic analysis and ergodic theory. The boundary of complex hyperbolic geometry is spherical CR geometry or *Heisenberg geometry.* Many treatments of analysis on bounded domains, Kähler manifolds and analysis on the Heisenberg group currently exist in the literature, but there does not seem to be a comprehensive treatment of the *geometry* of complex hyperbolic space or its boundary.

Largely motivated by applications to geometric structures, moduli spaces and discrete groups, this book does not attempt a thorough discussion of any of these topics. Nor does it attempt a thorough treatment of the analytic aspects listed above. Instead, this book is a "user's guide" to complex hyperbolic geometry, which I hope will stimulate research in this fascinating and important geometry.

This project began as the twin sibling of a computer program. In the early summer of 1988, Mark Phillips, Robert Miner and I began writing an interactive graphics program (called "HEISENBERG") for investigating discrete subgroups acting on complex hyperbolic 2-space. Quickly we discovered that the literature contained many different conventions concerning coordinates on complex hyperbolic space. (For example, we normalize the holomorphic sectional curvature of complex hyperbolic space to be -1, in which case the sectional curvatures κ range in the interval

$$-1 \leq \kappa \leq -1/4.$$

This differs from Epstein [48] where the sectional curvature lies between -1 and -4 and from Mostow [128] where the curvature lies between $-1/2$ and -2. Unfortunately, the range $-4 \leq \kappa \leq -1$ seems to be the most popular.) Computers, like humans, are not fond of inconsistent mathematical formulas. Therefore we must establish all of the formulas correctly once and for all. With an internally consistent exposition, we rest assured that the bugs in our programs are caused by our own stupidity and not by inconsistent formulas from the literature.

Several papers influenced my thinking at an early stage: G. D. Mostow's paper "On a remarkable class of polyhedra in complex hyperbolic space" ([128]) contains the first geometric construction of nonarithmetic lattices acting on complex hyperbolic space. The direct geometric techniques of this paper are very much in the spirit of this book. Élie Cartan's paper "Sur le groupe de la géométrie hypersphérique" ([21]) seems to be the first source on the synthetic Heisenberg geometry. Domingo Toledo's "Representations of surface groups on complex hyperbolic space" ([163]) applies complex hyperbolic geometry to representation theory of discrete groups and flat bundles over Riemann surfaces. As I became more familiar with these papers, the challenge of proving the results of one paper using techniques from the others became irresistible. For example, Mostow's paper described the geometric structure of bisectors (formerly called "equidistant" or "spinal" hypersurfaces) in terms of a decomposition into complex hyperplanes; our first computer experiments revealed the other foliation by totally real totally geodesic subspaces. This observation led to a short proof of one of the theorems of Cartan's paper [21] (see Theorem 5.3.4). This theorem beautifully examplifies synthetic Heisenberg geometry. Cartan's book [24] on the geometry of complex projective 3-space also influenced the evolution of the viewpoint here, as did the paper [19] of Burns and Shnider on spherical CR-structures and the paper [105] of Korányi–Reimann on the complex cross-ratio.

In the summer of 1993, my thinking on this subject changed significantly when Ossip Shvartsman introduced me to the (apparently neglected) work of Georges Giraud from 1915 to 1921. Strongly influenced by Picard, Giraud developed much of the theory of bisectors from a point of view very close to the present one. Indeed, Mostow's decomposition cited above may be found in Giraud's 1921 paper [65]. Furthermore Giraud proved a strong uniqueness theorem for intersections of bisectors (Theorem 8.3.3) which has strong consequences (as Giraud observes) for the structure of Dirichlet fundamental domains. Giraud's results are reproved and extended here. The complex-projective theory of *extors*, or "extended bisectors," is directly motivated by Giraud's ideas.

I have tried to present here a consistent set of formulas so that workers in this subject can perform calculations, both by machine and by hand. I have emphasized the interaction between the Kähler geometry on complex hyperbolic space and its degeneration to conformal Heisenberg geometry on the boundary. This is analogous to the more familiar degeneration of the Riemannian geometry of real hyperbolic space $\mathbf{H}^n_{\mathbb{R}}$ to Euclidean conformal geometry on its ideal boundary. We explore various aspects of the geometry: the structure of the automorphism group, the geometry of totally geodesic subspaces, chains, bisectors, etc. Bisectors play a central role here as do the spheres bounding them, which we call *spinal spheres.* Spinal spheres are ubiquitous: they are the analogue of isometric spheres for automorphisms of complex hyperbolic space, they arise as level sets of the Poisson kernel function, they are unions of boundaries of totally real geodesic subspaces level sets of the 3-point invariants defined by Cartan and Toledo and form faces of Dirichlet–Ford fundamental polyhedra of discrete groups. A key notion in the geometry at infinity is a "calibration" of the CR-structure (more traditionally called "pseudo-Hermitian structure" or "contact 1-form"). Playing a role analogous to conformal Riemannian metrics on the boundary of real hyperbolic space, calibrations of the CR-structure form a unifying bridge between spinal spheres at infinity and the potential theory and the symplectic geometry inside. (For the most part we have followed the conventions of Kobayashi–Nomizu [100] for differential geometry and Korányi–Reimann [106] for Heisenberg CR geometry.)

Although many of the results here are included in more general theorems about Lie groups or manifolds of nonpositive curvature, complex hyperbolic space is such an important special case that it merits individual attention. Complex hyperbolic space is the simplest example of a negatively curved Riemannian manifold not having constant curvature. It and its quotients therefore provide basic examples in Riemannian geometry and dynamical systems. The simplest examples of negatively curved Kähler manifolds, complex hyperbolic manifolds, provide important examples in complex analysis and algebraic geometry. General theories should be understood in terms of nontrivial examples. Therefore I have adopted the philosophy that explicit calculations are more valuable than quoting general theorems. For example, the root-space, Cartan and Bruhat decompositions of $\mathbf{PU}(n,1)$ are worked out explicitly, rather than simply quoted from Lie theory. Similarly, we have worked out detailed formulas for geodesics and orthogonal projections onto totally geodesic subspaces, although their qualitative properties are special cases of general results about Hadamard manifolds. This has led to a rather algebraic approach, since I succumbed to the temptation of trying to write everything down explicitly. This luxury is unavailable in a more general setting. The result is a rather long preliminary section, including such nonstandard prerequisites as real structures on complex vector spaces, Hermitian vector "cross" products, outer products and triple Hermitian products.

I have adopted the general point of view of symplectic topology, emphasizing the symplectic properties of complex hyperbolic space as they relate to its Riemannian and holomorphic properties. For example, complex hyperbolic space itself is constructed as a symplectic quotient. A whole section is devoted to the Hamiltonian functions on $\mathbf{H}^n_{\mathbb{C}}$ determining 1-parameter groups of automorphisms. Another section deals with the closely related functions arising as contact Hamiltonians which calibrate the CR-structure on $\partial\mathbf{H}^n_{\mathbb{C}}$ and other geometric entities (for example, totally geodesic submanifolds) on $\mathbf{H}^n_{\mathbb{C}}$. Both of these constructions relate to potential theory on $\mathbf{H}^n_{\mathbb{C}}$ as well as metric properties such as totally geodesic subspaces, orthogonal projections and involutions. The rich interplay between function theory and geometry is one unifying theme of this work.

A geometric theory of discrete groups requires appropriate notions of polyhedra, which becomes nontrivial in the absence of totally geodesic hypersurfaces from which to form a polyhedron's "faces." Owing to their ubiquity and many nice properties, bisectors (formerly called "equidistant hypersurfaces" or "spinal surfaces") form a reasonable substitute for totally geodesic hypersurfaces. Since the faces of fundamental polyhedra arising from the general constructions of Dirichlet–Poincaré and Ford are bisectors, it seems most natural to use bisectors as the building blocks for polyhedra in $\mathbf{H}^n_{\mathbb{C}}$. However, bisectors may be too rigid —Schwartz's very recent notion of "hybrid cones" ([153]) may be a much better class of submanifolds from which to build fundamental polyhedra.

For such a theory to be workable, it is necessary to understand how such hypersurfaces intersect. A large part of this text is devoted to this question. We prove that the intersection of two bisectors has two, one or none components. However, in an important case, bisector intersections are connected: if there is a single point in $\mathbf{H}^n_{\mathbb{C}}$ (respectively $\partial\mathbf{H}^n_{\mathbb{C}}$) from which two bisectors $\mathfrak{E}_1, \mathfrak{E}_2$ are "equidistant," we say that $\mathfrak{E}_1, \mathfrak{E}_2$ are *coequidistant* (respectively *covertical*). For example, if $\mathfrak{E}_1, \mathfrak{E}_2$ are the bisectors extending faces of a Dirichlet (respectively Ford) fundamental polyhedron, then they are coequidistant (respectively covertical). We prove that if $\mathfrak{E}_1, \mathfrak{E}_2$ are a pair of coequidistant or covertical bisectors, then $\mathfrak{E}_1 \cap \mathfrak{E}_2$ is connected. A statement equivalent to this is stated in [128], although the proof is flawed.

Bisectors are generalized and extended to real hypersurfaces in complex projective space, called *extors.* Bisectors are metric objects, while extors are projective objects. Furthermore intersections of extors are considerably simpler to classify than bisector intersections, and the projective theory illuminates the metric theory. We prove Giraud's theorem that the there is at most one other bisector containing the intersection of a generic pair of bisectors in the context of extors. As noted by Giraud, this implies that, under very general conditions, cycles of side-pairing transformations of Dirichlet-Ford fundamental domains have length at most 3. This phenomenon seems "generic," unlike the familiar cases of constant curvature, when side-pairing cycles can be arbitrarily long.

Summary

Chapter 1 reviews the one-dimensional complex geometry of from a point of view which generalizes to higher dimensions. This material—complex projective, elliptic and hyperbolic geometry—should be quite familiar to most readers. Most of the proofs are given as exercises. However, this introductory chapter sets the stage for the more complicated general theory and hopefully motivates some of the later constructions.

Chapter 2 develops the algebraic, analytic and geometric prerequisites needed later. The linear algebra needed in the sequel requires particular attention to the relation between complex and real vector spaces (§2.1). Hermitian structures on complex vector spaces relate to orthogonal and symplectic structures on the underlying real vector spaces (§2.2). In particular real subspaces are studied, and the principal invariant of a real subspace of a Hermitian vector space—the angle of holomorphy—is developed in detail (§2.2.2). Hermitian triple products—an invariant of triples of vectors in a Hermitian product which leads to Cartan's "invariant angulaire"—are introduced (§2.2.5). Hermitian outer products (§2.2.6) and cross-products (§2.2.7) are useful calculational tools and are introduced in this section.

The general point of view is heavily influenced by symplectic geometry (§2.3), since one obtains much of the structure of complex projective space by viewing it as a symplectic (or Kähler) quotient. Its momentum map explicitly described by Hermitian outer products. The basic differential operators ∂ and $\bar{\partial}$ on a complex manifold (§2.4) are used to construct Kähler potentials (§2.4.1), and also to deduce special properties of curvature in the presence of a Kähler structure

(§2.4.2). The boundary of a symplectic manifold enjoys *contact geometry* (§2.5) and the boundary of a complex manifold enjoys *CR geometry* (§2.5.4). A *contact structure* is a field of tangent hyperplanes and is defined by 1-forms, which *calibrate* the contact structure. Geometric objects in $\mathbf{H}_{\mathbb{C}}^n$ correspond to special families of calibrations of the CR-structure on $\partial\mathbf{H}_{\mathbb{C}}^n$, and in general calibrations relate to defining functions of real hypersurfaces in Kähler manifolds (§2.5.5). An abstract notion of a *Heisenberg space*—a space with a simply transitive action of the Heisenberg group—is defined (§2.6). Explicit contact flows on Heisenberg space are described, and the section concludes (§2.6.4) with a differential operator reminiscent of the "curl" due to Rumin [148, 149].

Chapter 3 begins the study of complex hyperbolic space $\mathbf{H}_{\mathbb{C}}^n$ with its model as the unit ball in $\mathbb{C}^n$. However, the full symmetry of $\mathbf{H}_{\mathbb{C}}^n$ is not apparent until $\mathbb{C}^n$ is completed to complex projective space. Complex reflections are defined, from which the homogeneity of $\mathbf{H}_{\mathbb{C}}^n$ follows. We derive a formula for the midpoint of two points (§3.1.2), a momentum mapping expressing $\mathbf{H}_{\mathbb{C}}^n$ as a symplectic quotient (§3.1.3), a synthetic description of the Bergman metric (§3.1.6), a formula for the Bergman distance function (§3.1.7) and a Kähler potential for the Bergman metric and its relation to the Bergman kernel (§3.1.8).

Complex projective subspaces of $\mathbb{P}_{\mathbb{C}}^n$ meet $\mathbf{H}_{\mathbb{C}}^n$ in *complex-linear* totally geodesic submanifolds (§3.1.4). Such a totally geodesic submanifold is an isometrically embedded copy of a lower-dimensional complex hyperbolic space. In particular a *complex geodesic*—a complex-linear totally geodesic subspace of complex dimension one—has the Poincaré model of hyperbolic geometry. (For example, it is a round disc with geodesics given by semicircular arcs orthogonal to the bounding circle.) The projection of $\mathbb{P}_{\mathbb{C}}^n$ onto a complex-linear totally geodesic subspace defines the *orthogonal projection* in $\mathbf{H}_{\mathbb{C}}^n$ (§3.1.5). A basic property of $\mathbf{H}_{\mathbb{C}}^n$ is that two distinct points span a unique *complex geodesic,* that is a complex-linear totally geodesic subspace of complex dimension one.

The other totally geodesic submanifolds are totally real (§3.1.9). These are isometrically embedded copies of real hyperbolic spaces. In contrast to a complex-linear totally geodesic submanifold, a totally real totally geodesic subspace has the Klein model of hyperbolic geometry. For each k, the automorphism group $\mathbf{PU}(n, 1)$ acts transitively on totally real totally geodesic subspaces of dimension k. Figures 3.4 and 3.3 depict Fermi coordinates on the two types of totally geodesic suspaces. The proof that the only totally geodesic submanifolds of $\mathbf{H}_{\mathbb{C}}^n$ are either totally real or complex-linear is sketched (§3.1.11).

Trigonometry in complex hyperbolic space is developed (§3.2), following Giraud [65], Hsiang [89], Brehm [17], and Leuzinger [111, 112], and earlier work by Blaschke and Terheggen [12, 158]. The cosine and sine laws, as well as the Pythagorean theorem and the formulas for the angle of parallelism, are generalized from real hyperbolic geometry to complex hyperbolic geometry. A new feature here is that associated to a vertex of a triangle are two "angles"—the ordinary Riemannian angle between the geodesic sides—as well as a "complex angle" defined as the angle between the complex geodesics containing the sides.

We view trigonometric formulas in $\mathbf{H}^n_{\mathbb{C}}$ as *deformations* of the trigonometric formulas in real hyperbolic space, where triangles in real hyperbolic space are those triangles which lie in totally geodesic subspaces. All of these formulas have analogues in complex elliptic space. The trigonometric formulas are used to compare real hyperbolic space with complex hyperbolic space (§3.2.5). The effect of the varying sectional curvature is explicitly related to distortion of distance. The angle of holomorphy is interpreted in terms of large triangles (§3.2.6). The *complex altitudes* of a triangle are defined (§3.2.7), and Cartan's theorem ([24]) that the complex altitudes of a triangle are concurrent if and only if the triangle lies in a totally geodesic subspace is proved.

We derive explicit formulas for some of the objects in $\mathbf{H}^n_{\mathbb{C}}$ (§3.3: in particular, formulas for polar vectors of complex hyperplanes (§3.3.1), orthogonal projections (§3.3.2), the exponential map (§3.3.3), the growth rate of volume (§3.3.4) and the geodesic between two points (§3.3.5). While orthogonal projection onto complex-linear totally geodesic subspaces is given by a linear projection, orthogonal projection onto totally real totally geodesic subspaces is considerably more complicated (§3.3.6). Figure 3.8 depicts the image of a real hypersurace (a spinal sphere) under this projection.

Chapter 4 introduces the second projective model of $\mathbf{H}^n_{\mathbb{C}}$, the *paraboloid* or *Siegel domain* model. This model is analogous the upper half plane model of the hyperbolic plane. While the ball model is what $\mathbf{H}^n_{\mathbb{C}}$ looks like "from inside," the paraboloid model is the view "from infinity." Just as the stabilizer of the origin in $\mathbf{H}^n_{\mathbb{C}}$ is represented by the linear unitary group $\mathbf{U}(n)$, the stabilizer of ∞ in the Siegel domain is represented by a group of affine transformations of $\mathbb{C}^n$. The Cayley transform (§4.1.1) relates these two models. Horospheres are introduced as level sets of Busemann functions (§4.1.2). The root-space decomposition of $\mathfrak{su}(n,1)$ leads to a description of the automorphisms of $\mathbf{H}^n_{\mathbb{C}}$ in the Siegel model (§4.1.3) and a Hermitian algebraic description (§4.1.4). From the root-space decomposition derive a very useful set of coordinates—*horospherical coordinates*–on $\mathbf{H}^n_{\mathbb{C}}$ (§4.2.1).

The resulting *Heisenberg geometry* is the conformal CR geometry of the Heisenberg group, the natural geometry on the boundary of $\mathbf{H}^n_{\mathbb{C}}$ (§4.2, §2.6).

Chains—the boundaries of complex geodesics—are described in Heisenberg coordinates as certain ellipses whose vertical projections are Euclidean circles (§4.3.1). We derive a formula for the inversion in a chain (§4.3.2), and use this to describe a moduli space of chains (§4.3.4). Orthogonal projection is described in Heisenberg geometry (§4.3.6) and is used to analyze the complex hyperbolic surfaces which are quotients of Fuchsian groups preserving a chain (§4.3.7). A synthetic geometry of chains "lifts" the synthetic geometry of circles in the Euclidean plane (§4.3.8). Using Heisenberg geometry, we reprove Cartan's theorem ([21]) that a chain-preserving transformation of Heisenberg space is an automorphism (§4.3.9).

The other 1-dimensional objects in Heisenberg geometry are $\mathbb{R}$-circles, the boundaries of totally real totally geodesic 2-planes (§4.4). In Heisenberg geometry, $\mathbb{R}$-circles are either CR-horizontal straight lines (§4.4.3) or CR-horizontal lifts of lemniscates (§4.4.4). Parameters for $\mathbb{R}$-circles analogous to center-radius coordinates for circles and chains are defined, obtaining a moduli space for $\mathbb{R}$-circle(§4.4.8). Unlike chains, infinitely many $\mathbb{R}$-circles pass through a pair of points (or through a point with given horizontal direction) (§4.4.11). Discrete groups whose limit set is an $\mathbb{R}$-circle are described (§4.4.13) and depicted in Figures 4.10–4.12.

Chapter 5 develops the theory of bisectors and spinal spheres. Bisectors enjoy two foliations by totally geodesic submanifolds (§5.1) and their boundaries—*spinal spheres*—enjoy corresponding foliations by hyperchains and $\mathbb{R}$-spheres. Any two bisectors in $\mathbf{H}^n_{\mathbb{C}}$ are equivalent by an element of $\mathbf{PU}(n, 1)$, and in a very precise sense bisectors (real codimension one) are *dual* to geodesics (real dimension one) (§5.1.4). Following Mostow [128], the geodesic associated to a bisector or spinal sphere is called its *spine.* Since a geodesic is determined by its pair of endpoints on $\partial\mathbf{H}^n_{\mathbb{C}}$, a bisector or spinal sphere is determined by its spine's endpoints, which we call its *vertices.*

Various examples of spinal spheres are computed in Heisenberg coordinates. The horizontal plane $v = 0$ is perhaps the simplest spinal sphere (§5.1.7) and has vertices at the origin and ∞. More generally, spinal spheres with one vertex at ∞ are the The *unit spinal sphere* defined by

$$||\zeta||^4 + v^2 = 1$$

(and depicted in Fig. 5.3) does not contain ∞ (§5.1.8). Applying Heisenberg similarities to this spinal sphere, one obtains the class of *vertical spinal spheres,* characterized as those whose complex spine is a vertical chain (§5.1.9). These spinal spheres are metric spheres with respect to a useful metric on Heisenberg space first defined by Cygan [37] (§5.1.9). *Great spinal spheres* play a role analogous to great circles in one-dimensional geometry (§5.1.10). *Infinite spinal spheres* are spinal spheres passing through ∞ (but not having ∞ as a vertex) (§5.1.12); such spinal spheres are depicted in Figs. 5.5–5.8 and Fig. 5.4.

The group of automorphisms of a given bisector $\mathfrak{E}$ one has principal orbits of codimension one (§5.2). The inversion in each slice preserves $\mathfrak{E}$, interchanging its vertices (§5.2.2). Similary $\mathfrak{E}$ is invariant under inversion in each meridian (§5.2.3). Analogous to the fact that two distinct points in $\mathbf{H}^n_{\mathbb{C}}$ determine a unique geodesic, two ultraparallel hyperplanes are slices of a unique bisector (§5.2.4). The slices of $\mathfrak{E}$ are characterized as those hyperplanes whose inversions interchange the vertices of $\mathfrak{E}$ (§5.2.5). The meridians of a spinal sphere are also characterized in terms of inversions in slices, leading to lemmas on tangent $\mathbb{R}$-circles (§5.2.6) which are used in the proof that some bisector intersections are connected.

How bisectors intersect is a principal theme in this work. The simplest case arises when the spines lie in the same complex geodesic (§5.3.1). Another important case arises when two bisectors contain a common meridian (§5.3.5). Particularly interesting is when the spines are ultraparallel geodesics in a totally real totally geodesic subspace P; then the bisectors intersect in a union of P with a complex geodesic. These pairs of bisectors are used in a new proof of Cartan's theorem ([21]) on a remarkable configuration of 7 $\mathbb{R}$-circles and 1 chain (§5.3.4). Following Cartan, this result implies that an $\mathbb{R}$-circle-preserving transformation of Heisenberg space must be an automorphism (§5.3.3).

Spinal spheres also arise analytically as level sets of functions related to calibrations of the CR-structure (§5.4). If ω is a 1-form calibrating the CR-structure at ∞ and g is an automorphism, then the set $\mathcal{S}$ of points for which $g^*\omega = \omega$ is defined as the *isometric sphere* of g with respect to ω. When ω is the calibration associated to a point O in $\mathbf{H}^n_{\mathbb{C}}$ (corresponding to the Kähler potential associated to O (§3.1.8)) then $\mathcal{S}$ is the spinal sphere bounding the bisector equidistant from O and $g^{-1}(O)$. Since such bisectors arise in the construction of Dirichlet fundamental domains, we call such spinal spheres *Dirichlet isometric spheres* (§5.4.2). These spinal spheres arise as the level sets of the Poisson kernel function (§5.4.3). As O tends to ∞, the Dirichlet isometric spheres converge to spinal spheres analogous to the isometric circles defined classically by Ford, and we call these spinal spheres *Ford isometric spheres* (§5.4.4). The association of a Ford isometric sphere to an automorphism closely relates to the Bruhat decomposition of $\mathbf{SU}(n,1)$ (§5.4.5). Just as for interior points and boundary points, complex-linear totally geodesic subspaces (§5.4.6) and totally real totally subspaces (§5.4.7) determine calibrations.

Although bisectors are not totally geodesic, their two foliations by totally geodesic submanifolds severely constrain their local geometry. The horospherical coordinates developed in §4.2.1 define "geographical coordinates" on a bisector, in which many of the components of the second fundamental form vanish (§5.5). All of these results have analogues in complex elliptic space (§5.5.4). Finally, a bisector is highly nonconvex, and its geographical coordinates describe qualitatively the regions containing geodesics with endpoints on it (§5.5.5).

Chapter 6 further pursues the automorphisms of $\mathbf{H}^n_{\mathbb{C}}$ and divides into two independent parts. The first part (§6.1) studies $\mathfrak{su}(n,1)$ from the point of view of the symplectic geometry of $\mathbf{H}^n_{\mathbb{C}}$. In particular, one-parameter subgroups of $\mathbf{PU}(n,1)$ define Hamiltonian flows (using the momentum map calculations from §2.3.1). The prototype of this is the explicit symplectic duality between bisectors and geodesics (§6.1.3), which leads to distance formulas for geodesics, totally geodesic submanifolds and bisectors (§6.1.4, §6.1.5).

The second part of Chapter 6 concerns conjugacy classes in $\mathbf{SU}(2,1)$. Algebraic calculations in $\mathbf{SL}(2,\mathbb{C})$ involve the trace function $\mathbf{SL}(2,\mathbb{C}) \to \mathbb{C}$; generically two elements of $\mathbf{SL}(2,\mathbb{C})$ are conjugate if and only if their traces agree. The trace function $\mathbf{SU}(n,1) \to \mathbb{C}$ satisfies analogous properties (§6.2.3), and provides an analogous trichotomy into elliptic, parabolic and hyperbolic conjugacy classes (§6.2.1). As for $\mathbf{SL}(2,\mathbb{R})$—but not $\mathbf{SL}(2,\mathbb{C})$—nonelementary groups containing no elliptic elements are discrete (§6.2.2).

Chapter 7 discusses three important numerical invariants. The angular invariant of Cartan [21] parametrizes triples of ideal points (§7.1) and admits a geometric interpretation as an area of an orthogonal projection of a triangle

(§7.1.2). Ideal triples are fully symmetric and have a unique *barycenter* (§7.1.3). Cartan's invariant relates to the characteristic class of flat bundles (§7.1.4) studied by Toledo [163]. This invariant also relates to a notion of "parallel transport" in Heisenberg geometry (§7.1.5) and geodesic projections (§7.1.6). Ideal triangle groups are also parametrized by the angular invariant (§7.1.7).

Generalizing the ordinary cross-ratio, Korányi and Reimann introduced in [105] a complex number associated to an ordered 4-tuple in Heisenberg space. We describe this invariant (§7.2) from a somewhat different point of view, interpret it geometrically (§7.2.1), determine the meaning of its reality (§7.2.2), and describe a general setting for invariants of its type (§7.2.3).

Pairs consisting of a real geodesic γ and a complex hyperplane c are parametrized in terms of a similarly definied invariant denoted $\eta(q_1, q_1; c)$ where q_1 and q_2 are the endpoints of γ (§7.3). This invariant relates to Cartan's invariant (§7.3.2), to the distance $\rho(\gamma, c)$ (§7.3.3) and to the distance $\rho(\gamma_{\mathbb{C}}, c)$ where $\gamma_{\mathbb{C}}$ is the complex geodesic containing γ (§7.3.4). The orthogonal projection $\Pi_c(\gamma)$ of γ onto c is an arc of a hypercycle, whose length and curvature are functions of $\eta(q_1, q_1; c)$ (§7.3.5). Similarly, as observed by Mostow [128], bisectors meet complex geodesics in hypercycles, whose geodesic curvature can be computed in terms of this invariant (§7.3.6). The condition that the bisector $\mathfrak{E}$ with spine γ meets c is equivalent to the condition that the complex number $\eta = \eta(q_1, q_1; c)$ lies inside the parabolic region $\mathfrak{P}$ of $\mathbb{C}$ defined by

$$\mathrm{Im}(\eta)^2 + 2\mathrm{Re}(\eta) < 1.$$

η is real precisely when $\Pi_c(\gamma)$ is a geodesic, or equivalently when $\gamma \cap c$ is a geodesic (§7.3.7). The distance $\rho(\mathfrak{E}, c)$ is a function of η (§7.3.8) and η admits an expression by complex cross ratios (§7.3.9).

The invariant η was developed to study intersections of bisectors. Convexity of orthogonal projections of bisectors onto complex geodesics closely relates to the the connectedness of bisector intersections (§7.3.10). Examples of these projections are depicted in Fig. 7.4–7.9.

Chapter 8 begins the general theory of *extors* in $\mathbb{P}^n_{\mathbb{C}}$. Extors extend and generalize metric bisectors in both complex hyperbolic and elliptic geometry just as circles in $\mathbb{P}^1_{\mathbb{C}}$ extend and generalize geodesics in $\mathbf{H}^1_{\mathbb{C}}$ and $\mathbf{E}^1_{\mathbb{C}}$ (§8.1). Extors are not smooth, but their singular strata under homogeneous (§8.2). Although defined in terms of a "slice decomposition," they also admit a "meridian decomposition," generalizing the two decompositions of bisectors in $\mathbf{H}^n_{\mathbb{C}}$ and $\mathbf{E}^n_{\mathbb{C}}$ (§8.2.3). Pairs of extors are classified into 4 basic types (§8.3) which organize the more complicated intersection phenomena of metric bisectors. In complex dimension two, extors generically intersect in *Clifford tori* (§8.3.4). We reprove Giraud's theorem [65] that for such a generic pair of extors $\mathfrak{E}_1, \mathfrak{E}_2$, there is a unique extor $\mathfrak{E}_3 \neq \mathfrak{E}_1, \mathfrak{E}_2$ containing $\mathfrak{E}_1 \cap \mathfrak{E}_2$ (§8.3.5). In higher dimensions, a generic pair of extors will have transverse foci (§8.3.7) but such pairs will not arise from Dirichlet-Ford constructions.

Chapter 9 develops the theory of bisector intersections in $\mathbf{H}^n_{\mathbb{C}}$. Using ideas of Mostow [128], each component of a bisector intersection is contractible. Examples of disconnected bisector intersections are constructed (§9.1.2) and interpreted in terms of orthogonal projections of bisectors (§9.1.3). Associated to a bisector and a complex geodesic is a component of a hyperbola of invariants η. The intersection of this curve with $\mathfrak{P}$ is the invariant used to prove that bisector intersections have at most two components (§9.1.4).

The proof that that intersections of coequidistant/covertical pairs is connected is a continuity argument (§9.2). First, tangencies between spinal spheres are analyzed using Heisenberg geometry (§9.2.1). Then a coequidistant or covertical pair is deformed to a pair of bisectors with orthogonal complex spines, where the intersection can be analyzed explicitly (§9.2.4). The basic transversality follows from the the bisectors not having a slice in common. The "cotranchal" case—when two bisectors possess a common slice— must be handled separately (§9.2.5). These properties of bisectors constrain the combinatorics of Dirichlet-Ford fundamental polyhedra (§9.3). Giraud's theorem implies that cycles of side-pairing transformations have length at most 3, and the above connectedness theorem implies that for some discrete subgroups of $\mathbf{PO}(2,1)$ the combinatorics of the Dirichlet-Ford polyhedra in $\mathbf{H}^2_{\mathbb{C}}$ correspond to the combinatorics in $\mathbf{H}^2_{\mathbb{R}}$. A summary of Giraud's paper [65] is appended for the reader's convenience.

CONTENTS

SOME PICTURES

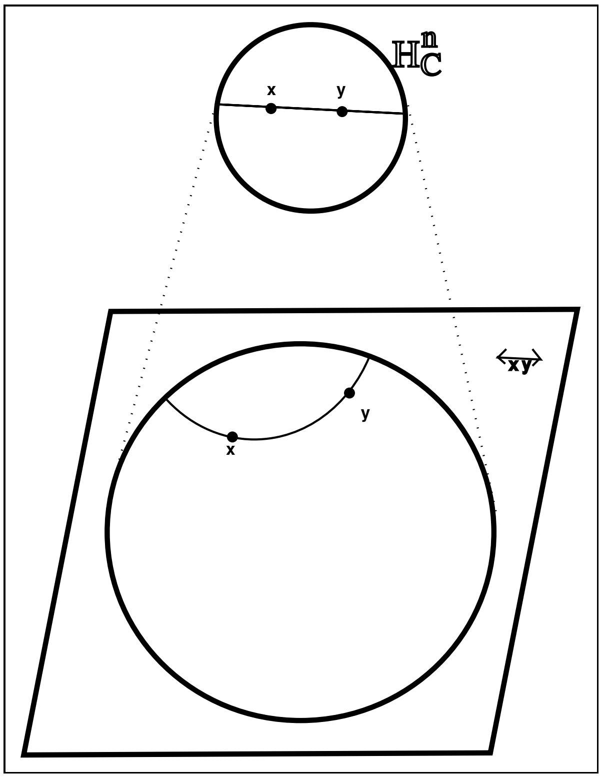

Fig. 0.1. The Bergman metric

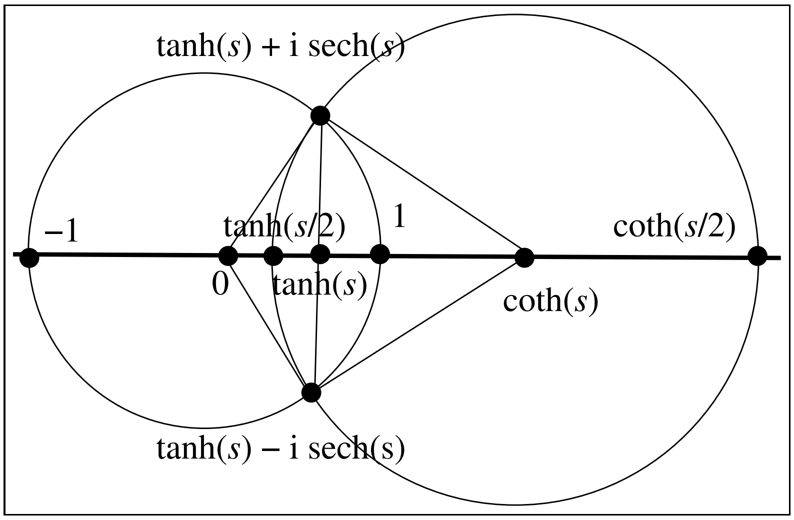

FIG. 0.2. Metric relations for the unit circle and an orthogonal circle

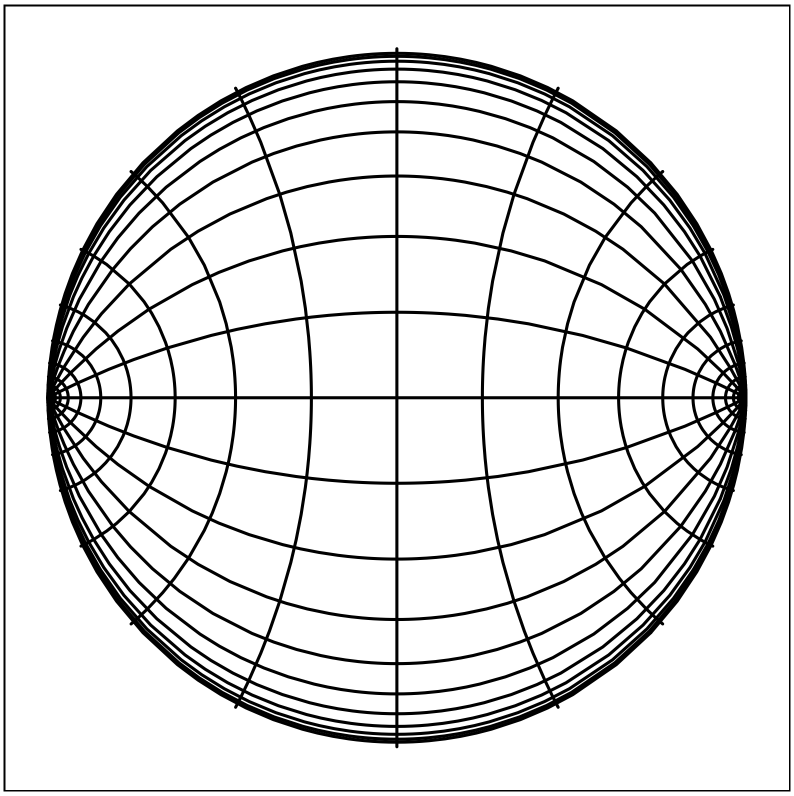

FIG. 0.3. Fermi coordinates in a complex geodesic

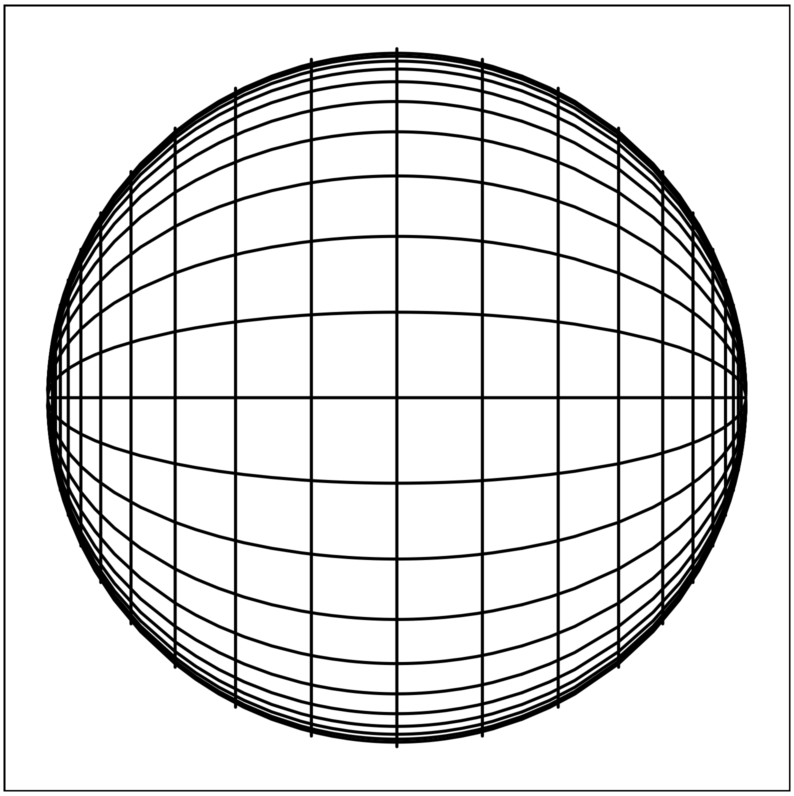

Fig. 0.4. Fermi coordinates in a totally real 2-plane

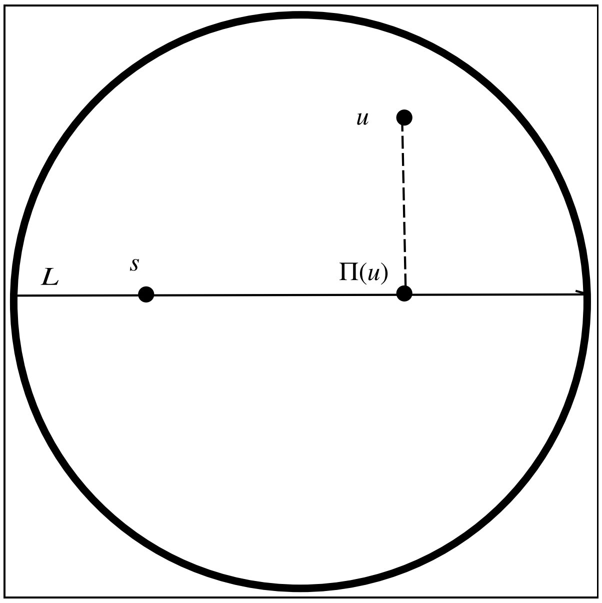

FIG. 0.5. Proof of Lemma 3.2.12

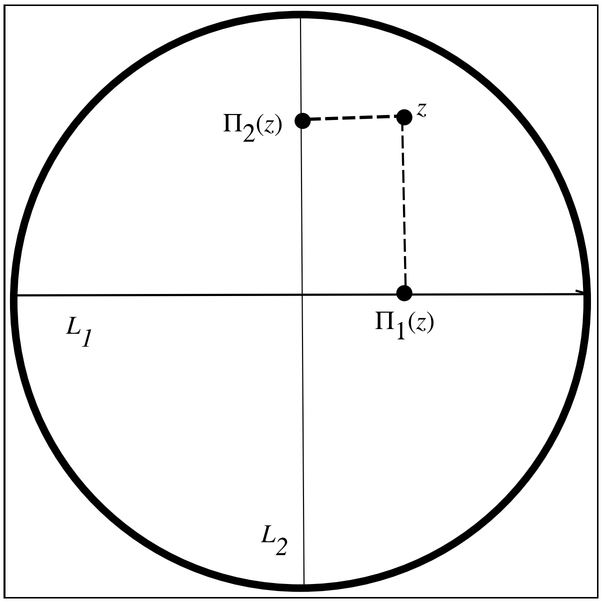

FIG. 0.6. Proof of Lemma 3.2.13

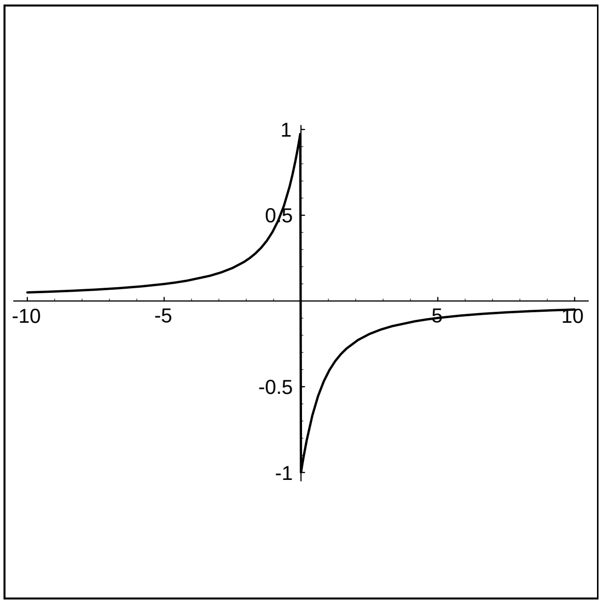

FIG. 0.7. Graph of $\beta = \pm(|\alpha| - \sqrt{\alpha^2 + 1})$

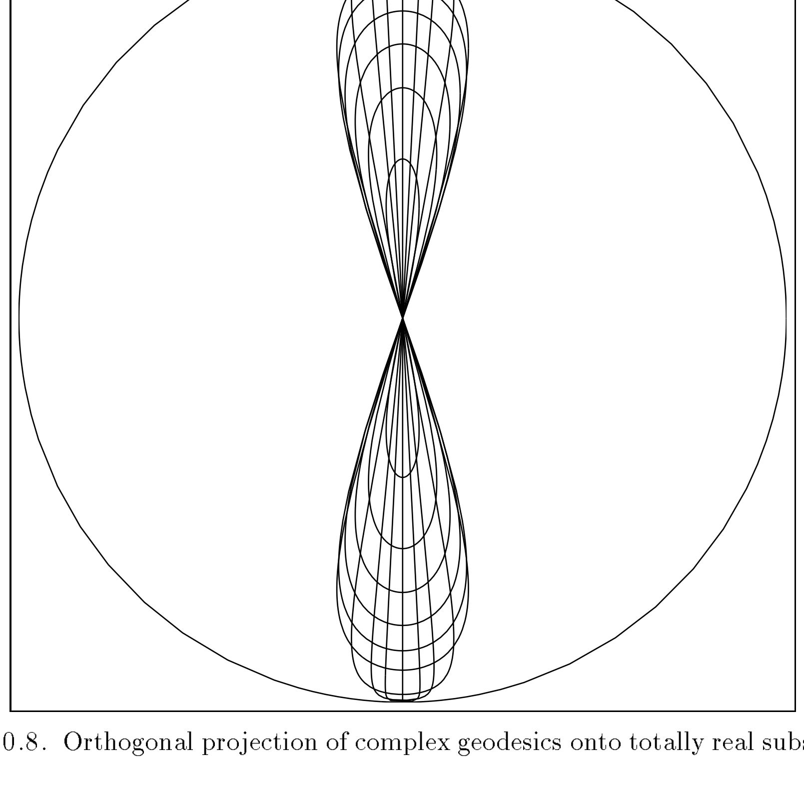

FIG. 0.8. Orthogonal projection of complex geodesics onto totally real subspace

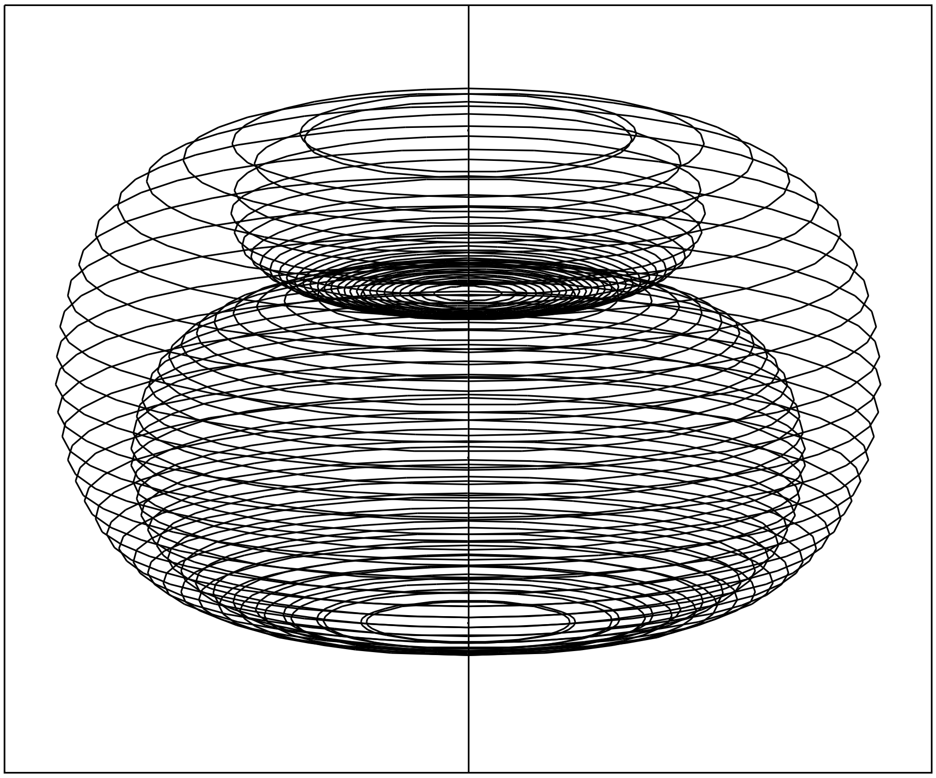

FIG. 0.9. Chains orthogonal to vertical axis

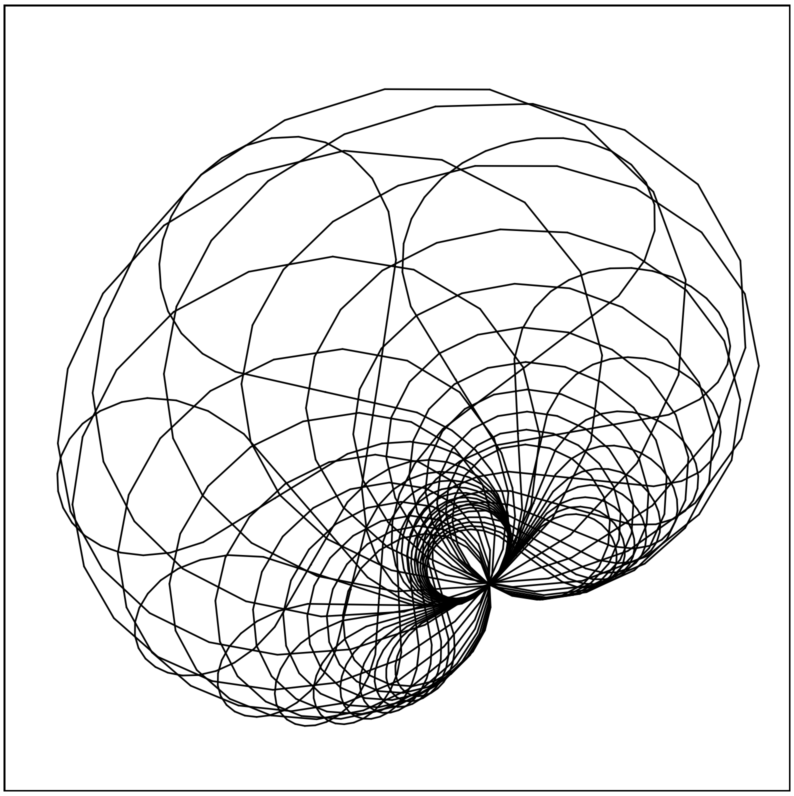

Fig. 0.10. Fan with vertex at the origin

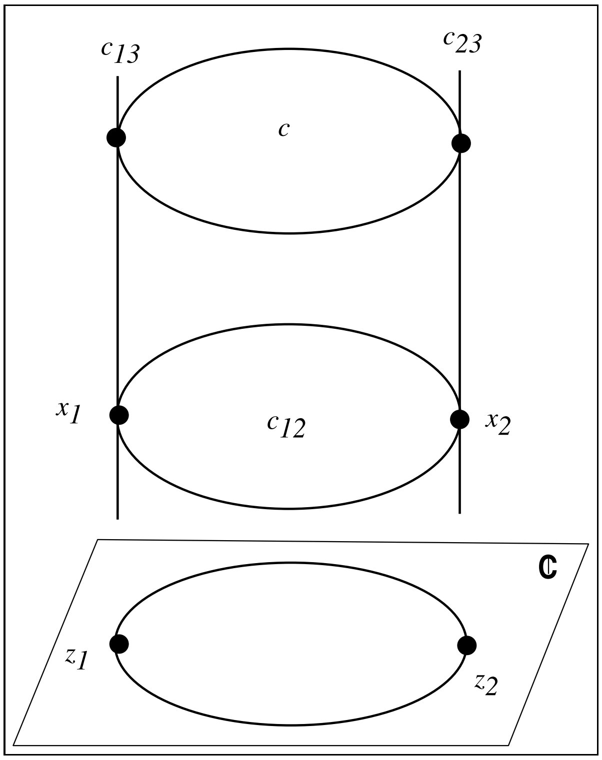

FIG. 0.11. Sketch of proof of Theorem 4.3.12

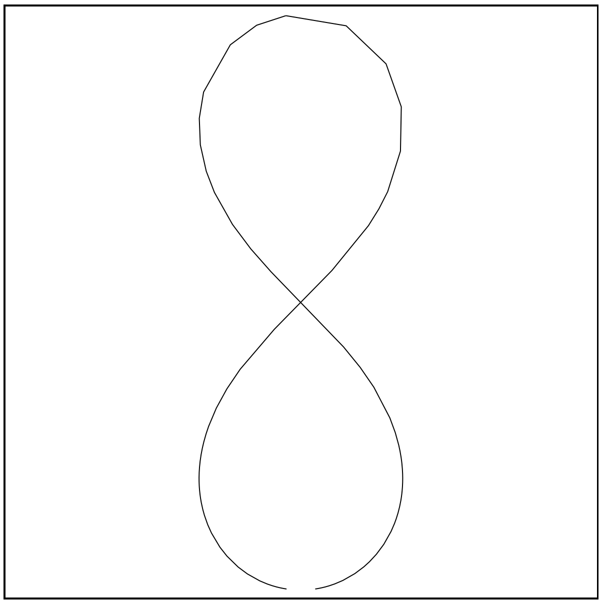

FIG. 0.12. Vertical projection of the purely imaginary $\mathbb{R}$-circle

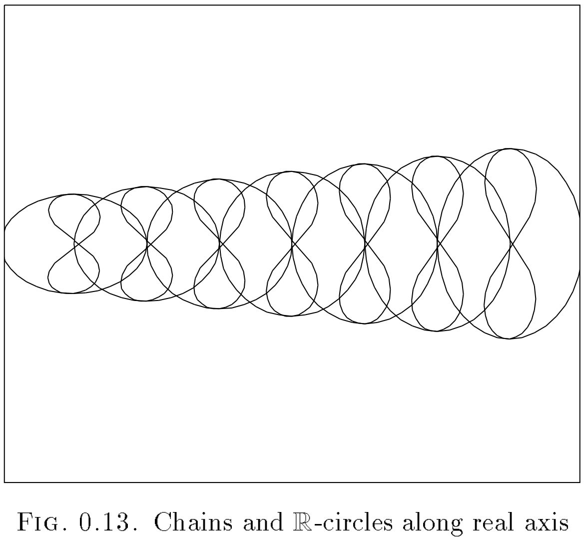

FIG. 0.13. Chains and $\mathbb{R}$-circles along real axis

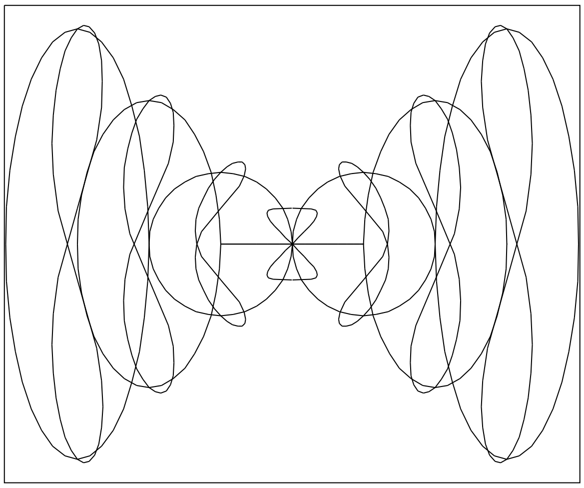

Fig. 0.14. Chains and $\mathbb{R}$-circles along real axis: another view

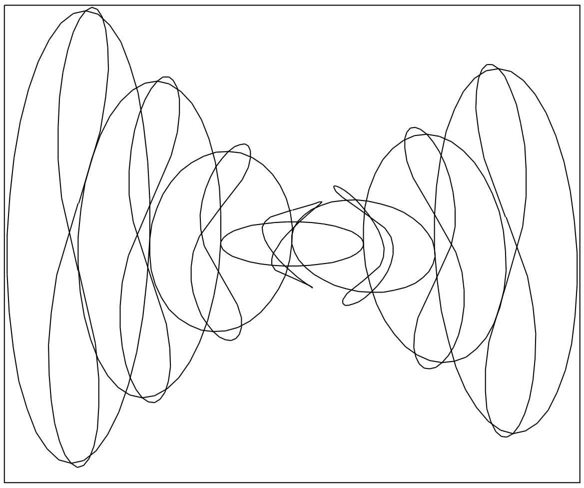

Fig. 0.15. Chains and $\mathbb{R}$-circles along real axis: yet another view

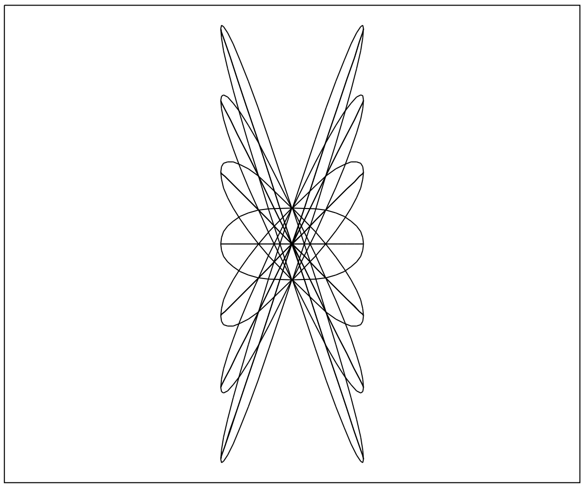

FIG. 0.16. Chains and $\mathbb{R}$-circles: still another view

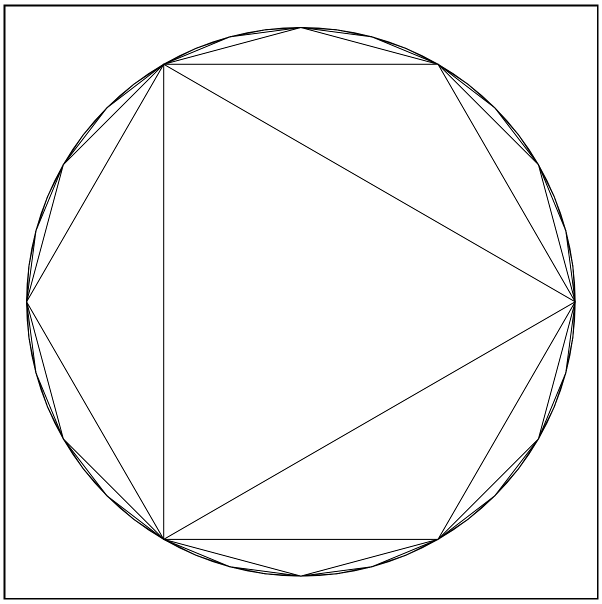

FIG. 0.17. Tiling $\mathbf{H}^2_{\mathbb{R}}$ by ideal triangles

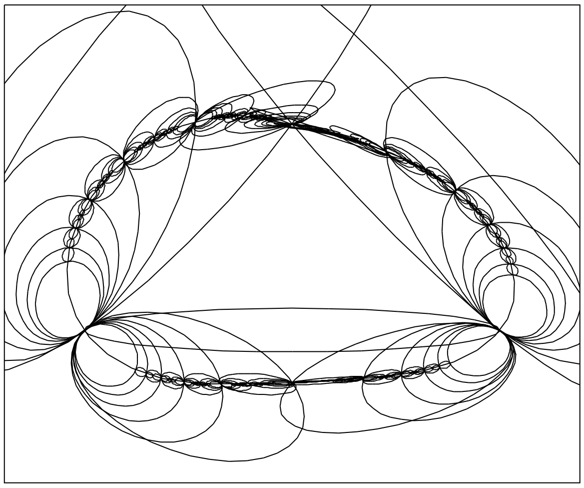

FIG. 0.18. An ideal triangle group in $\mathbf{PO}(2,1) \subset \mathbf{PU}(2,1)$

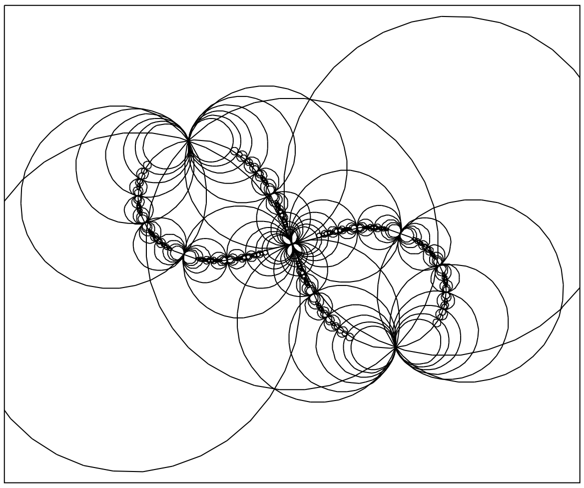

FIG. 0.19. An ideal triangle group: top view

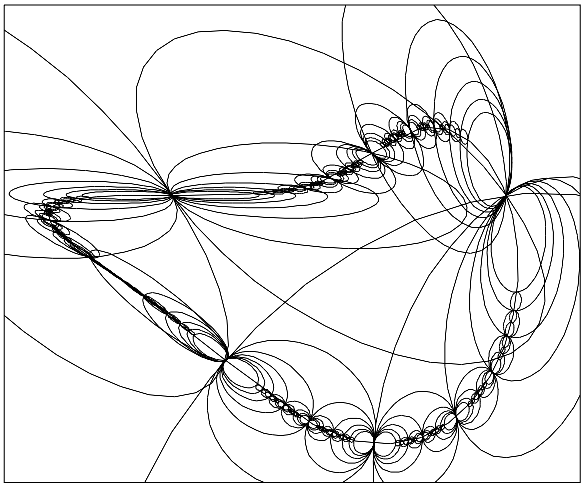

FIG. 0.20. An ideal triangle group: generic view

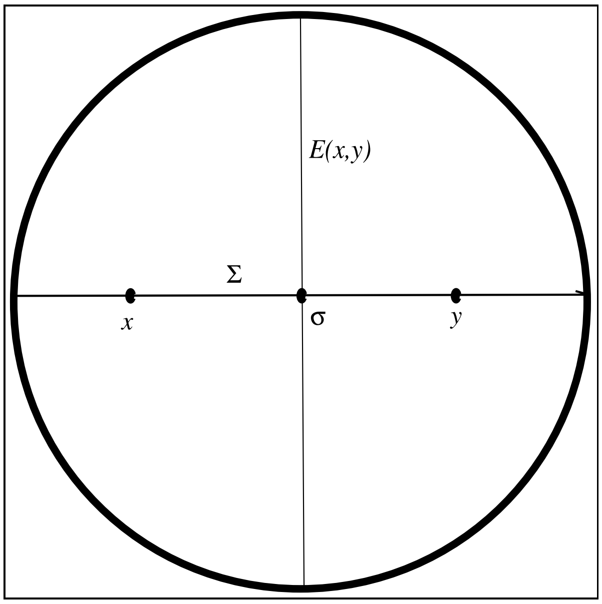

Fig. 0.21. The slice decomposition of a bisector

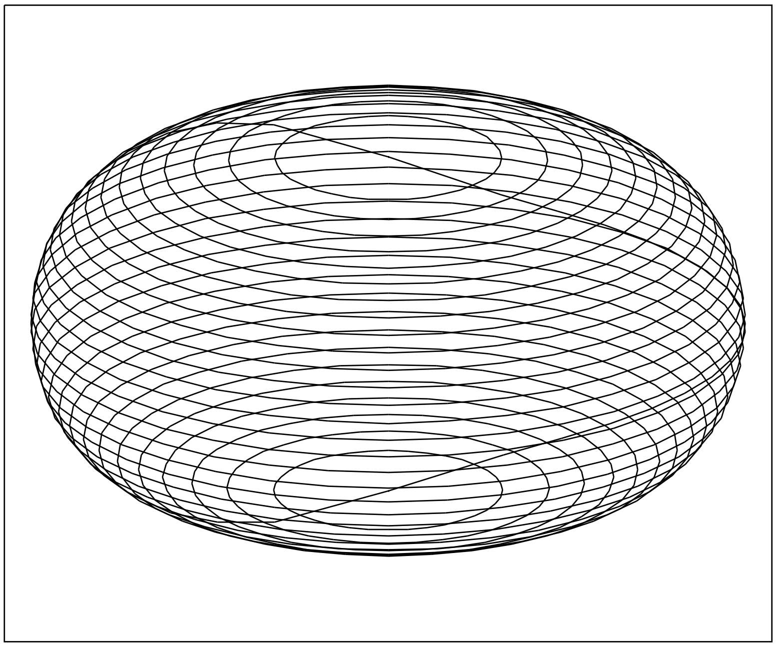

Fig. 0.22. A meridian and several slices of the unit spinal sphere

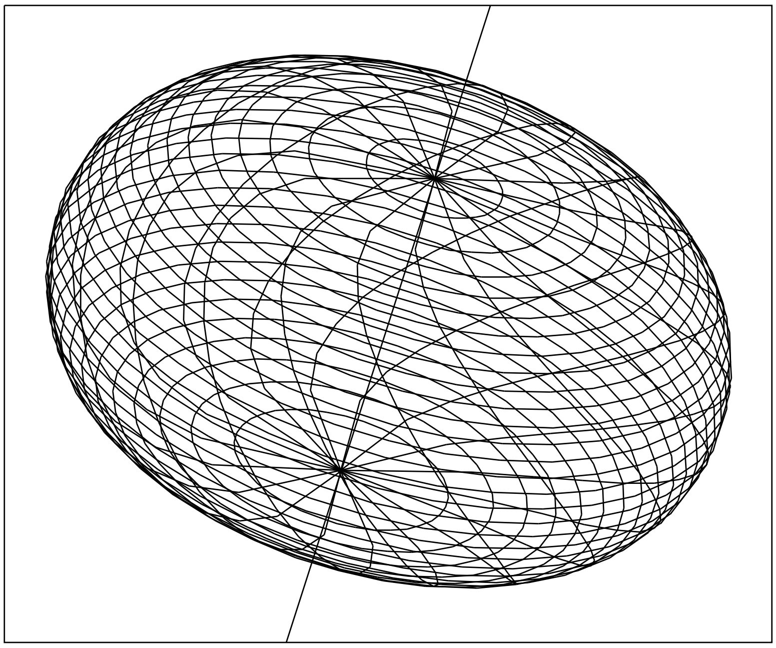

FIG. 0.23. Slices, meridian and complex spine of a vertical spinal sphere

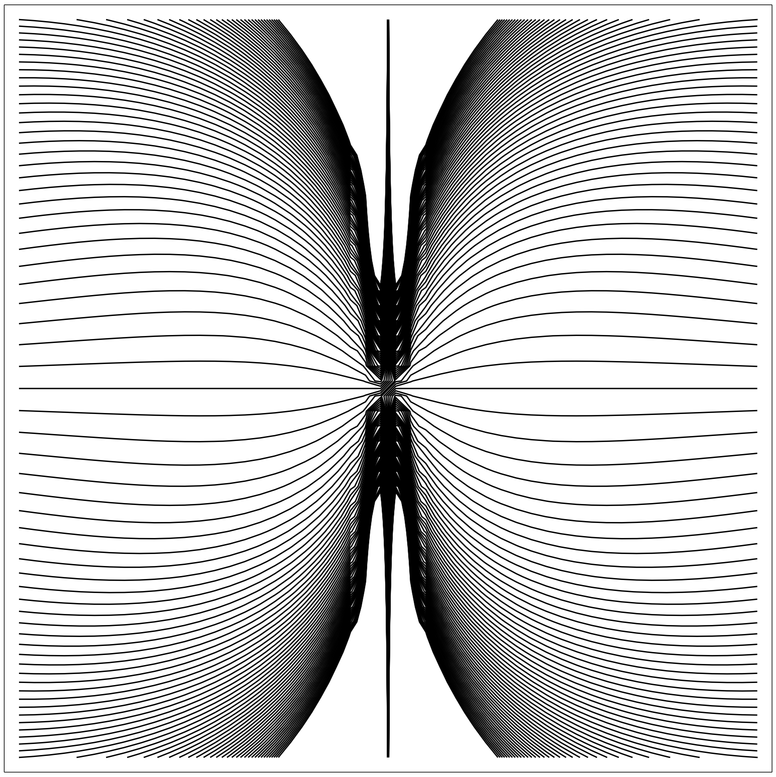

Fig. 0.24. Level sets of function defining spinal with vertices $(\pm 1, 0)$

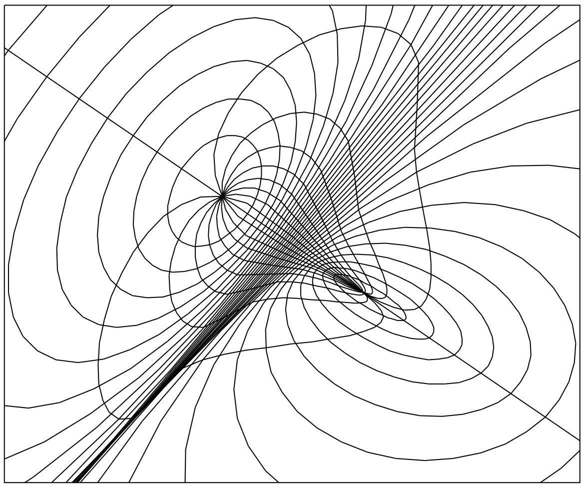

Fig. 0.25. A spinal sphere passing through infinity: another view

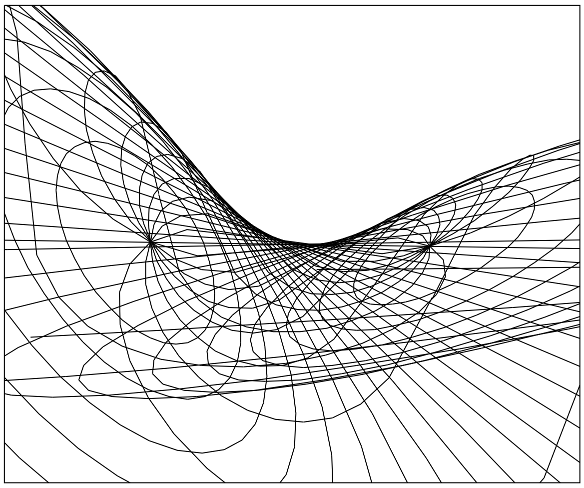

FIG. 0.26. A spinal sphere passing through infinity

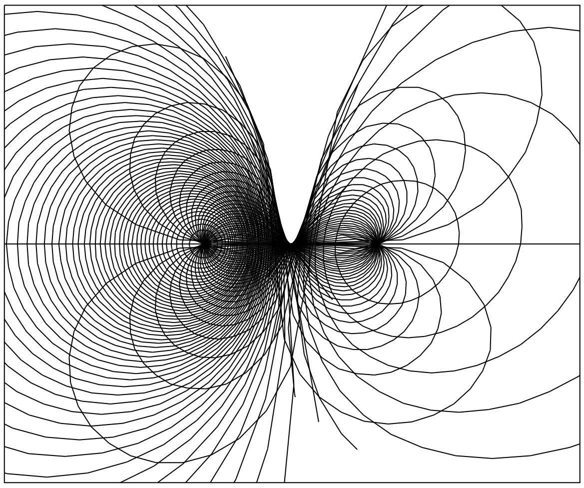

FIG. 0.27. Top view of a spinal sphere with real axis meridian

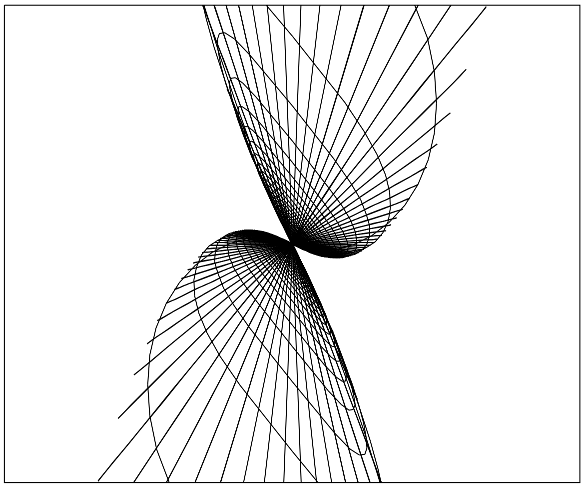

Fig. 0.28. Side view of a spinal sphere with real axis meridian

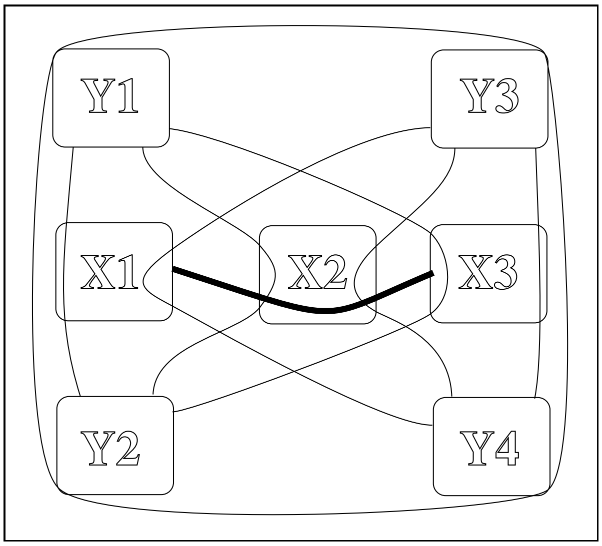

Fig. 0.29. Cartan's configuration: a schematic diagram

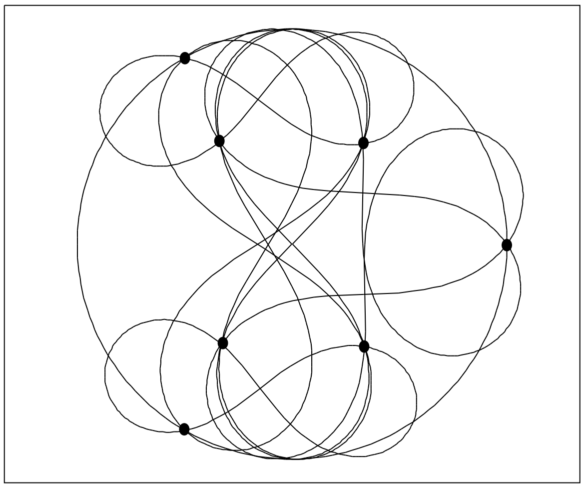

FIG. 0.30. Vertical view of Cartan's configuration

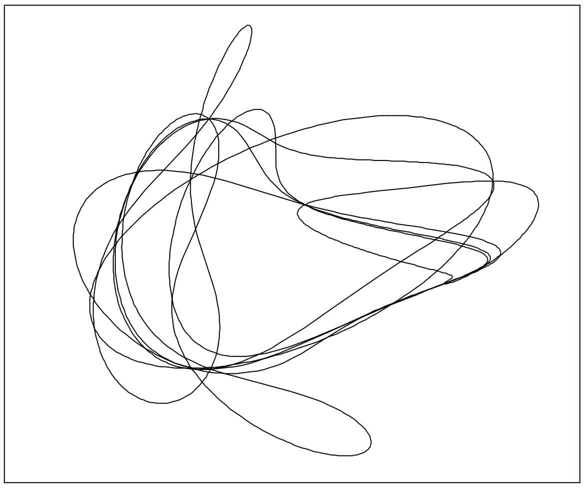

FIG. 0.31. Generic view of Cartan's configuration

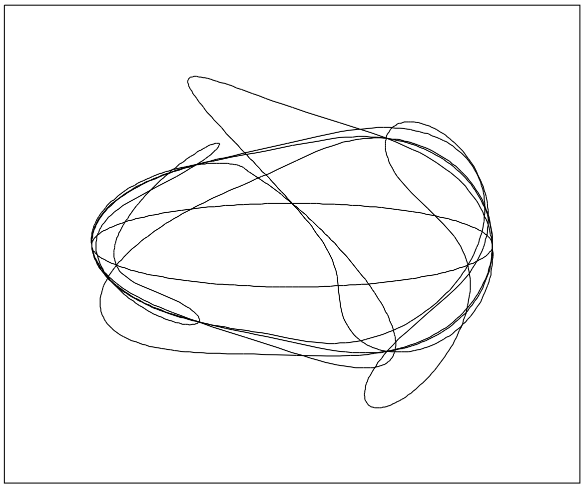

FIG. 0.32. Side view of Cartan's configuration

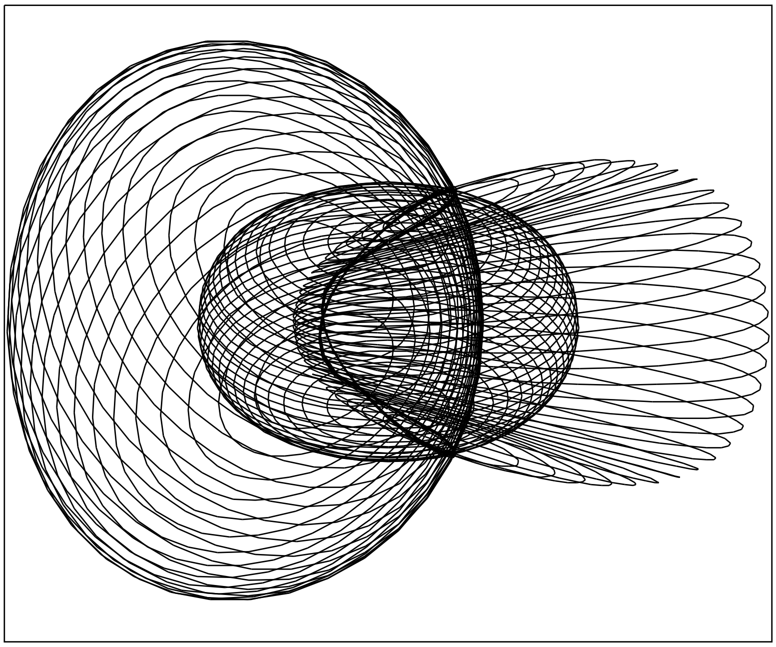

FIG. 0.33. Spinal spheres with a common meridian

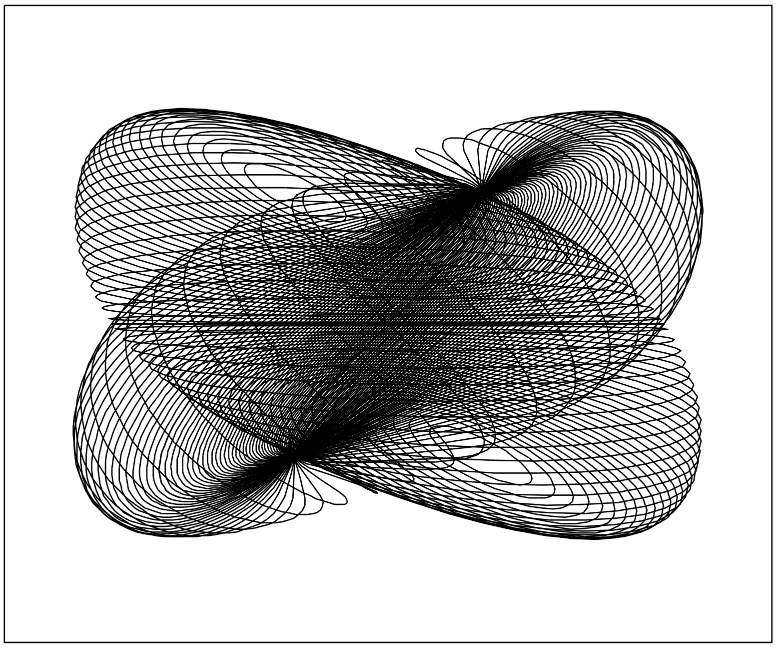

FIG. 0.34. A pair of spinal spheres with both a common slice and meridian

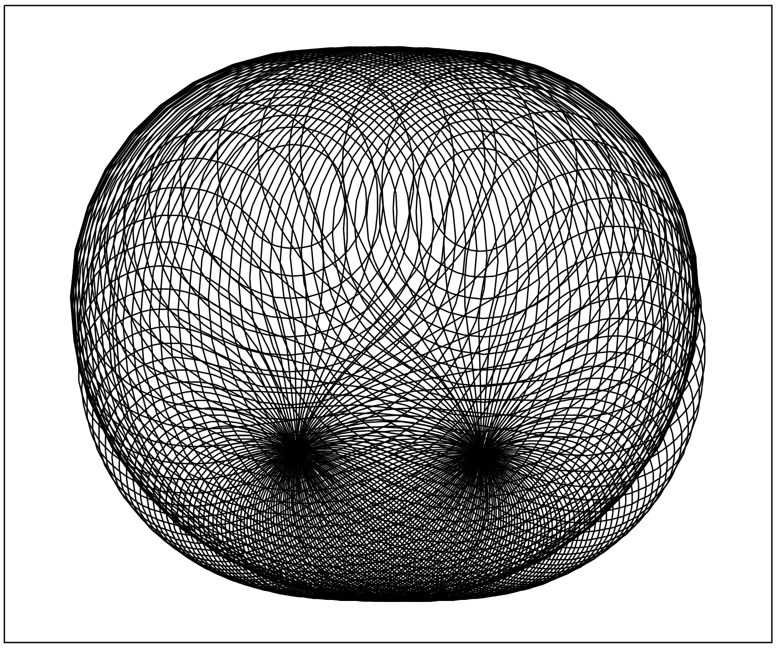

FIG. 0.35. Vertical projection of pair of spinal spheres with common slice and meridian

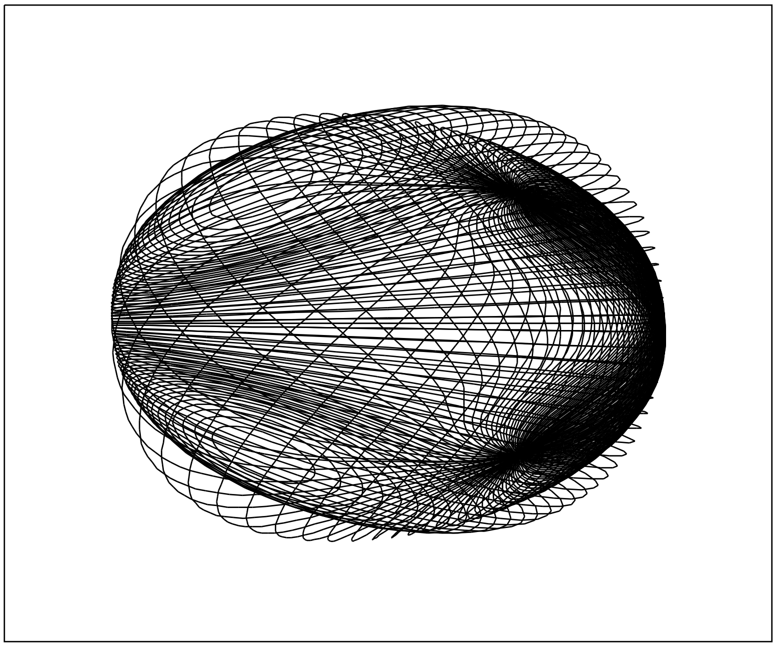

FIG. 0.36. Another view of pair of spinal spheres with common slice and meridian

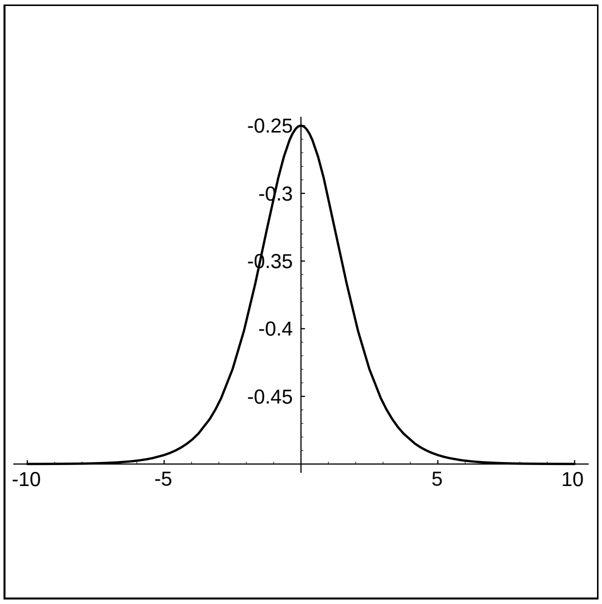

FIG. 0.37. Graph of the function $\tanh(u/2)/(2\tanh(u))$

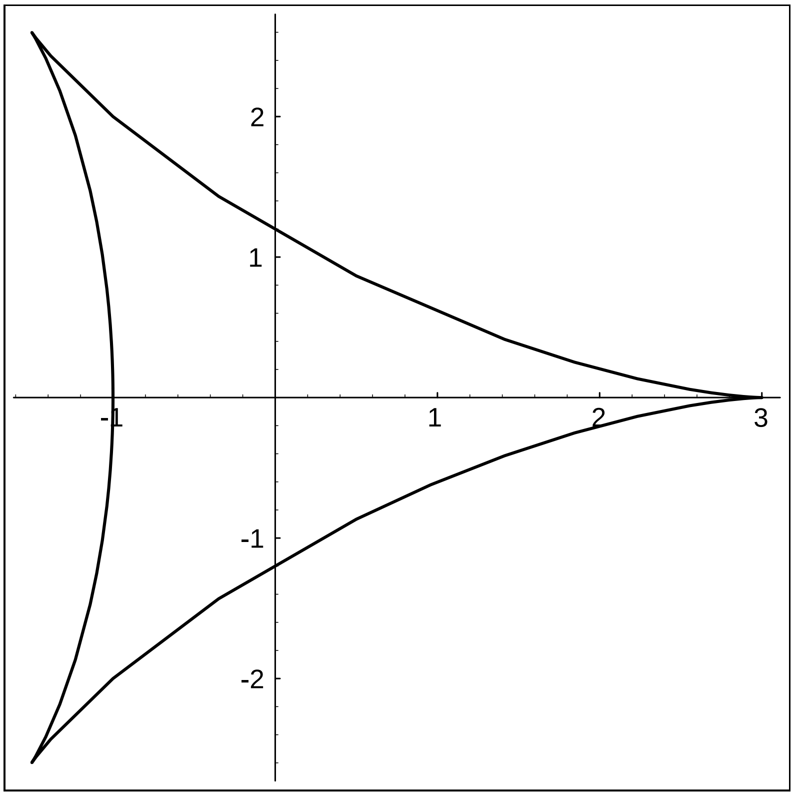

FIG. 0.38. Traces of elliptic elements

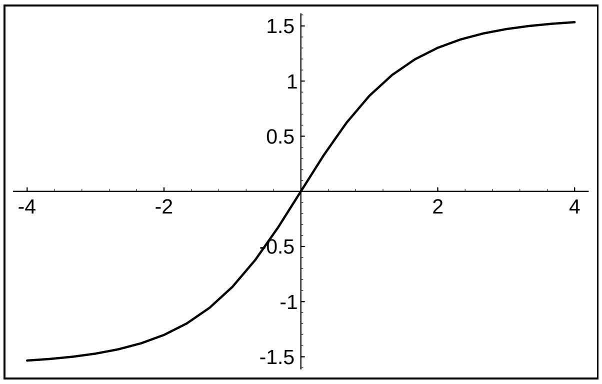

FIG. 0.39. Graph of $\tan^{-1}(\sinh(\mathbb{A}))$

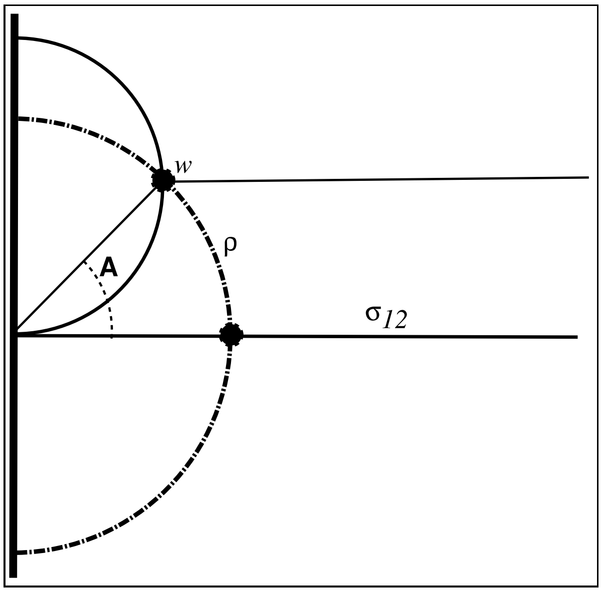

FIG. 0.40. Orthogonal projection of ideal triangle to complex geodesic

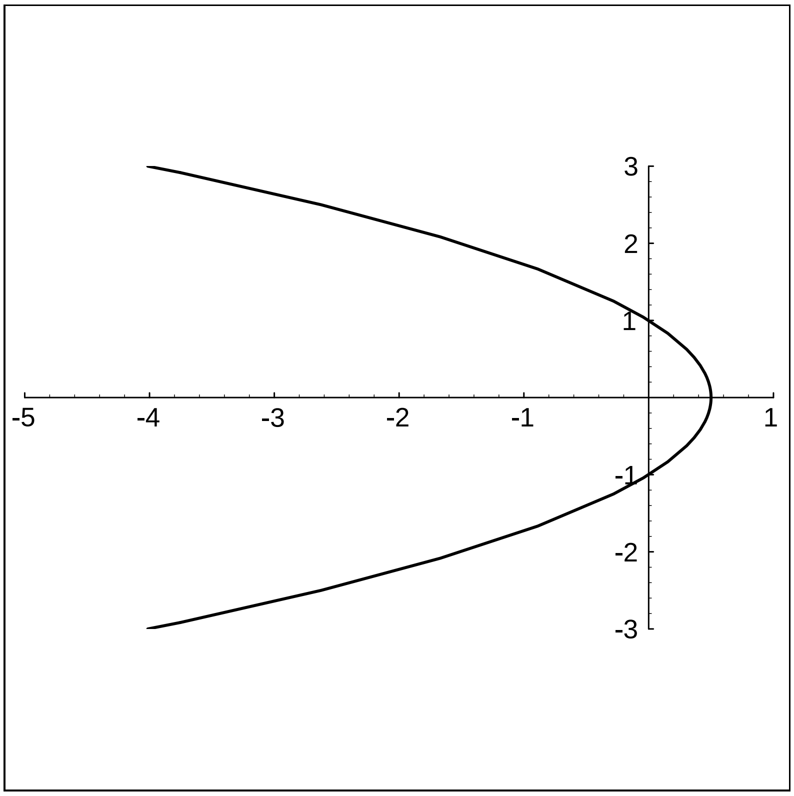

FIG. 0.41. Parabolic region $\mathfrak{P}$ of allowable $\eta(s)$

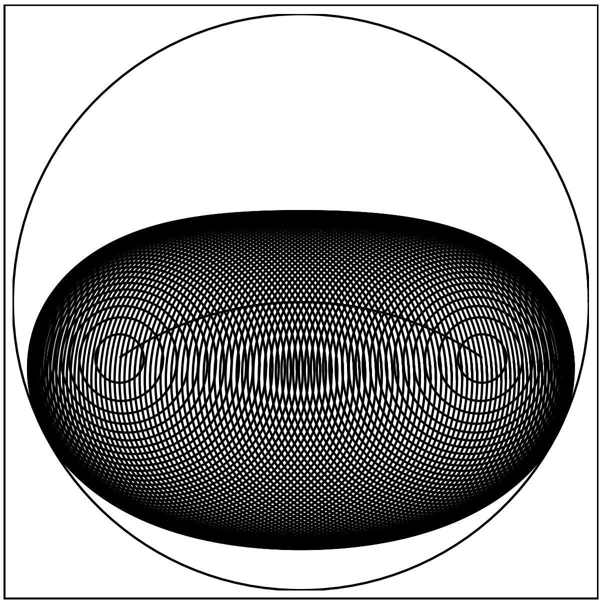

FIG. 0.42. Projection of bisector with $\eta = -0.44i$

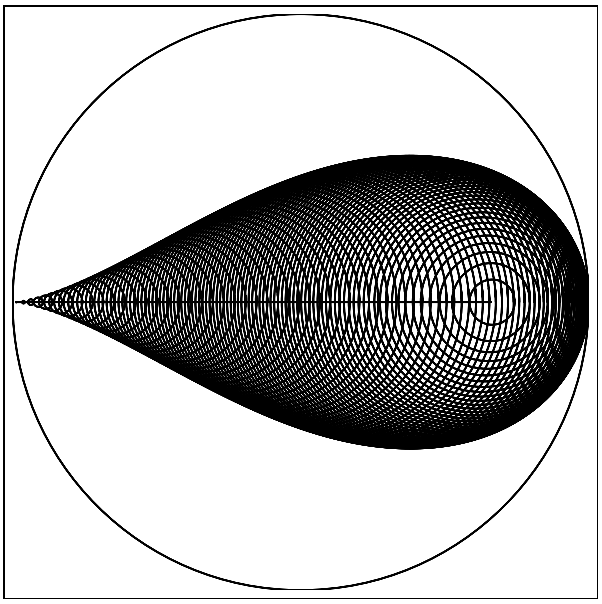

FIG. 0.43. Projection of bisector with $\eta = 0$

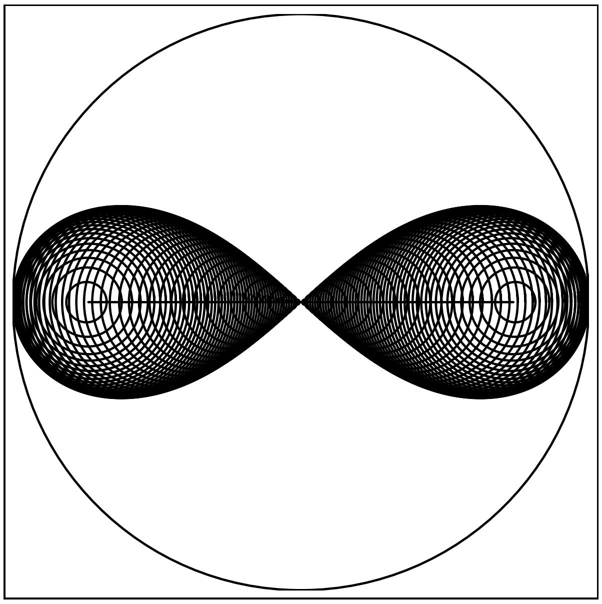

FIG. 0.44. Projection of bisector with $\eta = -0.4$

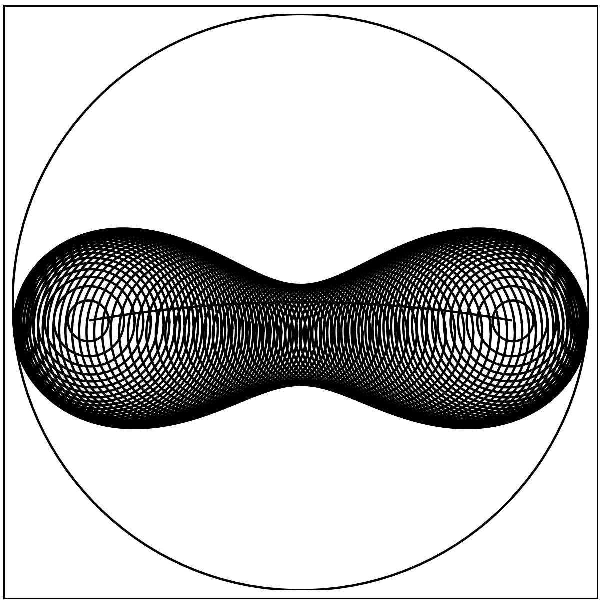

FIG. 0.45. Projection of bisector with $\eta = -0.37 - 0.16i$

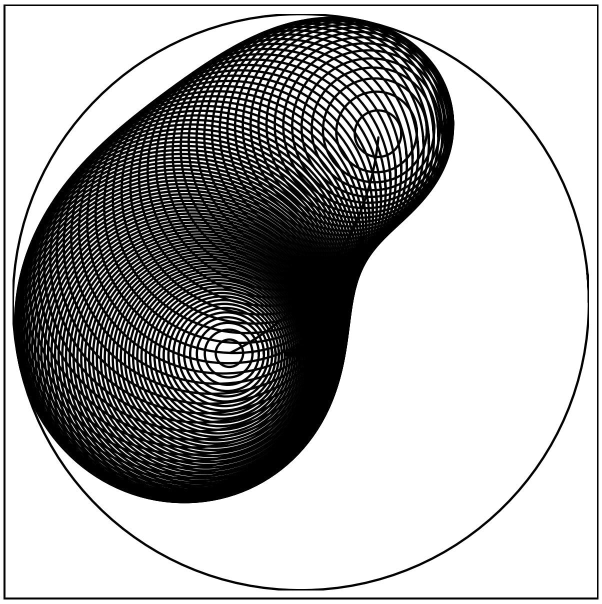

FIG. 0.46. Projection of bisector with $\eta = -1.03 + 0.94i$

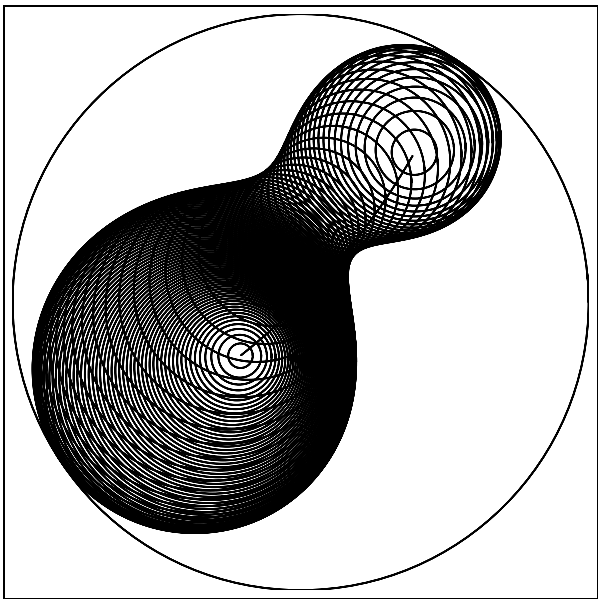

FIG. 0.47. Projection of bisector with $\eta = -1.54 + 0.42i$

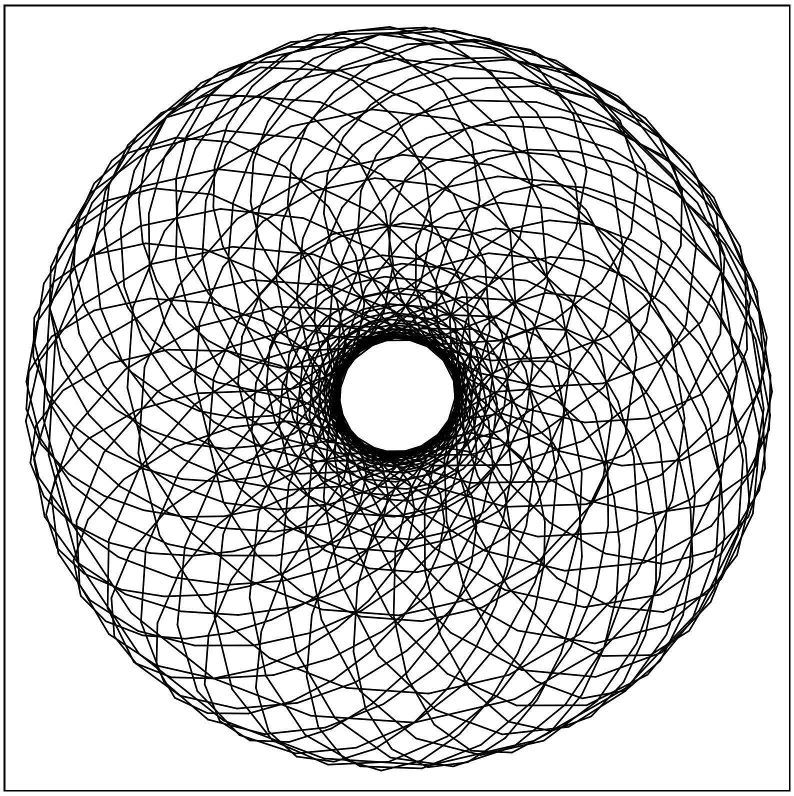

FIG. 0.48. Vertical projection of a Clifford torus

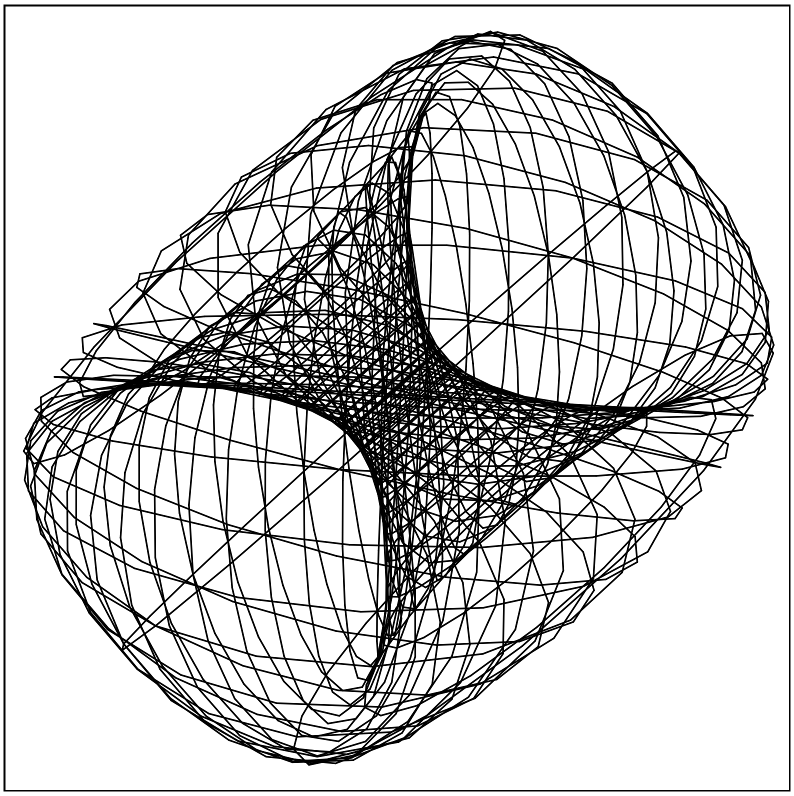

Fig. 0.49. Side view of a Clifford torus

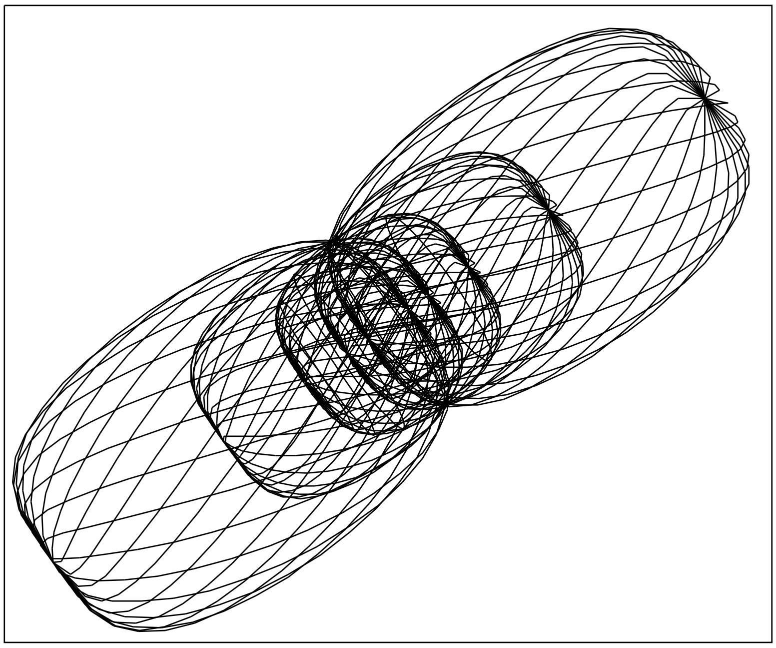

FIG. 0.50. Spinal spheres with a common slice

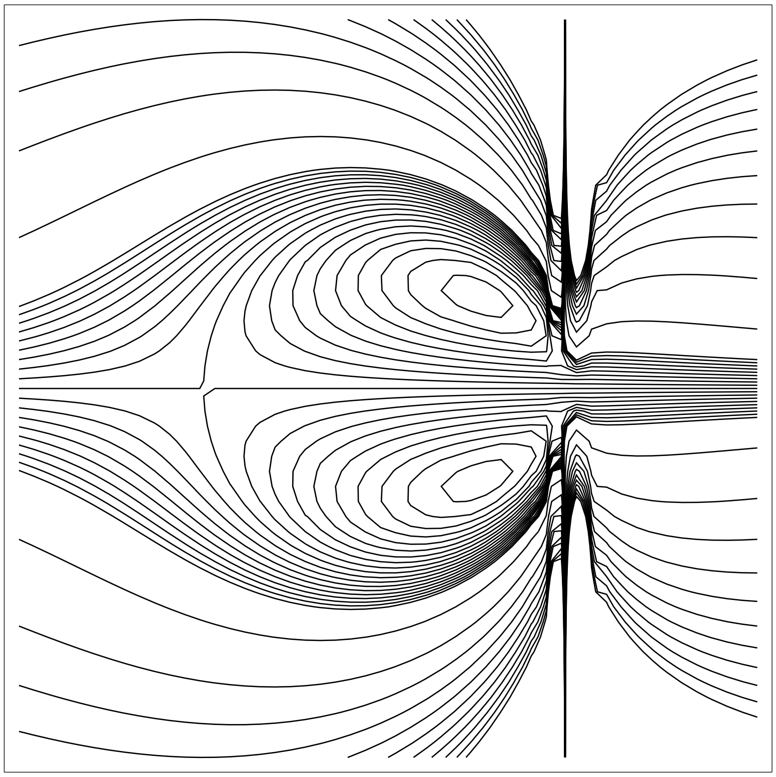

Fig. 0.51. Level sets of function defining an infinite spinal

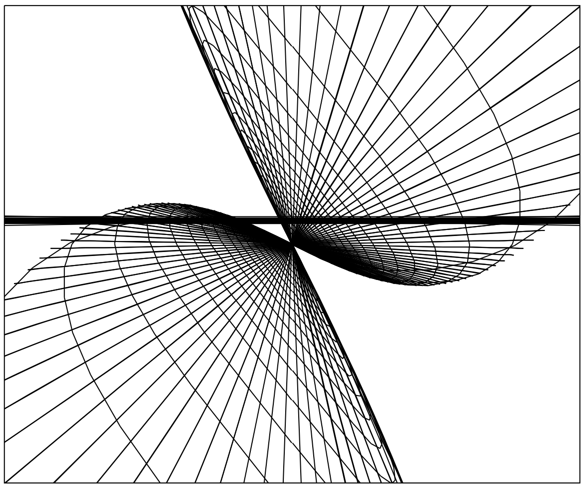

Fig. 0.52. A disconnected intersection of two spinal spheres

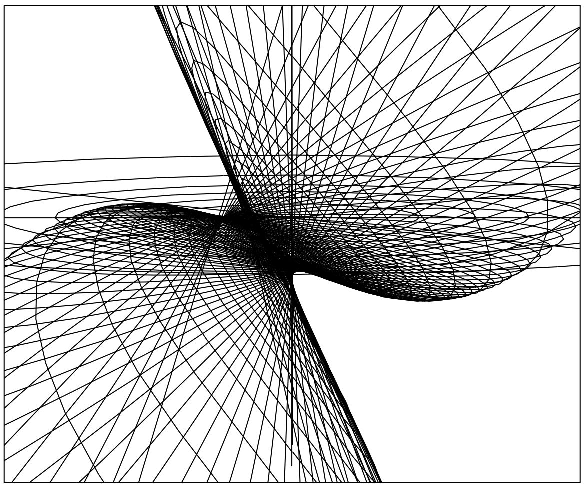

Fig. 0.53. Disconnected intersection of spinal spheres: another view

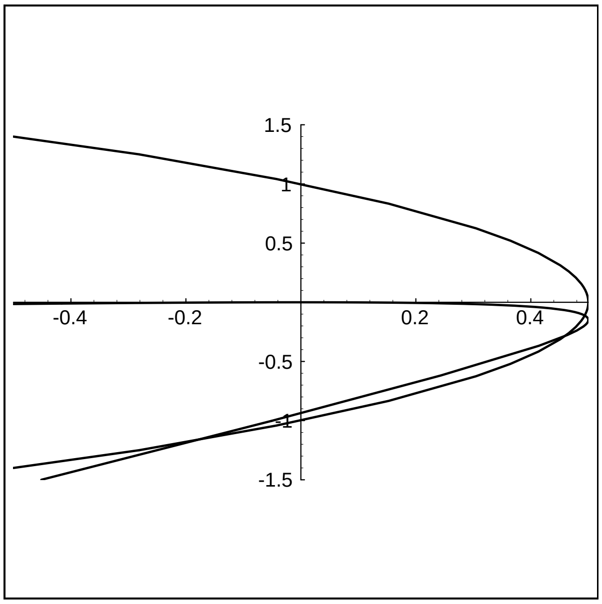

FIG. 0.54. η-hyperbola corresponding to disconnected bisector intersection

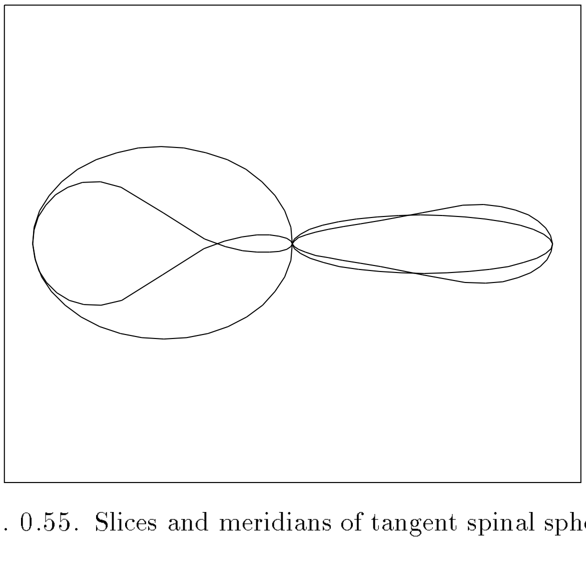

FIG. 0.55. Slices and meridians of tangent spinal spheres

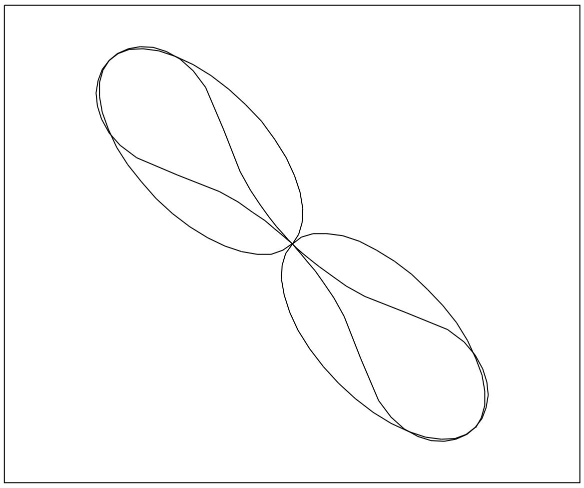

FIG. 0.56. Slices and meridians: another view

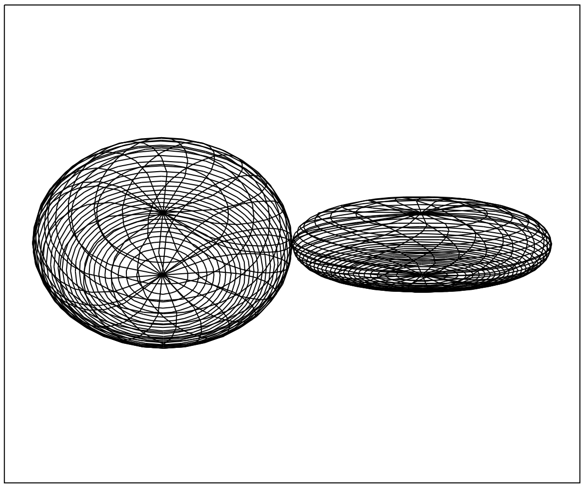

FIG. 0.57. Pair of tangent spinal spheres

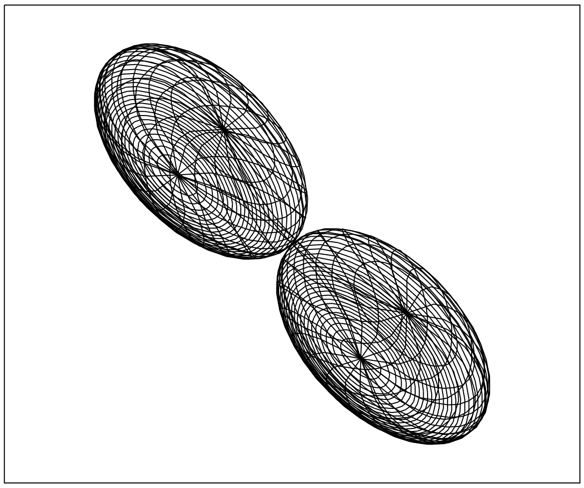

FIG. 0.58. Tangent spinal spheres: another view

1

THE COMPLEX PROJECTIVE LINE

This preliminary section discusses 1-dimensional complex geometry from the point of view adopted in this book. We begin with projective geometry, from which both complex elliptic and hyperbolic 1-dimensional geometry derive. We base our treatment on the books of Cartan [24] and Coolidge [31]. For an excellent treatment of projective geometry in the real case, see Coxeter [34]. A principal geometric object is a *circle* (which Cartan calls a "chain" in [24]) which is just a Euclidean circle or straight line. We find several algebraic ways of describing circles, which will be useful in the sequel. In particular we describe circles in terms of 2×2 Hermitian matrices. We shall make essential use of the *null polarity* (the unique object in projective geometry corresponding to the symplectic structure on $\mathbb{C}^2$) to identify anti-polarities with anti-involutions. Furthermore, elliptic and hyperbolic geometry are developed in terms of *anti-polarities* which are the projective counterpart of Hermitian forms.

1.1 Projective geometry

1.1.1 *Linear algebra in* $\mathbb{C}^2$

If $\mathbb{V}$ is a vector space, a *line* in $\mathbb{V}$ is a 1-dimensional linear subspace. The *projective space* $\mathbb{P}(\mathbb{V})$ *associated to* $\mathbb{V}$ is the space of all lines in $\mathbb{V}$. Since a nonzero vector spans a unique line, and two nonzero vectors span the same line, $\mathbb{P}(\mathbb{V})$ consists of the orbits of the group of nonzero scalars acting by scalar multiplication on the set of nonzero vectors in $\mathbb{V}$. We denote the point in $\mathbb{P}(\mathbb{V})$ corresponding to the line spanned by a nonzero vector $v \in \mathbb{V}$ by $[v]$. In this way, we may give $\mathbb{P}(\mathbb{V})$ the quotient topology.

The *complex projective line* $\mathbb{P}^1_{\mathbb{C}}$ is the projective space $\mathbb{P}(\mathbb{C}^2)$ associated to a 2-dimensional complex vector space. Thus a point in $\mathbb{P}^1_{\mathbb{C}}$ corresponds to a (complex) line in $\mathbb{V}$. We shall work in the standard affine chart whereby

$$\mathbb{P}^1_{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$$

via the embedding

$$\begin{aligned} \mathbb{C} &\longrightarrow \mathbb{P}^1_{\mathbb{C}} \\ z &\longmapsto \begin{bmatrix} z \\ 1 \end{bmatrix} \\ \infty &\longmapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \end{aligned}$$

This is the classical *Gaussian representation* of $\mathbb{P}^1_{\mathbb{C}}$.

1.1.2 *Collineations*

The projective automorphism group of $\mathbb{P}^1_{\mathbb{C}}$ is the group of transformations of $\mathbb{P}^1_{\mathbb{C}}$ induced by the linear automorphism group $\mathbf{GL}(2,\mathbb{C})$ of $\mathbb{V}$. Projective transformations are sometimes also called *collineations* or *homographies.* For example, a 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{GL}(2,\mathbb{C})$$

acts on points in the Gaussian representation by the *linear fractional transformation*

$$z \longmapsto \frac{az+b}{cz+d}.$$

This action is not faithful; its kernel consists of all scalar matrices

$$\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}, a \neq 0$$

and its image is denoted by $\mathbf{PGL}(2,\mathbb{C})$.

Exercise 1.1.1 *The Jordan normal form for 2×2 matrices furnishes a normal form for elements of $\mathbf{SL}(2,\mathbb{C})$. Let $A \in \mathbf{SL}(2,\mathbb{C})$. Then there are the following possibilities:*

1. *$A = \pm I$. In that case A acts trivially on $\mathbb{P}^1_{\mathbb{C}}$.*
2. *$A \neq \pm I$ but A has only one eigenvalue. In that case the eigenvalue must be ± 1. Then A has exactly one fixed point on $\mathbb{P}^1_{\mathbb{C}}$ and is conjugate to the matrix*

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

   which represents the translation $z \mapsto z+1$ in affine coordinates. In this case A is said to be parabolic.
3. *A has two distinct unreal eigenvalues of unit length. In that case the eigenvalues are complex conjugates $e^{\pm i\theta}$ and A has exactly two fixed points on $\mathbb{P}^1_{\mathbb{C}}$ and is conjugate to the diagonal matrix*

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

   representing a rotation $z \mapsto e^{2i\theta} z$ in affine coordinates. In this case A is said to be elliptic.

4. *A has two distinct real eigenvalues, necessarily reciprocals. Then A has exactly two fixed points on* $\mathbb{P}^1_{\mathbb{C}}$ *and is conjugate to the diagonal matrix*

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

representing a dilation $z \mapsto \lambda^2 z$ *in affine coordinates. In this case A is said to be (purely) hyperbolic.*

5. *A has two distinct eigenvalues, neither real nor of unit length. Then A has exactly two fixed points on* $\mathbb{P}^1_{\mathbb{C}}$ *and is conjugate to the diagonal matrix*

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

representing a dilation $z \mapsto \lambda^2 z$ *in affine coordinates. In this case A is said to be loxodromic.*

The collineation associated to A is parabolic or the identity if and only if $\mathrm{trace}(A) = \pm 2$*; it is elliptic if and only if* $-2 < \mathrm{trace}(A) < 2$*; it is purely hyperbolic if and only if* $\mathrm{trace}(A) \in \mathbb{R}$ *and* $|\mathrm{trace}(A)| > 2$*. Otherwise it is loxodromic.*

1.1.3 *Anti-collineations*

Projective transformations of $\mathbb{P}^1_{\mathbb{C}}$ are necessarily complex analytic—their differentials preserve the complex structure on the tangent spaces to projective space. In addition to the projective transformations are the *anti-projective transformations,* (sometimes called *anti-collineations* or *anti-homographies,*) which are the anti-holomorphic transformations induced by conjugate-linear mappings of $\mathbb{V}$. An anti-collineation of order 2 is called an *anti-involution.* The simplest anti-involution is defined by complex conjugation on $\mathbb{C}^2$:

$$\begin{aligned} \rho : \mathbb{C} \cup \{\infty\} &\longrightarrow \mathbb{C} \cup \{\infty\} \\ z &\longmapsto \bar{z} \end{aligned}$$

in inhomogeneous coordinates and by

$$\begin{aligned} \rho : \mathbb{P}^1_{\mathbb{C}} &\longrightarrow \mathbb{P}^1_{\mathbb{C}} \\ \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} &\longmapsto \begin{bmatrix} \bar{Z}_1 \\ \bar{Z}_2 \end{bmatrix} \end{aligned} \tag{1.1}$$

in homogeneous coordinates. The group of transformations of $\mathbb{P}^1_{\mathbb{C}}$ consisting of collineations and anti-collineations is generated by ρ and $\mathbf{PGL}(2, \mathbb{C})$. Another anti-involution is *inversion* in the unit circle, introduced by Steiner. In affine coordinates this inversion

$$z \longmapsto 1/\bar{z} = \frac{z}{|z|^2}$$

is induced by the anti-linear involution of $\mathbb{C}^2$:

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \longmapsto \begin{bmatrix} \bar{Z}_2 \\ \bar{Z}_1 \end{bmatrix}. \tag{1.2}$$

Both these anti-involutions fix points in $\mathbb{P}^1_{\mathbb{C}}$.

Exercise 1.1.2 *Find an anti-involution of $\mathbb{P}^1_{\mathbb{C}}$ which has no fixed points.*

Just as collineations are represented by matrices, an anti-linear map of $\mathbb{C}^2$ is determined by a matrix

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

acting on $Z \in \mathbb{C}^2$ by

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \longmapsto \begin{bmatrix} a\bar{Z}_1 + b\bar{Z}_2 \\ c\bar{Z}_1 + d\bar{Z}_2. \end{bmatrix}$$

Exercise 1.1.3 *Determine, in terms of algebraic properties of A, when the anti-collineation induced by A*

1. *is an anti-involution with fixed points;*
2. *is an anti-involution without fixed points.*

Anti-involutions with fixed points we call *hyperbolic* and one without fixed points we call *elliptic.* (Cartan [24] calls these anti-involutions "de première espèce" and "de seconde espèce" respectively.)

1.1.4 *Dual projective space*

Let $\mathbb{V}$ be a vector space. The vector space $\mathbb{V}^*$ *dual to* $\mathbb{V}$ consists of linear functionals $\mathbb{V} \longrightarrow \mathbb{C}$. We interpret the projective space $\mathbb{P}(\mathbb{V}^*)$ in terms of *hyperplanes* in $\mathbb{V}$ as follows. A *hyperplane* in a vector space $\mathbb{V}$ is a linear subspace of codimension 1. If $\psi \in \mathbb{V}^*$ is nonzero, then $\mathrm{Ker}(\psi)$ is a hyperplane and two linear functionals define the same hyperplane if and only if they are nonzero scalar multiples of one another. Thus the lines in $\mathbb{V}^*$ correspond bijectively to the hyperplanes in $\mathbb{V}$.

We also consider isomorphisms of a projective space with its dual. Consider a map $f : \mathbb{P}(\mathbb{V}) \longrightarrow \mathbb{P}(\mathbb{V}^*)$, a rule which associates to every point in $\mathbb{P}(\mathbb{V})$ a hyperplane in $\mathbb{P}(\mathbb{V})$. Such a map is required to be "projective" in the sense that it preserves incidences: if $l_1, l_2, l_3 \in \mathbb{P}(\mathbb{V})$ are collinear points (that is, if $l_i = [v_i]$ for $v \in \mathbb{V}$, then v_1, v_2, v_3 are linearly dependent), then the hyperplanes $f(l_1), f(l_2), f(l_3)$ intersect in a projective subspace of codimension 2. It is a remarkable fact that such projective isomorphisms (called *correlations*) correspond to linear isomorphisms $\mathbb{V} \longrightarrow \mathbb{V}^*$. (See Coxeter [35].)

When $\mathbb{V}$ has dimension 2, then lines and hyperplanes agree. There results a *canonical* identification of a 1-dimensional projective space with its dual.

This correlation is called the *null polarity* and is defined as follows. If $\mathbb{V}$ is 2-dimensional, then a *symplectic structure* on $\mathbb{V}$, that is, a nondegenerate skew-symmetric pairing

$$\nu : \mathbb{V} \times \mathbb{V} \longrightarrow \mathbb{C},$$

induces an isomorphism

$$\hat{\nu} : \mathbb{V} \longrightarrow \mathbb{V}^*$$

having the property that

$$\hat{\nu}(v)(v) = 0.$$

Furthermore since $\dim(\mathbb{V}) = 2$, the space of skew-symmetric pairings is itself 1-dimensional, so that any nondegenerate skew-symmetric pairing ν is a nonzero multiple of a fixed one ν_0.

A linear map $f : \mathbb{V} \longrightarrow \mathbb{V}^*$ identifies with a bilinear form $\tilde{f} : \mathbb{V} \times \mathbb{V} \longrightarrow \mathbb{C}$ via the rule

$$\tilde{f}(v_1, v_2) = f(v_1)(v_2).$$

Such bilinear forms correspond to square matrices F by

$$\tilde{f}(v_1, v_2) = v_2^{\dagger} F v_1.$$

The correlation is an isomorphism if and only if the corresponding bilinear form is nondegenerate.

For example the null polarity corresponds to the skew-symmetric matrix

$$J_0 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{1.3}$$

(or, more accurately, any nonzero scalar multiple of J_0). The correlation $\mathbb{P}(\mathbb{V}) \longrightarrow \mathbb{P}(\mathbb{V}^*)$ corresponding to a bilinear form $\tilde{f}$ defines (by composition with the inverse of the null polarity $\mathbb{P}(\mathbb{V}^*) \longrightarrow \mathbb{P}(\mathbb{V})$) a collineation $\mathbb{P}(\mathbb{V}) \longrightarrow \mathbb{P}(\mathbb{V})$ corresponding to the linear transformation defined by the matrix $F \circ J_0$.

Definition 1.1.4 *$p \in \mathbb{P}(\mathbb{V})$ and $H \in \mathbb{P}(\mathbb{V}^*)$ are incident if $p \in H$.*

Exercise 1.1.5 *A polarity is a correlation $c : \mathbb{P}(\mathbb{V}) \longrightarrow \mathbb{P}(\mathbb{V}^*)$ whose inverse $c^{-1} : \mathbb{P}(\mathbb{V}^*) \longrightarrow \mathbb{P}(\mathbb{V})$ "corresponds to itself;" that is, $p \in \mathbb{P}(\mathbb{V})$ and $H \in \mathbb{P}(\mathbb{V}^*)$ are incident if and only if $c^{-1}(H) \in \mathbb{P}(V)$ and $c(p) \in \mathbb{P}(V^*)$ are incident. Then a nondegenerate bilinear form B defines a polarity if and only if B is symmetric or skew-symmetric. Compare [35].*

Exercise 1.1.6 *Let $A \in \mathbf{GL}(2, \mathbb{C})$. Then*

$$A^{\dagger} J_0 A = \det(A) J_0.$$

1.1.5 *The conjugate projective space*

One may also define the *conjugate* projective space $\overline{\mathbb{P}^1_{\mathbb{C}}}$, which has the same underlying set as $\mathbb{P}^1_{\mathbb{C}}$ but enjoys the opposite complex structure. In other words, the identity mapping $\overline{\mathbb{P}^1_{\mathbb{C}}} \longrightarrow \mathbb{P}^1_{\mathbb{C}}$ induces an anti-projective isomorphism. In this way an anti-projective mapping is defined as a projective mapping defined on the *conjugate* projective space.

An *anti-correlation* of a projective space $\mathbb{P}(\mathbb{V})$ is thus a projective isomorphism from $\mathbb{P}(\mathbb{V})$ to its conjugate dual $\overline{\mathbb{P}(\mathbb{V}^*)}$. An anti-correlation $\mathbb{P}(\mathbb{V}) \longrightarrow \mathbb{P}(\bar{\mathbb{V}}^*)$ corresponds to a *sesquilinear form,* that is a form $H : \mathbb{V} \times \mathbb{V} \longrightarrow \mathbb{C}$ which is bilinear over $\mathbb{R}$ and is *Hermitian-symmetric:*

$$H(\lambda_1 v_1, \lambda_2 v_2) = \lambda_1 \overline{\lambda_2} H(v_1, v_2)$$

for $v_1, v_2 \in \mathbb{V}$ and $\lambda_1, \lambda_2 \in \mathbb{C}$. Such a form is described by a matrix F as follows:

$$H(v_1, v_2) = (\bar{v}_2)^\dagger F v_1.$$

Composition of an anti-correlation $\phi : \mathbb{P}(\mathbb{V}) \longrightarrow \mathbb{P}(\mathbb{V}^*)$ with the inverse of the null polarity $\nu^{-1} : \mathbb{P}(\mathbb{V}^*) \longrightarrow \mathbb{P}(\mathbb{V})$ yields an anti-collineation $\nu^{-1} \circ \phi : \mathbb{P}(\mathbb{V}) \longrightarrow \mathbb{P}(\mathbb{V})$.

Exercise 1.1.7 *Show that the anti-collineation ϕ is an anti-polarity (that is, equals its own inverse) if and only if the anti-collineation $\nu^{-1} \circ \phi$ is an anti-involution.*

The map on projective spaces is an isomorphism if and only if H is a nondegenerate. The anti-correlation is an anti-polarity if and only if H is a Hermitian form. In that case F is a *Hermitian matrix:*

$$\bar{F}^\dagger = F$$

Exercise 1.1.8 *Let ϕ be the anti-polarity defined by the standard Hermitian form*

$$\langle\!\langle u, v \rangle\!\rangle = u_1 \bar{v_1} + u_2 \bar{v_2}$$

on $\mathbb{C}^2$. Show that the corresponding anti-involution on the $\mathbb{P}^1_{\mathbb{C}}$ equals

$$\nu^{-1} \circ \phi : z \longmapsto 1/\bar{z}$$

If F is the Hermitian matrix, then the matrix A representing the anti-involution is given by

$$F = A^\dagger J_0. \tag{1.4}$$

The condition that $z \mapsto A\bar{z}$ has order 2 is equivalent to $A\bar{A} = I$, which is equivalent to F being Hermitian.

1.1.6 *The cross-ratio*

Let $\mathcal{C}_4(\mathbb{P}^1_{\mathbb{C}})$ denote the subset of $\mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}}$ consisting of quadruples of distinct points. The *cross-ratio* is the mapping

$$\mathcal{C}_4(\mathbb{P}^1_{\mathbb{C}}) \longrightarrow \mathbb{P}^1_{\mathbb{C}}$$

defined in Gaussian coordinates by

$$\mathbf{X}\{z_1, z_2, z_3, z_4\} = \frac{z_4 - z_1}{z_4 - z_2} \Big/ \frac{z_3 - z_1}{z_3 - z_2} \tag{1.5}$$

and is invariant under the diagonal action of $\mathbf{PGL}(2, \mathbb{C})$ on $\mathcal{C}_4(\mathbb{P}^1_{\mathbb{C}})$. (For a geometric description of this classical invariant, see §1.1.7). In particular,

$$\mathbf{X}\{0, \infty, 1, z\} = z. \tag{1.6}$$

The symmetric group $\mathfrak{S}_4$, consisting of all permutations of $\{1, 2, 3, 4\}$, acts on $\mathcal{C}_4(\mathbb{P}^1_{\mathbb{C}})$ by

$$\sigma : (z_1, z_2, z_3, z_4) \longmapsto (z_{\sigma(1)}, z_{\sigma(2)}, z_{\sigma(3)}, z_{\sigma(4)})$$

The normal subgroup

$$\mathfrak{D} \cong \mathbb{Z}/2 \times \mathbb{Z}/2 \lhd \mathfrak{S}_4$$

comprising products of pairs of disjoint transpositions leaves the cross-ratio invariant. $\mathfrak{S}_4$ decomposes as a semidirect product $\mathfrak{S}_4 \cong \mathfrak{S}_3 \ltimes \mathfrak{D}$ where $\mathfrak{S}_3 \subset \mathfrak{S}_4$ is the subgroup fixing 4. In particular $\mathfrak{S}_3$ is a system of coset representatives for $\mathfrak{S}_4/\mathfrak{D}$ so the equivariance of cross-ratio with respect to $\mathfrak{S}_4$ is described by the six formulas:

$$\begin{aligned}
\mathbf{X}\{0, \infty, 1, z\} &= z \\
\mathbf{X}\{\infty, 0, 1, z\} &= 1/z \\
\mathbf{X}\{1, \infty, 0, z\} &= 1 - z \\
\mathbf{X}\{\infty, 1, 0, z\} &= 1/(1 - z) \\
\mathbf{X}\{0, 1, \infty, z\} &= z/(z - 1) \\
\mathbf{X}\{1, 0, \infty, z\} &= (z - 1)/z.
\end{aligned}$$

Exercise 1.1.9 *A transformation $f : \mathbb{P}^1_{\mathbb{C}} \longrightarrow \mathbb{P}^1_{\mathbb{C}}$ is*

1. *a collineation if and only if*

$$\mathbf{X}\{f(z_1), f(z_2), f(z_3), f(z_4)\} = \mathbf{X}\{z_1, z_2, z_3, z_4\};$$

2. *an anti-collineation if and only if*

$$\mathbf{X}\{f(z_1), f(z_2), f(z_3), f(z_4)\} = \overline{\mathbf{X}\{z_1, z_2, z_3, z_4\}}.$$

Exercise 1.1.10 *Let $z_0, z_1, z_\infty \in \mathbb{P}^1_{\mathbb{C}}$ be distinct. Show that*

$$z \longmapsto \mathbf{X}\{z, z_0, z_1, z_\infty\}$$

is the unique collineation mapping z_0 to 0, z_1 to 1 and z_∞ to ∞.

1.1.7 *Cross-ratio and slopes*

Here is a geometric description of the cross-ratio. The points z_1, z_2, z_3, z_4 in $\mathbb{P}^1_{\mathbb{C}}$ correspond to distinct lines $L_i \subset \mathbb{C}^2$. Since $L_1 \neq L_2$, the vector space $\mathbb{C}^2$ decomposes as a direct sum

$$\mathbb{C}^2 = L_1 \oplus L_2.$$

Every line in $\mathbb{C}^2$ distinct from L_1, L_2 is the graph of a nonzero linear map $f : L_1 \longrightarrow L_2$:

$$L = \mathrm{graph}(f) = \{(u, f(u)) \mid u \in L_1\} \subset L_1 \oplus L_2.$$

Thus $L_3 = \mathrm{graph}(f_3)$ and $L_4 = \mathrm{graph}(f_4)$ correspond to nonzero linear maps $f_3, f_4 : L_1 \longrightarrow L_2$. Two nonzero linear maps between lines differ by multiplication by a scalar, which will be the cross-ratio above. If

$$f_4(x) = \lambda_{4,3} f_3(x)$$

for $x \in L_1$, then we define

$$\mathbf{X}\{z_1, z_2, z_3, zp_4\} = \lambda_{4,3}.$$

Exercise 1.1.11 *Cross-ratio extends to the larger subset of* $\mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}} \times \mathbb{P}^1_{\mathbb{C}}$ *consisting of ordered quadruples where at most two entries are equal.*

Explicitly, suppose that $z_1, z_2 \in \mathbb{C}$ define vectors

$$e_1 = \begin{bmatrix} z_1 \\ 1 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} z_2 \\ 1 \end{bmatrix}$$

spanning lines $L_1, L_2 \subset \mathbb{C}^2$. A linear map $f : L_1 \longrightarrow L_2$ is determined by a scalar λ (called the *multiplier*) such that

$$f(e_1) = \lambda e_2. \tag{1.7}$$

Let L be the line spanned by

$$\begin{bmatrix} z \\ 1 \end{bmatrix} \longleftrightarrow z \in \mathbb{P}^1_{\mathbb{C}}.$$

Then $L = \mathrm{graph}(f)$ where f is defined as in (1.7) and

$$\lambda = \frac{z - z_1}{z - z_2}.$$

Thus if L_3 corresponds to z_3 and L_4 corresponds to z_4 their respective multipliers are

$$\lambda_{4,3} = \frac{\lambda_3}{\lambda_4} = \frac{z_4 - z_1}{z_4 - z_2} \bigg/ \frac{z_3 - z_1}{z_3 - z_2}.$$

1.2 Circles

It follows from Exercise 1.1.9 that if four points are fixed by an anti-collineation, then their cross-ratio is real. We define a *circle* in $\mathbb{P}^1_{\mathbb{C}}$ to be the fixed-point set of an anti-collineation. In the usual Gaussian representation a circle is either a Euclidean straight line or a Euclidean circle. A circle passing through ∞ is a straight line, that is a degenerate Euclidean circle. (Compare Coolidge [31] or Young [171].) These ubiquitous geometric entities arise as geodesics, metric circles and bisectors in elliptic and hyperbolic geometry. Points in $\mathbb{P}^1_{\mathbb{C}}$ are limits of degenerating sequences of circles (circles of zero radius). Cartan [24] refers to these objects as *chains,* but since we wish to use the terminology "chain" later on for different objects (following Cartan [21]), we call these objects *circles* rather than chains. We reserve the terminology "Euclidean circle" and "metric circle" for the more specialized (and common) usages of this term.

1.2.1 *Circles and anti-involutions*

Let $\phi : \mathbb{P}(\mathbb{V}) \longrightarrow \mathbb{P}(\bar{\mathbb{V}}^*)$ be an anti-polarity. Its *null locus* $N(\phi)$ consists of all points $x \in \mathbb{P}(\mathbb{V})$ such that x is incident to $\phi(x)$. If ϕ is defined by a nondegenerate form $\Phi : \mathbb{V} \times \bar{\mathbb{V}} \longrightarrow \mathbb{C}$, then

$$N(\phi) = \{[v] \mid v \in \mathbb{V}, \quad \Phi(v, v) = 0\}$$

An anti-polarity is *hyperbolic* (respectively *elliptic*) if $N(\phi)$ is nonempty (respectively empty).

Exercise 1.2.1 *Suppose* $\dim \mathbb{V} = 2$ *and let* ν *denote the null polarity. The null locus* $N(\phi)$ *equals the fixed-point set of the corresponding anti-involution* $\nu^{-1} \circ \phi$.

Exercise 1.2.2 *Let* $C \subset \mathbb{P}^1_{\mathbb{C}}$ *be a 1-dimensional (real) submanifold of* $\mathbb{P}^1_{\mathbb{C}}$. *Then the following are equivalent:*

1. *C is a circle.*
2. *There is a (necessarily) unique anti-involution ι whose fixed-point set is C.*
3. *For all distinct $z_1, z_2, z_3, z_4 \in C$, the cross-ratio $\mathbf{X}\{z_1, z_2, z_3, z_4\} \in \mathbb{R}$.*
4. *There exists a collineation $\phi \in \mathbf{PGL}(2, \mathbb{C})$ such that $\phi(C) = \mathbb{P}^1_{\mathbb{R}}$.*

Suppose $z_1, z_2, z_3 \in \mathbb{P}^1_{\mathbb{C}}$ are distinct points. Then

$$\{z \in \mathbb{P}^1_{\mathbb{C}} \mid \mathbf{X}\{z_1, z_2, z_3, z\} \in \mathbb{R}\} \cup \{z_1, z_2, z_3\}$$

is the unique circle containing z_1, z_2, z_3.

Exercise 1.2.3 *Derive a formula for the anti-involution fixing this circle.*

Exercise 1.2.4 *Let* $C \in \mathbb{P}^1_{\mathbb{C}}$ *be a circle and let* $c_1, c_2 \in C$ *be distinct points. Let* $v_1, v_2 \in \mathbb{C}^2$ *be vectors such that* $c_i = [v_i]$. *Let* $S = \mathbb{R}v_1 + \mathbb{R}v_2$ *be the* $\mathbb{R}$*-linear 2-plane spanned by* v_1, v_2. *Then* C *consists of all* $[v]$ *where* $v \in S$.

1.2.2 *Circles and Hermitian matrices*

Since circles correspond to anti-polarities (as well as anti-involutions), they can be effectively represented by matrices. A circle in $\mathbb{P}^1_{\mathbb{C}}$ is the null locus of a hyperbolic anti-polarity, which is represented by an *indefinite Hermitian matrix*. For example, the anti-polarity will null locus the real axis corresponds to the Hermitian form

$$H(Z,W) = i(Z_1\bar{W}_2 - Z_2\bar{W}_1)$$

(the corresponding anti-involution is complex conjugation; see (1.1)). The corresponding matrices are

$$F = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The anti-polarity will null locus the unit circle axis corresponds to the Hermitian form

$$H(Z,W) = Z_1\bar{W}_1 - Z_2\bar{W}_2$$

(the corresponding anti-involution is given in (1.2)) and the corresponding matrices are

$$F = \mathbb{I}_{1,1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Suppose that F is a Hermitian matrix representing C. That is, $C = N(\phi)$ where ϕ is the anti-polarity defined by F. Suppose that $T \in \mathbf{GL}(2,\mathbb{C})$ is a matrix representing a collineation of $\mathbb{P}^1_{\mathbb{C}}$, which we also denote by T. Then the Hermitian matrix corresponding to $T(C)$ is $\bar{T}^\dagger F T$.

Exercise 1.2.5 *The Hermitian matrix corresponding to the line $\zeta_0 + e^{i\theta}\mathbb{R}$ is*

$$\begin{bmatrix} 1 & 0 \\ -\bar{\zeta}_0 & 1 \end{bmatrix} \begin{bmatrix} 0 & ie^{i\theta} \\ -ie^{-i\theta} & 0 \end{bmatrix} \begin{bmatrix} 1 & -\zeta_0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & ie^{i\theta} \\ -ie^{-i\theta} & -2\mathrm{Im}(e^{-i\theta}\zeta_0) \end{bmatrix}$$

and a Hermitian matrix corresponding to the circle centered at ζ_0 with radius r_0 is

$$\begin{bmatrix} 1 & 0 \\ -\bar{\zeta}_0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -r_0^2 \end{bmatrix} \begin{bmatrix} 1 & -\zeta_0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\zeta_0 \\ -\bar{\zeta}_0 & \zeta_0\bar{\zeta}_0 - r_0^2 \end{bmatrix}.$$

1.2.3 *Pairs of circles*

Two distinct circles may intersect in none, one or two points. The space of equivalence classes of pairs of circles is 1-dimensional. If C_1, C_2 are circles intersecting in two points, their intersection is transverse, and the angles of intersection at each of the two points of $C_1 \cap C_2$ are equal, giving a well-defined invariant. If

$C_1 \cap C_2$ is a single point, then C_1 and C_2 are tangent. If C_1 and C_2 are disjoint, then there exists a circle c which is orthogonal to both C_1 and C_2.

This trichotomy can be readily understood in terms of the hyperbolic anti-involutions (inversions) α_i determined by C_i. The product $\alpha_1 \circ \alpha_2$ is a collineation which fixes $C_1 \cap C_2$. If $\alpha_1 \circ \alpha_2$ is elliptic, then it is represented by an element of $\mathbf{SL}(2, \mathbb{C})$ with trace $\pm 2\cos(\theta)$, where θ is the angle between C_1 and C_2. $\alpha_1 \circ \alpha_2$ is parabolic if and only if $C_1 \cap C_2$ is a single point. $\alpha_1 \circ \alpha_2$ is hyperbolic if and only if $C_1 \cap C_2 = \emptyset$.

The collineation group $\mathbf{PGL}(2, \mathbb{C})$ acts transitively on circles. Circles C_1, C_2 corresponding to Hermitian matrices F_1, F_2 respectively meet at angle θ if and only if

$$\frac{\operatorname{trace}(J_0 \bar{F}_1 \ J_0 F_2)^2}{\det(J_0 \bar{F}_1 J_0 H_2)} = 4\cos^2(\theta) \tag{1.8}$$

(where J_0 is defined in (1.3)). To prove (1.8), consider the product of the anti-involutions corresponding to the circles C_1, C_2. By (1.4), the anti-involution corresponding to C_i is represented by the linear map

$$\tilde{\rho}_i : v \mapsto (J_0 \bar{h}_i)\bar{v}.$$

Their composition $\rho_1 \circ \rho_2$ is a collineation which is represented by the matrix

$$J_0 \bar{h}_1 \ J_0 h_2 \ \in \mathbf{GL}(2, \mathbb{C})$$

where J_0 is defined in (1.3). The collineation $\rho_1 \circ \rho_2$ is represented by an elliptic element of $\mathbf{PGL}(2, \mathbb{C})$ fixing $C_1 \cap C_2$, and is conjugate to a rotation of angle 2θ represented by the matrix

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \sim \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}.$$

having trace (2θ) when the determinant is normalized to equal one.

1.3 Elliptic geometry

Now we develop the Riemannian geometry on $\mathbb{P}^1_{\mathbb{C}} \approx S^2$ induced by the embedding of S^2 as the sphere $S^2(r)$ of radius $r > 0$ in Euclidean space $\mathbb{R}^3$. To illustrate the formal properties of these metrics, the radius r is a parameter throughout our discussion. Recall that the curvature of $S^2(r)$ equals r^{-2}.

For a detailed discussion of the geometry of S^2, see Coxeter [34], §VI.

1.3.1 *Isometries*

The isometry group of $S^2(r)$ is the orthogonal group $O(3)$ consisting of two components. The identity component $SO(3)$ consists of rotations. The other component consists of orientation-reversion isometries, and contains two kinds of involutions: reflections in hyperplanes, and the antipodal map

$$\alpha_O : \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \longmapsto \begin{bmatrix} -x_1 \\ -x_2 \\ -x_3 \end{bmatrix}$$

Exercise 1.3.1 *The general orientation-reversing isometry of S^2 is a composition of a reflection in a plane and a rotation leaving invariant the plane. For example,*

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

is a reflection in the xy-plane for $\theta = 0$ and the antipodal map for $\theta = \pm\pi$.

If H is a plane in $\mathbb{R}^3$ containing the origin O, then the Euclidean reflection ι_H in H is the unique isometry whose fixed-point set is H. It follows that the *great circle* $H \cap S^2(r)$ is a totally geodesic subset of $S^2(r)$ and thus a geodesic. Since through any point on $S^2(r)$ in any direction tangent to $S^2(r)$ there exists a great circle, the geodesics on $S^2(r)$ are precisely the great circles.

Suppose that $x, y \in S^2(r)$. Their distance—*as measured on* $S^2(r)$— equals the shortest length of an arc of the great circle on $S^2(r)$ joining x to y. Thus $d(x, y)$ equals the radius r multiplied by the angle subtended by the lines spanned by x and y:

$$\cos\left(\frac{d(x,y)}{r}\right) = |x/||x|| \cdot y/||y|||$$

which we rewrite as

$$\cos^2\left(\frac{d(x,y)}{r}\right) = \frac{(x\cdot y)(y\cdot x)}{(x\cdot x)(y\cdot y)}. \tag{1.9}$$

This expression is *homogeneous* in x, y, that is it depends only on the lines $[x], [y]$ spanned by x, y respectively. (Such expressions for non-Euclidean distance were considered by Cayley [26]; compare the discussion in Coxeter [34].)

In particular, if $x, y \in S^2(r)$ then $d(x, y) \leq \pi r$, with equality if and only if x, y are antipodal. Thus the (Riemannian) *diameter* of $S^2(r)$ equals πr.

Every geodesic through a point x also contains its antipode $\alpha_O(x)$. Furthermore, if $x, y \in S^2(r)$ are not antipodal, then the plane they generate meets $S^2(r)$ in the unique geodesic joining x, y. Thus, folowing Felix Klein, define the *elliptic plane* $\mathbf{E}^2_{\mathbb{R}}$ as the quotient of $S^2(r)$ by the antipodal map. $\mathbf{E}^2_{\mathbb{R}}$ identifies with the real projective plane. The restriction ds^2 of the Euclidean metric to $S^2(r)$ is invariant under α_O, and defines a Riemannian metric on $\mathbf{E}^2_{\mathbb{R}}$ of curvature r^{-2}. The diameter of $\mathbf{E}^2_{\mathbb{R}}$ equals $\pi r/2$ and the area is $2\pi r^2$.

1.3.2 *Stereographic projection*

To better visualize $\mathbb{P}^1_{\mathbb{C}}$ we relate it to the sphere $S^2(r)$ by stereographic projection. We give $\mathbb{P}^1_{\mathbb{C}}$ the Riemannian metric induced from the Euclidean metric on this sphere. With this metric, $\mathbb{P}^1_{\mathbb{C}}$ is the *complex elliptic line* and denoted $\mathbf{E}^1_{\mathbb{C}}$. Choose a parameter $r > 0$.

Exercise 1.3.2 *Embed $\mathbb{C}$ as the affine plane in $\mathbb{R}^3$ by*

$$\zeta = x + iy \longmapsto \begin{bmatrix} rx \\ ry \\ 0 \end{bmatrix}. \tag{1.10}$$

The South pole

$$S = \begin{bmatrix} 0 \\ 0 \\ -r \end{bmatrix}$$

corresponds to the origin in $\mathbb{C}$. The "North pole"

$$N = \begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix}$$

corresponds to $\infty \in \mathbb{P}^1_{\mathbb{C}}$. Define stereographic projection $\Sigma : \mathbb{C} \cup \{\infty\} \longrightarrow S^2(r)$ as follows. If $\zeta \in \mathbb{C}$, then $\Sigma(\zeta)$ is the unique point in which the line joining the North pole to the point corresponding to ζ in $\mathbb{R}^3$ intersects $S^2(r)$. Explicitly

$$\begin{aligned} \Sigma : \mathbb{P}^1_{\mathbb{C}} &\longrightarrow S^2(r) \\ \zeta &\longmapsto \frac{r}{|\zeta|^2 + 1} \begin{bmatrix} \mathrm{Re}(\zeta) \\ \mathrm{Im}(\zeta) \\ |\zeta|^2 - 1 \end{bmatrix} \\ \infty &\longmapsto \begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix} \end{aligned}$$

and its inverse is

$$\begin{aligned} \Sigma^{-1} : S^2 &\longrightarrow \mathbb{P}^1_{\mathbb{C}} \\ \begin{bmatrix} x \\ y \\ z \end{bmatrix} &\longmapsto \frac{x + iy}{r - z} \end{aligned}$$

We have chosen the embedding (1.10) so that the equator $S^2(r) \cap (\mathbb{R}^2 \times \{0\})$ corresponds to the unit circle $|\zeta| = 1$ in $\mathbb{C}$.

Here is the Riemannian metric on $\mathbb{P}^1_{\mathbb{C}}$ induced by stereographic projection:

Exercise 1.3.3 *Let ds^2 denote the Euclidean metric tensor (first fundamental form) on $S^2(r)$. Show that the Riemannian metric on $\mathbb{P}^1_{\mathbb{C}}$ induced by stereographic projection Σ equals*

$$\Sigma^*(ds^2) = \frac{4r^2}{(1+\zeta\bar{\zeta})^2} d\zeta d\bar{\zeta} \tag{1.11}$$

In particular the restriction of Σ to $\mathbb{C}$ is conformal with respect to the Euclidean metric on $\mathbb{C}$. As $r \longrightarrow \infty$, these Riemannian metrics do not converge to a Riemannian metric.

Exercise 1.3.4 *Instead of the embedding defined in (1.10), use*

$$\zeta = x + iy \longmapsto \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

to define an embedding $\Sigma' : \mathbb{C} \cup \{\infty\} \longrightarrow S^2(r)$ which no longer maps the unit circle to the equator.

$$(\Sigma')^*(ds^2) = \frac{4r^4}{(r^2+\zeta\bar{\zeta})^2} d\zeta d\bar{\zeta}$$

which converges to (four times) the Euclidean metric as $r \longrightarrow \infty$.

Here is how the antipodal map on $S^2(r)$ appears in the stereographic picture of $\mathbb{P}^1_{\mathbb{C}}$:

Exercise 1.3.5 *Show that the diagram*

$$\begin{array}{ccc} \mathbb{P}^1_{\mathbb{C}} & \xrightarrow{\Sigma} & S^2(r) \\ {\scriptstyle \alpha_{\mathfrak{E}}}\downarrow & & \downarrow{\scriptstyle \alpha_O} \\ \mathbb{P}^1_{\mathbb{C}} & \xrightarrow[\Sigma]{} & S^2(r) \end{array}$$

commutes, where $\alpha_{\mathfrak{E}}$ denotes the elliptic anti-involution

$$\begin{aligned} \alpha_{\mathfrak{E}} : \mathbb{P}^1_{\mathbb{C}} &\longrightarrow \mathbb{P}^1_{\mathbb{C}} \\ \zeta &\longmapsto -1/\bar{\zeta} = -\frac{\zeta}{|\zeta|^2}. \end{aligned}$$

Given a point $u \in E^2(r)$, the Euclidean reflection ι_u in the plane $u^\perp$ is an isometry and hence its fixed point set—the great circle $u^\perp \cap S^2(r)$—is a totally geodesic submanifold. Since it has dimension 1, it is actually a geodesic. Furthermore the reflection ι_u commutes with the antipodal map. Thus the hyperbolic anti-involution

$$\Sigma^{-1} \circ \alpha_u \circ \Sigma$$

of $\mathbb{P}^1_{\mathbb{C}}$ commutes with the elliptic anti-involution

$$\Sigma^{-1} \circ \alpha_O \circ \Sigma.$$

Here is a collection of characterizations of great circles:

Exercise 1.3.6 *Let C be a circle. Then the following conditions are equivalent:*

1. *$\Sigma(C)$ is a great circle on $S^2(r)$.*
2. *C is invariant under the elliptic anti-involution $\alpha_{\mathfrak{E}}$.*
3. *The hyperbolic involution fixing C commutes with $\alpha_{\mathfrak{E}}$.*
4. *$C = \mathbb{P}(F)$ where $F \subset \mathbb{V}$ is totally real 2-plane. that is a plane such that the restriction of the Hermitian form $\langle\langle , \rangle\rangle$ to F is real.*
5. *The indefinite Hermitian matrix h_C defining the antipolarity corresponding to C has trace 0.*
6. *C is either a straight line containing the origin or a Euclidean circle whose center ζ and radius R satisfy*

$$R^2 = 1 + |\zeta|^2.$$

In the stereographic representation of $\mathbb{P}^1_{\mathbb{C}}$ the spherical distance can be computed using (1.9). For simplicity we consider the distance between the origin O (corresponding to the South pole) and the point $\Sigma(\zeta) \in S^2(r)$.

Exercise 1.3.7 *The spherical distance between $\Sigma(0)$ and $\Sigma(\zeta)$ equals $2r\tan^{-1}|\zeta|$.*

The distance may also be expressed in terms of the cross-ratio and the elliptic anti-involution:

Exercise 1.3.8 *If $x \neq y \in \mathbb{P}^1_{\mathbb{C}}$ are distinct, then $x, y, \alpha_{\mathfrak{E}}(x), \alpha_{\mathfrak{E}}(y)$ lie on a circle and the spherical distance $d(x, y)$ between $\Sigma(x), \Sigma(y)$ satisfies*

$$\mathbf{X}\{x, y, \alpha_{\mathfrak{E}}(x), \alpha_{\mathfrak{E}}(y)\} = \cos^2\left(\frac{d}{2r}\right). \tag{1.12}$$

1.3.3 *The Fubini–Study metric*

Yet another formula for the spherical distance involves the Hermitian structure on $\mathbb{C}^2$, similar to Cayley's formula (1.9). This formula generalizes to complex hyperbolic space (see (3.5)) and we discuss it here for complex elliptic space in arbitrary dimension n.

Consider $\mathbb{C}^{n+1}$ with the standard (positive definite) Hermitian inner product:

$$\langle\langle x, y \rangle\rangle = \sum_{j=1}^{n+1} x_j \bar{y}_j.$$

(We use double angled brackets to denote the standard positive definite Hermitian form.) Then, analogous to (1.9), define a map $d : \mathbb{P}^n_{\mathbb{C}} \times \mathbb{P}^n_{\mathbb{C}} \longrightarrow \mathbb{R}$ by

$$\cos^2\left(\frac{d(x,y)}{r}\right) = \frac{\langle\!\langle x,y\rangle\!\rangle\langle\!\langle y,x\rangle\!\rangle}{\langle\!\langle x,x\rangle\!\rangle\langle\!\langle y,y\rangle\!\rangle} \tag{1.13}$$

where $x, y \in \mathbb{C}^{n+1} - \{0\}$ are nonzero vectors representing elements $[x],[y] \in \mathbb{P}^n_{\mathbb{C}}$ respectively. This expression is homogeneous and hence is well defined on $\mathbb{P}^n_{\mathbb{C}}$.

Exercise 1.3.9 *Prove that d satisfies the triangle inequality and defines the structure of a metric space on $\mathbb{P}^n_{\mathbb{C}}$. Furthermore show this distance function is the distance function defined by the Riemannian structure (1.11) (for $r = 1$).*

This distance was introduced by Fubini [54] and Study [157], and discussed in Coolidge [30]. The resulting *Fubini–Study metric* extends to a Kähler structure on complex projective space. *Complex elliptic n-space* $\mathbf{E}^n_{\mathbb{C}}$ is complex projective space with this structure.

For example, if $\zeta \in \mathbb{C}$, then the Fubini–Study distance between ζ and the origin O equals

$$\rho(0,\zeta) = 2\tan^{-1}|\zeta|.$$

The *metric circle* centered at x of radius ρ consists of all points at distance ρ from x. The preceding discussion implies that the metric circle C_ρ centered at O having Fubini–Study radius ρ is the Euclidean circle centered at O having (Euclidean) radius

$$R = \tan\left(\frac{\rho}{2r}\right).$$

By (1.11), the circumference of C_ρ equals

$$\int_{C_\rho} ds = 2\pi r \sin\left(\frac{\rho}{r}\right). \tag{1.14}$$

Integrating (1.14) with respect to ρ, the area of the disc D_ρ enclosed by C_ρ equals

$$\int_{D_\rho} dA = 2\pi r^2\left(1 - \cos\left(\frac{\rho}{r}\right)\right) = 4\pi r^2 \sin^2\left(\frac{\rho}{2r}\right).$$

We compute the geodesic curvature κ of C_ρ using the Gauss–Bonnet theorem. Since C_ρ is invariant under a transitive group of rotations about O, its geodesic curvature is constant. Since the Gaussian curvature K equals r^{-2}, the Gauss-Bonnet theorem states

$$\begin{aligned} 2\pi = 2\pi\chi(D_\rho) &= \int_{C_\rho} \kappa\, ds + \int_{D_\rho} K\, dA \\ &= 2\pi r \sin\left(\frac{\rho}{r}\right)\kappa + 4\pi \sin^2\left(\frac{\rho}{2r}\right). \end{aligned}$$

so

$$\kappa = r^{-1} \cot\left(\frac{\rho}{r}\right).$$

In particular when $\rho = \pi r/2$, the metric circle C_ρ is itself a geodesic.

By Exercise 1.2.4, the circle C_ρ is associated to a 2-dimensional real subspace $S \subset \mathbb{C}^2$. Such subspaces S possess an invariant, the *holomorphy angle,* which relates to the distance ρ and the geodesic curvature κ. (This invariant is described in more detail in §2.2.1.) Suppose that $S \subset \mathbb{C}^2$ is an $\mathbb{R}$-linear subspace of (real) dimension 2. The real part of the Hermitian form $\langle\langle , \rangle\rangle$

$$((u, v)) = \mathrm{Re}\langle\langle u, v\rangle\rangle$$

is the positive definite $\mathbb{R}$-valued inner product on $\mathbb{C}^2$. Choose a basis $v_1, v_2 \in S$ which is orthonormal with respect to $((,))$:

$$((v_i, v_j)) = \delta_{ij}.$$

Then $\langle\langle v_1, v_2\rangle\rangle$ is purely imaginary and there exists μ with $0 \leq \mu \leq \pi/2$ such that

$$\langle\langle v_1, v_2\rangle\rangle = \pm i \cos(\mu).$$

(μ is the smallest angle between vectors in S and vectors in the image iS of S under the complex structure on $\mathbb{C}^2$; compare §2.2.1 for a more complete discussion.)

The holomorphy angle μ equals 0 if and only if S is a complex line (in which case $\mathbb{P}(S)$ is a single point). $\mu = \pi/2$ if and only if f is totally real (see Exercise 1.3.6), that is $\mathbb{P}(S)$ is a great circle.

Theorem 1.3.10 *The radius of the metric circle $\mathbb{P}(S)$ equals the holomorphy angle μ.*

Proof In a suitable set of coordinates we may assume

$$v_1 = \begin{bmatrix} \sin(\mu/2) \\ \cos(\mu/2) \end{bmatrix}, \qquad v_2 = \begin{bmatrix} -i\sin(\mu/2) \\ i\cos(\mu/2) \end{bmatrix}$$

and S consists of all

$$\begin{bmatrix} \sin(\mu/2)\bar{\xi} \\ \cos(\mu/2)\xi \end{bmatrix}$$

where $\xi \in \mathbb{C}$. Its projectivization consists of all

$$z = \tan\left(\frac{\mu}{2}\right)\frac{\bar{\xi}}{\xi}$$

where ξ ranges over the nonzero complex numbers. Evidently $[f]$ consists of all z with $|z| = \tan(\mu/2)$, which is the metric circle of radius μ about O. $\square$

1.3.4 *Bisectors*

In complex dimension 1, the *metric bisectors* are also geodesics. Let (X, d) be a metric space and let $a, b \in X$ be distinct points. The *bisector equidistant from a and b* is defined to be the set

$$\mathfrak{E}\{a, b\} = \{z \in X \mid \rho(a, z) = \rho(b, z)\}$$

of point of equal distance from a and b. It follows from the distance formula (1.13) that when $(X, d) = \mathbf{E}^1_{\mathbb{C}}$

$$\mathfrak{E}\{[A], [B]\} = \big\{[Z] \mid Z \in \mathbb{C}^2, \quad H_{A,B}(Z, Z) = 0\big\}$$

where $H_{A,B}$ is the Hermitian form on $\mathbb{C}^2$ defined by

$$H_{A,B}(Z, W) = \langle\!\langle A, A\rangle\!\rangle^{-1}\langle\!\langle Z, A\rangle\!\rangle\langle\!\langle A, W\rangle\!\rangle - \langle\!\langle B, B\rangle\!\rangle^{-1}\langle\!\langle Z, B\rangle\!\rangle\langle\!\langle B, W\rangle\!\rangle.$$

In particular $\mathfrak{E}\{[A], [B]\}$ corresponds to the Hermitian matrix

$$F_{A,B} = \langle\!\langle A, A\rangle\!\rangle^{-1}\bar{A}^{\dagger}A - \langle\!\langle B, B\rangle\!\rangle^{-1}\bar{B}^{\dagger}B.$$

The Hermitian matrix corresponding to $H_{A,B}$ has trace zero. Observe that this condition is not invariant under $\mathbf{GL}(2, \mathbb{C})$, but only under the $\mathbf{U}(2)$, the stabilizer of the elliptic anti-polarity. The trace of the matrix representing a bilinear (or Hermitian) form is not intrinsic. However, the trace of a *linear transformation* is intrinsic. Thus if F_1, F_2 are Hermitian matrices, with one of them, say F_2, nondegenerate, then the quantity

$$\operatorname{trace}(F_1 F_2^{-1})$$

is invariant under the action of $\mathbf{GL}(2, \mathbb{C})$ on such pairs (F_1, F_2). Moreover, taking $F_2 = \mathbb{I}$ and $F_1 = H$, we see that $\operatorname{trace}(H)$ is invariant under the subgroup of $\mathbf{GL}(2, \mathbb{C})$ stabilizing the Hermitian form represented by the identity matrix.

Hermitian matrices corresponding to the great circles, the circle

$$|\zeta - \zeta_0| = \sqrt{1 + \zeta_0\bar{\zeta}_0}$$

and the straight line $e^{i\theta}\mathbb{R}$, respectively, are

$$F_{\zeta_0} = \begin{bmatrix} 1 & \bar{\zeta}_0 \\ \zeta_0 & 1 \end{bmatrix}, \qquad F_{\theta} = \begin{bmatrix} 0 & ie^{-i\theta} \\ -ie^{i\theta} & 0 \end{bmatrix}.$$

1.3.5 *Trigonometry*

We briefly review trigonometry in the complex elliptic line. Consider a triangle $\triangle$ with vertices $[\tilde{A}], [\tilde{B}], [\tilde{C}]$ represented by vectors $\tilde{A}, \tilde{B}, \tilde{C} \in \mathbb{V}$. Denote the side lengths by a, b, c and the angles by $0 < \alpha, \beta, \gamma < \pi$. Applying an automorphism we may assume that

$$\tilde{C} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \tilde{A} = \begin{bmatrix} \sin(b/2) \\ \cos(b/2) \end{bmatrix}, \qquad \tilde{B} = \begin{bmatrix} \sin(a/2)e^{i\gamma} \\ \cos(a/2) \end{bmatrix}.$$

The *first cosine law* is obtained from computing $c = d(A, B)$ by the distance formula (1.13):

$$\begin{aligned}
\frac{1+\cos(c)}{2} &= \cos^2(c/2) \\
&= |\sin(a/2)\sin(b/2)e^{i\gamma} + \cos(a/2)\cos(b/2)|^2 \\
&= (\sin(a/2)\sin(b/2)\cos(\gamma) + \cos(a/2)\cos(b/2))^2 \\
&\qquad + (\sin(a/2)\sin(b/2)\sin(\gamma))^2 \\
&= \sin^2(a/2)\sin^2(b/2) \\
&\qquad + 2\sin(a/2)\sin(b/2)\cos(\gamma)\cos(a/2)\cos(b/2) \\
&\qquad\qquad + \cos^2(a/2)\cos^2(b/2))^2 \\
&= \frac{1-\cos(a)}{2}\frac{1-\cos(b)}{2} \\
&\qquad + \frac{\cos(\gamma)\sin(a)\sin(b)}{2} \\
&\qquad\qquad + \frac{1+\cos(a)}{2}\frac{1+\cos(b)}{2} \\
&= \frac{1}{2} + \frac{1}{2}\left(\cos(a)\cos(b) + \cos(\gamma)\sin(a)\sin(b)\right)
\end{aligned}$$

obtaining

$$\cos(c) = \cos(a)\cos(b) + \cos(\gamma)\sin(a)\sin(b).$$

The Pythagorean theorem now follows, by taking $\gamma = \pi/2$:

$$\cos(c) = \cos(a)\cos(b)$$

for any right triangle with hypotenuse c and other sides a, b.

The law of sines can be deduced from the cosine law as follows (see Beardon [9], §7.12, upon which this is based):

$$\begin{aligned}\left(\frac{\sin(c)}{\sin(\gamma)}\right)^2 &= \frac{\sin^2(c)}{1-((\cos(a)\cos(b)-\cos(c))/(\sin(a)\sin(b)))^2}\\ &= \frac{\sin^2(a)\sin^2(b)\sin^2(c)}{\sin^2(a)\sin^2(b)-(\cos(a)\cos(b)-\cos(c))^2}\\ &= \frac{\sin^2(a)\sin^2(b)\sin^2(c)}{(1-\cos^2(a))(1-\cos^2(b))-(\cos(a)\cos(b)-\cos(c))^2}\\ &= \frac{\sin^2(a)\sin^2(b)\sin^2(c)}{1-\cos^2(a)-\cos^2(b)-\cos^2(c)+2\cos(a)\cos(b)\cos(c)}\end{aligned}$$

is symmetrical in a, b, c and therefore

$$\frac{\sin(c)}{\sin(\gamma)} = \frac{\sin(a)}{\sin(\alpha)} = \frac{\sin(b)}{\sin(\beta)}.$$

1.3.6 *An area formula*

Here is an algebraic formula for the area for a geodesic triangle in $\mathbf{E}^1_{\mathbb{C}}$. Suppose that A, B, C are points in $\mathbf{E}^1_{\mathbb{C}}$, which are distinct and no two of them are antipodal. As no pair of vertices is antipodal, there is a unique minimizing geodesic segment between any pair of vertices.

The area of a geodesic triangle in $\mathbf{E}^1_{\mathbb{C}}$ equals the *angular defect*

$$(\alpha + \beta + \gamma) - \pi$$

(compare the discussion in Coxeter [35], [34], [146]). Here is an alternative formula, for which I have been unable to find a reference. The quantity

$$\langle\langle \tilde{A}, \tilde{B}, \tilde{C}\rangle\rangle := \langle\langle \tilde{A}, \tilde{B}\rangle\rangle\langle\langle \tilde{B}, \tilde{C}\rangle\rangle\langle\langle \tilde{C}, \tilde{A}\rangle\rangle \tag{1.15}$$

is a nonzero complex number which scales by the positive real number $|\alpha\beta\gamma|^2$ when $\tilde{A}, \tilde{B}, \tilde{C}$ are replaced by $\alpha\tilde{A}, \beta\tilde{B}, \gamma\tilde{C}$ respectively. Thus its argument

$$\arg\langle\langle \tilde{A}, \tilde{B}, \tilde{C}\rangle\rangle$$

depends only on the equivalence classes $A, B, C \in \mathbb{P}^1_{\mathbb{C}}$, and is denoted $\mathbb{A}(A, B, C)$. (See §2.2.5 for a more detailed discussion.)

Exercise 1.3.11 *Let $A, B, C \in \mathbb{P}^1_{\mathbb{C}}$ be three points, no two of which are antipodal. Let $\triangle$ be the triangle with vertices A, B, C and sides the unique minimizing geodesic segments joining them. Then the area of $\triangle$ equals $2|\mathbb{A}(A, B, C)|$.*

1.4 Hyperbolic geometry

1.4.1 *The complex hyperbolic line*

The *complex hyperbolic line* $\mathbf{H}^1_{\mathbb{C}}$ (or *complex hyperbolic 1-space*) is defined analogously to $\mathbf{E}^1_{\mathbb{C}}$, except that the elliptic anti-polarity is replaced by a hyperbolic anti-polarity η. The null locus $N(\eta)$ is a circle, the *absolute* of this model of $\mathbf{H}^1_{\mathbb{C}}$.

A hyperbolic anti-polarity corresponds to an indefinite Hermitian form, which we now fix as the Hermitian form

$$\langle Z, W\rangle = Z_1\bar{W}_1 - Z_2\bar{W}_2 = \bar{W}^{\dagger}\mathbb{I}_{1,1}Z$$

where

$$\mathbb{I}_{1,1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

We refer to the vector space $\mathbb{C}^2$ with the structure defined by this Hermitian form as the *Hermitian vector space* $\mathbb{C}^{1,1}$.

A vector $Z \in \mathbb{C}^{1,1}$ is said to be *negative* (respectively *null, positive*) if and only if $\langle Z, Z\rangle < 0$ (respectively $\langle Z, Z\rangle = 0$, $\langle Z, Z\rangle > 0$). A line $l \subset \mathbb{C}^{1,1}$ is *negative* (respectively *null, positive*) if l consists of negative (respectively null, positive) vectors. The subset of $\mathbb{P}(\mathbb{C}^{1,1})$ consisting of negative lines (respectively null lines) is defined to be the *complex hyperbolic 1-space* $\mathbf{H}^1_{\mathbb{C}}$ (respectively the *absolute* $\partial\mathbf{H}^1_{\mathbb{C}}$). In the Gaussian representation of $\mathbb{P}(\mathbb{C}^{1,1}) = \mathbb{P}^1_{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, the complex hyperbolic line is the *unit disc* $|\zeta| < 1$.

Exercise 1.4.1 *If $v, w \in \mathbb{C}^{1,1}$ are negative vectors, then*

$$\frac{\langle v, w\rangle\langle w, v\rangle}{\langle v, v\rangle\langle w, w\rangle} \geq 1 \tag{1.16}$$

and depends only on the lines $[v], [w] \in \mathbf{H}^1_{\mathbb{C}}$.

In analogy with (1.13), define the *hyperbolic distance* between two points $[u], [v] \in \mathbf{H}^1_{\mathbb{C}}$ by

$$\cosh^2\left(\frac{d([u],[v])}{2}\right) = \frac{\langle v, w\rangle\langle w, v\rangle}{\langle v, v\rangle\langle w, w\rangle}$$

This defines a metric on $\mathbf{H}^1_{\mathbb{C}}$ whose infinitesimal form is *the Poincaré metric*

$$\frac{4d\zeta d\bar{\zeta}}{(1-\zeta\bar{\zeta})^2}$$

which has constant curvature -1.

Through every $U, V \in \mathbf{H}^1_{\mathbb{C}}$, there is a circle $c(U, V)$ which is orthogonal to the absolute $\partial\mathbb{C}^{1,1}$. The unique anti-involution fixing that circle preserves the absolute and thus the intersection $\mathbf{H}^1_{\mathbb{C}} \cap c(U, V)$ is a geodesic in $\mathbf{H}^1_{\mathbb{C}}$. Furthermore

$$c(U, V) \cap \partial\mathbf{H}^1_{\mathbb{C}}$$

consists of two points U_∞, V_∞ and the hyperbolic distance $d(U, V)$ can be expressed in terms of the cross-ratio:

$$\mathbf{X}\{U_\infty, V_\infty, U, V\} = e^{d(U,V)}.$$

Exercise 1.4.2 *Find a formula, similar to (1.12), expressing the distance in terms of the hyperbolic anti-involution and cross-ratio.*

The subgroup $\mathbf{U}(1,1)$ of unitary automorphisms of $\mathbb{C}^{1,1}$ acts by projective transformations preserving the absolute and acts isometrically on $\mathbf{H}^1_{\mathbb{C}}$.

Exercise 1.4.3 *Sometimes an indefinite Hermitian form given by a matrix which is anti-diagonal is more convenient than a diagonal Hermitian matrix. For example, for the indefinite Hermitian form*

$$\langle Z, W\rangle = Z_1\bar{W}_2 + Z_2\bar{W}_1$$

the corresponding unitary group identifies with $\mathrm{SL}(2,\mathbb{R})$ *and the absolute corresponds to the real projective line* $\mathbb{P}^1_{\mathbb{R}} \subset \mathbb{P}^1_{\mathbb{C}}$.

Geodesics in $\mathbf{H}^1_{\mathbb{C}}$ are also represented by indefinite Hermitian matrices (corresponding to hyperbolic anti-polarities) in a way similar to the parametrization of great circles by indefinite Hermitian matrices. For example, the Hermitian matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

represents the geodesic γ_I corresponding to the imaginary axis $i\mathbb{R}$ in affine coordinates.

Suppose $\gamma \subset \mathbf{H}^1_{\mathbb{C}}$ is an arbitrary geodesic. The origin $O \in \mathbf{H}^1_{\mathbb{C}}$ is represented by 0 in affine coordinates or, equivalently, the negative vector

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

Let $p = p(\gamma)$ be the point on γ closest to O. Let δ be the distance from O to p and let θ be the angle that the geodesic from O to p makes with the intersection

$$\mathbf{H}^1_{\mathbb{R}} = \mathbf{H}^1_{\mathbb{C}} \cap \mathbb{P}^1_{\mathbb{R}}$$

(if $p = O$, then $\pi/2 - \theta$ is defined as the angle between $\mathbf{H}^1_{\mathbb{R}}$ and γ). Then a Hermitian matrix F_γ such that $\gamma \subset N(F_\gamma)$ is

$$F_\gamma = \bar{A}^\dagger \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A = \begin{bmatrix} -\sinh(\delta) & e^{i\theta}\cosh(\delta) \\ e^{-i\theta}\cosh(\delta) & -\sinh(\delta) \end{bmatrix}$$

where the element

$$A = \begin{bmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{bmatrix} \begin{bmatrix} \cosh(\delta/2) & -\sinh(\delta/2) \\ -\sinh(\delta/2) & \cosh(\delta/2) \end{bmatrix}$$

of $\mathbf{U}(1,1)$ represents an isometry of $\mathbf{H}^1_{\mathbb{C}}$ taking the vertical geodesic γ_I in affine coordinates) to γ.

This gives a convenient model for the geodesics in $\mathbf{H}^1_{\mathbb{C}}$. If $\delta > 0$, then p completely determines θ. If $\delta = 0$, then θ is arbitrary, with ambiguity determined up to ± 1. It follows that the *2-dimensional real manifold consisting of geodesics in* $\mathbf{H}^1_{\mathbb{C}}$ *identifies with the unit disc in* $\mathbb{C}$ *blown up at the origin.*

1.4.2 *Hypercycles, horocycles and metric circles*

Metric circles in $\mathbf{H}^1_{\mathbb{C}}$ are represented by Euclidean circles. A metric circle of radius ρ has geodesic curvature $\coth(\rho)$.

Exercise 1.4.4 *Let $u : \mathbf{H}^1_{\mathbb{C}} \longrightarrow \mathbb{R}$ denote distance to the origin and θ the usual angular coordinate on the unit disc. Describe the Poincaré metric tensor*

$$\mathrm{g} = \frac{4dzd\bar{z}}{(1 - z\bar{z})^2} = du^2 + \sinh(u)^2 d\theta^2$$

to obtain hyperbolic polar coordinates

$$z = e^{i\theta}\tanh(u/2)$$

on $\mathbf{H}^1_{\mathbb{C}}$.

Metric circles limit to *horocycles,* which we may regard as "circles centered at infinity." For example, choose a point $p \in \mathbf{H}^1_{\mathbb{C}}$ and a line l in the tangent space at p. Let $z_n \in \mathbf{H}^1_{\mathbb{C}}$ be a path in $\mathbf{H}^1_{\mathbb{C}}$ which converges to a point z_∞ on the absolute. Let C_n be the unique circle centered at z_n which passes through p and is tangent to l at p. Then the sequence C_n converges (for example, in the Hausdorff topology) to $C_\infty \cup \{z_\infty\}$ where C_∞ is a horocycle.

Horocycles have geodesic curvature 1 and can be regarded as limits of *hypercycles* or *equidistant curves.* Let γ be a geodesic. For any $\delta > 0$, the set of points $x \in \mathbf{H}^1_{\mathbb{C}}$ with $d(x, \gamma) < \delta$ has two boundary curves, each of which is a δ-equidistant curve to γ. Such a curve C will be called a *hypercycle* and has geodesic curvature $\tanh(\delta)$. The two endpoints of a hypercycle are the two endpoints of γ and we say that C and γ are *parallel.*

A circle C in $\mathbb{P}^1_{\mathbb{C}}$ intersects $\mathbf{H}^1_{\mathbb{C}}$ in one of four possibilities:

1. The empty set.
2. A metric circle if $C \subset \mathbf{H}^1_{\mathbb{C}}$ is disjoint from the absolute.
3. A horocycle if $C \neq \partial\mathbf{H}^1_{\mathbb{C}}$ and $C \subset \mathbf{H}^1_{\mathbb{C}} \cup \partial\mathbf{H}^1_{\mathbb{C}}$. (In this case C intersects the absolute in one point.)
4. A hypercycle if C meets each of the two components of $\mathbb{P}^1_{\mathbb{C}} - \partial\mathbf{H}^1_{\mathbb{C}}$.

In the last case C intersects the absolute in two points. The angle of intersection θ is related to the distance δ by the following formulas:

$$\delta = \log(\cot(\theta/2)), \qquad \cot(\theta) = \sinh(\delta).$$

Exercise 1.4.5 *Given a geodesic γ, and a point $p \in \mathbf{H}^1_{\mathbb{C}}$, there exists a unique hypercycle $C(p;\gamma)$ through p and parallel to γ. If γ_n is a sequence of geodesics converging to an ideal point $q_\infty \in \partial\mathbf{H}^1_{\mathbb{C}}$, then show that the hypercycles $C(p;\gamma_n)$ converge to the unique horocycle centered at q_∞ passing through p.*

Exercise 1.4.6 *Let S be a 1-parameter subgroup of $\mathbf{PU}(1,1)$ and let $p \in \mathbf{H}^1_{\mathbb{C}}$. Show that the orbit $S(p)$ is*

1. *A metric circle (possibly the point $\{p\}$) if S is elliptic.*
2. *A horocycle if S is parabolic.*
3. *A hypercycle (possibly a geodesic) if S is hyperbolic.*

In the first two cases, the center of $S(p)$ is the fixed point of S, and in the last case, $S(p)$ is a hypercycle parallel to the unique geodesic invariant under S.

1.4.3 *Trigonometry*

The formulas of elliptic trigonometry have analogues in hyperbolic trigonometry, and we briefly present these formulas. The proofs are analogous to those in §1.3.5 and left as an exercise. We consider the complex hyperbolic line $\mathbf{H}^1_{\mathbb{C}}(R)$ with curvature $-r^2$. For more details, see Coxeter [35, 34], Fenchel [52] and Thurston [161]. The *first cosine law* in $\mathbf{H}^1_{\mathbb{C}}$ is

$$-\cosh(c/r) + \cosh(a/r)\cosh(b/r) = \cos(\gamma)\sinh(a/r)\sinh(b/r)$$

(which approaches the Euclidean cosine law

$$c^2 = a^2 + b^2 - 2\cos(\gamma)ab$$

as $r \longrightarrow \infty$). Setting $\gamma = \pi/2$ gives the Pythagorean theorem

$$\cosh(c/r) = \cosh(a/r)\cosh(b/r)$$

for any right triangle with hypotenuse c and other sides a, b. The *second cosine law* in $\mathbf{H}^1_{\mathbb{C}}$ is

$$\cos(\gamma) + \cos(\alpha)\cos(\beta) = \cosh(c/r)\sin(\alpha)\sin(\beta) \tag{1.17}$$

whose limit as $r \longrightarrow \infty$ reduces to

$$\cos(\gamma) = \cos(\alpha + \beta)$$

which follows from the Euclidean relation

$$\alpha + \beta + \gamma = \pi.$$

From the cosine law follows the sine law:

$$\frac{\sinh(c/r)}{\sin(\gamma)} = \frac{\sinh(a/r)}{\sin(\alpha)} = \frac{\sinh(b/r)}{\sin(\beta)}.$$

Here is an algebraic formula for the area for a geodesic triangle in $\mathbf{H}^1_{\mathbb{C}}$. Suppose that A, B, C are points in $\mathbf{H}^1_{\mathbb{C}}$, spanning a triangle $\triangle(A, B, C)$. Its area equals the *angular defect:*

$$\pi - (\alpha + \beta + \gamma)$$

(compare the discussion in Coxeter [35], [34], [146]). Here is an alternative formula, analogous to 1.3.6.

Exercise 1.4.7 *Let $\tilde{A}, \tilde{B}, \tilde{C} \in \mathbb{C}^{n-1}$ be negative vectors representing the vertices A, B, C. As in (1.15),*

$$\langle\!\langle \tilde{A}, \tilde{B}, \tilde{C} \rangle\!\rangle := \langle\!\langle \tilde{A}, \tilde{B} \rangle\!\rangle \langle\!\langle \tilde{B}, \tilde{C} \rangle\!\rangle \langle\!\langle \tilde{C}, \tilde{A} \rangle\!\rangle$$

depends only on the points $A, B, C \in \mathbf{H}^1_{\mathbb{C}} \subset \mathbb{P}^1_{\mathbb{C}}$. Let

$$\mathbb{A}(A, B, C) = \arg \langle\!\langle \tilde{A}, \tilde{B}, \tilde{C} \rangle\!\rangle.$$

The area of $\triangle(A, B, C)$ equals $2|\mathbb{A}(A, B, C)|$.

1.4.3.1 *The angle of parallelism* The duality between angle and length is concretely manifest by the relation of *angle of parallelism.* Namely, choose a point $O \in \mathbf{H}^1_{\mathbb{C}}$ and a complete geodesic $l \subset \mathbf{H}^1_{\mathbb{C}}$. Then the angle of parallelism is the angle 2θ at O between the two rays asymptotic to l, which is related to the distance $d = d(O, l)$ from O to l.

Alternatively, choose the point p on l closest to O; then O and p are two vertices of two ideal right triangle whose sides are rays of l. Then the angle at p of this triangle equals θ, and the finite side has length d.

The angle of parallelism θ and the length d are related by the suggestive formulas

$$\begin{aligned} \sinh(d) &= \cot(2\theta) \\ \cosh(d) &= \csc(2\theta) \\ \tanh(d) &= \cos(2\theta) \\ \coth(d) &= \sec(2\theta) \\ \operatorname{sech}(d) &= \sin(2\theta) \\ \operatorname{csch}(d) &= \tan(2\theta) \\ e^d &= \cot(\theta) \\ e^{-d} &= \tan(\theta). \end{aligned} \tag{1.18}$$

1.4.4 *The right half-plane model*

A closely related model for hyperbolic geometry is the right half-plane $\mathfrak{H}^1$ described by $\operatorname{Re}(w) > 0$. (Although the upper half-plane is more standard, the right half-plane seems to be more convenient for some of the higher-dimensional generalizations.) The embedding

$$\mathbf{B} : w \mapsto \begin{bmatrix} 1/2 - w \\ 1/2 + w \end{bmatrix}$$

maps $\mathfrak{H}^1 \longrightarrow \mathbb{C}^{1,1}$. Now

$$\langle \mathbf{B}(w_1), \mathbf{B}(w_2) \rangle = -(w_1 + \bar{w}_2),$$

and in particular $\langle \mathbf{B}(w), \mathbf{B}(w)\rangle = -2\mathrm{Re}(w)$. We choose coordinates $t, \delta \in \mathbb{R}$ such that

$$w = e^t(\mathrm{sech}(\delta) + i\tanh(\delta)).$$

The Riemannian metric on $\mathfrak{H}^1$ is

$$\mathbf{g} = \frac{|dw|^2}{\mathrm{Re}(w)^2} = \cosh^2(\delta)dt^2 + d\delta^2.$$

The positive real axis $\mathbb{R}^+$ is the geodesic γ_0 from O to ∞ and $\delta(w)$ is the distance from w to γ_0.

A closely related set of coordinates on $\mathfrak{H}^1$ uses the same coordinate δ but replaces t by τ where

$$e^\tau = e^t \mathrm{sech}(\delta)$$

so that

$$w = e^\tau(1 + i\sinh(\delta))$$

and

$$\mathbf{g} = \frac{|dw|^2}{\mathrm{Re}(w)^2} = d\tau^2 + (\sinh(\delta)d\tau + \cosh(\delta)d\delta)^2.$$

While the level sets of t are geodesics orthogonal to γ_0, the level sets of τ are horocycles orthogonal to γ_0.

In this model the hypercycles parallel to γ_0 are rays through the origin. In particular the ray γ_m of slope $-\infty < m < \infty$ satisfies

$$m = \sinh(\delta)$$

where δ denotes the (signed) distance from γ_m to γ_0. The distance along the hypercycle γ_m from parameter t for $t_1 < t < t_2$ is given by

$$\cosh(\delta)(t_2 - t_1)$$

whereas the distance between the points w_1, w_2 is given by

$$\sinh\left(\frac{\rho(w_1, w_2)}{2}\right) = \cosh(\delta)\sinh\left(\frac{(t_2 - t_1)}{2}\right).$$

1.4.5 *The real hyperbolic plane*

An alternative model for an isomorphic geometry is *real hyperbolic 2-space.* This model is quite different, arising from the *real* projective *plane* $\mathbb{P}^2_{\mathbb{R}}$, rather than the *complex* projective *line*. Consider the Lorentzian vector space $\mathbb{R}^{2,1}$, that is a 3-dimensional real vector space with inner product

$$\langle x, y\rangle = x_1y_1 + x_2y_2 - x_3y_3.$$

The corresponding affine space $\mathbb{E}$ has a parallel flat Lorentzian metric

$$g = dx_1dy_1 + dx_2dy_2 - dx_3dy_3$$

and with this structure (E, g) is a simply connected complete Lorentzian manifold (*Minkowski (2+1)-space*) of zero curvature. The set of all points satisfying

$$\langle x, x\rangle = -1$$

is a two-sheeted hyperboloid for which the restriction of g is a Riemannian metric of constant curvature -1.

The Siegel domain for $\mathbf{H}^2_{\mathbb{R}}$ consists of $w = (w_1, w_2) \in \mathbb{R}^2$ satisfying $2w_1 - w_2^2 < 0$. Let

$$f(w) = 2w_1 - w_2^2.$$

Then

$$\begin{aligned} d\log f(w) &= (2w_1 - w_2^2)^{-1}(2dw_1 - 2w_2dw_2) \\ Dd\log f(w) &= D(f^{-1}df) = f^{-1}d^2f - f^{-2}(df)^2 \\ &= \frac{4}{(2w_1 - w_2^2)^2}\left\{(dw_2 - w_1dw_1)^2 + (2w_1 - w_2^2)dw_2^2\right\} \end{aligned}$$

is the metric on $\mathbf{H}^2_{\mathbb{R}}$. Two vector fields whose flows generate the affine automorphism group are

$$w_2\frac{\partial}{\partial w_1} + \frac{\partial}{\partial w_2}, \qquad 2w_2\frac{\partial}{\partial w_1} + w_2\frac{\partial}{\partial w_2}.$$

Exercise 1.4.8 *A projectively equivalent model for $\mathbf{H}^2_{\mathbb{R}}$ is the unit ball $(x_1, x_2) \in \mathbb{R}^2$ where*

$$x_1^2 + x_2^2 < 1.$$

Show that in this model the metric tensor is given by

$$\frac{4}{(1 - x_1^2 - x_2^2)^2}(x_1dx_1 + x_2dx_2)^2 + \frac{4}{1 - x_1^2 - x_2^2}(dx_1^2 + dx_2^2).$$

The trigonometric formulas—in particular the laws of cosines—in $\mathbf{H}^2_{\mathbb{R}}$ can be expressed succinctly and abstractly in terms of 3×3 symmetric matrices as follows. Associated to a triangle are two triples, the *side lengths* a, b, c and *vertex angles*

α, β, γ. These invariants define symmetric bilinear forms on dual vector spaces, which we describe in terms of Gram matrices. The *edge matrix* is the symmetric matrix

$$E = \begin{bmatrix} 1 & \cosh(c) & \cosh(b) \\ \cosh(c) & 1 & \cosh(a) \\ \cosh(b) & \cosh(a) & 1 \end{bmatrix}$$

and the *vertex matrix* equals

$$V = \begin{bmatrix} 1 & -\cos(\gamma) & -\cos(\beta) \\ -\cos(\gamma) & 1 & -\cos(\alpha) \\ -\cos(\beta) & -\cos(\alpha) & 1 \end{bmatrix}.$$

Writing

$$D = \begin{bmatrix} \sinh(a) & 0 & 0 \\ 0 & \sinh(b) & 0 \\ 0 & 0 & \sinh(c) \end{bmatrix}$$

the *matrix form of the law of cosines* is

$$E^{-1} = \det(E)^{-1} D^{\dagger} V D$$

where

$$\det(E) = 1 - \cosh^2(a) - \cosh^2(b) - \cosh^2(c) + 2\cosh(a)\cosh(b)\cosh(c).$$

This matrix equation includes both versions of the laws of cosines, six equations in all. (Compare Coxeter [34], Thurston [161].)