Graduate Texts in Mathematics

John G. Ratcliffe

Foundations of Hyperbolic Manifolds

Springer Science+Business Media, LLC

Graduate Texts in Mathematics 149

Graduate Texts in Mathematics

continued after index

John G. Ratcliffe

Foundations of Hyperbolic Manifolds

With 164 Illustrations

Springer Science+Business Media, LLC

John G. Ratcliffe
Department of Mathematics
Vanderbilt University
Nashville, TN 37240
USA

ISBN 978-0-387-94348-0 ISBN 978-1-4757-4013-4 (eBook)
DOI 10.1007/978-1-4757-4013-4

Mathematics Subject Classifications (1991): 51-01, 51M10, 52A20, 57M50, 20H10, 30F40

Library of Congress Cataloging-in-Publication Data
Ratcliffe, John G., 1948–
Foundations of hyberbolic manifolds / John G. Ratcliffe.
p. cm. – (Graduate texts in mathematics)
Includes bibliographical references and index.

1. Geometry, Hyberbolic. 2. Hyperbolic spaces. I. Title.
II. Series.
QA685.R22 1994
516.9–dc20 94-8958

Printed on acid-free paper.

Production managed by Hal Henglein; manufacturing supervised by Gail Simon.
Photocomposed copy prepared from the author's TeX file.

9 8 7 6 5 4 3 2 1

To Susan, Kimberly, and Thomas

Preface

This book is an exposition of the theoretical foundations of hyperbolic manifolds. It is intended to be used both as a textbook and as a reference. Particular emphasis has been placed on readability and completeness of argument. The treatment of the material is for the most part elementary and self-contained. The reader is assumed to have a basic knowledge of algebra and topology at the first-year graduate level of an American university.

The book is divided into three parts. The first part, consisting of Chapters 1-7, is concerned with hyperbolic geometry and basic properties of discrete groups of isometries of hyperbolic space. The main results are the existence theorem for discrete reflection groups, the Bieberbach theorems, and Selberg's lemma. The second part, consisting of Chapters 8-12, is devoted to the theory of hyperbolic manifolds. The main results are Mostow's rigidity theorem and the determination of the structure of geometrically finite hyperbolic manifolds. The third part, consisting of Chapter 13, integrates the first two parts in a development of the theory of hyperbolic orbifolds. The main results are the construction of the universal orbifold covering space and Poincaré's fundamental polyhedron theorem.

This book was written as a textbook for a one-year course. Chapters 1-7 can be covered in one semester, and selected topics from Chapters 8-12 can be covered in the second semester. For a one-semester course on hyperbolic manifolds, the first two sections of Chapter 1 and selected topics from Chapters 8-12 are recommended. Since complete arguments are given in the text, the instructor should try to cover the material as quickly as possible by summarizing the basic ideas and drawing lots of pictures. If all the details are covered, there is probably enough material in this book for a two-year sequence of courses.

There are over 500 exercises in this book which should be read as part of the text. These exercises range in difficulty from elementary to moderately difficult, with the more difficult ones occurring toward the end of each set of exercises. There is much to be gained by working on these exercises.

An honest effort has been made to give references to the original published sources of the material in this book. Most of these original papers are well worth reading. The references are collected at the end of each chapter in the section on historical notes.

This book is a complete revision of my lecture notes for a one-year course on hyperbolic manifolds that I gave at the University of Illinois during 1984.

I wish to express my gratitude to:

(1) James Cannon for allowing me to attend his course on Kleinian groups at the University of Wisconsin during the fall of 1980;

(2) William Thurston for allowing me to attend his course on hyperbolic 3-manifolds at Princeton University during the academic year 1981-82 and for allowing me to include his unpublished material on hyperbolic Dehn surgery in Chapter 10;

(3) my colleagues at the University of Illinois who attended my course on hyperbolic manifolds, Kenneth Appel, Richard Bishop, Robert Craggs, George Francis, Mary-Elizabeth Hamstrom, and Joseph Rotman, for their many valuable comments and observations;

(4) my colleagues at Vanderbilt University who attended my ongoing seminar on hyperbolic geometry over the last seven years, Mark Baker, Bruce Hughes, Christine Kinsey, Michael Mihalik, Efstratios Prassidis, Barry Spieler, and Steven Tschantz, for their many valuable observations and suggestions;

(5) my colleagues and friends, William Abikoff, Colin Adams, Boris Apanasov, Richard Arenstorf, William Harvey, Linda Keen, Ruth Kellerhals, Victor Klee, Bernard Maskit, Hans Munkholm, Walter Neumann, Alan Reid, Robert Riley, Richard Skora, John Stillwell, Perry Susskind, and Jeffrey Weeks, for their helpful conversations and correspondence;

(6) the library staff at Vanderbilt University for helping me find the references for this book;

(7) Ruby Moore for typing up my manuscript;

(8) the editorial staff at Springer-Verlag New York for the careful editing of this book.

I especially wish to thank my colleague, Steven Tschantz, for helping me prepare this book on my computer and for drawing most of the 3-dimensional figures on his computer.

Finally, I would like to encourage the reader to send me your comments and corrections concerning the text, exercises, and historical notes.

Nashville, June, 1994 JOHN G. RATCLIFFE

Contents

CHAPTER 1

Euclidean Geometry

In this chapter, we review Euclidean geometry. We begin with an informal historical account of how criticism of Euclid's parallel postulate led to the discovery of hyperbolic geometry. In Section 1.2, the proof of the independence of the parallel postulate by the construction of a Euclidean model of the hyperbolic plane is discussed and all four basic models of the hyperbolic plane are introduced. In Section 1.3, we begin our formal study with a review of n-dimensional Euclidean geometry. The metrical properties of curves are studied in Sections 1.4 and 1.5. In particular, the concepts of geodesic and arc length are introduced.

§1.1. Euclid's Parallel Postulate

Euclid wrote his famous Elements around 300 B.C. In this thirteen-volume work, he brilliantly organized and presented the fundamental propositions of Greek geometry and number theory. In the first book of the Elements, Euclid develops plane geometry starting with basic assumptions consisting of a list of definitions of geometric terms, five "common notions" concerning magnitudes, and the following five postulates:

(1) *A straight line may be drawn from any point to any other point.*

(2) *A finite straight line may be extended continuously in a straight line.*

(3) *A circle may be drawn with any center and any radius.*

(4) *All right angles are equal.*

(5) *If a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if extended indefinitely, meet on the side on which the angles are less than two right angles.*

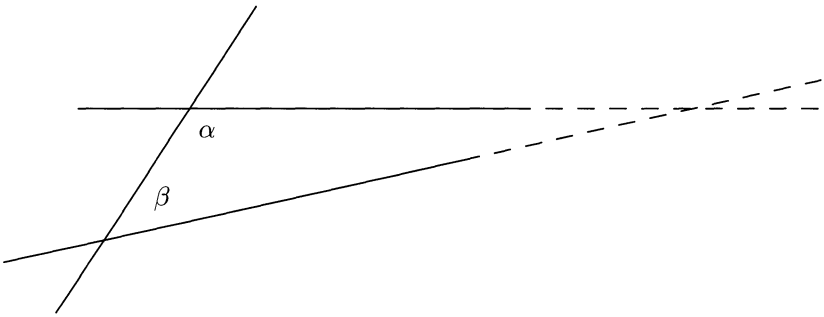


Figure 1.1.1. Euclid's parallel postulate

The first four postulates are simple and easily grasped, whereas the fifth is complicated and not so easily understood. Figure 1.1.1 illustrates the fifth postulate. When one tries to visualize all the possible cases of the postulate, one sees that it possesses an elusive infinite nature. As the sum of the two interior angles $\alpha + \beta$ approaches $180°$, the point of intersection in Figure 1.1.1 moves towards infinity. Euclid's fifth postulate is equivalent to the modern parallel postulate of Euclidean geometry:

> *Through a point outside a given infinite straight line there is one and only one infinite straight line parallel to the given line.*

From the very beginning, Euclid's presentation of geometry in his Elements was greatly admired, and The Thirteen Books of Euclid's Elements became the standard treatise of geometry and remained so for over two thousand years; however, even the earliest commentators on the Elements criticized the fifth postulate. The main criticism was that it is not sufficiently self-evident to be accepted without proof. Adding support to this belief is the fact that the converse of the fifth postulate (the sum of two angles of a triangle is less than $180°$) is one of the propositions proved by Euclid (Proposition 17, Book I). How could a postulate, whose converse can be proved, be unprovable? Another curious fact is that most of plane geometry can be proved without the fifth postulate. It is not used until Proposition 29 of Book I. This suggests that the fifth postulate is not really necessary.

Because of this criticism, it was believed by many that the fifth postulate could be derived from the other four postulates, and for over two thousand years geometers attempted to prove the fifth postulate. It was not until the nineteenth century that the fifth postulate was finally shown to be independent of the other postulates of plane geometry. The proof of this independence was the result of a completely unexpected discovery. The denial of the fifth postulate leads to a new consistent geometry. It was Carl Friedrich Gauss who first made this remarkable discovery.

Gauss began his meditations on the theory of parallels about 1792. After trying to prove the fifth postulate for over twenty years, Gauss discovered that the denial of the fifth postulate leads to a new strange geometry, which he called *non-Euclidean geometry.* After investigating its properties for over ten years and discovering no inconsistencies, Gauss was fully convinced of its consistency. In a letter to F. A. Taurinus, in 1824, he wrote: "The assumption that the sum of the three angles (of a triangle) is smaller than 180° leads to a geometry which is quite different from our (Euclidean) geometry, but which is in itself completely consistent." Gauss's assumption that the sum of the angles of a triangle is less than 180° is equivalent to the denial of Euclid's fifth postulate. Unfortunately, Gauss never published his results on non-Euclidean geometry.

Only a few years passed before non-Euclidean geometry was rediscovered independently by Nikolai Lobachevsky and János Bolyai. Lobachevsky published the first account of non-Euclidean geometry in 1829 in a paper entitled *On the principles of geometry.* A few years later, in 1832, Bolyai published an independent account of non-Euclidean geometry in a paper entitled *The absolute science of space.*

The strongest evidence given by the founders of non-Euclidean geometry for its consistency is the duality between non-Euclidean and spherical trigonometries. In this duality, the hyperbolic trigonometric functions play the same role in non-Euclidean trigonometry as the ordinary trigonometric functions play in spherical trigonometry. Today, the non-Euclidean geometry of Gauss, Lobachevsky, and Bolyai is called *hyperbolic geometry,* and the term *non-Euclidean geometry* refers to any geometry that is not Euclidean.

Spherical-Hyperbolic Duality

Spherical and hyperbolic geometries are oppositely dual geometries. This duality begins with the opposite nature of the parallel postulate in each geometry. The analogue of an infinite straight line in spherical geometry is a great circle of a sphere. Figure 1.1.2 illustrates three great circles on a sphere. For simplicity, we shall use the term *line* for either an infinite straight line in hyperbolic geometry or a great circle in spherical geometry. In spherical geometry, the parallel postulate takes the form:

> *Through a point outside a given line there is no line parallel to the given line.*

The parallel postulate in hyperbolic geometry has the opposite form:

> *Through a point outside a given line there are infinitely many lines parallel to the given line.*

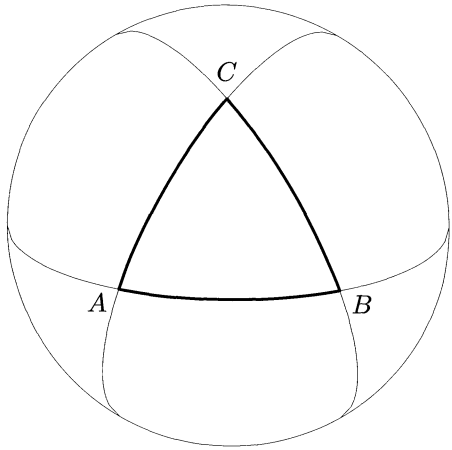


Figure 1.1.2. A spherical equilateral triangle ABC

The duality between spherical and hyperbolic geometries is further evident in the opposite shape of triangles in each geometry. The sum of the angles of a spherical triangle is always greater than 180°, whereas the sum of the angles of a hyperbolic triangle is always less than 180°. As the sum of the angles of a Euclidean triangle is 180°, one can say that Euclidean geometry is midway between spherical and hyperbolic geometries. See Figures 1.1.2, 1.1.3, and 1.1.5 for an example of an equilateral triangle in each geometry.

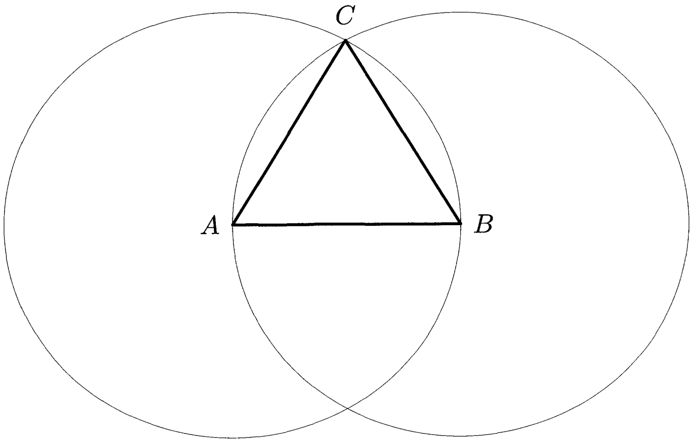


Figure 1.1.3. A Euclidean equilateral triangle ABC

Curvature

Strictly speaking, spherical geometry is not one geometry but a continuum of geometries. The geometries of two spheres of different radii are not metrically equivalent; although they are equivalent under a change of scale. The geometric invariant that best distinguishes the various spherical geometries is Gaussian *curvature*. A sphere of radius r has constant positive curvature $1/r^2$. Two spheres are metrically equivalent if and only if they have the same curvature.

The duality between spherical and hyperbolic geometries continues. Hyperbolic geometry is not one geometry but a continuum of geometries. Curvature distinguishes the various hyperbolic geometries. A *hyperbolic plane* has constant negative curvature, and every negative curvature is realized by some hyperbolic plane. Two hyperbolic planes are metrically equivalent if and only if they have the same curvature. Any two hyperbolic planes with different curvatures are equivalent under a change of scale.

For convenience, we shall adopt the unit sphere as our model for spherical geometry. The unit sphere has constant curvature equal to 1. Likewise, for convenience, we shall work exclusively with models for hyperbolic geometry whose constant curvature is -1. It is not surprising that a Euclidean plane is of constant curvature 0, which is midway between -1 and 1.

The simplest example of a surface of negative curvature is the saddle surface in $\mathbb{R}^3$ defined by the equation $z = xy$. The curvature of this surface at a point (x, y, z) is given by the formula

$$\kappa(x,y,z) = \frac{-1}{(1+x^2+y^2)^2}. \tag{1.1.1}$$

In particular, the curvature of the surface has a unique minimum value of -1 at the saddle point $(0,0,0)$.

There is a well-known surface in $\mathbb{R}^3$ of constant curvature -1. If one starts at $(0,0)$ on the xy-plane and walks along the y-axis pulling a small wagon that started at $(1,0)$ and has a handle of length 1, then the wagon would follow the graph of the *tractrix* (L. trahere, to pull) defined by the equation

$$y = \cosh^{-1}\left(\frac{1}{x}\right) - \sqrt{1-x^2}. \tag{1.1.2}$$

This curve has the property that the distance from the point of contact of a tangent to the point where it cuts the y-axis is 1. See Figure 1.1.4. The surface S obtained by revolving the tractrix about the y-axis in $\mathbb{R}^3$ is called the *tractroid*. The tractroid S has constant negative curvature -1; consequently, the local geometry of S is the same as that of a hyperbolic plane of curvature -1. Figure 1.1.5 illustrates a hyperbolic equilateral triangle on the tractroid S.

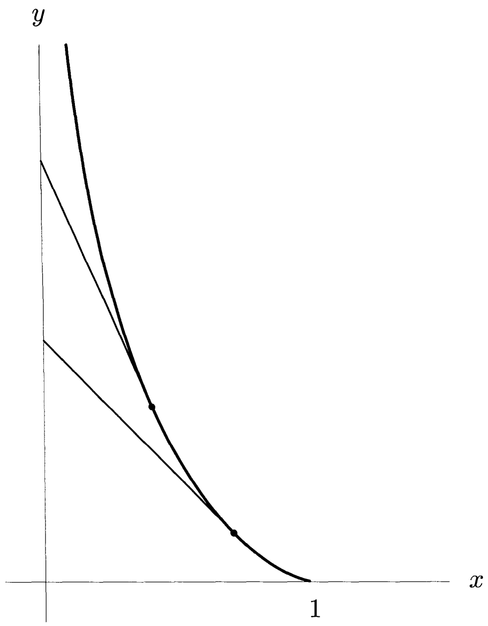


Figure 1.1.4. Two tangents to the graph of the tractrix

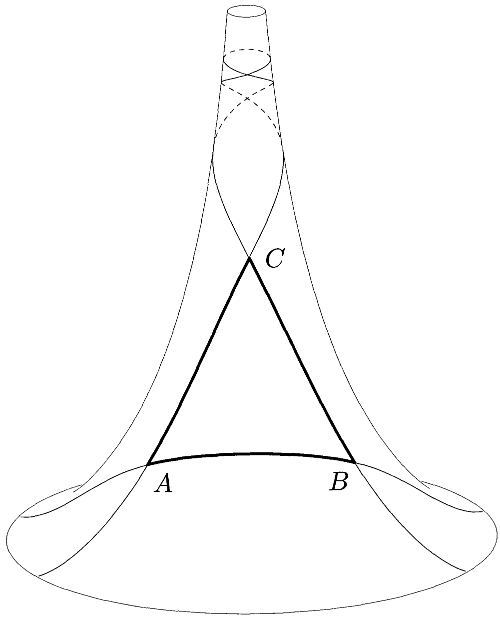


Figure 1.1.5. A hyperbolic equilateral triangle ABC on the tractroid

§1.2. Independence of the Parallel Postulate

After enduring twenty centuries of criticism, Euclid's theory of parallels was fully vindicated in 1868 when Eugenio Beltrami proved the independence of Euclid's parallel postulate by constructing a Euclidean model of the hyperbolic plane. The points of the model are the points inside a fixed circle, in a Euclidean plane, called the *circle at infinity.* The lines of the model are the open chords of the circle at infinity. It is clear from Figure 1.2.1 that *Beltrami's model* has the property that through a point P outside a line L there is more than one line parallel to L. Using differential geometry, Beltrami showed that his model satisfies all the axioms of hyperbolic plane geometry. As Beltrami's model is defined entirely in terms of Euclidean plane geometry, it follows that hyperbolic plane geometry is consistent if Euclidean plane geometry is consistent. Thus, Euclid's parallel postulate is independent of the other postulates of plane geometry.

In 1871, Felix Klein gave an interpretation of Beltrami's model in terms of projective geometry. In particular, Beltrami and Klein showed that the congruence transformations of Beltrami's model correspond by restriction to the projective transformations of the extended Euclidean plane that leave the model invariant. For example, a rotation about the center of the circle at infinity restricts to a congruence transformation of Beltrami's model. Because of Klein's interpretation, Beltrami's model is also called *Klein's model* of the hyperbolic plane. We shall take a neutral position and call this model the *projective disk model* of the hyperbolic plane.

The projective disk model has the advantage that its lines are straight, but it has the disadvantage that its angles are not necessarily the Euclidean angles. This is best illustrated by examining *right angles* in the model.

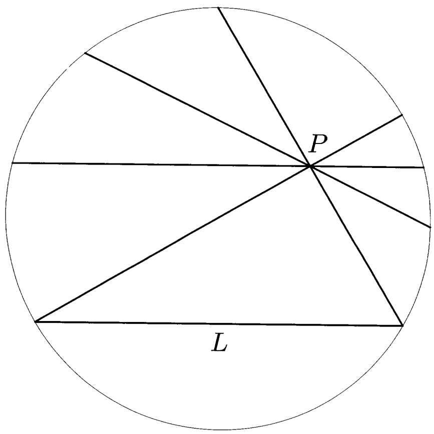


Figure 1.2.1. Lines passing through a point P parallel to a line L

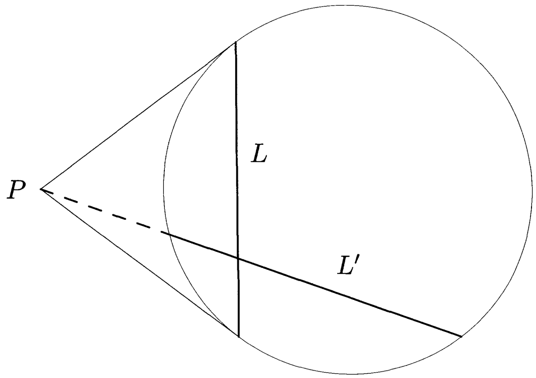


Figure 1.2.2. Two perpendicular lines L and L' of the projective disk model

Let L be a line of the model which is not a diameter, and let P be the intersection of the tangents to the circle at infinity at the endpoints of L as illustrated in Figure 1.2.2. Then a line L' of the model is *perpendicular* to L if and only if the Euclidean line extending L' passes through P. In particular, the Euclidean midpoint of L is the only point on L at which the right angle formed by L and its perpendicular is a Euclidean right angle. We shall study the projective disk model in detail in Chapter 6.

The Conformal Disk Model

There is another model of the hyperbolic plane whose points are the points inside a fixed circle in a Euclidean plane, but whose angles are the Euclidean angles. This model is called the *conformal disk model*, since its angles conform with the Euclidean angles. The lines of this model are the open diameters of the boundary circle together with the open circular arcs orthogonal to the boundary circle. See Figures 1.2.3 and 1.2.4. The hyperbolic geometry of the conformal disk model is the underlying geometry of M.C. Escher's famous circle prints. Figure 1.2.5 is Escher's Circle Limit IV. All the devils (angels) in Figure 1.2.5 are congruent with respect to the underlying hyperbolic geometry. Some appear larger than others because the model distorts distances. We shall study the conformal disk model in detail in Chapter 4.

The projective and conformal disk models both exhibit Euclidean rotational symmetry with respect to their Euclidean centers. Rotational symmetry is one of the two basic forms of Euclidean symmetry; the other is translational symmetry. There is another conformal model of the hyperbolic plane which exhibits Euclidean translational symmetry. This model is called the *upper half-plane model*.

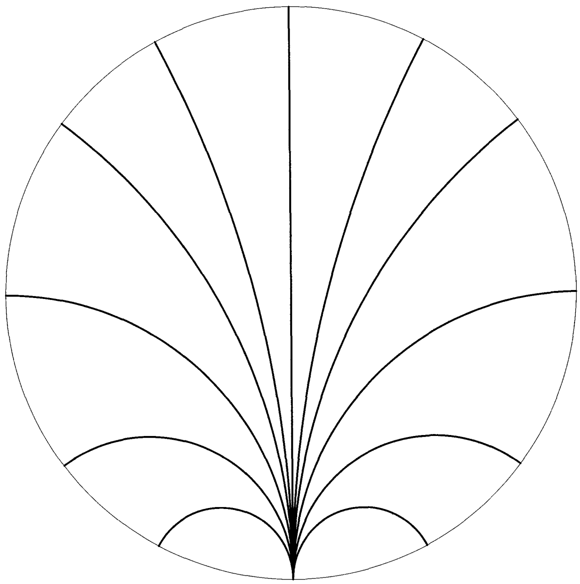


Figure 1.2.3. Asymptotic parallel lines of the conformal disk model

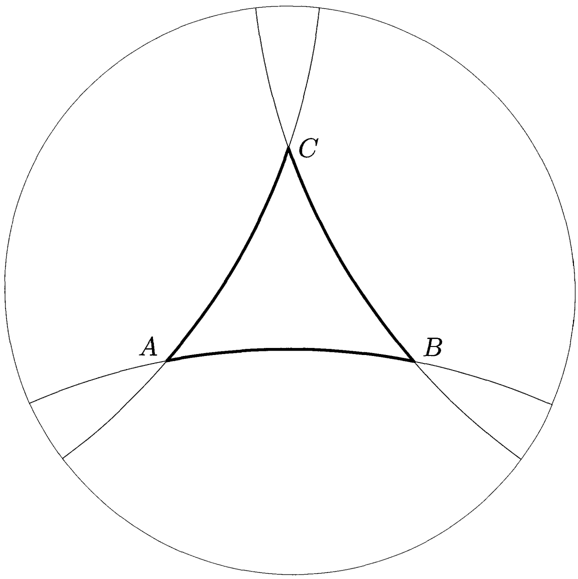


Figure 1.2.4. An equilateral triangle ABC in the conformal disk model

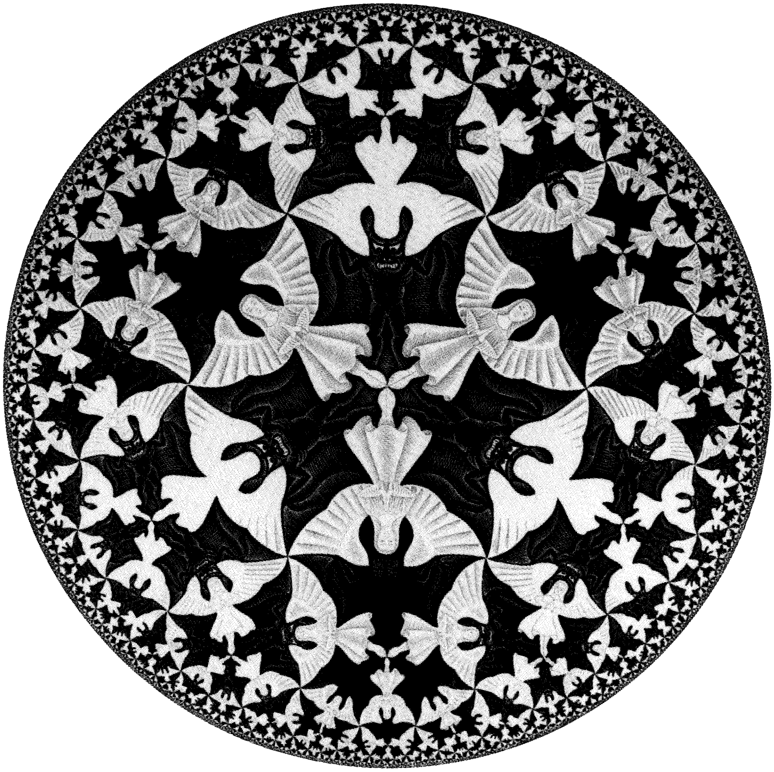


Figure 1.2.5. M. C. Escher: Circle Limit IV

The Upper Half-Plane Model

The points of the upper half-plane model are the complex numbers above the real axis in the complex plane. The lines of the model are the open rays orthogonal to the real axis together with the open semicircles orthogonal to the real axis. See Figures 1.2.6 and 1.2.7. The orientation preserving congruence transformations of the upper half-plane model are the linear fractional transformations of the form

$$\phi(z) = \frac{az+b}{cz+d} \quad \text{with } a, b, c, d \text{ real and } ad - bc > 0.$$

In particular, a Euclidean translation $\tau(z) = z + b$ is a congruence transformation. The upper half-plane model exhibits Euclidean translational symmetry at the expense of an unlimited amount of distortion. Any magnification $\mu(z) = az$, with $a > 1$, is a congruence transformation. We shall study the upper half-plane model in detail in Chapter 4.

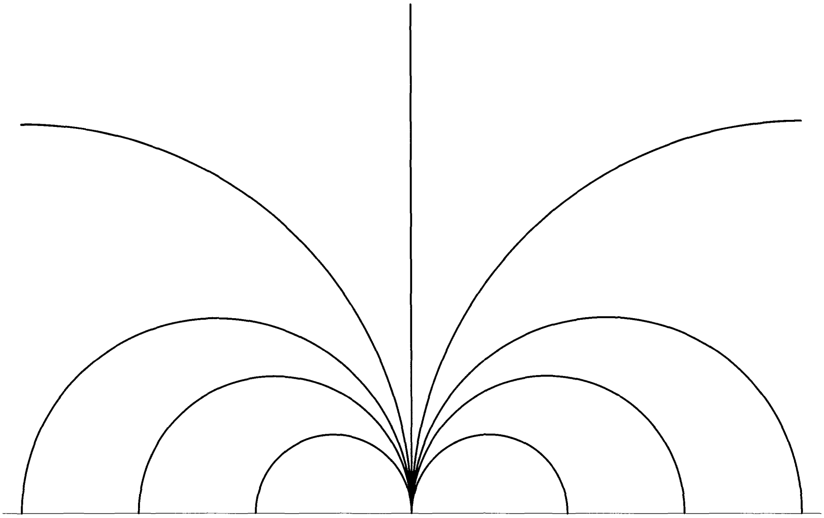


Figure 1.2.6. Asymptotic parallel lines of the upper half-plane model

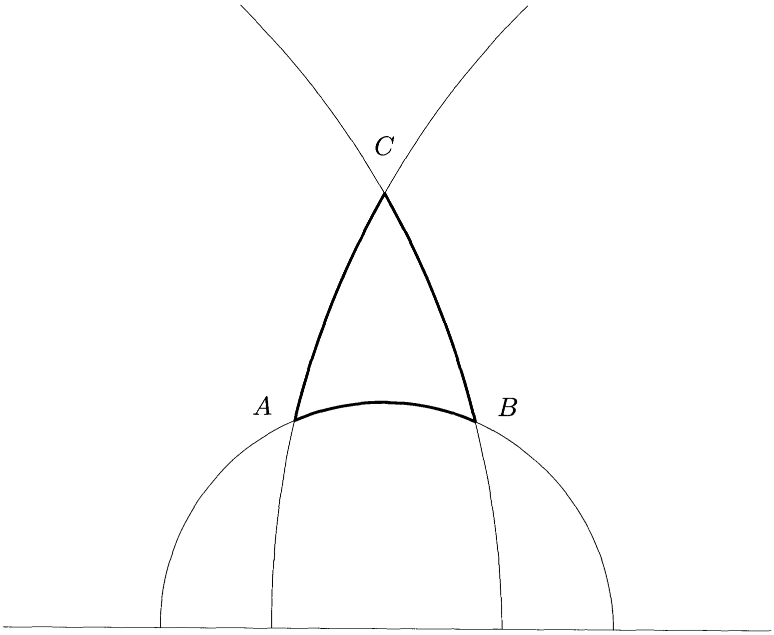


Figure 1.2.7. An equilateral triangle ABC in the upper half-plane model

The Hyperboloid Model

All the models of the hyperbolic plane we have described distort distances. Unfortunately, there is no way we can avoid distortion in a useful Euclidean model of the hyperbolic plane because of a remarkable theorem of David Hilbert that there is no complete C^2 surface of constant negative curvature in $\mathbb{R}^3$. Hilbert's theorem implies that there is no reasonable distortion-free model of the hyperbolic plane in Euclidean 3-space.

Nevertheless, there is an analytic distortion-free model of the hyperbolic plane in Lorentzian 3-space. This model is called the *hyperboloid model* of the hyperbolic plane. Lorentzian 3-space is $\mathbb{R}^3$ with a non-Euclidean geometry (described in Chapter 3). Even though the geometry of Lorentzian 3-space is non-Euclidean, it still has physical significance. Lorentzian 4-space is the model of space-time in the theory of special relativity.

The points of the hyperboloid model are the points of the positive sheet ($x > 0$) of the hyperboloid in $\mathbb{R}^3$ defined by the equation

$$x^2 - y^2 - z^2 = 1. \tag{1.2.1}$$

A line of the model is a branch of a hyperbola obtained by intersecting the model with a Euclidean plane passing through the origin. The angles in the hyperboloid model conform with the angles in Lorentzian 3-space. In Chapter 3, we shall adopt the hyperboloid model as our basic model of hyperbolic geometry because it most naturally exhibits the duality between spherical and hyperbolic geometries.

Exercise 1.2

1. Let P be a point outside a line L in the projective disk model. Show that there exists two lines L_1 and L_2 passing through P parallel to L such that every line passing through P parallel to L lies between L_1 and L_2. The two lines L_1 and L_2 are called the *parallels* to L at P. All the other lines passing through P parallel to L are called *ultraparallels* to L at P. Conclude that there are infinitely many ultraparallels to L at P.

2. Prove that any right triangle in the conformal disk model, with its right angle at the center of the model, has angle sum less than $180°$.

3. Let u, v be distinct points of the upper half-plane model. Show how to construct the hyperbolic line joining u and v with a Euclidean ruler and compass.

4. Let $\phi(z) = \frac{az+b}{cz+d}$ with a, b, c, d in $\mathbb{R}$ and $ad - bc > 0$. Prove that ϕ maps the complex upper half-plane bijectively onto itself.

5. Show that the intersection of the hyperboloid $x^2 - y^2 - z^2 = 1$ with a Euclidean plane passing through the origin is either empty or a hyperbola.

§1.3. Euclidean n-Space

The standard analytic model for n-dimensional Euclidean geometry is the n-dimensional real vector space $\mathbb{R}^n$. A *vector* in $\mathbb{R}^n$ is an ordered n-tuple $x = (x_1, \ldots, x_n)$ of real numbers. Let x and y be vectors in $\mathbb{R}^n$. The *Euclidean inner product* of x and y is defined to be the real number

$$x \cdot y = x_1y_1 + \cdots + x_ny_n. \tag{1.3.1}$$

The Euclidean inner product is the prototype for the following definition:

Definition: An *inner product* on a real vector space V is a function from $V \times V$ to $\mathbb{R}$, denoted by $(v, w) \mapsto \langle v, w \rangle$, such that for all v, w in V,

(1) $\langle v, \ \rangle$ and $\langle \ , w \rangle$ are linear functions from V to $\mathbb{R}$ (bilinearity);

(2) $\langle v, w \rangle = \langle w, v \rangle$ (symmetry); and

(3) if $v \neq 0$, then there is a $w \neq 0$ such that $\langle v, w \rangle \neq 0$ (nondegeneracy).

The Euclidean inner product on $\mathbb{R}^n$ is obviously bilinear and symmetric. Observe that if $x \neq 0$ in $\mathbb{R}^n$, then $x \cdot x > 0$, and so the Euclidean inner product is also nondegenerate.

An inner product $\langle \ , \ \rangle$ on a real vector space V is said to be *positive definite* if and only if $\langle v, v \rangle > 0$ for all nonzero v in V. The Euclidean inner product on $\mathbb{R}^n$ is an example of a positive definite inner product.

Let $\langle \ , \ \rangle$ be a positive definite inner product on V. The *norm* of v in V, with respect to $\langle \ , \ \rangle$, is defined to be the real number

$$\|v\| = \langle v, v \rangle^{\frac{1}{2}}. \tag{1.3.2}$$

The norm of x in $\mathbb{R}^n$, with respect to the Euclidean inner product, is called the *Euclidean norm* and is denoted by $|x|$.

Theorem 1.3.1. (Cauchy's inequality) *Let $\langle \ , \ \rangle$ be a positive definite inner product on a real vector space V. If v, w are vectors in V, then*

$$|\langle v, w \rangle| \leq \|v\| \ \|w\|$$

with equality if and only if v and w are linearly dependent.

Proof: If v and w are linearly dependent, then equality clearly holds. Suppose that v and w are linearly independent. Then $tv - w \neq 0$ for all t in $\mathbb{R}$, and so

$$\begin{aligned} 0 \ < \ \|tv - w\|^2 &= \langle tv - w, tv - w \rangle \\ &= t^2\|v\|^2 - 2t\langle v, w \rangle + \|w\|^2. \end{aligned}$$

The last expression is a quadratic polynomial in t with no real roots, and so its discriminant must be negative. Thus

$$4\langle v, w \rangle^2 - 4\|v\|^2\|w\|^2 < 0. \qquad \square$$

Let x, y be nonzero vectors in $\mathbb{R}^n$. By Cauchy's inequality, there is a unique real number $\theta(x, y)$ between 0 and π such that

$$x \cdot y = |x|\,|y| \cos\theta(x, y). \tag{1.3.3}$$

The *Euclidean angle* between x and y is defined to be $\theta(x, y)$.

Two vectors x, y in $\mathbb{R}^n$ are said to be *orthogonal* if and only if $x \cdot y = 0$. As $\cos(\pi/2) = 0$, two nonzero vectors x, y in $\mathbb{R}^n$ are orthogonal if and only if $\theta(x, y) = \pi/2$.

Corollary 1. (The triangle inequality) *If x and y are vectors in $\mathbb{R}^n$, then*

$$|x + y| \leq |x| + |y|$$

with equality if and only if x and y are linearly dependent.

Proof: Observe that

$$\begin{aligned} |x + y|^2 &= (x + y) \cdot (x + y) \\ &= |x|^2 + 2x \cdot y + |y|^2 \\ &\leq |x|^2 + 2|x|\,|y| + |y|^2 \\ &= (|x| + |y|)^2 \end{aligned}$$

with equality if and only if x and y are linearly dependent. □

Metric Spaces

The *Euclidean distance* between vectors x and y in $\mathbb{R}^n$ is defined to be

$$d_E(x, y) = |x - y|. \tag{1.3.4}$$

The distance function d_E is the prototype for the following definition:

Definition: A *metric* on a set X is a function $d : X \times X \to \mathbb{R}$ such that for all x, y, z in X,

(1) $d(x, y) \geq 0$ (nonnegativity);

(2) $d(x, y) = 0$ if and only if $x = y$ (nondegeneracy);

(3) $d(x, y) = d(y, x)$ (symmetry); and

(4) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

The Euclidean distance function d_E obviously satisfies the first three axioms for a metric on $\mathbb{R}^n$. By Corollary 1, we have

$$|x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z|.$$

Therefore d_E satisfies the triangle inequality. Thus d_E is a metric on $\mathbb{R}^n$, called the *Euclidean metric.*

Definition: : A *metric space* is a set X together with a metric d on X.

Example: *Euclidean n-space* E^n is the metric space consisting of $\mathbb{R}^n$ together with the Euclidean metric d_E.

An element of a metric space is called a *point.* Let X be a metric space with metric d. The *open ball* of radius $r > 0$, centered at the point a of X, is defined to be the set

$$B(a, r) = \{x \in X : d(a, x) < r\}.$$

The *closed ball* of radius $r > 0$, centered at the point a of X, is defined to be the set

$$C(a, r) = \{x \in X : d(a, x) \leq r\}.$$

A subset U of X is *open* in X if and only if for each point x of U, there is an $r > 0$ such that U contains $B(x, r)$. In particular, if S is a subset of X and $r > 0$, then the *r-neighborhood* of S in X, defined by

$$N(S, r) = \cup\{B(x, r) : x \in S\},$$

is a open in X.

The collection of all open subsets of a metric space X is a topology on X, called the *metric topology* of X. A metric space is always assumed to be topologized with its metric topology. The metric topology of E^n is called the *Euclidean topology* of $\mathbb{R}^n$. We shall assume that $\mathbb{R}^n$ is topologized with the Euclidean topology.

Isometries

A function $\phi : X \to Y$ between metric spaces *preserves distances* if and only if

$$d_Y(\phi(x), \phi(y)) = d_X(x, y) \quad \text{for all } x, y \text{ in } X.$$

Note that a distance preserving function is a continuous injection.

Definition: An *isometry* from a metric space X to a metric space Y is a distance preserving bijection $\phi : X \to Y$.

The inverse of an isometry is obviously an isometry, and the composite of two isometries is an isometry. Two metric spaces X and Y are said to be *isometric* (or *metrically equivalent*) if and only if there is an isometry $\phi : X \to Y$. Clearly, being isometric is an equivalence relation among the class of all metric spaces.

The set of isometries from a metric space X to itself, together with multiplication defined by composition, forms a group $\mathrm{I}(X)$, called the *group of isometries* of X. An isometry from E^n to itself is called a *Euclidean isometry.*

Example: Let a be a point of E^n. The function $\tau_a : E^n \to E^n$, defined by the formula

$$\tau_a(x) = a + x,$$

is called the *translation* of E^n by a. The function τ_a is an isometry, since τ_a is a bijection with inverse τ_{-a} and

$$|\tau_a(x) - \tau_a(y)| = |(a + x) - (a + y)| = |x - y|.$$

Definition: A metric space X is *homogeneous* if and only if for each pair of points x, y of X, there is an isometry ϕ of X such that $\phi(x) = y$.

Example: Euclidean n-space E^n is homogeneous, since for each pair of points x, y of E^n, the translation of E^n by $y - x$ translates x to y.

Orthogonal Transformations

Definition: A function $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is an *orthogonal transformation* if and only if

$$\phi(x) \cdot \phi(y) = x \cdot y \quad \text{for all } x, y \text{ in } \mathbb{R}^n.$$

Example: The *antipodal transformation* α of $\mathbb{R}^n$, defined by $\alpha(x) = -x$, is an orthogonal transformation, since

$$\alpha(x) \cdot \alpha(y) = -x \cdot -y = x \cdot y.$$

Definition: A basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$ is *orthonormal* if and only if

$$v_i \cdot v_j = \delta_{ij} \text{ (Kronecker's delta) for all } i, j.$$

Example: Let e_i be the vector in $\mathbb{R}^n$ whose coordinates are all zero, except for the ith, which is one. Then $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $\mathbb{R}^n$ called the *standard basis* of $\mathbb{R}^n$.

Theorem 1.3.2. *A function $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal transformation if and only if ϕ is linear and $\{\phi(e_1), \ldots, \phi(e_n)\}$ is an orthonormal basis of $\mathbb{R}^n$.*

Proof: Suppose that ϕ is an orthogonal transformation of $\mathbb{R}^n$. Then

$$\phi(e_i) \cdot \phi(e_j) = e_i \cdot e_j = \delta_{ij}.$$

To see that $\phi(e_1), \ldots, \phi(e_n)$ are linearly independent, suppose that

$$\sum_{i=1}^{n} c_i \phi(e_i) = 0.$$

Upon taking the inner product of this equation with $\phi(e_j)$, we find that $c_j = 0$ for each j. Hence $\{\phi(e_1), \ldots, \phi(e_n)\}$ is an orthonormal basis of $\mathbb{R}^n$.

Let x be in $\mathbb{R}^n$. Then there are coefficients $c_1, \ldots, c_n$ in $\mathbb{R}$ such that

$$\phi(x) = \sum_{i=1}^{n} c_i \phi(e_i).$$

As $\{\phi(e_1), \ldots, \phi(e_n)\}$ is an orthonormal basis, we have

$$c_j = \phi(x) \cdot \phi(e_j) = x \cdot e_j = x_j.$$

Then ϕ is linear, since

$$\phi\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i \phi(e_i).$$

Conversely, suppose that ϕ is linear and $\{\phi(e_1), \ldots, \phi(e_n)\}$ is an orthonormal basis of $\mathbb{R}^n$. Then ϕ is orthogonal, since

$$\begin{aligned}
\phi(x) \cdot \phi(y) &= \phi\left(\sum_{i=1}^{n} x_i e_i\right) \cdot \phi\left(\sum_{j=1}^{n} y_j e_j\right) \\
&= \left(\sum_{i=1}^{n} x_i \phi(e_i)\right) \cdot \left(\sum_{j=1}^{n} y_j \phi(e_j)\right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \phi(e_i) \cdot \phi(e_j) \\
&= \sum_{i=1}^{n} x_i y_i \quad = \quad x \cdot y.
\end{aligned}$$

□

Corollary 2. *Every orthogonal transformation is a Euclidean isometry.*

Proof: Let $\phi : \mathbb{R}^n \to \mathbb{R}^n$ be an orthogonal transformation. Then ϕ preserves Euclidean norms, since

$$|\phi(x)|^2 = \phi(x) \cdot \phi(x) = x \cdot x = |x|^2.$$

Consequently ϕ preserves distances, since

$$|\phi(x) - \phi(y)| = |\phi(x - y)| = |x - y|.$$

By Theorem 1.3.2, the map ϕ is bijective. Therefore ϕ is a Euclidean isometry.

□

A real $n \times n$ matrix A is said to be *orthogonal* if and only if the associated linear transformation $A : \mathbb{R}^n \to \mathbb{R}^n$, defined by $A(x) = Ax$, is orthogonal. The set of all orthogonal $n \times n$ matrices together with matrix multiplication forms a group $\mathrm{O}(n)$, called the *orthogonal group* of $n \times n$ matrices. By Theorem 1.3.2, the group $\mathrm{O}(n)$ is naturally isomorphic to the group of orthogonal transformations of $\mathbb{R}^n$.

The next theorem follows immediately from Theorem 1.3.2.

Theorem 1.3.3. *Let A be a real $n \times n$ matrix. Then the following are equivalent:*

(1) *The matrix A is orthogonal.*

(2) *The columns of A form an orthonormal basis of $\mathbb{R}^n$.*

(3) *The matrix A satisfies the equation $A^tA = I$.*

(4) *The matrix A satisfies the equation $AA^t = I$.*

(5) *The rows of A form an orthonormal basis of $\mathbb{R}^n$.*

Let A be an orthogonal matrix. As $A^tA = I$, we have that $(\det A)^2 = 1$. Thus $\det A = \pm 1$. If $\det A = 1$, then A is called a *rotation*. Let $\mathrm{SO}(n)$ be the set of all rotations in $\mathrm{O}(n)$. Then $\mathrm{SO}(n)$ is a subgroup of index two in $\mathrm{O}(n)$. The group $\mathrm{SO}(n)$ is called the *special orthogonal group* of $n \times n$ matrices.

Group Actions

Definition: A group G *acts* on a set X if and only if there is a function from $G \times X$ to X, written $(g, x) \mapsto gx$, such that for all g, h in G and x in X, we have

(1) $1 \cdot x = x$ and

(2) $g(hx) = (gh)x$.

A function from $G \times X$ to X satisfying conditions (1) and (2) is called an *action* of G on X.

Example: If X is a metric space, then the group $\mathrm{I}(X)$ of isometries of X acts on X by $\phi x = \phi(x)$.

Definition: An action of a group G on a set X is *transitive* if and only if for each x, y in X, there is a g in G such that $gx = y$.

Theorem 1.3.4. *For each dimension m, the natural action of $\mathrm{O}(n)$ on the set of m-dimensional vector subspaces of $\mathbb{R}^n$ is transitive.*

Proof: Let V be an m-dimensional vector subspace of $\mathbb{R}^n$ with $m > 0$. Identify $\mathbb{R}^m$ with the subspace of $\mathbb{R}^n$ spanned by the vectors $e_1, \ldots, e_m$. It suffices to show that there is an A in $\mathrm{O}(n)$ such that $A(\mathbb{R}^m) = V$.

Choose a basis $\{u_1, \ldots, u_n\}$ of $\mathbb{R}^n$ such that $\{u_1, \ldots, u_m\}$ is a basis of V. We now perform the Gram-Schmidt process on $\{u_1, \ldots, u_n\}$. Let $w_1 = u_1/|u_1|$. Then $|w_1| = 1$. Next, let $v_2 = u_2 - (u_2 \cdot w_1)w_1$. Then v_2 is nonzero, since u_1 and u_2 are linearly independent; moreover,

$$w_1 \cdot v_2 = w_1 \cdot u_2 - (u_2 \cdot w_1)(w_1 \cdot w_1) = 0.$$

Now let

$$\begin{aligned}
w_2 &= v_2/|v_2|, \\
v_3 &= u_3 - (u_3 \cdot w_1)w_1 - (u_3 \cdot w_2)w_2, \\
w_3 &= v_3/|v_3|, \\
&\vdots \\
v_n &= u_n - (u_n \cdot w_1)w_1 - (u_n \cdot w_2)w_2 - \cdots - (u_n \cdot w_{n-1})w_{n-1}, \\
w_n &= v_n/|v_n|.
\end{aligned}$$

Then $\{w_1, \ldots, w_n\}$ is an orthonormal basis of $\mathbb{R}^n$ with $\{w_1, \ldots, w_m\}$ a basis of V. Let A be the $n \times n$ matrix whose columns are $w_1, \ldots, w_n$. Then A is orthogonal by Theorem 1.3.3, and $A(\mathbb{R}^m) = V$. □

Definition: Two subsets S and T of a metric space X are *congruent* in X if and only if there is a isometry ϕ of X such that $\phi(S) = T$.

Being congruent is obviously an equivalence relation on the set of all subsets of X. An isometry of a metric space X is also called a *congruence transformation* of X.

Definition: An *m-plane* of E^n is a coset $a+V$ of an m-dimensional vector subspace V of $\mathbb{R}^n$.

Corollary 3. *All the m-planes of E^n are congruent.*

Proof: Let $a+V$ and $b+W$ be m-planes of E^n. By Theorem 1.3.4, there is a matrix A in $\mathrm{O}(n)$ such that $A(V) = W$. Define $\phi : E^n \to E^n$ by

$$\phi(x) = (b - Aa) + Ax.$$

Then ϕ is an isometry and

$$\phi(a + V) = b + W.$$

Thus $a + V$ and $b + W$ are congruent. □

Characterization of Euclidean Isometries

The following theorem characterizes an isometry of E^n.

Theorem 1.3.5. *Let $\phi : E^n \to E^n$ be a function. Then the following are equivalent:*

(1) *The function ϕ is an isometry.*

(2) *The function ϕ preserves distances.*

(3) *The function ϕ is of the form $\phi(x) = a+Ax$, where A is an orthogonal matrix and $a = \phi(0)$.*

Proof: By definition, (1) implies (2). Suppose that ϕ preserves distances. Then $A = \phi - \phi(0)$ also preserves distances and $A(0) = 0$. Therefore A preserves Euclidean norms, since

$$|Ax| = |A(x) - A(0)| = |x - 0| = |x|.$$

Consequently A is orthogonal, since

$$\begin{aligned} 2Ax \cdot Ay &= |Ax|^2 + |Ay|^2 - |Ax - Ay|^2 \\ &= |x|^2 + |y|^2 - |x - y|^2 \quad = \quad 2x \cdot y. \end{aligned}$$

Thus, there is an orthogonal $n \times n$ matrix A such that $\phi(x) = \phi(0) + Ax$, and so (2) implies (3). If ϕ is in the form given in (3), then ϕ is the composite of an orthogonal transformation followed by a translation, and so ϕ is an isometry. Thus (3) implies (1). □

Remark: Theorem 1.3.5 states that every isometry of E^n is the composite of an orthogonal transformation followed by a translation. It is worth noting that such a decomposition is unique.

Similarities

A function $\phi : X \to Y$ between metric spaces is a *change of scale* if and only if there is a real number $k > 0$ such that

$$d_Y(\phi(x), \phi(y)) = k\, d_X(x, y) \quad \text{for all } x, y \text{ in } X.$$

The positive constant k is called the *scale factor* of ϕ. Note that a change of scale is a continuous injection.

Definition: A *similarity* from a metric space X to a metric space Y is a bijective change of scale $\phi : X \to Y$.

The inverse of a similarity, with scale factor k, is a similarity with scale factor $1/k$. Therefore, a similarity is also a homeomorphism. Two metric spaces X and Y are said to be *similar* (or *equivalent under a change of scale*) if and only if there is a similarity $\phi : X \to Y$. Clearly, being similar is an equivalence relation among the class of all metric spaces. The set of similarities from a metric space X to itself, together with multiplication defined by composition, forms a group $\mathrm{S}(X)$, called the *group of similarities* of X. The group of similarities $\mathrm{S}(X)$ contains the group of isometries $\mathrm{I}(X)$ as a subgroup. A similarity from E^n to itself is called a *Euclidean similarity.*

Example: Let $k > 1$. The function $\mu_k : E^n \to E^n$, defined by $\mu_k(x) = kx$, is called the *magnification* of E^n by the factor k. Clearly, the magnification μ_k is a similarity with scale factor k.

The next theorem follows easily from Theorem 1.3.5.

Theorem 1.3.6. *Let $\phi : E^n \to E^n$ be a function. Then the following are equivalent:*

(1) *The function ϕ is a similarity.*

(2) *The function ϕ is a change of scale.*

(3) *The function ϕ is of the form $\phi(x) = a + kAx$, where A is an orthogonal matrix, k is a positive constant, and $a = \phi(0)$.*

Given a geometry on a space X, its *principal group* is the group of all transformations of X under which all the theorems of the geometry remain true. In his famous *Erlanger Program*, Klein proposed that the study of a geometry should be viewed as the study of the invariants of its principal group. The principal group of n-dimensional Euclidean geometry is the group $\mathrm{S}(E^n)$ of similarities of E^n.

Exercise 1.3

1. Let $v_0, \ldots, v_m$ be vectors in $\mathbb{R}^n$ such that $v_1 - v_0, \ldots, v_m - v_0$ are linearly independent. Show that there is a unique m-plane of E^n containing $v_0, \ldots, v_m$. Conclude that there is a unique 1-plane of E^n containing any two distinct points of E^n.

2. A *line* of E^n is defined to be a 1-plane of E^n. Let x, y be distinct points of E^n. Show that the unique line of E^n containing x and y is the set
$$\{x + t(y - x) : t \in \mathbb{R}\}.$$
The *line segment* in E^n joining x to y is defined to be the set
$$\{x + t(y - x) : 0 \leq t \leq 1\}.$$
Conclude that every line segment in E^n extends to a unique line of E^n.

3. Two m-planes of E^n are said to be *parallel* if and only if they are cosets of the same m-dimensional vector subspace of $\mathbb{R}^n$. Let x be a point of E^n outside of an m-plane P of E^n. Show that there is a unique m-plane of E^n containing x parallel to P.

4. Two m-planes of E^n are said to be *coplanar* if and only if there is an $(m+1)$-plane of E^n containing both m-planes. Show that two distinct m-planes of E^n are parallel if and only if they are coplanar and disjoint.

5. A *hyperplane* of E^n is defined to be an $(n-1)$-plane of E^n. Let x_0 be a point of a subset P of E^n. Prove that P is a hyperplane of E^n if and only if there is a unit vector a in $\mathbb{R}^n$, which is unique up to sign, such that
$$P = \{x \in E^n : a \cdot (x - x_0) = 0\}.$$

6. The *orthogonal complement* of an m-dimensional vector subspace V of $\mathbb{R}^n$ is defined to be the set
$$V^\perp = \{x \in \mathbb{R}^n : x \cdot y = 0 \quad \text{for all } y \text{ in } V\}.$$

Prove that $V^\perp$ is an $(n-m)$-dimensional vector subspace of $\mathbb{R}^n$ and that and each vector x in $\mathbb{R}^n$ can be written uniquely as $x = y + z$ with y in V and z in $V^\perp$. In other words, $\mathbb{R}^n = V \oplus V^\perp$.

7. A line and a hyperplane of E^n are said to be *orthogonal* if and only if their associated vector spaces are orthogonal complements. Let y be a point of E^n outside of a hyperplane P of E^n. Show that there is a unique point x_0 in P nearest to y and that the line passing through x_0 and y is the unique line of E^n passing through y orthogonal to P.

8. Let $u_0, \ldots, u_n$ be vectors in $\mathbb{R}^n$ such that $u_1 - u_0, \ldots, u_n - u_0$ are linearly independent, let $v_0, \ldots, v_n$ be vectors in $\mathbb{R}^n$ such that $v_1 - v_0, \ldots, v_n - v_0$ are linearly independent, and suppose that
$$|u_i - u_j| = |v_i - v_j| \quad \text{for all } i, j.$$
Show that there is a unique isometry ϕ of E^n such that $\phi(u_i) = v_i$ for each $i = 1, \ldots, n$.

9. Prove that E^m and E^n are isometric if and only if $m = n$.

10. Let $\|\ \ \|$ be the norm of a positive definite inner product $\langle\ ,\ \rangle$ on an n-dimensional real vector space V. Define a metric d on V by the formula $d(v, w) = \|v - w\|$. Show that d is a metric on V and prove that the metric space (V, d) is isometric to E^n.

§1.4. Geodesics

In this section, we study the metrical properties of lines of Euclidean n-space E^n. In order to prepare for later applications, all the basic definitions in this section are in the general context of curves in a metric space X.

Definition: A *curve* in a space X is a continuous function $\gamma : [a, b] \to X$ where $[a, b]$ is a closed interval in $\mathbb{R}$ with $a < b$.

Let $\gamma : [a, b] \to X$ be a curve. Then $\gamma(a)$ is called the *initial point* of γ and $\gamma(b)$ is called the *terminal point.* We say that γ is a curve in X from $\gamma(a)$ to $\gamma(b)$. If $X = E^n$, then γ is said to be *linear* if and only if
$$\gamma(a + t(b - a)) = \gamma(a) + t(\gamma(b) - \gamma(a))$$
for all t in $[0, 1]$.

Example: Let x, y be points of E^n. Define $\gamma : [0, 1] \to E^n$ by
$$\gamma(t) = x + t(y - x).$$
Then γ is a linear curve in E^n from x to y.

The proof of the next theorem is straightforward and is left to the reader.

Theorem 1.4.1. *Let $\gamma : [a,b] \to E^n$ be a curve. Then the following are equivalent:*

(1) *The curve γ is linear.*

(2) *The curve γ satisfies the equation*
$$\gamma(t) = \gamma(a) + \left(\frac{t-a}{b-a}\right)(\gamma(b) - \gamma(a)).$$

(3) *The curve γ has a constant first derivative $\gamma' : [a,b] \to E^n$.*

Definition: Three points x, y, z of E^n are *collinear*, with y between x and z, if and only if there is a real number t between 0 and 1 such that $y = x + t(z - x)$.

The proof of the next lemma is elementary and is left to the reader.

Lemma 1. *Three points x, y, z of E^n are collinear, with y between x and z, if and only if*
$$|z - x| = |y - x| + |z - y|.$$

Geodesic Arcs

Definition: A *geodesic arc* in a metric space X is a distance preserving function $\alpha : [a,b] \to X$, with $a < b$ in $\mathbb{R}$.

Note that a geodesic arc $\alpha : [a,b] \to X$ is a continuous injection and so is a curve.

Theorem 1.4.2. *A curve $\alpha : [a,b] \to E^n$ is a geodesic arc if and only if α is linear and $|\alpha'(t)| = 1$ for all t in $[a,b]$.*

Proof: Suppose that α is linear and $|\alpha'(t)| = 1$. Then by Theorem 1.4.1,
$$\alpha(t) = \alpha(a) + \left(\frac{t-a}{b-a}\right)(\alpha(b) - \alpha(a)),$$
and since $|\alpha'(t)| = 1$, we have
$$|\alpha(b) - \alpha(a)| = b - a.$$
Therefore
$$|\alpha(t) - \alpha(s)| = \frac{|t-s|}{b-a}|\alpha(b) - \alpha(a)| = |t - s|.$$
Thus α is a geodesic arc.

Conversely, suppose that α is a geodesic arc. Let t be in $[a,b]$. Then
$$\begin{aligned}|\alpha(b) - \alpha(a)| &= b - a \\ &= b - t + t - a \\ &= |\alpha(b) - \alpha(t)| + |\alpha(t) - \alpha(a)|.\end{aligned}$$

By Lemma 1, we have that $\alpha(a), \alpha(t), \alpha(b)$ are collinear with $\alpha(t)$ between $\alpha(a)$ and $\alpha(b)$. Therefore, there is some $f(t)$ in [0,1] such that

$$\alpha(t) = \alpha(a) + f(t)(\alpha(b) - \alpha(a)).$$

Now, since

$$f(t) = \frac{|\alpha(t) - \alpha(a)|}{|\alpha(b) - \alpha(a)|} = \frac{t-a}{b-a},$$

the curve α is linear by Theorem 1.4.1 and

$$|\alpha'(t)| = \frac{|\alpha(b) - \alpha(a)|}{b-a} = 1. \qquad \square$$

Definition: A *geodesic segment* joining a point x to a point y in a metric space X is the image of a geodesic arc $\alpha : [a, b] \to X$ whose initial point is x and terminal point is y.

Corollary 1. *The geodesic segments of E^n are its line segments.*

Theorem 1.4.3. *Let $[x, y]$ and $[y, z]$ be geodesic segments joining x to y and y to z, respectively, in a metric space X. Then the set $[x, y] \cup [y, z]$ is a geodesic segment joining x to z in X if and only if*

$$d(x, z) = d(x, y) + d(y, z).$$

Proof: If $[x, y] \cup [y, z]$ is a geodesic segment joining x to z, then obviously

$$d(x, z) = d(x, y) + d(y, z).$$

Conversely, suppose that the above equation holds. Let $\alpha : [a, b] \to X$ and $\beta : [b, c] \to X$ be geodesic arcs from x to y and y to z, respectively. Define $\gamma : [a, c] \to X$ by $\gamma(t) = \alpha(t)$ if $a \leq t \leq b$ and $\gamma(t) = \beta(t)$ if $b \leq t \leq c$. Suppose that $a \leq s < t \leq c$. If $t \leq b$, then

$$d(\gamma(s), \gamma(t)) = d(\alpha(s), \alpha(t)) = t - s.$$

If $b \leq s$, then

$$d(\gamma(s), \gamma(t)) = d(\beta(s), \beta(t)) = t - s.$$

If $s < b < t$, then

$$\begin{aligned} d(\gamma(s), \gamma(t)) &\leq d(\gamma(s), \gamma(b)) + d(\gamma(b), \gamma(t)) \\ &= (b - s) + (t - b) \quad = \quad t - s. \end{aligned}$$

Moreover

$$\begin{aligned} d(\gamma(s), \gamma(t)) &\geq d(\gamma(a), \gamma(c)) - d(\gamma(a), \gamma(s)) - d(\gamma(t), \gamma(c)) \\ &= d(x, z) - (s - a) - (c - t) \\ &= d(x, y) + d(y, z) - (c - a) + (t - s) \\ &= (b - a) + (c - b) - (c - a) + (t - s) \quad = \quad t - s. \end{aligned}$$

Therefore, we have

$$d(\gamma(s), \gamma(t)) = t - s.$$

Hence γ is a geodesic arc from x to z whose image is the set $[x, y] \cup [y, z]$. Thus $[x, y] \cup [y, z]$ is a geodesic segment joining x to y. □

A subset C of E^n is said to be *convex* if and only if for each pair of distinct points x, y in C, the line segment joining x to y is contained in C. The notion of convexity in E^n is the prototype for the following definition:

Definition: A metric space X is *geodesically convex* if and only if for each pair of distinct points x, y of X, there is a unique geodesic segment in X joining x to y.

Example: Euclidean n-space E^n is geodesically convex.

Remark: The modern interpretation of Euclid's first axiom is that a Euclidean plane is geodesically convex.

Definition: A metric space X is *geodesically connected* if and only if each pair of distinct points of X are joined by a geodesic segment in X.

A geodesically convex metric space is geodesically connected, but a geodesically connected metric space is not necessarily geodesically convex.

Definition: A *geodesic curve* in a metric space X is a locally distance preserving curve $\gamma : [a, b] \to X$.

A geodesic arc is a geodesic curve, but a geodesic curve is not necessarily a geodesic arc.

Definition: A *geodesic section* in a metric space X is the image of an injective geodesic curve $\gamma : [a, b] \to X$.

A geodesic segment is a geodesic section, but a geodesic section is not necessarily a geodesic segment.

Definition: A *geodesic half-line* in a metric space X is a locally distance preserving function $\eta : [0, +\infty) \to X$.

A geodesic half-line is continuous, since it is locally continuous.

Definition: A *geodesic ray* in a metric space X is the image of a geodesic half-line $\eta : [0, +\infty) \to X$.

Geodesic Lines

Definition: A *geodesic line* in a metric space X is a locally distance preserving function $\lambda : \mathbb{R} \to X$.

A geodesic line is continuous, since it is locally continuous.

Theorem 1.4.4. *A function $\lambda : \mathbb{R} \to E^n$ is a geodesic line if and only if $\lambda(t) = \lambda(0) + t(\lambda(1) - \lambda(0))$ for all t and $|\lambda(1) - \lambda(0)| = 1$.*

Proof: Suppose that $\lambda(t) = \lambda(0) + t(\lambda(1) - \lambda(0))$ and $|\lambda(1) - \lambda(0)| = 1$. Then $\lambda'(t)$ is constant and of norm one. Hence, the restriction of λ to any interval is a geodesic arc by Theorems 1.4.1 and 1.4.2. Thus λ is a geodesic line.

Conversely, suppose that λ is a geodesic line. By Theorems 1.4.1 and 1.4.2, the function λ is differentiable and λ' is a constant unit vector. Hence

$$\lambda(t) = \lambda(0) + t(\lambda(1) - \lambda(0))$$

for all t and $|\lambda(1) - \lambda(0)| = 1$. □

Definition: A *geodesic* in a metric space X is the image of a geodesic line $\lambda : \mathbb{R} \to X$.

Corollary 2. *The geodesics of E^n are its lines.*

Definition: A metric space X is *geodesically complete* if and only if each geodesic arc $\alpha : [a, b] \to X$ extends to a unique geodesic line $\lambda : \mathbb{R} \to X$.

Example: Euclidean n-space E^n is geodesically complete.

Remark: The modern interpretation of Euclid's second axiom is that a Euclidean plane is geodesically complete.

Definition: A metric space X is *totally geodesic* if and only if for each pair of distinct points x, y of X there is a geodesic of X containing both x and y.

Example: Euclidean n-space E^n is totally geodesic.

Definition: A *coordinate frame* of E^n is an n-tuple $(\lambda_1, \ldots, \lambda_n)$ of functions such that

(1) the function $\lambda_i : \mathbb{R} \to E^n$ is a geodesic line for each $i = 1, \ldots, n$;

(2) there is a point a of E^n such that $\lambda_i(0) = a$ for all i; and

(3) the set $\{\lambda_1'(0), \ldots, \lambda_n'(0)\}$ is an orthonormal basis of $\mathbb{R}^n$.

Example: Define $\varepsilon_i : \mathbb{R} \to E^n$ by $\varepsilon_i(t) = te_i$. Then $(\varepsilon_1, \ldots, \varepsilon_n)$ is a coordinate frame of E^n, called the the *standard coordinate frame* of E^n.

Theorem 1.4.5. *The action of* $\mathrm{I}(E^n)$ *on the set of coordinate frames of* E^n, *given by* $\phi(\lambda_1, \ldots, \lambda_n) = (\phi\lambda_1, \ldots, \phi\lambda_n)$, *is transitive.*

Proof: Let $(\lambda_1, \ldots, \lambda_n)$ be a coordinate frame of E^n. It suffices to show that there is a ϕ in $\mathrm{I}(E^n)$ such that $\phi(\varepsilon_1, \ldots, \varepsilon_n) = (\lambda_1, \ldots, \lambda_n)$. Let A be the $n \times n$ matrix whose columns are $\lambda_1'(0), \ldots, \lambda_n'(0)$. Then A is orthogonal by Theorem 1.3.3. Let $a = \lambda_i(0)$ and define $\phi : E^n \to E^n$ by $\phi(x) = a + Ax$. Then ϕ is an isometry. As $\phi\varepsilon_i(0) = \lambda_i(0)$ and $(\phi\varepsilon_i)'(0) = \lambda_i'(0)$, we have that $\phi(\varepsilon_1, \ldots, \varepsilon_n) = (\lambda_1, \ldots, \lambda_n)$. □

Remark: The modern interpretation of Euclid's fourth axiom is that the group of isometries of a Euclidean plane acts transitively on the set of all its coordinate frames.

Exercise 1.4

1. Prove Theorem 1.4.1.
2. Prove Lemma 1.
3. A subset X of E^n is said to be *affine* if and only if X is a totally geodesic metric subspace of E^n. Prove that an arbitrary intersection of affine subsets of E^n is affine.
4. An *affine combination* of points $v_1, \ldots, v_m$ of E^n is a linear combination of the form $t_1v_1 + \cdots + t_mv_m$ such that $t_1 + \cdots + t_m = 1$. Prove that a subset X of E^n is affine if and only if X contains every affine combination of points of X.
5. The *affine hull* of a subset S of E^n is defined to be the intersection $A(S)$ of all the affine subsets of E^n containing S. Prove that $A(S)$ is the set of all affine combinations of points of S.
6. A set $\{v_0, \ldots, v_m\}$ of points of E^n is said to be *affinely independent* if and only if $t_0v_0 + \cdots + t_mv_m = 0$ and $t_0 + \cdots + t_m = 0$ imply that $t_i = 0$ for all $i = 0, \ldots, m$. Prove that $\{v_0, \ldots, v_m\}$ is affinely independent if and only if the vectors $v_1 - v_0, \ldots, v_m - v_0$ are linearly independent.
7. An *affine basis* of an affine subset X of E^n is an affinely independent set of points $\{v_0, \ldots, v_m\}$ such that X is the affine hull of $\{v_0, \ldots, v_m\}$. Prove that every nonempty affine subset of E^n has an affine basis.
8. Prove that a nonempty subset X of E^n is affine if and only if X is an m-plane of E^n for some m.
9. A function $\phi : E^n \to E^n$ is said to be *affine* if and only if
$$\phi((1-t)x + ty) = (1-t)\phi(x) + t\phi(y)$$
for all x, y in E^n and t in $\mathbb{R}$. Show that an affine transformation of E^n maps affine sets to affine sets and convex sets to convex sets.

10. Prove that a function $\phi : E^n \to E^n$ is affine if and only if there is an $n \times n$ matrix A and a point a of E^n such that $\phi(x) = a + Ax$ for all x in E^n.

11. Prove that an arbitrary intersection of convex subsets of E^n is convex.

12. A *convex combination* of points $v_1, \ldots, v_m$ of E^n is a linear combination of the form $t_1v_1 + \cdots + t_mv_m$ such that $t_1 + \cdots + t_m = 1$ and $t_i \geq 0$ for all $i = 1, \ldots, m$. Prove that a subset C of E^n is convex if and only if C contains every convex combination of points of C.

13. The *convex hull* of a subset S of E^n is defined to be the intersection $C(S)$ of all the convex subsets of E^n containing S. Prove that $C(S)$ is the set of all convex combinations of points of S.

14. Let S be a subset of E^n. Prove that every element of $C(S)$ is a convex combination of at most $n+1$ points of S.

15. Let K be a compact subset of E^n. Prove that $C(K)$ is compact.

16. Let C be a convex subset of E^n. Prove that for all $r > 0$, the r-neighborhood $N(C, r)$ of C in E^n is convex.

17. A subset of S of E^n is *locally convex* if and only if for each x in S, there is an $r > 0$ so that $B(x, r) \cap S$ is convex. Prove that a closed, connected, locally convex subset of E^n is convex.

18. Prove that a geodesic section in a metric space X can be subdivided into a finite number of geodesic segments.

§1.5. Arc Length

Let a and b be real numbers such that $a < b$. A *partition* P of the closed interval $[a, b]$ is a finite sequence $\{t_0, \ldots, t_m\}$ of real numbers such that

$$a = t_0 < t_1 < \cdots < t_m = b.$$

The *norm* of the partition P is defined to be the real number

$$|P| = \max\{t_i - t_{i-1} : i = 1, \ldots, m\}.$$

Let $\mathcal{P}[a, b]$ be the set of all partitions of $[a, b]$. If P, Q are in $\mathcal{P}[a, b]$, then Q is said to *refine* P if and only if each term of P is a term of Q. Define a partial ordering of $\mathcal{P}[a, b]$ by $Q \leq P$ if and only if Q refines P.

Let $\gamma : [a, b] \to X$ be a curve in a metric space X and let

$$P = \{t_0, \ldots, t_m\}$$

be a partition of $[a, b]$. The *P-inscribed length* of γ is defined to be

$$\ell(\gamma, P) = \sum_{i=1}^{m} d(\gamma(t_{i-1}), \gamma(t_i)).$$

It follows from the triangle inequality that if $Q \leq P$, then $\ell(\gamma, P) \leq \ell(\gamma, Q)$.

Definition: The *length* of a curve $\gamma : [a,b] \to X$ is

$$|\gamma| = \sup\{\ell(\gamma, P) : P \in \mathcal{P}[a,b]\}.$$

Note that since $\{a,b\}$ is a partition of $[a,b]$, we have

$$d(\gamma(a), \gamma(b)) \leq |\gamma| \leq \infty.$$

Definition: A curve γ is *rectifiable* if and only if $|\gamma| < \infty$.

Example: Let $\gamma : [a,b] \to X$ be a geodesic arc and let P be a partition of $[a,b]$. Then

$$\begin{aligned} \ell(\gamma, P) &= \sum_{i=1}^{m} d(\gamma(t_{i-1}), \gamma(t_i)) \\ &= \sum_{i=1}^{m} (t_i - t_{i-1}) \quad = \quad b - a. \end{aligned}$$

Therefore γ is rectifiable and

$$|\gamma| = d(\gamma(a), \gamma(b)).$$

Theorem 1.5.1. *Let $\gamma : [a,c] \to X$ be a curve, let b be a number between a and c, and let $\alpha : [a,b] \to X$ and $\beta : [b,c] \to X$ be the restrictions of γ. Then we have*

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

Moreover γ is rectifiable if and only if α and β are rectifiable.

Proof: Let P be a partition of $[a,b]$ and let Q be a partition of $[b,c]$. Then $P \cup Q$ is a partition of $[a,c]$ and

$$\ell(\alpha, P) + \ell(\beta, Q) = \ell(\gamma, P \cup Q).$$

Therefore, we have

$$|\alpha| + |\beta| \leq |\gamma|.$$

Let R be a partition of $[a,c]$. Then $R' = R \cup \{b\}$ is a partition of $[a,c]$ and $R' = P \cup Q$, where P is a partition of $[a,b]$ and Q is a partition of $[b,c]$. Now

$$\ell(\gamma, R) \leq \ell(\gamma, R') = \ell(\alpha, P) + \ell(\beta, Q).$$

Therefore, we have

$$|\gamma| \leq |\alpha| + |\beta|.$$

Thus, we have

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

Moreover γ is rectifiable if and only if α and β are rectifiable. □

Let X be a geodesically connected metric space and let $\gamma : [a, b] \to X$ be a curve from x to y. Then $|\gamma| \geq d(x, y)$ with equality if γ is a geodesic arc. Thus $d(x, y)$ is the shortest possible length of γ. It is an exercise to show that $|\gamma| = d(x, y)$ if and only if γ maps $[a, b]$ onto a geodesic segment joining x to y and $d(x, \gamma(t))$ is an increasing function of t. Thus, a shortest path from x to y is along a geodesic segment joining x to y.

Let $\{t_0, \ldots, t_m\}$ be a partition of $[a, b]$ and let $\gamma_i : [t_{i-1}, t_i] \to X$, for $i = 1, \ldots, m$, be a sequence of curves such that the terminal point of γ_{i-1} is the initial point of γ_i. The *product* of $\gamma_1, \ldots, \gamma_m$ is the curve

$$\gamma_1 \cdots \gamma_m : [a, b] \to X$$

defined by

$$\gamma_1 \cdots \gamma_m(t) = \gamma_i(t) \quad \text{for } t_{i-1} \leq t \leq t_i.$$

If each γ_i is a geodesic arc, then $\gamma_1 \cdots \gamma_m$ is called a *piecewise geodesic curve*. By Theorem 1.5.1, a piecewise geodesic curve $\gamma_1 \cdots \gamma_m$ is rectifiable and

$$|\gamma_1 \cdots \gamma_m| = |\gamma_1| + \cdots + |\gamma_m|.$$

Let $\gamma : [a, b] \to X$ be a curve in a geodesically connected metric space X and let

$$P = \{t_0, \ldots, t_m\}$$

be a partition of $[a, b]$. Then there is a piecewise geodesic curve

$$\gamma_1 \cdots \gamma_m : [0, \ell] \to X$$

such that γ_i is a geodesic arc from $\gamma(t_{i-1})$ to $\gamma(t_i)$. The piecewise geodesic curve $\gamma_1 \cdots \gamma_m$ is said to be *inscribed* on γ. See Figure 1.5.1. Notice that

$$\ell(\gamma, P) = |\gamma_1 \cdots \gamma_m|.$$

Thus, the length of γ is the supremum of the lengths of all the piecewise geodesic curves inscribed on γ.

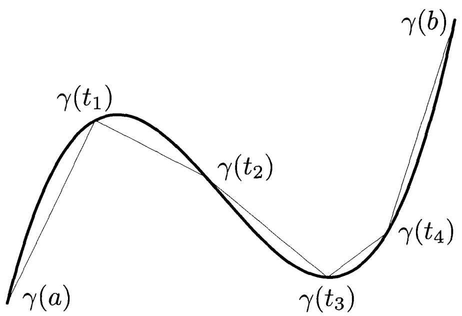


Figure 1.5.1. A piecewise geodesic curve inscribed on a curve γ

Euclidean Arc Length

A C^1 *curve* in E^n is defined to be a differentiable curve $\gamma : [a,b] \to E^n$ with a continuous derivative $\gamma' : [a,b] \to E^n$. Here $\gamma'(a)$ is the right-hand derivative of γ at a, and $\gamma'(b)$ is the left-hand derivative of γ at b.

Theorem 1.5.2. *If $\gamma : [a,b] \to E^n$ is a C^1 curve, then γ is rectifiable and the length of γ is given by the formula*

$$|\gamma| = \int_a^b |\gamma'(t)|dt.$$

Proof: Let $P = \{t_0, \ldots, t_m\}$ be a partition of $[a,b]$. Then we have

$$\begin{aligned} \ell(\gamma, P) &= \sum_{i=1}^m |\gamma(t_i) - \gamma(t_{i-1})| \\ &= \sum_{i=1}^m \left| \int_{t_{i-1}}^{t_i} \gamma'(t)dt \right| \\ &\leq \sum_{i=1}^m \int_{t_{i-1}}^{t_i} |\gamma'(t)|dt \quad = \quad \int_a^b |\gamma'(t)|dt. \end{aligned}$$

Therefore γ is rectifiable and

$$|\gamma| \leq \int_a^b |\gamma'(t)|dt.$$

If $a \leq c < d \leq b$, let $\gamma_{c,d}$ be the restriction of γ to the interval $[c,d]$. Define functions $\lambda, \mu : [a,b] \to \mathbb{R}$ by $\lambda(a) = 0$, $\lambda(t) = |\gamma_{a,t}|$ if $t > a$, and

$$\mu(t) = \int_a^t |\gamma'(t)|dt.$$

Then $\mu'(t) = |\gamma'(t)|$ by the fundamental theorem of calculus.

Suppose that $a \leq t < t + h \leq b$. Then by Theorem 1.5.1, we have

$$|\gamma(t+h) - \gamma(t)| \leq |\gamma_{t,t+h}| = \lambda(t+h) - \lambda(t).$$

Hence, by the first part of the proof applied to $\gamma_{t,t+h}$, we have

$$\left| \frac{\gamma(t+h) - \gamma(t)}{h} \right| \leq \frac{\lambda(t+h) - \lambda(t)}{h} \leq \frac{1}{h} \int_t^{t+h} |\gamma'(t)|dt = \frac{\mu(t+h) - \mu(t)}{h}.$$

Likewise, these inequalities also hold for $a \leq t + h < t \leq b$. Letting $h \to 0$, we conclude that

$$|\gamma'(t)| = \lambda'(t) = \mu'(t).$$

Therefore, we have

$$|\gamma| = \lambda(b) = \mu(b) = \int_a^b |\gamma'(t)|dt.$$

□

Let $\gamma : [a,b] \to E^n$ be a curve. Set

$$dx = (dx_1, \ldots, dx_n)$$

and

$$|dx| = (dx_1^2 + \cdots + dx_n^2)^{\frac{1}{2}}.$$

Then by definition, we have

$$\int_\gamma |dx| = |\gamma|.$$

Moreover, if γ is a C^1 curve, then by Theorem 1.5.2, we have

$$\int_\gamma |dx| = \int_a^b |\gamma'(t)|dt.$$

The differential $|dx|$ is called the *element of Euclidean arc length* of E^n.

Exercise 1.5

1. Let $\gamma : [a,b] \to X$ be a curve in a metric space X and let P, Q be partitions of $[a,b]$ such that Q refines P. Show that $\ell(\gamma, P) \leq \ell(\gamma, Q)$.

2. Let $\gamma : [a,b] \to X$ be a rectifiable curve in a metric space X. For each t in $[a,b]$, let $\gamma_{a,t}$ be the restriction of γ to $[a,t]$. Define a function $\lambda : [a,b] \to \mathbb{R}$ by $\lambda(a) = 0$ and $\lambda(t) = |\gamma_{a,t}|$ if $t > a$. Prove that λ is continuous.

3. Let $\gamma : [a,b] \to X$ be a curve from x to y in a metric space X with $x \neq y$. Prove that $|\gamma| = d(x,y)$ if and only if γ maps $[a,b]$ onto a geodesic segment joining x to y and $d(x, \gamma(t))$ is an increasing function of t.

4. Let $\gamma = (\gamma_1, \ldots, \gamma_n)$ be a curve in E^n. Prove that γ is rectifiable in E^n if and only if each of its component functions γ_i is rectifiable in $\mathbb{R}$.

5. Define $\gamma : [0,1] \to \mathbb{R}$ by $\gamma(0) = 0$ and $\gamma(t) = t\sin(1/t)$ if $t > 0$. Show that γ is a nonrectifiable curve in $\mathbb{R}$.

6. Let $\gamma : [a,b] \to X$ be a curve in a metric space X. Define $\gamma^{-1} : [a,b] \to X$ by $\gamma^{-1}(t) = \gamma(a+b-t)$. Show that $|\gamma^{-1}| = |\gamma|$.

7. Let $\gamma : [a,b] \to X$ be a curve in a metric space X and let $\eta : [a,b] \to [c,d]$ be an increasing homeomorphism. The curve $\gamma\eta^{-1} : [c,d] \to X$ is called a *reparameterization* of γ. Show that $|\gamma\eta^{-1}| = |\gamma|$.

8. Let $\gamma : [a,b] \to E^n$ be a C^1 curve. Show that γ has a reparameterization, given by $\eta : [a,b] \to [a,b]$, so that $\gamma\eta^{-1}$ is a C^1 curve and

$$(\gamma\eta^{-1})'(a) = 0 = (\gamma\eta^{-1})'(b).$$

Conclude that a piecewise C^1 curve can be reparameterized into a C^1 curve.

§1.6. Historical Notes

§1.1. For commentary on Euclid's fifth postulate, see Heath's translation of Euclid's *Elements* [118]. Gauss's correspondence and notes on non-Euclidean geometry can be found in Vol. VIII of his *Werke* [150]. For a translation of Gauss's 1824 letter to Taurinus, see Greenberg's 1974 text *Euclidean and non-Euclidean Geometries* [166]. A German translation of Lobachevsky's 1829-1830 Russian paper *On the principles of geometry* can be found in Engel's 1898 treatise *N. I. Lobatschefskij* [262]. Bolyai's 1832 paper *Scientiam spatii absolute veram exhibens*, with commentary, can be found in the 1987 translation *Appendix* [51]. Hyperbolic geometry is also called *Lobachevskian geometry.*

For the early history of non-Euclidean geometry, see Bonola's 1912 treatise *Non-Euclidean Geometry* [52]. See also Gray's 1979 article *Non-Euclidean geometry – a re-interpretation* [159], Gray's 1987 article *The discovery of non-Euclidean geometry* [161], Milnor's 1982 article *Hyperbolic geometry: the first 150 years* [290], and Houzel's 1992 article *The birth of non-Euclidean geometry* [200]. A comprehensive history of non-Euclidean geometry can be found in Rosenfeld's 1988 treatise *A History of Non-Euclidean Geometry* [353]. For a list of the early literature on non-Euclidean geometry, see Sommerville's 1970 *Bibliography of Non-Euclidean Geometry* [377].

For an explanation of the duality between spherical and hyperbolic geometries, see Chapter 5 of Helgason's 1978 treatise *Differential Geometry, Lie Groups, and Symmetric Spaces* [188]. The intrinsic curvature of a surface was formulated by Gauss in his 1828 treatise *Disquisitiones generales circa superficies curvas.* For a translation, with commentary, see Dombrowski's 1979 treatise *150 years after Gauss' "disquisitiones generales circa superficies curvas"* [148]. Commentary on Gauss's treatise and the derivation of Formula 1.1.1 can be found in Vol. II of Spivak's 1979 treatise *Differential Geometry* [378]. The tractroid was shown to have constant negative curvature by Minding in his 1839 paper *Wie sich entscheiden läfst, ob zwei gegebene krumme Flächen auf einander abwickelbar sind oder nicht* [292].

§1.2. Beltrami introduced the projective disk model of the hyperbolic plane in his 1868 paper *Saggio di interpetrazione della geometria non-euclidea* [38]. In this paper, Beltrami concluded that the intrinsic geometry of a surface of constant negative curvature is non-Euclidean. Klein's interpretation of hyperbolic geometry in terms of projective geometry appeared in his 1871 paper *Ueber die sogenannte Nicht-Euklidische Geometrie* [224]. In this paper, Klein introduced the term *hyperbolic geometry.* Beltrami introduced the conformal disk and upper half-plane models of the hyperbolic plane in his 1868 paper *Teoria fondamentale degli spazii di curvatura costante* [39]. The mathematical basis of Escher's circle prints is explained in Coxeter's 1981 article *Angels and devils* [94]. See also the proceedings of the 1985 M. C. Escher congress *M. C. Escher: Art and Science* [117]. Poincaré identified the linear fractional transformations of the complex upper half-plane with the congruence transformations of the hyperbolic plane in his 1882 memoir *Théorie des groupes fuchsiens* [330]. Hilbert's nonimbedding theorem for smooth complete surfaces of constant negative curvature appeared in his 1901 paper *Ueber Flächen von constanter Gaussscher Krümmung* [190]. For a proof of Hilbert's nonimbedding theorem for C^2 surfaces, see Milnor's 1972 paper *Efimov's theorem about complete immersed surfaces of negative curvature* [291].

§1.3. The study of n-dimensional geometry was initiated by Cayley in his 1843 paper *Chapters in the analytical geometry of* (n) *dimensions* [74]. Vectors in n-dimensions were introduced by Grassmann in his 1844 treatise *Die lineale Ausdehnungslehre* [156]. The Euclidean inner product appeared in Grassmann's 1862 revision of the *Ausdehnungslehre* [157], [158]. The Euclidean norm of an n-tuple of real numbers and Cauchy's inequality for the Euclidean inner product appeared in Cauchy's 1821 treatise *Cours d'Analyse* [71]. Formula 1.3.3 appeared in Schläfli's 1858 paper *On the multiple integral* $\int dxdy\cdots dz$ [360]. The triangle inequality is essentially Proposition 20 in Book I of Euclid's *Elements* [118]. The Euclidean distance between points in n-dimensional space was defined by Cauchy in his 1847 paper *Mémoire sur les lieux analytiques* [73]. The early history of n-dimensional Euclidean geometry can be found in Rosenfeld's 1988 treatise [353]. For the history of vectors, see Crowe's 1967 treatise *A History of Vector Analysis* [97].

The notion of a metric was introduced by Fréchet in his 1906 paper *Sur quelques points du calcul fonctionnel* [137]. Metric spaces were defined by Hausdorff in his 1914 treatise *Grundzüge der Mengenlehre* [181]. Orthogonal transformations in n-dimensions were first considered implicitly by Euler in his 1771 paper *Problema algebraicum ob affectiones prorsus singulares memorabile* [124]. Orthogonal transformations in n-dimensions were considered explicitly by Cauchy in his 1829 paper *Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes* [72]. The term *orthogonal transformation* appeared in Schläfli's 1855 paper *Réduction d'une intégrale multiple, qui comprend l'arc de cercle et l'aire du triangle sphérique comme cas particuliers* [359]. The term *group* was introduced by Galois in his 1831 paper *Mémoire sur les conditions de résolubilité des équations par radicaux* [146], which was published posthumously in 1846. The group of rotations of Euclidean 3-space appeared in Jordan's 1867 paper *Sur les groupes de mouvements* [205]. For the early history of group theory, see Wussing's 1984 history *The Genesis of the Abstract Group Concept* [418].

All the essential material in §1.3 in dimension three appeared in Euler's 1771 paper [124] and in his 1776 paper *Formulae generales pro translatione quacunque corporum rigidorum* [126]. See also Lagrange's 1773 papers *Nouvelle solution du problème du mouvement de rotation* [249] and

Sur l'attraction des sphéroides elliptiques [250]. The group of orientation preserving isometries of Euclidean 3-space appeared in Jordan's 1867 paper [205]. The group of similarities of Euclidean n-space appeared in Klein's 1872 *Erlanger Program* [226]. For commentary on Klein's Erlanger Program, see Hawkins' 1984 paper *The Erlanger Programm of Felix Klein* [185], Birkhoff and Bennett's 1988 article *Felix Klein and his "Erlanger Programm"* [48], and Rowe's 1992 paper *Klein, Lie, and the "Erlanger Programm"* [354]. Isometries of Euclidean n-space were studied by Jordan in his 1875 paper *Essai sur la géométrie à n dimensions* [207]. For an overview of the development of geometry and group theory in the nineteenth century, see Klein's 1928 historical treatise *Development of Mathematics in the 19th Century* [238] and Yaglom's 1988 monograph *Felix Klein and Sophus Lie* [420].

§1.4. The hypothesis that a line segment is the shortest path between two points was taken as a basic assumption by Archimedes in his third century B.C. treatise *On the sphere and cylinder* [23]. The concept of a geodesic arose out of the problem of finding a shortest path between two points on a surface at the end of the seventeenth century. Euler first published the differential equation satisfied by a geodesic on a surface in his 1732 paper *De linea brevissima in superficie quacunque duo quaelibet puncta jungente* [119]. For the history of geodesics, see Stäckel's 1893 article *Bemerkungen zur Geschichte der geodätischen Linien* [379]. The general theory of geodesics in metric spaces can be found in Busemann's 1955 treatise *The Geometry of Geodesics* [63].

§1.5. Archimedes approximated the length of a circle by the perimeters of inscribed and circumscribed regular polygons in his third century B.C. treatise *On the Measurement of the Circle* [23]. Latin translation of the works of Archimedes and Apollonius in the Middle Ages and the introduction of analytic geometry by Fermat and Descartes around 1637 spurred the development of geometric techniques for finding tangents and quadratures of plane curves in the first half of the seventeenth century. This led to a series of geometric rectifications of curves in the middle of the seventeenth century. In particular, the first algebraic formula for the length of a nonlinear curve, $y^2 = x^3$, was found independently by Neil, van Heuraet, and Fermat around 1658. In the last third of the seventeenth century, calculus was created independently by Newton and Leibniz. In particular, they discovered the element of Euclidean arc length and used integration to find the length of plane curves. For a concise history of arc length, see Boyer's 1964 article *Early rectifications of curves* [57]. A comprehensive history of arc length can be found in Traub's 1984 thesis *The Development of the Mathematical Analysis of Curve Length from Archimedes to Lebesgue* [391]. All the essential material in §1.5 appeared in Vol. I of Jordan's 1893 treatise *Cours d'Analyse* [210]. Arc length in metric spaces was introduced by Menger in his 1930 paper *Zur Metrik der Kurven* [286]. For the general theory of arc length in metric spaces, see Busemann's 1955 treatise [63].

CHAPTER 2

Spherical Geometry

In this chapter, we study spherical geometry. In order to emphasize the duality between spherical and hyperbolic geometries, a parallel development of hyperbolic geometry will be given in Chapter 3. In many cases, the arguments will be the same except for minor changes. As spherical geometry is much easier to understand, it is advantageous to first study spherical geometry before taking up hyperbolic geometry. We begin by studying spherical n-space. Elliptic n-space is considered in Section 2.2. Spherical arc length and volume are studied in Sections 2.3 and 2.4. The chapter ends with a section on spherical trigonometry.

§2.1. Spherical n-Space

The standard model for n-dimensional spherical geometry is the unit sphere S^n of $\mathbb{R}^{n+1}$ defined by

$$S^n = \{x \in \mathbb{R}^{n+1} : |x| = 1\}.$$

The *Euclidean metric* d_E on S^n is defined by the formula

$$d_E(x, y) = |x - y|. \tag{2.1.1}$$

The Euclidean metric on S^n is sufficient for most purposes, but it is not intrinsic to S^n, since it is defined in terms of the vector space structure of $\mathbb{R}^{n+1}$. We shall define an intrinsic metric on S^n, but first we need to review cross products in $\mathbb{R}^3$.

Cross Products

Let x, y be vectors in $\mathbb{R}^3$. The *cross product* of x and y is defined to be

$$x \times y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1). \tag{2.1.2}$$

The proof of the next theorem is routine and is left to the reader.

Theorem 2.1.1. *If w, x, y, z are vectors in $\mathbb{R}^3$, then*

(1) $\quad x \times y = -y \times x,$

(2) $$(x \times y) \cdot z = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix},$$

(3) $\quad (x \times y) \times z = (x \cdot z)y - (y \cdot z)x,$

(4) $$(x \times y) \cdot (z \times w) = \begin{vmatrix} x \cdot z & x \cdot w \\ y \cdot z & y \cdot w \end{vmatrix}.$$

Let x, y, z be vectors in $\mathbb{R}^3$. The real number $(x \times y) \cdot z$ is called the *scalar triple product* of x, y, z. It follows from Theorem 2.1.1(2) that

$$(x \times y) \cdot z = (y \times z) \cdot x = (z \times x) \cdot y. \tag{2.1.3}$$

Thus, the value of the scalar triple product of x, y, z remains unchanged when the vectors are cyclically permuted. Consequently

$$(x \times y) \cdot x = (x \times x) \cdot y = 0$$

and

$$(x \times y) \cdot y = (y \times y) \cdot x = 0.$$

Hence $x \times y$ is orthogonal to both x and y. It follows from Theorem 2.1.1 (4) and Formula 1.3.3 that if x and y are nonzero, then

$$|x \times y| = |x|\,|y| \sin\theta(x, y), \tag{2.1.4}$$

where $\theta(x, y)$ is the Euclidean angle between x and y.

Let A be in $\mathrm{O}(3)$. Then a straightforward calculation shows that

$$A(x \times y) = (\det A)(Ax \times Ay). \tag{2.1.5}$$

In particular, a rotation of $\mathbb{R}^3$ preserves cross products. Consequently, the direction of $x \times y$ relative to x and y is given by the right-hand rule, since $e_1 \times e_2 = e_3$.

The Spherical Metric

Let x, y be vectors in S^n and let $\theta(x, y)$ be the Euclidean angle between x and y. The *spherical distance* between x and y is defined to be the real number

$$d_S(x, y) = \theta(x, y). \tag{2.1.6}$$

Note that

$$0 \leq d_S(x, y) \leq \pi$$

and $d_S(x, y) = \pi$ if and only if $y = -x$. Two vectors x, y in S^n are said to be *antipodal* if and only if $y = -x$.

Theorem 2.1.2. *The spherical distance function d_S is a metric on S^n.*

Proof: The function d_S is obviously nonnegative, nondegenerate, and symmetric. It remains only to prove the triangle inequality. The orthogonal transformations of $\mathbb{R}^{n+1}$ act on S^n and obviously preserve spherical distances. Thus, we are free to transform x, y, z by an orthogonal transformation. Now the three vectors x, y, z span a vector subspace of $\mathbb{R}^{n+1}$ of dimension at most three. By Theorem 1.3.4, we may assume that x, y, z are in the subspace of $\mathbb{R}^{n+1}$ spanned by e_1, e_2, e_3. In other words, we may assume that $n = 2$. Then we have

$$\begin{aligned}
&\cos(\theta(x,y) + \theta(y,z)) \\
&\quad = \cos\theta(x,y)\cos\theta(y,z) - \sin\theta(x,y)\sin\theta(y,z) \\
&\quad = (x\cdot y)(y\cdot z) - |x\times y|\,|y\times z| \\
&\quad \leq (x\cdot y)(y\cdot z) - (x\times y)\cdot(y\times z) \\
&\quad = (x\cdot y)(y\cdot z) - ((x\cdot y)(y\cdot z) - (x\cdot z)(y\cdot y)) \\
&\quad = x\cdot z \\
&\quad = \cos\theta(x,z).
\end{aligned}$$

Thus, we have that $\theta(x,z) \leq \theta(x,y) + \theta(y,z)$. □

The metric d_S on S^n is called the *spherical metric.* The metric topology of S^n determined by d_S is the same as the metric topology of S^n determined by d_E. The metric space consisting of S^n together with its spherical metric d_S is called *spherical n-space.* Henceforth S^n will denote spherical n-space. An isometry from S^n to itself is called a *spherical isometry.*

Remark: A function $\phi : S^n \to S^n$ is an isometry if and only if it is an isometry with respect to the Euclidean metric on S^n because of the following identity on S^n:

$$x\cdot y = 1 - \frac{1}{2}|x-y|^2.$$

Theorem 2.1.3. *Every orthogonal transformation of $\mathbb{R}^{n+1}$ restricts to an isometry of S^n, and every isometry of S^n extends to a unique orthogonal transformation of $\mathbb{R}^{n+1}$.*

Proof: Clearly, a function $\phi : S^n \to S^n$ is an isometry if and only if it preserves Euclidean inner products on S^n. Therefore, an orthogonal transformation of $\mathbb{R}^{n+1}$ restricts to an isometry of S^n. The same argument as in the proof of Theorem 1.3.2 shows that an isometry of S^n extends to a unique orthogonal transformation of $\mathbb{R}^{n+1}$. □

Corollary 1. *The group of spherical isometries $\mathrm{I}(S^n)$ is isomorphic to the orthogonal group $\mathrm{O}(n+1)$.*

Spherical Geodesics

Definition: A *great circle* of S^n is the intersection of S^n with a 2-dimensional vector subspace of $\mathbb{R}^{n+1}$.

Let x and y be distinct points of S^n. If x and y are linearly independent, then x and y span a 2-dimensional subspace $V(x,y)$ of $\mathbb{R}^{n+1}$, and so the set $S(x,y) = S^n \cap V(x,y)$ is the unique great circle of S^n containing both x and y. If x and y are linearly dependent, then $y = -x$. Note that if $n > 1$, then there is a continuum of great circles of S^n containing both x and $-x$, since every great circle of S^n containing x also contains $-x$.

Definition: Three points x, y, z of S^n are *spherically collinear* if and only if there is a great circle of S^n containing x, y, z.

Lemma 1. *If x, y, z are in S^n and*

$$\theta(x,y) + \theta(y,z) = \theta(x,z),$$

then x, y, z are spherically collinear.

Proof: As x, y, z span a vector subspace of $\mathbb{R}^{n+1}$ of dimension at most 3, we may assume that $n = 2$. From the proof of Theorem 2.1.2, we have

$$(x \times y) \cdot (y \times z) = |x \times y|\,|y \times z|.$$

Hence $x \times y$ and $y \times z$ are linearly dependent by Theorem 1.3.1. Therefore $(x \times y) \times (y \times z) = 0$. As

$$(x \times y) \times (y \times z) = (x \cdot (y \times z))y,$$

we have that x, y, z are linearly dependent by Theorem 2.1.1(2). Hence x, y, z lie on a 2-dimensional vector subspace of $\mathbb{R}^{n+1}$ and so are spherically collinear. □

Theorem 2.1.4. *Let $\alpha : [a,b] \to S^n$ be a curve with $b - a < \pi$. Then the following are equivalent:*

1. *The curve α is a geodesic arc.*
2. *There are orthogonal vectors x, y in S^n such that*
$$\alpha(t) = (\cos(t-a))x + (\sin(t-a))y.$$
3. *The curve α satisfies the differential equation $\alpha'' + \alpha = 0$.*

Proof: Let A be an orthogonal transformation of $\mathbb{R}^{n+1}$. Then we have that $(A\alpha)' = A\alpha'$. Consequently α satisfies (3) if and only if $A\alpha$ does. Hence we are free to transform α by an orthogonal transformation. Suppose that α is a geodesic arc. Let t be in the interval $[a,b]$. Then we have

$$\begin{aligned}
\theta(\alpha(a), \alpha(b)) &= b - a \\
&= (t-a) + (b-t) \\
&= \theta(\alpha(a), \alpha(t)) + \theta(\alpha(t), \alpha(b)).
\end{aligned}$$

By Lemma 1, we have that $\alpha(a), \alpha(t), \alpha(b)$ are spherically collinear. As

$$\theta(\alpha(a), \alpha(b)) = b - a < \pi,$$

the points $\alpha(a)$ and $\alpha(b)$ are not antipodal. Hence $\alpha(a)$ and $\alpha(b)$ lie on a unique great circle S of S^n. Therefore, the image of α is contained in S. Hence, we may assume that $n = 1$. By applying a rotation of the form

$$\begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix}$$

we can rotate $\alpha(a)$ to e_1, so we may assume that $\alpha(a) = e_1$. Then

$$e_1 \cdot \alpha(t) = \alpha(a) \cdot \alpha(t) = \cos\theta(\alpha(a), \alpha(t)) = \cos(t - a).$$

Therefore $e_2 \cdot \alpha(t) = \pm\sin(t - a)$. As α is continuous and $b - a < \pi$, we have that either

$$e_2 \cdot \alpha(t) \quad = \quad \sin(t-a) \quad \text{for all } t$$

or

$$e_2 \cdot \alpha(t) \quad = \quad -\sin(t-a) \quad \text{for all } t.$$

In the latter case, we can apply the reflection

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

and so we may assume that

$$\alpha(t) = (\cos(t-a))e_1 + (\sin(t-a))e_2.$$

Thus (1) implies (2).

Next, suppose there are orthogonal vectors x, y in S^n such that

$$\alpha(t) = (\cos(t-a))x + (\sin(t-a))y.$$

Let s and t be such that $a \leq s \leq t \leq b$. Then we have

$$\begin{aligned} \cos\theta(\alpha(s), \alpha(t)) &= \alpha(s) \cdot \alpha(t) \\ &= \cos(s-a)\cos(t-a) + \sin(s-a)\sin(t-a) \\ &= \cos(t-s). \end{aligned}$$

As $t - s < \pi$, we have that $\theta(\alpha(s), \alpha(t)) = t - s$. Thus α is a geodesic arc. Hence (2) implies (1).

Clearly (2) implies (3). Suppose that (3) holds. Then

$$\alpha(t) = \cos(t-a)\alpha(a) + \sin(t-a)\alpha'(a).$$

Upon differentiating the equation $\alpha(t)\cdot\alpha(t) = 1$, we see that $\alpha(t)\cdot\alpha'(t) = 0$. Thus $\alpha(t)$ and $\alpha'(t)$ are orthogonal for all t. In particular, $\alpha(a)$ and $\alpha'(a)$ are orthogonal. Observe that

$$|\alpha(t)|^2 = \cos^2(t-a) + \sin^2(t-a)|\alpha'(a)|^2.$$

As $|\alpha(t)| = 1$, we have that $|\alpha'(a)| = 1$. Thus (3) implies (2). □

The next theorem follows easily from Theorem 2.1.4.

Theorem 2.1.5. *A function $\lambda : \mathbb{R} \to S^n$ is a geodesic line if and only if there are orthogonal vectors x, y in S^n such that*

$$\lambda(t) = (\cos t)x + (\sin t)y.$$

Corollary 2. *The geodesics of S^n are its great circles.*

Exercise 2.1

1. Show that the metric topology of S^n determined by the spherical metric is the same as the metric topology of S^n determined by the Euclidean metric.

2. Let A be a real $n \times n$ matrix. Prove that the following are equivalent:

   (1) A is orthogonal.

   (2) $|Ax| = |x|$ for all x in $\mathbb{R}^n$.

   (3) A preserves the quadratic form $f(x) = x_1^2 + \cdots + x_n^2$.

3. Show that every matrix in SO(2) is of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

4. Show that a curve $\alpha : [a, b] \to S^n$ is a geodesic arc if and only if there are orthogonal vectors x, y in S^n such that

$$\alpha(t) = (\cos(t-a))x + (\sin(t-a))y \quad \text{and} \quad b - a \leq \pi.$$

   Conclude that S^n, with $n > 0$, is geodesically connected but not geodesically convex.

5. Prove Theorem 2.1.5. Conclude that S^n is geodesically complete.

6. A *great m-sphere* of S^n is the intersection of S^n with an $(m+1)$-dimensional vector subspace of $\mathbb{R}^{n+1}$. Show that a subset X of S^n, with more than one point, is totally geodesic if and only if X is a great m-sphere of S^n for some $m > 0$.

7. Let $u_0, \ldots, u_n$ be linearly independent vectors in S^n, let $v_0, \ldots, v_n$ be linearly independent vectors in S^n, and suppose that $\theta(u_i, u_j) = \theta(v_i, v_j)$ for all i, j. Show that there is a unique isometry ϕ of S^n such that $\phi(u_i) = v_i$ for each $i = 0, \ldots, n$.

8. Prove that every similarity of S^n is an isometry.

9. A *tangent vector* to S^n at a point x of S^n is defined to be the derivative at 0 of a differentiable curve $\gamma : [-b, b] \to S^n$ such that $\gamma(0) = x$. Let $T_x = T_x(S^n)$ be the set of all tangent vectors to S^n at x. Show that

$$T_x = \{y \in \mathbb{R}^{n+1} : x \cdot y = 0\}.$$

   Conclude that T_x is an n-dimensional vector subspace of $\mathbb{R}^{n+1}$. The vector space T_x is called the *tangent space* of S^n at x.

10. A *coordinate frame* of S^n is a n-tuple $(\lambda_1, \ldots, \lambda_n)$ of functions such that

    (1) the function $\lambda_i : \mathbb{R} \to S^n$ is a geodesic line for each $i = 1, \ldots, n$;

    (2) there is a point x of S^n such that $\lambda_i(0) = x$ for all i; and

    (3) the set $\{\lambda_1'(0), \ldots, \lambda_n'(0)\}$ is an orthonormal basis of $T_x(S^n)$.

    Show that the action of $\mathrm{I}(S^n)$ on the set of coordinate frames of S^n, given by $\phi(\lambda_1, \ldots, \lambda_n) = (\phi\lambda_1, \ldots, \phi\lambda_n)$, is transitive.

§2.2. Elliptic n-Space

The antipodal map $\alpha : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, defined by $\alpha(x) = -x$, obviously commutes with every orthogonal transformation of $\mathbb{R}^{n+1}$; consequently, spherical geometry is antipodally symmetric. The antipodal symmetry of spherical geometry leads to a duplication of geometric information. For example, if three great circles of S^2 form the sides of a spherical triangle, then they also form the sides of the antipodal image of the triangle. See Figure 2.5.3 for an illustration of this duplication.

The antipodal duplication in spherical geometry is easily eliminated by identifying each pair of antipodal points $x, -x$ of S^n to one point $\pm x$. The resulting quotient space is called *real projective n-space*. P^n The spherical metric d_S on S^n induces a metric d_P on P^n defined by

$$d_P(\pm x, \pm y) = \min\{d_S(x, y), d_S(x, -y)\}. \tag{2.2.1}$$

Notice that $d_P(\pm x, \pm y)$ is just the spherical distance from the set $\{x, -x\}$ to the set $\{y, -y\}$ in S^n. The metric space consisting of P^n and the metric d_P is called *elliptic n-space*. The lines (geodesics) of P^n are the images of the geodesics of S^n with respect to the natural projection $\eta : S^n \to P^n$. As η is a double covering, each line of P^n is a circle that is double covered by a great circle of S^n. Elliptic geometry, unlike spherical geometry, shares with Euclidean geometry the property that there is a unique line passing through each pair of distinct points.

Gnomonic Projection

Identify $\mathbb{R}^n$ with $\mathbb{R}^n \times \{0\}$ in $\mathbb{R}^{n+1}$. The *gnomonic projection*

$$\nu : \mathbb{R}^n \to S^n$$

is defined to be the composition of the vertical translation of $\mathbb{R}^n$ by e_{n+1} followed by radial projection to S^n. See Figure 2.2.1. An explicit formula for ν is given by

$$\nu(x) = \frac{x + e_{n+1}}{|x + e_{n+1}|}. \tag{2.2.2}$$

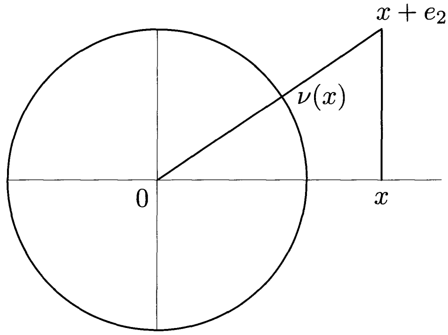


Figure 2.2.1. The gnomonic projection ν of $\mathbb{R}$ into S^1

The function ν maps $\mathbb{R}^n$ bijectively onto the upper hemisphere of S^n. Hence, the function $\eta\nu : \mathbb{R}^n \to P^n$ is an injection. The complement of $\eta\nu(\mathbb{R}^n)$ in P^n is P^{n-1}, which corresponds to the equator of S^n with antipodal points identified.

Classical real projective n-space is the set

$$\overline{\mathbb{R}}^n = \mathbb{R}^n \cup P^{n-1}$$

with P^{n-1} adjoined to $\mathbb{R}^n$ at infinity. In $\overline{\mathbb{R}}^n$, a point at infinity in P^{n-1} is adjoined to each line of $\mathbb{R}^n$ forming a finite line. Two finite lines intersect if and only if they intersect in $\mathbb{R}^n$ or they are parallel in $\mathbb{R}^n$, in which case they intersect at their common point at infinity. Besides the finite lines, there are the lines of P^{n-1} at infinity. When $n = 2$, there is exactly one line at infinity. Classically, the real projective plane refers to the Euclidean plane $\mathbb{R}^2$ together with one line at infinity adjoined to it so that lines intersect as described above.

The injection $\eta\nu : \mathbb{R}^n \to P^n$ extends by the identity map on P^{n-1} to a bijection $\overline{\nu} : \overline{\mathbb{R}}^n \to P^n$ that maps the lines of $\overline{\mathbb{R}}^n$ to the lines of P^n. Classical real projective n-space is useful in understanding elliptic geometry, since the finite lines of $\overline{\mathbb{R}}^n$ correspond to the lines of $\mathbb{R}^n$.

Exercise 2.2

1. Prove that d_P is a metric on P^n.
2. Let $\eta : S^n \to P^n$ be the natural projection. Show that if x is in S^n and $r > 0$, then $\eta(B(x,r)) = B(\eta(x), r)$.
3. Show that η maps the open hemisphere $B(x, \pi/2)$ homeomorphically onto $B(\eta(x), \pi/2)$. Conclude that η is a double covering.
4. Show that η maps $B(x, \pi/4)$ isometrically onto $B(\eta(x), \pi/4)$.
5. Prove that the geodesics of P^n are the images of the great circles of S^n with respect to η.

6. Show that P^1 is isometric to $\frac{1}{2}S^1$.
7. Show that the complement in P^2 of an open ball $B(x,r)$, with $r < \pi/2$, is a Möbius band.
8. Let x be a point of P^3 at a distance $s > 0$ from a geodesic L of P^3. Show that there is a geodesic L' of P^3 passing through x such that each point in L' is at a distance s from L. The geodesics L and L' are called *Clifford parallels.*
9. Let $S^n_+ = \{x \in S^n : x_{n+1} > 0\}$. Define $\phi : S^n_+ \to \mathbb{R}^n$ by
$$\phi(x_1, \dots, x_{n+1}) = (x_1/x_{n+1}, \dots, x_n/x_{n+1}).$$
Show that ϕ is inverse to $\nu : \mathbb{R}^n \to S^n$. Conclude that ν maps $\mathbb{R}^n$ homeomorphically onto S^n_+.
10. Define an *m-plane* Q of P^n to be the image of a great m-sphere of S^n with respect to the natural projection $\eta : S^n \to P^n$. Show that the intersection of a corresponding m-plane Q of $\overline{\mathbb{R}}^n$ with $\mathbb{R}^n$ is either an m-plane of E^n or the empty set, in which case Q is an m-plane at infinity in P^{n-1}.

§2.3. Spherical Arc Length

In this section, we determine the element of spherical arc length of S^n.

Theorem 2.3.1. *A curve $\gamma : [a,b] \to S^n$ is rectifiable in S^n if and only if γ is rectifiable in $\mathbb{R}^{n+1}$; moreover, the spherical length of γ is the same as the Euclidean length of γ.*

Proof: The following inequality holds for all θ:
$$1 - \theta^2/2 \leq \cos\theta \leq 1 - \theta^2/2 + \theta^4/24.$$
Hence, we have that
$$\theta^2 - \theta^4/12 \leq 2(1 - \cos\theta) \leq \theta^2.$$
Let x, y be in S^n. Then
$$|x - y|^2 = 2(1 - \cos\theta(x,y)).$$
Consequently
$$|x - y| \leq \theta(x,y) \leq \frac{|x-y|}{\sqrt{1 - \theta^2(x,y)/12}}.$$
As $0 \leq \theta(x,y) \leq \pi$, we have
$$|x - y| \leq \theta(x,y) \leq \frac{|x-y|}{\sqrt{1 - \pi^2/12}}.$$

Let P be a partition of $[a,b]$ and let $\ell_S(\gamma,P)$ and $\ell_E(\gamma,P)$ be the spherical and Euclidean P-inscribed length of γ, respectively. Then we have

$$\ell_E(\gamma,P) \leq \ell_S(\gamma,P) \leq \frac{\ell_E(\gamma,P)}{\sqrt{1-\pi^2/12}}.$$

Let $|\gamma|_S$ and $|\gamma|_E$ be the spherical and Euclidean length of γ, respectively. Then we have that

$$|\gamma|_E \leq |\gamma|_S \leq \frac{|\gamma|_E}{\sqrt{1-\pi^2/12}}.$$

Therefore γ is rectifiable in S^n if and only if γ is rectifiable in $\mathbb{R}^{n+1}$.

Suppose that $|P| \leq \delta$ and set

$$\mu(\gamma,\delta) = \sup\{\theta(\gamma(s),\gamma(t)) : |t-s| \leq \delta\}.$$

Then we have that

$$\ell_S(\gamma,P) \leq \frac{\ell_E(\gamma,P)}{\sqrt{1-\mu^2/12}}.$$

Hence, we have that

$$|\gamma|_S \leq \frac{|\gamma|_E}{\sqrt{1-\mu^2/12}}.$$

As $\gamma : [a,b] \to S^n$ is uniformly continuous, $\mu(\gamma,\delta)$ goes to zero with δ. Therefore $|\gamma|_S \leq |\gamma|_E$. Thus $|\gamma|_S = |\gamma|_E$. □

Corollary 1. *The element of spherical arc length of S^n is the element of Euclidean arc length of $\mathbb{R}^{n+1}$ restricted to S^n.*

§2.4. Spherical Volume

Let x be a vector in $\mathbb{R}^{n+1}$ such that x_n and x_{n+1} are not both zero. The *spherical coordinates* $(\rho, \theta_1, \ldots, \theta_n)$ of x are defined as follows:

(1) $\rho = |x|$,

(2) $\theta_i = \theta(e_i, x_i e_i + x_{i+1}e_{i+1} + \cdots + x_{n+1}e_{n+1})$ if $i < n$,

(3) θ_n is the polar angle from e_n to $x_n e_n + x_{n+1}e_{n+1}$.

The spherical coordinates of x satisfy the system of equations

$$\begin{aligned} x_1 &= \rho\cos\theta_1, \\ x_2 &= \rho\sin\theta_1\cos\theta_2, \\ &\vdots \\ x_n &= \rho\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-1}\cos\theta_n, \\ x_{n+1} &= \rho\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-1}\sin\theta_n. \end{aligned} \tag{2.4.1}$$

A straightforward calculation shows that

$$\text{(1)} \qquad \frac{\partial x}{\partial \rho} = \frac{x}{|x|}, \tag{2.4.2}$$

$$\text{(2)} \qquad \left|\frac{\partial x}{\partial \theta_i}\right| = \rho \sin\theta_1 \cdots \sin\theta_{i-1}, \tag{2.4.3}$$

$$\text{(3)} \qquad \frac{\partial x}{\partial \rho}, \frac{\partial x}{\partial \theta_1}, \ldots, \frac{\partial x}{\partial \theta_n} \text{ are orthogonal.} \tag{2.4.4}$$

This implies that the Jacobian of the spherical coordinate transformation $(\rho, \theta_1, \ldots, \theta_n) \mapsto (x_1, \ldots, x_{n+1})$ is $\rho^{n-1} \sin^{n-1}\theta_1 \sin^{n-2}\theta_2 \cdots \sin\theta_{n-1}$.

The *spherical coordinate parameterization* of S^n is the map

$$g : [0, \pi]^{n-1} \times [0, 2\pi] \to S^n$$

defined by $g(\theta_1, \ldots, \theta_n) = (x_1, \ldots, x_{n+1})$, where x_i is expressed in terms of $\theta_1, \ldots, \theta_n$ by the system of Equations (2.4.1). The map g is surjective, and injective on the open set $(0, \pi)^{n-1} \times (0, 2\pi)$.

A subset X of S^n is said to be *measurable* in S^n if and only if $g^{-1}(X)$ is measurable in $\mathbb{R}^n$. In particular, all the Borel subsets of S^n are measurable in S^n. If X is measurable in S^n, then the *spherical volume* of X is defined to be

$$\operatorname{Vol}(X) = \int_{g^{-1}(X)} \sin^{n-1}\theta_1 \sin^{n-2}\theta_2 \cdots \sin\theta_{n-1} d\theta_1 \cdots d\theta_n. \tag{2.4.5}$$

The motivation for Formula 2.4.5 is as follows: Subdivide the rectangular solid $[0, \pi]^{n-1} \times [0, 2\pi]$ into a rectangular grid. Each grid rectangular solid of volume $\Delta\theta_1 \cdots \Delta\theta_n$ that meets $g^{-1}(X)$ corresponds under g to a region in S^n that meets X. This region is approximated by the rectangular solid spanned by the vectors $\frac{\partial g}{\partial \theta_1}\Delta\theta_1, \ldots, \frac{\partial g}{\partial \theta_n}\Delta\theta_n$. Its volume is given by

$$\left|\frac{\partial g}{\partial \theta_1}\Delta\theta_1\right| \cdots \left|\frac{\partial g}{\partial \theta_n}\Delta\theta_n\right| = \sin^{n-1}\theta_1 \sin^{n-2}\theta_2 \cdots \sin\theta_{n-1}\Delta\theta_1 \cdots \Delta\theta_n.$$

As the mesh of the subdivision goes to zero, the sum of the volumes of the approximating rectangular solids approaches the volume of X as a limit.

Let X be a measurable subset of S^n and let ϕ be an orthogonal transformation of $\mathbb{R}^{n+1}$. It is a basic fact of advanced calculus that $\phi(X)$ is also measurable in S^n, and the volume of $\phi(X)$ can be measured with respect to the new parameterization ϕg of S^n. As ϕ maps the rectangular solid spanned by the vectors $\frac{\partial g}{\partial \theta_1}\Delta\theta_1, \ldots, \frac{\partial g}{\partial \theta_n}\Delta\theta_n$ onto the rectangular solid spanned by the vectors $\frac{\partial \phi g}{\partial \theta_1}\Delta\theta_1, \ldots, \frac{\partial \phi g}{\partial \theta_n}\Delta\theta_n$, we deduce that

$$\operatorname{Vol}(\phi(X)) = \operatorname{Vol}(X).$$

In other words, spherical volume is an isometry-invariant measure on S^n.

It is clear from Formula 2.4.5 that spherical volume is countably additive, that is, if $\{X_i\}_{i=1}^{\infty}$ is a sequence of disjoint measurable subsets of S^n, then $X = \bigcup_{i=1}^{\infty} X_i$ is also measurable in S^n and

$$\mathrm{Vol}(X) = \sum_{i=1}^{\infty} \mathrm{Vol}(X_i).$$

Theorem 2.4.1. *The element of spherical volume for the upper hemisphere $x_{n+1} > 0$ of S^n, with respect to the Euclidean coordinates $x_1, \ldots, x_n$, is*

$$\frac{dx_1 \cdots dx_n}{[1-(x_1^2+\cdots+x_n^2)]^{\frac{1}{2}}}.$$

Proof: It is more convenient for us to show that the element of spherical volume for the hemisphere $x_1 > 0$, with respect to the coordinates $x_2, \ldots, x_{n+1}$, is

$$\frac{dx_2 \cdots dx_{n+1}}{[1-(x_2^2+\cdots+x_{n+1}^2)]^{\frac{1}{2}}}.$$

The desired result will then follow by a simple change of coordinates.

Consider the transformation

$$\overline{g} : (0,\pi/2) \times (0,\pi)^{n-2} \times (0, 2\pi) \to \mathbb{R}^n$$

defined by

$$\overline{g}(\theta_1, \ldots, \theta_n) = (x_2, \ldots, x_{n+1}),$$

where x_i is given by (2.4.1). Then by (2.4.4), the vectors $\frac{\partial \overline{g}}{\partial \theta_1}, \ldots, \frac{\partial \overline{g}}{\partial \theta_n}$ are orthogonal. Hence, the Jacobian of the transformation $\overline{g}$ is given by

$$\begin{aligned} J\overline{g}(\theta_1, \ldots, \theta_n) &= \left|\frac{\partial \overline{g}}{\partial \theta_1}\right| \cdots \left|\frac{\partial \overline{g}}{\partial \theta_n}\right| \\ &= \cos\theta_1 \sin\theta_1^{n-1} \sin^{n-2}\theta_2 \cdots \sin\theta_{n-1}. \end{aligned}$$

By changing variables via $\overline{g}$, we have

$$\int_{g^{-1}(X)} \sin^{n-1}\theta_1 \sin^{n-2}\theta_2 \cdots \sin\theta_{n-1} d\theta_1 \cdots d\theta_n$$

$$= \int_{\overline{g}g^{-1}(X)} \frac{dx_2 \cdots dx_{n+1}}{x_1}$$

$$= \int_{p(X)} \frac{dx_2 \cdots dx_{n+1}}{[1-(x_2^2+\cdots+x_{n+1}^2)]^{\frac{1}{2}}},$$

where $p : S^n \to \mathbb{R}^n$ is the projection

$$p(x_1, \ldots x_{n+1}) = (x_2, \ldots, x_{n+1}).$$

□

Exercise 2.4

1. Show that the spherical coordinates of a vector x in $\mathbb{R}^{n+1}$ satisfy the system of Equations (2.4.1).
2. Show that the spherical coordinate transformation satisfies the Equations (2.4.2)-(2.4.4).
3. Show that the element of spherical arc length dx in spherical coordinates is given by
$$dx^2 = d\theta_1^2 + \sin^2\theta_1 d\theta_2^2 + \cdots + \sin^2\theta_1 \cdots \sin^2\theta_{n-1} d\theta_n^2.$$
4. Let $B(x,r)$ be the spherical disk centered at a point x of S^2 of spherical radius r. Show that the circumference of $B(x,r)$ is $2\pi \sin r$ and the area of $B(x,r)$ is $2\pi(1-\cos r)$. Conclude that $B(x,r)$ has less area than a Euclidean disk of radius r.
5. Show that
$$\text{(1)} \qquad \mathrm{Vol}(S^{2n-1}) = \frac{2\pi^n}{(n-1)!},$$
$$\text{(2)} \qquad \mathrm{Vol}(S^{2n}) = \frac{2^{n+1}\pi^n}{(2n-1)(2n-3)\cdots 3\cdot 1}.$$

§2.5. Spherical Trigonometry

Let x, y, z be three spherically noncollinear points of S^2. Then no two of x, y, z are antipodal. Let $S(x, y)$ be the unique great circle of S^2 containing x and y, and let $H(x, y, z)$ be the closed hemisphere of S^2 with $S(x, y)$ as its boundary and z in its interior. The *spherical triangle* with vertices x, y, z is defined to be
$$T(x,y,z) = H(x,y,z) \cap H(y,z,x) \cap H(z,x,y).$$
We shall assume that the vertices of $T(x, y, z)$ are labeled in positive order as in Figure 2.5.1.

Let $[x, y]$ be the minor arc of $S(x, y)$ joining x to y. The *sides* of $T(x, y, z)$ are defined to be $[x, y], [y, z]$, and $[z, x]$. Let $a = \theta(y, z)$, $b = \theta(z, x)$, and $c = \theta(x, y)$. Then a, b, c is the length of $[y, z], [z, x], [x, y]$, respectively. Let
$$f : [0, a] \to S^2, \ g : [0, b] \to S^2, \ h : [0, c] \to S^2$$
be the geodesic arc from y to z, z to x, and x to y, respectively.

The *angle* α between the sides $[z, x]$ and $[x, y]$ is defined to be the angle between $-g'(b)$ and $h'(0)$. Likewise, the *angle* β between the sides $[x, y]$ and $[y, z]$ is defined to be the angle between $-h'(c)$ and $f'(0)$, and the *angle* γ between the sides $[y, z]$ and $[z, x]$ is defined to be the angle between $-f'(a)$ and $g'(0)$. The angles α, β, γ are called the *angles* of $T(x, y, z)$. The side $[y, z], [z, x], [x, y]$ is said to be *opposite* the angle α, β, γ, respectively.

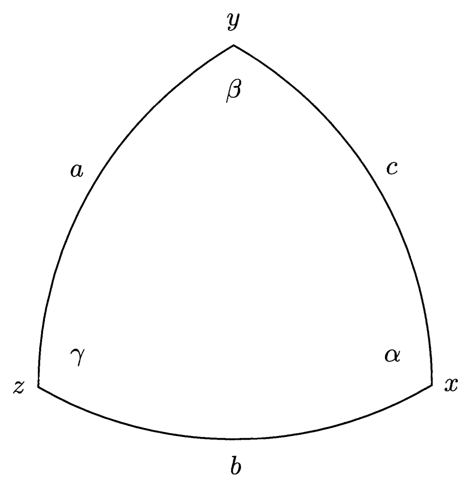


Figure 2.5.1. A spherical triangle $T(x, y, z)$

Lemma 1. *If α, β, γ are the angles of a spherical triangle $T(x, y, z)$, then*

(1) $\theta(z \times x, x \times y) = \pi - \alpha$,

(2) $\theta(x \times y, y \times z) = \pi - \beta$,

(3) $\theta(y \times z, z \times x) = \pi - \gamma$.

Proof: The proof of (1) is evident from Figure 2.5.2. The proof of (2), and (3), is similar. □

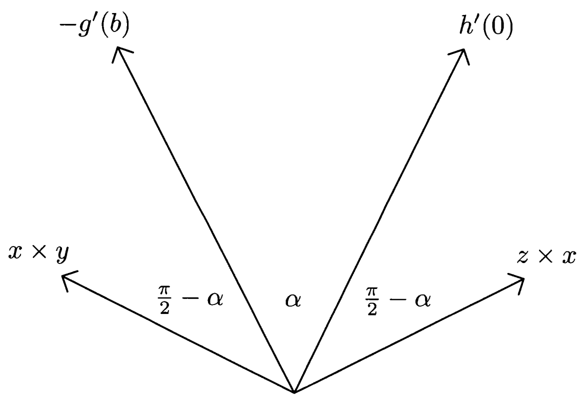


Figure 2.5.2. Four vectors on the tangent plane T_x with $\alpha < \pi/2$

Theorem 2.5.1. *If α, β, γ are the angles of a spherical triangle, then*

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

Proof: Let α, β, γ be the angles of a spherical triangle $T(x, y, z)$. Then

$$\begin{aligned} &((x \times y) \times (z \times y)) \cdot (z \times x) \\ &= [(x \cdot (z \times y))y - (y \cdot (z \times y))x] \cdot (z \times x) \\ &= (x \cdot (z \times y))(y \cdot (z \times x)) \\ &= -(y \cdot (z \times x))^2 \\ &< 0. \end{aligned}$$

By Theorem 2.1.1(2), the vectors $x \times y, z \times y, z \times x$ are linearly independent, and so their associated unit vectors are spherically noncollinear. By Lemma 1 of §2.1, we have

$$\theta(x \times y, z \times x) < \theta(x \times y, z \times y) + \theta(z \times y, z \times x).$$

Now by Lemma 1, we have

$$\pi - \alpha < \beta + \gamma.$$ □

Theorem 2.5.2. (The Law of Sines) *If α, β, γ are the angles of a spherical triangle and a, b, c are the lengths of the opposite sides, then*

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

Proof: Upon taking norms of both sides of the equations

$$\begin{aligned} (z \times x) \times (x \times y) &= (z \cdot (x \times y))x, \\ (x \times y) \times (y \times z) &= (x \cdot (y \times z))y, \\ (y \times z) \times (z \times x) &= (y \cdot (z \times x))z, \end{aligned}$$

we find that

$$\begin{aligned} \sin b \sin c \sin \alpha &= x \cdot (y \times z), \\ \sin c \sin a \sin \beta &= x \cdot (y \times z), \\ \sin a \sin b \sin \gamma &= x \cdot (y \times z). \end{aligned}$$ □

Theorem 2.5.3. (The First Law of Cosines) *If α, β, γ are the angles of a spherical triangle and a, b, c are the lengths of the opposite sides, then*

$$\cos \gamma = \frac{\cos c - \cos a \cos b}{\sin a \sin b}.$$

Proof: Since

$$(y \times z) \cdot (x \times z) = \begin{vmatrix} y \cdot x & y \cdot z \\ x \cdot z & z \cdot z \end{vmatrix},$$

we have that

$$\sin a \sin b \cos \gamma = \cos c - \cos a \cos b.$$ □

Let $T(x,y,z)$ be a spherical triangle. By the same argument as in the proof of Theorem 2.5.1, the vectors $z\times x, x\times y, y\times z$ are linearly independent, and so the associated unit vectors are spherically noncollinear. The spherical triangle

$$T' = T\left(\frac{y\times z}{|y\times z|}, \frac{z\times x}{|z\times x|}, \frac{x\times y}{|x\times y|}\right) \tag{2.5.1}$$

is called the *polar triangle* of $T(x,y,z)$. Let a', b', c' be the lengths of the sides of T' and let α', β', γ' be the opposite angles. By Lemma 1, we have

$$a' = \pi - \alpha,\ b' = \pi - \beta,\ c' = \pi - \gamma.$$

As $T(x,y,z)$ is the polar triangle of T', we have

$$\alpha' = \pi - a,\ \beta' = \pi - b,\ \gamma' = \pi - c.$$

Theorem 2.5.4. (The Second Law of Cosines) *If α, β, γ are the angles of a spherical triangle and a, b, c are the lengths of the opposite sides, then*

$$\cos c = \frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta}.$$

Proof: By the first law of cosines applied to the polar triangle, we have

$$\cos(\pi - c) = \frac{\cos(\pi-\gamma) - \cos(\pi-\alpha)\cos(\pi-\beta)}{\sin(\pi-\alpha)\sin(\pi-\beta)}. \qquad \square$$

Area of Spherical Triangles

A *lune* of S^2 is defined to be the intersection of two distinct, nonopposite hemispheres of S^2. Any lune of S^2 is congruent to a lune $L(\alpha)$ defined in terms of spherical coordinates (ϕ, θ) by the inequalities $0 \leq \theta \leq \alpha$. Here α is the angle formed by the two sides of $L(\alpha)$ at each of its two vertices. See Figure 2.5.3. By Formula 2.4.5, we have

$$\mathrm{Area}(L(\alpha)) = \int_0^\alpha \int_0^\pi \sin\phi\, d\phi d\theta = 2\alpha.$$

As $L(\pi/2)$ is a quarter-sphere, the area of S^2 is 4π.

Theorem 2.5.5. *If α, β, γ are the angles of a spherical triangle T, then*

$$\mathrm{Area}(T) = (\alpha + \beta + \gamma) - \pi.$$

Proof: The three great circles extending the sides of T subdivide S^2 into eight triangular regions which are paired off antipodally. Two of the regions are T and $-T$, and the other six regions are labeled $A, -A, B, -B, C, -C$ in Figure 2.5.4. Any two of the sides of T form a lune with angle α, β, or γ. The lune with angle α is the union of T and A. Hence, we have

$$\mathrm{Area}(T) + \mathrm{Area}(A) = 2\alpha.$$

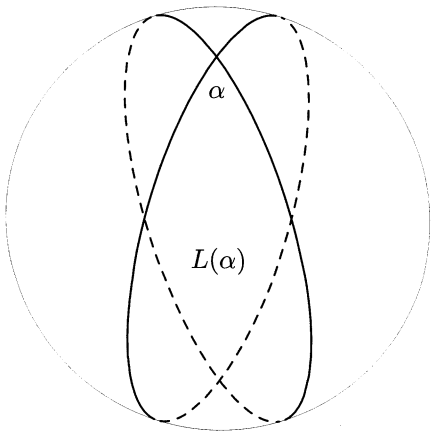


Figure 2.5.3. A lune $L(\alpha)$ of S^2

Likewise, we have that

$$\text{Area}(T) + \text{Area}(B) = 2\beta,$$

$$\text{Area}(T) + \text{Area}(C) = 2\gamma.$$

Adding these three equations and subtracting the equation

$$\text{Area}(T) + \text{Area}(A) + \text{Area}(B) + \text{Area}(C) = 2\pi$$

gives $\text{Area}(T) = \alpha + \beta + \gamma - \pi$. □

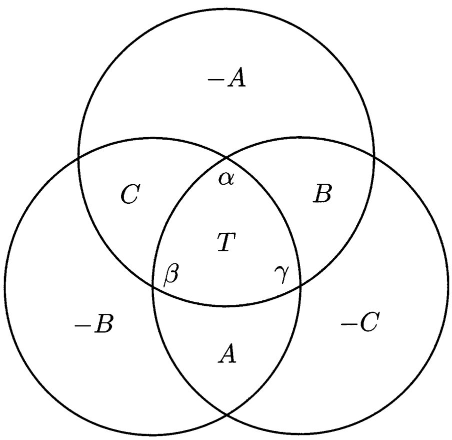


Figure 2.5.4. The subdivision of S^2 into eight triangular regions

Exercise 2.5

1. Let α, β, γ be the angles of a spherical triangle and let a, b, c be the lengths of the opposite sides. Show that

$$\begin{array}{llll}
(1) & \cos a & = & \cos b \cos c + \sin b \sin c \cos \alpha, \\
 & \cos b & = & \cos a \cos c + \sin a \sin c \cos \beta, \\
 & \cos c & = & \cos a \cos b + \sin a \sin b \cos \gamma, \\
(2) & \cos \alpha & = & -\cos \beta \cos \gamma + \sin \beta \sin \gamma \cos a, \\
 & \cos \beta & = & -\cos \alpha \cos \gamma + \sin \alpha \sin \gamma \cos b, \\
 & \cos \gamma & = & -\cos \alpha \cos \beta + \sin \alpha \sin \beta \cos c.
\end{array}$$

2. Let $\alpha, \beta, \pi/2$ be the angles of a spherical right triangle and let a, b, c be the lengths of the opposite sides. Show that

$$(1) \qquad \cos c = \cos a \cos b,$$

$$(2) \qquad \cos c = \cot \alpha \cot \beta,$$

$$(3) \qquad \begin{aligned} \sin a &= \sin c \sin \alpha, \\ \sin b &= \sin c \sin \beta, \end{aligned}$$

$$(4) \qquad \begin{aligned} \cos \alpha &= \tan b \cot c, \\ \cos \beta &= \tan a \cot c, \end{aligned}$$

$$(5) \qquad \begin{aligned} \sin a &= \tan b \cot \beta, \\ \sin b &= \tan a \cot \alpha, \end{aligned}$$

$$(6) \qquad \begin{aligned} \cos \alpha &= \cos a \sin \beta, \\ \cos \beta &= \cos b \sin \alpha. \end{aligned}$$

3. Let α, β, γ be the angles of a spherical triangle such that $\alpha, \beta, \gamma \leq \pi/2$ and let a, b, c be the lengths of the opposite sides. Prove that $a, b, c \leq \pi/2$ and that $a \leq b \leq c$ if and only if $\alpha \leq \beta \leq \gamma$.
4. Let $\alpha, \beta, \pi/2$ be the angles of a spherical right triangle, and let a, b, c be the lengths of the opposite sides. Prove that $\alpha, \beta < \pi/2$ if and only if $a, b, c < \pi/2$.
5. Prove that a spherical triangle is equilateral if and only if it is equiangular.
6. Let $T(x, y, z)$ be a spherical triangle labeled as in Figure 2.5.1 such that $\alpha, \beta < \pi/2$. Prove that a or $b < \pi/2$ and that the point on the great circle through x and y nearest to z lies in the interior of the side $[x, y]$.
7. Let α, β, γ be real numbers in the interval $(0, \pi/2]$ such that $\alpha + \beta + \gamma > \pi$. Prove that there is a spherical triangle with angles α, β, γ.
8. Prove that two spherical triangles are congruent if and only if they have the same angles.

§2.6. Historical Notes

§2.1. Spherical geometry in n-dimensions was first studied by Schläfli in his 1852 treatise *Theorie der vielfachen Kontinuität* [362], which was published posthumously in 1901. The most important results of Schläfli's treatise were published in his 1855 paper *Réduction d'une intégrale multiple, qui comprend l'arc de cercle et l'aire du triangle sphérique comme cas particuliers* [359] and in his 1858-1860 paper *On the multiple integral* $\int dxdy \cdots dz$ [360], [361]. In particular, n-dimensional spheres were defined by Schläfli in this paper [360]. The differential geometry of spherical n-space was first considered by Riemann in his 1854 lecture *Über die Hypothesen, welch der Geometrie zu Grunde liegen* [349], which was published posthumously in 1867. For a translation with commentary, see Vol. II of Spivak's 1979 treatise *Differential Geometry* [378]

The cross product appeared implicitly in Lagrange's 1773 paper *Nouvelle solution du problème du mouvement de rotation* [249]. The cross product evolved in the nineteenth century out of Grassmann's outer product defined in his 1844 *Ausdehnungslehre* [156] and Hamilton's vector product defined in his 1844-1850 paper *On Quaternions* [177]. The basic properties of cross products, in particular, Theorem 2.1.1, appeared in Hamilton's paper *On Quaternions* [177]. The cross product was defined by Gibbs in his 1881 monograph *Elements of Vector Analysis* [152]. The triple scalar product was defined by Hamilton in his paper *On Quaternions* [177]. According to Heath's 1921 treatise *A History of Greek Mathematics* [186], the triangle inequality for spherical geometry is Proposition 5 in Book I of the first century *Sphaerica* of Menelaus. That the geodesics of a sphere are its great circles was affirmed by Euler in his 1732 paper *De linea brevissima in superficie quacunque duo quaelibet puncta jungente* [119].

§2.2. Classical real projective space was introduced by Desargues in his 1639 monograph *Brouillon project d'une atteinte aux événements des recontres du cone avec un plan* [104]. Classical projective geometry was systematically developed by Poncelet in his 1822 treatise *Traité des propriétés projectives des figures* [341]. The metric for the elliptic plane was defined by Cayley in his 1859 paper *A sixth memoir upon quantics* [76]. Moreover, the idea of identifying antipodal points of a sphere to form real projective 2-space appeared in this paper. The term *elliptic geometry* was introduced by Klein in his 1871 paper *Ueber die sogenannte Nicht-Euklidische Geometrie* [224]. Three-dimensional Elliptic geometry was developed by Clifford in his 1873 paper *Preliminary sketch of biquaternions* [82] and by Newcomb in his 1877 paper *Elementary theorems relating to the geometry of a space of three dimensions and of uniform positive curvature* [315]. Real projective 3-space appeared in Killing's 1878 paper *Ueber zwei Raumformen mit constanter positiver Krümmung* [219]. Real projective n-space appeared in Killing's 1885 monograph *Nicht-Euklidischen Raumformen* [221].

§2.3. The element of spherical arc length for the unit sphere was derived by Euler in his 1755 paper *Principes de la trigonométrie sphérique tirés de la méthode des plus grands et plus petits* [120].

§2.4. Spherical coordinates and the element of spherical volume for the unit n-sphere appeared in Jacobi's 1834 paper *Functionibus homogeneis secundi ordinis* [202] and in Green's 1835 paper *On the determination of the exterior and interior attractions of ellipsoids of variable densities* [162]. Moreover, the volume of an n-dimensional sphere was implicitly determined by Jacobi and Green in these papers. Spherical coordinates for Euclidean n-space appeared in Schläfli's 1858 paper [360]. For the theory of measure on manifolds in Euclidean n-space, see Fleming's 1977 text *Functions of Several Variables* [133].

§2.5. According to Heath's 1921 treatise *A History of Greek Mathematics* [186], spherical triangles first appeared in the first century *Sphaerica* of Menelaus. In Book I of the *Sphaerica*, the theorem that the sum of the angles of a spherical triangle exceeds two right angles was established. According to Rosenfeld's 1988 study *A History of Non-Euclidean Geometry* [353], rules equivalent to the spherical sine and cosine laws first appeared in Indian astronomical works of the fifth-eighth centuries. In the ninth century, these rules appeared in the Arabic astronomical treatises of al-Khowarizmi, known in medieval Europe as Algorithmus. The spherical law of sines was proved by Ibn Iraq and Abu l-Wafa in the tenth century. The polar triangle and Lemma 1 appeared in the thirteenth century Arabic treatise *Disclosing the secrets of the figure of secants* by al-Tusi. The first law of cosines appeared in the fifteenth century treatise *De triangulis omnimodis libri quinque* of Regiomontanus, which was published posthumously in 1533. The vector proof of Theorem 2.5.3 (first law of cosines) was given by Hamilton in his paper *On Quaternions* [177]. The second law of cosines appeared in Viète's 1593 treatise *Variorum de rebus mathematicis responsorum liber VIII.* According to Lohne's 1979 article *Essays on Thomas Harriot* [270], the formula for the area of a spherical triangle in terms of the angular excess and its remarkably simple proof was first discovered by Harriot in 1603. However, Theorem 2.5.5 was first published by Girard in his 1629 paper *De la mesure de la superfice des triangles et polygones sphériques* with a more complicated proof. The simple proof of Theorem 2.5.5 appeared in Euler's 1781 paper *De mensura angulorum solidorum* [127]. Spherical trigonometry was thoroughly developed in modern form by Euler in his 1782 paper *Trigonometria sphaerica universa ex primis principiis breviter et dilucide derivata* [128].

CHAPTER 3

Hyperbolic Geometry

We now begin the study of hyperbolic geometry. The first step is to define a new inner product on $\mathbb{R}^n$, called the Lorentzian inner product. This leads to a new concept of length. In particular, imaginary lengths are possible. In Section 3.2, hyperbolic n-space is defined to be the positive half of the sphere of unit imaginary radius in $\mathbb{R}^{n+1}$. The elements of hyperbolic arc length and volume are determined in Sections 3.3 and 3.4. The chapter ends with a section on hyperbolic trigonometry.

§3.1. Lorentzian n-Space

Let x and y be vectors in $\mathbb{R}^n$ with $n > 1$. The *Lorentzian inner product* of x and y is defined to be the real number

$$x \circ y = -x_1y_1 + x_2y_2 + \cdots + x_ny_n. \tag{3.1.1}$$

The Lorentzian inner product is obviously an inner product on $\mathbb{R}^n$. The inner product space consisting of the vector space $\mathbb{R}^n$ together with the Lorentzian inner product is called *Lorentzian n-space*, and is denoted by $\mathbb{R}^{1,n-1}$. Sometimes it is desirable to replace the Lorentzian inner product on $\mathbb{R}^n$ by the equivalent inner product

$$\langle x, y\rangle = x_1y_1 + \cdots + x_{n-1}y_{n-1} - x_ny_n. \tag{3.1.2}$$

The inner product space consisting of $\mathbb{R}^n$ together with this new inner product is also called *Lorentzian n-space* but is denoted by $\mathbb{R}^{n-1,1}$. For example, in the theory of special relativity, $\mathbb{R}^{3,1}$ is a model for space-time. The first three coordinates of a vector $x = (x_1, x_2, x_3, x_4)$ in $\mathbb{R}^{3,1}$ are the space coordinates, and the last is the time coordinate. In this chapter, we shall work in $\mathbb{R}^{1,n-1}$, and for simplicity we shall continue to use the notation $\mathbb{R}^n$ for the underlying vector space of $\mathbb{R}^{1,n-1}$.

The *Lorentzian norm* of a vector x in $\mathbb{R}^n$ is defined to be the complex number

$$\|x\| = (x \circ x)^{\frac{1}{2}}. \tag{3.1.3}$$

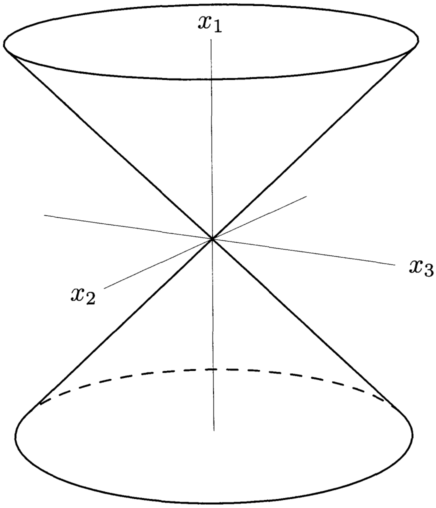


Figure 3.1.1. The light cone C^2 of $\mathbb{R}^{1,2}$

Here $\|x\|$ is either positive, zero, or positive imaginary. If $\|x\|$ is positive imaginary, we denote its absolute value (modulus) by $|||x|||$.

The *Lorentzian distance* between vectors x and y in $\mathbb{R}^n$ is defined to be the complex number

$$d_L(x,y) = \|x - y\|. \tag{3.1.4}$$

Note that $d_L(x,y)$ is either positive, zero, or positive imaginary. The set of all x in $\mathbb{R}^n$ such that $\|x\| = 0$ is the hypercone C^{n-1} defined by the equation

$$x_1^2 = x_2^2 + \cdots + x_n^2. \tag{3.1.5}$$

The hypercone C^{n-1} is called the *light cone* of $\mathbb{R}^n$. See Figure 3.1.1. If $\|x\| = 0$, then x is said to be *light-like*. A light-like vector x is said to be *positive* (resp. *negative*) if and only if $x_1 > 0$ (resp. $x_1 < 0$).

If $\|x\| > 0$, then x is said to be *space-like*. Note that x is space-like if and only if its coordinates satisfy the inequality

$$x_1^2 < x_2^2 + \cdots + x_n^2.$$

The *exterior* of C^{n-1} in $\mathbb{R}^n$ is the open subset of $\mathbb{R}^n$ consisting of all the space-like vectors.

If $\|x\|$ is imaginary, then x is said to be *time-like*. Note that x is time-like if and only if its coordinates satisfy the inequality

$$x_1^2 > x_2^2 + \cdots + x_n^2.$$

A time-like vector x is said to be *positive* (resp. *negative*) if and only if $x_1 > 0$ (resp. $x_1 < 0$). The *interior* of C^{n-1} in $\mathbb{R}^n$ is the open subset of $\mathbb{R}^n$ consisting of all the time-like vectors.

Theorem 3.1.1. *If x and y are positive (resp. negative) time-like vectors in $\mathbb{R}^n$ and $t > 0$, then*

(1) *the vector tx is positive (resp. negative) time-like;*

(2) *the vector $x + y$ is positive (resp. negative) time-like.*

Proof: (1) $\|tx\|^2 = t^2\|x\|^2 < 0$.

(2) $(x_1 + y_1)^2 = x_1^2 + 2x_1y_1 + y_1^2$
$> (x_2^2 + \cdots + x_n^2) + 2(x_2^2 + \cdots + x_n^2)^{\frac{1}{2}}(y_2^2 + \cdots + y_n^2)^{\frac{1}{2}} + (y_2^2 + \cdots + y_n^2)$
$\geq (x_2^2 + \cdots + x_n^2) + 2(x_2y_2 + \cdots + x_ny_n) + (y_2^2 + \cdots + y_n^2)$
$= (x_2 + y_2)^2 + \cdots + (x_n + y_n)^2$. □

Corollary 1. *The set of positive (resp. negative) time-like vectors is a convex subset of $\mathbb{R}^n$.*

Proof: If x and y are positive (resp. negative) time-like vectors in $\mathbb{R}^n$ and $0 < t < 1$, then $(1 - t)x + ty$ is positive (resp. negative) time-like by Theorem 3.1.1. □

Lorentz Transformations

Definition: A function $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is a *Lorentz transformation* if and only if

$$\phi(x) \circ \phi(y) = x \circ y \quad \text{for all } x, y \text{ in } \mathbb{R}^n.$$

A basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$ is said to be *Lorentz orthonormal* if and only if $v_1 \circ v_1 = -1$ and $v_i \circ v_j = \delta_{ij}$ otherwise. Note that the standard basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$ is Lorentz orthonormal.

Theorem 3.1.2. *A function $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is a Lorentz transformation if and only if ϕ is linear and $\{\phi(e_1), \ldots, \phi(e_n)\}$ is a Lorentz orthonormal basis of $\mathbb{R}^n$.*

Proof: Suppose that ϕ is a Lorentz transformation of $\mathbb{R}^n$. Then

$$\phi(e_1) \circ \phi(e_1) = e_1 \circ e_1 = -1$$

and

$$\phi(e_i) \circ \phi(e_j) = e_i \circ e_j = \delta_{ij} \quad \text{otherwise.}$$

This clearly implies that $\phi(e_1), \ldots, \phi(e_n)$ are linearly independent. Hence $\{\phi(e_1), \ldots, \phi(e_n)\}$ is a Lorentz orthonormal basis of $\mathbb{R}^n$.

Let x be in $\mathbb{R}^n$. Then there are coefficients $c_1, \ldots, c_n$ in $\mathbb{R}$ such that

$$\phi(x) = \sum_{i=1}^{n} c_i\phi(e_i).$$

As $\{\phi(e_1), \ldots, \phi(e_n)\}$ is a Lorentz orthonormal basis, we have

$$-c_1 \;=\; \phi(x) \circ \phi(e_1) \;=\; x \circ e_1 \;=\; -x_1$$

and

$$c_j \;=\; \phi(x) \circ \phi(e_j) \;=\; x \circ e_j \;=\; x_j \quad \text{for } j > 1.$$

Then ϕ is linear, since

$$\phi\Big(\sum_{i=1}^{n} x_i e_i\Big) = \sum_{i=1}^{n} x_i \phi(e_i).$$

Conversely, suppose that ϕ is linear and $\{\phi(e_1), \ldots, \phi(e_n)\}$ is a Lorentz orthonormal basis of $\mathbb{R}^n$. Then ϕ is a Lorentz transformation, since

$$\begin{aligned}
\phi(x) \circ \phi(y) &= \phi\left(\sum_{i=1}^{n} x_i e_i\right) \circ \phi\left(\sum_{j=1}^{n} y_j e_j\right) \\
&= \left(\sum_{i=1}^{n} x_i \phi(e_i)\right) \circ \left(\sum_{j=1}^{n} y_j \phi(e_j)\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_i \phi(e_i) \circ \phi(e_j) \\
&= -x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \;=\; x \circ y. \qquad \square
\end{aligned}$$

A real $n \times n$ matrix A is said to be *Lorentzian* if and only if the associated linear transformation $A : \mathbb{R}^n \to \mathbb{R}^n$, defined by $A(x) = Ax$, is Lorentzian. The set of all Lorentzian $n \times n$ matrices together with matrix multiplication forms a group $\mathrm{O}(1, n-1)$, called the *Lorentz group* of $n \times n$ matrices. By Theorem 3.1.2, the group $\mathrm{O}(1, n-1)$ is naturally isomorphic to the group of Lorentz transformations of $\mathbb{R}^n$. The next theorem follows immediately from Theorem 3.1.2.

Theorem 3.1.3. *Let A be a real $n \times n$ matrix. Then the following are equivalent:*

(1) *The matrix A is Lorentzian.*

(2) *The columns of A form a Lorentz orthonormal basis of $\mathbb{R}^n$.*

(3) *The matrix A satisfies the equation $A^t J A = J$, where*

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

(4) *The matrix A satisfies the equation $AJA^t = J$.*

(5) *The rows of A form a Lorentz orthonormal basis of $\mathbb{R}^n$.*

Let A be a Lorentzian matrix. As $A^t JA = J$, we have that $(\det A)^2 = 1$. Thus $\det A = \pm 1$. Let $\mathrm{SO}(1, n-1)$ be the set of all A in $\mathrm{O}(1, n-1)$ such that $\det A = 1$. Then $\mathrm{SO}(1, n-1)$ is a subgroup of index two in $\mathrm{O}(1, n-1)$. The group $\mathrm{SO}(1, n-1)$ is called the *special Lorentz group.*

By Corollary 1, the set of all time-like vectors in $\mathbb{R}^n$ has two connected components, the set of positive time-like vectors and the set of negative time-like vectors. A Lorentzian matrix A is said to be *positive* (resp. *negative*) if and only if A transforms positive time-like vectors into positive (resp. negative) time-like vectors. For example, the matrix J is negative. By continuity, a Lorentzian matrix is either positive or negative.

Let $\mathrm{PO}(1, n-1)$ be the set of all positive matrices in $\mathrm{O}(1, n-1)$. Then $\mathrm{PO}(1, n-1)$ is a subgroup of index two in $\mathrm{O}(1, n-1)$. The group of positive matrices $\mathrm{PO}(1, n-1)$ is called the *positive Lorentz group.* Likewise, let $\mathrm{PSO}(1, n-1)$ be the set of all positive matrices in $\mathrm{SO}(1, n-1)$. Then $\mathrm{PSO}(1, n-1)$ is a subgroup of index two in $\mathrm{SO}(1, n-1)$. The group $\mathrm{PSO}(1, n-1)$ is called the *positive special Lorentz group.*

Definition: Two vectors x, y in $\mathbb{R}^n$ are *Lorentz orthogonal* if and only if $x \circ y = 0$.

Theorem 3.1.4. *Let x and y be nonzero Lorentz orthogonal vectors in $\mathbb{R}^n$. If x is time-like, then y is space-like.*

Proof: As x is time-like, we have that $x_1^2 > x_2^2 + \cdots + x_n^2$. Hence, we have

$$1 > \left(\sum_{i=2}^n x_i^2\right) x_1^{-2}.$$

As $x \circ y = 0$, we have that $x_1 y_1 = x_2 y_2 + \cdots + x_n y_n$. Hence

$$y_1 = \left(\sum_{i=2}^n x_i y_i\right) x_1^{-1}.$$

Observe that

$$\begin{aligned}
\|y\|^2 &= -y_1^2 + y_2^2 + \cdots + y_n^2 \\
&= -\left[\left(\sum_{i=2}^n x_i y_i\right) x_1^{-1}\right]^2 + \sum_{i=2}^n y_i^2 \\
&\geq -\left(\sum_{i=2}^n x_i^2\right)\left(\sum_{i=2}^n y_i^2\right) x_1^{-2} + \sum_{i=2}^n y_i^2 \\
&= \left(\sum_{i=2}^n y_i^2\right)\left[1 - \left(\sum_{i=2}^n x_i^2\right) x_1^{-2}\right] \\
&\geq 0.
\end{aligned}$$

Moreover, if $\|y\|^2 = 0$, then $\sum_{i=2}^n y_i^2 = 0$, and so $y_i = 0$ for $i = 2, \ldots, n$. As $y_1 = (\sum_{i=2}^n x_i y_i)\, x_1^{-1}$, we have $y = 0$. But $y \neq 0$ and so $\|y\| > 0$. □

Definition: Let V be a vector subspace of $\mathbb{R}^n$. Then V is said to be

(1) *time-like* if and only if V has a time-like vector,

(2) *space-like* if and only if every nonzero vector in V is space-like, or

(3) *light-like* otherwise.

Theorem 3.1.5. *For each dimension m, the natural action of $\mathrm{PO}(1, n-1)$ on the set of m-dimensional time-like vector subspaces of $\mathbb{R}^n$ is transitive.*

Proof: Let V be an m-dimensional, time-like, vector subspace of $\mathbb{R}^n$. Identify $\mathbb{R}^m$ with the subspace of $\mathbb{R}^n$ spanned by the vectors $e_1, \ldots, e_m$. It suffices to show that there is an A in $\mathrm{PO}(1, n-1)$ such that $A(\mathbb{R}^m) = V$. Choose a basis $\{u_1, \ldots, u_n\}$ of $\mathbb{R}^n$ such that u_1 is a positive time-like vector in V and $\{u_1, \ldots, u_m\}$ is a basis for V. Let $w_1 = u_1/|||u_1|||$. Then we have that $w_1 \circ w_1 = -1$. Next, let $v_2 = u_2 + (u_2 \circ w_1)w_1$. Then v_2 is nonzero, since u_1 and u_2 are linearly independent; moreover

$$w_1 \circ v_2 = w_1 \circ u_2 + (u_2 \circ w_1)(w_1 \circ w_1) = 0.$$

Therefore v_2 is space-like by Theorem 3.1.4. Now let

$$\begin{aligned}
w_2 &= v_2/\|v_2\|, \\
v_3 &= u_3 + (u_3 \circ w_1)w_1 - (u_3 \circ w_2)w_2, \\
w_3 &= v_3/\|v_3\|, \\
&\vdots \\
v_n &= u_n + (u_n \circ w_1)w_1 - (u_n \circ w_2)w_2 - \cdots - (u_n \circ w_{n-1})w_{n-1}, \\
w_n &= v_n/\|v_n\|.
\end{aligned}$$

Then we have that $\{w_1, \ldots, w_n\}$ is a Lorentz orthonormal basis of $\mathbb{R}^n$ and $\{w_1, \ldots, w_m\}$ is a basis of V. Let A be the $n \times n$ matrix whose columns are $w_1, \ldots, w_n$. Then A is Lorentzian by Theorem 3.1.3, and $A(\mathbb{R}^m) = V$; moreover, A is positive, since $A(e_1) = w_1$ is positive time-like. □

Theorem 3.1.6. *Let x, y be positive (negative) time-like vectors in $\mathbb{R}^n$. Then $x \circ y \leq \|x\|\,\|y\|$ with equality if and only if x and y are linearly dependent.*

Proof: By Theorem 3.1.5, there is an A in $\mathrm{PO}(1, n-1)$ such that $Ax = te_1$. As A preserves Lorentzian inner products, we can replace x and y by Ax and Ay. Thus, we may assume, without loss of generality, that $x = x_1 e_1$.

Then we have

$$\begin{aligned}
\|x\|^2\|y\|^2 &= -x_1^2(-y_1^2+y_2^2+\cdots+y_n^2)\\
&= x_1^2y_1^2 - x_1^2(y_2^2+\cdots+y_n^2)\\
&\leq x_1^2y_1^2\\
&= (x\circ y)^2
\end{aligned}$$

with equality if and only if

$$y_2^2+\cdots+y_n^2=0,$$

that is, $y = y_1e_1$. Now since

$$x\circ y = -x_1y_1 < 0,$$

we have that

$$x\circ y \leq \|x\|\,\|y\|$$

with equality if and only if x and y are linearly dependent. □

The Time-Like Angle between Time-Like Vectors

Let x and y be positive (negative) time-like vectors in $\mathbb{R}^n$. By Theorem 3.1.6, there is a unique nonnegative real number $\eta(x,y)$ such that

$$x\circ y = \|x\|\,\|y\|\cosh\eta(x,y). \tag{3.1.6}$$

The *Lorentzian time-like angle* between x and y is defined to be $\eta(x,y)$. Note that $\eta(x,y)=0$ if and only if x and y are positive scalar multiples of each other.

Exercise 3.1

1. Let A be a real $n\times n$ matrix. Prove that the following are equivalent:

   (1) A is Lorentzian.

   (2) $\|Ax\| = \|x\|$ for all x in $\mathbb{R}^n$.

   (3) A preserves the quadratic form $q(x) = -x_1^2+x_2^2+\cdots+x_n^2$.

2. Prove algebraically that every Lorentzian $n\times n$ matrix is either positive or negative.

3. Show that $\mathrm{PO}(1,n-1)$ is naturally isomorphic to the *projective Lorentz group* $\mathrm{O}(1,n-1)/\{\pm I\}$.

4. The *Lorentzian complement* of a vector subspace V of $\mathbb{R}^n$ is defined to be the set
$$V^L = \{x\in\mathbb{R}^n : x\circ y = 0 \quad \text{for all } y \text{ in } V\}.$$
Show that $V^L = J(V)^\perp$.

5. Let V be a vector subspace of $\mathbb{R}^n$. Show that the following are equivalent:

(1) The subspace V is time-like.

(2) The subspace V^L is space-like.

(3) The subspace $V^\perp$ is space-like.

6. Let V be a vector subspace of $\mathbb{R}^n$. Show that V is light-like if and only if $V \cap C^{n-1}$ is a line passing through the origin.

7. Show that $\mathrm{PO}(1, n-1)$ acts transitively on the hyperboloid G^{n-1} in $\mathbb{R}^n$ defined by the equation $-x_1^2 + x_2^2 + \cdots + x_n^2 = 1$.

8. Show that $\mathrm{PO}(1, n-1)$ acts transitively on

   (1) the set of m-dimensional space-like subspaces of $\mathbb{R}^n$, and

   (2) the set of m-dimensional light-like subspaces of $\mathbb{R}^n$.

9. Let V be a 2-dimensional time-like subspace of $\mathbb{R}^n$. Show that $V \cap C^{n-1}$ is the union of two lines that intersect at the origin.

10. Let x and y be linearly independent space-like vectors in $\mathbb{R}^n$ and let V be the 2-dimensional vector subspace spanned by x and y. Show that

   (1) $|x \circ y| < \|x\|\,\|y\|$ if and only if V is space-like,

   (2) $|x \circ y| = \|x\|\,\|y\|$ if and only if V is light-like,

   (3) $|x \circ y| > \|x\|\,\|y\|$ if and only if V is time-like.

§3.2. Hyperbolic n-Space

Since a sphere of radius r in $\mathbb{R}^{n+1}$ is of constant curvature $1/r^2$ and hyperbolic n-space is of constant negative curvature, the duality between spherical and hyperbolic geometries suggests that hyperbolic n-space should be a sphere of imaginary radius. As imaginary distances are possible in Lorentzian $(n+1)$-space, we should take as our model for hyperbolic n-space the sphere of unit imaginary radius

$$F^n = \{x \in \mathbb{R}^{n+1} : \|x\|^2 = -1\}.$$

The only problem is that the set F^n is disconnected. The set F^n is a hyperboloid of two sheets defined by the equation

$$x_1^2 - (x_2^2 + \cdots + x_{n+1}^2) = 1.$$

The subset of all x in F^n such that $x_1 > 0$ (resp. $x_1 < 0$) is called the *positive* (resp. *negative*) *sheet* of F^n. We get around this problem by identifying antipodal vectors of F^n or equivalently by discarding the negative sheet of F^n. The *hyperboloid model* H^n of hyperbolic n-space is defined to be the positive sheet of F^n. See Figure 3.2.1. Note that hyperbolic geometry is actually dual to elliptic geometry rather than spherical geometry.

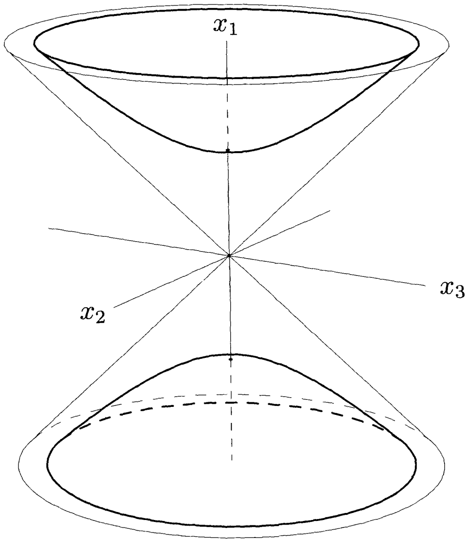


Figure 3.2.1. The hyperboloid F^2 inside C^2

Let x, y be vectors in H^n and let $\eta(x, y)$ be the Lorentzian time-like angle between x and y. The *hyperbolic distance* between x and y is defined to be the real number

$$d_H(x, y) = \eta(x, y). \tag{3.2.1}$$

As $x \circ y = \|x\|\ \|y\| \cosh \eta(x, y)$, we have the equation

$$\cosh d_H(x, y) = -x \circ y. \tag{3.2.2}$$

We shall prove that d_H is a metric on H^n, but first we need some preliminary results concerning cross products in $\mathbb{R}^3$.

Lorentzian Cross Products

Let x, y be vectors in $\mathbb{R}^3$ and let

$$J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.2.3}$$

The *Lorentzian cross product* of x and y is defined to be

$$x \otimes y = J(x \times y). \tag{3.2.4}$$

Observe that

$$\begin{aligned} x \circ (x \otimes y) &= x \circ J(x \times y) &= x \cdot (x \times y) &= 0, \\ y \circ (x \otimes y) &= y \circ J(x \times y) &= y \cdot (x \times y) &= 0. \end{aligned}$$

Hence $x \otimes y$ is Lorentz orthogonal to both x and y.

Lemma 1. *If x, y are vectors in $\mathbb{R}^3$, then $x \otimes y = J(y) \times J(x)$.*

Proof: As J is an orientation reversing orthogonal transformation, we have that

$$J(x \times y) = J(y) \times J(x). \qquad \square$$

Theorem 3.2.1. *If w, x, y, z are vectors in $\mathbb{R}^3$, then*

$$\begin{aligned}
&(1) \quad & x \otimes y &= -y \otimes x, \\
&(2) \quad & (x \otimes y) \circ z &= \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}, \\
&(3) \quad & x \otimes (y \otimes z) &= (x \circ y)z - (z \circ x)y, \\
&(4) \quad & (x \otimes y) \circ (z \otimes w) &= \begin{vmatrix} x \circ w & x \circ z \\ y \circ w & y \circ z \end{vmatrix}.
\end{aligned}$$

Proof: Observe that

$$\begin{aligned}
(1) \qquad x \otimes y &= J(y) \times J(x) \\
&= -J(x) \times J(y) \\
&= -y \otimes x.
\end{aligned}$$

$$\begin{aligned}
(2) \qquad (x \otimes y) \circ z &= J(x \times y) \cdot J(z) \\
&= (x \times y) \cdot z.
\end{aligned}$$

$$\begin{aligned}
(3) \qquad x \otimes (y \otimes z) &= J(y \otimes z) \times J(x) \\
&= (y \times z) \times J(x) \\
&= (y \cdot J(x))z - (z \cdot J(x))y \\
&= (x \circ y)z - (z \circ x)y.
\end{aligned}$$

$$\begin{aligned}
(4) \quad (x \otimes y) \circ (z \otimes w) &= J(x \times y) \circ J(z \times w) \\
&= (x \times y) \circ (z \times w) \\
&= (x \times y) \cdot J(z \times w) \\
&= (x \times y) \cdot (J(w) \times J(z)) \\
&= \begin{vmatrix} x \cdot J(w) & x \cdot J(z) \\ y \cdot J(w) & y \cdot J(z) \end{vmatrix} \\
&= \begin{vmatrix} x \circ w & x \circ z \\ y \circ w & y \circ z \end{vmatrix}. \qquad \square
\end{aligned}$$

Corollary 1. *If x, y are positive (negative) time-like vectors in $\mathbb{R}^3$, then $x \otimes y$ is space-like and $\|x \otimes y\| = -\|x\|\ \|y\| \sinh \eta(x, y)$.*

Proof: By Theorem 3.2.1(4), we have

$$\begin{aligned}\|x \otimes y\|^2 &= (x \circ y)^2 - \|x\|^2\|y\|^2 \\ &= \|x\|^2\|y\|^2 \cosh^2 \eta(x,y) - \|x\|^2\|y\|^2 \\ &= \|x\|^2\|y\|^2 \sinh^2 \eta(x,y).\end{aligned}$$

□

Corollary 2. *If x, y are space-like vectors in $\mathbb{R}^3$, then*

(1) $|x \circ y| < \|x\|\ \|y\|$ *if and only if $x \otimes y$ is time-like,*

(2) $|x \circ y| = \|x\|\ \|y\|$ *if and only if $x \otimes y$ is light-like,*

(3) $|x \circ y| > \|x\|\ \|y\|$ *if and only if $x \otimes y$ is space-like.*

Proof: By Theorem 3.2.1(4), we have $\|x \otimes y\|^2 = (x \circ y)^2 - \|x\|^2\|y\|^2$. □

Theorem 3.2.2. *The hyperbolic distance function d_H is a metric on H^n.*

Proof: The function d_H is obviously nonnegative and symmetric, and nondegenerate by Theorem 3.1.6. It remains only to prove the triangle inequality

$$d_H(x,z) \leq d_H(x,y) + d_H(y,z).$$

The positive Lorentz transformations of $\mathbb{R}^{n+1}$ act on H^n and obviously preserve hyperbolic distances. Thus, we are free to transform x, y, z by a positive Lorentz transformation. Now the three vectors x, y, z span a vector subspace of $\mathbb{R}^{n+1}$ of dimension at most three. By Theorem 3.1.5, we may assume that x, y, z are in the subspace of $\mathbb{R}^{n+1}$ spanned by e_1, e_2, e_3. In other words, we may assume that $n = 2$. By Corollary 1, we have

$$\|x \otimes y\| = \sinh \eta(x,y) \quad \text{and} \quad \|y \otimes z\| = \sinh \eta(y,z).$$

As y is Lorentz orthogonal to both $x \otimes y$ and $y \otimes z$, the vectors y and $(x \otimes y) \otimes (y \otimes z)$ are linearly dependent. Therefore, the latter is either zero or time-like. By Corollary 2, we have

$$|(x \otimes y) \circ (y \otimes z)| \leq \|x \otimes y\|\ \|y \otimes z\|.$$

Putting this all together, we have

$$\begin{aligned}&\cosh(\eta(x,y) + \eta(y,z)) \\ &\quad = \cosh \eta(x,y) \cosh \eta(y,z) + \sinh \eta(x,y) \sinh \eta(y,z) \\ &\quad = (x \circ y)(y \circ z) + \|x \otimes y\|\ \|y \otimes z\| \\ &\quad \geq (x \circ y)(y \circ z) + (x \otimes y) \circ (y \otimes z) \\ &\quad = (x \circ y)(y \circ z) + ((x \circ z)(y \circ y) - (x \circ y)(y \circ z)) \\ &\quad = -x \circ z \\ &\quad = \cosh \eta(x,z).\end{aligned}$$

Thus, we have that $\eta(x,z) \leq \eta(x,y) + \eta(y,z)$. □

The metric d_H on H^n is called the *hyperbolic metric.* The metric topology of H^n determined by d_H is the same as the metric topology determined by the Euclidean metric d_E on H^n defined by

$$d_E(x,y) = |x-y|. \tag{3.2.5}$$

The metric space consisting of H^n together with its hyperbolic metric d_H is called *hyperbolic n-space.* Henceforth H^n will denote hyperbolic n-space. An isometry from H^n to itself is called a *hyperbolic isometry.*

Theorem 3.2.3. *Every positive Lorentz transformation of $\mathbb{R}^{n+1}$ restricts to an isometry of H^n, and every isometry of H^n extends to a unique positive Lorentz transformation of $\mathbb{R}^{n+1}$.*

Proof: Clearly, a function $\phi : H^n \to H^n$ is an isometry if and only if it preserves Lorentzian inner products on H^n. Therefore, a positive Lorentz transformation of $\mathbb{R}^{n+1}$ restricts to an isometry of H^n.

Conversely, suppose that $\phi : H^n \to H^n$ is an isometry. Assume first that ϕ fixes e_1. Let $\phi_1, \ldots, \phi_{n+1}$ be the components of ϕ. Then

$$\begin{aligned} \phi_1(x) &= -\phi(x)\circ e_1 \\ &= -\phi(x)\circ\phi(e_1) \\ &= -x\circ e_1 \;=\; x_1. \end{aligned}$$

Thus $\phi(x) = (x_1, \phi_2(x), \ldots, \phi_{n+1}(x))$.

Let $p : H^n \to \mathbb{R}^n$ be defined by $p(x) = \overline{x}$, where $\overline{x} = (x_2, \ldots, x_{n+1})$. Then p is a bijection. Define $\overline{\phi} : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\overline{\phi}(u) = (\phi_2(p^{-1}(u)), \ldots, \phi_{n+1}(p^{-1}(u))).$$

Then $\overline{\phi}(\overline{x}) = \overline{\phi(x)}$ for all x in H^n. As $\phi(x)\circ\phi(y) = x\circ y$, we have

$$-x_1y_1 + \overline{\phi}(\overline{x})\cdot\overline{\phi}(\overline{y}) = -x_1y_1 + \overline{x}\cdot\overline{y}.$$

Therefore $\overline{\phi}(\overline{x})\cdot\overline{\phi}(\overline{y}) = \overline{x}\cdot\overline{y}$. Thus $\overline{\phi}$ is an orthogonal transformation. By Theorem 1.3.2, there is an orthogonal $n\times n$ matrix A such that $Au = \overline{\phi}(u)$ for all u in $\mathbb{R}^n$. Let $\hat{A}$ be the matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix}.$$

Then $\hat{A}$ is positive Lorentzian and $\hat{A}x = \phi(x)$ for all x in H^n.

Now assume that ϕ is an arbitrary isometry of H^n. By Theorem 3.1.5, there is a B in $\mathrm{PO}(1,n)$ such that $B\phi(e_1) = e_1$. As $B\phi$ extends to a positive Lorentz transformation of $\mathbb{R}^{n+1}$, the same is true of ϕ. Suppose that C and D are in $\mathrm{PO}(1,n)$ and extend ϕ. Then CD^{-1} fixes each point of H^n. As H^n is not contained in any proper vector subspace of $\mathbb{R}^{n+1}$, we have that CD^{-1} fixes all of $\mathbb{R}^{n+1}$. Therefore $C = D$. Thus ϕ extends to a unique positive Lorentz transformation of $\mathbb{R}^{n+1}$. □

Corollary 3. *The group of hyperbolic isometries* $\mathrm{I}(H^n)$ *is isomorphic to the positive Lorentz group* $\mathrm{PO}(1,n)$.

Hyperbolic Geodesics

Definition: A *hyperbolic line* of H^n is the intersection of H^n with a 2-dimensional time-like vector subspace of $\mathbb{R}^{n+1}$.

Let x and y be distinct points of H^n. Then x and y span a 2-dimensional time-like subspace $V(x,y)$ of $\mathbb{R}^{n+1}$, and so

$$L(x,y) = H^n \cap V(x,y)$$

is the unique hyperbolic line of H^n containing both x and y. Note that $L(x,y)$ is a branch of a hyperbola.

Definition: Three points x, y, z of H^n are *hyperbolically collinear* if and only if there is a hyperbolic line L of H^n containing x, y, z.

Lemma 2. *If* x, y, z *are points of* H^n *and*

$$\eta(x,y) + \eta(y,z) = \eta(x,z),$$

then x, y, z *are hyperbolically collinear.*

Proof: As x, y, z span a time-like vector subspace of $\mathbb{R}^{n+1}$ of dimension at most 3, we may assume that $n = 2$. From the proof of Theorem 3.2.2, we have that

$$(x \otimes y) \circ (y \otimes z) = \|x \otimes y\| \, \|y \otimes z\|.$$

By Corollary 2, we have that $(x \otimes y) \otimes (y \otimes z)$ is light-like. Now since

$$(x \otimes y) \otimes (y \otimes z) = -((x \otimes y) \circ z)y$$

and y is time-like, we have that $(x \otimes y) \circ z = 0$. Consequently x, y, z are linearly dependent by Theorem 3.2.1(2). Hence x, y, z lie on a 2-dimensional time-like vector subspace of $\mathbb{R}^3$ and so are hyperbolically collinear. □

Definition: Two vectors x, y in $\mathbb{R}^{n+1}$ are *Lorentz orthonormal* if and only if $\|x\|^2 = -1$ and $x \circ y = 0$ and $\|y\|^2 = 1$.

Theorem 3.2.4. *Let* $\alpha : [a,b] \to H^n$ *be a curve. Then the following are equivalent:*

(1) *The curve* α *is a geodesic arc.*

(2) *There are Lorentz orthonormal vectors* x, y *in* $\mathbb{R}^{n+1}$ *such that*

$$\alpha(t) = (\cosh(t-a))x + (\sinh(t-a))y.$$

(3) *The curve* α *satisfies the differential equation* $\alpha'' - \alpha = 0$.

Proof: Let A be a Lorentz transformation of $\mathbb{R}^{n+1}$. Then $(A\alpha)' = A\alpha'$. Consequently α satisfies (3) if and only if $A\alpha$ does. Hence, we are free to transform α by a Lorentz transformation. Suppose that α is a geodesic arc. Let t be in the interval $[a, b]$. Then we have

$$\begin{aligned} \eta(\alpha(a), \alpha(b)) &= b - a \\ &= (t-a) + (b-t) \\ &= \eta(\alpha(a), \alpha(t)) + \eta(\alpha(t), \alpha(b)). \end{aligned}$$

By Lemma 2, we have that $\alpha(a), \alpha(t), \alpha(b)$ are hyperbolically collinear. Consequently, the image of α is contained in a hyperbolic line L of H^n. Hence, we may assume that $n = 1$. By applying a Lorentz transformation of the form

$$\begin{pmatrix} \cosh s & \sinh s \\ \sinh s & \cosh s \end{pmatrix}$$

we can transform $\alpha(a)$ to e_1, and so we may assume that $\alpha(a) = e_1$. Then

$$\begin{aligned} e_1 \cdot \alpha(t) &= -\alpha(a) \circ \alpha(t) \\ &= \cosh \eta(\alpha(a), \alpha(t)) \\ &= \cosh(t-a). \end{aligned}$$

Therefore $e_2 \cdot \alpha(t) = \pm \sinh(t-a)$. As α is continuous, we have either

$$e_2 \cdot \alpha(t) = \sinh(t-a) \quad \text{for all } t$$

or

$$e_2 \cdot \alpha(t) = -\sinh(t-a) \quad \text{for all } t.$$

In the latter case, we can apply the reflection

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

and so we may assume that

$$\alpha(t) = (\cosh(t-a))e_1 + (\sinh(t-a))e_2.$$

Thus (1) implies (2).

Next, suppose there are Lorentz orthonormal vectors x, y in $\mathbb{R}^{n+1}$ such that

$$\alpha(t) = (\cosh(t-a))x + (\sinh(t-a))y.$$

Let s and t be such that $a \leq s \leq t \leq b$. Then we have

$$\begin{aligned} \cosh \eta(\alpha(s), \alpha(t)) &= -\alpha(s) \circ \alpha(t) \\ &= \cosh(s-a)\cosh(t-a) - \sinh(s-a)\sinh(t-a) \\ &= \cosh(t-s). \end{aligned}$$

Therefore $\eta(\alpha(s), \alpha(t)) = t - s$. Thus α is a geodesic arc. Hence (2) implies (1). Clearly (2) implies (3). Suppose that (3) holds. Then

$$\alpha(t) = \cosh(t-a)\alpha(a) + \sinh(t-a)\alpha'(a).$$

On differentiating the equation $\alpha(t)\circ\alpha(t) = -1$, we see that $\alpha(t)\circ\alpha'(t) = 0$. In particular, $\alpha(a)\circ\alpha'(a) = 0$. Observe that

$$\|\alpha(t)\|^2 = -\cosh^2(t-a) + \sinh^2(t-a)\|\alpha'(a)\|^2.$$

As $\|\alpha(t)\|^2 = -1$, we have that $\|\alpha'(a)\|^2 = 1$. Therefore $\alpha(a), \alpha'(a)$ are Lorentz orthonormal. Thus (3) implies (2). □

Theorem 3.2.5. *A function $\lambda : \mathbb{R} \to H^n$ is a geodesic line if and only if there are Lorentz orthonormal vectors x, y in $\mathbb{R}^{n+1}$ such that*

$$\lambda(t) = (\cosh t)x + (\sinh t)y.$$

Proof: Suppose there are Lorentz orthonormal vectors x, y in $\mathbb{R}^{n+1}$ such that $\lambda(t) = (\cosh t)x + (\sinh t)y$. Then λ satisfies the differential equation $\lambda'' - \lambda = 0$. Hence, the restriction of λ to any interval $[a, b]$, with $a < b$, is a geodesic arc by Theorem 3.2.4. Thus λ is a geodesic line.

Conversely, suppose that λ is a geodesic line. By Theorem 3.2.4, the function λ satisfies the differential equation $\lambda'' - \lambda = 0$. Consequently

$$\lambda(t) = (\cosh t)\lambda(0) + (\sinh t)\lambda'(0).$$

The same argument as in the proof of Theorem 3.2.4 shows that $\lambda(0), \lambda'(0)$ are Lorentz orthonormal. □

Corollary 4. *The geodesics of H^n are its hyperbolic lines.*

Proof: By Theorem 3.2.5, every geodesic of H^n is a hyperbolic line. Conversely, let L be a hyperbolic line of H^n. By Theorem 3.1.5, we may assume that $n = 1$. Then $L = H^1$. Define $\lambda : \mathbb{R} \to H^1$ by

$$\lambda(t) = (\cosh t)e_1 + (\sinh t)e_2.$$

Then λ is a geodesic line mapping onto H^1. Thus L is a geodesic. □

Hyperplanes

We now consider the geometry of hyperplanes of H^n.

Definition: A *hyperbolic m-plane* of H^n is the intersection of H^n with an $(m+1)$-dimensional time-like vector subspace of $\mathbb{R}^{n+1}$.

Note that a hyperbolic 1-plane of H^n is the same as a hyperbolic line of H^n. A hyperbolic $(n-1)$-plane of H^n is called a *hyperplane* of H^n.

Let x be a space-like vector in $\mathbb{R}^{n+1}$. Then the Lorentzian complement of the vector subspace $\langle x\rangle$ spanned by x is an n-dimensional time-like vector subspace of $\mathbb{R}^{n+1}$. Hence $P = \langle x\rangle^L \cap H^n$ is a hyperplane of H^n. The hyperplane P is called the hyperplane of H^n Lorentz orthogonal to x.

Theorem 3.2.6. *Let x and y be linearly independent space-like vectors in $\mathbb{R}^{n+1}$. Then the following are equivalent:*

(1) *The vectors x and y satisfy the equation $|x \circ y| < \|x\| \; \|y\|$.*

(2) *The vector subspace V spanned by x and y is space-like.*

(3) *The hyperplanes P and Q of H^n Lorentz orthogonal to x and y, respectively, intersect.*

Proof: Assume that (1) holds. Then for nonzero real numbers s and t, we have that

$$\begin{aligned} \|sx + ty\|^2 &= \|sx\|^2 + 2st(x \circ y) + \|ty\|^2 \\ &> \|sx\|^2 - 2|st| \; \|x\| \; \|y\| + \|ty\|^2 \\ &= (\|sx\| - \|ty\|)^2 \\ &\geq 0. \end{aligned}$$

Thus V is space-like.

Conversely, if (2) holds, then the Lorentzian inner product on V is positive definite. Hence, Cauchy's inequality holds in V, and so (1) holds. Thus (1) and (2) are equivalent. Now (2) and (3) are equivalent, since $V^L = \langle x \rangle^L \cap \langle y \rangle^L$. □

The Space-Like Angle between Space-Like Vectors

Let x and y be space-like vectors in $\mathbb{R}^{n+1}$ that span a space-like vector subspace. Then by Theorem 3.2.6, we have that

$$|x \circ y| \leq \|x\| \; \|y\|$$

with equality if and only if x and y are linearly dependent. Hence, there is a unique real number $\eta(x, y)$ between 0 and π such that

$$x \circ y = \|x\| \; \|y\| \cos \eta(x, y). \tag{3.2.6}$$

The *Lorentzian space-like angle* between x and y is defined to be $\eta(x, y)$. Note that $\eta(x, y) = 0$ if and only if x and y are positive scalar multiples of each other, $\eta(x, y) = \pi/2$ if and only if x and y are Lorentz orthogonal, and $\eta(x, y) = \pi$ if and only if x and y are negative scalar multiples of each other.

Let $\lambda, \mu = \mathbb{R} \to H^n$ be geodesic lines such $\lambda(0) = \mu(0)$. Then $\lambda'(0)$ and $\mu'(0)$ span a space-like vector subspace of $\mathbb{R}^{n+1}$. The *hyperbolic angle* between λ and μ is defined to be the Lorentzian space-like angle between $\lambda'(0)$ and $\mu'(0)$.

Let P be a hyperplane of H^n and let $\lambda : \mathbb{R} \to H^n$ be a geodesic line such that $\lambda(0)$ is in P. Then the hyperbolic line $L = \lambda(\mathbb{R})$ is said to be *Lorentz orthogonal* to P if and only if P is the hyperplane of H^n Lorentz orthogonal to $\lambda'(0)$.

Theorem 3.2.7. *Let x and y be linearly independent space-like vectors in $\mathbb{R}^{n+1}$. Then the following are equivalent:*

(1) *The vectors x and y satisfy the inequality $|x \circ y| > \|x\|\,\|y\|$.*

(2) *The vector subspace V spanned by x and y is time-like.*

(3) *The hyperplanes P and Q of H^n Lorentz orthogonal to x and y, respectively, are disjoint and have a common Lorentz orthogonal hyperbolic line.*

Proof: Except for scalar multiples of x, every element of V is a scalar multiple of an element of the form $tx + y$ for some real number t. Observe that the expression

$$\|tx + y\|^2 = t^2\|x\|^2 + 2t(x \circ y) + \|y\|^2$$

is a quadratic polynomial in t. This polynomial takes on negative values if and only if its discriminant

$$4(x \circ y)^2 - 4\|x\|^2\|y\|^2$$

is positive. Thus (1) and (2) are equivalent.

Suppose that V is time-like. Then V^L is space-like. Now since $V^L = \langle x\rangle^L \cap \langle y\rangle^L$, we have that P and Q are disjoint. Observe that $N = V \cap H^n$ is a hyperbolic line and $V \cap \langle x\rangle^L$ is a 1-dimensional subspace of $\mathbb{R}^{n+1}$. Moreover, the equation

$$(tx + y) \circ x = 0$$

has the unique solution

$$t = -x \circ y/\|x\|^2.$$

Furthermore

$$\|tx + y\|^2 = -\frac{(x \circ y)^2}{\|x\|^2} + \|y\|^2 < 0.$$

Hence $V \cap \langle x\rangle^L$ is time-like. Thus $N \cap P$ is the single point

$$u = \frac{-(x \circ y)(x/\|x\|) + \|x\|y}{\pm\sqrt{(x \circ y)^2 - \|x\|^2\|y\|^2}},$$

where the plus or minus sign is choosen so that u is positive time-like. Likewise $N \cap Q$ is a single point v. Let $\lambda : \mathbb{R} \to H^n$ be a geodesic line such that $\lambda(0) = u$ and $\lambda(\mathbb{R}) = N$. As $\lambda'(0)$ and x are both Lorentz orthogonal to u in V, we have that $\lambda'(0)$ is a scalar multiple of x. Thus N is Lorentz orthogonal to P. Likewise N is Lorentz orthogonal to Q.

Conversely, assume that (3) holds. Let N be the common Lorentz orthogonal hyperbolic line to P and Q. Then there is a 2-dimensional time-like vector subspace W of $\mathbb{R}^{n+1}$ such that $N = W \cap H^n$. As N is Lorentz orthogonal to P, we have that x is in W. Likewise y is in W. Hence $V = W$, and so V is time-like. □

Remark: The proof of Theorem 3.2.7 shows that if P and Q are disjoint hyperplanes of H^n, with a common Lorentz orthogonal hyperbolic line N, then N is unique; moreover, if x, y are space-like vectors in $\mathbb{R}^{n+1}$ Lorentz orthogonal to P, Q, respectively, then x and y are tangent vectors of N.

The Time-Like Angle between Space-Like Vectors

Let x and y be space-like vectors in $\mathbb{R}^{n+1}$ that span a time-like vector subspace. Then by Theorem 3.2.7, we have that $|x \circ y| > \|x\| \, \|y\|$. Hence, there is a unique positive real number $\eta(x, y)$ such that

$$|x \circ y| = \|x\| \, \|y\| \cosh \eta(x, y). \tag{3.2.7}$$

The *Lorentzian time-like angle* between x and y is defined to be $\eta(x, y)$. We now give a geometric interpretation of $\eta(x, y)$.

Theorem 3.2.8. *Let x and y be space-like vectors in $\mathbb{R}^{n+1}$ that span a time-like vector subspace, and let P, Q be the hyperplanes of H^n Lorentz orthogonal to x, y, respectively. Then $\eta(x, y)$ is the hyperbolic distance from P to Q measured along the hyperbolic line N Lorentz orthogonal to P and Q. Moreover $x \circ y < 0$ if and only if x and y are oppositely oriented tangent vectors of N.*

Proof: From the proof of Theorem 3.2.7, we have that $P \cap N$ is the point

$$u = \frac{-(x \circ y)(x/\|x\|) + \|x\| y}{\pm\sqrt{(x \circ y)^2 - \|x\|^2 \|y\|^2}}$$

and $Q \cap N$ is the point

$$v = \frac{\|y\| x - (x \circ y)(y/\|y\|)}{\pm\sqrt{(x \circ y)^2 - \|x\|^2 \|y\|^2}}.$$

Now

$$\begin{aligned}
\cosh d_H(u, v) &= -u \circ v \\
&= \frac{-((x \circ y)^3/\|x\| \, \|y\|) + (x \circ y)\|x\| \, \|y\|}{\pm((x \circ y)^2 - \|x\|^2 \|y\|^2)} \\
&= \frac{-((x \circ y)^3 + (x \circ y)\|x\|^2 \|y\|^2)/\|x\| \, \|y\|}{\pm((x \circ y)^2 - \|x\|^2 \|y\|^2)} \\
&= \frac{-(x \circ y)}{\pm\|x\| \, \|y\|} \\
&= \frac{|x \circ y|}{\|x\| \, \|y\|} \\
&= \cosh \eta(x, y).
\end{aligned}$$

Moreover, the calculation of $-u \circ v$ shows that u and v have the same sign if and only if $x \circ y < 0$. Observe that u and v are in the 2-dimensional time-like subspace V spanned by x and y. Evidently u and v are in the quadrant of V between x and y or $-x$ and $-y$ if and only if the coefficient $-x \circ y$ of u and v is positive. Thus x and y are oppositely oriented tangent vectors of N if and only if $x \circ y < 0$. □

Let x and y be space-like vectors in $\mathbb{R}^{n+1}$ and let P, Q be the hyperplanes of H^n Lorentz orthogonal to x, y, respectively. Then P and Q are said to *meet at infinity* if and only if $\langle x\rangle^L \cap \langle y\rangle^L$ is light-like. If P and Q meet at infinity, then P and Q are disjoint, but when viewed from the origin, they appear to meet at the positive ideal endpoint of the 1-dimensional light-like subspace of $\langle x\rangle^L \cap \langle y\rangle^L$.

Theorem 3.2.9. *Let x and y be linearly independent space-like vectors in $\mathbb{R}^{n+1}$. Then the following are equivalent:*

1. *The vectors x and y satisfy the equation $|x \circ y| = \|x\|\ \|y\|$.*
2. *The vector subspace V spanned by x and y is light-like.*
3. *The hyperplanes P and Q of H^n Lorentz orthogonal to x and y, respectively, meet at infinity.*

Proof: (1) and (2) are equivalent by Theorems 3.2.6 and 3.2.7, and (2) and (3) are equivalent, since $V^L = \langle x\rangle^L \cap \langle y\rangle^L$. □

Theorem 3.2.10. *Let x and y be linearly independent space-like vectors in $\mathbb{R}^{n+1}$ such that the vector subspace V spanned by x and y is light-like. Then $x \circ y < 0$ if and only if x and y are on opposite sides of the 1-dimensional light-like subspace of V.*

Proof: The equation $\|tx + y\| = 0$ is equivalent to the quadratic equation

$$t^2\|x\|^2 + 2(x \circ y)t + \|y\|^2 = 0,$$

which by Theorem 3.2.9 has the unique solution

$$t = -(x \circ y)/\|x\|^2.$$

Observe that the light-like vector

$$-(x \circ y)(x/\|x\|^2) + y$$

is in the quadrant of V between x and y if and only if $x \circ y < 0$. Hence x and y are on opposite sides of the 1-dimensional light-like subspace of V if and only if $x \circ y < 0$. □

Theorem 3.2.11. *Let y be a point of H^n and let P be a hyperplane of H^n. Then there is a unique hyperbolic line N of H^n passing through y and Lorentz orthogonal to P.*

Proof: Let x be a unit space-like vector Lorentz orthogonal to P, and let V be the subspace spanned by x and y. Then $N = V \cap H^n$ is a hyperbolic line passing through y. Now the equation

$$(tx + y) \circ x = 0$$

has the solution $t = -x \circ y$. Hence

$$w = \frac{-(x \circ y)x + y}{\pm\sqrt{(x \circ y)^2 + 1}}$$

is a point of $P \cap N$. Let $\lambda : \mathbb{R} \to H^n$ be a geodesic line such that $\lambda(\mathbb{R}) = N$ and $\lambda(0) = w$. As w, x are Lorentz orthonormal vectors, we have

$$\lambda(t) = (\cosh t)w \pm (\sinh t)x.$$

Hence $\lambda'(0) = \pm x$. Thus N is Lorentz orthogonal to P.

Suppose that N is a hyperbolic line passing through y and Lorentz orthogonal to P. Let $\lambda : \mathbb{R} \to H^n$ be a geodesic line such that $\lambda(\mathbb{R}) = N$ and $\lambda(0)$ is in P. Then $\lambda'(0)$ is Lorentz orthogonal to P. Hence $\lambda'(0) = \pm x$. Let W be the 2-dimensional time-like subspace such that $N = W \cap H^n$. As x and y are in W, we have that $W = V$. Thus N is unique. □

The Angle between Space-Like and Time-Like Vectors

Let x be a space-like vector and y a positive time-like vector in $\mathbb{R}^{n+1}$. Then there is a unique nonnegative real number $\eta(x, y)$ such that

$$|x \circ y| = \|x\| \; |||y||| \sinh \eta(x, y). \tag{3.2.8}$$

The *Lorentzian time-like angle* between x and y is defined to be $\eta(x, y)$. We now give a geometric interpretation of $\eta(x, y)$.

Theorem 3.2.12. *Let x be a space-like vector and y a positive time-like vector in $\mathbb{R}^{n+1}$, and let P be the hyperplane of H^n Lorentz orthogonal to x. Then $\eta(x, y)$ is the hyperbolic distance from $y/|||y|||$ to P measured along the hyperbolic line N passing through $y/|||y|||$ Lorentz orthogonal to P. Moreover $x \circ y < 0$ if and only if x and y are on opposite sides of the hyperplane of $\mathbb{R}^{n+1}$ spanned by P.*

Proof: As in the proof of Theorem 3.2.8, we have that $P \cap N$ is the point

$$u = \frac{-(x \circ y)(x/\|x\|) + \|x\|y}{\pm\sqrt{(x \circ y)^2 - \|x\|^2\|y\|^2}}.$$

Let $v = y/|||y|||$. Then

$$\begin{aligned} \cosh d_H(u, v) &= -u \circ v \\ &= \frac{\sqrt{(x \circ y)^2 - \|x\|^2\|y\|^2}}{\|x\| \; |||y|||} \\ &= \cosh \eta(x, y). \end{aligned}$$

Moreover, the calculation of $-u \circ v$ shows that u has the plus sign. Observe that u is in the 2-dimensional time-like subspace V spanned by x and y. Evidently u is in the quadrant of V between x and y if and only if the coefficient $-x \circ y$ of u is positive. Thus x and y are on opposite sides of the hyperplane of $\mathbb{R}^{n+1}$ spanned by P if and only if $x \circ y < 0$. □

Exercise 3.2

1. Show that the metric topology of H^n determined by the hyperbolic metric is the same as the metric topology of H^n determined by the Euclidian metric.
2. Prove that H^n is homeomorphic to E^n.
3. Show that every matrix in $\mathrm{PSO}(1,1)$ is of the form
$$\begin{pmatrix} \cosh s & \sinh s \\ \sinh s & \cosh s \end{pmatrix}.$$
4. Let A be in $\mathrm{PO}(1,2)$. Prove that $A(x \otimes y) = (\det A)(Ax \otimes Ay)$.
5. Show that every hyperbolic line of H^n is the branch of a hyperbola whose asymptotes are 1-dimensional time-like vector subspaces.
6. Prove that H^n is geodesically complete.
7. Two hyperbolic lines of H^n are said to be *parallel* if and only if there is a hyperbolic 2-plane containing both lines and the lines are disjoint. Show that for each point x of H^n outside a hyperbolic line L, there are infinitely many hyperbolic lines passing through x parallel to L.
8. Prove that a nonempty subset X of H^n is totally geodesic if and only if X is a hyperbolic m-plane of H^n for some m.
9. Prove that H^1 is isometric to E^1, but H^n is not isometric to E^n for $n > 1$.
10. Let $u_0, \ldots, u_n$ be linearly independent vectors in H^n, let $v_0, \ldots, v_n$ be linearly independent vectors in H^n, and suppose that $\eta(u_i, u_j) = \eta(v_i, v_j)$ for all i, j. Prove that there is a unique hyperbolic isometry ϕ of H^n such that $\phi(u_i) = v_i$ for each $i = 0, \ldots, n$.
11. A *tangent vector* to H^n at a point x of H^n is defined to be the derivative at 0 of a differentiable curve $\gamma : [-b, b] \to H^n$ such that $\gamma(0) = x$. Let $T_x = T_x(H^n)$ be the set of all tangent vectors to H^n at x. Show that
$$T_x = \{y \in \mathbb{R}^{n+1} : x \circ y = 0\}.$$
Conclude that T_x is an n-dimensional space-like vector subspace of $\mathbb{R}^{n+1}$. The vector space T_x is called the *tangent space* of H^n at x.
12. A *coordinate frame* of H^n is an n-tuple of functions $(\lambda_1, \ldots, \lambda_n)$ such that
    (1) the function $\lambda_i : \mathbb{R} \to H^n$ is a geodesic line for each $i = 1, \ldots, n$;
    (2) there is a point x of H^n such that $\lambda_i(0) = x$ for all i; and
    (3) the set $\{\lambda_1'(0), \ldots, \lambda_n'(0)\}$ is a Lorentz orthonormal basis of $T_x(H^n)$.

    Show that the action of $\mathrm{I}(H^n)$ on the set of coordinate frames of H^n, given by $\phi(\lambda_1, \ldots, \lambda_n) = (\phi\lambda_1, \ldots, \phi\lambda_n)$, is transitive.

§3.3. Hyperbolic Arc Length

In this section, we compare the hyperbolic length of a curve γ in H^n with its Lorentzian length in $\mathbb{R}^{n+1}$ and show that they are the same. In the process, we find the element of hyperbolic arc length of H^n.

Let x, y be points of H^n. By Theorem 3.1.6, we have

$$\begin{aligned} \|x-y\|^2 &= \|x\|^2 - 2x \circ y + \|y\|^2 \\ &\geq -2 - 2\|x\|\,\|y\| \;=\; 0 \end{aligned}$$

with equality if and only if $x = y$. Hence, the Lorentzian distance function

$$d_L(x,y) = \|x-y\|$$

satisfies the first three axioms for a metric on H^n. Unfortunately, d_L does not satisfy the triangle inequality. Nevertheless, we can still use d_L to define the length of a curve in H^n.

Let $\gamma : [a,b] \to H^n$ be a curve and let $P = \{t_0, \ldots, t_m\}$ be a partition of $[a,b]$. The *Lorentzian P-inscribed length* of γ is defined to be

$$\ell_L(\gamma, P) = \sum_{i=1}^{m} \|\gamma(t_i) - \gamma(t_{i-1})\|.$$

The curve γ is said to be *Lorentz rectifiable* if and only if there is a real number $\ell(\gamma)$ such that for each $\epsilon > 0$ there is a partition P of $[a,b]$ such that if $Q \leq P$, then

$$\big|\ell(\gamma) - \ell_L(\gamma, Q)\big| < \epsilon.$$

If $\ell(\gamma)$ exists, then it is unique, since if P and Q are partitions of $[a,b]$, then there is a partition R of $[a,b]$ such that $R \leq P$, Q.

The *Lorentzian length* $\|\gamma\|$ of γ is defined to be $\ell(\gamma)$ if γ is Lorentz rectifiable or ∞ otherwise.

Theorem 3.3.1. *Let $\gamma : [a,b] \to H^n$ be a curve. Then γ is rectifiable in H^n if and only if γ is Lorentz rectifiable; moreover, the hyperbolic length of γ is the same as the Lorentzian length of γ.*

Proof: Let x, y be in H^n. Then we have

$$\begin{aligned} \|x-y\|^2 &= \|x\|^2 - 2x \circ y + \|y\|^2 \\ &= 2(\cosh \eta(x,y) - 1). \end{aligned}$$

Now since

$$\cosh \eta \geq 1 + (\eta^2/2),$$

we have that

$$\|x-y\| \geq \eta(x,y).$$

Suppose that γ is Lorentz rectifiable. Then there is a partition P of $[a,b]$ such that if $Q \leq P$, then

$$\big|\, \|\gamma\| - \ell_L(\gamma, Q)\big| < 1.$$

Hence, for all $Q \leq P$, we have

$$\ell_H(\gamma, Q) \leq \ell_L(\gamma, Q) \leq \|\gamma\| + 1.$$

Thus γ is rectifiable. By Taylor's theorem, we have

$$\cosh \eta \leq 1 + \frac{\eta^2}{2} + \frac{\eta^4}{24} \cosh \eta.$$

Hence, if $\cosh \eta(x,y) \leq 12$, we have

$$\|x - y\| \leq \eta(x,y)\sqrt{1 + \eta^2(x,y)}.$$

Now suppose that γ is rectifiable and $\epsilon > 0$. Then there is a partition P of $[a,b]$ such that

$$|\gamma|_H - \ell_H(\gamma, P) < \epsilon.$$

Let $\delta > 0$ and set

$$\mu(\gamma, \delta) = \sup\{\eta(\gamma(s), \gamma(t)) : |s - t| \leq \delta\}.$$

As γ is uniformly continuous, $\mu(\gamma, \delta)$ goes to zero with δ. Hence, there is a $\delta > 0$ such that $\cosh \mu(\gamma, \delta) \leq 12$ and

$$|\gamma|_H \sqrt{1 + \mu^2(\gamma, \delta)} < |\gamma|_H + \epsilon.$$

Now we may assume that $|P| \leq \delta$. Then for all $Q \leq P$, we have

$$\begin{aligned} |\gamma|_H - \epsilon &< \ell_H(\gamma, Q) \\ &\leq \ell_L(\gamma, Q) \\ &\leq \ell_H(\gamma, Q)\sqrt{1 + \mu^2} \\ &\leq |\gamma|_H \sqrt{1 + \mu^2} \\ &< |\gamma|_H + \epsilon. \end{aligned}$$

Hence, we have

$$\big|\, |\gamma|_H - \ell_L(\gamma, Q)\big| < \epsilon \ \text{ for all } \ Q \leq P.$$

Thus γ is Lorentz rectifiable and $\|\gamma\| = |\gamma|_H$. □

Let $\gamma : [a,b] \to H^n$ be a differentiable curve. As $\gamma(t) \circ \gamma(t) = -1$, we have $\gamma(t) \circ \gamma'(t) = 0$. Hence $\gamma'(t)$ is space-like for all t by Theorem 3.1.4.

Theorem 3.3.2. *Let $\gamma : [a,b] \to H^n$ be a C^1 curve. Then γ is rectifiable and the hyperbolic length of γ is given by the formula*

$$\|\gamma\| = \int_a^b \|\gamma'(t)\| dt.$$

Proof: Define $f : [a,b]^{n+1} \to \mathbb{R}$ by the formula

$$f(x) = |-\gamma_1'(x_1)^2 + \gamma_2'(x_2)^2 + \cdots + \gamma_{n+1}'(x_{n+1})^2|^{\frac{1}{2}}.$$

Then f is continuous. Observe that the set

$$\{|f(x)-f(y)| : x,y \in [a,b]^{n+1}\}$$

is bounded, since $[a,b]^{n+1}$ is compact. Let $\delta > 0$ and set

$$\mu(f,\delta) = \sup\big\{|f(x)-f(y)| : |x_i - y_i| \leq \delta \ \text{ for } i = 1,\ldots,n+1\big\}.$$

Let $P = \{t_0,\ldots,t_m\}$ be a partition of $[a,b]$ such that $|P| \leq \delta$. By the mean value theorem, there is a real number s_{ij} between t_{j-1} and t_j such that

$$\gamma_i(t_j) - \gamma_i(t_{j-1}) = \gamma_i'(s_{ij})(t_j - t_{j-1}).$$

Then we have

$$\|\gamma(t_j) - \gamma(t_{j-1})\| = f(s_j)(t_j - t_{j-1}),$$

where $s_j = (s_{1,j},\ldots,s_{n+1,j})$. Hence

$$\begin{aligned} &\big|\, \|\gamma(t_j) - \gamma(t_{j-1})\| - \|\gamma'(t_j)\|(t_j - t_{j-1})\big| \\ &= \big|f(s_j) - \|\gamma'(t_j)\|\,\big|(t_j - t_{j-1}) \\ &\leq \mu(f,\delta)(t_j - t_{j-1}). \end{aligned}$$

Set

$$S(\gamma,P) = \sum_{j=1}^{m} \|\gamma'(t_j)\|(t_j - t_{j-1}).$$

Then we have

$$\begin{aligned} &\big|\ell_L(\gamma,P) - S(\gamma,P)\big| \\ &\leq \sum_{j=1}^{m} \big|\, \|\gamma(t_j) - \gamma(t_{j-1})\| - \|\gamma'(t_j)\|(t_j - t_{j-1})\big| \\ &\leq \sum_{j=1}^{m} \mu(f,\delta)(t_j - t_{j-1}) \ = \ \mu(f,\delta)(b-a). \end{aligned}$$

Next, observe that

$$\begin{aligned} &\left|\int_a^b \|\gamma'(t)\|dt - S(\gamma,P)\right| \\ &= \left|\sum_{j=1}^{m} \int_{t_{j-1}}^{t_j} (\|\gamma'(t)\| - \|\gamma'(t_j)\|)dt\right| \\ &\leq \sum_{j=1}^{m} \left|\int_{t_{j-1}}^{t_j} (\|\gamma'(t)\| - \|\gamma'(t_j)\|)dt\right| \\ &\leq \sum_{j=1}^{m} \int_{t_{j-1}}^{t_j} \big|\, \|\gamma'(t)\| - \|\gamma'(t_j)\|\, \big|dt \\ &\leq \sum_{j=1}^{m} \int_{t_{j-1}}^{t_j} \mu(f,\delta)dt \ = \ \mu(f,\delta)(b-a). \end{aligned}$$

Thus

$$\begin{aligned}
&\left|\int_a^b \|\gamma'(t)\|dt - \ell_L(\gamma, P)\right| \\
&\leq \left|\int_a^b \|\gamma'(t)\|dt - S(\gamma, P)\right| + |S(\gamma, P) - \ell_L(\gamma, P)| \\
&\leq 2\mu(f, \delta)(b - a).
\end{aligned}$$

Now $f : [a, b]^{n+1} \to \mathbb{R}$ is uniformly continuous, since $[a, b]^{n+1}$ is compact. Therefore $\mu(f, \delta)$ goes to zero with δ. Hence

$$\lim_{|P|\to 0} \ell_L(\gamma, P) = \int_a^b \|\gamma'(t)\|dt. \qquad \square$$

Let $\gamma : [a, b] \to H^n$ be a curve. Set $dx = (dx_1, \ldots, dx_{n+1})$ and

$$\|dx\| = (-dx_1^2 + dx_2^2 + \cdots + dx_{n+1}^2)^{\frac{1}{2}}.$$

Then by definition, we have

$$\int_\gamma \|dx\| = \|\gamma\|.$$

Moreover, if γ is a C^1 curve, then by Theorem 3.3.2, we have

$$\int_\gamma \|dx\| = \int_a^b \|\gamma'(t)\|dt.$$

The differential $\|dx\|$ is called the *element of hyperbolic arc length* of H^n.

Exercise 3.3

1. Show that the Lorentzian distance function d_L is not a metric on H^n.
2. Let $\gamma : [a, b] \to H^n$ be a curve that is rectifiable in E^{n+1}. Prove that γ is rectifiable in H^n.
3. Let $\gamma : [a, b] \to H^n$ be a rectifiable curve. Prove that γ is rectifiable in E^{n+1}.

§3.4. Hyperbolic Volume

Let x be a point of H^n, with $n > 1$, such that x_n and x_{n+1} are not both zero. The *hyperbolic coordinates* $(\eta_1, \ldots, \eta_n)$ of x are defined as follows:

$$\eta_i = \eta(e_i, x_ie_i + x_{i+1}e_{i+1} + \cdots + x_{n+1}e_{n+1}) \quad \text{if } i < n,$$

η_n is the polar angle from e_n to $x_ne_n + x_{n+1}e_{n+1}$.

The hyperbolic coordinates of x satisfy the system of equations

$$\begin{aligned} x_1 &= \cosh\eta_1, \\ x_2 &= \sinh\eta_1\cos\eta_2, \\ &\vdots \\ x_n &= \sinh\eta_1\sin\eta_2\cdots\sin\eta_{n-1}\cos\eta_n, \\ x_{n+1} &= \sinh\eta_1\sin\eta_2\cdots\sin\eta_{n-1}\sin\eta_n. \end{aligned} \tag{3.4.1}$$

The *hyperbolic coordinate parameterization* of H^n is the map

$$h : \mathbb{R}\times[0,\pi]^{n-2}\times[0,2\pi]\to H^n$$

defined by

$$h(\eta_1,\ldots,\eta_n) = (x_1,\ldots,x_{n+1}),$$

where x_i is expressed in terms of the hyperbolic coordinates $\eta_1,\ldots,\eta_n$ by the system of Equations (3.4.1). The map h is surjective, and injective on the open set $\mathbb{R}\times(0,\pi)^{n-2}\times(0,2\pi)$. A straightforward calculation shows that

(1) $$\left\|\frac{\partial h}{\partial\eta_1}\right\| = 1, \tag{3.4.2}$$

(2) $$\left\|\frac{\partial h}{\partial\eta_i}\right\| = \sinh\eta_1\sin\eta_2\cdots\sin\eta_{i-1} \quad \text{for } i>1, \tag{3.4.3}$$

(3) $$\frac{\partial h}{\partial\eta_i}\circ\frac{\partial h}{\partial\eta_j} = 0 \quad \text{for } i<j. \tag{3.4.4}$$

A subset X of H^n is said to be *measurable* in H^n if and only if $h^{-1}(X)$ is measurable in $\mathbb{R}^n$. In particular, all the Borel subsets of H^n are measurable in H^n. If X is measurable in H^n, then the *hyperbolic volume* of X is defined by the formula

$$\mathrm{Vol}(X) = \int_{h^{-1}(X)} \sinh^{n-1}\eta_1\sin^{n-2}\eta_2\cdots\sin\eta_{n-1}d\eta_1\cdots d\eta_n. \tag{3.4.5}$$

The motivation for Formula 3.4.5 is as follows: Subdivide $\mathbb{R}^n$ into a rectangular grid pattern parallel to the coordinate axes. Each grid rectangular solid of volume $\Delta\eta_1\cdots\Delta\eta_n$ that meets $h^{-1}(X)$ corresponds under h to a region in H^n that meets X. This region is approximated by the Lorentzian rectangular solid spanned by the vectors $\frac{\partial h}{\partial\eta_1}\Delta\eta_1,\ldots,\frac{\partial h}{\partial\eta_n}\Delta\eta_n$. Its Lorentzian volume is

$$\left\|\frac{\partial h}{\partial\eta_1}\Delta\eta_1\right\|\cdots\left\|\frac{\partial h}{\partial\eta_n}\Delta\eta_n\right\| = \sinh^{n-1}\eta_1\sin^{n-2}\eta_2\cdots\sin\eta_{n-1}\Delta\eta_1\cdots\Delta\eta_n.$$

As the mesh of the subdivision goes to zero, the sum of the volumes of the approximating rectangular solids approaches the volume of X as a limit.

Let X be a measurable subset of H^n and let ϕ be a positive Lorentz transformation of $\mathbb{R}^{n+1}$. It is a basic fact of differential geometry that

$\phi(X)$ is also measurable in H^n and the hyperbolic volume of $\phi(X)$ can be measured with respect to the new parameterization ϕh of H^n. As ϕ maps the Lorentzian rectangular solid spanned by the vectors

$$\frac{\partial h}{\partial \eta_1}\Delta\eta_1, \ldots, \frac{\partial h}{\partial \eta_n}\Delta\eta_n$$

onto the Lorentzian rectangular solid spanned by the vectors

$$\frac{\partial \phi h}{\partial \eta_1}\Delta\eta_1, \ldots, \frac{\partial \phi h}{\partial \eta_n}\Delta\eta_n,$$

we deduce that

$$\mathrm{Vol}(\phi(X)) = \mathrm{Vol}(X).$$

In other words, hyperbolic volume is an isometry-invariant measure on H^n.

It is clear from Formula 3.4.5 that hyperbolic volume is countably additive, that is, if $\{X_i\}_{i=1}^{\infty}$ is a sequence of disjoint measurable subsets of H^n, then $X = \bigcup_{i=1}^{\infty} X_i$ is also measurable in H^n and

$$\mathrm{Vol}(X) = \sum_{i=1}^{\infty} \mathrm{Vol}(X_i).$$

Theorem 3.4.1. *The element of hyperbolic volume of H^n with respect to the Euclidean coordinates $x_1, \ldots, x_n$ in $\mathbb{R}^{n,1}$ is*

$$\frac{dx_1 \cdots dx_n}{[1 + (x_1^2 + \cdots + x_n^2)]^{\frac{1}{2}}}.$$

Proof: It is more convenient for us to work in $\mathbb{R}^{1,n}$ and show that the element of hyperbolic volume of H^n with respect to the coordinates $x_2, \ldots, x_{n+1}$ is

$$\frac{dx_2 \cdots dx_{n+1}}{[1 + (x_2^2 + \cdots + x_{n+1}^2)]^{\frac{1}{2}}}.$$

The desired result will then follow by a simple change of coordinates. Consider the transformation $\overline{h} : \mathbb{R}^{n-1} \times (0, 2\pi) \to \mathbb{R}^n$ defined by

$$\overline{h}(\eta_1, \ldots, \eta_n) = (x_2, \ldots, x_{n+1}),$$

where x_i is given by the system of Equations (3.4.1). Then by Formula 3.4.4, the vectors $\frac{\partial \overline{h}}{\partial \eta_1}, \ldots, \frac{\partial \overline{h}}{\partial \eta_n}$ are orthogonal. Hence, the Jacobian of the transformation $\overline{h}$ is given by

$$\begin{aligned} J\overline{h}(\eta_1, \ldots, \eta_n) &= \left|\frac{\partial \overline{h}}{\partial \eta_1}\right| \cdots \left|\frac{\partial \overline{h}}{\partial \eta_n}\right| \\ &= \cosh \eta_1 \sinh^{n-1} \eta_1 \sin^{n-2} \eta_2 \cdots \sin \eta_{n-1}. \end{aligned}$$

By changing variables via $\overline{h}$, we have

$$\int_{h^{-1}(X)} \sinh^{n-1}\eta_1 \sin^{n-2}\eta_2 \cdots \sin\eta_{n-1} d\eta_1 \cdots d\eta_n$$

$$= \int_{\overline{h}h^{-1}(X)} \frac{dx_2 \cdots dx_{n+1}}{\cosh\eta_1}$$

$$= \int_{p(X)} \frac{dx_2 \cdots dx_{n+1}}{x_1},$$

where $p : H^n \to \mathbb{R}^n$ is the projection

$$p(x_1, \ldots, x_{n+1}) = (x_2, \ldots, x_{n+1}).$$

□

Exercise 3.4

1. Show that the hyperbolic coordinates of a point x of H^n satisfy the system of Equations (3.4.1).
2. Show that the hyperbolic coordinates parameterization h satisfies Equations (3.4.2)-(3.4.4).
3. Show that the element of hyperbolic arc length $\|dx\|$ in hyperbolic coordinates is given by
$$\|dx\|^2 = d\eta_1^2 + \sinh^2\eta_1 d\eta_2^2 + \cdots + \sinh^2\eta_1 \sin^2\eta_2 \cdots \sin^2\eta_{n-1} d\eta_n^2.$$
4. Let $B(x, r)$ be the hyperbolic disk centered at a point x of H^2 of hyperbolic radius r. Show that the circumference of $B(x, r)$ is $2\pi \sinh r$ and the area of $B(x, r)$ is $2\pi(\cosh r - 1)$. Conclude that $B(x, r)$ has more area than a Euclidean disk of radius r.
5. Let $B(x, r)$ be the hyperbolic ball centered at a point x of H^3 of hyperbolic radius r. Show that the volume of $B(x, r)$ is $\pi(\sinh 2r - 2r)$.
6. Prove that every similarity of H^n, with $n > 1$, is an isometry.

§3.5. Hyperbolic Trigonometry

Let x, y, z be three hyperbolically noncollinear points of H^2. Let $L(x, y)$ be the unique hyperbolic line of H^2 containing x and y, and let $H(x, y, z)$ be the closed half-plane of H^2 with $L(x, y)$ as its boundary and z in its interior. The *hyperbolic triangle* with vertices x, y, z is defined to be

$$T(x, y, z) = H(x, y, z) \cap H(y, z, x) \cap H(z, x, y).$$

We shall assume that the vertices of $T(x, y, z)$ are labeled in negative order as in Figure 3.5.1.

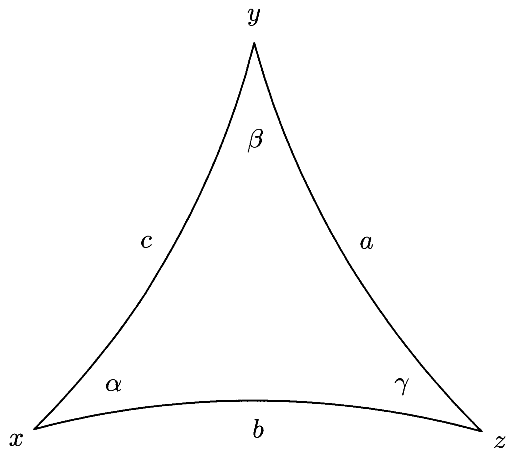


Figure 3.5.1. A hyperbolic triangle $T(x, y, z)$

Let $[x, y]$ be the segment of $L(x, y)$ joining x to y. The *sides* of $T(x, y, z)$ are defined to be $[x, y]$, $[y, z]$, and $[z, x]$. Let $a = \eta(y, z)$, $b = \eta(z, x)$, and $c = \eta(x, y)$. Then a, b, c is the hyperbolic length of $[y, z]$, $[z, x]$, $[x, z]$, respectively. Let

$$f : [0, a] \to H^2, \ g : [0, b] \to H^2, \ h : [0, c] \to H^2$$

be geodesic arcs from y to z, z to x, and x to y, respectively.

The *angle* α between the sides $[z, x]$ and $[x, y]$ of $T(x, y, z)$ is defined to be the Lorentzian angle between $-g'(b)$ and $h'(0)$. The *angle* β between the sides $[x, y]$ and $[y, z]$ of $T(x, y, z)$ is defined to be the Lorentzian angle between $-h'(c)$ and $f'(0)$. The *angle* γ between the sides $[y, z]$ and $[z, x]$ of $T(x, y, z)$ is defined to be the Lorentzian angle between $-f'(a)$ and $g'(0)$. The angles α, β, γ are called the *angles* of $T(x, y, z)$. The side $[y, z]$, $[z, x]$, $[x, y]$ is said to be *opposite* the angle α, β, γ, respectively.

Lemma 1. *If α, β, γ are the angles of a hyperbolic triangle $T(x, y, z)$, then*

(1) $\eta(z \otimes x,\ x \otimes y) = \pi - \alpha$,

(2) $\eta(x \otimes y,\ y \otimes z) = \pi - \beta$,

(3) $\eta(y \otimes z,\ z \otimes x) = \pi - \gamma$.

Proof: Without loss of generality, we may assume that $x = e_1$. The proof of (1) is evident from Figure 2.5.2. The proof of (2), and (3), is similar. □

Lemma 2. *Let x, y be space-like vectors in $\mathbb{R}^3$. If $x \otimes y$ is time-like, then*

$$|||x \otimes y||| = \|x\| \ \|y\| \sin \eta(x, y).$$

Proof: As $x \otimes y$ is time-like, the vector subspace of $\mathbb{R}^3$ spanned by x and y is space-like. By Theorem 3.2.1(4), we have

$$\begin{aligned}\|x \otimes y\|^2 &= (x \circ y)^2 - \|x\|^2\ \|y\|^2 \\ &= \|x\|^2\ \|y\|^2 \cos^2 \eta(x,y) - \|x\|^2\ \|y\|^2 \\ &= -\|x\|^2 \|y\|^2 \sin^2 \eta(x,y).\end{aligned}$$

□

Theorem 3.5.1. *If α, β, γ are the angles of a hyperbolic triangle, then*

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

Proof: Let α, β, γ be the angles of a hyperbolic triangle $T(x, y, z)$. By the same argument as in Theorem 2.5.1, the vectors $x \otimes y$, $z \otimes y$, $z \otimes x$ are linearly independent. Let

$$u = \frac{x \otimes y}{\|x \otimes y\|}, \quad v = \frac{z \otimes y}{\|z \otimes y\|}, \quad w = \frac{z \otimes x}{\|z \otimes x\|}.$$

Now as

$$(x \otimes y) \otimes (z \otimes y) = ((x \otimes y) \circ z)y$$

and

$$(z \otimes y) \otimes (z \otimes x) = ((x \otimes y) \circ z)z,$$

we have that both $u \otimes v$ and $v \otimes w$ are negative time-like vectors. By Lemma 2 and Theorems 3.1.6 and 3.2.1(4), we have

$$\begin{aligned}&\cos(\eta(u,v) + \eta(v,w)) \\ &= \cos \eta(u,v) \cos \eta(v,w) - \sin \eta(u,v) \sin \eta(v,w) \\ &= (u \circ v)(v \circ w) + \|u \otimes v\|\ \|v \otimes w\| \\ &> (u \circ v)(v \circ w) + ((u \otimes v) \circ (v \otimes w)) \\ &= (u \circ v)(v \circ w) + ((u \circ w)(v \circ v) - (v \circ w)(u \circ v)) \\ &= u \circ w \\ &= \cos \eta(u,w).\end{aligned}$$

Hence, either

$$\eta(u,w) > \eta(u,v) + \eta(v,w)$$

or

$$2\pi - \eta(u,w) < \eta(u,v) + \eta(v,w).$$

By Lemma 1, we have that $\eta(u,w) = \pi - \alpha$, $\eta(u,v) = \beta$, and $\eta(v,w) = \gamma$. Thus, either $\pi > \alpha + \beta + \gamma$ or $\pi + \alpha < \beta + \gamma$. Without loss of generality, we may assume that α is the largest angle. If $\pi + \alpha < \beta + \gamma$, we have the contradiction

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

Therefore, we have that

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

□

Theorem 3.5.2. (Law of Sines) *If α, β, γ are the angles of a hyperbolic triangle and a, b, c are the lengths of the opposite sides, then*

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

Proof: Upon taking norms of both sides of the equations

$$\begin{aligned}(z \otimes x) \otimes (x \otimes y) &= -((z \otimes x) \circ y)x,\\ (x \otimes y) \otimes (y \otimes z) &= -((x \otimes y) \circ z)y,\\ (y \otimes z) \otimes (z \otimes x) &= -((y \otimes z) \circ x)z,\end{aligned}$$

we find that

$$\begin{aligned}\sinh b\ \sinh c\ \sin \alpha &= |(x \otimes y) \circ z|,\\ \sinh c\ \sinh a\ \sin \beta &= |(x \otimes y) \circ z|,\\ \sinh a\ \sinh b\ \sin \gamma &= |(x \otimes y) \circ z|.\end{aligned}$$ □

Theorem 3.5.3. (The First Law of Cosines) *If α, β, γ are the angles of a hyperbolic triangle and a, b, c are the lengths of the opposite sides, then*

$$\cos \gamma = \frac{\cosh a \cosh b - \cosh c}{\sinh a\ \sinh b}.$$

Proof: Since

$$(y \otimes z) \circ (x \otimes z) = \begin{vmatrix} y \circ z & y \circ x \\ z \circ z & z \circ x \end{vmatrix},$$

we have that

$$\sinh a\ \sinh b \cos \gamma = \cosh a \cosh b - \cosh c.$$ □

Theorem 3.5.4. (The Second Law of Cosines) *If α, β, γ are the angles of a hyperbolic triangle and a, b, c are the lengths of the opposite sides, then*

$$\cosh c = \frac{\cos \alpha \cos \beta + \cos \gamma}{\sin \alpha \sin \beta}.$$

Proof: Let

$$x' = \frac{y \otimes z}{\|y \otimes z\|}, \quad y' = \frac{z \otimes x}{\|z \otimes x\|}, \quad z' = \frac{x \otimes y}{\|x \otimes y\|}.$$

Then

$$x = \frac{y' \otimes z'}{|||y' \otimes z'|||} \quad \text{and} \quad y = \frac{z' \otimes x'}{|||z' \otimes x'|||}.$$

Now since

$$(y' \otimes z') \circ (z' \otimes x') = \begin{vmatrix} y' \circ x' & y' \circ z' \\ z' \circ x' & z' \circ z' \end{vmatrix},$$

we have

$$-\sin(\pi - \alpha) \sin(\pi - \beta) \cosh c = -\cos(\pi - \gamma) - \cos(\pi - \alpha)\cos(\pi - \beta).$$ □

It is interesting to compare the hyperbolic sine law

$$\frac{\sinh a}{\sin \alpha} = \frac{\sinh b}{\sin \beta} = \frac{\sinh c}{\sin \gamma}$$

with the spherical sine law

$$\frac{\sin a}{\sin \alpha} = \frac{\sin b}{\sin \beta} = \frac{\sin c}{\sin \gamma},$$

and the hyperbolic cosine laws

$$\begin{aligned}\cos \gamma &= \frac{\cosh a \cosh b - \cosh c}{\sinh a \sinh b},\\ \cosh c &= \frac{\cos \alpha \cos \beta + \cos \gamma}{\sin \alpha \sin \beta}\end{aligned}$$

with the spherical cosine laws

$$\begin{aligned}\cos \gamma &= \frac{\cos c - \cos a \cos b}{\sin a \sin b},\\ \cos c &= \frac{\cos \alpha \cos \beta + \cos \gamma}{\sin \alpha \sin \beta}.\end{aligned}$$

Recall that

$$\sin ia = i \sinh a \quad \text{and} \quad \cos ia = \cosh a.$$

Hence, the hyperbolic trigonometry formulas can be obtained from their spherical counterparts by replacing a, b, c by ia, ib, ic, respectively.

Area of Hyperbolic Triangles

A *sector* of H^2 is defined to be the intersection of two distinct nonopposite half-planes of H^2. Any sector of H^2 is congruent to a sector $S(\alpha)$ defined in terms of hyperbolic coordinates (η, θ) by the inequalities

$$-\alpha/2 \leq \theta \leq \alpha/2.$$

Here α is the angle formed by the two sides of $S(\alpha)$ at its vertex e_1.

Let $\beta = \alpha/2$. Then the geodesic rays that form the sides of $S(\alpha)$ are represented in parametric form by

$$\begin{aligned}&(\cosh t)e_1 + (\sinh t)((\cos \beta)e_2 + (\sin \beta)e_3) \quad \text{for } t \geq 0,\\ &(\cosh t)e_1 + (\sinh t)((\cos \beta)e_2 - (\sin \beta)e_3) \quad \text{for } t \geq 0.\end{aligned}$$

These geodesic rays are asymptotic to the 1-dimensional light-like vector subspaces spanned by the vectors $(1, \cos \beta, \sin \beta)$ and $(1, \cos \beta, -\sin \beta)$, respectively. These two light-like vectors span a 2-dimensional vector subspace V that intersects H^2 in a hyperbolic line L. Let $T(\alpha)$ be the intersection of $S(\alpha)$ and the closed half-plane bounded by L and containing e_1. See Figure 3.5.2. It is an interesting fact, which will be proved in Chapter 4, that H^2 viewed from the origin looks like the projective disk model with the point e_1 at its center. Observe that the two sides of the sector $S(\alpha)$

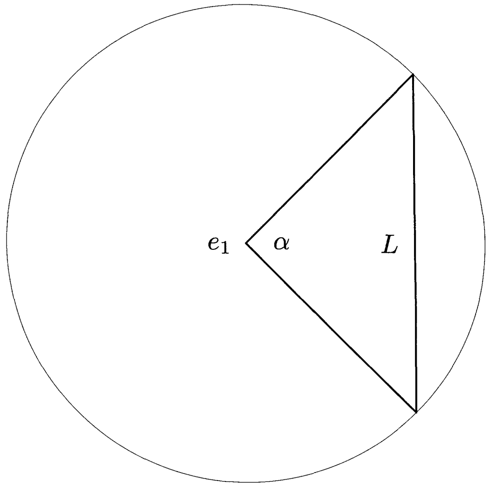


Figure 3.5.2. A generalized triangle with two ideal vertices

meet the hyperbolic line L at infinity. From this perspective, it is natural to regard $T(\alpha)$ as a hyperbolic triangle with two ideal vertices at infinity.

A *generalized hyperbolic triangle* in H^2 is defined in the same way that we defined a hyperbolic triangle in H^2 except that some of its vertices may be ideal. When viewed from the origin, a generalized hyperbolic triangle in H^2 appears to be a Euclidean triangle in the projective disk model with its ideal vertices on the circle at infinity. See Figure 3.5.2. The *angle* of a generalized hyperbolic triangle at an ideal vertex is defined to be zero.

An *infinite hyperbolic triangle* is a generalized hyperbolic triangle with at least one ideal vertex. An infinite hyperbolic triangle with three ideal vertices is called an *ideal hyperbolic triangle*. Obviously, any infinite hyperbolic triangle with exactly two ideal vertices is congruent to $T(\alpha)$ for some angle α.

We now find a parametric representation for the side L of $T(\alpha)$ in terms of hyperbolic coordinates (η, θ). To begin with, the vector

$$(1, \cos\beta, \sin\beta) \times (1, \cos\beta, -\sin\beta) = (-2\cos\beta\sin\beta, 2\sin\beta, 0)$$

is normal to the 2-dimensional vector subspace V whose intersection with H^2 is L. Hence, the vectors in V satisfy the equation

$$(\cos\beta)x_1 - x_2 = 0.$$

Now the points of H^2 satisfy the system of equations

$$\begin{cases} x_1 = \cosh\eta, \\ x_2 = \sinh\eta\cos\theta, \\ x_3 = \sinh\eta\sin\theta. \end{cases}$$

Hence, the points of L satisfy the equation

$$x_1 = \sec\beta\cos\theta\sqrt{x_1^2 - 1}.$$

Solving for x_1, we find that

$$x_1 = \frac{\cos\theta}{\sqrt{\cos^2\theta - \cos^2\beta}}.$$

Therefore

$$x_2 = \frac{\cos\theta\cos\beta}{\sqrt{\cos^2\theta - \cos^2\beta}}$$

and

$$x_3 = \frac{\sin\theta\cos\beta}{\sqrt{\cos^2\theta - \cos^2\beta}}.$$

Lemma 3. Area $T(\alpha) = \pi - \alpha$.

Proof: Let

$$x(\theta) = (x_1(\theta), x_2(\theta), x_3(\theta))$$

be the polar angle parameterization of L that we have just found. Then by Formula 3.4.5, we have

$$\begin{aligned}
\text{Area } T(\alpha) &= \int_{-\beta}^{\beta}\int_0^{\eta(e_1,x(\theta))} \sinh\eta\, d\eta d\theta \\
&= \int_{-\beta}^{\beta} (\cosh\eta(e_1, x(\theta)) - 1)d\theta \\
&= \int_{-\beta}^{\beta} x_1(\theta)d\theta \; - \; \alpha
\end{aligned}$$

and

$$\begin{aligned}
\int_{-\beta}^{\beta} x_1(\theta)d\theta &= \int_{-\beta}^{\beta} \frac{\cos\theta d\theta}{\sqrt{\cos^2\theta - \cos^2\beta}} \\
&= \int_{-\beta}^{\beta} \frac{\cos\theta d\theta}{\sqrt{\sin^2\beta - \sin^2\theta}} \\
&= \int_{-1}^{1} \frac{du}{\sqrt{1-u^2}}, \quad \text{where } u = \frac{\sin\theta}{\sin\beta} \\
&= \text{Arc}\sin u\Big|_{-1}^{1} \; = \; \pi.
\end{aligned}$$

Thus, we have that

$$\text{Area } T(\alpha) = \pi - \alpha.$$

□

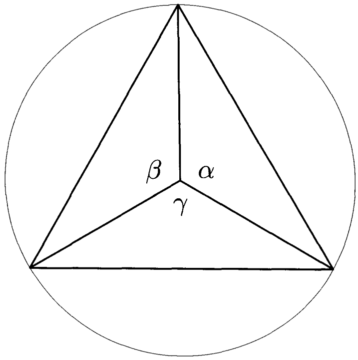


Figure 3.5.3. An ideal triangle subdivided into three infinite triangles

Lemma 4. *The area of an ideal hyperbolic triangle is* π.

Proof: Let T be any ideal hyperbolic triangle and let x be any point in the interior of T. Then T can be subdivided into three infinite hyperbolic triangles each of which has x as its only finite vertex. See Figure 3.5.3. Let α, β, γ be the angles of the triangles at the vertex x. Then

$$\text{Area}(T) = (\pi - \alpha) + (\pi - \beta) + (\pi - \gamma) = \pi. \qquad \square$$

Theorem 3.5.5. *If* α, β, γ *are the angles of a generalized hyperbolic triangle* T, *then*

$$\text{Area}(T) = \pi - (\alpha + \beta + \gamma).$$

Proof: By Lemmas 3 and 4, the formula holds if T has two or three ideal vertices. Suppose that T has only two finite vertices x and y with angles α and β. By extending the finite side of T, as in Figure 3.5.4, we see that T is the difference of two infinite hyperbolic triangles T_x and T_y with only one finite vertex x and y, respectively. Consequently

$$\text{Area}(T) = \text{Area}(T_x) - \text{Area}(T_y) = (\pi - \alpha) - \beta.$$

Now suppose that T has three finite vertices x, y, z with angles α, β, γ. By extending the sides of T, as in Figure 3.5.5, we can find an ideal hyperbolic triangle T' that can be subdivided into four regions, one of which is T, and the others are infinite hyperbolic triangles T_x, T_y, T_z with only one finite vertex x, y, z, respectively. Consequently, we have

$$\text{Area}(T') \;=\; \text{Area}(T) + \text{Area}(T_x) + \text{Area}(T_y) + \text{Area}(T_z).$$

Thus

$$\pi \;=\; \text{Area}(T) + \alpha + \beta + \gamma. \qquad \square$$

Corollary 1. *If* α, β, γ *are the angles of a generalized hyperbolic triangle, then*

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

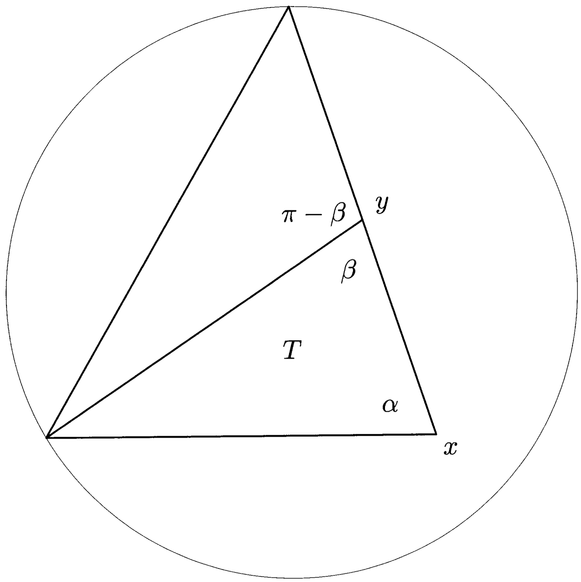


Figure 3.5.4. An infinite triangle T expressed as the difference of two triangles

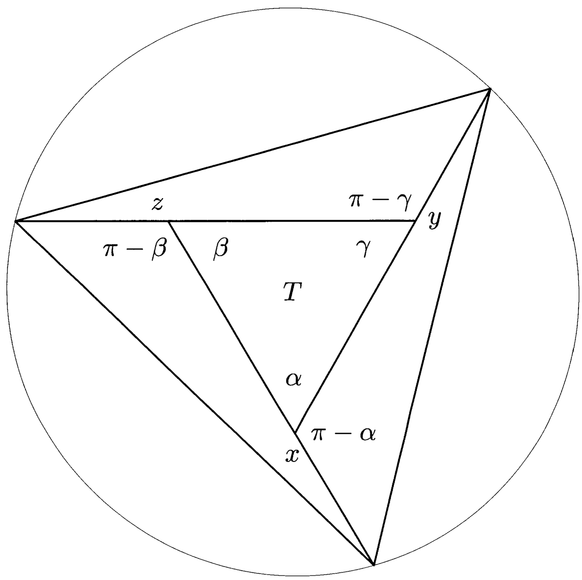


Figure 3.5.5. The ideal triangle found by extending the sides of $T(x, y, z)$

Existence of Hyperbolic Triangles

The next theorem extends Theorem 3.5.4 to the case $\gamma = 0$.

Theorem 3.5.6. *If $\alpha, \beta, 0$ are the angles of an infinite hyperbolic triangle with exactly one ideal vertex and c is the length of the finite side, then*

$$\cosh c = \frac{1 + \cos\alpha\cos\beta}{\sin\alpha\sin\beta}.$$

Proof: Let $T(x, y, z)$ be an infinite hyperbolic triangle with one ideal vertex z. We represent z by a positive light-like vector. Let

$$x' = \frac{y \otimes z}{\|y \otimes z\|}, \quad y' = \frac{z \otimes x}{\|z \otimes x\|}, \quad z' = \frac{x \otimes y}{\|x \otimes y\|}.$$

Then

$$x = \frac{y' \otimes z'}{|||y' \otimes z'|||} \quad \text{and} \quad y = \frac{z' \otimes x'}{|||z' \otimes x'|||}.$$

Let u be a point in the interior of the side $[x, z)$ and let v be a point in the interior of the side $[y, z)$. By Lemma 1, we have

$$\eta(u \otimes x,\ x \otimes y) = \pi - \alpha,$$
$$\eta(x \otimes y,\ y \otimes v) = \pi - \beta.$$

Hence, we have

$$\eta(z \otimes x,\ x \otimes y) = \pi - \alpha,$$
$$\eta(x \otimes y,\ y \otimes z) = \pi - \beta.$$

Now z is in the subspace V spanned by x' and y', and x' and y' are on opposite sides of $\langle z\rangle$ in V. Hence $x' \circ y' = -1$ by Theorems 3.2.9 and 3.2.10. Now since

$$(y' \otimes z') \circ (z' \otimes x') = \begin{vmatrix} y' \circ x' & y' \circ z' \\ z' \circ x' & z' \circ z' \end{vmatrix},$$

we have

$$-\sin(\pi - \alpha)\sin(\pi - \beta)\cosh c = -1 - \cos(\pi - \alpha)\cos(\pi - \beta). \qquad \square$$

We next prove a law of cosines for a hyperbolic quadrilateral with two adjacent right angles. See Figure 3.5.6.

Theorem 3.5.7. *Let Q be a hyperbolic convex quadrilateral with two adjacent right angles, opposite angles α, β, and sides of length c, d between α, β and the right angles, respectively. Then*

$$\cosh c = \frac{\cos\alpha\cos\beta + \cosh d}{\sin\alpha\sin\beta}.$$

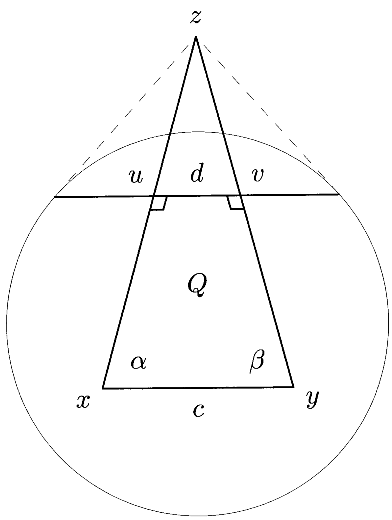


Figure 3.5.6. A hyperbolic quadrilateral Q with two adjacent right angles

Proof: Let x, y be the vertices of Q at α, β, and let z be the unit space-like vector Lorentz orthogonal and exterior to the side of Q of length d. Let

$$x' = \frac{y \otimes z}{\|y \otimes z\|}, \quad y' = \frac{z \otimes x}{\|z \otimes x\|}, \quad z' = \frac{x \otimes y}{\|x \otimes y\|}.$$

Then

$$x = \frac{y' \otimes z'}{|||y' \otimes z'|||} \quad \text{and} \quad y = \frac{z' \otimes x'}{|||z' \otimes x'|||}.$$

Now since

$$(y' \otimes z') \circ (z' \otimes x') = \begin{vmatrix} y' \circ x' & y' \circ z' \\ z' \circ x' & z' \circ z' \end{vmatrix},$$

we have

$$-\sin(\pi - \alpha)\sin(\pi - \beta)\cosh c = -\cosh d - \cos(\pi - \alpha)\cos(\pi - \beta). \qquad \square$$

Theorem 3.5.8. *Let Q be a hyperbolic convex quadrilateral with two adjacent right angles and opposite angles α, β. Then $\alpha + \beta < \pi$.*

Proof: Subdivide Q into two triangles with angles $\alpha, \beta_1, \gamma_1$ and $\beta_2, \gamma_2, \pi/2$ such that $\beta_1 + \beta_2 = \beta$ and $\gamma_1 + \gamma_2 = \pi/2$. Then

$$\begin{aligned} \text{Area}(Q) &= \pi - \alpha - \beta_1 - \gamma_1 + \pi - \beta_2 - \gamma_2 - \pi/2 \\ &= \pi - \alpha - \beta. \end{aligned} \qquad \square$$

We next prove the existence theorem for hyperbolic triangles.

Theorem 3.5.9. *Let α, β, γ be positive real numbers such that*

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

Then there is a hyperbolic triangle, unique up to congruence, with angles α, β, γ.

Proof: We shall only prove existence. The proof of uniqueness is left as an exercise for the reader. We may assume, without loss of generality, that $\alpha, \beta < \pi/2$. Now since

$$\alpha + \beta < \pi - \gamma,$$

we have that

$$\cos(\alpha + \beta) > \cos(\pi - \gamma).$$

Hence

$$\cos\alpha\cos\beta - \sin\alpha\sin\beta > -\cos\gamma,$$

and so

$$\cos\alpha\cos\beta + \cos\gamma > \sin\alpha\sin\beta.$$

Thus, we have that

$$\frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta} > 1.$$

Hence, there is a unique positive real number c satisfying the equation

$$\cosh c = \frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta}.$$

Let $[x, y]$ be a geodesic segment in H^2 of length c joining a point x to a point y, and let L_b, L_a be the hyperbolic lines passing through the points x, y, respectively, making an angle α, β, respectively, with $[x, y]$ on the same side of $[x, y]$. We claim that L_a and L_b meet on the same side of the hyperbolic line L_c, containing $[x, y]$, as α, β. The proof is by contradiction.

Assume first that L_a and L_b meet, possibly at infinity, on the opposite side of L_c than the angles α, β. Then the lines L_a, L_b, L_c form a generalized hyperbolic triangle two of whose angles are $\pi - \alpha$ and $\pi - \beta$, but

$$(\pi - \alpha) + (\pi - \beta) > \pi,$$

which contradicts Corollary 1.

Assume next that L_a and L_b do not meet, even at infinity. Then L_a and L_b have a common perpendicular hyperbolic line L_d joining a point u of L_b to a point v of L_a. Assume first that $u \neq x$, $v \neq y$ and that $[u, v]$ is on the opposite side of L_c. See Figure 3.5.7. Then u, v, x, y are the vertices of a hyperbolic quadrilateral with two adjacent right angles and opposite angles $\pi - \alpha$ and $\pi - \beta$, but

$$(\pi - \alpha) + (\pi - \beta) > \pi,$$

which contradicts Theorem 3.5.8.

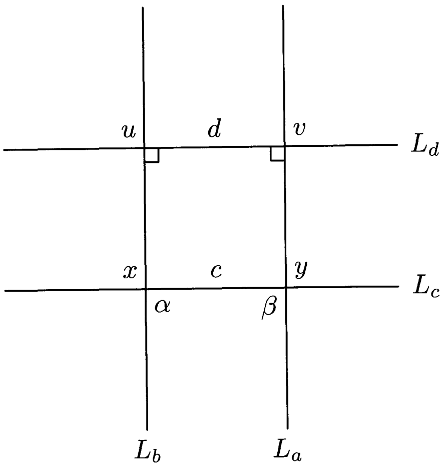


Figure 3.5.7. The four lines in the proof of Theorem 3.5.9

Next, assume that $u = x, v \neq y$ and that v is on the opposite side of L_c. Then x, y, v are the vertices of a hyperbolic triangle with angles $\pi/2 - \alpha, \pi - \beta, \pi/2$, but

$$(\pi/2 - \alpha) + (\pi - \beta) + \pi/2 > \pi,$$

which contradicts Corollary 1. Likewise, if $v = y$ and u is on the opposite side of L_c, we also have a contradiction.

Next, assume that $u \neq x$ and that u is on the same side of L_c as α, and $v \neq y$ and v is on the opposite side of L_c. Then the lines L_a, L_b, L_c, L_d form two hyperbolic triangles two of whose angles are $\alpha, \pi/2$ and $\pi - \beta, \pi/2$, respectively. As $\beta < \pi/2$, we have

$$\pi - \beta + \pi/2 > \pi,$$

which contradicts Corollary 1. Likewise, if $v \neq y$ and v is on the same side of L_c as β, and $u \neq x$ and u is on the opposite side of L_c, we also have a contradiction.

Next, assume that $v = y$, $u \neq x$ and that u is on the same side of L_c as α. Then x, y, u are the vertices of a hyperbolic triangle with angles $\alpha, \beta - \pi/2, \pi/2$, but $\beta < \pi/2$, which is a contradiction. Likewise, if $u = x$ and $v \neq y$ and v is on the same side of L_c as β, we also have a contradiction.

Finally, assume that $u \neq x, v \neq y$, and $[u, v]$ is on the same side of L_c as α, β. Then u, v, x, y are the vertices of a hyperbolic quadrilateral with two adjacent right angles and opposite angles α, β. By Theorem 3.5.7, we have

$$\cosh c = \frac{\cos\alpha\cos\beta + \cosh d}{\sin\alpha\sin\beta},$$

which is a contradiction, since $\cosh d > \cos\gamma$.

It follows that L_a and L_b meet, possibly at infinity, on the same side of L_c as α, β. Therefore, the lines L_a, L_b, L_c form a generalized hyperbolic triangle T with angles α, β, δ. By Theorems 3.5.4 and 3.5.6, we have

$$\cosh c = \frac{\cos\alpha\cos\beta + \cos\delta}{\sin\alpha\sin\beta}.$$

Hence $\cos\delta = \cos\gamma$ and therefore $\delta = \gamma$. Thus T is the desired triangle. □

Almost Rectangular Quadrilaterals and Pentagons

Theorem 3.5.10. *Let Q be a hyperbolic convex quadrilateral with three right angles and fourth angle γ, and let a, b the lengths of the sides opposite the angle γ. Then*

$$\cos\gamma = \sinh a \sinh b.$$

Proof: Let x, y be space-like vectors Lorentz orthogonal and exterior to the sides of Q of length a, b, respectively. Let z be the vertex of Q of angle γ and z' the opposite vertex. Let u, v be the vertices of Q between x, z and y, z, respectively. See Figure 3.5.8. By Lemma 1, we have

$$\eta(v \otimes z, z \otimes u) = \pi - \gamma.$$

Hence, we have

$$\eta(y \otimes z, z \otimes x) = \pi - \gamma.$$

Likewise $\eta(x, y) = \pi/2$.

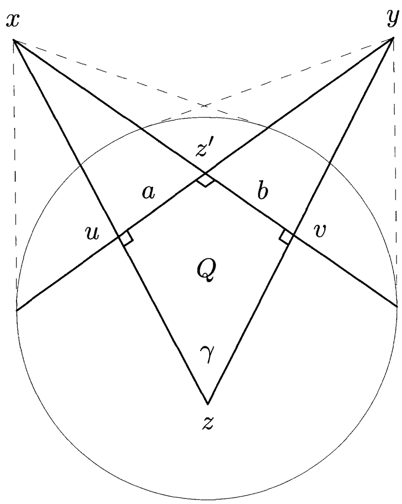


Figure 3.5.8. A hyperbolic quadrilateral Q with three right angles

Let
$$x' = \frac{y \otimes z}{\|y \otimes z\|} \quad \text{and} \quad y' = \frac{z \otimes x}{\|z \otimes x\|}.$$
Then
$$x = \frac{y' \otimes z'}{\|y' \otimes z'\|} \quad \text{and} \quad y = \frac{z' \otimes x'}{\|z' \otimes x'\|}.$$

Now since
$$(y' \otimes z') \circ (z' \otimes x') = \begin{vmatrix} y' \circ x' & y' \circ z' \\ z' \circ x' & z' \circ z' \end{vmatrix},$$
we have by Theorem 3.2.12 that
$$0 = -\cos(\pi - \gamma) - \sinh a \sinh b. \qquad \square$$

Theorem 3.5.11. *Let P be a hyperbolic convex pentagon with four right angles and fifth angle γ, let c' be the length of the side of P opposite γ, and let a, b be the lengths of the sides of P adjacent to the side opposite γ. Then*
$$\cosh c' = \frac{\cosh a \cosh b + \cos \gamma}{\sinh a \sinh b}.$$
Moreover, the above formula also holds if the vertex of P of angle γ is at infinity.

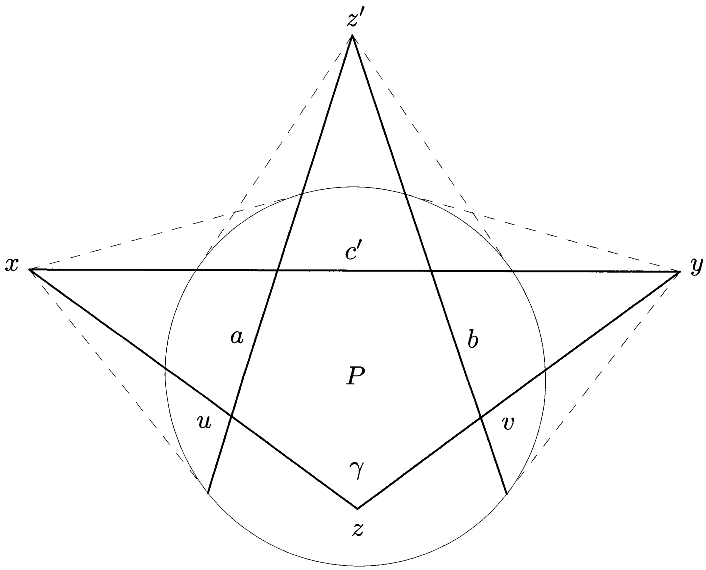


Figure 3.5.9. A hyperbolic pentagon P with four right angles

Proof: Assume first that the vertex z of angle γ is finite. Let x, y, z' be unit space-like vectors Lorentz orthogonal and exterior to the sides of P of length a, b, c', respectively. Let u, v be the vertices of P between x, z, and y, z, respectively. See Figure 3.5.9. By Lemma 1, we have

$$\eta(v \otimes z, z \otimes u) = \pi - \gamma.$$

Hence, we have

$$\eta(y \otimes z, z \otimes x) = \pi - \gamma.$$

Let

$$x' = \frac{y \otimes z}{\|y \otimes z\|} \quad \text{and} \quad y' = \frac{z \otimes x}{\|z \otimes x\|}.$$

Then

$$x = \frac{y' \otimes z'}{\|y' \otimes z'\|} \quad \text{and} \quad y = \frac{z' \circ x'}{\|z' \otimes x'\|}.$$

Now since

$$(y' \otimes z') \circ (z' \otimes x') = \begin{vmatrix} y' \circ x' & y' \circ z' \\ z' \circ x' & z' \circ z' \end{vmatrix},$$

we have

$$-\sinh a \sinh b \cosh c = -\cos \gamma - \cosh a \cosh b.$$

Assume now that z is at infinity. We can then represent z by a positive light-like vector. Let x' and y' be as above. Then z is in the subspace V spanned by x' and y', and x' and y' are on opposite sides of $\langle z \rangle$ in V. Hence $x' \circ y' = -1$ by Theorems 3.2.9 and 3.2.10. As before, we have

$$-\sinh a \sinh b \cosh c = -1 - \cosh a \cosh b. \qquad \square$$

Right-Angled Hyperbolic Hexagons

Let H be a right-angled hyperbolic convex hexagon in the projective disk model D^2 of the hyperbolic plane. Without loss of generality, we may assume that the center of D^2 is in the interior of H. Then no side of H is part of a diameter of D^2. As all the perpendiculars to a nondiameter line of D^2 meet in a common point outside of D^2, the three Euclidean lines extending three alternate sides of H meet pairwise in three points x, y, z outside of D. Likewise, the three Euclidean lines extending the opposite three alternate sides of H meet pairwise in three points x', y', z' outside of D^2. See Figure 3.5.10. The points x', y', z' are determined by the points x, y, z. To understand why, we switch to the hyperbolic model H^2. We can then represent x, y, z as unit space-like vectors that are Lorentz orthogonal and exterior to three alternate sides of H. Then

$$x' = \frac{y \otimes z}{\|y \otimes z\|}, \quad y' = \frac{z \otimes x}{\|z \otimes x\|}, \quad z' = \frac{x \otimes y}{\|x \otimes y\|}.$$

In other words $T(x', y', z')$ is the *polar triangle* of the *ultra-ideal triangle* $T(x, y, z)$. Compare with Formula 2.5.1. See also Figure 1.2.2.

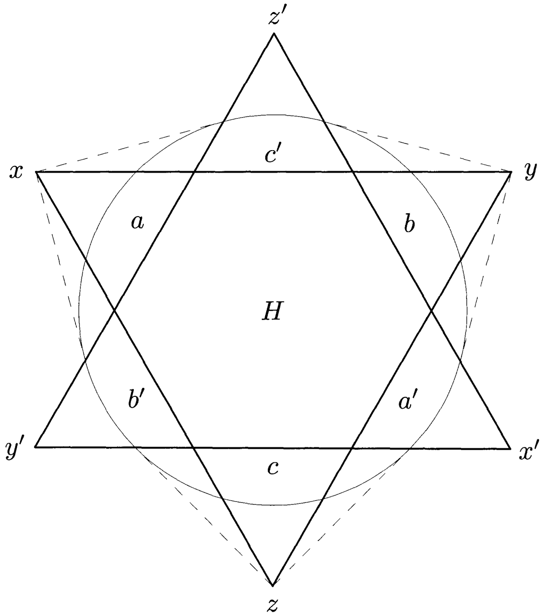


Figure 3.5.10. A right-angled hyperbolic hexagon H

Lemma 5. *Let x, y be space-like vectors in $\mathbb{R}^3$. If $x \otimes y$ is space-like, then*

$$\|x \otimes y\| = \|x\| \; \|y\| \sinh \eta(x, y).$$

Proof: As $x \otimes y$ is space-like, the vector subspace of $\mathbb{R}^3$ spanned by x and y is time-like. Hence

$$|x \circ y| = \|x\| \; \|y\| \cosh \eta(x, y).$$

By Theorem 3.2.1(4), we have

$$\begin{aligned} \|x \otimes y\|^2 &= (x \circ y)^2 - \|x\|^2 \|y\|^2 \\ &= \|x\|^2 \|y\|^2 \cosh^2 \eta(x, y) - \|x\|^2 \|y\|^2 \\ &= \|x\|^2 \|y\|^2 \sinh^2 \eta(x, y). \end{aligned}$$

□

Theorem 3.5.12. (Law of Sines for right-angled hyperbolic hexagons) *If a, b, c are the lengths of alternate sides of a right-angled hyperbolic convex hexagon and a', b', c' are the lengths of the opposite sides, then*

$$\frac{\sinh a}{\sinh a'} = \frac{\sinh b}{\sinh b'} = \frac{\sinh c}{\sinh c'}.$$

Proof: By Theorem 3.2.8, we have

$$\begin{aligned} a' &= \eta(y, z), \; b' = \eta(z, x), \; c' = \eta(y, z), \\ a &= \eta(y', z'), \; b = \eta(z', x'), \; c = \eta(y', z'). \end{aligned}$$

Upon taking norms of both sides of the equations

$$\begin{aligned}(z \otimes x) \otimes (x \otimes y) &= -((z \otimes x) \circ y)x, \\ (x \otimes y) \otimes (y \otimes z) &= -((x \otimes y) \circ z)y, \\ (y \otimes z) \otimes (z \otimes x) &= -((y \otimes z) \circ x)z,\end{aligned}$$

we find that

$$\begin{aligned}\sinh b' \sinh c' \sinh a &= |(x \otimes y) \circ z|, \\ \sinh c' \sinh a' \sinh b &= |(x \otimes y) \circ z|, \\ \sinh a' \sinh b' \sinh c &= |(x \otimes y) \circ z|.\end{aligned}$$

□

Theorem 3.5.13. (Law of Cosines for right-angled hyperbolic hexagons) *If a, b, c are the lengths of alternate sides of a right-angled hyperbolic convex hexagon and a', b', c' are the lengths of the opposite sides, then*

$$\cosh c' = \frac{\cosh a \cosh b + \cosh c}{\sinh a \sinh b}.$$

Proof: Since

$$(y \otimes z) \circ (z \otimes x) = \begin{vmatrix} y \circ x & y \circ z \\ z \circ x & z \circ z \end{vmatrix},$$

we have by Theorem 3.2.8 that

$$-\sinh a' \sinh b' \cosh c = -\cosh c' - \cosh a' \cosh b'.$$

□

Corollary 2. *The lengths of three alternate sides of a right-angled hyperbolic hexagon are determined by the lengths of the opposite three sides.*

We now prove the existence theorem for right-angled hexagons.

Theorem 3.5.14. *Let a, b, c be positive real numbers. Then there is a right-angled hyperbolic convex hexagon, unique up to congruence, with alternate sides of length a, b, c, respectively.*

Proof: Let c' be the unique positive real number that satisfies the equation

$$\cosh c' = \frac{\cosh a \cosh b + \cosh c}{\sinh a \sinh b}$$

and let $S_{c'}$ be a geodesic segment in H^2 of length c'. Erect perpendicular geodesic segments S_a and S_b of length a and b, respectively, at the endpoints of $S_{c'}$ on the same side of $S_{c'}$. Let $L_{a'}$ and $L_{b'}$ be the hyperbolic lines perpendicular to S_b and S_a, respectively, at the endpoint of S_b and S_a, respectively, opposite the endpoint of $S_{c'}$. See Figure 3.5.10.

Without loss of geomerality, we may assume that $c \geq a, b$. Then $L_{b'}$ does not meet S_b; otherwise, we would have a quadrilateral with three right angles and fourth angle γ, and opposite sides of length a and c', and so by Theorem 3.5.10, we would have

$$\sinh a \sinh c' = \cos \gamma,$$

but

$$\begin{aligned}
\sinh^2 a \sinh^2 c' &= \sinh^2 a(\cosh^2 c' - 1) \\
&= \frac{(\cosh a \cosh b + \cosh c)^2 - \sinh^2 a \sinh^2 b}{\sinh^2 b} \\
&> \frac{\cosh^2 c}{\sinh^2 b} \\
&> 1,
\end{aligned}$$

which is a contradiction. Likewise $L_{a'}$ does not meet S_a. Moreover $L_{a'}$ does not meet $L_{b'}$, even at infinity; otherwise, we would have a pentagon with four right-angles and fifth angle γ as in Figure 3.5.9, and so by Theorem 3.5.11, we would have

$$\cosh c' = \frac{\cosh a \cosh b + \cos\gamma}{\sinh a \sinh b},$$

which is a contradiction, since $\cosh c > \cos\gamma$.

By Theorems 3.2.6-3.2.9, the hyperbolic lines $L_{a'}$ and $L_{b'}$ have a common perpendicular hyperbolic line L_c. Let L_a, L_b be the hyperbolic line of H^2 containing S_a, S_b, respectively. Then L_c is on the same side of L_a as $S_{c'}$, since L_c meets $L_{a'}$ and $L_{a'}$ is on the same side of L_a as $S_{c'}$. Likewise L_c is on the same side of L_b as $S_{c'}$. Let S_c be the segment of L_c joining $L_{a'}$ to $L_{b'}$. Then we have a right-angled convex hexagon H with alternate sides S_a, S_b, S_c. Let d be the length of S_c. Then by Theorem 3.5.13, we have

$$\cosh c' = \frac{\cosh a \cosh b + \cosh d}{\sinh a \sinh b}.$$

Hence $d = c$. Thus H has alternate sides of length a, b, c. The proof that H is unique up to congruence is left as an exercise for the reader. □

Exercise 3.5

1. Let α, β, γ be the angles of a hyperbolic triangle and let a, b, c be the lengths of the opposite sides. Prove that $a \leq b \leq c$ if and only if $\alpha \leq \beta \leq \gamma$.

2. Let α, β, γ be the angles of a hyperbolic triangle and let a, b, c be the lengths of the opposite sides. Show that

(1)
$$\cosh a = \cosh b \cosh c - \sinh b \sinh c \cos\alpha, \tag{3.5.1}$$
$$\cosh b = \cosh a \cosh c - \sinh a \sinh c \cos\beta, \tag{3.5.2}$$
$$\cosh c = \cosh a \cosh b - \sinh a \sinh b \cos\gamma, \tag{3.5.3}$$

(2)
$$\cos\alpha = -\cos\beta\cos\gamma + \sin\beta\sin\gamma\cosh a, \tag{3.5.4}$$
$$\cos\beta = -\cos\alpha\cos\gamma + \sin\alpha\sin\gamma\cosh b, \tag{3.5.5}$$
$$\cos\gamma = -\cos\alpha\cos\beta + \sin\alpha\sin\beta\cosh c. \tag{3.5.6}$$

3. Let $\alpha, \beta, \pi/2$ be the angles of a hyperbolic right triangle and let a, b, c be the lengths of the opposite sides. Show that

$$\text{(1)} \qquad \cosh c = \cosh a \cosh b, \tag{3.5.7}$$

$$\text{(2)} \qquad \cosh c = \cot\alpha \cot\beta, \tag{3.5.8}$$

$$\text{(3)} \qquad \sinh a = \sinh c \sin\alpha, \tag{3.5.9}$$
$$\sinh b = \sinh c \sin\beta, \tag{3.5.10}$$

$$\text{(4)} \qquad \cos\alpha = \tanh b \coth c, \tag{3.5.11}$$
$$\cos\beta = \tanh a \coth c, \tag{3.5.12}$$

$$\text{(5)} \qquad \sinh a = \tanh b \cot\beta, \tag{3.5.13}$$
$$\sinh b = \tanh a \cot\alpha, \tag{3.5.14}$$

$$\text{(6)} \qquad \cos\alpha = \cosh a \sin\beta, \tag{3.5.15}$$
$$\cos\beta = \cosh b \sin\alpha. \tag{3.5.16}$$

4. Let $\alpha, \beta, 0$ be the angles of an infinite hyperbolic triangle with exactly one ideal vertex and let c be the length of the finite side. Show that

$$\sinh c = \frac{\cos\alpha + \cos\beta}{\sin\alpha \sin\beta}. \tag{3.5.17}$$

5. Prove that a generalized hyperbolic triangle is equilateral if and only if it is equiangular.

6. Show that for a hyperbolic equilateral triangle of angle α and side length a,

$$\cosh(a/2)\sin(\alpha/2) = 1/2. \tag{3.5.18}$$

7. Prove that an angle bisector of a hyperbolic triangle T bisects the opposite side of T if and only if the other two sides of T have the same length.

8. Prove that the three angle bisectors of a hyperbolic triangle T meet in a common point inside T equidistant from each of the three sides of T.

9. Let $T(x, y, z)$ be a hyperbolic triangle labeled as in Figure 3.5.1 such that $\alpha, \beta < \pi/2$. Prove that the point on the hyperbolic line through x and y nearest to z lies in the interior of the side $[x, y]$.

10. Let α, β, γ be nonnegative real numbers such that $\alpha + \beta + \gamma < \pi$. Prove that there is a generalized hyperbolic triangle with angles α, β, γ.

11. Prove that two generalized hyperbolic triangles are congruent if and only if they have the same angles.

12. Prove that two right-angled hyperbolic convex hexagons are congruent if and only if they have the same three lengths for alternate sides.

§3.6. Historical Notes

§3.1. Lorentzian geometry was introduced by Klein in his 1873 paper *Ueber die sogenannte Nicht-Euklidische Geometrie* [227] and was developed by Killing in his 1885 treatise *Nicht-Euklidischen Raumformen* [221]. Three-dimensional Lorentzian geometry was described by Poincaré in his 1887 paper *Sur les hypothèses fondamentales de la géométrie* [335]. See also Bianchi's 1888 paper *Sulle forme differenziali quadratiche indefinite* [45]. Lorentzian 4-dimensional space was introduced by Poincaré as a model for space-time in his 1906 paper *Sur la dynamique de l'électron* [338]. For commentary on Poincaré's paper, see Miller's 1973 article *A study of Henri Poincaré's "Sur la dynamique de l'électron"* [288]. Lorentzian 4-dimensional space was proposed as a model for space-time in the theory of special relativity by Minkowski in his 1907 lecture *Das Relativitätsprinzip* [296]. For commentary, see Pyenson's 1977 article *Hermann Minkowski and Einstein's Special Theory of Relativity* [345]. Lorentzian geometry was developed by Minkowski in his 1908 paper *Die Grundgleichungen für die elektromagnetischen Vorgänge in bewegten Körpern* [293] and in his 1909 paper *Raum und Zeit* [294]. Lorentzian 4-space is also called *Minkowski space-time.* Lorentz transformations of n-space were first considered by Killing in his 1885 treatise [221]. In particular, Theorem 3.1.3 appeared in Killing's treatise. Lorentz transformations of space-time were introduced by Lorentz in his 1904 paper *Electromagnetic phenomena in a system moving with any velocity less than that of light* [271]. The terms *Lorentz transformation* and *Lorentz group* were introduced by Poincaré in his 1906 paper [338]. The geometry of the Lorentz group was studied by Klein in his 1910 paper *Über die geometrichen Grundlagen der Lorentzgruppe* [236]. For a discussion of the role played by Lorentzian geometry in the theory of relativity, see Penrose's 1978 article *The geometry of the universe* [325] and Naber's 1992 monograph *The Geometry of Minkowski Spacetime* [313].

§3.2. The hyperboloid model of hyperbolic space and Formula 3.2.2 appeared in Killing's 1878 paper *Ueber zwei Raumformen mit constanter positiver Krümmung* [219]. The time-like and space-like angles were essentially defined by Klein in his 1871 paper *Ueber die sogenannte Nicht-Euklidische Geometrie* [224]. Most of the material in §3.2 appeared in Killing's 1885 treatise [221]. Other references for this section are Klein's 1928 treatise *Vorlesungen über nich-euklidisch Geometrie* [237], Coxeter's 1942 treatise *Non-Euclidean Geometry* [91], Busemann and Kelly's 1953 treatise *Projective Geometry and Projective Metrics* [64], and Thurston's 1979 lecture notes *The Geometry and Topology of 3-Manifolds* [389].

§3.3. The element of hyperbolic arc length of the hyperboloid model of hyperbolic space appeared in Killing's 1880 paper *Die Rechnung in den Nicht-Euklidischen Raumformen* [220]. The Lorentzian length of a hyperbolic line segment was defined by Yaglom in his 1979 monograph *A Simple Non-Euclidean Geometry and Its Physical Basis* [419].

§3.4. Two-dimensional hyperbolic coordinates appeared as polar coordinates in Lobachevski's 1829-30 paper *On the principles of geometry* [262]. Two-dimensional hyperbolic coordinates were defined by Cox in terms of Euclidean coordinates in his 1882 paper *Homogeneous coordinates in imaginary geometry* [84]. Moreover, Cox gave the element of hyperbolic area in both hyperbolic and Euclidean coordinates in this paper. Hyperbolic coordinates in n-dimensions and Formula 3.4.5 appeared in Böhm and Hertel's 1981 treatise *Polyedergeometrie in n-dimensionalen Räumen konstanter Krümmung* [50].

§3.5. That the sum of the angles of a hyperbolic triangle is less than two right angles was proved by Saccheri, under his acute angle hypotheis, in his 1733 treatise *Euclides ab omni naevo vindicatus* [355]. Formulas equivalent to the hyperbolic sine and cosine laws appeared in Lobachevski's 1829-30 paper [262]. See also his 1837 paper *Géométrie imaginaire* [264]. The law of sines appeared in a form that is valid in spherical, Euclidean, and hyperbolic geometries in Bolyai's 1832 paper *Scientiam spatii absolute veram exhibens* [51]. The duality between hyperbolic and spherical trigonometries was developed by Lambert in his 1770 memoire *Observations trigonométriques* [251]. Taurinus proposed that the duality between hyperbolic and spherical trigonometries infers the existence of a geometry opposite to spherical geometry and studied its properties in his 1826 treatise *Geometriae prima elementa* [386]. That the area of a hyperbolic triangle is proportional to its angle defect first appeared in Lambert's monograph *Theorie der Parallellinien* [252], which was published posthumously in 1786. For a translation of the relevant passages, see Rosenfeld's 1988 treatise *A History of Non-Euclidean Geometry* [353]. The elegant proof of Theorem 3.5.5 was communicated to Bolyai's father by Gauss in his letter of March 6, 1832. For a translation, see Coxeter's 1977 article *Gauss as a geometer* [93].

The law of cosines for quadrilaterals with two adjacent right angles appeared in Ranum's 1912 paper *Lobachefskian polygons trigonometrically equivalent to the triangle* [346]. The cosine law for trirectangular quadrilaterals appeared in Barbarin's 1901 treatise *Études de géométrie analytique non Euclidienne* [30]. The law of cosines for quadrectangular pentagons appeared in Ranum's 1912 paper [346]. That the formulas of spherical trigonometry with pure imaginary arguments admit an interpretation as formulas for right-angled hyperbolic hexagons appeared implicitly in Schilling's 1891 note *Ueber die geometrische Bedeutung der Formeln der sphärischen Trigonometrie im Falle complexer Argumente* [357]. The sine and cosine laws for right-angled hyperbolic hexagons appeared implicitly in Schilling's 1894 paper *Beiträge zur geometrischen Theorie der Schwarz'schen s-Function* [358] and explicitly in Ranum's 1912 paper [346]. References for hyperbolic trigonometry are Beardon's 1983 treatise *The Geometry of Discrete Groups* [34] and Fenchel's 1989 treatise *Elementary Geometry in Hyperbolic Space* [132].

CHAPTER 4

Inversive Geometry

In this chapter, we study the group of transformations of E^n generated by reflections in hyperplanes and inversions in spheres. It turns out that this group is isomorphic to the group of isometries of H^{n+1}. This leads to a deeper understanding of hyperbolic geometry. In Sections 4.5 and 4.6, the conformal ball and upper half-space models of hyperbolic n-space are introduced. The chapter ends with a geometric analysis of the isometries of hyperbolic n-space.

§4.1. Reflections

Let a be a unit vector in E^n and let t be a real number. Consider the hyperplane of E^n defined by

$$P(a,t) = \{x \in E^n : a \cdot x = t\}.$$

Observe that every point x in $P(a,t)$ satisfies the equation

$$a \cdot (x - ta) = 0.$$

Hence $P(a,t)$ is the hyperplane of E^n with unit normal vector a passing through the point ta. One can easily show that every hyperplane of E^n is of this form, and every hyperplane has exactly two representations $P(a,t)$ and $P(-a,-t)$.

The *reflection* ρ of E^n in the plane $P(a,t)$ is defined by the formula

$$\rho(x) = x + sa,$$

where s is a real scalar so that $x + \frac{1}{2}sa$ is in $P(a,t)$. This leads to the explicit formula

$$\rho(x) = x + 2(t - a \cdot x)a. \tag{4.1.1}$$

The proof of the following theorem is routine and is left as an exercise for the reader.

Theorem 4.1.1. *If ρ is the reflection of E^n in the plane $P(a,t)$, then*

(1) $\rho(x) = x$ *if and only if* x *is in* $P(a,t)$*;*

(2) $\rho^2(x) = x$ *for all* x *in* E^n*; and*

(3) ρ *is an isometry.*

Theorem 4.1.2. *Every isometry of E^n is a composition of at most $n+1$ reflections in hyperplanes.*

Proof: Let $\phi : E^n \to E^n$ be an isometry and set $v_0 = \phi(0)$. Let ρ_0 be the identity if $v_0 = 0$, or the reflection in the plane $P(v_0/|v_0|, |v_0|/2)$ otherwise. Then $\rho_0(v_0) = 0$ and so $\rho_0\phi(0) = 0$. By Theorem 1.3.5, the map $\phi_0 = \rho_0\phi$ is an orthogonal transformation.

Now suppose that ϕ_{k-1} is an orthogonal transformation of E^n that fixes $e_1, \ldots, e_{k-1}$. Let $v_k = \phi_{k-1}(e_k) - e_k$ and let ρ_k be the identity if $v_k = 0$, or the reflection in the plane $P(v_k/|v_k|, 0)$ otherwise. Then $\rho_k\phi_{k-1}$ fixes e_k. See Figure 4.1.1. Also, for each $j = 1, \ldots, k-1$, we have

$$\begin{aligned} v_k \cdot e_j &= (\phi_{k-1}(e_k) - e_k) \cdot e_j \\ &= \phi_{k-1}(e_k) \cdot e_j \\ &= \phi_{k-1}(e_k) \cdot \phi_{k-1}(e_j) \\ &= e_k \cdot e_j \\ &= 0. \end{aligned}$$

Therefore e_j is in the plane $P(v_k/|v_k|, 0)$ and so is fixed by ρ_k. Thus, we have that $\phi_k = \rho_k\phi_{k-1}$ fixes $e_1, \ldots, e_k$. It follows by induction that there are maps $\rho_0, \ldots, \rho_n$ such that each ρ_i is either the identity or a reflection and $\rho_n \cdots \rho_0\phi$ fixes $0, e_1, \ldots, e_n$. Therefore $\rho_n \cdots \rho_0\phi$ is the identity and we have that $\phi = \rho_0 \cdots \rho_n$. □

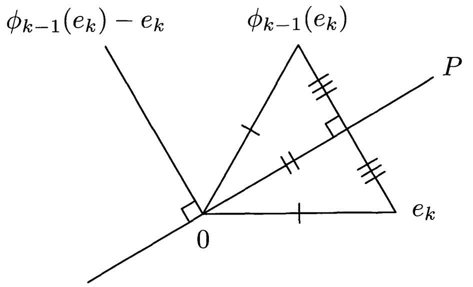


Figure 4.1.1. The reflection of the point $\phi_{k-1}(e_k)$ in the plane P

Inversions

Let a be a point of E^n and let r be a positive real number. The *sphere* of E^n of *radius* r *centered at* a is defined to be the set

$$S(a,r) = \{x \in E^n : |x-a| = r\}.$$

The *reflection* (or *inversion*) σ of E^n in the sphere $S(a,r)$ is defined by the formula

$$\sigma(x) = a + s(x-a),$$

where s is a positive scalar so that

$$|\sigma(x) - a|\ |x-a| = r^2.$$

This leads to the explicit formula

$$\sigma(x) = a + \left(\frac{r}{|x-a|}\right)^2 (x-a). \tag{4.1.2}$$

There is a nice geometric construction of the point $\sigma(x)$. Assume first that x is inside $S(a,r)$. Erect a chord of $S(a,r)$ passing through x perpendicular to the line joining a to x. Let u and v be the endpoints of the chord. Then $\sigma(x)$ is the point x' of intersection of the lines tangent to $S(a,r)$ at the points u and v in the plane containing a, u, and v, as in Figure 4.1.2. Observe that the right triangles $T(a,x,v)$ and $T(a,v,x')$ are similar. Consequently, we have

$$\frac{|x'-a|}{r} = \frac{r}{|x-a|}.$$

Therefore $x' = \sigma(x)$ as claimed.

Now assume that x is outside $S(a,r)$. Let y be the midpoint of the line segment $[a,x]$ and let C be the circle centered at y of radius $|x-y|$. Then C intersects $S(a,r)$ in two points u, v, and $\sigma(x)$ is the point x' of intersection of the line segments $[a,x]$ and $[u,v]$, as in Figure 4.1.3.

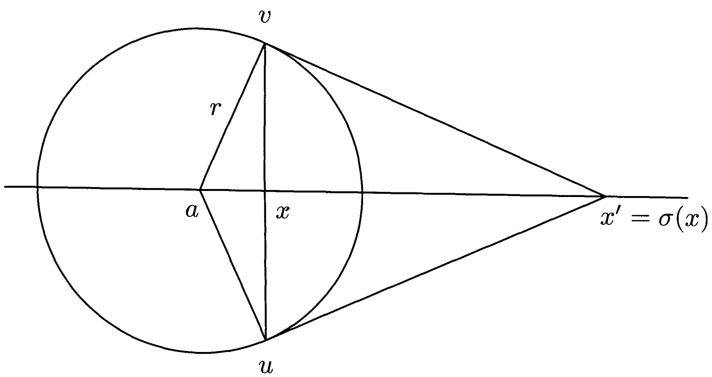


Figure 4.1.2. The construction of the reflection of a point x in a sphere $S(a,r)$

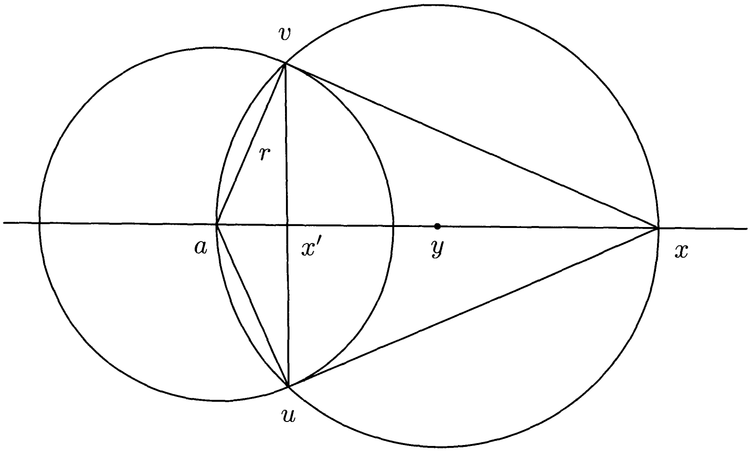


Figure 4.1.3. The construction of the reflection of a point x in a sphere $S(a,r)$

Theorem 4.1.3. *If σ is the reflection of E^n in the sphere $S(a,r)$, then*

(1) *$\sigma(x) = x$ if and only if x is in $S(a,r)$;*

(2) *$\sigma^2(x) = x$ for all $x \neq a$; and*

(3) *for all $x, y \neq a$,*

$$|\sigma(x) - \sigma(y)| = \frac{r^2|x-y|}{|x-a|\,|y-a|}.$$

Proof: (1) Since

$$|\sigma(x) - a|\,|x-a| = r^2,$$

we have that $\sigma(x) = x$ if and only if $|x-a| = r$.

(2) Observe that

$$\begin{aligned}
\sigma^2(x) &= a + \left(\frac{r}{|\sigma(x)-a|}\right)^2 (\sigma(x)-a) \\
&= a + \left(\frac{|x-a|}{r}\right)^2 \left(\frac{r}{|x-a|}\right)^2 (x-a) \\
&= x.
\end{aligned}$$

(3) Observe that

$$\begin{aligned}
|\sigma(x) - \sigma(y)| &= r^2 \left| \frac{(x-a)}{|x-a|^2} - \frac{(y-a)}{|y-a|^2} \right| \\
&= r^2 \left[\frac{1}{|x-a|^2} - \frac{2(x-a)\cdot(y-a)}{|x-a|^2\,|y-a|^2} + \frac{1}{|y-a|^2} \right]^{1/2} \\
&= \frac{r^2|x-y|}{|x-a|\,|y-a|}.
\end{aligned}$$

□

Conformal Transformations

Let U be an open subset of E^n and let $\phi : U \to E^n$ be a differentiable function. Let $\phi'(x)$ be the matrix $\left(\frac{\partial \phi_i}{\partial x_j}(x)\right)$ of partial derivatives of ϕ. The function ϕ is said to be *conformal* if and only if there is a function

$$\kappa : U \to \mathbb{R}_+,$$

called the *scale factor* of ϕ, such that $\kappa(x)^{-1}\phi'(x)$ is an orthogonal matrix for each x in U. Notice that the scale factor κ of a conformal function ϕ is uniquely determined by ϕ, since $[\kappa(x)]^n = |\det \phi'(x)|$.

Lemma 1. *Let A be a real $n \times n$ matrix. Then there is a positive scalar k such that $k^{-1}A$ is an orthogonal matrix if and only if A preserves angles between nonzero vectors.*

Proof: Suppose there is a $k > 0$ such that $k^{-1}A$ is an orthogonal matrix. Then A is nonsingular. Let x and y be nonzero vectors in E^n. Then Ax and Ay are nonzero, and A preserves angles, since

$$\begin{aligned} \cos\theta(Ax, Ay) &= \frac{Ax \cdot Ay}{|Ax|\,|Ay|} \\ &= \frac{k^{-1}Ax \cdot k^{-1}Ay}{|k^{-1}Ax|\,|k^{-1}Ay|} \\ &= \frac{x \cdot y}{|x|\,|y|} \quad = \quad \cos\theta(x, y). \end{aligned}$$

Conversely, suppose that A preserves angles between nonzero vectors. Then A is nonsingular. As $\theta(Ae_i, Ae_j) = \theta(e_i, e_j) = 0$ for all $i \neq j$, the vectors $Ae_1, \ldots, Ae_n$ are orthogonal. Let B be the orthogonal matrix such that $Be_i = Ae_i/|Ae_i|$ for each i. Then $B^{-1}A$ also preserves angles and $B^{-1}Ae_i = c_ie_i$ where $c_i = |Ae_i|$. Thus, we may assume, without loss of generality, that $Ae_i = c_ie_i$, with $c_i > 0$, for each $i = 1, \ldots, n$. As

$$\theta(A(e_i + e_j), Ae_j) = \theta(e_i + e_j, e_j)$$

for all $i \neq j$, we have

$$\frac{(c_ie_i + c_je_j) \cdot c_je_j}{(c_i^2 + c_j^2)^{1/2}c_j} = \frac{1}{\sqrt{2}}.$$

Thus $2c_j^2 = c_i^2 + c_j^2$ and so $c_i = c_j$ for all i and j. Therefore, the common value of the c_i is a positive scalar k such that $k^{-1}A$ is orthogonal. □

Let $\alpha, \beta : [-b, b] \to E^n$ be differentiable curves such that $\alpha(0) = \beta(0)$ and $\alpha'(0), \beta'(0)$ are both nonzero. The *angle* between α and β at 0 is defined to be the angle between $\alpha'(0)$ and $\beta'(0)$.

Theorem 4.1.4. *Let U be an open subset of E^n and let $\phi : U \to E^n$ be a differentiable function. Then ϕ is conformal if and only if ϕ preserves angles between differentiable curves in U.*

Proof: Suppose that the function ϕ is conformal. Then there is a function $\kappa : U \to \mathbb{R}_+$ such that $\kappa(x)^{-1}\phi'(x)$ is orthogonal for each x in U. Let $\alpha, \beta : [-b, b] \to U$ be differentiable curves such that $\alpha(0) = \beta(0)$ and $\alpha'(0), \beta'(0)$ are both nonzero. Then by Lemma 1, we have

$$\begin{aligned} &\theta((\phi\alpha)'(0), (\phi\beta)'(0)) \\ &\quad = \theta(\phi'(\alpha(0))\alpha'(0), \phi'(\beta(0))\beta'(0)) \\ &\quad = \theta(\alpha'(0), \beta'(0)). \end{aligned}$$

Hence, the angle between $\phi\alpha$ and $\phi\beta$ at 0 is the same as the angle between α and β at 0.

Conversely, suppose that ϕ preserves angles between differentiable curves in U. Then the matrix $\phi'(x)$ preserves angles between nonzero vectors for each x. By Lemma 1, there is a positive scalar $\kappa(x)$ such that $\kappa(x)^{-1}\phi'(x)$ is orthogonal for each x in U. Thus ϕ is conformal. □

Let U be an open subset of E^n and let $\phi : U \to E^n$ be a differentiable function. Then ϕ is said to *preserve* (resp. *reverse*) *orientation at a point* x of U if and only if $\det \phi'(x) > 0$ (resp. $\det \phi'(x) < 0$). The function ϕ is said to *preserve* (resp. *reverse*) *orientation* if and only if ϕ preserves (resp. reverses) orientation at each point x of U.

Theorem 4.1.5. *Every reflection of E^n in a hyperplane or sphere is conformal and reverses orientation.*

Proof: Let ρ be the reflection of E^n in the plane $P(a, t)$. Then

$$\begin{aligned} \rho(x) &= x + 2(t - a \cdot x)a, \\ \rho'(x) &= (\delta_{ij} - 2a_i a_j) = I - 2A, \end{aligned}$$

where A is the matrix $(a_i a_j)$. As $\rho'(x)$ is independent of t, we may assume without loss of generality that $t = 0$. Then ρ is an orthogonal transformation and

$$\rho(x) = (I - 2A)x.$$

Thus $I - 2A$ is an orthogonal matrix, and so ρ is conformal.

By Theorem 1.3.4, there is an orthogonal transformation ϕ such that $\phi(a) = e_1$. Then

$$\begin{aligned} \phi\rho\phi^{-1}(x) &= \phi(\phi^{-1}(x) - 2(a \cdot \phi^{-1}(x))a) \\ &= x - 2(a \cdot \phi^{-1}(x))e_1 \\ &= x - 2(\phi(a) \cdot x)e_1 \\ &= x - 2(e_1 \cdot x)e_1. \end{aligned}$$

Therefore $\phi\rho\phi^{-1}$ is the reflection in $P(e_1, 0)$. By the chain rule,

$$\det(\phi\rho\phi^{-1})'(x) = \det \rho'(x).$$

To compute the determinant of $\rho'(x)$, we may assume that $a = e_1$. Then

$$I - 2A = \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}.$$

Thus $\det \rho'(x) = -1$, and so ρ reverses orientation.

Let σ_r be the reflection of E^n in the sphere $S(0,r)$. Then

$$\sigma_r(x) = \frac{r^2 x}{|x|^2}$$

and so

$$\sigma_r'(x) = r^2\left(\frac{\delta_{ij}}{|x|^2} - \frac{2x_i x_j}{|x|^4}\right) = \frac{r^2}{|x|^2}(I - 2A),$$

where A is the matrix $(x_i x_j/|x|^2)$. We have already shown that $I - 2A$ is orthogonal, and so σ_r is conformal; moreover σ_r reverses orientation, since

$$\begin{aligned} \det \sigma_r'(x) &= \left(\frac{r}{|x|}\right)^{2n} \det(I - 2A) \\ &= -\left(\frac{r}{|x|}\right)^{2n} < 0. \end{aligned}$$

Now let σ be the reflection with respect to $S(a,r)$ and let τ be the translation by a. Then $\tau'(x) = I$ and $\sigma = \tau\sigma_r\tau^{-1}$. Hence $\sigma'(x) = \sigma_r'(x-a)$. Thus σ is conformal and reverses orientation. □

Exercise 4.1

1. Prove Theorem 4.1.1.

2. Show that the reflections of E^n in the planes $P(a,0)$ and $P(b,0)$ commute if and only if their normal vectors a and b are orthogonal.

3. Show that a real $n \times n$ matrix A preserves angles between nonzero vectors if and only if there is a positive scalar k such that $|Ax| = k|x|$ for all x in E^n.

4. Let U be an open connected subset of E^n and let $\phi : U \to E^n$ be a C^1 function such that $\phi'(x)$ is nonsingular for all x in U. Show that ϕ either preserves orientation or reverses orientation.

5. Let U be an open connected subset of $\mathbb{C}$. Prove that a function $\phi : U \to \mathbb{C}$ is conformal if and only if either ϕ is analytic and $\phi'(z) \neq 0$ for all z in U or $\overline{\phi}$ is analytic and $\overline{\phi}'(z) \neq 0$ for all z in U.

§4.2. Stereographic Projection

Identify E^n with $E^n \times \{0\}$ in E^{n+1}. The *stereographic projection* π of E^n onto $S^n - \{e_{n+1}\}$ is defined by projecting x in E^n towards (or away from) e_{n+1} until it meets the sphere S^n in the unique point $\pi(x)$ other than e_{n+1}. See Figure 4.2.1. As $\pi(x)$ is on the line passing through x in the direction of $e_{n+1} - x$, there is a scalar s such that

$$\pi(x) = x + s(e_{n+1} - x).$$

The condition $|\pi(x)|^2 = 1$ leads to the value

$$s = \frac{|x|^2 - 1}{|x|^2 + 1}$$

and the explicit formula

$$\pi(x) = \left(\frac{2x_1}{1+|x|^2}, \ldots, \frac{2x_n}{1+|x|^2}, \frac{|x|^2-1}{|x|^2+1}\right). \tag{4.2.1}$$

The map π is a bijection of E^n onto $S^n - \{e_{n+1}\}$.

There is a nice interpretation of stereographic projection in terms of inversive geometry. Let σ be the reflection of E^{n+1} in the sphere $S(e_{n+1}, \sqrt{2})$. Then

$$\sigma(x) = e_{n+1} + \frac{2(x - e_{n+1})}{|x - e_{n+1}|^2}. \tag{4.2.2}$$

If x is in E^n, then

$$\begin{aligned}\sigma(x) &= e_{n+1} + \frac{2}{1+|x|^2}(x_1, \ldots, x_n, -1) \\ &= \left(\frac{2x_1}{1+|x|^2}, \ldots, \frac{2x_n}{1+|x|^2}, \frac{|x|^2-1}{|x|^2+1}\right).\end{aligned}$$

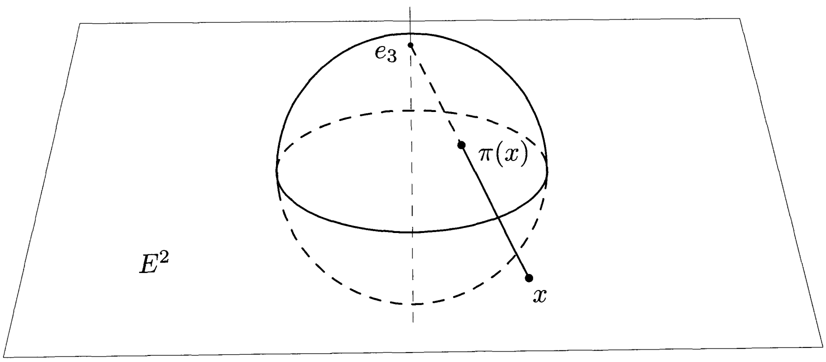


Figure 4.2.1. The stereographic projection π of E^2 into S^2

Thus, the restriction of σ to E^n is stereographic projection

$$\pi : E^n \to S^n - \{e_{n+1}\}.$$

As σ is its own inverse, we can compute the inverse of π from Formula 4.2.2. If y is in $S^n - \{e_{n+1}\}$, then

$$\begin{aligned}\sigma(y) &= e_{n+1} + \frac{2(y - e_{n+1})}{|y|^2 - 2y \cdot e_{n+1} + 1} \\ &= e_{n+1} + \frac{1}{1 - y_{n+1}}(y_1, \ldots, y_n, y_{n+1} - 1) \\ &= \left(\frac{y_1}{1 - y_{n+1}}, \ldots, \frac{y_n}{1 - y_{n+1}}, 0\right).\end{aligned}$$

Hence

$$\pi^{-1}(y) = \left(\frac{y_1}{1 - y_{n+1}}, \ldots, \frac{y_n}{1 - y_{n+1}}\right). \tag{4.2.3}$$

Let ∞ be a point not in E^{n+1} and define $\hat{E}^n = E^n \cup \{\infty\}$. Now extend π to a bijection $\hat{\pi} : \hat{E}^n \to S^n$ by setting $\hat{\pi}(\infty) = e_{n+1}$, and define a metric d on $\hat{E}^n$ by the formula

$$d(x, y) = |\hat{\pi}(x) - \hat{\pi}(y)|. \tag{4.2.4}$$

The metric d is called the *chordal metric* on $\hat{E}^n$. By definition, the map $\hat{\pi}$ is an isometry from $\hat{E}^n$, with the chordal metric, to S^n with the Euclidean metric. The metric topology on E^n determined by the chordal metric is the same as the Euclidean topology, since π maps E^n homeomorphically onto the open subset $S^n - \{e_{n+1}\}$ of S^n. The metric space $\hat{E}^n$ is compact and is obtained from E^n by adjoining one point at infinity. For this reason, $\hat{E}^n$ is called the *one-point compactification* of E^n. The one-point compactification of the complex plane $\mathbb{C}$ is called the *Riemann sphere* $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

Theorem 4.2.1. *If x, y are in E^n, then*

(1) $$d(x, \infty) = \frac{2}{(1 + |x|^2)^{1/2}},$$

(2) $$d(x, y) = \frac{2|x - y|}{(1 + |x|^2)^{1/2}(1 + |y|^2)^{1/2}}.$$

Proof: (1) Observe that

$$\begin{aligned}d(x, \infty) &= |\hat{\pi}(x) - \hat{\pi}(\infty)| \\ &= |\pi(x) - e_{n+1}| \\ &= \left|\left(\frac{2x_1}{1 + |x|^2}, \ldots, \frac{2x_n}{1 + |x|^2}, \frac{-2}{1 + |x|^2}\right)\right| \\ &= \frac{2}{(1 + |x|^2)^{1/2}}.\end{aligned}$$

(2) By Theorem 4.1.3, we have

$$\begin{aligned} d(x,y) &= \frac{2|x-y|}{|x-e_{n+1}|\ |y-e_{n+1}|} \\ &= \frac{2|x-y|}{(1+|x|^2)^{1/2}(1+|y|^2)^{1/2}}. \end{aligned}$$ □

By Theorem 4.2.1, the distance $d(x,\infty)$ depends only on $|x|$. Consequently, every open ball $B_d(\infty,r)$ is of the form $\hat{E}^n - \overline{B(0,s)}$ for some $s > 0$. Therefore, a basis for the topology of $\hat{E}^n$ consists of all the open balls $B(x,r)$ of E^n together with all the neighborhoods of ∞ of the form

$$N(\infty,s) = \hat{E}^n - \overline{B(0,s)}.$$

In particular, this implies that a function $f : \hat{E}^n \to \hat{E}^n$ is continuous at a point a of $\hat{E}^n$ if and only $\lim_{x\to a} f(x) = f(a)$ in the usual Euclidean sense.

Let $P(a,t)$ be a hyperplane of E^n. Define

$$\hat{P}(a,t) = P(a,t) \cup \{\infty\}.$$

Note that the subspace $\hat{P}(a,t)$ of $\hat{E}^n$ is homeomorphic to S^{n-1}. Let ρ be the reflection of E^n in $P(a,t)$ and let $\hat{\rho} : \hat{E}^n \to \hat{E}^n$ be the extension of ρ obtained by setting $\hat{\rho}(\infty) = \infty$. Then $\hat{\rho}(x) = x$ for all x in $\hat{P}(a,t)$ and $\hat{\rho}^2$ is the identity. The map $\hat{\rho}$ is called the *reflection* of $\hat{E}^n$ in the extended hyperplane $\hat{P}(a,t)$.

Theorem 4.2.2. *Every reflection of $\hat{E}^n$ in an extended hyperplane is a homeomorphism.*

Proof: Let ρ be the reflection of E^n in a hyperplane. Then ρ is continuous. As $\lim_{x\to\infty} \rho(x) = \infty$, we have that $\hat{\rho}$ is continuous at ∞. Therefore $\hat{\rho}$ is a continuous function. As $\hat{\rho}$ is its own inverse, it is a homeomorphism. □

Let σ be the reflection of E^n in the sphere $S(a,r)$. Extend σ to a map $\hat{\sigma} : \hat{E}^n \to \hat{E}^n$ by setting $\hat{\sigma}(a) = \infty$ and $\hat{\sigma}(\infty) = a$. Then $\hat{\sigma}(x) = x$ for all x in $S(a,r)$ and $\hat{\sigma}^2$ is the identity. The map $\hat{\sigma}$ is called the *reflection* of $\hat{E}^n$ in the sphere $S(a,r)$.

Theorem 4.2.3. *Every reflection of $\hat{E}^n$ in a sphere of E^n is a homeomorphism.*

Proof: Let σ be the reflection of E^n in the sphere $S(a,r)$ and let $\hat{\sigma}$ be the extended reflection of $\hat{E}^n$. As $\hat{\sigma}^2$ is the identity, $\hat{\sigma}$ is a bijection with inverse $\hat{\sigma}$. The map $\hat{\sigma}$ is continuous, since σ is continuous, $\lim_{x\to a} \sigma(x) = \infty$, and $\lim_{x\to\infty} \sigma(x) = a$. Thus $\hat{\sigma}$ is a homeomorphism. □

Cross Ratio

Let u, v, x, y be points of $\hat{E}^n$ such that $u \neq v$ and $x \neq y$. The *cross ratio* of these points is defined to be the real number

$$[u, v, x, y] = \frac{d(u,x)d(v,y)}{d(u,v)d(x,y)}. \tag{4.2.5}$$

The cross ratio is a continuous function of four variables, since the metric $d : \hat{E}^n \times \hat{E}^n \to \mathbb{R}$ is a continuous function. The following theorem follows immediately from Theorem 4.2.1.

Theorem 4.2.4. *If u, v, x, y are points of E^n such that $u \neq v$ and $x \neq y$, then*

$$(1) \qquad [u, v, x, y] = \frac{|u-x|\,|v-y|}{|u-v|\,|x-y|},$$

$$(2) \qquad [\infty, v, x, y] = \frac{|v-y|}{|x-y|},$$

$$(3) \qquad [u, \infty, x, y] = \frac{|u-x|}{|x-y|},$$

$$(4) \qquad [u, v, \infty, y] = \frac{|v-y|}{|u-v|},$$

$$(5) \qquad [u, v, x, \infty] = \frac{|u-x|}{|u-v|}.$$

Exercise 4.2

1. Derive Formula 4.2.1.
2. Let U be a subset of $\hat{E}^n$ containing ∞. Show that U is open in $\hat{E}^n$ if and only if U is of the form $\hat{E}^n - K$, where K is a compact subset of E^n.
3. Let $\eta : E^n \to E^n$ be a homeomorphism and let $\hat{\eta} : \hat{E}^n \to \hat{E}^n$ be the extension obtained by setting $\hat{\eta}(\infty) = \infty$. Prove that $\hat{\eta}$ is a homeomorphism.
4. Prove that the Euclidean metric on E^n does not extend to a metric $\hat{d}$ on $\hat{E}^n$ so that the metric space $(\hat{E}^n, \hat{d})$ is compact or connected.
5. Let $P(a, t)$ be a hyperplane of E^n. Show that the extended plane $\hat{P}(a, t)$ is homeomorphic to S^{n-1}.

§4.3. Möbius Transformations

A *sphere* Σ of $\hat{E}^n$ is defined to be either a Euclidean sphere $S(a, r)$ or an extended plane $\hat{P}(a, t) = P(a, t) \cup \{\infty\}$. It is worth noting that $\hat{P}(a, t)$ is topologically a sphere.

Definition: A *Möbius transformation* of $\hat{E}^n$ is a finite composition of reflections of $\hat{E}^n$ in spheres.

Let $\mathrm{M}(\hat{E}^n)$ be the set of all Möbius transformations of $\hat{E}^n$. Then $\mathrm{M}(\hat{E}^n)$ obviously forms a group under composition. By Theorem 4.1.2, every isometry of E^n extends in a unique way to a Möbius transformation of $\hat{E}^n$. Thus, we may regard the group of Euclidean isometries $\mathrm{I}(E^n)$ as a subgroup of $\mathrm{M}(\hat{E}^n)$.

Let k be a positive constant and let $\mu_k : \hat{E}^n \to \hat{E}^n$ be the function defined by $\mu_k(x) = kx$. Then μ_k is a Möbius transformation, since μ_k is the composite of the reflection in $S(0,1)$ followed by the reflection in $S(0,\sqrt{k})$. As every similarity of E^n is the composite of an isometry followed by μ_k for some k, every similarity of E^n extends in a unique way to a Möbius transformation of $\hat{E}^n$. Thus, we may also regard the group of Euclidean similarities $\mathrm{S}(E^n)$ as a subgroup of $\mathrm{M}(\hat{E}^n)$.

In order to simplify notation, we shall no longer use a hat to denote the extension of a map to $\hat{E}^n$.

Lemma 1. *If σ is the reflection of $\hat{E}^n$ in the sphere $S(a,r)$ and σ_1 is the reflection in $S(0,1)$, and $\phi : \hat{E}^n \to \hat{E}^n$ is defined by $\phi(x) = a + rx$, then $\sigma = \phi\sigma_1\phi^{-1}$.*

Proof: Observe that

$$\begin{aligned}
\sigma(x) &= a + \left(\frac{r}{|x-a|}\right)^2 (x-a) \\
&= \phi\left(\frac{r(x-a)}{|x-a|^2}\right) \\
&= \phi\sigma_1\left(\frac{x-a}{r}\right) \quad = \quad \phi\sigma_1\phi^{-1}(x).
\end{aligned}$$

□

Theorem 4.3.1. *A function $\phi : \hat{E}^n \to \hat{E}^n$ is a Möbius transformation if and only if it preserves cross ratios.*

Proof: Let ϕ be a Möbius transformation. As ϕ is a composition of reflections, we may assume that ϕ is a reflection. A Euclidean similarity obviously preserves cross ratios, and so we may assume by Lemma 1 that $\phi(x) = x/|x|^2$. By Theorem 4.1.3, we have

$$|\phi(x) - \phi(y)| = \frac{|x-y|}{|x|\,|y|}.$$

By Theorem 4.2.4, we deduce that

$$[\phi(u), \phi(v), \phi(x), \phi(y)] = [u, v, x, y]$$

if u, v, x, y are all finite and nonzero. The remaining cases follow by continuity. Thus ϕ preserves cross ratios.

Conversely, suppose that ϕ preserves cross ratios. By composing ϕ with a Möbius transformation, we may assume that $\phi(\infty) = \infty$. Let u, v, x, y be points of E^n such that $u \neq v$, $x \neq y$, and $(u, v) \neq (x, y)$. Then either $u \neq x$ or $v \neq y$. Assume first that $u \neq x$. As $[\phi(u), \infty, \phi(x), \phi(y)] = [u, \infty, x, y]$, we have

$$\frac{|\phi(u) - \phi(x)|}{|\phi(x) - \phi(y)|} = \frac{|u - x|}{|x - y|},$$

and since $[\phi(u), \phi(v), \phi(x), \infty] = [u, v, x, \infty]$, we have

$$\frac{|\phi(u) - \phi(x)|}{|\phi(u) - \phi(v)|} = \frac{|u - x|}{|u - v|}.$$

Hence

$$\frac{|\phi(u) - \phi(v)|}{|u - v|} = \frac{|\phi(u) - \phi(x)|}{|u - x|} = \frac{|\phi(x) - \phi(y)|}{|x - y|}.$$

Similarly, if $v \neq y$, then

$$\frac{|\phi(u) - \phi(v)|}{|u - v|} = \frac{|\phi(x) - \phi(y)|}{|x - y|}.$$

Hence, there is a positive constant k such that $|\phi(x) - \phi(y)| = k|x - y|$ for all x, y in E^n. By Theorem 1.3.6, we have that ϕ is a Euclidean similarity. Thus ϕ is a Möbius transformation. □

From the proof of Theorem 4.3.1, we deduce the following theorem.

Theorem 4.3.2. *A Möbius transformation ϕ of $\hat{E}^n$ fixes ∞ if and only if ϕ is a similarity of E^n.*

The Isometric Sphere

Let ϕ be a Möbius transformation of $\hat{E}^n$ with $\phi(\infty) \neq \infty$. Let $a = \phi^{-1}(\infty)$ and let σ be the reflection of $\hat{E}^n$ in the sphere $S(a, r)$. Then $\phi\sigma$ fixes ∞. Hence $\phi\sigma$ is a similarity of E^n by Theorem 4.3.2. Therefore, there is a point b of E^n, a scalar $k > 0$, and an orthogonal transformation A of E^n such that

$$\phi(x) = b + kA\sigma(x). \tag{4.3.1}$$

By Theorem 4.1.3, we have

$$|\phi(x) - \phi(y)| = \frac{kr^2|x - y|}{|x - a|\ |y - a|}.$$

Now suppose that x, y are in $S(a, t)$. Then $|\phi(x) - \phi(y)| = |x - y|$ if and only if $t = r\sqrt{k}$. Thus ϕ acts as an isometry on the sphere $S(a, r\sqrt{k})$, and $S(a, r\sqrt{k})$ is unique with this property among the spheres of E^n centered at a. For this reason, $S(a, r\sqrt{k})$ is called the *isometric sphere* of ϕ.

Theorem 4.3.3. *Let ϕ be a Möbius transformation of $\hat{E}^n$ with $\phi(\infty) \neq \infty$. Then there is a unique reflection σ in a Euclidean sphere Σ and a unique Euclidean isometry ψ such that $\phi = \psi\sigma$. Moreover Σ is the isometric sphere of ϕ.*

Proof: Let σ be the reflection in the isometric sphere $S(a,r)$ of ϕ. Then $a = \phi^{-1}(\infty)$ and $\phi\sigma(\infty) = \infty$. By Theorem 4.3.2, we have that $\phi\sigma$ is a Euclidean similarity. Let x, y be in $S(a,r)$. Then we have

$$|\phi\sigma(x) - \phi\sigma(y)| = |\phi(x) - \phi(y)| = |x - y|.$$

Thus $\psi = \phi\sigma$ is a Euclidean isometry and $\phi = \psi\sigma$.

Conversely, suppose that σ is a reflection in a sphere $S(a,r)$ and ψ is a Euclidean isometry such that $\phi = \psi\sigma$. Then $\phi(a) = \infty$ and ϕ acts as an isometry on $S(a,r)$. Therefore $S(a,r)$ is the isometric sphere of ϕ. As $\psi = \phi\sigma$, both σ and ψ are unique. □

Preservation of Spheres

The equation defining a sphere $S(a,r)$ or $\hat{P}(a,t)$ in $\hat{E}^n$ is

$$|x|^2 - 2a \cdot x + |a|^2 - r^2 = 0 \tag{4.3.2}$$

or

$$-\,2a \cdot x + 2t = 0, \tag{4.3.3}$$

respectively, and these can be written in the common form

$$a_0|x|^2 - 2a \cdot x + a_{n+1} = 0 \quad \text{with } |a|^2 > a_0 a_{n+1}.$$

Conversely, any vector $(a_0, \ldots, a_{n+1})$ in $\mathbb{R}^{n+2}$ such that $|a|^2 > a_0 a_{n+1}$, where $a = (a_1, \ldots, a_n)$ determines a sphere Σ of $\hat{E}^n$ satisfying the equation

$$a_0|x|^2 - 2a \cdot x + a_{n+1} = 0.$$

If $a_0 \neq 0$, then

$$\Sigma \;=\; S\left(\frac{a}{a_0}, \frac{(|a|^2 - a_0 a_{n+1})^{\frac{1}{2}}}{|a_0|}\right).$$

If $a_0 = 0$, then

$$\Sigma \;=\; \hat{P}\left(\frac{a}{|a|}, \frac{a_{n+1}}{2|a|}\right).$$

The vector $(a_0, \ldots, a_{n+1})$ is called a *coefficient vector* for Σ, and it is uniquely determined by Σ up to multiplication by a nonzero scalar.

Theorem 4.3.4. *Let ϕ be a Möbius transformation of $\hat{E}^n$. If Σ is a sphere of $\hat{E}^n$, then $\phi(\Sigma)$ is also a sphere of $\hat{E}^n$.*

Proof: Let ϕ be a Möbius transformation, and let Σ be a sphere. As ϕ is a composition of reflections, we may assume that ϕ is a reflection.

A Euclidean similarity obviously maps spheres to spheres, and so we may assume by Lemma 1 that $\phi(x) = x/|x|^2$.

Let $(a_0, \ldots, a_{n+1})$ be a coefficient vector for Σ. Then Σ satisfies the equation

$$a_0|x|^2 - 2a \cdot x + a_{n+1} = 0.$$

Let $y = \phi(x)$. Then y satisfies the equation

$$a_0 - 2a \cdot y + a_{n+1}|y|^2 = 0.$$

But this is the equation of another sphere Σ'. Hence ϕ maps Σ into Σ'. The same argument shows that ϕ maps Σ' into Σ. Therefore $\phi(\Sigma) = \Sigma'$. □

Theorem 4.3.5. *The natural action of* $\mathrm{M}(\hat{E}^n)$ *on the set of spheres of* $\hat{E}^n$ *is transitive.*

Proof: Let Σ be a sphere of $\hat{E}^n$. It suffices to show that there is a Möbius transformation ϕ such that $\phi(\Sigma) = \hat{E}^{n-1}$. As the group of Euclidean isometries $\mathrm{I}(E^n)$ acts transitively on the set of hyperplanes of E^n, we may assume that Σ is a Euclidean sphere. As the group of Euclidean similarities $\mathrm{S}(E^n)$ acts transitively on the set of spheres of E^n, we may assume that $\Sigma = S^{n-1}$. Let σ be the reflection in the sphere $S(e_n, \sqrt{2})$. Then we have that $\sigma(S^{n-1}) = \hat{E}^{n-1}$ by stereographic projection. □

Theorem 4.3.6. *If ϕ is a Möbius transformation of $\hat{E}^n$ that fixes each point of a sphere Σ of $\hat{E}^n$, then ϕ is either the identity map of $\hat{E}^n$ or the reflection in Σ.*

Proof: Assume first that $\Sigma = \hat{E}^{n-1}$. Then $\phi(\infty) = \infty$. By Theorem 4.3.2, we have that ϕ is a Euclidean similarity. As $\phi(0) = 0$ and $\phi(e_1) = e_1$, we have that ϕ is an orthogonal transformation. Moreover, since ϕ fixes $e_1, \ldots, e_{n-1}$, we have that $\phi(e_n) = \pm e_n$. Thus ϕ is either the identity or the reflection in $P(e_n, 0)$.

Now assume that Σ is arbitrary. By Theorem 4.3.5, there is a Möbius transformation ψ such that $\psi(\Sigma) = \hat{E}^{n-1}$. As $\psi\phi\psi^{-1}$ fixes each point of $\hat{E}^{n-1}$, we find that $\psi\phi\psi^{-1}$ is either the identity or the reflection ρ in $\hat{E}^{n-1}$. Hence ϕ is either the identity or $\psi^{-1}\rho\psi$. Let σ be the reflection in Σ. As $\psi\sigma\psi^{-1}$ fixes each point of $\hat{E}^{n-1}$ and is not the identity, we have that $\psi\sigma\psi^{-1} = \rho$. Hence $\sigma = \psi^{-1}\rho\psi$. Thus ϕ is either the identity or σ. □

Definition: Given a reflection σ in a sphere Σ of $\hat{E}^n$, two points x and y of $\hat{E}^n$ are said to be *inverse points* with respect to Σ if and only if $y = \sigma(x)$.

Theorem 4.3.7. *Let ϕ be a Möbius transformation of $\hat{E}^n$. If x and y are inverse points with respect to a sphere Σ of $\hat{E}^n$, then $\phi(x)$ and $\phi(y)$ are inverse points with respect to $\phi(\Sigma)$.*

Proof: Let σ be the reflection in Σ. Then $\phi\sigma\phi^{-1}$ fixes each point of $\phi(\Sigma)$ and is not the identity. By Theorem 4.3.6, we have that $\phi\sigma\phi^{-1}$ is the reflection in $\phi(\Sigma)$. As $\phi\sigma\phi^{-1}(\phi(x)) = \phi(y)$, we have that $\phi(x)$ and $\phi(y)$ are inverse points with respect to $\phi(\Sigma)$. $\square$

Exercise 4.3

1. Show that a Möbius transformation of $\hat{E}^n$ either preserves or reverses orientation depending on whether it is the composition of an even or odd number of reflections. Let $\mathrm{M}_0(\hat{E}^n)$ be the set of all orientation preserving Möbius transformations of $\hat{E}^n$. Conclude that $\mathrm{M}_0(\hat{E}^n)$ is a subgroup of $\mathrm{M}(\hat{E}^n)$ of index two.

2. A *linear fractional transformation* of the Riemann sphere $\hat{\mathbb{C}}$ is a continuous map $\phi : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ of the form $\phi(z) = \frac{az+b}{cz+d}$, where a, b, c, d are in $\mathbb{C}$ and $ad - bc \neq 0$. Show that every linear fractional transformation of $\hat{\mathbb{C}}$ is an orientation preserving Möbius transformation of $\hat{\mathbb{C}}$.

3. Let $\mathrm{LF}(\hat{\mathbb{C}})$ be the set of all linear fractional transformations of $\hat{\mathbb{C}}$. Show that $\mathrm{LF}(\hat{\mathbb{C}})$ is a group under composition.

4. Let $\mathrm{GL}(2,\mathbb{C})$ be the group of all invertible complex 2×2 matrices, and let $\mathrm{PGL}(2,\mathbb{C})$ be the quotient group of $\mathrm{GL}(2,\mathbb{C})$ by the normal subgroup $\{kI : k \in \mathbb{C}^*\}$. Show that the map $\Xi : \mathrm{GL}(2,\mathbb{C}) \to \mathrm{LF}(\hat{\mathbb{C}})$, defined by
$$\Xi\begin{pmatrix} a & b \\ c & d \end{pmatrix}(z) = \frac{az+b}{cz+d},$$
induces an isomorphism from $\mathrm{PGL}(2,\mathbb{C})$ to $\mathrm{LF}(\hat{\mathbb{C}})$.

5. Let $\rho(z) = \overline{z}$ be complex conjugation. Show that
$$\mathrm{M}(\hat{\mathbb{C}}) = \mathrm{LF}(\hat{\mathbb{C}}) \cup \mathrm{LF}(\hat{\mathbb{C}})\rho.$$
Deduce that $\mathrm{LF}(\hat{\mathbb{C}}) = \mathrm{M}_0(\hat{\mathbb{C}})$.

6. Let $\phi(z) = \frac{az+b}{cz+d}$ be a linear fractional transformation of $\hat{\mathbb{C}}$ with $\phi(\infty) \neq \infty$. Show that the *isometric circle* of ϕ is the set
$$\left\{z \in \mathbb{C} : |cz+d| = |ad-bc|^{\frac{1}{2}}\right\}.$$

7. Let ϕ be a Möbius transformation of $\hat{E}^n$ with $\phi(\infty) \neq \infty$, and let Σ_ϕ be the isometric sphere of ϕ. Prove that $\phi(\Sigma_\phi) = \Sigma_{\phi^{-1}}$.

8. Let ϕ be a Möbius transformation of $\hat{E}^n$ with $\phi(\infty) \neq \infty$, and let $\phi'(x)$ be the matrix of partial derivatives of ϕ. Prove that the isometric sphere of ϕ is the set $\{x \in E^n : \phi'(x) \text{ is orthogonal}\}$.

§4.4. Poincaré Extension

Under the identification of E^{n-1} with $E^{n-1} \times \{0\}$ in E^n, a point x of E^{n-1} corresponds to the point $\tilde{x} = (x, 0)$ of E^n. Let ϕ be a Möbius transformation of $\hat{E}^{n-1}$. We shall extend ϕ to a Möbius transformation $\tilde{\phi}$ of $\hat{E}^n$ as follows. If ϕ is the reflection of $\hat{E}^{n-1}$ in $\hat{P}(a, t)$, then $\tilde{\phi}$ is the reflection of $\hat{E}^n$ in $\hat{P}(\tilde{a}, t)$. If ϕ is the reflection of $\hat{E}^{n-1}$ in $S(a, r)$, then $\tilde{\phi}$ is the reflection of $\hat{E}^n$ in $S(\tilde{a}, r)$. In both these cases

$$\tilde{\phi}(x, 0) = (\phi(x), 0) \quad \text{for all } x \text{ in } E^{n-1}.$$

Thus $\tilde{\phi}$ extends ϕ. In particular $\tilde{\phi}$ leaves $\hat{E}^{n-1}$ invariant. It is also clear that $\tilde{\phi}$ leaves invariant *upper half-space*

$$U^n = \{(x_1, \ldots, x_n) \in E^n : x_n > 0\}.$$

Now assume that ϕ is an arbitrary Möbius transformation of $\hat{E}^{n-1}$. Then ϕ is the composition $\phi = \sigma_1 \cdots \sigma_m$ of reflections. Let $\tilde{\phi} = \tilde{\sigma}_1 \cdots \tilde{\sigma}_m$. Then $\tilde{\phi}$ extends ϕ and leaves U^n invariant. Suppose that $\tilde{\phi}_1$ and $\tilde{\phi}_2$ are two such Möbius transformations. Then $\tilde{\phi}_1\tilde{\phi}_2^{-1}$ fixes each point of $\hat{E}^{n-1}$ and leaves U^n invariant. By Theorem 4.3.6, we have that $\tilde{\phi}_1\tilde{\phi}_2^{-1}$ is the identity and so $\tilde{\phi}_1 = \tilde{\phi}_2$. Thus $\tilde{\phi}$ depends only on ϕ and not on the decomposition $\phi = \sigma_1 \cdots \sigma_m$. The map $\tilde{\phi}$ is called the *Poincaré extension* of ϕ.

Theorem 4.4.1. *A Möbius transformation ϕ of $\hat{E}^n$ leaves upper half-space U^n invariant if and only if ϕ is the Poincaré extension of a Möbius transformation of $\hat{E}^{n-1}$.*

Proof: Let ϕ be a Möbius transformation of $\hat{E}^n$ that leaves U^n invariant. As ϕ is a homeomorphism, it also leaves the boundary of U^n invariant. Hence ϕ restricts to a homeomorphism $\overline{\phi}$ of $\hat{E}^{n-1}$. As ϕ preserves cross ratios in $\hat{E}^n$, we have that $\overline{\phi}$ preserves cross ratios in $\hat{E}^{n-1}$. Therefore $\overline{\phi}$ is a Möbius transformation of $\hat{E}^{n-1}$ by Theorem 4.3.1. Let $\tilde{\overline{\phi}}$ be the Poincaré extension of $\overline{\phi}$. Then $\tilde{\overline{\phi}}\phi^{-1}$ fixes each point of $\hat{E}^{n-1}$ and leaves U^n invariant. Therefore $\phi = \tilde{\overline{\phi}}$ by Theorem 4.3.6. □

Möbius Transformations of Upper Half-Space

Definition: A *Möbius transformation of upper half-space* U^n is a Möbius transformation of $\hat{E}^n$ that leaves U^n invariant.

Let $\mathrm{M}(U^n)$ be the set of all Möbius transformations of U^n. Then $\mathrm{M}(U^n)$ is a subgroup of $\mathrm{M}(\hat{E}^n)$. The next corollary follows immediately from Theorem 4.4.1.

Corollary 1. *The group $\mathrm{M}(U^n)$ of Möbius transformations of U^n is isomorphic to $\mathrm{M}(\hat{E}^{n-1})$.*

Two spheres Σ and Σ' of $\hat{E}^n$ are said to be *orthogonal* if and only if they intersect in E^n and at each point of intersection in E^n their normal lines are orthogonal.

Corollary 2. *Every Möbius transformation of U^n is the composition of reflections of $\hat{E}^n$ in spheres orthogonal to $\hat{E}^{n-1}$.*

Proof: Let ψ be a Möbius transformation of U^n. Then ψ is the Poincaré extension $\tilde{\phi}$ of a Möbius transformation ϕ of $\hat{E}^{n-1}$. The map ϕ is the composition $\sigma_1 \cdots \sigma_m$ of reflections of $\hat{E}^{n-1}$ in spheres. The Poincaré extension of the reflection σ_i is a reflection of $\hat{E}^n$ in a sphere orthogonal to $\hat{E}^{n-1}$. As $\tilde{\phi} = \tilde{\sigma}_1 \cdots \tilde{\sigma}_m$, we have that ψ is the composition of reflections of $\hat{E}^n$ in spheres orthogonal to $\hat{E}^{n-1}$. □

Theorem 4.4.2. *Two spheres of $\hat{E}^n$ are orthogonal under the following conditions:*

1. *The spheres $\hat{P}(a, r)$ and $\hat{P}(b, s)$ are orthogonal if and only if a and b are orthogonal.*
2. *The spheres $S(a, r)$ and $\hat{P}(b, s)$ are orthogonal if and only if a is in $P(b, s)$.*
3. *The spheres $S(a, r)$ and $S(b, s)$ are orthogonal if and only if r and s satisfy the equation $|a - b|^2 = r^2 + s^2$.*

Proof: Part (1) is obvious. The proof of (2) is left to the reader. The proof of (3) goes as follows: At each point of intersection x of $S(a, r)$ and $S(b, s)$, the normal lines have the equations

$$\begin{cases} u = a + t(x - a), \\ v = b + t(x - b), \end{cases}$$

where t is a real parameter. These lines are orthogonal if and only if their direction vectors $x - a$ and $x - b$ are orthogonal. Observe that

$$\begin{aligned} |a - b|^2 &= |(x - b) - (x - a)|^2 \\ &= |x - a|^2 - 2(x - b) \cdot (x - a) + |x - a|^2 \\ &= s^2 - 2(x - b) \cdot (x - a) + r^2. \end{aligned}$$

Hence $(x - a)$ and $(x - b)$ are orthogonal if and only if

$$|a - b|^2 = r^2 + s^2.$$

Thus, if the spheres are orthogonal, then

$$|a - b|^2 = r^2 + s^2.$$

Conversely, suppose that $|a - b|^2 = r^2 + s^2$. Then there is a right triangle in E^n with vertices a, b, x such that $|x - a| = r$ and $|x - b| = s$. Consequently, x is a point of intersection of $S(a, r)$ and $S(b, s)$, and the spheres are orthogonal. See Figure 4.4.1. □

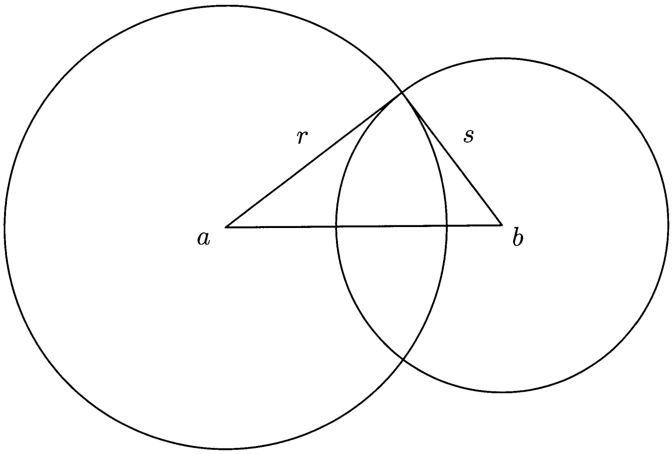


Figure 4.4.1. Orthogonal circles $S(a,r)$ and $S(b,s)$

Remark: It is clear from the proof of Theorem 4.4.2 that two spheres Σ and Σ' of $\hat{E}^n$ are orthogonal if and only if they are orthogonal at a single point of intersection in E^n.

Theorem 4.4.3. *A reflection σ of $\hat{E}^n$ in a sphere Σ leaves upper half-space U^n invariant if and only if $\hat{E}^{n-1}$ and Σ are orthogonal.*

Proof: Let $\Sigma = \hat{P}(a,t)$ or $S(a,r)$. By Theorem 4.4.2, we have that $\hat{E}^{n-1}$ and Σ are orthogonal if and only if $a_n = 0$. Let x be in E^n and set $y = \sigma(x)$. Then for all finite values of y, we have

$$y_n = \begin{cases} x_n + 2(t - a\cdot x)a_n & \text{if } \Sigma = \hat{P}(a,t), \\ \left(\frac{r}{|x-a|}\right)^2 x_n + \left(1 - \left(\frac{r}{|x-a|}\right)^2\right) a_n & \text{if } \Sigma = S(a,r). \end{cases}$$

Assume that $a_n = 0$ and $x_n > 0$. Then $x \neq a$, and so y is finite and $y_n > 0$. Thus σ leaves U^n invariant.

Conversely, assume that σ leaves U^n invariant. Then σ leaves $\hat{E}^{n-1}$ invariant. As the reflection in $\hat{E}^{n-1}$ switches U^n and $-U^n$, we may assume that Σ is not $\hat{E}^{n-1}$. Let x be in $\hat{E}^{n-1} - \Sigma$ with y finite. Then $x_n = 0 = y_n$. As x is not in Σ, the coefficient of a_n in the above expression for y_n is nonzero. Hence $a_n = 0$. □

Theorem 4.4.4. *Let ϕ be a Möbius transformation of U^n. If $\phi(\infty) = \infty$, then ϕ is a Euclidean similarity. If $\phi(\infty) \neq \infty$, then the isometric sphere Σ of ϕ is orthogonal to E^{n-1} and $\phi = \psi\sigma$, where σ is the reflection in Σ and ψ is a Euclidean isometry that leaves U^n invariant.*

Proof: If $\phi(\infty) = \infty$, then ϕ is a Euclidean similarity by Theorem 4.3.2. Now assume that $\phi(\infty) \neq \infty$. Then ϕ is the Poincaré extension of a Möbius transformation $\overline{\phi}$ of $\hat{E}^{n-1}$ by Theorem 4.4.1. Let $\overline{\sigma}$ be the reflection of $\hat{E}^{n-1}$ in the isometric sphere $\overline{\Sigma}$ of $\overline{\phi}$. Then there is a Euclidean isometry $\overline{\psi}$ of $\hat{E}^{n-1}$ such that $\overline{\phi} = \overline{\psi}\overline{\sigma}$ by Theorem 4.3.3. Let σ, ψ be the Poincaré extensions of $\overline{\sigma}, \overline{\psi}$, respectively. Then σ is a reflection in a sphere Σ of E^n orthogonal to $\hat{E}^{n-1}$, and ψ is an isometry of E^n that leaves U^n invariant. As $\overline{\phi} = \overline{\psi}\overline{\sigma}$, we have that $\phi = \psi\sigma$. Therefore Σ is the isometric sphere of ϕ by Theorem 4.3.3. □

Möbius Transformations of the Unit n-Ball

Let σ be the reflection of $\hat{E}^n$ in the sphere $S(e_n, \sqrt{2})$. Then

$$\sigma(x) \quad = \quad e_n + \frac{2(x - e_n)}{|x - e_n|^2}. \tag{4.4.1}$$

Therefore

$$|\sigma(x)|^2 \quad = \quad 1 + \frac{4e_n \cdot (x - e_n)}{|x - e_n|^2} + \frac{4}{|x - e_n|^2}.$$

Thus

$$|\sigma(x)|^2 \quad = \quad 1 + \frac{4x_n}{|x - e_n|^2}. \tag{4.4.2}$$

This implies that σ maps lower half-space $-U^n$ into the *open unit n-ball*

$$B^n = \{x \in E^n : |x| < 1\}.$$

As σ is a homeomorphism of $\hat{E}^n$, it maps each component of $\hat{E}^n - \hat{E}^{n-1}$ homeomorphically onto a component of $\hat{E}^n - S^{n-1}$. Thus σ maps $-U^n$ homeomorphically onto B^n and vice versa.

Let ρ be the reflection of $\hat{E}^n$ in $\hat{E}^{n-1}$ and define $\eta = \sigma\rho$. Then η maps U^n homeomorphically onto B^n. The Möbius transformation η is called the *standard transformation* from U^n to B^n.

Definition: A *Möbius transformation* of S^n is a function $\phi : S^n \to S^n$ such that $\pi^{-1}\phi\pi$ is a Möbius transformation of $\hat{E}^n$, where $\pi : \hat{E}^n \to S^n$ is stereographic projection.

Let $\mathrm{M}(S^n)$ be the set of all Möbius transformations of S^n. Then $\mathrm{M}(S^n)$ forms a group under composition. The mapping $\psi \mapsto \pi\psi\pi^{-1}$ is an isomorphism from $\mathrm{M}(\hat{E}^n)$ to $\mathrm{M}(S^n)$.

Let ϕ be a Möbius transformation of S^{n-1}. The *Poincaré extension* of ϕ is the Möbius transformation $\tilde{\phi}$ of $\hat{E}^n$ defined by $\tilde{\phi} = \eta\tilde{\psi}\eta^{-1}$, where $\tilde{\psi}$ is the Poincaré extension of $\psi = \pi^{-1}\phi\pi$ and η is the standard transformation from U^n to B^n. The Möbius transformation $\tilde{\phi}$ obviously extends ϕ and leaves B^n invariant; moreover, $\tilde{\phi}$ is unique with this property. The following theorem follows immediately from Theorem 4.4.1.

Theorem 4.4.5. *A Möbius transformation ϕ of $\hat{E}^n$ leaves the open unit ball B^n invariant if and only if ϕ is the Poincaré extension of a Möbius transformation of S^{n-1}.*

Definition: A *Möbius transformation of the open unit ball* B^n is a Möbius transformation of $\hat{E}^n$ that leaves B^n invariant.

Let $\mathrm{M}(B^n)$ be the set of all Möbius transformations of B^n. Then $\mathrm{M}(B^n)$ is a subgroup of $\mathrm{M}(\hat{E}^n)$. The next corollary follows immediately from Theorem 4.4.5.

Corollary 3. *The group $\mathrm{M}(B^n)$ of Möbius transformations of B^n is isomorphic to $\mathrm{M}(S^{n-1})$.*

The following corollary follows immediately from Corollary 2.

Corollary 4. *Every Möbius transformation of B^n is the composition of reflections of $\hat{E}^n$ in spheres orthogonal to S^{n-1}.*

Theorem 4.4.6. *A reflection σ of $\hat{E}^n$ in a sphere Σ leaves the open unit ball B^n invariant if and only if S^{n-1} and Σ are orthogonal.*

Proof: Let η be the standard transformation from U^n to B^n. Then $\Sigma' = \eta^{-1}(\Sigma)$ is a sphere of $\hat{E}^n$ by Theorem 4.3.4, and $\sigma' = \eta^{-1}\sigma\eta$ is the reflection in Σ' by Theorem 4.3.6. As η maps U^n bijectively onto B^n, the map σ leaves B^n invariant if and only if σ' leaves U^n invariant. By Theorem 4.4.3, this is the case if and only if $\hat{E}^{n-1}$ and Σ' are orthogonal. By Theorem 4.1.5, the map η is conformal and so it preserves angles. Hence $\hat{E}^{n-1}$ and Σ' are orthogonal if and only if S^{n-1} and Σ are orthogonal. □

Theorem 4.4.7. *Let ϕ be a Möbius transformation of B^n. If $\phi(\infty) = \infty$, then ϕ is orthogonal. If $\phi(\infty) \neq \infty$, then the isometric sphere Σ of ϕ is orthogonal to S^{n-1} and $\phi = \psi\sigma$, where σ is the reflection in Σ and ψ is an orthogonal transformation.*

Proof: Assume first that $\phi(\infty) = \infty$. Then ϕ is a Euclidean similarity by Theorem 4.3.2. As $\phi(0) = 0$, we have that $\phi(x) = kAx$, where $k > 0$ and A is an orthogonal matrix. As ϕ leaves S^{n-1} invariant, we must have that $k = 1$. Thus ϕ is orthogonal.

Now assume that $\phi(\infty) \neq \infty$. Let σ be the reflection in the sphere $S(a,r)$, where $a = \phi^{-1}(\infty)$ and $r^2 = 1 - |a|^2$. Then $S(a,r)$ is orthogonal to S^{n-1} by Theorem 4.4.2. Hence σ leaves B^n invariant by Theorem 4.4.6. Now $\phi\sigma(\infty) = \phi(a) = \infty$. Hence $\phi\sigma$ is an orthogonal transformation ψ, and $\phi = \psi\sigma$. By Theorem 4.3.3, the isometric sphere of ϕ is $S(a,r)$. □

Theorem 4.4.8. *Let ϕ be a Möbius transformation of B^n. Then $\phi(0) = 0$ if and only if ϕ is an orthogonal transformation of E^n.*

Proof: As 0 and ∞ are inverse points with respect to S^{n-1}, and ϕ leaves S^{n-1} invariant, $\phi(0)$ and $\phi(\infty)$ are inverse points with respect to S^{n-1}. Therefore ϕ fixes 0 if and only it fixes ∞. The theorem now follows from Theorem 4.4.7. □

Exercise 4.4

1. Identify the upper half-plane U^2 with the set of complex numbers
$$\{z \in \mathbb{C} : \ \mathrm{Im}\, z > 0\}.$$
Show that a linear fractional transformation ϕ of $\hat{\mathbb{C}}$ leaves U^2 invariant if and only if there exists real numbers a, b, c, d, with $ad - bc > 0$, such that
$$\phi(z) = \frac{az + b}{cz + d}.$$

2. Let ϕ be in $\mathrm{LF}(\hat{\mathbb{C}})$. Show that there are complex numbers a, b, c, d such that $\phi(z) = \frac{az+b}{cz+d}$ and $ad - bc = 1$.

3. Let $\mathrm{SL}(2, \mathbb{C})$ be the group of all complex 2×2 matrices of determinant one, and let $\mathrm{PSL}(2, \mathbb{C})$ be the quotient of $\mathrm{SL}(2, \mathbb{C})$ by the normal subgroup $\{\pm I\}$. Show that the inclusion of $\mathrm{SL}(2, \mathbb{C})$ into $\mathrm{GL}(2, \mathbb{C})$ induces an isomorphism from $\mathrm{PSL}(2, \mathbb{C})$ to $\mathrm{PGL}(2, \mathbb{C})$. Deduce that $\mathrm{PSL}(2, \mathbb{C})$ and $\mathrm{LF}(\hat{\mathbb{C}})$ are isomorphic groups.

4. Show that the standard transformation $\eta : U^2 \to B^2$ is given by
$$\eta(z) = \frac{iz + 1}{z + i}.$$

5. Identify the open unit disk B^2 with the open unit disk in $\mathbb{C}$,
$$\{z \in \mathbb{C} : |z| < 1\}.$$
Let $\phi(z) = \frac{az+b}{cz+d}$ be in $\mathrm{LF}(\hat{\mathbb{C}})$ normalized so that $ad - bc = 1$. Show that ϕ leaves B^2 invariant if and only if $c = \overline{b}$ and $d = \overline{a}$.

6. Identify upper half-space U^3 with the set of quaternions
$$\{z + tj : z \in \mathbb{C} \ \text{ and } \ t > 0\}.$$
Let $\phi(z) = \frac{az+b}{cz+d}$ be a linear fractional transformation of $\hat{\mathbb{C}}$ normalized so that $ad - bc = 1$. Show that the Poincaré extension of ϕ is given by
$$\tilde{\phi}(w) = (aw + b)(cw + d)^{-1}, \ \text{ where } \ w = z + tj.$$

7. Prove that Poincaré extension induces a monomorphism
$$\Upsilon : \mathrm{M}(B^{n-1}) \to \mathrm{M}(B^n)$$
mapping $\mathrm{M}(B^{n-1})$ onto the subgroup $\tilde{\mathrm{M}}(B^{n-1})$ of elements of $\mathrm{M}(B^n)$ that leave B^{n-1} and each component of $B^n - B^{n-1}$ invariant.

§4.5. The Conformal Ball Model

Henceforth, we shall work with hyperbolic n-space H^n in $\mathbb{R}^{n,1}$. We now redefine the Lorentzian inner product on $\mathbb{R}^{n+1}$ to be

$$x \circ y = x_1y_1 + \cdots + x_ny_n - x_{n+1}y_{n+1}. \tag{4.5.1}$$

All the results of Chapter 3 remain true after one reverses the order of the coordinates of $\mathbb{R}^{n+1}$. The Lorentz group of $\mathbb{R}^{n,1}$ is denoted by $\mathrm{O}(n,1)$.

Identify $\mathbb{R}^n$ with $\mathbb{R}^n \times \{0\}$ in $\mathbb{R}^{n+1}$. The *stereographic projection* ζ of the open unit ball B^n onto hyperbolic space H^n is defined by projecting x in B^n away from $-e_{n+1}$ until it meets H^n in the unique point $\zeta(x)$. See Figure 4.5.1. As $\zeta(x)$ is on the line passing through x in the direction of $x + e_{n+1}$, there is a scalar s such that

$$\zeta(x) = x + s(x + e_{n+1}).$$

The condition $\|\zeta(x)\|^2 = -1$ leads to the value

$$s = \frac{1+|x|^2}{1-|x|^2}$$

and the explicit formula

$$\zeta(x) = \left(\frac{2x_1}{1-|x|^2}, \ldots, \frac{2x_n}{1-|x|^2}, \frac{1+|x|^2}{1-|x|^2}\right). \tag{4.5.2}$$

The map ζ is a bijection of B^n onto H^n. The inverse of ζ is given by

$$\zeta^{-1}(y) = \left(\frac{y_1}{1+y_{n+1}}, \ldots, \frac{y_n}{1+y_{n+1}}\right). \tag{4.5.3}$$

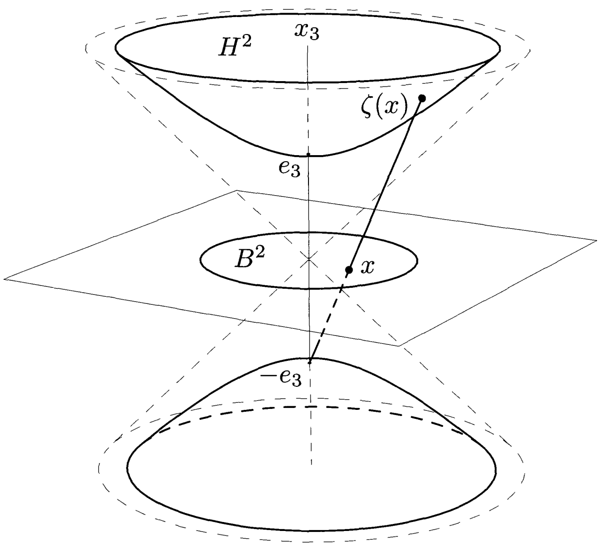


Figure 4.5.1. The stereographic projection ζ of B^2 onto H^2

Define a metric d_B on B^n by the formula

$$d_B(x,y) = d_H(\zeta(x), \zeta(y)). \tag{4.5.4}$$

The metric d_B is called the *Poincaré metric* on B^n. By definition, ζ is an isometry from B^n, with the metric d_B, to hyperbolic n-space H^n. The metric space consisting of B^n together with the metric d_B is called the *conformal ball model* of hyperbolic n-space.

Theorem 4.5.1. *The metric d_B on B^n is given by*

$$\cosh d_B(x,y) = 1 + \frac{2|x-y|^2}{(1-|x|^2)(1-|y|^2)}.$$

Proof: By Formula 3.2.2, we have

$$\begin{aligned}
\cosh d_H(\zeta(x), \zeta(y)) &= -\zeta(x) \circ \zeta(y) \\
&= \frac{-4x\cdot y + (1+|x|^2)(1+|y|^2)}{(1-|x|^2)(1-|y|^2)} \\
&= \frac{(1-|x|^2)(1-|y|^2) + 2(|x|^2+|y|^2) - 4x\cdot y}{(1-|x|^2)(1-|y|^2)} \\
&= 1 + \frac{2|x-y|^2}{(1-|x|^2)(1-|y|^2)}.
\end{aligned}$$

□

Lemma 1. *If ϕ is a Möbius transformation of B^n and x, y are in B^n, then*

$$\frac{|\phi(x)-\phi(y)|^2}{(1-|\phi(x)|^2)(1-|\phi(y)|^2)} = \frac{|x-y|^2}{(1-|x|^2)(1-|y|^2)}.$$

Proof: This is obvious if ϕ is an orthogonal transformation. By Theorem 4.4.6, we may assume that ϕ is a reflection in a sphere $S(a,r)$ orthogonal to S^{n-1}. By Theorem 4.1.3, we have

$$\frac{|\phi(x)-\phi(y)|}{|x-y|} = \frac{r^2}{|x-a|\;|y-a|}.$$

As $S(a,r)$ is orthogonal to S^{n-1}, we have that $r^2 = |a|^2 - 1$. Moreover

$$\phi(x) = a + \frac{r^2}{|x-a|^2}(x-a).$$

Hence

$$|\phi(x)|^2 = |a|^2 + \frac{2r^2}{|x-a|^2} a\cdot(x-a) + \frac{r^4}{|x-a|^2}.$$

Thus

$$\begin{aligned}
|\phi(x)|^2 - 1 &= \frac{(|a|^2-1)|x-a|^2 + 2r^2 a\cdot(x-a) + r^4}{|x-a|^2} \\
&= \frac{r^2[|x-a|^2 + 2a\cdot(x-a) + |a|^2 - 1]}{|x-a|^2} \\
&= \frac{r^2(|x|^2-1)}{|x-a|^2}.
\end{aligned}$$

Hence

$$\frac{1-|\phi(x)|^2}{1-|x|^2} = \frac{r^2}{|x-a|^2}.$$

Therefore

$$\frac{|\phi(x)-\phi(y)|^2}{|x-y|^2} = \frac{(1-|\phi(x)|^2)(1-|\phi(y)|^2)}{(1-|x|^2)(1-|y|^2)}. \qquad \square$$

Hyperbolic Translation

Let $S(a,r)$ be a sphere of E^n orthogonal to S^{n-1}. By Theorem 4.4.2, we have $r^2 = |a|^2 - 1$. Thus, the radius r is a function of a. Let σ_a be the the reflection in $S(a,r)$. Then σ_a leaves B^n invariant by Theorem 4.4.6. Let ρ_a be the reflection in the hyperplane $a \cdot x = 0$. Then ρ_a also leaves B^n invariant, and therefore the composite $\rho_a\sigma_a$ leaves B^n invariant. Let $a^* = a/|a|^2$. A straightforward calculation shows that

$$\rho_a\sigma_a(x) = \frac{(|a|^2-1)}{|x-a|^2}x - \frac{(|x|^2 - 2x\cdot a^* + 1)}{|x-a|^2}a.$$

In particular $\rho_a\sigma_a(0) = -a^*$.

Let b be a nonzero point of B^n and set $b' = -b^*$. By Theorem 4.4.2, the sphere $S(b', (|b'|^2-1)^{1/2})$ is orthogonal to S^{n-1}. Hence, we may define a Möbius transformation of B^n by the formula $\tau_b = \rho_{b'}\sigma_{b'}$. Then

$$\tau_b(x) = \frac{(|b^*|^2-1)}{|x+b^*|^2}x + \frac{(|x|^2+2x\cdot b+1)}{|x+b^*|^2}b^*.$$

In terms of b, we have the formula

$$\tau_b(x) = \frac{(1-|b|^2)}{(|b|^2|x|^2+2x\cdot b+1)}x + \frac{(|x|^2+2x\cdot b+1)}{(|b|^2|x|^2+2x\cdot b+1)}b. \tag{4.5.5}$$

As τ_b is the composite of two reflections in hyperplanes orthogonal to the line $(-b/|b|, b/|b|)$, the transformation τ_b acts as a translation along this line. We also define τ_0 to be the identity. Then $\tau_b(0) = b$ for all b in B^n. The map τ_b is called the *hyperbolic translation* of B^n by b.

Theorem 4.5.2. *Every Möbius transformation of B^n restricts to an isometry of the conformal ball model B^n, and every isometry of B^n extends to a unique Möbius transformation of B^n.*

Proof: That every Möbius transformation of B^n restricts to an isometry of B^n follows immediately from Theorem 4.5.1 and Lemma 1. Conversely, let $\phi: B^n \to B^n$ be an isometry. Define $\psi: B^n \to B^n$ by $\psi(x) = \tau^{-1}_{\phi(0)}\phi(x)$. Then $\psi(0) = 0$. By the first part of the theorem, ψ is an isometry of B^n.

Let x, y be points of B^n. From the relation

$$d_B(\psi(x), 0) = d_B(x, 0)$$

and Theorem 4.5.1, we have

$$\frac{|\psi(x)|^2}{1-|\psi(x)|^2} = \frac{|x|^2}{1-|x|^2}.$$

Hence $|\psi(x)| = |x|$. Likewise, we have

$$\frac{|\psi(x)-\psi(y)|^2}{(1-|\psi(x)|^2)(1-|\psi(y)|^2)} = \frac{|x-y|^2}{(1-|x|^2)(1-|y|^2)}.$$

Therefore, we have

$$|\psi(x)-\psi(y)| = |x-y|.$$

Thus ψ preserves Euclidean distances in B^n.

Now ψ maps each radius of B^n onto a radius of B^n. Therefore ψ extends to a function $\overline{\psi} : \overline{B}^n \to \overline{B}^n$ such that

$$\psi([0,x)) = [0, \overline{\psi}(x)) \quad \text{for each } x \text{ in } S^{n-1}.$$

Moreover $\overline{\psi}$ is continuous, since

$$\overline{\psi}(x) = 2\psi(x/2) \quad \text{for each } x \text{ in } \overline{B}^n.$$

Therefore $\overline{\psi}$ preserves Euclidean distances. Hence $\overline{\psi}$ preserves Euclidean inner products on $\overline{B}^n$. The same argument as in the proof of Theorem 1.3.2 shows that $\overline{\psi}$ is the restriction of an orthogonal transformation A of E^n. Therefore $\tau_{\phi(0)}A$ extends ϕ. Moreover $\tau_{\phi(0)}A$ is the only Möbius transformation of B^n extending ϕ, since any two Möbius transformations extending ϕ agree on $\overline{B}^n$ and so are the same by Theorem 4.3.6. □

By Theorem 4.5.2, we can identify the group $\mathrm{I}(B^n)$ of isometries of the conformal ball model with the group $\mathrm{M}(B^n)$ of Möbius transformations of B^n. In particular, we have the following corollary.

Corollary 1. *The groups* $\mathrm{I}(B^n)$ *and* $\mathrm{M}(B^n)$ *are isomorphic.*

An *m-sphere* of E^n is defined to be the intersection of a sphere $S(a,r)$ of E^n with an $(m+1)$-plane of E^n that contains the center a. An *m-sphere* of $\hat{E}^n$ is defined to be either an m-sphere or an extended m-plane $\hat{P}$ of $\hat{E}^n$.

Lemma 2. *The group* $\mathrm{M}(\hat{E}^n)$ *acts transitively on the set of all m-spheres of* $\hat{E}^n$.

Proof: Let V be the vector subspace of E^n spanned by $e_1, \ldots, e_m$. It suffices to show that for every m-sphere Σ of $\hat{E}^n$, there is a Möbius transformation ϕ of $\hat{E}^n$ such that $\phi(\hat{V}) = \Sigma$, and the image of $\hat{V}$ under every Möbius transformation of $\hat{E}^n$ is an m-sphere of $\hat{E}^n$.

Let Σ be an arbitrary m-sphere of $\hat{E}^n$. If Σ is an extended m-plane, then there is an isometry ϕ of E^n such that $\phi(\hat{V}) = \Sigma$, since $\mathrm{I}(E^n)$ acts transitively on the set of m-planes of E^n.

Now suppose that Σ is an m-sphere of E^n. As the group of similarities of E^n acts transitively on the set of m-spheres of E^n, we may assume that $\Sigma = S^m$. Then the reflection in the sphere $S(e_{m+1}, \sqrt{2})$ maps $\hat{V}$ onto Σ.

Let ϕ be a Möbius transformation of $\hat{E}^n$. If $\phi(\infty) = \infty$, then ϕ is a Euclidean similarity, and so $\phi(\hat{V})$ is an extended m-plane of $\hat{E}^n$. Now assume that $\phi(\infty) \neq \infty$. Then by Theorem 4.3.3, we have that $\phi = \psi\sigma$ where σ is the reflection in a sphere $S(a, r)$ and ψ is a Euclidean isometry. If a is in V, then σ leaves $\hat{V}$ invariant, and so $\phi(\hat{V})$ is an extended m-plane of $\hat{E}^n$.

Now assume that a is not in V. Then V and a span an $(m + 1)$-dimensional vector subspace W of E^n. Moreover $\hat{V}$ is a sphere in $\hat{W}$. As σ leaves $\hat{W}$ invariant, $\sigma(\hat{V})$ is a sphere in $\hat{W}$ by Theorem 4.3.4. The point ∞ is not in $\sigma(\hat{V})$, since a is not in $\hat{V}$. Hence $\sigma(\hat{V})$ is an m-sphere of E^n, and so $\phi(\hat{V})$ is an m-sphere of E^n. □

A subset P of B^n is said to be a *hyperbolic m-plane* of B^n if and only if $\zeta(P)$ is a hyperbolic m-plane of H^n. A p-sphere Σ and a q-sphere Σ' of $\hat{E}^n$ are said to be *orthogonal* if and only if they intersect and at each finite point of intersection their tangent planes are orthogonal.

Theorem 4.5.3. *A subset P of B^n is a hyperbolic m-plane of B^n if and only if P is the intersection of B^n with either an m-dimensional vector subspace of E^n or an m-sphere of E^n orthogonal to S^{n-1}.*

Proof: Let P be the intersection of B^n with the vector subspace V of E^n spanned by $e_1, \ldots, e_m$. Then obviously ζ maps P onto the hyperbolic m-plane of H^n obtained by intersecting H^n with the vector subspace spanned by V and e_{n+1}. Thus P is a hyperbolic m-plane of B^n.

Let P' be an arbitrary hyperbolic m-plane of B^n. By Theorem 3.1.5, the group $\mathrm{M}(B^n)$ acts transitively on the set of hyperbolic m-planes of B^n. Hence, there is a Möbius transformation ϕ of B^n such that $\phi(P) = P'$. By Lemma 2, the set $\phi(\hat{V})$ is an m-sphere of $\hat{E}^n$. As ϕ is conformal, $\phi(\hat{V})$ is orthogonal to $\phi(S^{n-1}) = S^{n-1}$. Therefore P' is the intersection of B^n with either an m-dimensional vector subspace of E^n or an m-sphere of E^n orthogonal to S^{n-1}.

Let Q be the intersection of B^n with either an m-dimensional vector subspace of E^n or an m-sphere of E^n orthogonal to S^{n-1}. Then the boundary of Q in S^{n-1} is an $(m-1)$-sphere Σ of E^n. By Lemma 2, there is a Möbius transformation ψ of S^{n-1} such that ψ maps the boundary of P in S^{n-1} onto Q. The Poincaré extension $\tilde{\psi}$ then maps P onto Q. Thus Q is a hyperbolic m-plane of B^n. □

A *hyperbolic line* of B^n is defined to be a hyperbolic 1-plane of B^n. The geodesics of B^n are its hyperbolic lines by Corollary 4 of §3.2.

Corollary 2. *A subset L of B^n is a hyperbolic line of B^n if and only if L is either an open diameter of B^n or the intersection of B^n with a circle orthogonal to S^{n-1}.*

It is clear from the geometric definition of the stereographic projection ζ of B^n onto H^n that ζ preserves the Euclidean angle between any two geodesic lines intersecting at the origin. As the hyperbolic angle between two geodesic lines in H^n intersecting at $\zeta(0) = e_{n+1}$ is the same as the Euclidean angle, the hyperbolic angle between two geodesic lines in B^n intersecting at the origin is the same as the Euclidean angle between the lines. Moreover, since the isometries of B^n are conformal, the hyperbolic angle between any two intersecting geodesic lines in B^n is the same as the Euclidean angle between the lines. Thus, the hyperbolic angles of B^n conform with the corresponding Euclidean angles. For this reason, B^n is called the conformal ball model of hyperbolic n-space.

The *hyperbolic sphere* of B^n, with center b and radius $r > 0$, is defined to be the set

$$S_B(b,r) = \{x \in B^n : d_B(b,x) = r\}.$$

Theorem 4.5.4. *A subset S of B^n is a hyperbolic sphere of B^n if and only if S is a Euclidean sphere of E^n that is contained in B^n.*

Proof: Let $S = S_B(b,r)$. Assume first that $b = 0$. By Theorem 4.5.1, the distance $d_B(0,x)$ is an invertible function of $|x|$. Therefore S is a Euclidean sphere centered at 0. Now assume that b is an arbitrary point of B^n. Then the hyperbolic translation τ_b maps $S_B(0,r)$ onto S. Therefore S is a Euclidean sphere by Theorem 4.3.4.

Conversely, suppose that S is a Euclidean sphere contained in B^n. If S is centered at 0, then S is a hyperbolic sphere, since $d_B(0,x)$ is an invertible function of $|x|$. Now assume that S is not centered at 0. Let x be the point of S nearest to 0, and let y be the point of S farthest from 0. Then the line segment $[x,y]$ is a diameter of S. The line segment $[x,y]$ is also a geodesic segment of B^n. Let b be the hyperbolic midpoint of $[x,y]$, and let r be the hyperbolic distance from b to x. Then τ_b maps $S_B(0,r)$ onto $S_B(b,r)$, and $S_B(b,r)$ is a Euclidean sphere by Theorem 4.3.4. Observe that τ_b maps a diameter of $S_B(0,r)$ onto $[x,y]$. Therefore $[x,y]$ is orthogonal to $S_B(b,r)$ at x and y, since τ_b is conformal. Hence $[x,y]$ is a Euclidean diameter of $S_B(b,r)$. Therefore $S = S_B(b,r)$. □

Let a be a point on a hyperbolic sphere S of B^n, and let R be the geodesic ray of B^n starting at a and passing through the center c of S. If we expand S by moving c away from a on R at a constant rate while keeping a on S, the sphere tends to a limiting hypersurface Σ in B^n containing a. By moving a to 0, we see that Σ is a Euclidean sphere minus the ideal endpoint b of R and that the Euclidean sphere $\overline{\Sigma}$ is tangent to S^{n-1} at b.

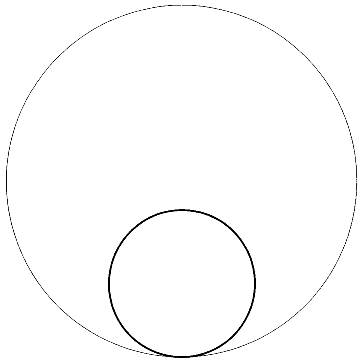


Figure 4.5.2. A horocycle of B^2

A *horosphere* Σ of B^n, based at a point b of S^{n-1}, is defined to be the intersection with B^n of a Euclidean sphere in $\overline{B}^n$ tangent to S^{n-1} at b. A horosphere in dimension two is also called a *horocycle*. See Figure 4.5.2. The interior of a horosphere is called a *horoball*. The interior of a horocycle is also called a *horodisk*.

Theorem 4.5.5. *The element of hyperbolic arc length of the conformal ball model B^n is*

$$\frac{2|dx|}{1-|x|^2}.$$

Proof: Let $y = \zeta(x)$. From the results of §3.3, the element of hyperbolic arc length of H^n is

$$\|dy\| = (dy_1^2 + \cdots + dy_n^2 - dy_{n+1}^2)^{\frac{1}{2}}.$$

Now since

$$y_i = \frac{2x_i}{1-|x|^2} \quad \text{for } i = 1, \ldots, n,$$

we have

$$dy_i = \frac{2dx_i}{1-|x|^2} + \frac{4x_i(x \cdot dx)}{(1-|x|^2)^2}.$$

Hence

$$dy_i^2 = \frac{4}{(1-|x|^2)^2}\left(dx_i^2 + \frac{4x_i dx_i(x \cdot dx)}{1-|x|^2} + \frac{4x_i^2(x \cdot dx)^2}{(1-|x|^2)^2}\right).$$

Thus

$$\begin{aligned}\sum_{i=1}^n dy_i^2 &= \frac{4}{(1-|x|^2)^2}\left(|dx|^2 + \frac{4(x \cdot dx)^2}{1-|x|^2} + \frac{4|x|^2(x \cdot dx)^2}{(1-|x|^2)^2}\right) \\ &= \frac{4}{(1-|x|^2)^2}\left(|dx|^2 + \frac{4(x \cdot dx)^2}{(1-|x|^2)^2}\right).\end{aligned}$$

Now since

$$y_{n+1} = \frac{1+|x|^2}{1-|x|^2},$$

we have that

$$dy_{n+1} = \frac{4x \cdot dx}{(1-|x|^2)^2}.$$

Thus

$$\sum_{i=1}^{n} dy_i^2 - dy_{n+1}^2 = \frac{4|dx|^2}{(1-|x|^2)^2}. \qquad \square$$

Theorem 4.5.6. *The element of hyperbolic volume of the conformal ball model B^n is*

$$\frac{2^n dx_1 \cdots dx_n}{(1-|x|^2)^n}.$$

Proof: An intuitive argument goes as follows: The element of hyperbolic arc length in the x_i-direction is

$$ds_i = \frac{2dx_i}{1-|x|^2}.$$

Therefore, the element of hyperbolic volume is

$$ds_1 \cdots ds_n = \frac{2^n dx_1 \cdots dx_n}{(1-|x|^2)^n}.$$

For a proof based on the definition of hyperbolic volume, start with the element of hyperbolic volume of H^n with respect to the Euclidean coordinates $y_1, \ldots, y_n$ given by Theorem 3.4.1,

$$\frac{dy_1 \cdots dy_n}{[1+(y_1^2+\cdots+y_n^2)]^{\frac{1}{2}}}.$$

Then change coordinates via the map $\overline{\zeta} : B^n \to E^n$ defined by

$$\overline{\zeta}(x) = \frac{2x}{1-|x|^2}.$$

Now since $\overline{\zeta}$ is a radial map, it is best to switch to spherical coordinates $(\rho, \theta_1, \ldots, \theta_{n-1})$ and decompose $\overline{\zeta}$ into the composite mapping

$$\begin{aligned}
(x_1, \ldots, x_n) &\mapsto (\rho, \theta_1, \ldots, \theta_{n-1}) \\
&\mapsto \left(\frac{2\rho}{1-\rho^2}, \theta_1, \ldots, \theta_{n-1}\right) \\
&\mapsto (y_1, \ldots, y_n).
\end{aligned}$$

Now since

$$\frac{d}{d\rho}\left(\frac{2\rho}{1-\rho^2}\right) = \frac{2(1+\rho^2)}{(1-\rho^2)^2},$$

the Jacobian of $\overline{\zeta}$ is

$$\frac{1}{\rho^{n-1}}\frac{2(1+\rho^2)}{(1-\rho^2)^2}\left(\frac{2\rho}{1-\rho^2}\right)^{n-1} = \frac{2^n(1+\rho^2)}{(1-\rho^2)^{n+1}}.$$

Let $y = \overline{\zeta}(x)$. Then

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

Therefore

$$\begin{aligned}\frac{dy_1\cdots dy_n}{(1+|y|^2)^{\frac{1}{2}}} &= \frac{2^n(1+|x|^2)}{(1-|x|^2)^{n+1}}\frac{(1-|x|^2)}{(1+|x|^2)}dx_1\cdots dx_n \\ &= \frac{2^n dx_1\cdots dx_n}{(1-|x|^2)^n}.\end{aligned}$$

□

Exercise 4.5

1. Show that if x is in B^n, then
$$d_B(0,x) = \log\left(\frac{1+|x|}{1-|x|}\right).$$

2. Let b be a nonzero point of B^n. Show that the hyperbolic translation τ_b of B^n acts as a hyperbolic translation along the hyperbolic line passing through 0 and b.

3. Let b be a point of B^n and let A be in $\mathrm{O}(n)$. Show that
   (1) $\tau_b^{-1} = \tau_{-b}$,
   (2) $A\tau_b A^{-1} = \tau_{Ab}$.

4. Show that $S_B(0,r) = S(0,\tanh(r/2))$.

5. Prove that the hyperbolic and Euclidean centers of a sphere of B^n coincide if and only if the sphere is centered at the origin.

6. Prove that the metric topology on B^n determined by d_B is the same as the Euclidean topology on B^n.

7. Prove that all the horospheres of B^n are congruent.

8. Let b be a point of B^n not on a hyperbolic m-plane P of B^n. Prove that there is a unique point a of P nearest to b and that the hyperbolic line passing through a and b is the unique hyperbolic line of B^n passing through b orthogonal to P. Hint: Move b to the origin.

9. Let b be a point of B^n not on a horosphere Σ of B^n. Prove that there is a unique point a of Σ nearest to b and the hyperbolic line passing through a and b is the unique hyperbolic line of B^n passing through b orthogonal to Σ.

10. Show that every isometry of B^2 is of the form
$$z \mapsto \frac{az+b}{\overline{b}z+\overline{a}} \quad \text{or} \quad z \mapsto \frac{a\overline{z}+b}{\overline{b}\overline{z}+\overline{a}} \quad \text{where } |a|^2-|b|^2 = 1.$$

§4.6. The Upper Half-Space Model

Let η be the standard transformation from upper half-space U^n to the open unit ball B^n. Then $\eta = \sigma\rho$, where ρ is the reflection of $\hat{E}^n$ in the hyperplane E^{n-1} and σ is the reflection of $\hat{E}^n$ in the sphere $S(e_n, \sqrt{2})$. Define a metric d_U on U^n by the formula

$$d_U(x,y) = d_B(\eta(x), \eta(y)). \tag{4.6.1}$$

The metric d_U is called the *Poincaré metric* on U^n. By definition, η is an isometry from U^n, with the metric d_U, to the conformal ball model B^n of hyperbolic n-space. The metric space consisting of U^n together with the metric d_U is called the *upper half-space model* of hyperbolic n-space.

Theorem 4.6.1. *The metric d_U on U^n is given by*

$$\cosh d_U(x,y) = 1 + \frac{|x-y|^2}{2x_n y_n}.$$

Proof: By Theorem 4.5.1, we have

$$\begin{aligned}\cosh d_U(x,y) &= \cosh d_B(\eta(x), \eta(y)) \\ &= 1 + \frac{2|\sigma\rho(x) - \sigma\rho(y)|^2}{(1-|\sigma\rho(x)|^2)(1-|\sigma\rho(y)|^2)}.\end{aligned}$$

By Theorem 4.1.3, we have

$$\begin{aligned}|\sigma\rho(x) - \sigma\rho(y)| &= \frac{2|\rho(x) - \rho(y)|}{|\rho(x) - e_n|\,|\rho(y) - e_n|} \\ &= \frac{2|x-y|}{|x+e_n|\,|y+e_n|},\end{aligned}$$

and by Formula 4.4.2, we have

$$1 - |\sigma\rho(x)|^2 = \frac{-4[\rho(x)]_n}{|\rho(x) - e_n|^2} = \frac{4x_n}{|x+e_n|^2}.$$

Therefore

$$\cosh d_U(x,y) = 1 + \frac{|x-y|^2}{2x_n y_n}.$$ □

The next theorem follows immediately from Theorem 4.5.2.

Theorem 4.6.2. *Every Möbius transformation of U^n restricts to an isometry of the upper half-space model U^n, and every isometry of U^n extends to a unique Möbius transformation of U^n.*

By Theorem 4.6.2, we can identify the group $\mathrm{I}(U^n)$ of isometries of the upper half-space model with the group $\mathrm{M}(U^n)$ of Möbius transformations of U^n.

Corollary 1. *The groups* $\mathrm{I}(U^n)$ *and* $\mathrm{M}(U^n)$ *are isomorphic.*

As the upper half-space model U^n is isometric to hyperbolic n-space H^n, we have that $\mathrm{I}(U^n)$ is isomorphic to $\mathrm{I}(H^n)$. By Corollary 1 of §4.4, the groups $\mathrm{M}(U^n)$ and $\mathrm{M}(\hat{E}^{n-1})$ are isomorphic. Thus, from Corollary 1, we have the following corollary.

Corollary 2. *The groups* $\mathrm{I}(H^n)$ *and* $\mathrm{M}(\hat{E}^{n-1})$ *are isomorphic.*

A subset P of U^n is said to be a *hyperbolic m-plane* of U^n if and only if $\eta(P)$ is a hyperbolic m-plane of B^n. The next theorem follows immediately from Theorem 4.5.3.

Theorem 4.6.3. *A subset P of U^n is a hyperbolic m-plane of U^n if and only if P is the intersection of U^n with either an m-plane of E^n orthogonal to E^{n-1} or an m-sphere of E^n orthogonal to E^{n-1}.*

A *hyperbolic line* of U^n is defined to be a hyperbolic 1-plane of U^n. The geodesics of U^n are its hyperbolic lines by Corollary 4 of §3.2.

Corollary 3. *A subset L of U^n is a hyperbolic line of U^n if and only if L is the intersection of U^n with either a straight line orthogonal to E^{n-1} or a circle orthogonal to E^{n-1}.*

The standard transformation $\eta : U^n \to B^n$ is conformal. Hence, the hyperbolic angle between any two intersecting geodesic lines of U^n conforms with the Euclidean angle between the lines, since this is the case in the conformal ball model B^n. Thus, the upper half-space model U^n is also a conformal model of hyperbolic n-space.

The *hyperbolic sphere* of U^n, with center a and radius $r > 0$, is defined to be the set

$$S_U(a,r) = \{x \in U^n : d_U(a,x) = r\}.$$

The next theorem follows immediately from Theorem 4.5.4

Theorem 4.6.4. *A subset S of U^n is a hyperbolic sphere of U^n if and only if S is a Euclidean sphere of E^n that is contained in U^n.*

A subset Σ of U^n is said to be a *horosphere* of U^n based at a point b of $\hat{E}^{n-1}$ if and only if $\eta(\Sigma)$ is a horosphere of B^n based at the point $\eta(b)$.

Theorem 4.6.5. *A subset Σ of U^n is a horosphere of U^n based at a point b of $\hat{E}^{n-1}$ if and only if Σ is either a Euclidean hyperplane in U^n parallel to E^{n-1} if $b = \infty$, or the intersection with U^n of a Euclidean sphere in $\overline{U}^n$ tangent to E^{n-1} at b if $b \neq \infty$.*

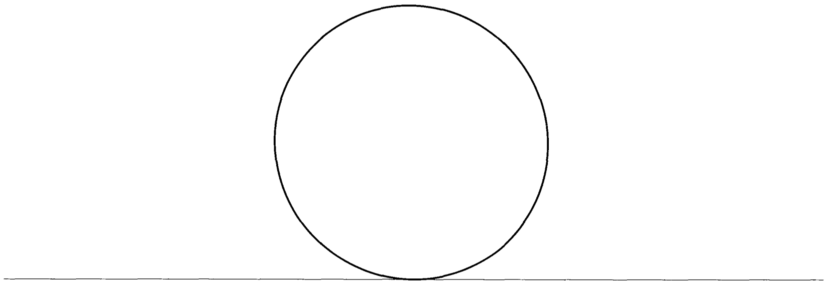


Figure 4.6.1. A horocycle of U^2

Proof: By Theorem 4.3.4, a subset Σ of U^n is a horosphere of U^n if and only if $\overline{\Sigma}$ is a sphere of $\hat{E}^n$ that is contained in $\overline{U}^n$ and meets $\hat{E}^{n-1}$ at exactly one point. Therefore Σ is a horosphere of U^n if and only if Σ is either a Euclidean hyperplane in U^n parallel to E^{n-1} or the intersection with U^n of a Euclidean sphere in $\overline{U}^n$ tangent to E^{n-1}. □

A horosphere in dimension two is also called a *horocycle.* See Figure 4.6.1. The interior of a horosphere is called a *horoball.* The interior of a horocycle is also called a *horodisk.*

Theorem 4.6.6. *The element of hyperbolic arc length of the upper half-space model U^n is*

$$\frac{|dx|}{x_n}.$$

Proof: Let $y = \eta(x)$. Then

$$y = e_n + \frac{2(\rho(x) - e_n)}{|x + e_n|^2}.$$

By Theorem 4.5.5, the element of arc length of B^n is $2|dy|/(1 - |y|^2)$. As

$$y_i = \frac{2x_i}{|x + e_n|^2} \quad \text{for } i = 1, \ldots, n-1,$$

we have

$$dy_i = \frac{2dx_i}{|x + e_n|^2} - \frac{4x_i(x + e_n) \cdot dx}{|x + e_n|^2}.$$

Hence

$$dy_i^2 = \frac{4}{|x + e_n|^4}\left[dx_i^2 - \frac{4x_i dx_i (x + e_n) \cdot dx}{|x + e_n|^2} + \frac{4x_i^2[(x + e_n) \cdot dx]^2}{|x + e_n|^4}\right].$$

Now since
$$y_n = 1 - \frac{2(x_n+1)}{|x+e_n|^2},$$
we have
$$dy_n = \frac{-2dx_n}{|x+e_n|^2} + \frac{4(x_n+1)(x+e_n)\cdot dx}{|x+e_n|^4}.$$
Hence
$$dy_n^2 = \frac{4}{|x+e_n|^4}\left[dx_n^2 - \frac{4(x_n+1)dx_n(x+e_n)\cdot dx}{|x+e_n|^2} + \frac{4(x_n+1)^2[(x+e_n)\cdot dx]^2}{|x+e_n|^4}\right].$$
Thus
$$\begin{aligned}|dy|^2 &= \frac{4}{|x+e_n|^4}\left[|dx|^2 - \frac{4[(x+e_n)\cdot dx]^2}{|x+e_n|^2} + \frac{4|x+e_n|^2[(x+e_n)\cdot dx]^2}{|x+e_n|^4}\right]\\ &= \frac{4|dx|^2}{|x+e_n|^4}.\end{aligned}$$
From the proof of Theorem 4.6.1, we have
$$1-|y|^2 = \frac{4x_n}{|x+e_n|^2}.$$
Therefore, we have
$$\frac{2|dy|}{1-|y|^2} = \frac{|dx|}{x_n}.$$
□

Theorem 4.6.7. *The element of hyperbolic volume of the upper half-space model U^n is*
$$\frac{dx_1\cdots dx_n}{(x_n)^n}.$$

Proof: An intuitive argument goes as follows: The element of hyperbolic arc length in the x_i-direction is
$$ds_i = \frac{dx_i}{x_n}.$$
Therefore, the element of hyperbolic volume is
$$ds_1\cdots ds_n = \frac{dx_1\cdots dx_n}{(x_n)^n}.$$
The element of hyperbolic volume of U^n can also be derived from the element of hyperbolic volume of B^n. Let $y=\eta(x)$. By Theorem 4.5.6, the element of hyperbolic volume of B^n is
$$\frac{2^n dy_1\cdots dy_n}{(1-|y|^2)^n}.$$

From the proof of Theorem 4.1.5, we see that the Jacobian of η is

$$\frac{(\sqrt{2})^{2n}}{|\rho(x)-e_n|^{2n}} = \frac{2^n}{|x+e_n|^{2n}}.$$

From the proof of Theorem 4.6.1, we have

$$1-|y|^2 = \frac{4x_n}{|x+e_n|^2}.$$

Therefore

$$\begin{aligned}\frac{2^n dy_1\cdots dy_n}{(1-|y|^2)^n} &= \left(\frac{2^n}{|x+e_n|^{2n}}\right)\left(\frac{2^n|x+e_n|^{2n}}{4^n(x_n)^n}\right)dx_1\cdots dx_n\\ &= \frac{dx_1\cdots dx_n}{(x_n)^n}.\end{aligned}$$

□

Exercise 4.6

1. Show that if $x = se_n$ and $y = te_n$, then $d_U(x,y) = |\log(s/t)|$.
2. Show that if $-1 < s < 1$ and x is in U^n, then
$$\eta^{-1}\tau_{se_n}\eta(x) = \left(\frac{1+s}{1-s}\right)x.$$
3. Let x be in U^n. Show that the nearest point to x on the positive nth axis is $|x|e_n$ and
$$\cosh d_U(x,|x|e_n) = |x|/x_n.$$
4. Show that $S_U(a,r) = S(a(r), a_n\sinh r)$, where
$$a(r) = (a_1,\ldots,a_{n-1},a_n\cosh r).$$
5. Prove that the metric topology on U^n determined by d_U is the same as the Euclidean topology.
6. Prove that all the horospheres of U^n are congruent.
7. Prove that any Möbius transformation ϕ of $\hat{E}^n$ that leaves the horosphere $\Sigma_1 = \{x\in U^n : x_n = 1\}$ invariant is a Euclidean isometry of E^n.
8. Show by changing coordinates that every Möbius transformation of U^n preserves hyperbolic volume.
9. Show that every isometry of U^2 is of the form
$$z\mapsto \frac{az+b}{cz+d} \quad\text{or}\quad z\mapsto\frac{a(-\overline{z})+b}{c(-\overline{z})+d},$$
where a,b,c,d are real and $ad-bc=1$. Conclude that the group $\mathrm{I}_0(U^2)$ of orientation preserving isometries of U^2 is isomorphic to $\mathrm{PSL}(2,\mathbb{R})$.

§4.7. Classification of Transformations

Let ϕ be a Möbius transformation of B^n. Then ϕ maps the closed ball $\overline{B}^n$ to itself. By the Brouwer fixed point theorem, ϕ has a fixed point in $\overline{B}^n$. The transformation ϕ is said to be

(1) *elliptic* if ϕ fixes a point of B^n;

(2) *parabolic* if ϕ fixes no point of B^n and fixes a unique point of S^{n-1};

(3) *hyperbolic* if ϕ fixes no point of B^n and fixes two points of S^{n-1}.

Let F_ϕ be the set of all the fixed points of ϕ in $\overline{B}^n$, and let ψ be a Möbius transformation of B^n. Then

$$F_{\psi\phi\psi^{-1}} = \psi(F_\phi).$$

Hence ϕ is elliptic, parabolic, or hyperbolic if and only if $\psi\phi\psi^{-1}$ is elliptic, parabolic, or hyperbolic, respectively. Thus, being elliptic, parabolic, or hyperbolic depends only on the conjugacy class of ϕ in $\mathrm{M}(B^n)$.

Elliptic Transformations

We now characterize the elliptic transformations of B^n.

Theorem 4.7.1. *A Möbius transformation ϕ of B^n is elliptic if and only if ϕ is conjugate in $\mathrm{M}(B^n)$ to an orthogonal transformation of E^n.*

Proof: Suppose that ϕ is elliptic. Then ϕ fixes a point b of B^n. Let τ_b be the hyperbolic translation of B^n by b. Then $\tau_b^{-1}\phi\tau_b$ fixes the origin. By Theorem 4.4.8, the map $\tau_b^{-1}\phi\tau_b$ is an orthogonal transformation A of E^n. Thus $\phi = \tau_b A \tau_b^{-1}$. Conversely, suppose that ϕ is conjugate in $\mathrm{M}(B^n)$ to an orthogonal transformation A of E^n. Then A is elliptic, since it fixes the origin. Therefore ϕ is elliptic. □

Let $S_B(b,r)$ be the hyperbolic sphere of B^n with center b and radius r. Let x, y be in $S_B(b,r)$ and let $\alpha, \beta : [0,r] \to B^n$ be geodesics arcs from b to x and y, respectively. Regard α and β as the hyperbolic radii of $S_B(b,r)$ from the center b to x and y. The sphere $S_B(b,r)$ has a natural spherical metric given by

$$d(x,y) = (\sinh r)\theta(\alpha'(0), \beta'(0)).$$

In other words, a hyperbolic sphere of radius r is isometric to a Euclidean sphere of radius $\sinh r$. If ϕ is an elliptic transformation of B^n, with b as a fixed point, then obviously ϕ leaves each hyperbolic sphere $S_B(b,r)$ centered at b invariant; moreover, ϕ acts as an isometry of the natural spherical metric on $S_B(b,r)$.

Parabolic Transformations

In order to analyze parabolic and hyperbolic transformations, it will be more convenient to work in the upper half-space model U^n of hyperbolic space. Elliptic, parabolic, and hyperbolic Möbius transformations of U^n are defined in the same manner as in the conformal ball model B^n. Let ϕ be a Möbius transformation of U^n. The transformation ϕ is said to be

(1) *elliptic* if ϕ fixes a point of U^n;

(2) *parabolic* if ϕ fixes no point of U^n and fixes a unique point of $\hat{E}^{n-1}$;

(3) *hyperbolic* if ϕ fixes no point of U^n and fixes two points of $\hat{E}^{n-1}$.

Note that being elliptic, parabolic, or hyperbolic depends only on the conjugacy class of ϕ in $\mathrm{M}(U^n)$.

We now characterize the parabolic transformations of U^n.

Theorem 4.7.2. *A Möbius transformation ϕ of U^n is parabolic if and only if ϕ is conjugate in $\mathrm{M}(U^n)$ to the Poincaré extension of a fixed point free isometry of E^{n-1}.*

Proof: Suppose that ϕ is parabolic. Then ϕ fixes a point a of $\hat{E}^{n-1}$. In the conformal ball model B^n, the point ∞ corresponds to e_n and an appropriate rotation will map any point of S^{n-1} to e_n. Hence, there is a Möbius transformation ψ of U^n such that $\psi(a) = \infty$. Then $\psi\phi\psi^{-1}$ fixes ∞. By Theorems 4.3.2 and 4.4.1, the map $\psi\phi\psi^{-1}$ is the Poincaré extension of a similarity of E^{n-1}. Hence, there is a point b in E^{n-1}, a scalar $k > 0$, and an orthogonal transformation A of E^{n-1} such that

$$\psi\phi\psi^{-1}(x) = b + k\tilde{A}x.$$

As ∞ is the only fixed point of $\psi\phi\psi^{-1}$, the fixed point equation

$$b + kAx = x$$

has no solutions in E^{n-1}. The above equation can be rewritten as

$$\left(A - \frac{1}{k}I\right)x = -\frac{b}{k}.$$

Since this equation has no solution, we have

$$\det\left(A - \frac{1}{k}I\right) = 0.$$

Thus $1/k$ is an eigenvalue of A. As A is orthogonal, $k = 1$. Thus $\psi\phi\psi^{-1}$ is the Poincaré extension of a fixed point free isometry of E^{n-1}.

Conversely, suppose that ϕ is conjugate in $\mathrm{M}(U^n)$ to the Poincaré extension of a fixed point free isometry ψ of E^{n-1}. Then the Poincaré extension $\tilde{\psi}$ is parabolic, since ∞ is its only fixed point. Thus ϕ is parabolic. $\square$

An important class of parabolic transformations of U^n are the nontrivial Euclidean translations of U^n. Such a transformation is of the form

$$x \mapsto a + x,$$

where a is a nonzero point of E^{n-1}. A Möbius transformation ϕ of U^n is said to be a *parabolic translation* if and only if ϕ is conjugate in $\mathrm{M}(U^n)$ to a nontrivial Euclidean translation of U^n.

Let Σ_1 be the horosphere of U^n defined by

$$\Sigma_1 = \{x \in U^n : x_n = 1\}.$$

The horosphere Σ_1 has a natural Euclidean metric given by

$$d(x, y) = |x - y|.$$

This metric is natural, since the element of hyperbolic arc length $|dx|/x_n$ of U^n restricts to the element of Euclidean arc length $|dx|$ on Σ_1.

Let Σ be any horosphere of U^n. Then there is a Möbius transformation ϕ of U^n such that $\phi(\Sigma) = \Sigma_1$. Define a Euclidean metric on Σ by

$$d(x, y) = |\phi(x) - \phi(y)|.$$

We claim that this metric is independent of the choice of ϕ. Suppose that ψ is another Möbius transformation of U^n such that $\psi(\Sigma) = \Sigma_1$. Then $\phi\psi^{-1}$ leaves Σ_1 invariant. This implies that $\phi\psi^{-1}$ is a Euclidean isometry. Therefore, if x, y are in Σ, then

$$|\phi(x) - \phi(y)| = |\phi\psi^{-1}\psi(x) - \phi\psi^{-1}\psi(y)| = |\psi(x) - \psi(y)|.$$

Thus, the metric d on Σ does not depend on ϕ. The metric d is called the *natural Euclidean metric* on Σ.

Theorem 4.7.3. *Let Σ and Σ' be horospheres of U^n and let ψ be a Möbius transformation of U^n such that $\psi(\Sigma) = \Sigma'$. Then ψ acts as an isometry with respect to the natural Euclidean metrics on Σ and Σ'.*

Proof: Let ϕ and ϕ' be Möbius transformations of U^n such that $\phi(\Sigma) = \Sigma_1$ and $\phi'(\Sigma') = \Sigma_1$. Then $\phi'\psi\phi^{-1}$ leaves Σ_1 invariant and so is a Euclidean isometry. Hence, if x, y are in Σ, then

$$\begin{aligned} d'(\psi(x), \psi(y)) &= |\phi'\psi(x) - \phi'\psi(y)| \\ &= |\phi'\psi\phi^{-1}\phi(x) - \phi'\psi\phi^{-1}\phi(y)| \\ &= |\phi(x) - \phi(y)| \\ &= d(x, y). \end{aligned}$$

□

Now let ϕ be a parabolic transformation of U^n with a as its unique fixed point in $\hat{E}^{n-1}$. By Theorem 4.7.2, the map ϕ leaves each horosphere of U^n based at a invariant. By Theorem 4.7.3, the map ϕ acts as an isometry of the natural Euclidean metric on each horosphere based at a.

Hyperbolic Transformations

We now characterize the hyperbolic transformations of U^n.

Theorem 4.7.4. *A Möbius transformation ϕ of U^n is hyperbolic if and only if ϕ is conjugate in* $\mathrm{M}(U^n)$ *to the Poincaré extension of a similarity ψ of E^{n-1} of the form $\psi(x) = kAx$, where $k > 1$ and A is an orthogonal transformation of E^{n-1}.*

Proof: Suppose that ϕ is hyperbolic. By conjugating ϕ, we may assume that one of the fixed points of ϕ is ∞. Let a in E^{n-1} be another fixed point and let τ be the translation of E^n by $-a$. Then $\tau\phi\tau^{-1}$ fixes both 0 and ∞. This implies that there is a scalar $k > 0$ and an orthogonal transformation A of E^{n-1} such that

$$\tau\phi\tau^{-1}(x) = k\tilde{A}x.$$

As $\tilde{A}$ fixes e_n and $\tau\phi\tau^{-1}$ has no fixed points in U^n, we must have $k \neq 1$. Let $\sigma(x) = x/|x|^2$. Then

$$\sigma\tau\phi\tau^{-1}\sigma^{-1}(x) = k^{-1}\tilde{A}x.$$

Hence, we may assume that $k > 1$.

Conversely, suppose that ϕ is conjugate in $\mathrm{M}(U^n)$ to the Poincaré extension of a similarity ψ of E^{n-1} of the form $\psi(x) = kAx$, where $k > 1$ and A is an orthogonal transformation of E^{n-1}. Then the Poincaré extension $\tilde{\psi}$ is hyperbolic, since 0 and ∞ are its only fixed points. Therefore ϕ is hyperbolic. □

Corollary 1. *A hyperbolic transformation has exactly two fixed points.*

The simplest class of hyperbolic transformations of U^n are the nontrivial magnifications of U^n. Such a transformation is of the form $x \mapsto kx$, where $k > 1$. Notice that a magnification of U^n leaves the positive nth axis invariant. Moreover, if $t > 0$, then

$$d_U(te_n, kte_n) = \log k.$$

Thus, a magnification of U^n acts as a hyperbolic translation along the positive nth axis. A Möbius transformation ϕ of U^n is said to be a *hyperbolic translation* if and only if ϕ is conjugate in $\mathrm{M}(U^n)$ to a magnification of U^n.

Now let ϕ be an arbitrary hyperbolic transformation of U^n with a and b its two fixed points, and let L be the hyperbolic line of U^n with endpoints a and b. By Theorem 4.7.4, the map ϕ is the composite of an elliptic transformation of U^n that fixes the line L followed by a hyperbolic translation along L. The line L is called the *axis* of the hyperbolic transformation ϕ. Note that a hyperbolic transformation acts as a translation along its axis.

Remark: We are not using the term *hyperbolic transformation* in its usual sense. Traditionally, a hyperbolic translation is called a hyperbolic transformation, and a hyperbolic transformation that is not a hyperbolic translation is called a *loxodromic transformation.*

Exercise 4.7

1. Prove that every element of $\mathrm{LF}(\hat{\mathbb{C}})$ has either one or two fixed points in $\hat{\mathbb{C}}$.
2. Let z_1, z_2, z_3 be distinct points of $\hat{\mathbb{C}}$ and let w_1, w_2, w_3 be distinct points of $\hat{\mathbb{C}}$. Show that there is a unique element ϕ of $\mathrm{M}(\hat{\mathbb{C}})$ such that $\phi(z_j) = w_j$ for $j = 1, 2, 3$.
3. For each nonzero k in $\mathbb{C}$, define μ_k in $\mathrm{LF}(\hat{\mathbb{C}})$ by $\mu_k(z) = kz$ if $k \neq 1$, and $\mu_1(z) = z + 1$. Prove that each nonidentity element of $\mathrm{LF}(\hat{\mathbb{C}})$ is conjugate to μ_k for some k.
4. Let
$$\phi(z) = \frac{az+b}{cz+d} \quad \text{with} \quad a, b, c, d \ \text{ in } \mathbb{C} \ \text{ and } \ ad - bc = 1.$$
Define
$$\mathrm{tr}^2(\phi) = (a+d)^2.$$
Show that two nonidentity elements ϕ, ψ of $\mathrm{LF}(\hat{\mathbb{C}})$ are conjugate if and only if $\mathrm{tr}^2(\phi) = \mathrm{tr}^2(\psi)$.
5. Let ϕ be in $\mathrm{LF}(\hat{\mathbb{C}})$ with $\phi \neq I$. Show that
   (1) $\tilde{\phi}$ is an elliptic transformation of U^3 if and only if $\mathrm{tr}^2(\phi)$ is in $[0, 4)$;
   (2) $\tilde{\phi}$ is a parabolic transformation of U^3 if and only if $\mathrm{tr}^2(\phi) = 4$;
   (3) $\tilde{\phi}$ is a hyperbolic translation of U^3 if and only if $\mathrm{tr}^2(\phi)$ is in $(4, +\infty)$.
6. Prove that the fixed set in B^n of an elliptic transformation of B^n is a hyperbolic m-plane.
7. Let a be the point of S^{n-1} fixed by a parabolic transformation ϕ of B^n. Prove that if x is in $\overline{B}^n$, then
$$\lim_{m\to\infty} \phi^m(x) = a.$$
In other words, a is an *attractive fixed point.*
8. Let a and b be the points of S^{n-1} fixed by a hyperbolic transformation ψ of B^n, and let L be the axis of ψ. Suppose that ψ translates L in the direction of a. Prove that if x is in $\overline{B}^n$ and $x \neq b$, then
$$\lim_{m\to\infty} \psi^m(x) = a.$$
In other words, a is an *attractive fixed point* and b is a *repulsive fixed point.*
9. Let A be in $\mathrm{PO}(n, 1)$. Prove that
   (1) A is elliptic if and only if A leaves invariant a 1-dimensional time-like vector subspace of $\mathbb{R}^{n,1}$;
   (2) A is parabolic if and only if A is not elliptic and A leaves invariant a unique 1-dimensional light-like vector subspace of $\mathbb{R}^{n,1}$;
   (3) A is hyperbolic if and only if A is not elliptic and A leaves invariant two 1-dimensional light-like vector subspaces of $\mathbb{R}^{n,1}$.
10. Let A be in $\mathrm{PO}(n, 1)$. Prove algebraically that A is either elliptic, parabolic, or hyperbolic.

§4.8. Historical Notes

§4.1. Jordan proved that a reflection of Euclidean n-space in a hyperplane is orientation reversing in his 1875 paper *Essai sur la géométrie à n dimensions* [207]. That an isometry of Euclidean n-space is the composition of at most $n+1$ reflections in hyperplanes appeared in Coxeter's 1948 treatise *Regular Polytopes* [92].

According to Rosenfeld's 1988 treatise *A History of Non-Euclidean Geometry* [353], Appollonius proved that an inversion in a circle maps circles to circles in his lost treatise *On plane loci.* A systematic development of inversion in a circle was first given by Plücker in his 1834 paper *Analytisch-geometrische Aphorismen* [326]. Inversion in a sphere was considered by Bellavitis in his 1836 paper *Teoria delle figure inverse, e loro uso nella geometria elementare* [37]. Theorem 4.1.3 appeared in Liouville's 1847 *Note au sujet de l'article précédent* (*de M. Thomson*) [259]. For the early history of inversion, see Patterson's 1933 article *The origins of the geometric principle of inversion* [324].

Conformal transformations of the plane appeared in Euler's 1770 paper *Considerationes de trajectoriis orthogonalibus* [123]. In particular, Euler considered linear fractional transformations of the complex plane in this paper. That inversion in a circle is conformal appeared in Plücker's 1834 paper [326]. That inversion in a sphere is conformal appeared in Thomson's 1845 letter to Liouville *Extrait d'une lettre de M. Thomson* [388].

§4.2. According to Heath's 1921 treatise *A History of Greek Mathematics* [186], stereographic projection was described by Ptolemy in his second century treatise *Planisphaerium.* That stereographic projection is the inversion of a sphere into a plane appeared in Bellavitis' 1836 paper [37]. The Riemann sphere was introduced by Riemann in his 1857 paper *Theorie der Abel'schen Functionen* [348]. The cross ratio of four points in the plane was introduced by Möbius in his 1852 paper *Ueber eine neue Verwandtschaft zwischen ebenen Figuren* [297].

§4.3. Möbius transformations of the plane were studied by Möbius in his 1855 paper *Theorie der Kreisverwandtschaft in rein geometrischer Darstellung* [298]. In particular, the 2-dimensional versions of Theorems 4.3.1 and 4.3.2 appeared in this paper. Möbius transformations of 3-space were considered by Liouville in his 1847 note [259]. Liouville proved the remarkable theorem that a smooth conformal transformation of 3-space is a Möbius transformation in his 1850 note *Extension au cas des trois dimensions de la question du tracé géographique* [261]. Liouville's theorem was extended to n-dimensions, $n > 2$, by Lie in his 1871 paper *Über diejenige Theorie eines Raumes mit beliebig vielen Dimensionen* [258]. The isometric circle of a linear fractional transformation of the complex plane was introduced by Ford in his 1927 paper *On the foundations of the theory of discontinuous groups* [135]. That inversion in a sphere maps inverse points to inverse points appeared in Thomson's 1845 letter to Liouville [388].

§4.4. The Poincaré extension of a Möbius transformation of the plane was defined by Poincaré in his 1881 note *Sur les groupes kleinéens* [329]. Möbius transformations of a sphere were considered by Möbius in his 1855 paper [298]. The 2-dimensional versions of Theorems 4.4.7 and 4.4.8 appeared in Ford's 1929 treatise *Automorphic Functions* [136].

§4.5. The conformal ball model of radius two was introduced by Beltrami in his 1868 paper *Saggio di interpetrazione della geometria non-euclidea* [38]. In particular, he derived its element of arc length and noted that this Riemannian metric had already been affirmed to be of constant negative curvature by Riemann in his 1854 lecture *Über die Hypothesen, welch der Geometrie zu Grunde liegen* [349]. For a discussion, see the introduction of Stillwell's 1985 translation of Poincaré's *Papers on Fuchsian Functions* [340]. The stereographic projection of Beltrami's conformal ball model onto hyperbolic space H^n appeared in Killing's 1878 paper *Ueber zwei Raumformen mit constanter positiver Krümmung* [219]. The 2-dimensional conformal ball model of radius one and curvature -4 appeared in Poincaré's 1882 paper *Sur les fonctions fuchsiennes* [331]. The 2-dimensional conformal ball model of radius one and curvature -1 appeared in Hausdorff's 1899 paper *Analytische Beiträge zur nichteuklidischen Geometrie* [180].

§4.6. The upper half-space model was introduced by Beltrami in his 1868 paper [38]. In particular, he derived its element of arc length and noted that this Riemannian metric in dimension two had already been shown to be of constant negative curvature by Liouville in his 1850 note *Sur le théorème de M. Gauss, concernant le produit des deux rayons de courbure principaux* [260]. That the group of Möbius transformations of n-space is isomorphic to the group of isometries of hyperbolic $(n+1)$-space follows immediately from observations of Klein in his 1872 paper *Ueber Liniengeometrie und metrische Geometrie* [225] and in his 1873 paper *Ueber die sogenannte Nicht-Euklidische Geometrie* [227].

§4.7. The classification of the isometries of the hyperbolic plane into three types according to the nature of their fixed points appeared in Klein's 1871 paper *Ueber die sogenannte Nicht-Euklidische Geometrie* [224]. The terms *elliptic, parabolic*, and *hyperbolic transformations* were introduced by Klein in his 1879 paper *Ueber die Transformation der elliptischen Functionen* [231] and were applied to isometries of hyperbolic n-space by Thurston in his 1979 lectures notes *The Geometry and Topology of 3-Manifolds* [389].

That the intrinsic geometry of a sphere in hyperbolic space is spherical is implicit in Lambert's remark in his 1786 monograph *Theorie der Parallellinien* [252] that spherical trigonometry is independent of Euclid's parallel postulate. This was proved by Bolyai in his 1832 paper *Scientiam spatii absolute veram exhibens* [51]. The corresponding fact in hyperbolic n-space appeared in Beltrami's 1868 paper *Teoria fondamentale degli spazii di curvatura costante* [39]. That the intrinsic geometry of a horosphere is Euclidean appeared in Lobachevski's 1829-30 paper *On the principles of geometry* [262] and in Bolyai's 1832 paper [51].

CHAPTER 5

Isometries of Hyperbolic Space

In this chapter, we study the topology of the group $\mathrm{I}(H^n)$ of isometries of hyperbolic space. The chapter begins with an introduction to topological groups. The topological group structure of $\mathrm{I}(H^n)$ is studied from various points of view in Section 5.2. The discrete subgroups of $\mathrm{I}(H^n)$ are of fundamental importance for the study of hyperbolic manifolds. The basic properties of the discrete subgroups of $\mathrm{I}(H^n)$ are examined in Section 5.3. A characterization of the discrete subgroups of $\mathrm{I}(E^n)$ is given in Section 5.4. The chapter ends with a characterization of all the elementary discrete subgroups of $\mathrm{I}(H^n)$.

§5.1. Topological Groups

Consider the n-dimensional complex vector space $\mathbb{C}^n$. A *vector* in $\mathbb{C}^n$ is an ordered n-tuple $z = (z_1, \ldots, z_n)$ of complex numbers. Let z and w be vectors in $\mathbb{C}^n$. The *Hermitian inner product* of z and w is defined to be the complex number

$$z * w = z_1\overline{w}_1 + \cdots + z_n\overline{w}_n, \tag{5.1.1}$$

where a bar denotes complex conjugation. The *Hermitian norm* of a vector z in $\mathbb{C}^n$ is defined to be the real number

$$|z| = (z * z)^{\frac{1}{2}}. \tag{5.1.2}$$

Obviously $|z| \geq 0$, since

$$|z| = (|z_1|^2 + \cdots + |z_n|^2)^{\frac{1}{2}}.$$

The Hermitian norm determines a metric on $\mathbb{C}^n$ in the usual way,

$$d_C(z, w) = |z - w|. \tag{5.1.3}$$

The metric space consisting of $\mathbb{C}^n$ together with the metric d_C is called *complex n-space*. Define $\phi : \mathbb{C}^n \to \mathbb{R}^{2n}$ by

$$\phi(z_1, \ldots, z_n) = (\mathrm{Re}\, z_1, \mathrm{Im}\, z_1, \ldots, \mathrm{Re}\, z_n, \mathrm{Im}\, z_n).$$

Then ϕ is obviously an isomorphism of real vector spaces. Moreover,

$$\phi(z) \cdot \phi(w) = \mathrm{Re}\,(z * w).$$

Consequently ϕ preserves norms. Therefore ϕ is an isometry. For this reason, we call d_C the *Euclidean metric* on $\mathbb{C}^n$.

Definition: A *topological group* is a group G that is also a topological space such that the multiplication $(g, h) \mapsto gh$ and inversion $g \mapsto g^{-1}$ in G are continuous functions.

The following are some familiar examples of topological groups:

(1) real n-space $\mathbb{R}^n$ with the operation of vector addition,

(2) complex n-space $\mathbb{C}^n$ with the operation of vector addition,

(3) the positive real numbers $\mathbb{R}_+$ with the operation of multiplication,

(4) the unit circle S^1 in the complex plane with the operation of complex multiplication,

(5) the nonzero complex numbers $\mathbb{C}^*$ with the operation of complex multiplication.

Definition: Two topological groups G and H are *isomorphic topological groups* if and only if there is an isomorphism $\phi : G \to H$ that is also a homeomorphism.

Example: The spaces $\mathbb{C}^n$ and $\mathbb{R}^{2n}$ are isomorphic topological groups.

The General Linear Group

Let $\mathrm{GL}(n, \mathbb{C})$ be the set of all invertible complex $n \times n$ matrices. Then $\mathrm{GL}(n, \mathbb{C})$ is a group under the operation of matrix multiplication. The group $\mathrm{GL}(n, \mathbb{C})$ is called the *general linear group* of complex $n \times n$ matrices.

The *norm* of a complex $n \times n$ matrix $A = (a_{ij})$ is defined to be the real number

$$|A| = \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{1/2}. \tag{5.1.4}$$

This norm determines a metric on $\mathrm{GL}(n, \mathbb{C})$ in the usual way,

$$d(A, B) = |A - B|. \tag{5.1.5}$$

Note that this is just the Euclidean metric on $\mathrm{GL}(n, \mathbb{C})$ regarded as a subset of $\mathbb{C}^{n^2}$. For this reason, we call d the *Euclidean metric* on $\mathrm{GL}(n, \mathbb{C})$.

Theorem 5.1.1. *The general linear group* $\mathrm{GL}(n,\mathbb{C})$, *with the Euclidean metric topology, is a topological group.*

Proof: Matrix multiplication $(A, B) \mapsto AB$ is continuous, since the entries of AB are polynomials in the entries of A and B. The determinant function

$$\det : \mathrm{GL}(n,\mathbb{C}) \to \mathbb{C}^*$$

is continuous, since $\det A$ is a polynomial in the entries of A. By the adjoint formula for A^{-1}, we have

$$(A^{-1})_{ji} = (-1)^{i+j}(\det A^{ij})/(\det A),$$

where A^{ij} is the matrix obtained from A by deleting the ith row and jth column. Consequently, each entry of A^{-1} is a rational function of the entries of A. Therefore, the inversion map $A \mapsto A^{-1}$ is continuous. Thus $\mathrm{GL}(n,\mathbb{C})$ is a topological group. □

Any subgroup H of a topological group G is a topological group with the subspace topology. Hence, each of the following subgroups of $\mathrm{GL}(n,\mathbb{C})$ is a topological group with the Euclidean metric topology:

(1) the special linear group $\mathrm{SL}(n,\mathbb{C})$ of all complex $n \times n$ matrices of determinant one,

(2) the general linear group $\mathrm{GL}(n,\mathbb{R})$ of all invertible real $n \times n$ matrices,

(3) the special linear group $\mathrm{SL}(n,\mathbb{R})$ of all real $n \times n$ matrices of determinant one,

(4) the orthogonal group $\mathrm{O}(n)$,

(5) the special orthogonal group $\mathrm{SO}(n)$,

(6) the Lorentz groups $\mathrm{O}(1, n-1)$ and $\mathrm{O}(n-1, 1)$,

(7) the positive Lorentz groups $\mathrm{PO}(1, n-1)$ and $\mathrm{PO}(n-1, 1)$.

The Unitary Group

A complex $n \times n$ matrix A is said to be *unitary* if and only if

$$(Az) * (Aw) = z * w$$

for all z, w in $\mathbb{C}^n$. Obviously, the set of all unitary matrices in $\mathrm{GL}(n,\mathbb{C})$ forms a subgroup $\mathrm{U}(n)$, called the *unitary group* of complex $n \times n$ matrices. A unitary matrix is real if and only if it is orthogonal. Therefore $\mathrm{U}(n)$ contains $\mathrm{O}(n)$ as a subgroup.

Two vectors z and w in $\mathbb{C}^n$ are said to be *orthogonal* if and only if $z * w = 0$. A basis $\{v_1, \ldots, v_n\}$ of $\mathbb{C}^n$ is said to be *orthonormal* if and only if $v_i * v_j = \delta_{ij}$ for all i, j. The next theorem characterizes a unitary matrix. The proof is left as an exercise for the reader.

Theorem 5.1.2. *Let A be a complex $n \times n$ matrix. Then the following are equivalent:*

(1) *The matrix A is unitary.*

(2) *The columns of A form an orthonormal basis of $\mathbb{C}^n$.*

(3) *The matrix A satisfies the equation $\overline{A}^t A = I$.*

(4) *The matrix A satisfies the equation $A\overline{A}^t = I$.*

(5) *The rows of A form an orthonormal basis of $\mathbb{C}^n$.*

Corollary 1. *A real matrix is unitary if and only if it is orthogonal.*

Let A be a unitary matrix. As $\overline{A}^t A = I$, we have that $|\det A| = 1$. Let $\mathrm{SU}(n)$ be the set of all A in $\mathrm{U}(n)$ such that $\det A = 1$. Then $\mathrm{SU}(n)$ is a subgroup of $\mathrm{U}(n)$. The group $\mathrm{SU}(n)$ is called the *special unitary group* of complex $n \times n$ matrices.

Theorem 5.1.3. *The unitary group* $\mathrm{U}(n)$ *is compact.*

Proof: If A is in $\mathrm{U}(n)$, then $|A|^2 = \sum_{j=1}^{n} |Ae_j|^2 = n$. Therefore $\mathrm{U}(n)$ is a bounded subset of $\mathbb{C}^{n^2}$. The function

$$f : \mathbb{C}^{n^2} \to \mathbb{C}^{n^2},$$

defined by $f(A) = \overline{A}^t A$, is continuous. Therefore $\mathrm{U}(n) = f^{-1}(I)$ is a closed subset of $\mathbb{C}^{n^2}$. Hence $\mathrm{U}(n)$ is a closed bounded subset of $\mathbb{C}^{n^2}$ and therefore is compact. □

Corollary 2. *The orthogonal group* $\mathrm{O}(n)$ *is compact.*

Proof: As $\mathbb{R}^{n^2}$ is closed in $\mathbb{C}^{n^2}$ and $\mathrm{O}(n) = \mathrm{U}(n) \cap \mathbb{R}^{n^2}$, we have that $\mathrm{O}(n)$ is closed in $\mathrm{U}(n)$, and so $\mathrm{O}(n)$ is compact. □

Quotient Topological Groups

Lemma 1. *If h is an element of a topological group G, then the maps*

$$g \mapsto hg \quad \text{and} \quad g \mapsto gh,$$

from G to itself, are homeomorphisms.

Proof: Both maps are continuous and have continuous inverses $g \mapsto h^{-1}g$ and $g \mapsto gh^{-1}$, respectively. □

Let H be a subgroup of a topological group G. The *coset space* G/H is the set of cosets $\{gH : g \in G\}$ with the quotient topology. The quotient map will be denoted by $\pi : G \to G/H$.

Lemma 2. *If H is a subgroup of a topological group G, then the quotient map $\pi : G \to G/H$ is an open map.*

Proof: Let U be open in G. Then $\pi(U)$ is open in G/H if and only if $\pi^{-1}(\pi(U))$ is open in G by the definition of the quotient topology on G/H. Now since

$$\pi^{-1}(\pi(U)) = UH = \bigcup_{h \in H} Uh,$$

we have that $\pi^{-1}(\pi(U))$ is open by Lemma 1. Thus π is an open map. □

Theorem 5.1.4. *Let N be a normal subgroup of a topological group G. Then G/N, with the quotient topology, is a topological group.*

Proof: Let $\pi : G \to G/N$ be the quotient map $g \mapsto gN$. Then we have a commutative diagram

$$\begin{array}{ccc} G & \xrightarrow{g \mapsto g^{-1}} & G \\ \pi \downarrow & & \downarrow \pi \\ G/N & \xrightarrow{gN \mapsto g^{-1}N} & G/N. \end{array}$$

This implies that the inversion map $gN \mapsto g^{-1}N$ is continuous.

Next, observe that we have a commutative diagram

$$\begin{array}{ccc} G \times G & \xrightarrow{(g,h) \mapsto gh} & G \\ \pi \times \pi \downarrow & & \downarrow \pi \\ G/N \times G/N & \xrightarrow{(gN,hN) \mapsto ghN} & G/N \end{array}$$

As π is an open map, $\pi \times \pi$ is also an open map. Consequently $\pi \times \pi$ is a quotient map. From the diagram, we deduce that the multiplication in G/N is continuous. □

By Theorem 5.1.4, the following quotient groups, with the quotient topology, are topological groups:

(1) the projective general linear group $\mathrm{PGL}(n, \mathbb{C}) = \mathrm{GL}(n, \mathbb{C})/N$, where N is the normal subgroup $\{kI : k \in \mathbb{C}^*\}$;

(2) the projective special linear group $\mathrm{PSL}(n, \mathbb{C}) = \mathrm{SL}(n, \mathbb{C})/N$, where N is the normal subgroup $\{wI : w \text{ is an } n\text{th root of unity}\}$;

(3) the projective general linear group $\mathrm{PGL}(n, \mathbb{R}) = \mathrm{GL}(n, \mathbb{R})/N$, where N is the normal subgroup $\{kI : k \in \mathbb{R}^*\}$;

(4) the projective special linear group $\mathrm{PSL}(2n,\mathbb{R}) = \mathrm{SL}(2n,\mathbb{R})/\{\pm I\}$;

(5) the projective special unitary group $\mathrm{PSU}(n) = \mathrm{SU}(n)/N$, where N is the normal subgroup $\{wI : w \text{ is an } n\text{th root of unity}\}$.

Theorem 5.1.5. *Let H be a subgroup of a topological group G, and let $\eta : G \to X$ be a continuous function such that $\eta^{-1}(\eta(g)) = gH$ for each g in G. If $\sigma : X \to G$ is a continuous right inverse of η, then the function $\phi : X \times H \to G$, defined by $\phi(x,h) = \sigma(x)h$, is a homeomorphism; moreover, the function $\overline{\eta} : G/H \to X$, induced by η, is a homeomorphism.*

Proof: The function ϕ is a composite of continuous functions and so is continuous. Let g be in G. As $\eta\sigma\eta(g) = \eta(g)$, we have that $\sigma\eta(g)$ is in gH, and so $g^{-1}\sigma\eta(g)$ is in H. Define a function

$$\psi : G \to X \times H$$

by the formula

$$\psi(g) = (\eta(g), [\sigma\eta(g)]^{-1}g).$$

The map ψ is the composite of continuous functions and so is continuous. Observe that

$$\begin{aligned} \phi\psi(g) &= \phi(\eta(g), [\sigma\eta(g)]^{-1}g) \\ &= \sigma\eta(g)[\sigma\eta(g)]^{-1}g \\ &= g \end{aligned}$$

and

$$\begin{aligned} \psi\phi(x,h) &= \psi(\sigma(x)h) \\ &= (\eta(\sigma(x)h), [\sigma\eta(\sigma(x)h)]^{-1}\sigma(x)h) \\ &= (\eta\sigma(x), [\sigma\eta\sigma(x)]^{-1}\sigma(x)h) \\ &= (x, [\sigma(x)]^{-1}\sigma(x)h) \\ &= (x,h). \end{aligned}$$

Thus ϕ is a homeomorphism with inverse ψ.

Let $\pi : G \to G/H$ be the quotient map. Then η induces a continuous bijection $\overline{\eta} : G/H \to X$ such that $\overline{\eta}\pi = \eta$. The map $\pi\sigma$ is a continuous inverse of $\overline{\eta}$, and so $\overline{\eta}$ is a homeomorphism. □

Exercise 5.1

1. Prove that $\mathbb{R}$ and $\mathbb{R}_+$ are isomorphic topological groups.
2. Prove that $\mathbb{C}^*$ and $\mathbb{R}_+ \times S^1$ are isomorphic topological groups.
3. Prove that S^1 and $\mathrm{SO}(2)$ are isomorphic topological groups.
4. Prove that $\mathbb{R}$ and $\mathrm{PSO}(1,1)$ are isomorphic topological groups.

5. Prove that if z, w are in $\mathbb{C}^n$, then $|z * w| \leq |z|\,|w|$ with equality if and only if z and w are linearly dependent.
6. Let A be a complex $n \times n$ matrix. Show that $|Az| \leq |A|\,|z|$ for all z in $\mathbb{C}^n$.
7. Let A, B be complex $n \times n$ matrices. Prove that $|AB| \leq |A|\,|B|$.
8. Let A, B be complex $n \times n$ matrices. Prove that $|A \pm B| \leq |A| + |B|$.
9. Prove Theorem 5.1.2.
10. Prove that a complex $n \times n$ matrix A is unitary if and only if $|Az| = |z|$ for all z in $\mathbb{C}^n$.
11. Let A be in $\mathrm{SL}(2, \mathbb{C})$. Show that the following are equivalent:

    (1) A is unitary;

    (2) $|A|^2 = 2$;

    (3) A is of the form $\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}$.
12. Let A be a complex 2×2 matrix. Show that $2|\det A| \leq |A|^2$.
13. Let $\pi : \mathrm{SL}(2, \mathbb{C}) \to \mathrm{PSL}(2, \mathbb{C})$ be the quotient map. Prove that π maps any open ball of radius $\sqrt{2}$ homeomorphically onto its image. Deduce that π is a double covering.
14. Prove that $\mathrm{PSL}(2, \mathbb{C})$ and $\mathrm{PGL}(2, \mathbb{C})$ are isomorphic topological groups.
15. Prove that $\mathrm{GL}(n, \mathbb{C})$ is homeomorphic to $\mathbb{C}^* \times \mathrm{SL}(n, \mathbb{C})$.

§5.2. Groups of Isometries

Let X be a metric space. Henceforth, we shall assume that the group $\mathrm{I}(X)$ of isometries of X and the group $\mathrm{S}(X)$ of similarities of X are topologized with the subspace topology inherited from the space $\mathrm{C}(X, X)$ of continuous self-maps of X with the compact-open topology.

Theorem 5.2.1. *A sequence $\{\phi_i\}$ of isometries of a metric space X converges in $\mathrm{I}(X)$ to an isometry ϕ if and only if $\{\phi_i(x)\}$ converges to $\phi(x)$ for each point x of X.*

Proof: It is a basic property of the compact-open topology of $\mathrm{C}(X, X)$ that $\phi_i \to \phi$ if and only if $\{\phi_i\}$ converges uniformly to ϕ on compact sets, that is, for each compact subset K of X and $\epsilon > 0$, there is an integer k such that $d(\phi_i(x), \phi(x)) < \epsilon$ for all $i \geq k$ and every x in K. If $\phi_i \to \phi$, then $\phi_i(x) \to \phi(x)$ for each x in X, since each point of X is compact.

Conversely, suppose that $\phi_i(x) \to \phi(x)$ for each x in X. Let K be a compact subset of X and let $\epsilon > 0$. On the contrary, suppose that $\{\phi_i\}$

does not converge uniformly on K. Then there is a subsequence $\{\phi_{i_j}\}$ of $\{\phi_i\}$ and a sequence $\{x_j\}$ of points of K such that for each j, we have

$$d(\phi_{i_j}(x_j), \phi(x_j)) \geq \epsilon.$$

By passing to a subsequence, we may assume that $\{x_j\}$ converges to a point x in K, since K is compact. Choose j large enough so that $d(x_j, x) < \epsilon/4$ and $d(\phi_{i_j}(x), \phi(x)) < \epsilon/2$. Then we have the contradiction

$$\begin{aligned} d(\phi_{i_j}(x_j), \phi(x_j)) &\leq d(\phi_{i_j}(x_j), \phi_{i_j}(x)) + d(\phi_{i_j}(x), \phi(x)) + d(\phi(x), \phi(x_j)) \\ &= 2d(x_j, x) + d(\phi_{i_j}(x), \phi(x)) \\ &< \epsilon. \end{aligned}$$

Therefore $\phi_i \to \phi$ uniformly on K. Thus $\phi_i \to \phi$. □

Definition: A metric space X is *finitely compact* if and only if all its closed balls are compact, that is,

$$C(a, r) = \{x \in X : d(a, x) \leq r\}$$

is compact for each point a of X and $r > 0$.

Theorem 5.2.2. *If X is a finitely compact metric space, then* $\mathrm{I}(X)$ *is a topological group.*

Proof: It is a basic property of the compact-open topology that the composition map $(\phi, \psi) \mapsto \phi\psi$ is continuous when X is locally compact. Now a finitely compact metric space has a countable basis. Consequently, $\mathrm{C}(X, X)$ and therefore $\mathrm{I}(X)$ has a countable basis. Hence, we can prove that the inversion map $\phi \mapsto \phi^{-1}$ is continuous using sequences. Suppose that $\phi_i \to \phi$ in $\mathrm{I}(X)$. Then $\phi_i(x) \to \phi(x)$ for each x in X. Let $\epsilon > 0$, let x be a point of X, and let $y = \phi^{-1}(x)$. Then there is an integer k such that for all $i \geq k$, we have $d(\phi_i(y), \phi(y)) < \epsilon$. Then for all $i \geq k$, we have

$$\begin{aligned} d(\phi_i^{-1}(x), \phi^{-1}(x)) &= d(x, \phi_i\phi^{-1}(x)) \\ &= d(\phi\phi^{-1}(x), \phi_i\phi^{-1}(x)) \\ &= d(\phi(y), \phi_i(y)) < \epsilon. \end{aligned}$$

Therefore $\phi_i^{-1}(x) \to \phi^{-1}(x)$. By Theorem 5.2.1, we have that $\phi_i^{-1} \to \phi^{-1}$. Hence, the inversion map is continuous. Thus $\mathrm{I}(X)$ is a topological group. □

Theorem 5.2.3. *The restriction map $\rho : \mathrm{O}(n) \to \mathrm{I}(S^n)$ is an isomorphism of topological groups.*

Proof: By Theorem 2.1.3, we have that ρ is an isomorphism. Thus, we only need to show that ρ is a homeomorphism. Suppose that $A_i \to A$ in $\mathrm{O}(n)$. Then obviously $A_i x \to Ax$ for all x in S^n. Therefore $A_i \to A$ in $\mathrm{I}(S^n)$ by Theorem 5.2.1. Conversely, suppose that $A_i \to A$ in $\mathrm{I}(S^n)$. Then $A_i e_j \to Ae_j$ for each $j = 1, \ldots, n+1$. Hence $A_i \to A$ in $\mathrm{O}(n)$. Thus ρ is a homeomorphism. □

Theorem 5.2.4. *The function $\Phi : E^n \times \mathrm{O}(n) \to \mathrm{I}(E^n)$, defined by the formula $\Phi(a, A) = a + A$, is a homeomorphism.*

Proof: Let $e : \mathrm{I}(E^n) \to E^n$ be the evaluation map defined by $e(\phi) = \phi(0)$. It is a basic property of the compact-open topology that the evaluation map e is continuous. Define $\tau : E^n \times E^n \to E^n$ by $\tau(a, x) = a + x$. Then τ is obviously continuous. It is a basic property of the compact-open topology that the corresponding function $\hat{\tau} : E^n \to \mathrm{I}(E^n)$, defined by $\hat{\tau}(a)(x) = a+x$, is also continuous. The map $\hat{\tau}$ is a right inverse for e.

We shall identify $\mathrm{O}(n)$ with the group of isometries of E^n that fix the origin. By the same argument as in the proof of Theorem 5.2.3, the compact-open topology on $\mathrm{O}(n)$ is the same as the Euclidean topology on $\mathrm{O}(n)$.

For each ϕ in $\mathrm{I}(E^n)$, we have

$$e^{-1}(e(\phi)) = \phi\mathrm{O}(n).$$

Therefore Φ is a homeomorphism by Theorem 5.1.5. □

The group $\mathrm{T}(E^n)$ of translations of E^n is a subgroup of $\mathrm{I}(E^n)$, and so $\mathrm{T}(E^n)$ is a topological group with the subspace topology. The next corollary follows immediately from Theorem 5.2.4.

Corollary 1. *The evaluation map $e : \mathrm{T}(E^n) \to E^n$, defined by the formula $e(\tau) = \tau(0)$, is an isomorphism of topological groups.*

Theorem 5.2.5. *The restriction map $\rho : \mathrm{PO}(n, 1) \to \mathrm{I}(H^n)$ is an isomorphism of topological groups.*

Proof: By Theorem 3.2.3, we have that ρ is an isomorphism. Thus, we only need to show that ρ is a homeomorphism. Suppose that $A_i \to A$ in $\mathrm{PO}(n, 1)$. Then obviously $A_i x \to Ax$ for all x in H^n. Therefore $A_i \to A$ in $\mathrm{I}(H^n)$ by Theorem 5.2.1. Conversely, suppose that $A_i \to A$ in $\mathrm{I}(H^n)$. Then $A_i e_{n+1} \to A e_{n+1}$. Now for each $j = 1, \ldots, n$, the vector $v_j = e_j + \sqrt{2} e_{n+1}$ is in H^n. Hence $A_i v_j \to A v_j$ for each $j = 1, \ldots, n$. Therefore, we have

$$A_i e_j + \sqrt{2} A_i e_{n+1} \to A e_j + \sqrt{2} A e_{n+1}.$$

Hence $A_i e_j \to A e_j$ for each $j = 1, \ldots, n$. Therefore $A_i \to A$ in $\mathrm{PO}(n, 1)$. Thus ρ is a homeomorphism. □

Groups of Möbius Transformations

Each Möbius transformation of B^n is completely determined by its action on $\partial B^n = S^{n-1}$ because of Poincaré extension. Consequently, the topology of S^{n-1} determines a natural topology on the group $\mathrm{M}(B^n)$. This topology is the metric topology defined by the metric

$$D_B(\phi, \psi) = \sup_{x \in S^{n-1}} |\phi(x) - \psi(x)|. \tag{5.2.1}$$

The metric topology determined by D_B on $\mathrm{M}(B^n)$ is a natural topology because it coincides with the compact-open topology inherited from the function space $\mathrm{C}(S^{n-1}, S^{n-1})$ of continuous self-maps of S^{n-1}.

Lemma 1. *If ϕ is in $\mathrm{M}(B^n)$, then*

$$\sup_{x,y\in S^{n-1}} \frac{|\phi(x)-\phi(y)|}{|x-y|} = \exp d_B(0,\phi(0)).$$

Proof: Suppose that $\phi(\infty)=\infty$. Then ϕ is orthogonal by Theorem 4.4.7. Hence, we have

$$\frac{|\phi(x)-\phi(y)|}{|x-y|} = 1 = \exp d_B(0,0).$$

Now suppose that $\phi(\infty)\neq\infty$. Then $\phi=\psi\sigma$, where σ is the reflection in a sphere $S(a,r)$ orthogonal to S^{n-1} and ψ is an orthogonal transformation. By Theorem 4.4.2(3), we have that $r^2=|a|^2-1$; and by Theorem 4.1.3,

$$\frac{|\phi(x)-\phi(y)|}{|x-y|} = \frac{r^2}{|x-a|\,|y-a|} = \frac{|a|^2-1}{|x-a|\,|y-a|}.$$

From the equation $|x-a|^2 = 1-2a\cdot x+|a|^2$, we see that the minimum value of $|x-a|$ occurs when $x=a/|a|$. Therefore

$$\sup_{x,y\in S^{n-1}} \frac{|\phi(x)-\phi(y)|}{|x-y|} = \frac{|a|^2-1}{(|a|-1)^2} = \frac{|a|+1}{|a|-1}.$$

Now since

$$\sigma(x) = a + \frac{|a|^2-1}{|x-a|^2}(x-a),$$

we have that $\sigma(0)=a/|a|^2$. Therefore $|a|=1/|\phi(0)|$. Hence

$$\frac{|a|+1}{|a|-1} = \frac{1+|\phi(0)|}{1-|\phi(0)|} = \exp d_B(0,\phi(0)).$$ □

Theorem 5.2.6. *The group $\mathrm{M}(B^n)$, with the metric topology determined by D_B, is a topological group.*

Proof: Let ϕ,ϕ_0,ψ,ψ_0 be in $\mathrm{M}(B^n)$. By Lemma 1, there is a positive constant $k(\phi)$ such that $|\phi(x)-\phi(y)|\leq k(\phi)|x-y|$ for all x,y in S^{n-1}. As ψ restricts to a bijection of S^{n-1}, we have $D(\phi\psi,\phi_0\psi)=D(\phi,\phi_0)$. Hence

$$\begin{aligned} D(\phi\psi,\phi_0\psi_0) &\leq D(\phi\psi,\phi_0\psi)+D(\phi_0\psi,\phi_0\psi_0)\\ &\leq D(\phi,\phi_0)+k(\phi_0)D(\psi,\psi_0). \end{aligned}$$

This implies that the composition map $(\phi,\psi)\mapsto\phi\psi$ is continuous at (ϕ_0,ψ_0). Similarly, the map $\phi\mapsto\phi^{-1}$ is continuous at ϕ_0, since

$$\begin{aligned} D(\phi^{-1},\phi_0^{-1}) &= D(\phi^{-1}\phi,\phi_0^{-1}\phi)\\ &= D(\phi_0^{-1}\phi_0,\phi_0^{-1}\phi)\\ &\leq k(\phi_0^{-1})D(\phi_0,\phi). \end{aligned}$$ □

Corollary 2. *The group* $\mathrm{M}(S^{n-1})$, *with the metric topology determined by* D_B, *is a topological group.*

Let η be the standard transformation from U^n to B^n. Then η induces an isomorphism $\eta_* : \mathrm{M}(U^n) \to \mathrm{M}(B^n)$ defined by $\eta_*(\phi) = \eta\phi\eta^{-1}$. The restriction of η to $\hat{E}^{n-1}$ is stereographic projection

$$\pi : \hat{E}^{n-1} \to S^{n-1}.$$

Let d be the chordal metric on $\hat{E}^{n-1}$. Define a metric D_U on $\mathrm{M}(U^n)$ by

$$D_U(\phi,\psi) = \sup_{x\in\hat{E}^{n-1}} d(\phi(x),\psi(x)). \tag{5.2.2}$$

Then

$$\begin{aligned} D_U(\phi,\psi) &= \sup_{x\in\hat{E}^{n-1}} |\pi\phi(x) - \pi\psi(x)| \\ &= \sup_{y\in S^{n-1}} |\pi\phi\pi^{-1}(y) - \pi\psi\pi^{-1}(y)| \\ &= D_B(\eta\phi\eta^{-1}, \eta\psi\eta^{-1}) \\ &= D_B(\eta_*(\phi), \eta_*(\psi)). \end{aligned}$$

Thus $\eta_* : \mathrm{M}(U^n) \to \mathrm{M}(B^n)$ is an isometry of metric spaces. The next theorem follows immediately from Theorem 5.2.6.

Theorem 5.2.7. *The group* $\mathrm{M}(U^n)$, *with the metric topology determined by* D_U, *is a topological group.*

Poincaré extension induces a homeomorphism from $\mathrm{M}(S^{n-1})$ to $\mathrm{M}(B^n)$. Therefore, Poincaré extension induces a homeomorphism from $\mathrm{M}(\hat{E}^{n-1})$ to $\mathrm{M}(U^n)$. This implies the following corollary.

Corollary 3. *The group* $\mathrm{M}(\hat{E}^{n-1})$, *with the metric topology determined by* D_U, *is a topological group.*

Theorem 5.2.8. *The function* $\Phi : B^n \times \mathrm{O}(n) \to \mathrm{M}(B^n)$, *defined by the formula* $\Phi(b,A) = \tau_b A$, *is a homeomorphism.*

Proof: Let $e : \mathrm{M}(B^n) \to B^n$ be the evaluation map defined by $e(\phi) = \phi(0)$. We now show that e is continuous. Suppose that $D(\phi, I) < r$. As each Euclidean diameter L_α of B^n is mapped by ϕ onto a hyperbolic line $\phi(L_\alpha)$ of B^n whose endpoints are a distance at most r from those of L_α, the Euclidean cylinder C_α with axis L_α and radius r contains $\phi(L_\alpha)$. Then e is continuous at the identity map I, since

$$\{\phi(0)\} \subset \bigcap_\alpha \phi(L_\alpha) \subset \bigcap_\alpha C_\alpha = \{x \in B^n : |x| < r\}.$$

Now suppose that $\{\phi_i\}$ is a sequence in $\mathrm{M}(B^n)$ converging to ϕ. Then $\phi^{-1}\phi_i$ converges to I, since $\mathrm{M}(B^n)$ is a topological group. As e is continuous at I, we have that $\phi^{-1}\phi_i(0)$ converges to 0. Therefore $\phi_i(0)$ converges to $\phi(0)$. Thus e is continuous.

Define $\partial\tau : B^n \times S^{n-1} \to S^{n-1}$ by $\partial\tau(b,x) = \tau_b(x)$. By Formula 4.5.5, we have that

$$\tau_b(x) = \frac{(1-|b|^2)}{|x+b|^2}x + \frac{2(1+x\cdot b)}{|x+b|^2}b.$$

Therefore $\partial\tau$ is continuous. Hence, the function $\partial\hat{\tau} : B^n \to \mathrm{M}(S^{n-1})$, defined by $\partial\hat{\tau}(b)(x) = \tau_b(x)$, is continuous, since the metric topology on $\mathrm{M}(S^{n-1})$, determined by D_B, is the same as the compact-open topology. Therefore, the function $\hat{\tau} : B^n \to \mathrm{M}(B^n)$, defined by $\hat{\tau}(b)(x) = \tau_b(x)$, is continuous, since the map from $\mathrm{M}(S^{n-1})$ to $\mathrm{M}(B^n)$, induced by Poincaré extension, is a homeomorphism. The map $\hat{\tau}$ is a right inverse of e.

Let ϕ be in $\mathrm{M}(B^n)$. Then clearly $\phi\mathrm{O}(n)$ is contained in $e^{-1}(e(\phi))$. Suppose that ψ is in $e^{-1}(e(\phi))$. Then $\psi(0) = \phi(0)$ and so $\phi^{-1}\psi(0) = 0$. By Theorem 4.4.8, we have that $\phi^{-1}\psi$ is in $\mathrm{O}(n)$. Therefore ψ is in $\phi\mathrm{O}(n)$. Thus $e^{-1}(e(\phi)) = \phi\mathrm{O}(n)$. Hence Φ is a homeomorphism by Theorem 5.1.5. □

Theorem 5.2.9. *The function $\Psi : B^n \times \mathrm{O}(n) \to \mathrm{I}(B^n)$, defined by the formula $\Psi(b, A) = \tau_b A$, is a homeomorphism.*

Proof: Let $e : \mathrm{I}(B^n) \to B^n$ be the evaluation map defined by $e(\phi) = \phi(0)$. Then e is continuous. Define $\tau : B^n \times B^n \to B^n$ by $\tau(b,x) = \tau_b(x)$. Let b and x be in B^n. Then by Formula 4.5.5, we have

$$\tau_b(x) = \frac{(1-|b|^2)}{(|b|^2|x|^2 + 2x\cdot b + 1)}x + \frac{(|x|^2 + 2x\cdot b + 1)}{(|b|^2|x|^2 + 2x\cdot b + 1)}b.$$

Hence τ is continuous. Therefore, the function $\hat{\tau} : B^n \to \mathrm{I}(B^n)$, defined by $\hat{\tau}(b)(x) = \tau_b(x)$, is continuous. The map $\hat{\tau}$ is a right inverse of e.

We shall identify $\mathrm{O}(n)$ with the group of all isometries of B^n that fix the origin. By the same argument as in the proof of Theorem 5.2.3, with e_j replaced by $e_j/2$, the compact-open topology on $\mathrm{O}(n)$ is the same as the Euclidean topology on $\mathrm{O}(n)$. As $e^{-1}(e(\phi)) = \phi\mathrm{O}(n)$, we have that Ψ is a homeomorphism by Theorem 5.1.5. □

Theorem 5.2.10. *The restriction map $\rho : \mathrm{M}(B^n) \to \mathrm{I}(B^n)$ is an isomorphism of topological groups.*

Proof: The map ρ is an isomorphism by Theorem 4.5.2. The functions $\Phi : B^n \times \mathrm{O}(n) \to \mathrm{M}(B^n)$ and $\Psi : B^n \times \mathrm{O}(n) \to \mathrm{I}(B^n)$ are homeomorphisms by Theorems 5.2.8 and 5.2.9. As $\rho = \Psi\Phi^{-1}$, we have that ρ is a homeomorphism. □

The next theorem follows immediately from Theorem 5.2.10.

Theorem 5.2.11. *The restriction map $\rho : \mathrm{M}(U^n) \to \mathrm{I}(U^n)$ is an isomorphism of topological groups.*

The group $\mathrm{S}(E^{n-1})$ of similarities of E^{n-1} is isomorphic, by extension to ∞, to the group $\mathrm{M}(\hat{E}^{n-1})_\infty$ of transformations in $\mathrm{M}(\hat{E}^{n-1})$ fixing ∞.

Theorem 5.2.12. *The restriction map $\rho : \mathrm{M}(\hat{E}^{n-1})_\infty \to \mathrm{S}(E^{n-1})$ is an isomorphism of topological groups.*

Proof: The metric topology on $\mathrm{M}(\hat{E}^{n-1})_\infty$ is the same as the compact-open topology, since $\hat{E}^{n-1}$ is compact. Suppose that $\psi_i \to \psi$ in $\mathrm{M}(\hat{E}^{n-1})_\infty$. Then $\psi_i(x) \to \psi(x)$ for each point x in E^{n-1}. By essentially the same argument as in the proof of Theorem 5.2.1 (see Exercise 5.2.2), we have that $\rho(\psi_i) \to \rho(\psi)$. Therefore ρ is continuous.

Suppose that $\phi_i \to \phi$ in $\mathrm{S}(E^{n-1})$. Then $\phi_i(x) \to \phi(x)$ for each point x in E^{n-1}. Let $\tilde{\phi}$ be the Poincaré extension of ϕ. Then obviously $\tilde{\phi}_i(x) \to \tilde{\phi}(x)$ for each point x in U^n. Hence $\tilde{\phi}_i \to \tilde{\phi}$ in $\mathrm{M}(U^n)$ by Theorems 5.2.1 and 5.2.11. Let $\hat{\phi} : \hat{E}^{n-1} \to \hat{E}^{n-1}$ be the extension of ϕ defined by $\hat{\phi}(\infty) = \infty$. Then $\hat{\phi}_i \to \hat{\phi}$, since Poincaré extension induces a homeomorphism from $\mathrm{M}(\hat{E}^{n-1})$ to $\mathrm{M}(U^n)$. As $\rho(\hat{\phi}) = \phi$, we have that $\rho^{-1}(\phi_i) \to \rho^{-1}(\phi)$. Hence ρ^{-1} is continuous. Thus ρ is a homeomorphism. □

Exercise 5.2

1. Let $\xi : X \to Y$ be an isometry of finitely compact metric spaces. Prove that the function $\xi_* : \mathrm{I}(X) \to \mathrm{I}(Y)$, defined by $\xi_*(\phi) = \xi\phi\xi^{-1}$, is an isomorphism of topological groups.
2. Let X be a metric space. Prove that $\phi_i \to \phi$ in $\mathrm{S}(X)$ if and only if $\phi_i(x) \to \phi(x)$ for each point x of X.
3. Let X be a finitely compact metric space. Prove that $\mathrm{S}(X)$ is a topological group.
4. Prove that the function $\Phi : E^n \times \mathbb{R}_+ \times \mathrm{O}(n) \to \mathrm{S}(E^n)$, defined by the formula $\Phi(a, k, A) = a + kA$, is a homeomorphism.
5. Let $\mathrm{S}(E^n)_0$ be the subgroup of $\mathrm{S}(E^n)$ of all similarities that fix the origin. Prove that the map $\Psi : \mathbb{R}_+ \times \mathrm{O}(n) \to \mathrm{S}(E^n)_0$, defined by $\Psi(k, A) = kA$, is an isomorphism of topological groups.
6. Let $E(n)$ be the group of all real $(n+1) \times (n+1)$ matrices of the form
$$A_a = \begin{pmatrix} & & & a_1 \\ & A & & \vdots \\ & & & a_n \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$
where A is an $n \times n$ orthogonal matrix and a is a point of E^n. Prove that the function $\eta : \mathrm{I}(E^n) \to E(n)$, defined by $\eta(a + A) = A_a$, is an isomorphism of topological groups.

7. Let $\Xi : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{LF}(\hat{\mathbb{C}})$ be defined by
$$\Xi\begin{pmatrix} a & b \\ c & d \end{pmatrix}(z) = \frac{az+b}{cz+d}.$$
Prove that Ξ is continuous. Here $\mathrm{SL}(2,\mathbb{C})$ has the Euclidean metric topology and $\mathrm{LF}(\hat{\mathbb{C}})$ has the compact-open topology.

8. Prove that a homomorphism $\eta : G \to H$ of topological groups is continuous if and only if η is continuous at the identity element 1 of G.

9. Let $\phi(z) = \frac{az+b}{cz+d}$ be in $\mathrm{LF}(\hat{\mathbb{C}})$ with $ad - bc = 1$ and $d \neq 0$. Show that
   (1) $d^2 = \frac{1}{\phi(1)-\phi(0)} - \frac{1}{\phi(\infty)-\phi(0)}$,
   (2) $cd = \frac{1}{\phi(\infty)-\phi(0)}$,
   (3) $b/d = \phi(0)$,
   (4) $ad = \frac{\phi(\infty)}{\phi(\infty)-\phi(0)}$.

10. Prove that $\mathrm{PSL}(2,\mathbb{C})$ and $\mathrm{LF}(\hat{\mathbb{C}})$ are isomorphic topological groups.

11. Let $\phi(z) = \frac{az+b}{cz+d}$ be in $\mathrm{LF}(\hat{\mathbb{C}})$ with $ad - bc = 1$. Prove that $\tilde{\phi}(j) = j$ in U^3 if and only if $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is unitary.

12. Prove that $\mathrm{PSU}(2)$ and $\mathrm{SO}(3)$ are isomorphic topological groups.

13. Let $\mathcal{H}$ be the set all matrices of the form $\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix}$ with z, w in $\mathbb{C}$. Show that $\mathcal{H}$, with matrix addition and multiplication, is isomorphic to the ring of quaternions $\mathbb{H}$ via the mapping
$$\begin{pmatrix} z & w \\ -\overline{w} & \overline{z} \end{pmatrix} \mapsto z + wj.$$

14. Prove that $\mathrm{SU}(2)$ and the group S^3 of unit quaternions are isomorphic topological groups.

15. Prove that $S^3/\{\pm 1\}$ and $\mathrm{SO}(3)$ are isomorphic topological groups.

§5.3. Discrete Groups

In this section, we study the basic properties of discrete groups of isometries of S^n, E^n, and H^n.

Definition: A *discrete group* is a topological group Γ all of whose points are open.

Lemma 1. *If Γ is a topological group, then Γ is discrete if and only if $\{1\}$ is open in Γ.*

Proof: If Γ is discrete, then $\{1\}$ is open. Conversely, suppose that $\{1\}$ is open. Let g be in Γ. Then left multiplication by g is a homeomorphism of Γ. Hence $g\{1\} = \{g\}$ is open in Γ. □

Any group Γ can be made into a discrete group by giving Γ the discrete topology. Therefore, the topology of a discrete group is not very interesting. What is interesting is the study of discrete subgroups of a continuous group like $\mathbb{R}^n$ or $\mathrm{GL}(n, \mathbb{C})$. Here are some examples of discrete subgroups of familiar continuous groups.

(1) The integers $\mathbb{Z}$ is a discrete subgroup of $\mathbb{R}$.

(2) The Gaussian integers $\mathbb{Z}[i] = \{m + ni : m, n \in \mathbb{Z}\}$ is a discrete subgroup of $\mathbb{C}$.

(3) The set $\{k^n : n \in \mathbb{Z}\}$ is a discrete subgroup of $\mathbb{R}_+$ for each $k > 0$.

(4) The group of nth roots of unity $\{\exp(i2\pi m/n) : m = 0, 1, \ldots, n-1\}$ is a discrete subgroup of S^1 for each positive integer n.

(5) The set $\{k^n : n \in \mathbb{Z}\}$ is a discrete subgroup of $\mathbb{C}^*$ for each k in $\mathbb{C}^* - S^1$.

Lemma 2. *A metric space X is discrete if and only if every convergent sequence $\{x_n\}$ in X is eventually constant.*

Proof: Suppose that X is discrete and $x_n \to x$ in X. Then there is an $r > 0$ such that $B(x, r) = \{x\}$. As $x_n \to x$, there is an integer m such that x_n is in $B(x, r)$ for all $n \geq m$. Thus $x_n = x$ for all $n \geq m$.

Conversely, suppose that every convergent sequence in X is eventually constant and X is not discrete. Then there is a point x such that $\{x\}$ is not open. Therefore $B(x, 1/n) \neq \{x\}$ for each integer $n > 0$. Choose x_n in $B(x, 1/n)$ different from x. Then $x_n \to x$, but $\{x_n\}$ is not eventually constant, which is a contradiction. Therefore X must be discrete. □

Lemma 3. *If G is a topological group with a metric topology, then every discrete subgroup of G is closed in G.*

Proof: Let Γ be a discrete subgroup of G and suppose that $G - \Gamma$ is not open. Then there is a g in $G - \Gamma$ and g_n in $B(g, 1/n) \cap \Gamma$ for each integer $n > 0$. As $g_n \to g$ in G, we have that $g_n g_{n+1}^{-1} \to 1$ in Γ. But $\{g_n g_{n+1}^{-1}\}$ is not eventually constant, which contradicts Lemma 2. Therefore, the set $G - \Gamma$ must be open, and so Γ is closed in G. □

Theorem 5.3.1. *A subgroup Γ of $\mathrm{U}(n)$ is discrete if and only if Γ is finite.*

Proof: If Γ is finite, then Γ is obviously discrete. Conversely, suppose that Γ is discrete. Then Γ is closed in $\mathrm{U}(n)$ by Lemma 3. Therefore Γ is compact, since $\mathrm{U}(n)$ is compact. As Γ is discrete, it must be finite. □

Corollary 1. *A subgroup Γ of $\mathrm{O}(n)$ is discrete if and only if Γ is finite.*

Definition: The *group of symmetries* of a subset S of a metric space X is the group of all isometries of X that leave S invariant.

Example 1. It has been known since antiquity that the five regular solids can be inscribed in a sphere; in fact, a construction is given in Book 13 of Euclid's Elements. The group of symmetries of a regular solid P inscribed in S^2 is a finite subgroup of $\mathrm{O}(3)$ whose order is

(1) 24 if P is a tetrahedron,

(2) 48 if P is a cube or octahedron,

(3) 120 if P is a dodecahedron or icosahedron.

Theorem 5.3.2. *A subgroup Γ of $\mathbb{R}^n$ is discrete if and only if Γ is generated by a set of linearly independent vectors.*

Proof: We may assume that Γ is nontrivial. Suppose that Γ is generated by a set $\{v_1, \ldots, v_m\}$ of linearly independent vectors. Then

$$\Gamma = \mathbb{Z}v_1 \oplus \cdots \oplus \mathbb{Z}v_m.$$

By applying a nonsingular linear transformation, we may assume that $v_i = e_i$ for each $i = 1, \ldots, m$. Then $\Gamma \cap B(0,1) = \{0\}$. Therefore Γ is discrete by Lemma 1.

Conversely, suppose that Γ is discrete. This part of the proof is by induction on n. Assume first that $n = 1$. Let $r > 0$ be such that $B(0,r)$ contains a nonzero element of Γ. Then $C(0,r) \cap \Gamma$ is a closed subset of $C(0,r)$ by Lemma 3. Hence $C(0,r) \cap \Gamma$ is a compact discrete space and therefore is finite. Thus, there is a nonzero element u in Γ nearest to 0. By replacing u by $-u$, if necessary, we may assume that u is positive. Let v be an arbitrary element in Γ. Then there is an integer k such that v is in the interval $[ku, (k+1)u)$. Hence $v - ku$ is in the set

$$\Gamma \cap [0, u) = \{0\}.$$

Therefore $v = ku$. Thus u generates Γ.

Now assume that $n > 1$ and every discrete subgroup of $\mathbb{R}^{n-1}$ is generated by a set of linearly independent vectors. As above, there is a nonzero element u in Γ nearest to 0 and

$$\Gamma \cap \mathbb{R}u = \mathbb{Z}u.$$

Let $u_1, \ldots, u_n$ be a basis of $\mathbb{R}^n$ with $u_n = u$, and let $\eta : \mathbb{R}^n \to \mathbb{R}^{n-1}$ be the linear transformation defined by $\eta(u_i) = e_i$ for $i = 1, \ldots, n-1$ and $\eta(u) = 0$. Then η is a continuous function such that $\eta^{-1}(\eta(x)) = x + \mathbb{R}u$ for all x in $\mathbb{R}^n$. Define a linear transformation $\sigma : \mathbb{R}^{n-1} \to \mathbb{R}^n$ by $\sigma(e_i) = u_i$ for $i = 1, \ldots, n-1$. Then σ is a continuous right inverse of η. By Theorem 5.1.5, the map $\overline{\eta} : \mathbb{R}^n/\mathbb{R}u \to \mathbb{R}^{n-1}$ induced by η is an isomorphism of topological groups.

Let $\pi : \mathbb{R}^n \to \mathbb{R}^n/\mathbb{R}u$ be the quotient map. We claim that $\pi(\Gamma)$ is a discrete subgroup of $\mathbb{R}^n/\mathbb{R}u$. Let $\{v_i\}$ be a sequence in Γ such that $\pi(v_i) \to 0$ in $\pi(\Gamma)$. Then $\overline{\eta}\pi(v_i) \to 0$ in $\mathbb{R}^{n-1}$ and so $\eta(v_i) \to 0$ in $\mathbb{R}^{n-1}$. Therefore $\sigma\eta(v_i) \to 0$ in $\mathbb{R}^n$. Hence $v_i \to 0 \pmod{\mathbb{R}u}$. Consequently, there are real numbers r_i such that $v_i - r_i u \to 0$ in $\mathbb{R}^n$. By adding a suitable integral multiple of u to v_i, we may assume that $|r_i| \leq 1/2$. For large enough i, we have that

$$|v_i - r_i u| < |u|/2.$$

Whence, we have

$$\begin{aligned} |v_i| &\leq |v_i - r_i u| + |r_i u| \\ &< |u|/2 + |u|/2 \quad = \quad |u|. \end{aligned}$$

Therefore $v_i = 0$ for all sufficiently large i. Consequently, every convergent sequence in $\pi(\Gamma)$ is eventually constant. Thus $\pi(\Gamma)$ is a discrete subgroup of $\mathbb{R}^n/\mathbb{R}u$ by Lemma 2. By the induction hypothesis, there are vectors $w_1, \ldots, w_m$ in Γ such that $\pi(w_1), \ldots, \pi(w_m)$ are linearly independent in $\mathbb{R}^n/\mathbb{R}u$ and generate $\pi(\Gamma)$. Therefore $u, w_1, \ldots, w_m$ are linearly independent in $\mathbb{R}^n$ and generate Γ. This completes the induction. □

Definition: A *lattice* of $\mathbb{R}^n$ is a subgroup generated by n linearly independent vectors of $\mathbb{R}^n$.

Corollary 2. *Every lattice of $\mathbb{R}^n$ is a discrete subgroup of $\mathbb{R}^n$.*

Example 2. Let Γ be the set of points of $\mathbb{R}^4$ of the form $\frac{1}{2}(m, n, p, q)$ where m, n, p, q are either all odd integers or all even integers. Then Γ is a lattice of $\mathbb{R}^4$. This lattice is interesting because it has 24 unit vectors $\pm e_i$ for $i = 1, 2, 3, 4$ and $\left(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}\right)$ all of which are a nearest neighbor to 0 in Γ. It is worth noting that these 24 points are the vertices of a regular polyhedron in $\mathbb{R}^4$ called the *24-cell.*

Let $\hat{\mathrm{SL}}(n, \mathbb{C})$ be the group of complex $n \times n$ matrices whose determinant is ± 1. Then $\hat{\mathrm{SL}}(n, \mathbb{C})$ is a subgroup of $\mathrm{GL}(n, \mathbb{C})$ containing $\mathrm{SL}(n, \mathbb{C})$ as a subgroup of index two.

Theorem 5.3.3. *A subgroup Γ of $\hat{\mathrm{SL}}(n, \mathbb{C})$ is discrete if and only if for each $r > 0$, the set $\{A \in \Gamma : |A| \leq r\}$ is finite.*

Proof: Suppose that $\{A \in \Gamma : |A| \leq r\}$ is finite for each $r > 0$. Let $B_j \to B$ in Γ. As the norm function is continuous, $|B_j| \to |B|$. Hence, there is an integer k such that $|\,|B_j| - |B|\,| \leq 1$ for all $j \geq k$. Now the set $\{A \in \Gamma : |A| \leq 1 + |B|\}$ is finite. Hence $\{B_j\}$ is eventually constant. Therefore Γ is discrete by Lemma 2.

Conversely, suppose that Γ is discrete and the set $\{A \in \Gamma : |A| \leq r\}$ is infinite for some $r > 0$. Then there is an infinite sequence $\{A_j\}$ of distinct elements in Γ such that $|A_j| \leq r$ for all j. As the set

$$\{x \in \mathbb{C}^{n^2} : |x| \leq r\}$$

is compact, the sequence $\{A_j\}$ contains a convergent subsequence. By passing to this subsequence, we may assume that $A_j \to A$. As the determinant function $\det : \mathbb{C}^{n^2} \to \mathbb{C}$ is continuous, the set

$$\hat{\mathrm{SL}}(n, \mathbb{C}) = \det{}^{-1}\{-1, 1\}$$

is closed in $\mathbb{C}^{n^2}$. Hence $\hat{\mathrm{SL}}(n, \mathbb{C})$ contains its limit point A. Consequently $A_j A_{j+1}^{-1} \to I$ in Γ. But the sequence $\{A_j A_{j+1}^{-1}\}$ is not eventually constant, contrary to Lemma 2. Thus, the set $\{A \in \Gamma : |A| \leq r\}$ is finite for all $r > 0$. □

Corollary 3. *Every discrete subgroup Γ of $\hat{\mathrm{SL}}(n, \mathbb{C})$ is countable.*

Proof: Let $\Gamma_m = \{A \in \Gamma : |A| \leq m\}$. Then $\Gamma = \bigcup_{m=1}^{\infty} \Gamma_m$ is countable. □

Example 3. Observe that the *modular group* $\mathrm{SL}(n, \mathbb{Z})$ and the *unimodular group* $\mathrm{GL}(n, \mathbb{Z})$ are discrete subgroups of $\hat{\mathrm{SL}}(n, \mathbb{C})$ by Theorem 5.3.3.

Discontinuous Groups

Let G be a group acting on a set X and let x be an element of X.

(1) The subgroup $G_x = \{g \in G : gx = x\}$ of G is called the *stabilizer* of x in G.

(2) The subset $Gx = \{gx : g \in G\}$ of X is called the *G-orbit* through x. The G-orbits partition X.

(3) Define a function $\phi : G/G_x \to Gx$ by $\phi(gG_x) = gx$. Then ϕ is a bijection. Therefore, the index of G_x in G is the cardinality of the orbit Gx.

Definition: A group G acts *discontinuously* on a topological space X if and only if G acts on X and for each compact subset K of X, the set $K \cap gK$ is nonempty for only finitely many g in G.

Lemma 4. *If a group G acts discontinuously on a topological space X, then each stabilizer subgroup of G is finite.*

Proof: Let x be a point of X. Then the stabilizer G_x of x in G is finite, since $\{x\}$ is compact. □

Definition: A collection $\mathcal{S}$ of subsets of a topological space X is *locally finite* if and only if for each point x of X, there is an open neighborhood U of x in X such that U meets only finitely many members of $\mathcal{S}$.

Clearly, any subcollection of a locally finite collection $\mathcal{S}$ is also locally finite. Another useful fact is that the union of the members of a locally finite collection $\mathcal{S}$ of closed sets is closed.

Lemma 5. *If a group G acts discontinuously on a metric space X, then each G-orbit is a closed discrete subset of X.*

Proof: Let x be a point of X. We now show that the collection of one-point subsets of Gx is locally finite. On the contrary, suppose that y is a point of X such that every neighborhood of y contains infinitely many points of Gx. Since X is a metric space, there is an infinite sequence $\{g_i\}$ of distinct elements of G such that $\{g_i x\}$ converges to y. Then

$$K = \{x, y, g_1x, g_2x, \ldots\}$$

is a compact subset of X. As $g_i x$ is in $K \cap g_i K$ for each i, we have a contradiction. Thus $\{\{gx\} : g \in G\}$ is a locally finite family of closed subsets of X. Hence, every subset of Gx is closed in X. Therefore Gx is a closed discrete subset of X. □

Definition: A group G of homeomorphisms of a topological space X is *discontinuous* if and only if G acts discontinuously on X.

Theorem 5.3.4. *Let Γ be a group of similarities of a metric space X. Then Γ is discontinuous if and only if*

(1) *each stabilizer subgroup of Γ is finite, and*

(2) *each Γ-orbit is a closed discrete subset of X.*

Proof: If Γ is discontinuous, then Γ satisfies (1) and (2) by Lemmas 4 and 5. Conversely, suppose that Γ satisfies (1) and (2). On the contrary, suppose that Γ is not discontinuous. Then there is a compact subset K of X and an infinite sequence $\{g_i\}$ of distinct elements of Γ such that K and $g_i K$ overlap. Now $g_i^{-1}K$ and K also overlap. By passing to a subsequence, we may assume that $g_i \neq g_j^{-1}$ for all $i \neq j$, and by replacing g_i with g_i^{-1}, if necessary, we may assume that the scale factor k_i of g_i is at most one. Now for each i, there is a point x_i in K such that $g_i x_i$ is in K. As K is compact, the sequence $\{x_i\}$ has a limit point x in K. By passing to a subsequence, we may assume that $\{x_i\}$ converges to x. Likewise, we may assume that $\{g_i x_i\}$ converges to a point y in K. Now observe that

$$\begin{aligned} d(g_i x, y) &\leq d(g_i x, g_i x_i) + d(g_i x_i, y) \\ &= k_i d(x, x_i) + d(g_i x_i, y). \end{aligned}$$

Hence $\{g_i x\}$ converges to y. For each i, there are only finitely many j such that $g_i x = g_j x$ by (1). Hence, there is an infinite subsequence of $\{g_i x\}$, whose terms are all distinct, converging to y; but this contradicts (2). Thus Γ is discontinuous. □

Lemma 6. *If X is a finitely compact metric space, then* $\mathrm{I}(X)$ *is closed in the space* $\mathrm{C}(X, X)$ *of all continuous self-maps of* X.

Proof: The space X has a countable basis, since X is finitely compact. Therefore $\mathrm{C}(X, X)$ has a countable basis. Hence $\mathrm{I}(X)$ is closed in $\mathrm{C}(X, X)$ if and only if every infinite sequence of elements of $\mathrm{I}(X)$ that converges in $\mathrm{C}(X, X)$ converges in $\mathrm{I}(X)$.

Let $\{\phi_i\}$ be a sequence in $\mathrm{I}(X)$ that converges to a map $\phi : X \to X$. Then for each pair of points x, y of X, we have that

$$d(\phi_i(x), \phi_i(y)) \to d(\phi(x), \phi(y)).$$

Therefore, we have

$$d(x, y) = d(\phi(x), \phi(y)).$$

Hence ϕ preserves distances.

We now show that ϕ is surjective. Let a be a base point of X and let $C(a, r)$ be the closed ball centered at a of radius $r > 0$. Then the set $\phi(C(a, 2r))$ is closed in X, since $C(a, 2r)$ is compact. On the contrary, suppose that y is a point of $C(\phi(a), r)$ that is not in $\phi(C(a, 2r))$. Set

$$s = \mathrm{dist}\big(y, \phi(C(a, 2r))\big).$$

Then $0 < s \leq r$. As $\phi_i \to \phi$ uniformly on $C(a, 2r)$, there is an index j such that

$$d(\phi_j(x), \phi(x)) < s$$

for each point x in $C(a, 2r)$. Observe that

$$d(y, \phi_j(a)) \leq d(y, \phi(a)) + d(\phi(a), \phi_j(a)) \leq r + s \leq 2r.$$

Therefore y is in $C(\phi_j(a), 2r)$. As ϕ_j maps $C(a, 2r)$ onto $C(\phi_j(a), 2r)$, there is a point x in $C(a, 2r)$ such that $\phi_j(x) = y$. Then we have the contradiction

$$d(y, \phi(x)) = d(\phi_j(x), \phi(x)) < s.$$

Therefore, we have that

$$C(\phi(a), r) \subset \phi(C(a, 2r)).$$

As r is arbitrary, ϕ must be surjective. Hence ϕ is an isometry. Therefore, the sequence $\{\phi_i\}$ converges in $\mathrm{I}(X)$. Thus $\mathrm{I}(X)$ is closed in $\mathrm{C}(X, X)$. □

Theorem 5.3.5. *Let X be a finitely compact metric space. Then a group Γ of isometries of X is discrete if and only if Γ is discontinuous.*

Proof: Suppose that Γ is discontinuous. Let x be a point of X. Then the orbit Γx is discrete and the stabilizer subgroup Γ_x is finite by Theorem 5.3.4. Let $\varepsilon_x : \Gamma \to \Gamma x$ be the evaluation map at x. Then ε_x is continuous. Hence, the set $\varepsilon_x^{-1}(x) = \Gamma_x$ is open in Γ. Therefore, the identity map of X is open in Γ, and so Γ is discrete by Lemma 1.

Conversely, suppose that Γ is discrete. Now X has a countable basis, since X is finitely compact. Therefore $\mathrm{C}(X, X)$ has a countable basis. Moreover $\mathrm{C}(X, X)$ is regular, since X is regular. Therefore $\mathrm{C}(X, X)$ is metrizable. Hence Γ is closed in $\mathrm{I}(X)$ by Lemma 3, and so Γ is closed in $\mathrm{C}(X, X)$ by Lemma 6.

On the contrary, suppose that Γ is not discontinuous. Then there is a point y of X and an infinite sequence $S = \{\phi_i\}$ of distinct elements of Γ such that the sequence $\{\phi_i(y)\}$ converges to a point of X. The set S is closed in $\mathrm{C}(X, X)$, since Γ is a closed discrete subset of $\mathrm{C}(X, X)$. The set S is equicontinuous on X, since for each x in X, $r > 0$, and i, we have

$$\phi_i(B(x, r)) = B(\phi_i(x), r)).$$

Let x be an arbitrary point of X. Then $\varepsilon_x(S) = \{\phi_i(x)\}$. Observe that for all i, we have that

$$d(\phi_i(x), \phi_i(y)) = d(x, y).$$

Let $r = d(x, y)$. Then we have that

$$\{\phi_i(x)\} \subset \overline{N}(\{\phi_i(y)\}, r),$$

which is compact, since $\{\phi_i(y)\}$ is bounded. Hence $\overline{\varepsilon_x(S)}$ is compact. Therefore S is compact by the Arzela-Ascoli theorem. As S is discrete, we have the contradiction that S is finite. Thus Γ is discontinuous. □

Exercise 5.3

1. Prove that a subgroup Γ of $\mathbb{R}_+$ is discrete if and only if there is a $k > 0$ such that $\Gamma = \{k^m : m \in \mathbb{Z}\}$.
2. Prove that a subgroup Γ of S^1 is discrete if and only if Γ is the group of nth roots of unity for some n.
3. Prove that every finite group of order $n + 1$ is isomorphic to a subgroup of $\mathrm{O}(n)$. Hint: Consider the group of symmetries of a regular n-simplex inscribed in S^{n-1}.
4. Prove that the *projective modular group* $\mathrm{PSL}(2n, \mathbb{Z}) = \mathrm{SL}(2n, \mathbb{Z})/\{\pm I\}$ is a discrete subgroup of $\mathrm{PSL}(2n, \mathbb{R})$.
5. Prove that the *elliptic modular group*, of all linear fractional transformations $\phi(z) = \frac{az+b}{cz+d}$ where a, b, c, d are integers and $ad-bc = 1$, is a discrete subgroup of $\mathrm{LF}(\hat{\mathbb{C}})$.

6. Prove that *Picard's group* $\mathrm{PSL}(2, \mathbb{Z}[i]) = \mathrm{SL}(2, \mathbb{Z}[i])/\{\pm I\}$ is a discrete subgroup of $\mathrm{PSL}(2, \mathbb{C})$.
7. Let G be a group acting on a set X. Prove that
   (1) the G-orbits partition X;
   (2) the function $\phi : G/G_x \to Gx$, defined by $\phi(gG_x) = gx$, is a bijection for each x in X.
8. Prove that a discrete group Γ of isometries of a finitely compact metric space X is countable.
9. Let Γ be the group generated by a magnification of E^n. Prove that
   (1) Γ is a discrete subgroup of $\mathrm{S}(E^n)$;
   (2) Γ does not act discontinuously on E^n;
   (3) Γ acts discontinuously on $E^n - \{0\}$.
10. Let $X = S^n, E^n$, or H^n. Prove that a subgroup Γ of $\mathrm{I}(X)$ is discrete if and only if every Γ-orbit is a discrete subset of X.

§5.4. Discrete Euclidean Groups

In this section, we characterize the discrete subgroups of the group $\mathrm{I}(E^n)$ of isometries of E^n.

Definition: An isometry ϕ of E^n is *elliptic* if and only if ϕ fixes a point of E^n; otherwise ϕ is *parabolic.*

Note that ϕ in $\mathrm{I}(E^n)$ is elliptic (resp. parabolic) if and only if its Poincaré extension $\tilde{\phi}$ in $\mathrm{M}(U^{n+1})$ is elliptic (resp. parabolic). Every element ϕ of $\mathrm{I}(E^n)$ is of the form $\phi(x) = a + Ax$ with a in E^n and A in $\mathrm{O}(n)$. We shall write simply $\phi = a + A$.

Theorem 5.4.1. *Let ϕ be in $\mathrm{I}(E^n)$. Then ϕ is parabolic if and only if there is a line L of E^n on which ϕ acts as a nontrivial translation.*

Proof: Suppose that $\phi = a + A$ is parabolic. Then ϕ has no fixed points in E^n by definition. Let V be the space of all vectors in E^n fixed by A, and let W be its orthogonal complement. Write $a = b + c$ with b in V and c in W. Let x be an arbitrary point of E^n and write $x = v + w$ with v in V and w in W. Now the orthogonal transformation A leaves the decomposition

$$E^n = V \oplus W$$

invariant. Hence $A - I$ maps W to itself. As V is the kernel of $A - I$ and $V \cap W = \{0\}$, we have that $A - I$ maps W isomorphically onto itself.

Next, observe that the fixed point equation

$$a + Ax = x$$

is equivalent to the equation

$$(b + c) + (v + Aw) = v + w,$$

which is equivalent to

$$(A - I)w = -b - c.$$

Consequently $b \neq 0$, otherwise we could solve the last equation for w and obtain a fixed point for ϕ. Choose y in E^n such that $(A - I)y = -c$. Let L be the line whose parametric form is $x = tb + y$, with t in $\mathbb{R}$. Then ϕ acts as a nontrivial translation on L, since

$$\begin{aligned}\phi(tb + y) &= a + A(tb + y)\\ &= a + tAb + Ay\\ &= a + tb + y - c\\ &= (t+1)b + y.\end{aligned}$$

Conversely, suppose there is a line L of E^n on which ϕ acts as a nontrivial translation. Then ϕ maps each hyperplane of E^n orthogonal to L to another hyperplane orthogonal to L. Consequently ϕ has no fixed points in E^n. Therefore ϕ is parabolic. □

Corollary 1. *If ϕ is a parabolic isometry of E^n, then there is a line L of E^n, an elliptic isometry ψ of E^n that fixes each point of L, and a nontrivial translation τ that leaves L invariant, such that $\phi = \tau\psi$.*

Proof: Let $\phi = a + A$ be parabolic. Write $a = b + c$ as in the proof of Theorem 5.4.1. Choose y such that $(A - I)y = -c$ and let L be the line

$$x = tb + y \quad \text{with } t \text{ in } \mathbb{R}.$$

Let $\psi = c + A$ and $\tau = b + I$. Then $\phi = \tau\psi$. Moreover, ψ fixes each point of L, and τ leaves L invariant. □

Corollary 2. *If ϕ is a parabolic isometry of E^n, then the subgroup Γ of $\mathrm{I}(E^n)$ generated by ϕ is discrete.*

Proof: By Theorem 5.4.1, there is a line L of E^n on which ϕ acts as a nontrivial translation. Let x be a point on L. Then the orbit Γx is a discrete set. As the map $e : \Gamma \to \Gamma x$, defined by $e(\phi^m) = \phi^m(x)$, is continuous, we have that $e^{-1}(x) = \{I\}$ is open in Γ, and so Γ is discrete. □

Remark: Let ϕ be an elliptic isometry of E^n. Then ϕ has a fixed point in E^n, and so ϕ is conjugate in $\mathrm{I}(E^n)$ to an element in $\mathrm{O}(n)$. Consequently, the subgroup generated by ϕ is discrete if and only if ϕ has finite order.

The next theorem is a basic result in linear algebra.

Theorem 5.4.2. *Let A be an orthogonal $n \times n$ matrix. Then there are angles $\theta_1, \ldots, \theta_m$, with $0 \leq \theta_1 \leq \cdots \leq \theta_m \leq \pi$, such that A is conjugate in* $\mathrm{O}(n)$ *to a block diagonal matrix of the form*

$$\begin{pmatrix} B(\theta_1) & & 0 \\ & \ddots & \\ 0 & & B(\theta_m) \end{pmatrix},$$

where $B(0) = 1$, $B(\pi) = -1$, *and* $B(\theta_j) = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix}$ *otherwise.*

The angles $\theta_1, \ldots, \theta_m$ in Theorem 5.4.2 are called the *angles of rotation* of A, and they completely determine the conjugacy class of A in $\mathrm{O}(n)$, since $e^{\pm i\theta_1}, \ldots, e^{\pm i\theta_m}$ are the eigenvalues of A, counting multiplicities. Furthermore, A is conjugate in $\mathrm{U}(n)$ to a diagonal matrix with diagonal entries $e^{\pm i\theta_1}, \ldots, e^{\pm i\theta_m}$. Note that A has finite order if and only if each angle of rotation of A is a rational multiple of π.

Commutivity in Discrete Euclidean Groups

If A and B are real $n \times n$ matrices and if x is a point of E^n, then

$$(1) \qquad |Ax| \leq |A|\,|x|, \tag{5.4.1}$$
$$(2) \qquad |AB| \leq |A|\,|B|, \tag{5.4.2}$$
$$(3) \qquad |A \pm B| \leq |A| + |B|; \tag{5.4.3}$$

if B is orthogonal, then

$$(4) \qquad |BA| = |A| = |AB|, \tag{5.4.4}$$
$$(5) \qquad |BAB^{-1} - I| = |A - I|. \tag{5.4.5}$$

Lemma 1. *If A is in* $\mathrm{O}(n)$ *and $|A - I| < 2$, then A is a rotation with all rotation angles less than $\pi/2$.*

Proof: By Formula 5.4.5, we may assume that A is in the block diagonal form of Theorem 5.4.2. Since $|A - I| < 2$, no rotation angle of A is equal to π, and so A is a rotation. Moreover, for each rotation angle $\theta > 0$, we have that

$$(\cos\theta - 1)^2 + \sin^2\theta + \sin^2\theta + (\cos\theta - 1)^2 < 4.$$

Hence, we have that

$$4 - 4\cos\theta < 4.$$

Therefore $\cos\theta > 0$. Thus $\theta < \pi/2$ for each rotation angle θ of A. □

Lemma 2. *Let A, B be in $\mathrm{GL}(n, \mathbb{C})$ with A conjugate to a diagonal matrix, and let $\mathbb{C}^n = V_1 \oplus \cdots \oplus V_m$ be the eigenspace decomposition of $\mathbb{C}^n$ relative to A. Then A and B commute if and only if $B(V_j) = V_j$ for each j.*

Proof: Let c_j be the eigenvalue associated to the eigenspace V_j for each j. Then $V_j = \ker(A - c_j I)$ by definition. Hence

$$\begin{aligned} B(V_j) &= \ker B(A - c_j I)B^{-1} \\ &= \ker(BAB^{-1} - c_j I). \end{aligned}$$

Therefore

$$\mathbb{C}^n = B(V_1) \oplus \cdots \oplus B(V_m)$$

is the eigenspace decomposition of $\mathbb{C}^n$ relative to BAB^{-1}.

Now suppose that A and B commute. Then $BAB^{-1} = A$ and therefore $B(V_j) = V_j$ for each j. Conversely, suppose that $B(V_j) = V_j$ for each j. Let v be an arbitrary vector in $\mathbb{C}^n$. Then we can write $v = v_1 + \cdots + v_m$ with v_j in V_j. Observe that

$$BAv_j = Bc_j v_j = c_j Bv_j$$

and

$$ABv_j = A(Bv_j) = c_j Bv_j.$$

But this implies that $BAv = ABv$, and so $BA = AB$. □

Lemma 3. *Let A, B be in $\mathrm{O}(n)$ with $|B - I| < 2$. If A commutes with BAB^{-1}, then A commutes with B.*

Proof: By Lemma 1, we have that B is a rotation with all angles less than $\pi/2$. Hence, all the eigenvalues of B have positive real parts. Let $\mathbb{C}^n = W_1 \oplus \cdots \oplus W_\ell$ be the eigenspace decomposition of $\mathbb{C}^n$ relative to B. Then the eigenspaces W_j are mutually orthogonal, since B is orthogonal. Let w be a nonzero vector in $\mathbb{C}^n$ and write $w = w_1 + \cdots + w_\ell$ with w_j in W_j. Let c_j be the eigenvalue of B corresponding to W_j. Then

$$\mathrm{Re}\left((Bw) * w\right) = \mathrm{Re}\left(\left(\sum c_j w_j\right) * \sum w_k\right) = \mathrm{Re} \sum c_j |w_j|^2 > 0.$$

Hence B cannot send any nonzero vector of $\mathbb{C}^n$ to an orthogonal vector.

Let $\mathbb{C}^n = V_1 \oplus \cdots \oplus V_m$ be the eigenspace decomposition of $\mathbb{C}^n$ relative to A. Then

$$\mathbb{C}^n = B(V_1) \oplus \cdots \oplus B(V_m)$$

is the eigenspace decomposition of $\mathbb{C}^n$ relative to BAB^{-1}. Now since BAB^{-1} and A commute, $A(B(V_j)) = B(V_j)$ for each j by Lemma 2. Consequently

$$B(V_j) = \bigoplus_k \left(B(V_j) \cap V_k\right)$$

is the eigenspace decomposition of $B(V_j)$ relative to A. Now, since B cannot send any nonzero vector of $\mathbb{C}^n$ to an orthogonal vector, we must have that $B(V_j) \cap V_k = \{0\}$ for $j \neq k$. Thus $B(V_j) = B(V_j) \cap V_j \subset V_j$. Hence $B(V_j) = V_j$ for all j, and so A commutes with B by Lemma 2. □

Lemma 4. *Let Γ be a discrete subgroup of* $\mathrm{I}(E^n)$ *and let $\phi = a + A$ and $\psi = b + B$ be in Γ. If $|A - I| < 1/2$ and $|B - I| < 2$, then A and B commute.*

Proof: On the contrary, suppose that $BA \neq AB$. Define a sequence $\{\psi_m\}$ in Γ by $\psi_0 = \psi$ and $\psi_{m+1} = \psi_m \phi \psi_m^{-1}$. Let $\psi_m = b_m + B_m$. Then we have

$$\begin{aligned}
\psi_{m+1} &= \psi_m \phi \psi_m^{-1} \\
&= \psi_m \phi(-B_m^{-1} b_m + B_m^{-1}) \\
&= \psi_m (a - AB_m^{-1} b_m + AB_m^{-1}) \\
&= b_m + B_m a - B_m A B_m^{-1} b_m + B_m A B_m^{-1}.
\end{aligned}$$

Hence $B_{m+1} = B_m A B_m^{-1}$. As $|B_0 - I| < 2$ and

$$|B_{m+1} - I| = |B_m A B_m^{-1} - I| = |A - I| < 1/2,$$

it follows by induction that $B_m A \neq AB_m$ for all m, since $B_0 A \neq AB_0$ and if $B_m A \neq AB_m$, then $(B_m A B_m^{-1})A \neq A(B_m A B_m^{-1})$ by Lemma 3. Hence $B_m \neq A$ for all m.

Next, observe that

$$\begin{aligned}
|A - B_{m+1}| &= |A - B_m A B_m^{-1}| \\
&= |AB_m - B_m A| \\
&= |(A - B_m)(A - I) - (A - I)(A - B_m)| \\
&\leq |(A - B_m)(A - I)| + |(A - I)(A - B_m)| \\
&\leq 2|A - I|\,|A - B_m| \\
&< |A - B_m|.
\end{aligned}$$

Thus B_{m+1} is nearer to A than B_m. Hence, the terms of the sequence $\{B_m\}$, and therefore of $\{\psi_m\}$, are distinct.

Next, observe that

$$b_{m+1} = (I - B_m A B_m^{-1}) b_m + B_m a$$

and so

$$|b_{m+1}| \leq \frac{1}{2}|b_m| + |a|.$$

Consequently $|b_m|$ is bounded by $2|a| + |b|$ for all m. Therefore, the sequence $\{b_m\}$ has a convergent subsequence $\{b_{m_j}\}$. Furthermore $\{B_{m_j}\}$ has a convergent subsequence, since $\mathrm{O}(n)$ is compact. Therefore $\{\psi_m\}$ has a subsequence that converges in $\mathrm{I}(E^n)$ by Theorem 5.2.4, and therefore in Γ, since Γ is closed in $\mathrm{I}(E^n)$. As the terms of $\{\psi_m\}$ are distinct, we have a contradiction to the discreteness of Γ by Lemma 2 of §5.3. □

Lemma 5. *Let Γ be a discrete subgroup of* $\mathrm{I}(E^n)$ *and let $\phi = a + A$ and $\psi = b + B$ be in Γ with $|A - I| < 1$ and $|B - I| < 1$. If A and B commute, then ϕ and ψ commute.*

Proof: Let $[\phi, \psi] = \phi\psi\phi^{-1}\psi^{-1}$. Then

$$\begin{aligned}
[\phi, \psi] &= \phi\psi\phi^{-1}(-B^{-1}b + B^{-1}) \\
&= \phi\psi(-A^{-1}a - A^{-1}B^{-1}b + A^{-1}B^{-1}) \\
&= \phi(b - BA^{-1}a - BA^{-1}B^{-1}b + BA^{-1}B^{-1}) \\
&= a + Ab - ABA^{-1}a - ABA^{-1}B^{-1}b + ABA^{-1}B^{-1} \\
&= (A - I)b + (I - B)a + I.
\end{aligned}$$

Now set

$$c = (A - I)b + (I - B)a.$$

Define a sequence $\{\phi_m\}$ in Γ by $\phi_1 = [\phi, [\phi, \psi]]$ and $\phi_m = [\phi, \phi_{m-1}]$. Then $\phi_1 = (A - I)c + I$, and in general $\phi_m = (A - I)^m c + I$. Now

$$|(A - I)^m c| \leq |A - I|^m |c|.$$

As $|A - I| < 1$, we have that $(A - I)^m c \to 0$ in E^n. Therefore $\phi_m \to I$ in Γ by Theorem 5.2.4. Hence, the sequence $\{\phi_m\}$ is eventually constant by Lemma 2 of §5.3. Therefore $(A - I)^m c = 0$ for some m.

Let V be the space of all vectors in E^n fixed by A and let W be its orthogonal complement. Write $c = v + w$ with v in V and w in W. Then

$$(A - I)^m c = (A - I)^m w.$$

As A is orthogonal, $A - I$ maps W isomorphically onto itself. Therefore $w = 0$. Hence c is fixed by A. The same argument, with the sequence $\{\psi_m\}$ defined by $\psi_1 = [\psi, [\phi, \psi]]$ and $\psi_m = [\psi, \psi_{m-1}]$, shows that c is also fixed by B.

Now observe that $(A - I)b$ is in W and so is orthogonal to c. Likewise $(I - B)a$ is orthogonal to c. As $c = (A - I)b + (I - B)a$, we have that c is orthogonal to itself, and so $c = 0$. Thus ϕ and ψ commute. □

Lemma 6. *If X is a compact metric space, then for each $r > 0$, there is a maximum number $k(r)$ of points of X with mutual distances at least r.*

Proof: On the contrary, suppose there is no upper bound to the number of points of X with mutual distances at least r. Since X is compact, it can be covered by finitely many balls of radius $r/2$, say $B(x_1, r/2), \ldots, B(x_m, r/2)$. Let $y_1, \ldots, y_{m+1}$ be $m + 1$ points of X with mutual distances at least r. Then some ball $B(x_i, r/2)$ contains two points y_j and y_k. But

$$d(y_j, y_k) \leq d(y_j, x_i) + d(x_i, y_k) < r/2 + r/2 = r,$$

which is a contradiction. □

Lemma 7. *Let Γ be a subgroup of $\mathrm{I}(E^n)$ and for each $r > 0$, let Γ_r be the subgroup of Γ generated by all elements $\phi = a + A$ in Γ, with $|A - I| < r$, and let $k_n(r)$ be the maximum number of elements of $\mathrm{O}(n)$ with mutual distances at least r relative to the metric $d(A, B) = |A - B|$. Then Γ_r is a normal subgroup of Γ and $[\Gamma : \Gamma_r] \leq k_n(r)$ for each $r > 0$.*

Proof: Let $\phi = a + A$ be in Γ_r, with $|A - I| < r$, and let $\psi = b + B$ be in Γ. Then $\psi\phi\psi^{-1} = c + BAB^{-1}$ for some c in E^n. Hence

$$|BAB^{-1} - I| = |A - I| < r.$$

Thus $\psi\phi\psi^{-1}$ is in Γ_r. Consequently Γ_r is a normal subgroup of Γ.

Let $\psi_i = b_i + B_i$, for $i = 1, \ldots, m$, be a maximal number of elements of Γ such that the mutual distances between $B_1, \ldots, B_m$ are at least r. Then $m \leq k_n(r)$. Let $\psi = b + B$ be an arbitrary element of Γ. Then there is an index j such that $|B - B_j| < r$; otherwise $\psi, \psi_1, \ldots, \psi_m$ would be $m + 1$ elements of Γ such that the mutual distances between $B, B_1, \ldots, B_m$ are at least r. Hence $|BB_j^{-1} - I| < r$. As $\psi\psi_j^{-1} = c + BB_j^{-1}$ for some c in E^n, we have that $\psi\psi_j^{-1}$ is in Γ_r. Therefore ψ is in the coset $\Gamma_r\psi_j$. Hence

$$\Gamma = \Gamma_r\psi_1 \cup \cdots \cup \Gamma_r\psi_m.$$

Thus $[\Gamma : \Gamma_r] \leq m \leq k_n(r)$. □

Theorem 5.4.3. *Let* Γ *be a discrete subgroup of* $\mathrm{I}(E^n)$. *Then* Γ *has an abelian normal subgroup* N *of finite index containing all the translations in* Γ *and the index of* N *in* Γ *is bounded by a number depending only on* n.

Proof: Let $\mathrm{N} = \Gamma_{\frac{1}{2}}$. Then we have that N is a normal subgroup of Γ with $[\Gamma : \mathrm{N}] \leq k_n(1/2)$ by Lemma 7; moreover, N is abelian by Lemmas 4 and 5. Clearly N contains every translation in Γ. □

Example: Let Γ be the group of symmetries of $\mathbb{Z}^n$ in E^n. Then $\Gamma 0 = \mathbb{Z}^n$; moreover, the stabilizer Γ_0 is the subgroup of $\mathrm{O}(n)$ of all matrices with integral entries. Clearly Γ_0 is a finite group. Now the map $e : \Gamma \to \mathbb{Z}^n$, defined by $e(\phi) = \phi(0)$, is continuous. Hence $e^{-1}(0) = \Gamma_0$ is open in Γ. As Γ_0 is finite, we have that $\{I\}$ is open in Γ. Therefore Γ is discrete.

If $\phi = a + A$ is in Γ, then obviously A is in Γ_0. Hence, the mapping $a + A \mapsto A$ determines a short exact sequence

$$1 \to \mathrm{T} \to \Gamma \to \Gamma_0 \to 1,$$

where T is the translation subgroup of Γ. The sequence splits, since Γ_0 is a subgroup of Γ. Therefore $\Gamma = \mathrm{T}\Gamma_0$ is a semi-direct product. In particular, the index of T in Γ is the order of Γ_0.

Definition: Let G be a group acting on a set X.

(1) An element g of G *acts trivially* on X if and only if $gx = x$ for all x in X.

(2) The group G *acts trivially* on X if and only if every element of G acts trivially on X.

(3) The group G *acts effectively* on G if and only if 1 is the only element of G acting trivially on X.

Theorem 5.4.4. *Let* Γ *be an abelian discrete subgroup of* $\mathrm{I}(E^n)$. *Then there are subgroups* H *and* K *of* Γ *and an* m*-plane* P *of* E^n *such that*

(1) *the group* Γ *has the direct sum decomposition* $\Gamma = \mathrm{K} \oplus \mathrm{H}$*;*

(2) *the group* K *is finite and acts trivially on* P*; and*

(3) *the group* H *is free abelian of rank* m *and acts effectively on* P *as a discrete group of translations.*

Proof: The proof is by induction on the dimension n. The theorem is trivial when $n = 0$. Assume that $n > 0$ and the theorem is true for all dimensions less than n. Choose $\phi = a + A$ in Γ such that the dimension of the space V of all vectors in E^n fixed by A is as small as possible. If $V = E^n$, then Γ is a group of translations and the theorem holds for Γ by Theorem 5.3.2 with $\mathrm{H} = \Gamma$ and P the vector space spanned by the orbit $\Gamma 0$.

Now assume that $\dim V < n$. Let W be the orthogonal complement of V in E^n. Write $a = v + w$ with v in V and w in W. Since the image of $A - I$ is W, there is a y such that $(A - I)y = w$. Let $\tau = y + I$. Then

$$\begin{aligned} \tau\phi\tau^{-1} &= \tau\phi(-y + I) \\ &= \tau(a - Ay + A) \\ &= y + a - Ay + A \\ &= a - w + A \\ &= v + A. \end{aligned}$$

Consequently, by conjugating the group Γ by τ, we may assume that A fixes a.

Let $\psi = b + B$ be in Γ. From the proof of Lemma 5, we have

$$[\phi, \psi] = (A - I)b + (I - B)a + I.$$

Hence $(A - I)b + (I - B)a = 0$. As A and B commute, $B(V) = V$ and so $(B - I)(V) \subset V$. From the equation

$$(B - I)a = (A - I)b,$$

we deduce that $(B - I)a$ is in $V \cap W = \{0\}$. Hence B fixes a and A fixes b. Thus b is in V. Consequently ψ, and therefore Γ, leaves V invariant.

By conjugating the group Γ by an appropriate rotation, we may assume that $V = E^k$ with $k < n$. Let $\overline{\Gamma}$ be the subgroup of $\mathrm{I}(E^k)$ obtained by restricting the isometries in Γ, and let $\rho : \Gamma \to \overline{\Gamma}$ be the restriction homomorphism. The kernel of ρ is a discrete subgroup of $\mathrm{O}(n)$ and is therefore finite by Theorem 5.3.1. As Γ acts discontinuously on E^k, the group $\overline{\Gamma}$ does also and is therefore discrete.

By the induction hypothesis, there are subgroups $\overline{\mathrm{H}}$ and $\overline{\mathrm{K}}$ of $\overline{\Gamma}$, and an m-plane P of E^k such that (1) $\overline{\Gamma} = \overline{\mathrm{K}} \oplus \overline{\mathrm{H}}$, (2) $\overline{\mathrm{K}}$ is finite and acts trivially on P, and (3) $\overline{\mathrm{H}}$ is free abelian of rank m and acts effectively on

P as a discrete group of translations. Let $\mathrm{K} = \rho^{-1}(\overline{\mathrm{K}})$. Then K is a finite subgroup of Γ, and K acts trivially on P. Moreover, there is an exact sequence

$$1 \to \mathrm{K} \to \Gamma \to \overline{\mathrm{H}} \to 1.$$

The sequence splits, since $\overline{\mathrm{H}}$ is free abelian. Hence, there is a subgroup H of Γ such that $\Gamma = \mathrm{K} \oplus \mathrm{H}$ and ρ maps H isomorphically onto $\overline{\mathrm{H}}$. Therefore H is free abelian of rank m and H acts effectively on P as a discrete group of translations. This completes the induction. □

Definition: A *lattice subgroup* Γ of $\mathrm{I}(E^n)$ is a group Γ generated by n linearly independent translations.

Corollary 3. *A subgroup Γ of $\mathrm{I}(E^n)$ is a lattice subgroup if and only if Γ is discrete and free abelian of rank n.*

Lemma 8. *Let H be a subgroup of finite index in a topological group Γ with a metric topology. If H is discrete, then Γ is discrete.*

Proof: Suppose that H is discrete. Then H is closed in Γ by Lemma 3 of §5.3. Since H is of finite index in Γ, there are elements $g_1, \ldots, g_m$ in Γ, with $g_1 = 1$, such that

$$\Gamma = g_1\mathrm{H} \cup \cdots \cup g_m\mathrm{H}.$$

Hence, we have

$$\mathrm{H} = \Gamma - g_2\mathrm{H} \cup \cdots \cup g_m\mathrm{H}.$$

As each coset $g_i\mathrm{H}$ is closed in Γ, we have that H is open in Γ. As $\{1\}$ is open in H, we have that $\{1\}$ is open in Γ. Thus Γ is discrete. □

The next theorem follows immediately from Theorems 5.4.3 and 5.4.4 and Lemma 8.

Theorem 5.4.5. *Let Γ be a subgroup of $\mathrm{I}(E^n)$. Then Γ is discrete if and only if Γ has a free abelian subgroup H of rank m and of finite index such that H acts effectively on an m-plane P of E^n as a discrete group of translations.*

We shall prove that the m-plane P in Theorem 5.4.5 can be chosen so that P is invariant under Γ. The next lemma takes care of the case $m = 0$.

Lemma 9. *If Γ is a finite subgroup of $\mathrm{I}(E^n)$, then Γ fixes a point of E^n.*

Proof: Let $m = |\Gamma|$ and set

$$a = \frac{1}{m} \sum_{\phi \in \Gamma} \phi(0).$$

Then for $\psi = b + B$ in Γ, we have

$$\begin{aligned}\psi(a) &= b + \frac{1}{m}\sum_{\phi\in\Gamma} B\phi(0)\\ &= \frac{1}{m}\sum_{\phi\in\Gamma} b + B\phi(0)\\ &= \frac{1}{m}\sum_{\phi\in\Gamma}\psi\phi(0)\\ &= \frac{1}{m}\sum_{\phi\in\Gamma}\phi(0) \;=\; a.\end{aligned}$$

□

Theorem 5.4.6. *Let Γ be a discrete subgroup of $\mathrm{I}(E^n)$. Then*

(1) *the group Γ has a free abelian subgroup H of rank m and finite index;*

(2) *there is an m-plane P of E^n such that H acts effectively on P as a discrete group of translations; and*

(3) *the m-plane P is invariant under Γ.*

Proof: By Theorem 5.4.3, the group Γ has an abelian normal subgroup N of finite index. By Theorem 5.4.4, the group N has a free abelian subgroup H of rank m and of finite index, there is an m-plane Q of E^n such that H acts effectively on Q as a discrete group of translations, and N acts on Q via translations. By conjugating Γ in $\mathrm{I}(E^n)$, we may assume that $Q = E^m$.

Let $\phi = a + A$ be an arbitrary element of N. As $\phi(0) = a$, we find that a is in E^m and ϕ acts on E^m by translation by a. Hence A fixes each point of E^m. Let V_ϕ be the subspace of E^n of elements fixed by A and set

$$V = \bigcap_{\phi\in\mathrm{N}} V_\phi.$$

Then $E^m \subset V$.

Let $\psi = b + B$ be an arbitrary element of Γ. We now show that ψ leaves V invariant. First of all, we have

$$\begin{aligned}B(V) &= B\left(\bigcap_{\phi\in\mathrm{N}} V_\phi\right)\\ &= \bigcap_{\phi\in\mathrm{N}} BV_\phi\\ &= \bigcap_{\phi\in\mathrm{N}} V_{\psi\phi\psi^{-1}}\\ &= \bigcap_{\phi\in\mathrm{N}} V_\phi \;=\; V.\end{aligned}$$

Thus B leaves V invariant. Let $\phi = a + A$ be in N. Then

$$\psi\phi\psi^{-1} = (I - BAB^{-1})b + Ba + BAB^{-1}.$$

As $\psi\phi\psi^{-1}$ is in N, there is a v in E^m such that

$$(I - BAB^{-1})b + Ba = v.$$

Let $W_{\psi\phi\psi^{-1}}$ be the orthogonal complement of $V_{\psi\phi\psi^{-1}}$. Write $b = c + d$ with c in $V_{\psi\phi\psi^{-1}}$ and d in $W_{\psi\phi\psi^{-1}}$. Then we have

$$(I - BAB^{-1})d + Ba = v.$$

Now observe that

$$Ba = v + (BAB^{-1} - I)d$$

is the orthogonal decomposition of Ba with respect to $V_{\psi\phi\psi^{-1}}$ and $W_{\psi\phi\psi^{-1}}$. As Ba is in V, we have that $(BAB^{-1} - I)d = 0$, and so $d = 0$. Therefore b is in $V_{\psi\phi\psi^{-1}}$ for each ϕ in N. Hence b is in V. Thus ψ leaves V invariant. Furthermore Ba is in E^m for each a in E^m. Hence B leaves E^m invariant.

Now by conjugating Γ by an appropriate rotation of E^n that leaves E^m fixed, we may assume that $V = E^\ell$ with $\ell \geq m$. Let $\eta : E^\ell \to E^{\ell-m}$ be the projection defined by

$$\eta(x_1, \ldots, x_\ell) = (x_{\ell-m+1}, \ldots, x_\ell).$$

Define $\sigma : E^{\ell-m} \to E^\ell$ by

$$\sigma(x_1, \ldots, x_{\ell-m}) = (0, \ldots, 0, x_1, \ldots, x_{\ell-m}).$$

Then σ is a right inverse for η. By Theorem 5.1.5, we have that η induces an isomorphism of topological groups

$$\overline{\eta} : E^\ell/E^m \to E^{\ell-m}.$$

Define a metric on E^ℓ/E^m by

$$d(x + E^m, y + E^m) = |\eta(x) - \eta(y)|.$$

Then $\overline{\eta}$ is an isometry.

We now define an action of Γ/N on E^ℓ/E^m by

$$(\mathrm{N}\psi)(x + E^m) = \psi(x) + E^m = b + Bx + E^m.$$

This action is well defined, since N acts on E^ℓ by translation by elements of E^m and B leaves E^m invariant. Moreover Γ/N acts on E^ℓ/E^m via isometries. By Lemma 9, the finite group Γ/N fixes a point $P = x + E^m$ of E^ℓ/E^m. Hence Γ leaves the m-plane P invariant, and H acts effectively on P as a discrete group of translations. □

Exercise 5.4

1. Let $\mathrm{I}_0(\mathbb{C})$ be the group of orientation preserving Euclidean isometries of $\mathbb{C}$. Show that every element of $\mathrm{I}_0(\mathbb{C})$ is of the form $\phi(z) = az + b$ with a in S^1 and b in $\mathbb{C}$.
2. Determine all the discrete subgroups of $\mathrm{I}_0(\mathbb{C})$.
3. Let A be a real $n \times n$ matrix. Prove that $|A|^2 = \mathrm{tr}(AA^t)$.
4. Prove Formula 5.4.4.

5. Prove Formula 5.4.5.
6. Let ϕ be a parabolic isometry of E^n and let L be a line of E^n on which ϕ acts as a translation. Show that the unit vector u pointing in the direction in which ϕ translates L is uniquely determined by ϕ. The vector u is called the *translation direction vector* of ϕ.
7. Let Γ be a discrete subgroup of $\mathrm{I}(E^n)$. Prove that the subgroup T of translations of Γ has finite index in Γ if and only if every isometry $\phi = a + A$ in Γ has the property that its $\mathrm{O}(n)$-component A has finite order.
8. Find an upper bound for $k_n(1/2)$.
9. Prove that the order of the group Γ_0, in the example after Theorem 5.4.3, is $2^n n!$.
10. Let Γ be a discrete subgroup of $\mathrm{I}(E^n)$ and let m be as in Theorem 5.4.6. Prove that any two Γ-invariant m-planes of E^n are parallel.

§5.5. Elementary Groups

In this section, we shall characterize the elementary discrete subgroups of $\mathrm{M}(B^n)$.

Definition: A subgroup G of $\mathrm{M}(B^n)$ is *elementary* if and only if G has a finite orbit in the closed ball $\overline{B}^n$.

We shall divide the elementary subgroups of $\mathrm{M}(B^n)$ into three types. Let G be an elementary subgroup of $\mathrm{M}(B^n)$.

(1) The group G is said to be of *elliptic type* if and only if G has a finite orbit in B^n.

(2) The group G is said to be of *parabolic type* if and only if G fixes a point of S^{n-1} and has no other finite orbits in $\overline{B}^n$.

(3) The group G is said to be of *hyperbolic type* if and only if G is neither of elliptic type nor of parabolic type.

Let ϕ be in $\mathrm{M}(B^n)$ and let x be a point of $\overline{B}^n$. Then
$$(\phi G \phi^{-1})\phi(x) = \phi(Gx).$$
In other words, the $\phi G\phi^{-1}$-orbit through $\phi(x)$ is the ϕ-image of the G-orbit through x. This implies that $\phi G\phi^{-1}$ is also elementary; moreover, G and $\phi G\phi^{-1}$ have the same type. Thus, the elementary type of G depends only on the conjugacy class of G.

Elementary Groups of Elliptic Type

Theorem 5.5.1. *Let G be an elementary subgroup of* $\mathrm{M}(B^n)$. *Then the following are equivalent:*

(1) *The group G is of elliptic type.*

(2) *The group G fixes a point of B^n.*

(3) *The group G is conjugate in* $\mathrm{M}(B^n)$ *to subgroup of* $\mathrm{O}(n)$.

Proof: Suppose that G is of elliptic type. We pass to the hyperboloid model H^n of hyperbolic space and regard G as a subgroup of $\mathrm{PO}(n,1)$. As G is of elliptic type, it has a finite orbit $\{v_1,\ldots,v_m\}$ in H^n. Let $v = v_1 + \cdots + v_m$. Then v is a positive time-like vector of $\mathbb{R}^{n,1}$ by Theorem 3.1.1. Now let $v_0 = v/|||v|||$. Then v_0 is in H^n. If A is in G, then A permutes the elements of $\{v_1,\ldots,v_m\}$ by left multiplication. Therefore, we have

$$\begin{aligned} Av_0 &= \frac{Av}{|||v|||} \\ &= \frac{Av_1 + \cdots + Av_m}{|||v|||} \\ &= \frac{v_1 + \cdots + v_m}{|||v|||} = v_0. \end{aligned}$$

Thus G fixes v_0. Hence (1) implies (2).

Suppose that G fixes a point b of B^n. Let ϕ be a Möbius transformation of B^n such that $\phi(0) = b$. Then $\phi^{-1}G\phi$ fixes 0. Consequently $\phi^{-1}G\phi$ is a subgroup of $\mathrm{O}(n)$ by Theorem 4.4.8. Thus (2) implies (3).

Suppose there is a ϕ in $\mathrm{M}(B^n)$ such that $\phi^{-1}G\phi$ is a subgroup of $\mathrm{O}(n)$. Then G fixes $\phi(0)$, and so (3) implies (1). □

The next theorem follows immediately from Theorems 5.3.1 and 5.5.1.

Theorem 5.5.2. *Let* Γ *be a subgroup of* $\mathrm{M}(B^n)$. *Then the following are equivalent:*

(1) *The group* Γ *is finite.*

(2) *The group* Γ *is conjugate in* $\mathrm{M}(B^n)$ *to a finite subgroup of* $\mathrm{O}(n)$.

(3) *The group* Γ *is an elementary discrete subgroup of elliptic type.*

Elementary Groups of Parabolic Type

In order to analyze elementary groups of parabolic and hyperbolic type, it will be more convenient to work in the upper half-space model U^n of hyperbolic space. Elementary subgroups of $\mathrm{M}(U^n)$ of elliptic, parabolic, and hyperbolic type are defined in the same manner as in the conformal ball model B^n. The main advantage of working in $\mathrm{M}(U^n)$ is that the group of Euclidean similarities $\mathrm{S}(E^{n-1})$ is isomorphic by Poincaré extension to the stabilizer of ∞ in $\mathrm{M}(U^n)$. Consequently, we may identify $\mathrm{S}(E^{n-1})$ with the stabilizer of ∞ in $\mathrm{M}(U^n)$.

Theorem 5.5.3. *Let G be an elementary subgroup of $\mathrm{M}(U^n)$. Then the following are equivalent:*

(1) *The group G is of parabolic type.*

(2) *The group G has a unique fixed point in $\hat{E}^{n-1}$.*

(3) *The group G is conjugate in $\mathrm{M}(U^n)$ to a subgroup of $\mathrm{S}(E^{n-1})$ that fixes no point of E^{n-1}.*

Proof: Obviously (1) implies (2), and (2) and (3) are equivalent. We shall prove that (2) implies (1) by contradiction. Suppose that G fixes a unique point a of $\hat{E}^{n-1}$ and G is not of parabolic type. Then G has a finite orbit $\{u_1, \ldots, u_m\}$ in $\overline{U}^n$ other than $\{a\}$. Assume first that $\{u_1, \ldots, u_m\}$ is in U^n. Then G is of elliptic type, and so it fixes a point u of U^n by Theorem 5.5.1. Consequently G fixes the hyperbolic line L starting at a and passing through u. But this implies that G fixes the other endpoint of L contrary to the uniqueness of a. Therefore $\{u_1, \ldots, u_m\}$ must be contained in $\hat{E}^{n-1}$.

As a is the only fixed point of G in $\hat{E}^{n-1}$, we must have $m \geq 2$. The index of each stabilizer G_{u_i} is m. Therefore $H = G_{u_1} \cap G_{u_2}$ is of finite index in G. Moreover, each element of H is elliptic, since H fixes the three points a, u_1, u_2. Therefore H fixes the hyperbolic line L joining a and u_1. Let u be any point on L. As G_u contains H, we have that G_u is of finite index in G. Consequently, the orbit Gu is finite. But we have already shown that this leads to a contradiction. Therefore G must be of parabolic type. Thus (2) implies (1). □

Theorem 5.5.4. *Let ϕ, ψ be in $\mathrm{M}(U^n)$ with ψ hyperbolic. If ϕ and ψ have exactly one fixed point in common, then the subgroup generated by ϕ and ψ is not discrete.*

Proof: By conjugating in $\mathrm{M}(U^n)$, we may assume that the common fixed point is ∞. Thus, we may regard ϕ and ψ to be in $\mathrm{S}(E^{n-1})$. By conjugating in $\mathrm{S}(E^{n-1})$, we may assume that ψ fixes 0. Then there are positive scalars r, s, matrices A, B in $\mathrm{O}(n-1)$, and a nonzero point a of E^{n-1} such that $\phi(x) = a + rAx$ and $\psi(x) = sBx$. By replacing ψ with ψ^{-1}, if necessary, we may also assume that $0 < s < 1$. Then we have

$$\psi^m\phi\psi^{-m}(x) = s^mB^ma + rB^mAB^{-m}x$$

for each positive integer m. The terms of the sequence $\{\psi^m\phi\psi^{-m}\}$ are all distinct, since

$$\psi^m\phi\psi^{-m}(0) = s^mB^ma \quad \text{with } a \neq 0.$$

As $\mathrm{O}(n-1)$ is compact, the sequence $\{B^mAB^{-m}\}$ has a convergent subsequence $\{B^{m_j}AB^{-m_j}\}$. Let τ_m be the translation of E^{n-1} by s^mB^ma. Then $\{\tau_m\}$ converges to I by Corollary 1 of §5.2. As

$$\psi^m\phi\psi^{-m} = \tau_m rB^mAB^{-m},$$

the sequence $\{\psi^{m_j}\phi\psi^{-m_j}\}$ converges but is not eventually constant. Therefore, the group $\langle\phi, \psi\rangle$ is not discrete by Lemma 2 of §5.3. □

Theorem 5.5.5. *A subgroup Γ of $\mathrm{M}(U^n)$ is an elementary discrete subgroup of parabolic type if and only if Γ is conjugate in $\mathrm{M}(U^n)$ to an infinite discrete subgroup of $\mathrm{I}(E^{n-1})$.*

Proof: Suppose that Γ is an elementary discrete subgroup of parabolic type. By Theorem 5.5.3, we may assume that Γ is a subgroup of $\mathrm{S}(E^{n-1})$ that fixes no point of E^{n-1}. By Theorem 5.5.4, the group Γ has no hyperbolic elements, otherwise Γ would fix a point of E^{n-1}. Consequently, every element ϕ of Γ is of the form $\phi(x) = a + Ax$, where A is in $\mathrm{O}(n-1)$ and a is in E^{n-1}. Thus Γ is a subgroup of $\mathrm{I}(E^{n-1})$. The group Γ must be infinite, otherwise Γ would be of elliptic type.

Conversely, suppose that Γ is an infinite discrete subgroup of $\mathrm{I}(E^{n-1})$. On the contrary, assume that Γ fixes a point of E^{n-1}. By conjugating in $\mathrm{I}(E^{n-1})$, we may assume that Γ fixes 0. Then Γ is a subgroup of $\mathrm{O}(n-1)$. But Γ is discrete, and so Γ must be finite, which is not the case. Therefore Γ fixes no point of E^{n-1}. Hence Γ is of parabolic type by Theorem 5.5.3. □

Elementary Groups of Hyperbolic Type

Let $\mathrm{S}(E^{n-1})_*$ be the subgroup of $\mathrm{M}(E^{n-1})$ of all transformations that leave invariant the set $\{0, \infty\}$. The group $\mathrm{S}(E^{n-1})_*$ contains the subgroup $\mathrm{S}(E^{n-1})_0$ of all similarities that fix both 0 and ∞ as a subgroup of index two. We shall identify $\mathrm{S}(E^{n-1})_*$ with the subgroup of $\mathrm{M}(U^n)$ of all transformations that leave $\{0, \infty\}$ invariant.

Theorem 5.5.6. *Let G be an elementary subgroup of $\mathrm{M}(U^n)$. Then the following are equivalent:*

(1) *The group G is of hyperbolic type.*

(2) *The union of all the finite orbits of G in $\overline{U}^n$ consists of two points in $\hat{E}^{n-1}$.*

(3) *The group G is conjugate in $\mathrm{M}(U^n)$ to a subgroup of $\mathrm{S}(E^{n-1})_*$ that fixes no point of the positive nth axis.*

Proof: Suppose that G is of hyperbolic type. Then all the finite orbits of G are contained in $\hat{E}^{n-1}$, since G is not of elliptic type. Let $\{u_1, \ldots, u_m\}$ be the union of a finite number of finite G-orbits. Then each of the stabilizers G_{u_i} is of finite index in G, since each of the orbits Gu_i is finite. Let

$$H = G_{u_1} \cap \cdots \cap G_{u_m}.$$

Then H is of finite index in G and fixes each u_i. If $m \geq 3$, the group H must be of elliptic type; but this implies that G is of elliptic type, which is not the case. Therefore m can be at most 2. The case of one finite orbit, consisting of a single point, is ruled out by Theorem 5.5.3. Therefore, either

G has one finite orbit consisting of two points or two finite orbits consisting of one point each. Thus (1) implies (2).

Obviously (2) implies (3). Suppose that G is a subgroup of $\mathrm{S}(E^{n-1})_*$ that fixes no point of the positive nth axis. Then either G fixes both 0 and ∞ or $\{0, \infty\}$ is a G-orbit. Consequently G is not of parabolic type.

On the contrary, assume that G is of elliptic type. If G fixes both 0 and ∞, then G fixes the positive nth axis, which is not the case. Therefore $\{0, \infty\}$ is a G-orbit. The stabilizer G_0 is of index two in G and fixes both 0 and ∞. Hence G_0 fixes the positive nth axis L. Let ϕ be in $G - G_0$. Then ϕ leaves L invariant and switches its endpoints. Consequently ϕ has a fixed point u on L. As G_0 and ϕ generate G, the group G fixes u, which is a contradiction. It follows that G is of hyperbolic type. Thus (3) implies (1). □

The next theorem follows immediately from Theorems 5.5.2 and 5.5.6.

Theorem 5.5.7. *A subgroup Γ of $\mathrm{M}(U^n)$ is an elementary discrete subgroup of hyperbolic type if and only if Γ is conjugate in $\mathrm{M}(U^n)$ to an infinite discrete subgroup of $\mathrm{S}(E^{n-1})_*$.*

The structure of an infinite discrete subgroup Γ of $\mathrm{S}(E^{n-1})_*$ is easy to describe. Let Γ_0 be the subgroup of Γ fixing 0. Then Γ_0 is of index 1 or 2 in Γ. Every element of Γ_0 is of the form kA, where k is a positive scalar and A is in $\mathrm{O}(n-1)$. Let

$$\rho : \Gamma_0 \to \mathbb{R}_+$$

be the homomorphism defined by $\rho(kA) = k$. The kernel of ρ is the group $\Gamma_0 \cap \mathrm{O}(n-1)$, which is finite. As the orbit $\Gamma_0 e_n$ is discrete, we find that the image of ρ is an infinite discrete subgroup of $\mathbb{R}_+$. Hence, there is a scalar $s > 1$ such that

$$\rho(\Gamma_0) = \{s^m : m \in \mathbb{Z}\}.$$

Thus Γ_0 is finite by infinite cyclic.

Let ψ be an element of Γ_0 such that $\rho(\psi) = s$. Then Γ_0 is the semidirect product of the finite subgroup $\Gamma_0 \cap \mathrm{O}(n-1)$ and the infinite cyclic subgroup generated by ψ. Consequently Γ has an infinite cyclic subgroup generated by a hyperbolic transformation as a subgroup of finite index. This leads to the next theorem.

Theorem 5.5.8. *A subgroup Γ of $\mathrm{M}(U^n)$ is an elementary discrete subgroup of hyperbolic type if and only if Γ contains an infinite cyclic subgroup generated by a hyperbolic transformation as a subgroup of finite index.*

Proof: Suppose that Γ has an infinite cyclic subgroup H generated by a hyperbolic transformation ψ as a subgroup of finite index. Let a and b be the fixed points of ψ. As Γ_a contains H, we have that Γ_a is of finite index in Γ. Therefore, the orbit Γa is finite. Likewise Γb is finite. Hence Γ is elementary. As H has no fixed points in U^n, the type of Γ is not elliptic by Theorem 5.5.1. Moreover Γ is not of parabolic type, since the union of all the finite orbits of Γ contains at least a and b. Therefore Γ must be of hyperbolic type. Let L be the axis of ψ and let x be a point on L. Then the orbit $\mathrm{H}x$ is a discrete set. As the map $e : \mathrm{H} \to \mathrm{H}x$, defined by $e(\psi^m) = \psi^m(x)$, is continuous, $e^{-1}(x) = \{I\}$ is open in H, and so H is discrete. Consequently Γ is discrete by Lemma 8 of §5.4. The converse follows from Theorem 5.5.7 and the discussion thereafter. □

Example: Let μ be the magnification of U^n defined by $\mu(x) = 2x$, and let σ be the inversion of U^n defined by $\sigma(x) = x/|x|^2$. Let Γ be the group generated by μ and σ. As $\sigma\mu\sigma = \mu^{-1}$, the infinite cyclic group $\langle\mu\rangle$ has index two in Γ. Therefore Γ is an elementary discrete subgroup of $\mathrm{M}(U^n)$ of hyperbolic type by Theorem 5.5.8. Observe that Γ leaves the set $\{0, \infty\}$ invariant but fixes neither 0 nor ∞.

Solvable Groups

Let F_ϕ be the set of all fixed points in $\overline{B}^n$ of a Möbius transformation ϕ of B^n. If ϕ, ψ are in $\mathrm{M}(B^n)$, then obviously

$$F_{\psi\phi\psi-1} = \psi(F_\phi).$$

This simple observation is the key to the proof of the next lemma.

Lemma 1. *Every abelian subgroup of* $\mathrm{M}(B^n)$ *is elementary.*

Proof: The proof is by induction on n. The theorem is trivial when $n = 0$, since $B^0 = \{0\}$ by definition. Now suppose that $n > 0$ and the theorem is true for all dimensions less than n. Let G be an abelian subgroup of $\mathrm{M}(B^n)$. Assume first that G has an element ϕ that is either parabolic or hyperbolic. Then F_ϕ consists of one or two points. As $\psi\phi\psi^{-1} = \phi$ for all ψ in G, we have that $\psi(F_\phi) = F_\phi$ for all ψ in G, and so G is elementary.

Now assume that all the elements of G are elliptic. Let ϕ be in G. Then F_ϕ is the closure in $\overline{B}^n$ of a hyperbolic m-plane of B^n, since ϕ is conjugate in $\mathrm{M}(B^n)$ to an element of $\mathrm{O}(n)$. Therefore F_ϕ is a closed m-disk. Choose ϕ in G such that the dimension of F_ϕ is as small as possible. If $\dim F_\phi = n$, then G is trivial, so assume that $\dim F_\phi < n$. By conjugating G in $\mathrm{M}(B^n)$, we may assume that $F_\phi = \overline{B}^m$ with $m < n$. As G is abelian, we have that $\psi(F_\phi) = F_\phi$ for all ψ in G; in other words, G leaves $\overline{B}^m$ invariant. Moreover G leaves $\hat{E}^m$ invariant by Theorem 4.3.7.

Let $\overline{G}$ be the group of transformations of $\hat{E}^m$ obtained by restricting the elements of G. Then $\overline{G}$ is a subgroup of $\mathrm{M}(B^m)$ by Theorem 4.3.1. Moreover $\overline{G}$ is abelian, since $\overline{G}$ is a homomorphic image of G. By the induction hypothesis, $\overline{G}$, and therefore G, has a finite orbit in $\overline{B}^m$. Thus G is elementary. This completes the induction. □

Theorem 5.5.9. *Let* Γ *be a discrete subgroup of* $\mathrm{M}(B^n)$. *Then* Γ *is elementary if and only if it has a abelian subgroup of finite index. Moreover, if* Γ *is elementary, then it has a free abelian subgroup of finite index whose rank is* 0 *if* Γ *is elliptic,* 1 *if* Γ *is hyperbolic, or* k, *with* $0 < k < n$, *if* Γ *is parabolic.*

Proof: If Γ is elementary, then it has a free abelian subgroup of finite index by Theorems 5.4.5, 5.5.2, 5.5.5, and 5.5.8 whose rank is 0 if Γ is elliptic, 1 if Γ is hyperbolic, or k, with $0 < k < n$, if Γ is parabolic.

Conversely, suppose that Γ has an abelian subgroup H of finite index. Then H is elementary by Lemma 1. Let x be a point in $\overline{B}^n$ such that Hx is finite. As $[\Gamma : \mathrm{H}]$ is finite, there are elements $\phi_1, \ldots, \phi_m$ in Γ such that

$$\Gamma = \phi_1\mathrm{H} \cup \cdots \cup \phi_m\mathrm{H}.$$

Hence, we have that

$$\Gamma x = \phi_1\mathrm{H}x \cup \cdots \cup \phi_m\mathrm{H}x$$

is finite. Therefore Γ is elementary. □

Theorem 5.5.10. *Every solvable subgroup of* $M(B^n)$ *is elementary.*

Proof: Let G be a solvable subgroup of $M(B^n)$. Define $G^{(0)} = G$ and $G^{(k)} = [G^{(k-1)}, G^{(k-1)}]$ for $k > 0$. Then $G^{(k)} = 1$ for some smallest k. We prove that G is elementary by induction on the solvability degree k. This is clear if $k = 0$, so assume that $k > 0$ and all subgroups of $M(B^n)$ of solvability degree $k - 1$ are elementary. As the solvability degree of $H = G^{(1)}$ is $k - 1$, we have that H is elementary.

Assume first that H is of parabolic or hyperbolic type. Then the union of the finite orbits of H in S^{n-1} is a one or two point set F. Let h be in H and g in G. Then $g^{-1}hg$ is in H, since H is a normal subgroup of G. Hence $g^{-1}hg(F) = F$. Therefore $hg(F) = g(F)$. Hence $g(F)$ is a union of finite orbits of H, and therefore $g(F) = F$. Hence G has a finite orbit and so G is elementary.

Now assume that H is elliptic. Let F be the set of all points of B^n fixed by H. Then F is an m-plane of B^n. By conjugating G in $M(B^n)$, we may assume that $F = B^m$. If x is in F, and h is in H, and g is in G, then $g^{-1}hgx = x$, and so $hgx = gx$, and therefore gx is in F. Hence G maps F to itself. Let $\overline{G}$ be the subgroup of $M(B^m)$ obtained by restricting the elements of G to F. Then H is a subgroup of the kernel of the restriction homomorphism $\rho : G \to \overline{G}$. Hence ρ induces a homomorphism from G/H onto $\overline{G}$. As G/H is abelian, $\overline{G}$ is abelian. Therefore $\overline{G}$ is elementary by Lemma 1. Hence $\overline{G}$, and therefore G, has a finite orbit in $\overline{F}$. Thus G is elementary. □

Theorem 5.5.11. *If* G *is a nonelementary subgroup of* $\mathrm{M}(B^n)$ *that leaves no proper* m*-plane of* B^n *invariant, then* G *has no nontrivial, elementary, normal subgroups.*

Proof: On the contrary, let H be a nontrivial, elementary, normal subgroup of G. Assume first that H is of elliptic type. Then the set F of all points of B^n fixed by H is a proper m-plane of B^n. Let x be a point of F, let ϕ be in H, and let ψ be in G. Then $\psi^{-1}\phi\psi(x) = x$, whence $\phi\psi(x) = \psi(x)$. Hence $\psi(x)$ is fixed by ϕ. As ϕ is arbitrary in H, we have that $\psi(x)$ is in F. As ψ is arbitrary in G, we deduce that G leaves F invariant, which is not the case.

Assume next that H is not of elliptic type. Then the union of all the finite orbits of H is a one or two point set F. Let ψ be in G. Then

$$\psi^{-1}H\psi(F) = HF = H.$$

Hence $H\psi(F) = \psi(F)$. Therefore $\psi(F) = F$. As ψ is arbitrary in G, we deduce that $GF = F$, which is not the case because G is nonelementary. Thus, we have a contradiction. □

Corollary 1. *If $n > 1$, then $\mathrm{M}(B^n)$ has no nontrivial, solvable, normal subgroups.*

Proof: By Theorem 3.1.5, we have that $\mathrm{M}(B^n)$ leaves no proper m-plane of B^n invariant. Furthermore, since $\mathrm{M}(B^n)$ acts transitively on S^{n-1}, we have that $\mathrm{M}(B^n)$ is nonelementary for $n > 1$. Therefore $\mathrm{M}(B^n)$ has no nontrivial, solvable, normal subgroups by Theorems 5.5.10 and 5.5.11. □

Remark: The group $\mathrm{M}(B^n)$ is isomorphic to $\mathrm{I}(H^n)$. Therefore $\mathrm{I}(H^n)$ has no nontrivial, solvable, normal subgroups for $n > 1$. In contrast, both $\mathrm{I}(S^n)$ and $\mathrm{I}(E^n)$ have nontrivial, abelian, normal subgroups.

The group $\mathrm{M}(B^n)$ has a nontrivial, abelian, quotient group because the subgroup $\mathrm{M}_0(B^n)$ of orientation preserving isometries of B^n has index two. It follows from the next theorem that $\mathrm{M}_0(B^n)$ is the only proper normal subgroup of $\mathrm{M}(B^n)$ whose group of cosets is abelian.

Theorem 5.5.12. *The group $\mathrm{M}_0(B^n)$ has no nontrivial, abelian, quotient groups.*

Proof: It suffices to show that $\mathrm{M}_0(B^n)$ is equal to its commutator subgroup. We pass to the upper half-space model U^n. The group $\mathrm{M}_0(U^n)$ is generated by all products $\gamma = \sigma_1\sigma_2$ of two reflections in spheres Σ_1 and Σ_2 of $\hat{E}^n$ that are orthogonal to E^{n-1}. There is a sphere Σ of $\hat{E}^n$ that is orthogonal to E^{n-1} and tangent to both Σ_1 and Σ_2. Let σ be the reflection in Σ. Then $\beta_1 = \sigma_1\sigma$ and $\beta_2 = \sigma\sigma_2$ are parabolic translations. This is clear upon positioning the spheres so that ∞ is the point of tangency. As $\gamma = \beta_1\beta_2$, we find that $\mathrm{M}_0(U^n)$ is generated by the set of all parabolic translations of U^n.

Now as any parabolic translation of U^n is conjugate in $\mathrm{M}_0(U^n)$ to the parabolic translation τ of U^n, defined by $\tau(x) = e_1 + x$, it suffices to show that τ is a commutator. Let μ be the magnification of U^n defined by $\mu(x) = 2x$. Then

$$\begin{aligned}
\mu\tau\mu^{-1}\tau^{-1}(x) &= \mu\tau\mu^{-1}(-e_1 + x) \\
&= \mu\tau(-e_1/2 + x/2) \\
&= \mu(e_1/2 + x/2) \\
&= e_1 + x.
\end{aligned}$$

Therefore $\tau = [\mu, \tau]$. □

We now define an elementary subgroup of $\mathrm{I}(H^n)$. Let $\zeta : B^n \to H^n$ be stereographic projection.

Definition: A subgroup Γ of $\mathrm{I}(H^n)$ is *elementary* if and only if the subgroup $\zeta^{-1}\Gamma\zeta$ of $\mathrm{I}(B^n)$ corresponds to an elementary subgroup of $\mathrm{M}(B^n)$ under the natural isomorphism from $\mathrm{I}(B^n)$ to $\mathrm{M}(B^n)$.

All the results of this section now apply to elementary subgroups of $\mathrm{I}(H^n)$. Furthermore, it is clear that we can define in a similar fashion elementary subgroups of the group of isometries of any model of hyperbolic space and all the results of this section apply to any model of hyperbolic space.

Exercise 5.5

1. Let G be an elementary subgroup of $\mathrm{M}(B^n)$ of hyperbolic type. Prove that G has a hyperbolic element and that every element of G is either elliptic or hyperbolic.

2. Let ϕ, ψ be elliptic elements in $\mathrm{M}(B^n)$. Prove that if ϕ and ψ commute, then either $F_\phi \subset F_\psi$ or $F_\psi \subset F_\phi$ or F_ϕ and F_ψ intersect orthogonally.

3. Let G be an abelian subgroup of $\mathrm{M}(B^n)$. Prove that

   (1) G is of elliptic type if and only if every element of G is elliptic,

   (2) G is of parabolic type if and only if G has a parabolic element, and

   (3) G is of hyperbolic type if and only if G has a hyperbolic element.

4. Let ϕ, ψ be in $\mathrm{M}(B^n)$ and suppose that ϕ and ψ have a common fixed point in $\overline{B}^n$. Prove that $[\phi, \psi]$ is either elliptic or parabolic.

5. Let G be a subgroup of $\mathrm{M}(B^n)$ with no nonidentity elliptic elements. Prove that G is elementary if and only if any two elements of G have a common fixed point.

§5.6. Historical Notes

§5.1. The quadratic form of the Hermitian inner product was introduced by Hermite in his 1854 paper *Sur la théorie des formes quadratiques* [189]. Complex n-space was described by Klein in his 1873 paper *Ueber die sogenannte Nicht-Euklidische Geometrie* [227]. The concept of a topological group evolved out of the notion of a continuous group of transformations of n-dimensional space as developed by Lie, Killing, and Cartan in the late nineteenth century. For an overview of the relationship between continuous groups and geometry, see Cartan's 1915 survey article *La théorie des groupes continus et la géométrie* [68]. Abstract topological groups were introduced by Schreier in his 1925 paper *Abstrakte kontinuierliche Gruppen* [368]. A systematic development of the algebra of matrices was first given by Cayley in his 1858 paper *A memoir on the theory of matrices* [75]. For the early history of matrix algebra, see Hawkins' 1977 articles *Another look at Cayley and the theory of matrices* [182] and *Weierstrass and the theory of matrices* [183]. Unitary transformations were studied by Frobenius in his 1883 paper *Über die principale Transformation der Thetafunctionen mehrerer Variabeln* [141]. The unitary group appeared in Autonne's 1902 paper *Sur l'Hermitien* [28]. Quotient topological groups were considered by Schreier in his 1925 paper [368]. Theorem 5.1.4 appeared in Pontrjagin's 1939 treatise *Topological Groups* [343]. The n-dimensional projective general linear group appeared in Klein's 1873 paper [227].

§5.2. The group of isometries of a finitely compact metric space was shown to have a natural topological group structure by van Dantzig and van der Waerden in their 1928 paper *Über metrisch homogene Räume* [393]. See also Koecher and Roelcke's 1959 paper *Diskontinuierliche und diskrete Gruppen von Isometrien metrischer Räume* [246]. As a reference for the compact-open topology, see Dugundji's 1966 text *Topology* [110]. Theorem 5.2.8 appeared in Beardon's 1983 text *The Geometry of Discrete Groups* [34].

§5.3. Discrete groups of Euclidean isometries were studied implicitly by crystallographers in the first half of the nineteenth century. For the early history of group theory in crystallography, see Scholz's 1989 articles *The rise of symmetry concepts in the atomistic and dynamistic schools of crystallography, 1815-1830* [364] and *Crystallographic symmetry concepts and group theory (1850-1880)* [365]. Discrete groups of Euclidean isometries were first studied explicitly by Jordan in his 1869 *Mémoire sur les groupes de mouvements* [206]. In particular, the 3-dimensional versions of Corollary 1 and Theorem 5.3.2 appeared in Jordan's paper. Lattices arose in crystallography, in the theory of quadratic forms, and in the theory of elliptic functions during the nineteenth century. Finite groups and subgroups of the elliptic modular group were the first discrete linear groups studied. In particular, Klein determined all the finite groups of linear fractional transformations of the complex plane in his 1876 paper *Ueber binäre Formen* *mit linearen Transformationen in sich selbst* [229]. Subgroups of the elliptic modular group were investigated by Klein in his 1879 paper *Ueber die Transformation der elliptischen Functionen* [231]. The term *discrete group* was used informally by Schreier in his 1925 paper [368]. A *discrete topological group* was defined by Pontrjagin in his 1939 treatise [343].

Poincaré defined a *discontinuous group* to be a group of linear fractional transformations of the complex plane that has no infinitesimal operations in his 1881 note *Sur les fonctions fuchsiennes* [327]. He defined a *Fuchsian group* to be a discontinuous group that leaves invariant a circle. Poincaré knew that a Fuchsian group is equivalent to a discrete group of isometries of the hyperbolic plane. Klein pointed out that there are discrete groups of linear fractional transformations of the complex plane that do not act discontinuously anywhere on the plane in his 1883 paper *Neue Beiträge zur Riemannschen Funktionentheorie* [233]. Poincaré then defined a *properly discontinuous group* to be a group of linear fractional transformations of the complex plane that acts discontinuously on a nonempty open subset of the plane in his 1883 *Mémoire sur les groupes kleinéens* [332]. He called such a group a *Kleinian group.* Poincaré knew that a Kleinian group acts as a discrete group of isometries of the upper half-space model of hyperbolic 3-space. See Poincaré's 1881 note *Sur les groupes kleinéens* [329]. In modern terminology, a *Kleinian group* is any discrete group of linear fractional transformations of the complex plane. Moreover, the terms discontinuous and properly discontinuous have been replaced by discrete and discontinuous, respectively. For the evolution of the definition of a discontinuous group, see Fenchel's 1957 article *Bemerkungen zur allgemeinen Theorie der diskontinuierlichen Transformationsgruppen* [131]. Theorem 5.3.3 appeared in Fubini's 1905 paper *Sulla teoria dei gruppi discontinui* [143]. Theorem 5.3.4 for groups of isometries appeared in Bers and Gardiner's 1986 paper *Fricke Spaces* [43]. Theorem 5.3.5 for groups of isometries of hyperbolic space was proved by Poincaré in his 1883 memoir [332]. Theorem 5.3.5 was essentially proved by Siegel in his 1943 paper *Discontinuous groups* [375]. See also Koecher and Roelcke's 1959 paper [246].

Poincaré was led to investigate discrete groups of isometries of the hyperbolic plane because of his work on differential equations of functions of a complex variable. In particular, Poincaré studied functions f of a complex variable z with the property that $f(\gamma z) = f(z)$ for all elements γ of a discrete group Γ of linear fractional transformations of the complex plane. Such a function f is called an *automorphic function* with respect to the group Γ. For the fascinating history of this line of research, see Gray's 1986 monograph *Linear Differential Equations and Group Theory from Riemann to Poincaré* [160]. References for the theory of Fuchsian and Kleinian groups are Fricke and Klein's 1897-1912 treatise *Vorlesungen über die Theorie der automorphen Functionen* [139], Ford's 1929 treatise *Automorphic Functions* [136], Lehner's 1964 treatise *Discontinuous Groups and Automorphic Functions* [255], Magnus' 1974 treatise *Noneuclidean Tesse-*

lations and their Groups [272], Beardon's 1983 text [34], and Maskit's 1988 treatise *Kleinian Groups* [282].

§5.4. The 3-dimensional version of Theorem 5.4.1 was first proved by Chasles in his 1831 paper *Note sur les propriétés générales du système de deux corps semblables entr'eux* [78]. Theorems 5.4.1 and 5.4.2 appeared in Jordan's 1875 paper *Essai sur la géométrie à n dimensions* [207]. Lemma 3 was proved by Frobenius in his 1911 paper *Über den von L. Bieberbach gefundenen Beweis eines Satzes von C. Jordan* [142]. Lemma 4 for finite subgroups of the orthogonal group also appeared in this paper. Lemmas 4, 5, and 7 appeared in Oliver's 1980 paper *On Bieberbach's analysis of discrete Euclidean groups* [323]. Theorem 5.4.3 was first proved for finite subgroups of the orthogonal group by Jordan in his 1878 *Mémoire sur les équations différentielles linéaires* [208] and in his 1880 paper *Sur la détermination des groupes d'ordre fini contenus dans le groupe linéaire* [209]. Theorem 5.4.3 follows easily from Jordan's theorem and Bieberbach's algebraic characterization of discrete Euclidean groups given in his 1911 paper *Über die Bewegungsgruppen der Euklidischen Räume* [46]. Likewise, Theorems 5.4.4-5.4.6 follow from Bieberbach's characterization in this paper.

§5.5. The concept of an elementary group is implicit in the classification of discontinuous groups of linear fractional transformations of the complex plane given by Fricke and Klein in Vol. I of their 1897 treatise [139]. The term *elementary group* was introduced by Ford in his 1929 treatise [136]. Our definition of an elementary group conforms with the definition of an elementary group in dimension three given by Beardon in his 1983 text [34]. The 2-dimensional version of Theorem 5.5.4 appeared on p. 118 in Vol. I of Fricke and Klein's 1897 treatise [139]. Theorem 5.5.5 appeared in Greenberg's 1974 paper *Commensurable groups of Moebius transformations* [165]. Theorems 5.5.7 and 5.5.8 were proved by Tukia in his 1985 paper *On isomorphisms of geometrically finite Möbius groups* [392]. Theorem 5.5.9 appeared in Martin's 1989 paper *On discrete Möbius groups in all dimensions* [280]. The 3-dimensional version of Theorem 5.5.10 was essentially proved by Myrberg in his 1941 paper *Die Kapazität der singulären Menge der linearen Gruppen* [312]. Theorem 5.5.11 was essentially proved by Chen and Greenberg in their 1974 paper *Hyperbolic spaces* [80]. Theorem 5.5.12 follows from the fact that $\mathrm{M}_0(B^n)$ is a simple Lie group. References for elementary groups are Ford's 1929 treatise [136], Beardon's 1983 text [34], Kulkarni's 1988 paper *Conjugacy classes in* $\mathrm{M}(n)$ [248], and Waterman's 1988 paper *Purely elliptic Möbius groups* [404].

CHAPTER 6

Geometry of Discrete Groups

In this chapter, we study the geometry of discrete groups of isometries of S^n, E^n, and H^n. The chapter begins with an introduction to the projective disk model of hyperbolic n-space. Convex sets, polyhedra, and polytopes in S^n, E^n, and H^n are studied in Sections 6.2, 6.3, and 6.4, respectively. The basic properties of fundamental domains for a discrete group are examined in Sections 6.5 and 6.6. The chapter ends with a study of the basic properties of tessellations of S^n, E^n, and H^n.

§6.1. The Projective Disk Model

The *open unit n-disk* in $\mathbb{R}^n$ is defined to be the set

$$D^n = \{x \in \mathbb{R}^n : |x| < 1\}.$$

Note that D^n is the same set as B^n. The reason for the new notation is that a new metric d_D on D^n will be defined so that D^n and B^n are different metric spaces.

Identify $\mathbb{R}^n$ with $\mathbb{R}^n \times \{0\}$ in $\mathbb{R}^{n+1}$. The *gnomonic projection* μ of D^n onto H^n is defined to be the composition of the vertical translation of D^n by e_{n+1} followed by radial projection to H^n. See Figure 6.1.1. An explicit formula for μ is given by

$$\mu(x) = \frac{x + e_{n+1}}{|||x + e_{n+1}|||}. \tag{6.1.1}$$

The map $\mu : D^n \to H^n$ is a bijection. The inverse of μ is given by

$$\mu^{-1}(x_1, \ldots, x_{n+1}) = (x_1/x_{n+1}, \ldots, x_n/x_{n+1}). \tag{6.1.2}$$

Define a metric d_D on D^n by

$$d_D(x, y) = d_H(\mu(x), \mu(y)). \tag{6.1.3}$$

By definition, μ is an isometry from D^n, with the metric d_D, to hyperbolic n-space H^n. The metric space consisting of D^n, together with the metric d_D, is called the *projective disk model* of hyperbolic n-space.

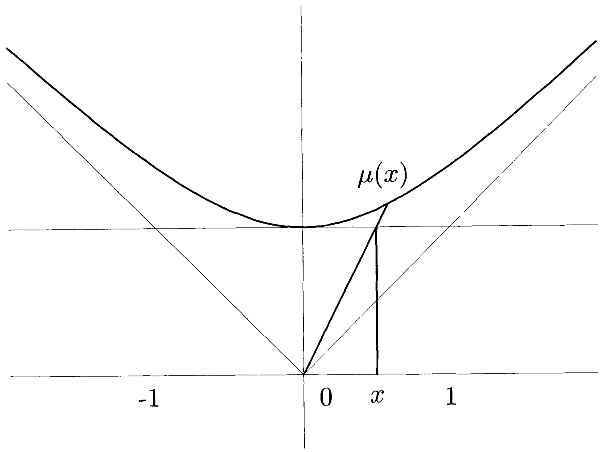


Figure 6.1.1. The gnomonic projection μ of D^1 onto H^1

Theorem 6.1.1. *The metric d_D on D^n is given by*

$$\cosh d_D(x,y) = \frac{1 - x\cdot y}{\sqrt{1-|x|^2}\sqrt{1-|y|^2}}.$$

Proof: By Formula 3.2.2, we have

$$\begin{aligned}
\cosh d_D(x,y) &= \cosh d_H(\mu(x),\mu(y)) \\
&= -\frac{x+e_{n+1}}{|||x+e_{n+1}|||} \circ \frac{y+e_{n+1}}{|||y+e_{n+1}|||} \\
&= \frac{1-x\cdot y}{\sqrt{1-|x|^2}\sqrt{1-|y|^2}}.
\end{aligned}$$

□

In order to understand the isometries of D^n, we need to introduce homogeneous coordinates for projective n-space P^n and classical projective n-space $\overline{\mathbb{R}}^n$. By definition, $P^n = S^n/\{\pm 1\}$. Thus, a point of P^n is a pair of antipodal points of S^n. The idea of homogeneous coordinates is to use any nonzero vector on the line passing through a pair $\pm x$ of antipodal points of S^n to represent the point $\{\pm x\}$ of P^n. With this in mind, we say that a nonzero vector x in $\mathbb{R}^{n+1}$ is a set of *homogeneous coordinates* for the point $\{\pm x/|x|\}$ of P^n. Notice that two nonzero vectors x, y in $\mathbb{R}^{n+1}$ are homogeneous coordinates for the same point of P^n if and only if each is a nonzero scalar multiple of the other. By definition, $\overline{\mathbb{R}}^n = \mathbb{R}^n \cup P^{n-1}$. Moreover, gnomonic projection $\nu : \mathbb{R}^n \to S^n$ induces a bijection $\overline{\nu} : \overline{\mathbb{R}}^n \to P^n$. A set of homogeneous coordinates for a point x of $\overline{\mathbb{R}}^n$ is a set of homogeneous coordinates for the point $\overline{\nu}(x)$. In particular, if $x_{n+1} \neq 0$, then $(x_1, \ldots, x_{n+1})$ is a set of homogeneous coordinates for the point $(x_1/x_{n+1}, \ldots, x_n/x_{n+1})$ of $\mathbb{R}^n$ in $\overline{\mathbb{R}}^n$.

A *projective transformation* of P^n is a bijection $\phi : P^n \to P^n$ that corresponds to a bijective linear transformation $\tilde{\phi} : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ with respect to homogeneous coordinates that is determined only up to multiplication by a nonzero scalar. In other words, a projective transformation of P^n corresponds to an element of $\mathrm{PGL}(n+1, \mathbb{R})$. Projective transformations of $\overline{\mathbb{R}}^n$ correspond to projective transformations of P^n via the bijection $\overline{\nu} : \overline{\mathbb{R}}^n \to P^n$.

Theorem 6.1.2. *Every isometry of D^n extends to a unique projective transformation of classical projective n-space $\overline{\mathbb{R}}^n$ and every projective transformation of $\overline{\mathbb{R}}^n$ that leaves D^n invariant restricts to an isometry of D^n.*

Proof: Let ϕ be a projective transformation of $\overline{\mathbb{R}}^n$. Then ϕ corresponds to a bijective linear transformation $\tilde{\phi}$ of $\mathbb{R}^{n+1}$ that is unique up to multiplication by a nonzero scalar. Let $(x_1, \ldots, x_{n+1})$, with $x_{n+1} \neq 0$, be a set of homogeneous coordinates for the vector $(x_1/x_{n+1}, \ldots, x_n/x_{n+1})$ in $\mathbb{R}^n$. Then

$$\left(\frac{x_1}{x_{n+1}}\right)^2 + \cdots + \left(\frac{x_n}{x_{n+1}}\right)^2 < 1$$

if and only if

$$x_1^2 + \cdots + x_n^2 < x_{n+1}^2.$$

Hence ϕ leaves D^n invariant if and only if $\tilde{\phi}$ leaves invariant the interior of the light cone C^n in $\mathbb{R}^{n,1}$ defined by the equation

$$x_1^2 + \cdots + x_n^2 = x_{n+1}^2.$$

Suppose that $\tilde{\phi}$ leaves invariant the interior of the light cone C^n. We claim that some nonzero scalar multiple of $\tilde{\phi}$ is a positive Lorentz transformation. Since $\tilde{\phi}$ is continuous, $\tilde{\phi}$ either leaves invariant the positive and negative components of the interior of C^n or permutes them. By multiplying $\tilde{\phi}$ by -1, if necessary, we may assume that $\tilde{\phi}$ leaves invariant the components of the interior of C^n. By composing $\tilde{\phi}$ with a positive Lorentz transformation, we may assume that $\tilde{\phi}$ leaves invariant the $(n+1)$st axis of $\mathbb{R}^{n+1}$. By multiplying $\tilde{\phi}$ by a positive scalar, we may assume that $\tilde{\phi}$ fixes the unit vector e_{n+1}. We now show that $\tilde{\phi}$ is an orthogonal transformation. Let x be a vector in $\mathbb{R}^{n+1}$ not on the $(n+1)$st axis of $\mathbb{R}^{n+1}$. It suffices to show that $|\tilde{\phi}(x)| = |x|$. Let V be the 2-dimensional vector subspace of $\mathbb{R}^{n+1}$ spanned by x and e_{n+1}. By composing $\tilde{\phi}$ with an orthogonal transformation of $\mathbb{R}^{n+1}$ that fixes e_{n+1}, we may assume that $\tilde{\phi}$ leaves V invariant. Consequently, we may assume that $n = 1$. Then the matrix for $\tilde{\phi}$ is of the form

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

Now since $\tilde{\phi}$ leaves invariant the light cone, and since

$$\begin{pmatrix} a & 0 \\ b & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ b+1 \end{pmatrix},$$

we have that $a = \pm(b+1)$. By composing $\tilde{\phi}$ with the reflection

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

if necessary, we may assume that $a = b+1$. Then we have

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

with $a = -b+1$. Hence $a = 1$ and $b = 0$. Therefore $\tilde{\phi}$ is the identity. Hence $\tilde{\phi}$ is an orthogonal transformation that fixes e_{n+1}. Therefore $\tilde{\phi}$ is a positive Lorentz transformation. Thus $\tilde{\phi}$ leaves the interior of the light cone C^n invariant if and only if some nonzero scalar multiple of $\tilde{\phi}$ is a positive Lorentz transformation.

Now every isometry of H^n extends to a unique positive Lorentz transformation of $\mathbb{R}^{n,1}$, and every positive Lorentz transformation of $\mathbb{R}^{n,1}$ restricts to an isometry of H^n by Theorem 3.2.3. Moreover, the isometries of H^n correspond via the isometry $\mu^{-1} : H^n \to D^n$, defined by

$$\mu^{-1}(x_1, \ldots, x_{n+1}) = (x_1/x_{n+1}, \ldots, x_n/x_{n+1}),$$

to the isometries of D^n. Therefore, every isometry of D^n extends to a unique projective transformation of $\overline{\mathbb{R}}^n$, and every projective transformation of $\overline{\mathbb{R}}^n$ that leaves D^n invariant restricts to an isometry of D^n. □

Theorem 6.1.3. *A function $\phi : D^n \to D^n$ fixing the origin is an isometry of D^n if and only if it is the restriction of an orthogonal transformation of $\mathbb{R}^n$.*

Proof: If ϕ is the restriction of an orthogonal transformation of $\mathbb{R}^n$, then ϕ is an isometry of D^n by Theorem 6.1.1. Now assume that ϕ is an isometry. Then ϕ extends to a projective transformation $\hat{\phi}$ of $\overline{\mathbb{R}}^n$ and $\hat{\phi}$ corresponds to a bijective linear transformation $\tilde{\phi}$ of $\mathbb{R}^{n+1}$ with respect to homogeneous coordinates that is unique up to multiplication by a nonzero scalar. The unit vector e_{n+1} in $\mathbb{R}^{n+1}$ is a set of homogeneous coordinates for the origin in D^n. Hence $\tilde{\phi}$ leaves the $(n+1)$st axis invariant. Thus, by multiplying $\tilde{\phi}$ by a nonzero scalar, we may assume that $\tilde{\phi}$ fixes the vector e_{n+1}. Now by the same argument as in the proof of Theorem 6.1.2, we deduce that $\tilde{\phi}$ is an orthogonal transformation of $\mathbb{R}^{n+1}$. Now since $\tilde{\phi}$ restricts to ϕ on D^n, we have that ϕ is the restriction of an orthogonal transformation of $\mathbb{R}^n$. □

A subset P of D^n is said to be a *hyperbolic m-plane* of D^n if and only if $\mu(P)$ is a hyperbolic m-plane of H^n.

Theorem 6.1.4. *A subset P of D^n is a hyperbolic m-plane of D^n if and only if P is the nonempty intersection of D^n with an m-plane of $\mathbb{R}^n$.*

Proof: Let Q be a hyperbolic m-plane of H^n. Then Q is the intersection of H^n with an $(m+1)$-dimensional time-like vector subspace V of $\mathbb{R}^{n+1}$. Observe that μ^{-1} is the composite of the radial projection of H^n onto the hyperplane $P(e_{n+1}, 1)$ followed by the translation by $-e_{n+1}$. Clearly, radial projection maps Q onto the intersection of the m-plane $V \cap P(e_{n+1}, 1)$ with the interior of the light-cone C^n of $\mathbb{R}^{n,1}$. Thus $\mu^{-1}(Q)$ is the nonempty intersection of D^n with an m-plane of $\mathbb{R}^n$. Clearly, we can reverse the argument and show that any nonempty intersection of D^n with an m-plane of $\mathbb{R}^n$ is the image under μ^{-1} of a hyperbolic m-plane of H^n. □

A *hyperbolic line* of D^n is defined to be a hyperbolic 1-plane of D^n.

Corollary 1. *The hyperbolic lines of D^n are the open chords of D^n.*

Remark: The fact that the hyperbolic m-planes of D^n conform with Euclidean m-planes makes the projective model very useful for convexity arguments. However, one must keep in mind that the hyperbolic angles of D^n do not necessarily conform with the Euclidean angles; in other words, D^n is not a conformal model of hyperbolic n-space.

Theorem 6.1.5. *The element of hyperbolic arc length of the projective disk model D^n is*

$$\frac{[(1-|x|^2)|dx|^2 + (x \cdot dx)^2]^{\frac{1}{2}}}{1-|x|^2}.$$

Proof: Let $y = \mu(x)$. From the results of §3.3, the element of hyperbolic arc length of H^n is

$$\|dy\| = (dy_1^2 + \cdots + dy_n^2 - dy_{n+1}^2)^{\frac{1}{2}}.$$

Now since

$$y_i = \frac{x_i}{(1-|x|^2)^{1/2}} \quad \text{for } i = 1, \ldots, n,$$

we have

$$dy_i = \frac{dx_i}{(1-|x|^2)^{1/2}} + \frac{x_i(x \cdot dx)}{(1-|x|^2)^{3/2}}.$$

Hence

$$dy_i^2 = \frac{1}{1-|x|^2}\left(dx_i^2 + \frac{2x_i dx_i (x \cdot dx)}{1-|x|^2} + \frac{x_i^2 (x \cdot dx)^2}{(1-|x|^2)^2}\right).$$

Thus

$$\begin{aligned} \sum_{i=1}^n dy_i &= \frac{1}{1-|x|^2}\left(|dx|^2 + \frac{2(x \cdot dx)^2}{1-|x|^2} + \frac{|x|^2(x \cdot dx)^2}{(1-|x|^2)^2}\right) \\ &= \frac{1}{1-|x|^2}\left(|dx|^2 + \frac{(2-|x|^2)(x \cdot dx)^2}{(1-|x|^2)^2}\right). \end{aligned}$$

Now since

$$y_{n+1} = \frac{1}{(1-|x|^2)^{1/2}},$$

we have that

$$dy_{n+1} = \frac{x \cdot dx}{(1-|x|^2)^{3/2}}.$$

Thus

$$\sum_{i=1}^{n} dy_i^2 - dy_{n+1}^2 = \frac{(1-|x|^2)|dx|^2 + (x \cdot dx)^2}{(1-|x|^2)^2}. \qquad \square$$

Theorem 6.1.6. *The element of hyperbolic volume of the projective disk model D^n is*

$$\frac{dx_1 \cdots dx_n}{(1-|x|^2)^{\frac{n+1}{2}}}.$$

Proof: By Theorem 3.4.1, the element of hyperbolic volume of H^n, with respect to the Euclidean coordinates $y_1, \ldots, y_n$, is given by

$$\frac{dy_1 \cdots dy_n}{[1 + (y_1^2 + \cdots + y_n^2)]^{\frac{1}{2}}}.$$

To find the element of hyperbolic volume of D^n, we change coordinates via the map $\overline{\mu} : D^n \to \mathbb{R}^n$ defined by

$$\overline{\mu}(x) = \frac{x}{(1-|x|^2)^{\frac{1}{2}}}.$$

As μ is a radial map, it is best to switch to spherical coordinates and decompose $\overline{\mu}$ into the composite

$$\begin{aligned} (x_1, \ldots, x_n) &\mapsto (\rho, \theta_1, \ldots, \theta_{n-1}) \\ &\mapsto \left(\frac{\rho}{(1-\rho^2)^{\frac{1}{2}}}, \theta_1, \ldots, \theta_{n-1}\right) \\ &\mapsto (y_1, \ldots, y_n). \end{aligned}$$

Now as

$$\frac{d}{d\rho}\left(\frac{\rho}{(1-\rho^2)^{\frac{1}{2}}}\right) = \frac{1}{(1-\rho^2)^{\frac{3}{2}}},$$

the Jacobian of $\overline{\mu}$ is

$$\frac{1}{\rho^{n-1}} \frac{1}{(1-\rho^2)^{\frac{3}{2}}} \left(\frac{\rho}{(1-\rho^2)^{\frac{1}{2}}}\right)^{n-1} = \frac{1}{(1-\rho^2)^{\frac{n+2}{2}}}.$$

Therefore

$$\begin{aligned} \frac{dy_1 \cdots dy_n}{[1 + (y_1^2 + \cdots + y_n^2)]^{\frac{1}{2}}} &= \frac{1}{(1-|x|^2)^{\frac{n+2}{2}}} \quad \frac{dx_1 \cdots dx_n}{\left(1 + \frac{|x|^2}{1-|x|^2}\right)^{\frac{1}{2}}} \\ &= \frac{dx_1 \cdots dx_n}{(1-|x|^2)^{\frac{n+1}{2}}}. \end{aligned} \qquad \square$$

Exercise 6.1

1. Show that the hyperbolic angle between any two geodesic lines of D^n intersecting at the origin conforms with the Euclidean angle between the lines. In other words, D^n is conformal at the origin.

2. Let P be a hyperplane of D^n. Prove that all the tangent lines of S^{n-1} at the points of $\overline{P} \cap S^{n-1}$ intersect in a unique point of classical real projective n-space $\overline{\mathbb{R}}^n$ called the *pole* of P. See Figure 1.2.2.

3. Prove that a line L of D^n is orthogonal to a hyperplane P of D^n if and only if the projective line extending L passes through the pole of P.

4. Prove that the correspondence between a hyperplane of D^n and its pole gives a one-to-one correspondence between the set of hyperplanes of D^n and the points of $\overline{\mathbb{R}}^n - D^n$.

5. Let x be a point of D^n. Define an inner product $\langle\ ,\ \rangle_x$ on $\mathbb{R}^n$ by
$$\langle e_i, e_j\rangle_x = \begin{cases} 1 - |x|^2 + x_i^2 & \text{if } i = j, \\ x_i x_j & \text{if } i \neq j. \end{cases}$$
Let $\lambda, \mu : \mathbb{R} \to D^n$ be geodesic lines such that $\lambda(0) = x = \mu(0)$, and let $u = \lambda'(0)$ and $v = \mu'(0)$. Show that the hyperbolic angle θ between λ and μ is given by the formula
$$\cos\theta = \frac{\langle u, v\rangle_x}{\langle u, u\rangle_x^{\frac{1}{2}} \langle v, v\rangle_x^{\frac{1}{2}}}.$$

§6.2. Convex Sets

Throughout this section, $X = S^n, E^n$, or H^n with $n > 0$. A pair of points x, y of X is said to be *proper* if and only if x, y are distinct and x, y are not antipodal points of $X = S^n$. If x, y are a proper pair of points of X, then there is a unique geodesic segment in X joining x to y. We shall denote this segment by $[x, y]$.

Definition: A subset C of X is *convex* if and only if for each pair of proper points x, y of C, the geodesic segment $[x, y]$ is contained in C.

In order to have uniformity in terminology, we shall define an m-plane of S^n to be a great m-sphere of S^n.

Example: Every m-plane of X is convex. In particular, every pair of antipodal points of S^n is convex!

Remark: It is obvious from the definition of convexity in X that an arbitrary intersection of convex subsets of X is convex.

Let C be a nonempty convex subset of X.

(1) The *dimension* of C is defined to be the least integer m such that C is contained in an m-plane of X.

(2) If $\dim C = m$, then clearly C is contained in a unique m-plane of X, which is denoted by $\langle C \rangle$.

(3) The *interior* of C is the topological interior of C in $\langle C \rangle$ and is denoted by C°.

(4) The *boundary* of C is the topological boundary of C in $\langle C \rangle$ and is denoted by ∂C.

(5) The *closure* of C is the topological closure of C in X and is denoted by $\overline{C}$. Note that $\overline{C}$ is also the topological closure of C in $\langle C \rangle$, since $\langle C \rangle$ is closed in X. Therefore $\overline{C}$ is the disjoint union of C° and ∂C.

If C is the empty set, then the dimension of C is undefined, and all the sets $\langle C \rangle$, C°, ∂C, and $\overline{C}$ are empty by definition.

Lemma 1. *Let x, y be a proper pair of points of X. Then there is an $r > 0$ such that if u is in $B(x, r)$ and v is in $B(y, r)$, then u, v is a proper pair.*

Proof: This is clear if $X = E^n$ or H^n. Assume that $X = S^n$. Observe that the sets $\{\pm x\}$ and $\{\pm y\}$ are disjoint, since x, y is a proper pair of points. Let r be half the distance from $\{\pm x\}$ to $\{\pm y\}$. Then $B(x, r)$, $B(y, r)$, and $B(-x, r)$ are mutually disjoint. As $-B(x, r) = B(-x, r)$, no point of $B(x, r)$ can be antipodal to a point of $B(y, r)$. □

Theorem 6.2.1. *If C is a convex subset of X, then so is $\overline{C}$.*

Proof: Let x, y be a proper pair of points in $\overline{C}$. By Lemma 1, there are proper pairs of points u_i, v_i, for $i = 1, 2, \ldots$, in C such that $u_i \to x$ and $v_i \to y$. Define a curve

$$\gamma : [0, 1] \to X$$

from x to y by

$$\gamma(t) = \begin{cases} (1-t)x + ty & \text{if } X = E^n, \\ \frac{(1-t)x+ty}{|(1-t)x+ty|} & \text{if } X = S^n, \\ \frac{(1-t)x+ty}{|||(1-t)x+ty|||} & \text{if } X = H^n. \end{cases}$$

Likewise, define a curve

$$\gamma_i(t) : [0, 1] \to C$$

from u_i to v_i for each i. Then clearly $\gamma_i(t) \to \gamma(t)$ for each t. Therefore $\gamma(t)$ is in $\overline{C}$ for each t. □

Given a proper pair of points x, y of X, let $[x, y)$ denote the segment $[x, y]$ minus its endpoint y.

Theorem 6.2.2. *Let C be a convex subset of X and let x, y be a proper pair of points in $\overline{C}$. If x is in C°, then $[x, y)$ is contained in C°.*

Proof: Without loss of generality, we may assume that $\langle C \rangle = X$. We first consider the case $X = E^n$. As x is in C°, there is an $r > 0$ such that $B(x, r)$ is contained in C. Let t be in the open interval $(0, 1)$ and let

$$z = (1-t)x + ty.$$

We need to show that z is in C°. Assume first that y is in C. Observe that z is in the set

$$(1-t)B(x, r) + ty = B(z, (1-t)r).$$

As $B(x, r)$ and y are both contained in C, we have that $B(z, (1-t)r)$ is contained in C, since C is convex. Thus z is in C°. See Figure 6.2.1.

Assume now that y is in ∂C. As y is in ∂C, the open ball $B(y, t^{-1}(1-t)r)$ contains a point v of C. Now since

$$B(y, t^{-1}(1-t)r) = t^{-1}(z - (1-t)B(x, r)),$$

there is a point u of $B(x, r)$ such that

$$v = t^{-1}(z - (1-t)u).$$

Then $z = (1-t)u + tv$. Let $w = (1-t)x + tv$. Then z is in the set

$$(1-t)B(x, r) + tv = B(w, (1-t)r).$$

As $B(x, r)$ and v are contained in C, we have that $B(w, (1-t)r)$ is contained in C. Therefore z is in C°. Thus (x, y) is contained in C°.

Next, assume that $X = H^n$. We now pass to the projective disk model D^n and regard C as a convex subset of D^n. Then C is also a convex subset of E^n. As D^n is open in E^n, we have that C° in D^n is the same as C° in E^n. Therefore $[x, y)$ is contained in C° by the Euclidean case.

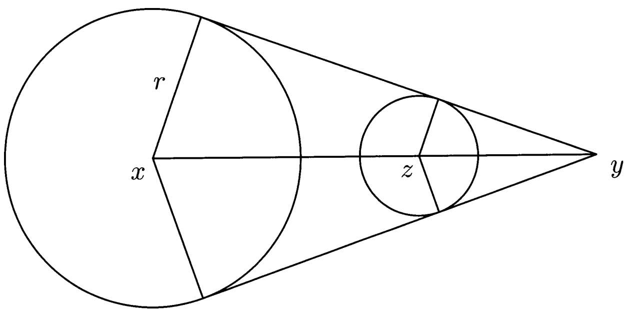


Figure 6.2.1. $B(z, (1-t)r) = (1-t)B(x, r) + ty$

Finally, assume that $X = S^n$. Let z be the midpoint of the geodesic segment $[x, y]$. Then $B(z, \pi/2)$ is an open hemisphere of S^n containing $[x, y]$. As x is in C°, we have that $C^\circ \cap B(z, \pi/2)$ is a nonempty open subset of S^n. Consequently

$$\langle C \cap B(z, \pi/2)\rangle = S^n.$$

By replacing C with $C \cap B(z, \pi/2)$, we may assume, without loss of generality, that C is contained in $B(z, \pi/2)$. We may also assume that $z = e_{n+1}$. Now by gnomonic projection, we can view C as a convex subset of E^n. Then $[x, y)$ is contained in C° by the Euclidean case. □

Theorem 6.2.3. *If C is a nonempty convex subset of X, then so is C°.*

Proof: That C° is convex follows immediately from Theorem 6.2.2. It remains to show that C° is nonempty. Without loss of generality, we may assume that $\langle C\rangle = X$. We first consider the case $X = E^n$. Then there exist $n+1$ vectors $v_0, \ldots, v_n$ in C such that $v_1 - v_0, \ldots, v_n - v_0$ are linearly independent. As C is convex, it contains every vector of the form $x = \sum_{i=0}^n t_i v_i$ with $t_i \geq 0$ and $\sum_{i=0}^n t_i = 1$. By applying an affine transformation of E^n, we may assume that $v_0 = 0$ and $v_i = e_i$ for $i > 0$.

Let $a = \left(\frac{1}{n+1}, \ldots, \frac{1}{n+1}\right)$ in E^n. We now show that $B\left(a, \frac{1}{n(n+1)}\right)$ is contained in C. Suppose that

$$|x - a| < \frac{1}{n(n+1)}.$$

Then we have

$$-\frac{1}{n(n+1)} < x_i - \frac{1}{n+1} < \frac{1}{n(n+1)}$$

and so

$$\frac{1}{(n+1)}\left(1 - \frac{1}{n}\right) < x_i < \frac{1}{(n+1)}\left(1 + \frac{1}{n}\right).$$

Therefore $0 < x_i < \frac{1}{n}$ for $i = 1, \ldots, n$. Hence $\sum_{i=1}^n x_i < 1$. This implies that x is in C. Consequently $B\left(a, \frac{1}{n(n+1)}\right)$ is contained in C. Thus a is in C° and so C° is nonempty.

Next, assume that $X = H^n$. We pass to the projective disk model D^n and regard C as a convex subset of D^n. Then C° is nonempty by the Euclidean case. Finally, assume that $X = S^n$. Then C contains a basis $v_1, \ldots, v_{n+1}$ of $\mathbb{R}^{n+1}$, since $\langle C\rangle = S^n$. Let P be the hyperplane of $\mathbb{R}^{n+1}$ containing $v_1, \ldots, v_{n+1}$. Then P does not contain the origin of $\mathbb{R}^{n+1}$. Let V be the n-dimensional vector subspace of $\mathbb{R}^{n+1}$ parallel to P, and let H be the open hemisphere of S^n whose boundary is $V \cap S^n$ and that contains $v_1, \ldots, v_{n+1}$. Then $\langle C \cap H\rangle = S^n$. By replacing C with $C \cap H$, we may assume that $C \subset H$. We may also assume that H is the upper hemisphere of S^n. Now by gnomonic projection, we can view C as a convex subset of E^n. Then C° is nonempty by the Euclidean case. □

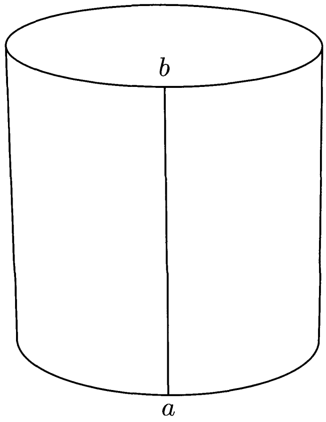


Figure 6.2.2. A right circular cylinder in E^3

Sides of a Convex Set

Definition: A *side* of a convex subset C of X is a nonempty, maximal, convex subset of ∂C.

Example: Let C be a right circular cylinder in E^3 situated as in Figure 6.2.2. Then the sides of C are the top and bottom of C and all the vertical line segments in ∂C joining the top to the bottom of C as $[a, b]$ in Figure 6.2.2. Notice that C has an uncountable number of sides.

Theorem 6.2.4. *If S is a side of a convex subset C of X, then*

$$\overline{C} \cap \langle S \rangle = S.$$

Proof: This is clear if $\dim S = 0$, so assume that $\dim S > 0$. We first show that C° and $\langle S \rangle$ are disjoint. Suppose that x is in both C° and $\langle S \rangle$. Now S° is nonempty by Theorem 6.2.3. As $\dim S > 0$, we can choose y in S° so that x and y are nonantipodal. As C° and ∂C are disjoint, $x \neq y$. Hence x, y is a proper pair of points. Now since y is in S°, there is an $r > 0$ such that

$$B(y, r) \cap \langle S \rangle \subset S.$$

By Theorem 6.2.2, the half-open geodesic segment $[x, y)$ is contained in C°. But observe that

$$[x, y) \cap B(y, r) \subset \langle S \rangle \cap B(y, r) \subset S \subset \partial C,$$

which is a contradiction. Therefore C° and $\langle S \rangle$ are disjoint.

Now as $\overline{C} = C^\circ \cup \partial C$, we have that $\overline{C} \cap \langle S \rangle \subset \partial C$. The set $\overline{C}$ is convex by Theorem 6.2.1. Hence $\overline{C} \cap \langle S \rangle$ is a convex subset of ∂C containing S. Therefore $\overline{C} \cap \langle S \rangle = S$ because of the maximality of S. □

Theorem 6.2.5. *Let P be an m-plane of X that contains an $(m-1)$-dimensional side S of a convex subset C of X. Then $C^\circ \cap P$ is contained in one of the components of $P - \langle S\rangle$; moreover, $\overline{C} \cap P$ is contained in one of the closed half-spaces of P bounded by $\langle S\rangle$.*

Proof: If $C^\circ \cap P = \emptyset$, then $\overline{C} \cap P = S$, since $\overline{C} \cap P$ is a convex subset of ∂C containing S. Hence, we may assume that $C^\circ \cap P \neq \emptyset$. Then $P \subset \langle C\rangle$, since $\langle S\rangle \subset P$ and P contains a point of C°. Therefore $C^\circ \cap P$ is a nonempty, open, convex subset of $P - \langle S\rangle$. On the contrary, suppose that x and y are points of $C^\circ \cap P$ contained in different components of $P - \langle S\rangle$. As $\dim(C^\circ \cap P) > 0$, we may assume that x and y are nonantipodal. Now since $[x, y]$ is connected, it must contain a point of $\langle S\rangle$. But $[x, y]$ is contained in C° by Theorem 6.2.3, and C° is disjoint from $\langle S\rangle$ by Theorem 6.2.4, which is a contradiction. Therefore $C^\circ \cap P$ is contained in a component of $P - \langle S\rangle$.

Clearly, we have
$$\overline{C^\circ \cap P} \subset \overline{C} \cap P.$$
Let y be in $\partial C \cap P$ and choose x in $C^\circ \cap P$ so that x, y are nonantipodal. By Theorem 6.2.2, the set $C^\circ \cap P$ contains $[x, y)$. Therefore y is in $\overline{C^\circ \cap P}$. Thus $\overline{C^\circ \cap P} = \overline{C} \cap P$. Consequently $\overline{C} \cap P$ is contained in one of the closed half-spaces of P bounded by $\langle S\rangle$ by the first part of the theorem. □

Theorem 6.2.6. *If C is a convex subset of X, then*

(1) *every nonempty convex subset of ∂C is contained in a side of C;*

(2) *every side of C is closed;*

(3) *the sides of C meet only along their boundaries.*

Proof: (1) Let K be a nonempty convex subset of ∂C and let $\mathcal{K}$ be the set of all convex subsets of ∂C containing K. Then $\mathcal{K}$ is partially ordered by inclusion and nonempty, since $\mathcal{K}$ contains K. Let $\mathcal{C}$ be a chain of $\mathcal{K}$. Then the union of the elements of $\mathcal{C}$ is obviously convex and an upper bound for $\mathcal{C}$. Therefore $\mathcal{K}$ has a maximal element by Zorn's lemma.

(2) Let S be a side of C. Then $\overline{S}$ is convex by Theorem 6.2.1. Also $\overline{S}$ is contained in ∂C, since ∂C is closed. Therefore $S = \overline{S}$ because of the maximality of S. Thus S is closed.

(3) Let S and T be distinct sides of C. On the contrary, suppose that x is in both S and T°. As S and T are distinct maximal convex subsets of ∂C, the side T is not contained in S. Hence, there is a point y of T not in S. By Theorem 6.2.4, we have that $\overline{C} \cap \langle S\rangle = S$, and so y is not in $\langle S\rangle$.

Assume first that $\dim T = 0$. Then x and y are antipodal. As S is not contained in T, it contains a point $z \neq x$. Let $S(x, z)$ be the unique great circle of S^n containing x and z. Then $S(x, z)$ also contains $y = -x$. As $S(x, z)$ is contained in $\langle S\rangle$, we find that y is also in $\langle S\rangle$, which is a contradiction.

Now assume that $\dim T > 0$. Then $T - S$ is an open subset of T by (2). Therefore, we may assume that y is not antipodal to x. Let L be the unique geodesic of X passing through x and y, and let P be the plane of X of dimension $1 + \dim S$ that contains $\langle S \rangle$ and L. As x is in T°, there is an $r > 0$ such that

$$B(x, r) \cap \langle T \rangle \subset T.$$

Observe that $B(x, r) \cap L$ is on both sides of $\langle S \rangle$ in P and

$$B(x, r) \cap L \subset B(x, r) \cap \langle T \rangle \subset T \subset \partial C.$$

Therefore, there are points of $\overline{C}$ on both sides of $\langle S \rangle$ in P contrary to Theorem 6.2.5. It follows that S and T° are disjoint. Thus S and T meet only along their boundaries. □

Exercise 6.2

1. Let C be a convex subset of X that is not a pair of antipodal points of S^n. Prove that C is connected.

2. Let C be a nonempty convex subset of S^n. Prove that C is a great m-sphere of S^n if and only if $-C = C$.

3. Let C be a convex subset of X that is not a closed great semicircle of S^n. Prove that C is geodesically convex if and only if C does not contain a pair of antipodal points.

4. Let C be a nonempty convex subset of X. Show that

   (1) $\overline{(C^\circ)} = \overline{C} = \overline{\overline{C}}$,

   (2) $\partial C^\circ = \partial C = \partial \overline{C}$,

   (3) $(C^\circ)^\circ = C^\circ = (\overline{C})^\circ$,

   (4) $\langle C^\circ \rangle = \langle C \rangle = \langle \overline{C} \rangle$,

   (5) $\dim C^\circ = \dim C = \dim \overline{C}$.

5. Let C be a proper, closed, convex subset of X. Prove that C is the intersection of all the closed half-spaces of X containing C.

6. Let C be a closed convex subset of S^n. Prove that C is contained in an open hemisphere of S^n if and only if C does not contain a pair of antipodal points.

7. Let C be a subset of S^n. Define $K(C)$ to be the union of all the geodesic rays in E^{n+1} from the origin passing through a point of C. Prove that C is a convex subset of S^n if and only if $K(C)$ is a convex subset of E^{n+1}.

8. Let C be a convex subset of S^n. Prove that a subset S of ∂C is a side of C if and only if $K(S)$ is a side of $K(C)$.

9. Let C be a bounded, n-dimensional, convex, proper subset of X. Prove that ∂C is homeomorphic to S^{n-1}.

§6.3. Convex Polyhedra

Throughout this section, $X = S^n, E^n$, or H^n with $n > 0$.

Definition: A *convex polyhedron* P in X is a nonempty, closed, convex subset of X such that the collection $\mathcal{S}$ of its sides is locally finite in X.

Remark: Locally finite in S^n is the same as finite, since S^n is compact; and every locally finite collection of subsets of E^n or H^n is countable, since E^n and H^n are finitely compact metric spaces.

Theorem 6.3.1. *Every side of an m-dimensional convex polyhedron P in X has dimension $m - 1$.*

Proof: We may assume that $m = n$. Let S be a side of P. Then there is a point x in S° by Theorem 6.2.3. Now as the collection of sides of P is locally finite, there is an $r > 0$ such that $B(x, r)$ meets only finitely many sides of P. By Theorem 6.2.6(3), the side S is the only side of P containing x. Hence, we may shrink $B(x, r)$ to avoid all the other sides of P, since the sides of P are closed. Consequently, we may assume that

$$B(x, r) \cap \partial P \subset S.$$

Moreover, we may assume that $r < \pi/2$. As x is in ∂P, the open ball $B(x, r)$ contains a point y of P° and a point z of $X - P$. Now y is not in $\langle S \rangle$ by Theorem 6.2.4. Let Q be the plane of X of dimension $1 + \dim S$ that contains y and $\langle S \rangle$. Since the geodesic segment $[y, z]$ is connected, it contains a point w of ∂P. As $[y, z] \subset B(x, r)$, the point w is in S. See Figure 6.3.1. Hence z is in Q. Consequently Q contains the nonempty open set $B(x, r) \cap (X - P)$. Therefore $Q = X$. Thus $\dim S = n - 1$. $\square$

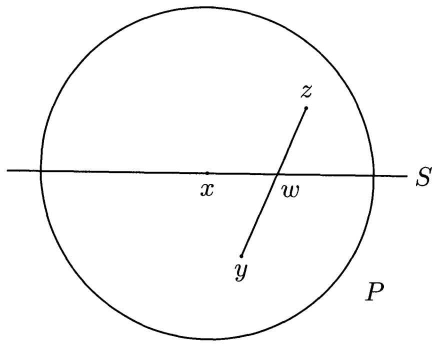


Figure 6.3.1. The four points w, x, y, z in the proof of Theorem 6.3.1

Theorem 6.3.2. *Let P be an n-dimensional convex polyhedron in X that is not all of X. For each side S of P, let H_S be the closed half-space of X such that $\partial H_S = \langle S \rangle$ and $P \subset H_S$. Then*

$$P = \cap\{H_S : S \text{ is a side of } P\}.$$

Proof: Let $K = \cap\{H_S : S \text{ is a side of } P\}$. Clearly, we have $P \subset K$. Let x be a point of $X - P$ and let y be a point of P° that is not antipodal to x. Then the segment $[x, y]$ contains a point z of ∂P, since $[x, y]$ is connected. Let S be a side of P that contains z. Then x and y are on opposite sides of the hyperplane $\langle S \rangle$. Hence y is not in H_S. Therefore $X - P \subset X - K$ and so $K \subset P$. Thus $P = K$. □

Theorem 6.3.3. *If x is a point in the boundary of a side S of a convex polyhedron P in X, then x is in the boundary of another side of P.*

Proof: We may assume that $\langle P \rangle = X$. On the contrary, suppose that x is not contained in any other side of P. Since the collection of sides of P is locally finite, there is an $r > 0$ such that $B(x, r)$ meets only finitely many sides of P. As S is the only side of P containing x, we can shrink $B(x, r)$ to avoid all the other sides of P, since the sides of P are closed. Therefore, we may assume that $B(x, r) \cap \partial P \subset S$. Moreover, we may assume that $r < \pi/2$. As x is in ∂P, the ball $B(x, r)$ contains a point y of P°. As x is in ∂S, the ball $B(x, r)$ contains a point z of $\langle S \rangle - S$. Now z is in $X - P$, since $P \cap \langle S \rangle = S$ by Theorem 6.2.4. Consequently, the geodesic segment $[y, z]$ contains a point w of ∂P. See Figure 6.3.2.

As $B(x, r) \cap \partial P \subset S$, the point w is in S. As z, w are in $\langle S \rangle$, we deduce that y is in $\langle S \rangle$, which is a contradiction, since $P \cap \langle S \rangle = S$. It follows that x is contained in some other side T of P; moreover, x must be in the boundary of T by Theorem 6.2.6(3). □

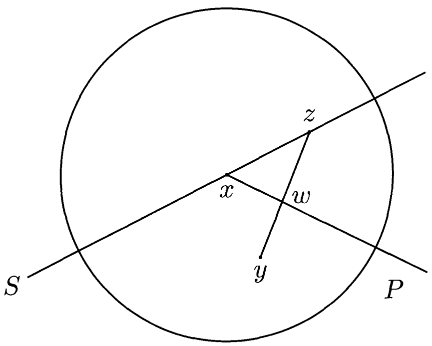


Figure 6.3.2. The four points w, x, y, z in the proof of Theorem 6.3.3

Theorem 6.3.4. *Let S and T be distinct sides of a convex subset C of X, and let x, y be a proper pair of points of C with x in S° and y in T°. Then the open geodesic segment (x, y) is contained in C°.*

Proof: Assume first that $[x, y]$ is contained in ∂C. Then $[x, y]$ is contained in a side R of C by Theorem 6.2.6(1). As R meets S° at x, we have that $R = S$ by Theorem 6.2.6(3). But R also meets T° at y, and so $R = T$, which is a contradiction. Therefore (x, y) contains a point z of C°. Furthermore, $(x, z]$ and $[z, y)$ are contained in C° by Theorem 6.2.2. Thus (x, y) is contained in C°. □

Theorem 6.3.5. *Every side of a convex polyhedron P in X is a convex polyhedron.*

Proof: Let S be a side of P. Then S is nonempty and convex by definition; moreover, S is closed by Theorem 6.2.6(2). Clearly S is a convex polyhedron if the dimension of S is either 0 or 1, so assume that $\dim S > 1$.

Let $\mathcal{R}$ be the collection of sides of S. We need to show that $\mathcal{R}$ is locally finite in X. Let x be a point of X. As the collection $\mathcal{S}$ of sides of P is locally finite, there is an $r > 0$ such that $B(x, r)$ meets only finitely many sides of P. We may assume that $r < \pi/2$. Let $\mathcal{R}_0$ be the collection of all the sides of S that meet $B(x, r)$. Suppose that R is in $\mathcal{R}_0$. Then $B(x, r)$ contains a point y of R°, since $B(x, r)$ is open. By Theorem 6.3.3, we can choose a side $f(R)$ of P other than S containing y.

We claim that the function $f : \mathcal{R}_0 \to \mathcal{S}$ is injective. On the contrary, let R_1 and R_2 be distinct sides of S in $\mathcal{R}_0$ such that $f(R_1) = f(R_2)$. Now $f(R_i)$ contains a point y_i of $R_i^\circ \cap B(x, r)$ for $i = 1, 2$. As $r < \pi/2$, we have that y_1 and y_2 are nonantipodal. By Theorem 6.3.4, the open geodesic segment (y_1, y_2) is contained in S°. But $[y_1, y_2]$ is contained in $f(R_i)$ because of the convexity of $f(R_i)$, which is a contradiction. Therefore f is injective.

As $B(x, r)$ meets only finitely many sides of P, the image of f is finite. Therefore $\mathcal{R}_0$ is finite. This shows that $\mathcal{R}$ is locally finite. Thus S is a convex polyhedron. □

Definition: A *ridge* of a convex polyhedron P is a side of a side of P.

Theorem 6.3.6. *If R is a ridge of a convex polyhedron P in X, then*

(1) *R° meets exactly two sides S_1 and S_2 of P;*

(2) *R is a side of both S_1 and S_2;*

(3) *$R = S_1 \cap S_2$.*

Proof: We may assume that $\langle P \rangle = X$. Let R be a side of a side S_1 of P. Choose a point x in R° and an $r > 0$ such that $B(x, r) \cap \langle R \rangle \subset R$. By Theorem 6.3.3, there is another side S_2 of P containing x in its boundary. By Theorem 6.3.1, both $\langle S_1 \rangle$ and $\langle S_2 \rangle$ are hyperplanes of X. Now by

Theorem 6.2.5, the convex set P is contained in one of the closed half-spaces of X bounded by $\langle S_2\rangle$. Hence, every diameter of $B(x,r)$ in R must lie in $\langle S_2\rangle$. Therefore $B(x,r)\cap R\subset\langle S_2\rangle$. Consequently, by Theorem 6.2.4, we have $B(x,r)\cap R\subset S_2$. Furthermore, by Theorem 6.2.6(3), we have $B(x,r)\cap R\subset\partial S_2$. Now by Theorem 6.2.6(1), the convex set $B(x,r)\cap R$ is contained in a side R_2 of S_2. Let $R_1=R$. Then by Theorems 6.3.1 and 6.3.5, both $\langle R_1\rangle$ and $\langle R_2\rangle$ have dimension $n-2$. As $B(x,r)\cap R_1\subset R_2$, we have that $\langle R_1\rangle=\langle R_2\rangle$. Now $\langle S_1\rangle\cap\langle S_2\rangle$ contains $\langle R\rangle$. Therefore

$$\dim\big(\langle S_1\rangle\cap\langle S_2\rangle\big)\geq n-2.$$

If the last equality were strict, then we would have $\langle S_1\rangle=\langle S_2\rangle$, which is not the case by Theorem 6.2.4. Therefore $\langle S_1\rangle\cap\langle S_2\rangle=\langle R\rangle$. Hence, for each i, we have

$$\begin{aligned} R_i &= S_i\cap\langle R\rangle\\ &= P\cap\langle S_i\rangle\cap\langle R\rangle\\ &= P\cap\langle S_1\rangle\cap\langle S_2\rangle \;=\; S_1\cap S_2.\end{aligned}$$

Thus $R_1=R_2$. Therefore R is a side of S_1 and S_2, and $R=S_1\cap S_2$.

Next, assume that R° meets a third side S_3 of P. Then the same argument as above shows that R is a side of S_3 and $R=S_1\cap S_3$. Furthermore $\langle S_3\rangle$ is also a hyperplane of X. Now the set $X-\langle S_1\rangle\cup\langle S_2\rangle$ has four components C_1,C_2,C_3,C_4, one of which, say C_1, contains P° by Theorem 6.2.5. Moreover P is contained in $\overline{C}_1$. As S_3 is in $\overline{C}_1$, the hyperplane $\langle S_3\rangle$ divides C_1 into two parts, that is, $C-\langle S_3\rangle$ has two components C_{11} and C_{12}. See Figure 6.3.3. Now by Theorem 6.2.5, we have that P° is contained in both C_{11} and C_{12}, which is a contradiction. Therefore R° meets exactly two sides of P. □

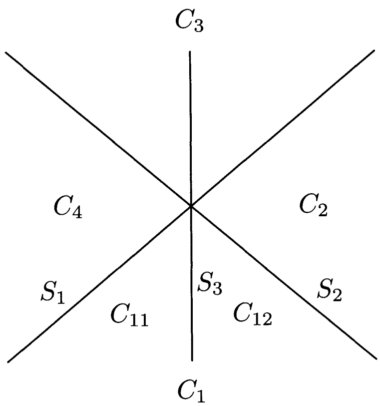


Figure 6.3.3. The subdivision of E^2 by three concurrent lines

Theorem 6.3.7. *An m-dimensional convex polyhedron P in E^n or H^n, with $m > 0$, is compact if and only if*

(1) *the polyhedron P has at least $m+1$ sides;*

(2) *the polyhedron P has only finitely many sides; and*

(3) *each side of P is compact.*

Proof: We may assume that $m = n$. The proof is by induction on n. The theorem is obviously true when $n = 1$, so assume that $n > 1$ and the theorem is true for $n - 1$. Let $Y = E^n$ or H^n.

Now suppose that P is compact. Then ∂P is nonempty; otherwise P would be Y, which is not the case. Therefore P has at least one side S by Theorem 6.2.6(1). Now S is an $(n-1)$-dimensional convex polyhedron by Theorems 6.3.1 and 6.3.5; moreover, S is compact, since S is a closed subset of P. Therefore S has at least n sides $R_1, \ldots, R_n$ by the induction hypothesis. By Theorem 6.3.6, each R_i is the side of another side S_i of P; moreover, the sides $S_1, \ldots, S_n$ are distinct, since $S \cap S_i = R_i$. Therefore P has at least $n+1$ sides.

Now, for each x in P, there is a $r(x) > 0$ such that $B(x, r(x))$ meets only finitely many sides of P. As P is compact, there is a finite subset $\{x_1, \ldots, x_k\}$ of P such that P is covered by the union of $B(x_i, r(x_i))$, for $i = 1, \ldots, k$. Therefore P has only finitely many sides; moreover, each side of P is compact, since each side of P is a closed subset of P.

Conversely, suppose that P satisfies properties (1), (2), (3). By Theorem 6.2.6(1), the boundary of P is the union of all the sides of P. Therefore ∂P is compact. Let x be a point in P°. Then there is an $r > 0$ such that $B(x, r)$ contains ∂P, since ∂P is bounded. Let y be a point on ∂P and let z be the endpoint of the radius of $B(x, r)$ passing through y. Then z is not in P because of Theorem 6.2.3. Therefore, the set $S(x, r) - P$ is nonempty. As the sphere $S(x, r)$ is connected for $n > 1$, the set $S(x, r) \cap P^\circ$ is empty. Hence $S(x, r)$ is contained in $Y - P$. As P is connected, $P \subset B(x, r)$. Thus P is bounded and so is compact. This completes the induction. □

Theorem 6.3.8. *Let P be an m-dimensional convex polyhedron in S^n, with $m > 0$. Then the following are equivalent:*

(1) *P is contained in an open hemisphere of S^n;*

(2) *P has at least $m+1$ sides and each side S of P is contained in an open hemisphere of $\langle S \rangle$;*

(3) *P has a side S that is contained in an open hemisphere of $\langle S \rangle$.*

Proof: Suppose that P is contained in an open hemisphere H of S^n. We may assume that H is the upper hemisphere of S^n. Then by gnomonic projection, we can view P as a compact convex polyhedron of E^n. Then P has at least $m+1$ sides by Theorem 6.3.7. If S is a side of P, then S is contained in the open hemisphere $H \cap \langle S \rangle$. Thus (1) implies (2). Clearly (2) implies (3).

Suppose that P has a side S that is contained in an open hemisphere of $\langle S \rangle$. On the contrary, assume that P is not contained in an open hemisphere of S^n. We may assume that $m = n$, $\langle S \rangle = S^{n-1}$, and P is contained in the closed southern hemisphere S^n_- of S^n. Then $\mathrm{dist}(e_n, P) = \pi/2$. Let y be a point of S^n. Then $\mathrm{dist}(y, P) \leq \pi/2$; otherwise P would be contained in the open hemisphere opposite y. Hence, there is a point x of P such that $\theta(x, y) \leq \pi/2$. Now assume that $y \neq \pm e_n$. Then x is in the n-dimensional lune $S^n_- \cap C(y, \pi/2)$. Consequently, there is a sequence of points $\{y_i\}$ of $[e_n, y]$ converging to e_n and a sequence of points $\{x_i\}$ of P such that x_i is in $S^n_- \cap C(y_i, \pi/2)$ for each i. As P is compact, the sequence $\{x_i\}$ has a limit point x_0 in $P \cap S^{n-1} = S$ that is contained in the closed hemisphere of S^{n-1} whose center is the intersection of the great circle through e_n and y with S^{n-1}. Thus S has the property that every closed hemisphere of $\langle S \rangle$ contains a point of S, which is a contradiction. Thus (3) implies (1). □

Faces of a Convex Polyhedron

Let P be an m-dimensional convex polyhedron in X. We now define a *k-face* of P for each $k = 0, 1, \ldots, m$ inductively as follows: The only m-face of P is P itself. Suppose that all the $(k+1)$-faces of P have been defined and each is a $(k+1)$-dimensional convex polyhedron in X. Then a k-face of P is a side of a $(k+1)$-face of P. By Theorems 6.3.1 and 6.3.5, a k-face of P is a k-dimensional convex polyhedron in X. A *proper face* of P is a k-face of P with $k < m$. Note that a proper face of P is just a side of a side ... of a side of P. Therefore, a face E of a face F of P is a face of P. In other words, the face relation is transitive.

Theorem 6.3.9. *If C is a convex subset of a convex polyhedron P in X such that C° meets a face E of P, then $C \subset E$.*

Proof: Let $m = \dim P$ and $k = \dim E$. The proof is by induction on $m - k$. This is certainly true if $k = m$, so assume that $k < m$ and the theorem is true for all $(k+1)$-faces of P. Now E is a side of a $(k+1)$-face F of P. By the induction hypothesis $C \subset F$. Let x be a point of $C^\circ \cap E$. Choose $r > 0$ so that

$$B(x, r) \cap \langle C \rangle \subset C.$$

By Theorem 6.2.5, the convex set F is contained in one of the closed half-spaces of $\langle F \rangle$ bounded by $\langle E \rangle$. Hence, every diameter of $B(x, r)$ in C must lie in $\langle E \rangle$. Therefore

$$B(x, r) \cap \langle C \rangle \subset \langle E \rangle.$$

Hence $\langle C \rangle \subset \langle E \rangle$. Therefore

$$C \subset F \cap \langle E \rangle = E.$$

□

Theorem 6.3.10. *The interiors of all the faces of a convex polyhedron P in X form a partition of P.*

Proof: Let $m = \dim P$. We first prove that P is the union of the interiors of all its faces by induction on m. This is certainly true if $m = 0$, so assume that $m > 0$ and any $(m-1)$-dimensional convex polyhedron in X is the union of the interiors of all its faces. Then each side of P is the union of the interiors of all its faces. As P is the union of ∂P and P°, we have that P is the union of the interiors of all its faces.

Now suppose that E and F are faces such that E° meets F°. Then $E \subset F$ and $F \subset E$ by Theorem 6.3.9. Hence $E = F$. Thus, the interiors of all the faces of P form a partition of P. □

Theorem 6.3.11. *If E and F are faces of a convex polyhedron P in X such that $E \subset F$, then E is a face of F.*

Proof: Let x be a point of E°. Then there is a face G of F such that x is in G° by Theorem 6.3.10. Now $E \subset G$ and $G \subset E$ by Theorem 6.3.9. Therefore $E = G$. Thus E is a face of F. □

Theorem 6.3.12. *The family of all the faces of a convex polyhedron P in X is locally finite.*

Proof: Let $m = \dim P$. The proof is by induction on m. This is certainly true if $m = 0$, so assume that $m > 0$ and the theorem is true for all $(m-1)$-dimensional polyhedra in X. Let x be a point of X. Then there is an $r_0 > 0$ such that $B(x, r_0)$ meets only finitely many sides of P, say $S_1, \ldots, S_k$. By the induction hypothesis, the family of all faces of S_i is locally finite in X for each $i = 1, \ldots, k$. Hence, there is an $r_i > 0$ such that $B(x, r_i)$ meets only finitely many faces of S_i for each $i = 1, \ldots, k$. Let

$$r = \min\{r_0, \ldots, r_k\}.$$

Then $B(x, r)$ meets only finitely many faces of P. □

Theorem 6.3.13. *If E is a k-face of an m-dimensional convex polyhedron P in X, then*

(1) *E is a side of every $(k+1)$-face of P that meets E°;*

(2) *E is a side of only finitely many $(k+1)$-faces of P;*

(3) *E is a side of at least $m-k$ $(k+1)$-faces of P;*

(4) *E is the intersection of any two $(k+1)$-faces of P that meet E°.*

Proof: (1) Suppose that F is a $(k+1)$-face of P that meets E°. Then $E \subset F$ by Theorem 6.3.9, moreover, E is a side of F by Theorem 6.3.11.

(2) Let x be a point of E°. Then there is an $r > 0$ such that $B(x, r)$ meets only finitely many $(k+1)$-faces of P by Theorem 6.3.12. Hence E is a side of only finitely many $(k+1)$-faces of P.

(3) We now prove that E is a side of at least $m-k$ $(k+1)$-faces of P by induction on $m-k$. This is certainly true if $k=m$, so assume that $k<m$ and the theorem is true for all $(k+1)$-faces of P. Now E is a side of a $(k+1)$-face F of P. By the induction hypothesis, F is a side of $m-k-1$ $(k+2)$-faces of P, say $G_1, \ldots, G_{m-k-1}$. By Theorem 6.3.6, we have that E is a side of exactly two sides F and F_i of G_i for each $i=1,\ldots,m-k-1$. Suppose that $i \neq j$. As $F \subset G_i \cap G_j$, we have

$$\dim(G_i \cap G_j) = k+1.$$

Therefore, we have

$$F^\circ \subset (G_i \cap G_j)^\circ.$$

By Theorem 6.3.9, we have that $G_i \cap G_j \subset F$. Thus $F = G_i \cap G_j$. Consequently $F_i \neq F_j$. Thus, the $m-k$ $(k+1)$-faces F, $F_1, \ldots, F_{m-k-1}$ are distinct.

(4) Let F_1 and F_2 be distinct $(k+1)$-faces of P that meet E°. Then $E \subset F_1 \cap F_2$ by (1). Hence

$$\dim(F_1 \cap F_2) = k.$$

Therefore, we have

$$E^\circ \subset (F_1 \cap F_2)^\circ.$$

By Theorem 6.3.9, we have that $F_1 \cap F_2 \subset E$. Thus $E = F_1 \cap F_2$. □

Theorem 6.3.14. *If E is a k-face of an m-dimensional convex polyhedron P in X, then*

(1) *E is a face of every side of P that meets E°;*

(2) *E is a face of only finitely many sides of P;*

(3) *E is a face of at least $m-k$ sides of P;*

(4) *E is the intersection of all the sides of P that meet E° or $E = P$.*

Proof: (1) Let S be a side of P that meets E°. Then $E \subset S$ by Theorem 6.3.9; moreover E is a face of S by Theorem 6.3.11. (2) Let x be a point of E°. Then there is an $r > 0$ such that $B(x,r)$ meets only finitely many sides of P. Hence E is a face of only finitely many sides of P.

We now prove (3) and (4) by induction on $m-k$. This is certainly true if $k = m-1$ or m, so assume that $k < m-1$ and the theorem is true for all $(k+1)$-faces of P. By Theorem 6.3.13, we have that E is a side of finitely many $(k+1)$-faces of P, say $F_1, \ldots, F_\ell$ with $\ell \geq m-k$. By the induction hypothesis and (2), we have that F_i is a face of only finitely many sides of P, say $S_{i1}, \ldots, S_{i\ell_i}$, and $\ell_i \geq m-k-1$ for each i and

$$F_i = \bigcap_{j=1}^{\ell_i} S_{ij}.$$

Now the sets $\{S_{1j}\}$ and $\{S_{2j}\}$ are not the same, since F_1 and F_2 are distinct $(k+1)$-faces of P. Hence, one of the sides in one of the sets is not in the other set. Therefore E is a face of at least $m-k$ sides of P. Clearly

$$\{S_{ij} : j = 1, \ldots, \ell_i \text{ and } i = 1, \ldots, \ell\}$$

is the set of all the sides of P that meet E°. By Theorem 6.3.13, we have that $F_i \cap F_j = E$ for all i, j such that $i \neq j$. Hence

$$E = \bigcap_{i=1}^{\ell} E_i = \bigcap_{i=1}^{\ell} \bigcap_{j=1}^{\ell_i} S_{ij}.$$

Thus E is the intersection of all the sides of P that meet E°. □

Theorem 6.3.15. *Every nonempty intersection of sides of a convex polyhedron P in X is a face of P.*

Proof: Let C be a nonempty intersection of sides of P. Then C° contains a point x by Theorem 6.2.3. The point x is in ∂P, since $C \subset \partial P$. By Theorem 6.3.10, there is a face E of P such that x is in E°. Then $C \subset E$ by Theorem 6.3.9. Now E is the intersection of all the sides of P that meet E° by Theorem 6.3.14. Therefore $E \subset C$. Thus $E = C$. □

Theorem 6.3.16. *Let P be an m-dimensional convex polyhedron in S^n. Then either*

(1) *the polyhedron P is a great m-sphere of S^n; or*

(2) *the intersection of all the sides of P is a great k-sphere of S^n; or*

(3) *the polyhedron P is contained in an open hemisphere of S^n.*

Proof: The proof is by induction on m. The theorem is certainly true for $m = 0$, so assume that $m > 0$ and the theorem is true for all $(m-1)$-dimensional convex polyhedra in S^n. If P has no sides, then (1) holds. Hence, we may assume that P has a side S.

Now assume that S is a great $(m-1)$-sphere of S^n. Then P is a closed hemisphere of $\langle P \rangle$, since a point of P° can be joined to any point of S by a geodesic arc. Therefore (2) holds. Thus, we may assume that no side of P is a great $(m-1)$-sphere of S^n.

If S is contained in an open hemisphere of $\langle S \rangle$, then (3) holds by Theorem 6.3.8. Hence, we may assume that no side of P is contained in an open hemisphere. By the induction hypothesis, the intersection of all the sides of a side of P is a great k-sphere of S^n.

We may assume that $m = n$, $\langle S \rangle = S^{n-1}$, and $P \subset S^n_+$. Let T_0 be the intersection of all the sides of a side T of P. Then T_0 is a great k-sphere of S^n. As $T_0 \subset S^n_+$, we must have

$$T_0 \subset P \cap S^{n-1} = S.$$

Now T_0 is a face of T by Theorem 6.3.15, and so T_0 is a face of P. Therefore T_0 is a face of S by Theorem 6.3.11. Now T_0 is the intersection of all the sides of S that meet T_0 by Theorem 6.3.14. Let S_0 be the intersection of all the sides of S. Then $S_0 \subset T_0 \subset T$. Let P_0 be the intersection of all the sides of P. Then $S_0 \subset P_0$. Now S_0 is a face of S by Theorem 6.3.15, and so S_0 is a face of P. Therefore S_0 is the intersection of all the sides of P that meet S_0 by Theorem 6.3.14. Hence $P_0 \subset S_0$. Thus $P_0 = S_0$. Hence P_0 is a great k-sphere of S^n. Thus (2) holds. This completes the induction. □

Vertices of a Convex Polyhedron

A 0-face of a convex polyhedron P in X consists either of a single point or a pair of antipodal points.

Definition: A *vertex* of a polyhedron P is a point in a 0-face of P.

Definition: The *convex hull* of a subset S of X is the intersection of all the convex subsets of X containing S.

Theorem 6.3.17. *A convex polyhedron P in E^n or H^n is compact if and only if P has only finitely many vertices and P is the convex hull of its vertices.*

Proof: Assume first that P is in E^n. The proof is by induction of the dimension m of P. The theorem is certainly true when $m = 0$, so assume that $m > 0$ and the theorem is true in dimension $m - 1$. Suppose that P is compact. Then by Theorem 6.3.7, the polyhedron P has only finitely many sides and each side is compact. By the induction hypothesis, each side of P has only finitely many vertices and is the convex hull of its vertices. Therefore P has only finitely many vertices. Let V be the set of vertices of P. Then the convex hull $C(V)$ is contained in P, since P is convex. Let x be a point of P. We claim that x is in $C(V)$. If x is in a side S of P, then x is a convex combination of the vertices of S by the induction hypothesis. Hence, we may assume that x is in P°. Let v_0 be a vertex of P. Then the ray from v_0 passing through x meets ∂P in a point y other than v_0, since P is bounded. By Theorem 6.2.2, the point x lies between v_0 and y. Hence, there is a real number t between 0 and 1 such that

$$x = (1 - t)v_0 + ty.$$

Let S be a side of P containing y. By the induction hypothesis, there are vertices $v_1, \ldots, v_k$ of S and positive real numbers $t_1, \ldots, t_k$ such that

$$y = \sum_{i=1}^{k} t_i v_i \quad \text{and} \quad \sum_{i=1}^{k} t_i = 1.$$

Observe that

$$x = (1 - t)v_0 + t\sum_{i=1}^{k} t_i v_i$$

is a convex combination of $v_0, \ldots, v_k$. Hence x is in $C(V)$. Therefore $P = C(V)$.

Conversely, suppose that P has only finitely many vertices and P is the convex hull of its vertices. Let $r > 0$ be such that the ball $B(0, r)$ contains the set V of vertices of P. Then $B(0, r)$ contains the convex hull $C(V)$, since $B(0, r)$ is convex. Hence P is bounded and so P is compact. This completes the induction.

Now assume that P is H^n. We pass to the projective disk model D^n. If P is compact, then P is a Euclidean polyhedron, and so P has only finitely many vertices and P is the convex hull of its vertices by the Euclidean case. Conversely, suppose that P has only finitely many vertices and P is the convex hull of its vertices. Then P is compact by the same argument as in the Euclidean case. □

Theorem 6.3.18. *An m-dimensional convex polyhedron P in S^n, with $m > 0$, is contained in an open hemisphere of S^n if and only if P is the convex hull of its vertices.*

Proof: Suppose that P is contained in an open hemisphere of S^n. We may assume that P is contained in the open northern hemisphere of S^n. Now by gnomonic projection, we can view P as a compact polyhedron in E^n. Then P is the convex hull of its vertices by Theorem 6.3.17.

Conversely, suppose that P is the convex hull of its vertices. On the contrary, suppose that P is not contained in an open hemisphere of S^n. Then the intersection P_0 of all the sides of P is a great k-sphere of S^n by Theorem 6.3.16. Now P_0 is contained in every 0-face of P, since a 0-face of P is the intersection of all the sides of P containing it by Theorem 6.3.14. Therefore $\dim P_0 = 0$, and so P_0 is a pair of antipodal points. Hence P has only two vertices. Therefore, the convex hull of the vertices of P is P_0, which is a contradiction, since $m > 0$. □

Links of a Convex Polyhedron

Let x be a point of a convex polyhedron P in X. Then there is a real number r such that $0 < r < \pi/2$ and r is less than the distance from x to any side of P not containing x, since the set of sides of P is locally finite. The set

$$L(x) = P \cap S(x, r)$$

is called a *link* of x in the polyhedron P. The spherical geometry of the link $L(x)$ is uniquely determined by x up to a change of scale induced by radial projection from x.

For simplicity, we have only considered spherical polyhedra in S^n. By a simple change of scale, the theory of spherical polyhedra in S^n generalizes to polyhedra in any sphere of X.

Theorem 6.3.19. *Let x be a point of an m-dimensional convex polyhedron P in X, with $m > 0$, and let r be a real number such that $0 < r < \pi/2$ and r is less than the distance from x to any side of P not containing x. Then the link*

$$L(x) = P \cap S(x, r)$$

of x in P is an $(m-1)$-dimensional convex polyhedron in the sphere $S(x,r)$. Moreover, if $\mathcal{S}(x)$ is the set of sides of P containing x, then

$$\{T \cap S(x,r) : T \in \mathcal{S}(x)\}$$

is the set of sides of $L(x)$.

Proof: The proof is by induction on m. The theorem is obviously true for $m = 1$, so assume that $m > 1$ and the theorem is true for all $(m-1)$-dimensional convex polyhedra in X. We may assume that $m = n$. If x is in P°, then $L(x) = S(x,r)$, so assume that x is in ∂P. Let $\mathcal{S}$ be the set of sides of P. For each T in $\mathcal{S}$, let H_T be the closed half-space of X bounded by the hyperplane $\langle T \rangle$ and containing P. Then we have

$$P = \bigcap_{T \in \mathcal{S}} H_T.$$

Now as $H_T \cap S(x,r) = S(x,r)$ for each T not containing x, we have

$$P \cap S(x,r) = \bigcap_{T \in \mathcal{S}(x)} \big(H_T \cap S(x,r)\big).$$

Now $H_T \cap S(x,r)$ is a closed hemisphere of $S(x,r)$ for each T in $\mathcal{S}(x)$. Therefore $L(x)$ is a closed convex subset of $S(x,r)$.

Let y be a point of P° such that y is not antipodal to x. By shrinking r, if necessary, we may assume that $d(x,y) \geq r$. Then the geodesic segment $[x,y]$ intersects $S(x,r)$ in a point z of P° by Theorem 6.2.2. Therefore $P^\circ \cap S(x,r)$ is a nonempty open subset of $S(x,r)$ contained in $L(x)$. Hence $\dim L(x) = n - 1$.

Now as

$$P^\circ \cap S(x,r) \subset L(x)^\circ,$$

we have that

$$\partial L(x) \subset \partial P \cap S(x,r).$$

Let T be a side of P containing x. By the induction hypothesis, $T \cap S(x,r)$ is an $(n-2)$-dimensional convex polyhedron in $S(x,r)$. Now since $P \subset H_T$, no point of $T \cap S(x,r)$ has an open neighborhood in $S(x,r)$ contained in $L(x)$. Therefore

$$T \cap S(x,r) \subset \partial L(x).$$

Hence, we have

$$\partial P \cap S(x,r) \subset \partial L(x).$$

Therefore, we have

$$\partial L(x) = \partial P \cap S(x,r).$$

The convex set $T \cap S(x,r)$ is contained in a side $\hat{T}$ of $L(x)$ by Theorem 6.2.6(1). Now as

$$\partial P \cap S(x,r) = \bigcup_{T\in\mathcal{S}(x)} \big(T \cap S(x,r)\big),$$

we have that

$$\partial L(x) = \bigcup_{T\in\mathcal{S}(x)} \hat{T}.$$

Therefore $\{\hat{T} : T \in \mathcal{S}(x)\}$ is the set of sides of $L(x)$ by Theorem 6.2.6(3). Hence $L(x)$ has only finitely many sides. Thus $L(x)$ is a convex polyhedron in $S(x,r)$.

Now by Theorem 6.2.6(3), we have

$$\hat{T}^\circ \subset T \cap S(x,r).$$

Therefore $\hat{T} = T \cap S(x,r)$ for each T in $\mathcal{S}(x)$. Thus

$$\{T \cap S(x,r) : T \in \mathcal{S}(x)\}$$

is the set of sides of $L(x)$. □

Theorem 6.3.20. *Let P be a convex polyhedron in D^n. Then its closure $\overline{P}$ in E^n is a convex subset of E^n such that $\overline{P} \cap D^n = P$ and*

$$\partial(\overline{P}) = \partial P \cup (\overline{P} \cap S^{n-1}).$$

Moreover, if S is a side of P, then its closure $\overline{S}$ in E^n is a side of $\overline{P}$, and if u is a point of $\partial(\overline{P})$ that is not in the Euclidean closure of a side of P, then $\{u\}$ is a side of $\overline{P}$.

Proof: We may assume that $\langle P\rangle = D^n$. As P is a convex subset of E^n, we have that $\overline{P}$ is a convex subset of E^n by Theorem 6.2.1. As D^n is open in E^n and P is closed in D^n, we have

$$\overline{P} \cap D^n = P, \qquad P^\circ \subset (\overline{P})^\circ, \quad \text{and} \quad \partial P \subset \partial(\overline{P}).$$

Clearly, we have that

$$\overline{P} \cap S^{n-1} \subset \partial(\overline{P}).$$

Therefore, we deduce that $P^\circ = (\overline{P})^\circ$ and

$$\partial(\overline{P}) = \partial P \cup (\overline{P} \cap S^{n-1}).$$

Let S be a side of P. Then S is contained in a side $\hat{S}$ of $\overline{P}$. Now $\hat{S} \cap D^n$ is a convex subset of ∂P containing S. Therefore $\hat{S} \cap D^n = S$. Clearly, we have that

$$\hat{S} \cap S^{n-1} \subset \partial(\hat{S}).$$

Therefore $\hat{S}^\circ \subset S$, and so $\hat{S} = \overline{S}$ by Theorem 6.2.2.

Let u be a point of $\partial(\overline{P})$ that is not in the closure of a side of P. Let U be a side of $\partial(\overline{P})$ containing u. Then U is not the closure of a side of P. Hence U° is disjoint from ∂P, and so $U^\circ \subset S^{n-1}$. Therefore $U = \{u\}$. □

Define $\kappa : D^n \to B^n$ by $\kappa = \zeta^{-1}\mu$. Then κ is an isometry from D^n to B^n. Observe that

$$\begin{aligned}\zeta^{-1}\mu(x) &= \zeta^{-1}\left(\frac{x+e_{n+1}}{|||x+e_{n+1}|||}\right)\\ &= \frac{x}{|||x+e_{n+1}|||}\,\frac{1}{(1+|||x+e_{n+1}|||^{-1})}\\ &= \frac{x}{|||x+e_{n+1}|||+1}.\end{aligned}$$

Hence, we have

$$\kappa(x) = \frac{x}{1+\sqrt{1-|x|^2}}. \tag{6.3.1}$$

The inverse of κ is given by

$$\kappa^{-1}(y) = \frac{2y}{1+|y|^2}. \tag{6.3.2}$$

Observe that κ extends to a homeomorphism

$$\overline{\kappa} : \overline{D}^n \to \overline{B}^n,$$

which is the identity on S^{n-1}.

Definition: An *ideal point* of a convex polyhedron P in B^n is a point u of $\overline{P} \cap S^{n-1}$, where $\overline{P}$ is the closure of P in E^n.

Theorem 6.3.21. *Let u be an ideal point of a convex polyhedron P in B^n. Then for each point x of P, there is a geodesic ray $[x, u)$ in P starting at x and ending at u.*

Proof: Since the isometry $\kappa : D^n \to B^n$ extends to a homeomorphism $\overline{\kappa} : \overline{D}^n \to \overline{B}^n$, we can pass to the projective disk model D^n of hyperbolic space. Let x be a point of P. Now $\overline{P}$ is a convex subset of E^n by Theorem 6.3.20. Therefore, the line segment $[x, u]$ is in $\overline{P}$. Now since

$$[x,u] \cap S^{n-1} = \{u\} \quad \text{and} \quad \overline{P} \cap D^n = P,$$

we have that $[x, u) \subset P$. □

Definition: A side S of a convex polyhedron P in B^n is *incident* with an ideal point u of P if and only if u is in the closure of S in E^n.

Theorem 6.3.22. *Let ∞ be an ideal point of a convex polyhedron P in U^n. Then a side S of P is incident with ∞ if and only S is vertical.*

Proof: Every hemispherical side of P is bounded in E^n. Therefore, if a side S of P is incident with ∞, then S must be vertical.

Conversely, suppose that S is a vertical side of P. Let x be a point of S. By Theorem 6.3.21, there is a geodesic ray $[x, \infty)$ in P starting at x and ending at ∞. Now since $[x, \infty)$ and $\langle S \rangle$ are vertical, we deduce that

$$[x, \infty) \subset \langle S \rangle \cap P = S.$$

Therefore S is incident with ∞. □

Definition: A *horopoint* of a convex polyhedron P in B^n is an ideal point u of P for which there is a horosphere Σ of B^n based at u such that Σ meets only the sides of P incident with u.

Note that if P is finite-sided, then every ideal point of P is a horopoint.

Example: Let P be a convex polyhedron in U^n all of whose sides are hemispherical hyperplanes of U^n such that P is the closed region above them. Then ∞ is an ideal point of P, and ∞ is a horopoint of P if and only if the set of radii of the sides is bounded.

Let u be a horopoint of a convex polyhedron P in B^n. Then there is a horosphere Σ of B^n based at u such that Σ meets only the sides of P incident with u. The set

$$L(u) = P \cap \Sigma$$

is called a *link* of u in the polyhedron P. The Euclidean geometry of the link $L(u)$ is uniquely determined by u up to a similarity induced by radial projection from u.

Theorem 6.3.23. *Let u be a horopoint of an m-dimensional convex polyhedron P in B^n, and let Σ be a horosphere of B^n based at u such that Σ meets only the sides of P incident with u. Then the link*

$$L(u) = P \cap \Sigma$$

of u in P is an $(m-1)$-dimensional convex polyhedron in the horosphere Σ. Moreover, if $\mathcal{S}(u)$ is the set of sides of P incident with u, then

$$\{S \cap \Sigma : S \in \mathcal{S}(u)\}$$

is the set of sides of $L(u)$.

Proof: We pass to the upper half-space model U^n of hyperbolic space. We may assume that $u = \infty$. The proof is by induction on m. The theorem is obviously true for $m = 1$, so assume that $m > 1$ and the theorem is true for all $(m-1)$-dimensional convex polyhedra in U^n. We may assume that $m = n$. By Theorem 6.3.22, a side of P is incident with ∞ if and only if it is vertical. If P has no vertical sides, then $L(u) = \Sigma$, so assume that P has a vertical side. Let $\mathcal{S}$ be the set of sides of P. For each S in $\mathcal{S}$, let H_S be the closed half-space of U^n bounded by the hyperplane $\langle S\rangle$ and containing P. Then we have

$$P = \bigcap_{S\in\mathcal{S}} H_S.$$

Now as $H_S \cap \Sigma = \Sigma$ for each hemispherical side S of P, we have

$$P\cap\Sigma = \bigcap_{T\in\mathcal{S}(u)} (H_S\cap\Sigma).$$

Now $H_S \cap \Sigma$ is a closed half-space of Σ for each S in $\mathcal{S}(u)$. Therefore $L(u)$ is a closed convex subset of Σ.

Let x be a point of P°. By shrinking Σ, if necessary, we may assume that x is not inside of Σ. Then the geodesic ray $[x,\infty)$ intersects Σ in a point y of P° by Theorem 6.2.2 applied to the Euclidean closure of P in the projective disk model. Therefore $P^\circ\cap\Sigma$ is a nonempty open subset of Σ contained in $L(u)$. Hence $\dim L(u) = n-1$.

Now as

$$P^\circ\cap\Sigma \subset L(u)^\circ,$$

we have that

$$\partial L(u) \subset \partial P\cap\Sigma.$$

Let S be a vertical side of P. By the induction hypothesis, $S\cap\Sigma$ is an $(n-2)$-dimensional convex polyhedron in Σ. Now since $P\subset H_S$, no point of $S\cap\Sigma$ has an open neighborhood in Σ contained in $L(u)$. Therefore

$$S\cap\Sigma\subset\partial L(u).$$

Hence, we have

$$\partial P\cap\Sigma\subset\partial L(u).$$

Therefore, we have

$$\partial L(u) = \partial P\cap\Sigma.$$

The convex set $S\cap\Sigma$ is contained in a side $\hat{S}$ of $L(u)$ by Theorem 6.2.6(1). Now as

$$\partial P\cap\Sigma = \bigcup_{S\in\mathcal{S}(u)} (S\cap\Sigma),$$

we have that

$$\partial L(u) = \bigcup_{S\in\mathcal{S}(u)} \hat{S}.$$

Therefore $\{\hat{S} : S\in\mathcal{S}(u)\}$ is the set of sides of $L(u)$ by Theorem 6.2.6(3).

Now by Theorem 6.2.6(3), we have

$$\hat{S}^\circ\subset S\cap\Sigma.$$

Therefore $\hat{S} = S\cap\Sigma$ for each S in $\mathcal{S}(u)$. Thus

$$\{S\cap\Sigma : S\in\mathcal{S}(u)\}$$

is the set of sides of $L(u)$. Moreover, the set of sides of $L(u)$ is locally finite in Σ, since the set of sides of P is locally finite in U^n. Thus $L(u)$ is a convex polyhedron in Σ. □

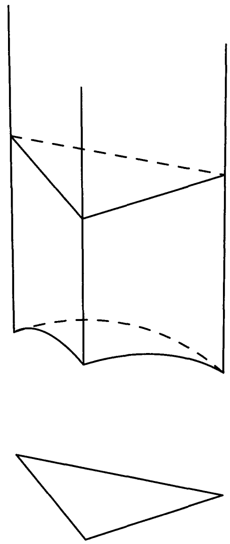


Figure 6.3.4. The link of ∞ in a polyhedron in U^3

There is a nice way of representing the link of a horopoint u of a polyhedron P in U^n. If we position P so that $u = \infty$, then the vertical projection

$$\nu : U^n \to E^{n-1}$$

projects $L(u)$ onto a similar polyhedron in E^{n-1} that does not depend on the choice of the horosphere Σ of U^n such that $L(u) = P \cap \Sigma$. See Figure 6.3.4.

Definition: An *ideal vertex* of a convex polyhedron P in B^n is a horopoint of P whose link is compact.

For example, the polyhedron in Figure 6.3.4 has an ideal vertex at ∞.

Theorem 6.3.24. *Let P be a convex polyhedron in D^n. Then its closure $\overline{P}$ in E^n is a convex polyhedron in E^n if and only if every ideal point of P is an ideal vertex of P.*

Proof: Let $m = \dim P$. We may assume that $m > 0$. Suppose that $\overline{P}$ is a convex polyhedron in E^n. Let u be an ideal point of P. We claim that u is a vertex of $\overline{P}$. On the contrary, suppose that u is not a vertex of $\overline{P}$. Then u is in the interior of a k-face F of $\overline{P}$ for some $k > 0$ by Theorem 6.3.10. Hence, there is an open Euclidean line segment in F containing u. But any such line segment cannot lie entirely in $\overline{D}^n$, since u is in S^{n-1}. Thus, we have a contradiction, and so u must be a vertex of $\overline{P}$.

If $m = 1$, the sides of $\overline{P}$ are the two endpoints of $\overline{P}$. If $m > 1$, the sides of $\overline{P}$ are the closures of the sides of P by Theorem 6.3.20. As $\overline{P}$ is compact, $\overline{P}$ has only finitely many sides. Therefore P has only finitely many sides. Let u be an ideal point of P. Then u is a horopoint of P. Let Σ be a horosphere of D^n based at u such that Σ meets only the sides of P incident with u. We claim that $P \cap \Sigma$ is compact. The proof is by induction on m. This is certainly true if $m = 1$, so assume that $m > 1$ and the claim is true for all $(m-1)$-dimensional convex polyhedra in D^n. Now the vertex u of $\overline{P}$ meets at least m sides of $\overline{P}$ by Theorem 6.3.14. Therefore $P \cap \Sigma$ has at least m sides by Theorem 6.2.23. If S is a side of P incident with u, then $S \cap \Sigma$ is compact by the induction hypothesis. Therefore $P \cap \Sigma$ is compact by Theorem 6.3.7. Thus u is an ideal vertex of P.

Conversely, suppose that every ideal point of P is an ideal vertex. We may assume that $m > 1$. Then every ideal point of P is in the closure of a side of P. Hence $\overline{P}$ is a closed convex subset of E^n whose sides are the closures of the sides of P by Theorem 6.3.20. We now show that the set of sides of $\overline{P}$ is locally finite in E^n. Let x be a point of E^n. We need to find an open neighborhood N of x in E^n that meets only finitely many sides of $\overline{P}$. If x is in $E^n - \overline{P}$, we may take $N = E^n - \overline{P}$. If x is in D^n, then such an N exists, since the set of sides of P is locally finite in D^n. Therefore, we may assume that x is an ideal vertex of P.

We pass to the upper half-space model U^n of hyperbolic space and position P so that $x = \infty$. Let Σ be a horizontal horosphere of U^n that meets only the the sides of P incident with ∞. Then $L(\infty) = P \cap \Sigma$ is compact. By Theorem 6.3.22, the sides of P incident with ∞ are the vertical sides of P. Let B be a ball in E^n centered at a point in E^{n-1} such that $L(\infty) \subset B$. Then B contains the closures of all the hemispherical sides of P, since all the hemispherical sides of P lie below $L(\infty)$. Therefore $N = \hat{E}^n - \overline{B}$ is an open neighborhood of ∞ in $\hat{E}^n$ that meets only the sides of $\overline{P}$ containing ∞. As $L(\infty)$ is compact, $L(\infty)$ has only finitely sides. Therefore P has only finitely many sides incident with ∞ by Theorem 6.3.23. Hence N meets only finitely many sides of $\overline{P}$. We pass back to the projective disk model D^n of hyperbolic space. Then the set of sides of $\overline{P}$ is locally finite in E^n. Therefore $\overline{P}$ is a convex polyhedron in E^n. □

Definition: A *generalized vertex* of a convex polyhedron P in B^n is either a vertex of P or an ideal vertex of P.

Definition: The *convex hull* in D^n of a subset S of $\overline{D}^n$ is the intersection of the convex hull of S in E^n with D^n.

Theorem 6.3.25. *Let P be a convex polyhedron in D^n. Then its closure $\overline{P}$ in E^n is a convex polyhedron in E^n if and only if P has only finitely many generalized vertices and P is the convex hull of its generalized vertices.*

Proof: Let $m = \dim P$. We may assume that $m > 0$. Suppose that $\overline{P}$ is a convex polyhedron in E^n. If $m = 1$, the sides of $\overline{P}$ are the two endpoints of $\overline{P}$. If $m > 1$, the sides of $\overline{P}$ are the closures of the sides of P by Theorem 6.3.20. We claim that the vertices of $\overline{P}$ are the generalized vertices of P. The proof is by induction on m. This is certainly true if $m = 1$, so assume that $m > 1$ and the claim is true for all $(m-1)$-dimensional convex polyhedra in D^n. Now the vertices of $\overline{P}$ are the vertices of the sides of $\overline{P}$. Therefore, the vertices of $\overline{P}$ are the generalized vertices of the sides of P by the induction hypothesis. Let v be a vertex of $\overline{P}$ in S^{n-1}. Then v is an ideal vertex of P by Theorem 6.3.24. Hence, every vertex of $\overline{P}$ is a generalized vertex of P. If v is an ideal vertex of P, then v is an ideal vertex of every side of P incident with v and therefore v is a vertex of $\overline{P}$. Hence, every generalized vertex of P is a vertex of $\overline{P}$. Thus, the vertices of $\overline{P}$ are the generalized vertices of P, which completes the induction.

Let V be the set of vertices of $\overline{P}$. As $\overline{P}$ is compact, V is finite and $\overline{P} = C(V)$ by Theorem 6.3.17. Hence P has only finitely many generalized vertices and P is the convex hull of its generalized vertices, since

$$P = \overline{P} \cap D^n = C(V) \cap D^n.$$

Conversely, suppose that P has only finitely many generalized vertices and P is the convex hull of its generalized vertices. Let V be the set of generalized vertices of P and let $C(V)$ be the convex hull of V in E^n. Then we have

$$P = C(V) \cap D^n.$$

As $V \subset \overline{D}^n$ and $\overline{D}^n$ is a convex subset of E^n, we have that $C(V) \subset \overline{D}^n$. Clearly, we have

$$C(V) \cap S^{n-1} \subset V.$$

Therefore, we have

$$C(V) = P \cup V.$$

Now $C(V)$ is a closed subset of E^n containing P, since V is finite. Therefore, we have

$$\overline{P} \subset C(V) = P \cup V \subset \overline{P}.$$

Hence, we have $\overline{P} = P \cup V$. Therefore, every ideal point of P is an ideal vertex of P. Hence $\overline{P}$ is a convex polyhedron in E^n by Theorem 6.3.24. □

Theorem 6.3.26. *Let P be an m-dimensional convex polyhedron in D^n, with $m > 1$. Then its closure $\overline{P}$ in E^n is a convex polyhedron in E^n if and only if P has only finitely many sides and P has finite volume in $\langle P \rangle$.*

Proof: We may assume that $m = n$. Suppose that $\overline{P}$ is a convex polyhedron in E^n. By Theorem 6.3.20, the sides of $\overline{P}$ are the closures of the sides of P. As $\overline{P}$ is compact, $\overline{P}$ has only finitely many sides. Therefore P has only finitely many sides.

By the argument in the proof of Theorem 6.3.24, every ideal point of P is a vertex of $\overline{P}$. As $\overline{P}$ is compact, $\overline{P}$ has only finitely many vertices. Therefore P has only finitely many ideal points. Now every ideal point of P is an ideal vertex of P by Theorem 6.3.24. Let $v_1, \ldots, v_k$ be the ideal vertices of P. For each i, choose a horoball B_i based at v_i such that $\overline{B}_i$ meets only the sides of P incident with v_i. Then the set

$$P - (B_1 \cup \cdots \cup B_k)$$

is compact and therefore has finite volume. Hence, it suffices to show that $P \cap B_i$ has finite volume for each $i = 1, \ldots, k$.

Let v be an ideal vertex of P and let B be the corresponding horoball. We now pass to the upper half-space model U^n. Without loss of generality, we may assume that $v = \infty$. Then B is of the form

$$\{x \in U^n : x_n > s\}$$

for some $s > 0$. Now all the sides of P incident with ∞ are vertical. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then by Theorem 4.6.7, we have

$$\begin{aligned}
\mathrm{Vol}(P \cap B) &= \int_{P\cap B} \frac{dx_1 \cdots dx_n}{(x_n)^n} \\
&= \int_s^\infty \left\{ \int_{\nu(P\cap\partial B)} dx_1 \cdots dx_{n-1} \right\} \frac{dx_n}{(x_n)^n} \\
&= \mathrm{Vol}(\nu(P \cap \partial B)) \left[\frac{1}{(n-1)} \frac{-1}{x^{n-1}} \right]_s^\infty \\
&= \frac{1}{n-1} \mathrm{Vol}(\nu(P \cap \partial B)) \frac{1}{s^{n-1}}.
\end{aligned}$$

Now the set $P \cap \partial B$ is compact, since v is an ideal vertex of P. Therefore $\mathrm{Vol}(P \cap B)$ is finite. Thus P has finite volume.

Conversely, suppose that P has only finitely many sides and P has finite volume in D^n. Then every ideal point of P is a horopoint of P. The above volume computation shows that the link of every ideal point of P has finite volume and is therefore compact. See Exercise 6.3.5. Hence, every ideal point of P is an ideal vertex. Therefore $\overline{P}$ is a convex polyhedron in E^n by Theorem 6.3.24. □

Exercise 6.3

1. Let P be a convex polyhedron in X. Prove that P has a countable number of sides; and if $X = S^n$, then P has a finite number of sides.
2. Prove that the intersection of a locally finite family of closed half-spaces of X is a convex polyhedron in X.

3. Find an example of a convex polyhedron in E^3 such that the family of half-spaces in Theorem 6.3.2 is not locally finite.

4. Let P be a convex polyhedron in E^n or H^n. Prove that P is compact if and only if P does not contain a geodesic ray.

5. Let P be an n-dimensional convex polyhedron in E^n. Prove that P is compact if and only if the volume of P is finite.

6. Let P be a subset of S^n. Prove that P is a convex polyhedron in S^n if and only if $K(P)$ is a convex polyhedron in E^{n+1}. See Exercise 6.2.7.

7. Let x be a point of an m-dimensional convex polyhedron P in X, with $m > 0$, let r be a real number such that $0 < r < \pi/2$ and r is less than the distance from x to any side of P not containing x, and let $L(x) = P \cap S(x, r)$. Prove that

   (1) the link $L(x)$ is a great $(m-1)$-sphere of $S(x,r)$ if and only if x is in P°;

   (2) the intersection of all the sides of $L(x)$ is a great $(k-1)$-sphere of $S(x,r)$ if and only if x is in the interior of a k-face of P with $0 < k < m$;

   (3) the link $L(x)$ is contained in an open hemisphere of $S(x,r)$ if and only if x is a vertex of P.

8. Find an example of a convex polygon in D^2 of finite area with an infinite number of sides.

9. Let P be an n-dimensional convex polyhedron in B^n of finite volume. Prove that P is has finitely many sides if and only if every ideal point of P is a horopoint of P.

§6.4. Polytopes

Throughout this section, $X = S^n, E^n$, or H^n with $n > 0$. We now consider the classical polyhedra in X.

Definition: A *polytope* in X is a convex polyhedron P in X such that

(1) P has only finitely many vertices;

(2) P is the convex hull of its vertices;

(3) P is not a pair of antipodal points of S^n.

Theorem 6.4.1. *A convex polyhedron P in X is a polytope in X if and only if P is compact, and if $X = S^n$, then P is contained in an open hemisphere of S^n.*

Proof: This follows immediately from Theorems 6.3.17 and 6.3.18. □

Theorem 6.4.2. *An m-dimensional polytope P in X has at least $m+1$ vertices.*

Proof: Assume first that P is in E^n. The proof is by induction on the dimension m. The theorem is certainly true when $m = 0$, so suppose that $m > 0$ and the theorem is true in dimension $m - 1$. Let S be a side of P. Then S is a polytope by Theorem 6.4.1. Hence, by the induction hypothesis, S has at least m vertices. Now since P is the convex hull of its vertices, S cannot contain all the vertices of P. Therefore P has at least $m + 1$ vertices. This completes the induction.

Now assume that P is in S^n. Then by gnomonic projection, we can view P as a Euclidean polyhedron. Therefore P has at least $m + 1$ vertices by the Euclidean case.

Now assume that P is H^n. We pass to the projective disk model D^n. Then P is a Euclidean polyhedron, since P is compact. Therefore P has at least $m + 1$ vertices by the Euclidean case. □

Definition: An *m-simplex* in X is an m-dimensional polytope in X with exactly $m + 1$ vertices.

We leave it as an exercise to prove that a subset S of E^n is an m-simplex if and only if S is the convex hull of an affinely independent subset of $m+1$ points $\{v_0, \ldots, v_m\}$ of E^n.

Example: The *standard m-simplex* Δ^m in E^n is the convex hull of the points $0, e_1, \ldots, e_m$ of E^n.

Theorem 6.4.3. *An m-dimensional polytope in X, with $m > 0$, has at least $m + 1$ sides.*

Proof: This follows from Theorems 6.3.7, 6.3.8, and 6.4.1. □

Theorem 6.4.4. *An m-dimensional polytope in X, with $m > 0$, is an m-simplex if and only if P has exactly $m + 1$ sides.*

Proof: The proof is by induction on m. The theorem is certainly true for $m = 1$, so assume that $m > 1$ and the theorem is true for all $(m-1)$-dimensional polytopes in X. Suppose that P is an m-simplex. Then P has at least $m + 1$ sides by Theorem 6.4.3. Let S be a side of P. Then S does not contain all the vertices of P, since P is the convex hull of its vertices. Therefore S has at most m vertices. As S is an $(m-1)$-dimensional polytope, S has at least m vertices by Theorem 6.4.2. Therefore S has exactly m vertices. Hence S is an $(m - 1)$-simplex. Thus, each side of P is an $(m - 1)$-simplex. Hence, each side of P is the convex hull of m vertices of P. Since the set of $m+1$ vertices of P has exactly $m+1$ subsets with m vertices, P has at most $m + 1$ sides. Therefore P has exactly $m + 1$ sides.

Conversely, suppose that P has exactly $m+1$ sides. Then P has at least $m+1$ vertices by Theorem 6.4.2. Now by Theorem 6.3.14, each vertex of P is the intersection of at least m sides of P. As the intersection of all the sides of P is contained in each vertex of P, the intersection of all the sides of P is empty. Therefore, each vertex of P is the intersection of exactly m sides of P. Since the set of $m+1$ sides of P has exactly $m+1$ subsets with m sides, P has at most $m+1$ vertices. Therefore P has exactly $m+1$ vertices. Thus P is an m-simplex. □

Theorem 6.4.5. *Let P be a polytope in X. Then the group of symmetries of P in $\langle P\rangle$ is finite.*

Proof: The proof is by induction on $\dim P = m$. The theorem is obviously true if $m = 0$, so assume that $m > 0$ and the theorem is true for all $(m-1)$-dimensional polytopes in X. Let Γ be the group of symmetries of P in $\langle P\rangle$. Then Γ acts on the finite set $\mathcal{S}$ of sides of P. Now $\mathcal{S}$ is nonempty by Theorem 6.4.3, and each side of P is an $(m-1)$-dimensional polytope by Theorem 6.4.1. By the induction hypothesis, the stabilizer of each side of P is finite. Therefore Γ is finite. □

Definition: The *centroid* of a polytope P in X with vertices $v_1, \ldots, v_k$ is the point

$$c = \begin{cases} (v_1 + \cdots + v_k)/k & \text{if } X = E^n, \\ \frac{(v_1+\cdots+v_k)/k}{|(v_1+\cdots+v_k)/k|} & \text{if } X = S^n, \\ \frac{(v_1+\cdots+v_k)/k}{|||(v_1+\cdots+v_k)/k|||} & \text{if } X = H^n. \end{cases}$$

Note that c is a well-defined point of X by Theorems 3.1.1 and 6.4.1. A polytope P in X contains its centroid c, since c is in the convex hull of the vertices of P. It is an exercise to prove that the centroid c of P is in the interior of P.

Theorem 6.4.6. *Let P be a polytope in X. Then every symmetry of P fixes the centroid of P.*

Proof: Let g be a symmetry of P. Then g permutes the vertices $v_1, \ldots, v_k$ of P. If $X = E^n$, then there is a point a of E^n and an A in $\mathrm{O}(n)$ such that $g = a + A$ by Theorem 1.3.5. If $X = S^n$ or H^n, then g is linear. Therefore, we have

$$g\left(\frac{v_1 + \cdots + v_k}{k}\right) = \frac{v_1 + \cdots + v_k}{k}.$$

Hence $gc = c$. □

Generalized Polytopes

We now generalize the concept of a polytope in H^n to allow ideal vertices on the sphere at infinity of H^n. It will be more convenient for us to work in a model of hyperbolic space that allows a direct representation of the sphere at infinity.

Definition: A *generalized polytope* in D^n is a convex polyhedron P in D^n such that P has only finitely many generalized vertices and P is the convex hull of its generalized vertices.

Theorem 6.4.7. *A convex polyhedron P in D^n is a generalized polytope in D^n if and only if its closure $\overline{P}$ in E^n is a polytope in E^n.*

Proof: This follows immediately from Theorems 6.3.25 and 6.4.1. □

Theorem 6.4.8. *Let P be an m-dimensional convex polyhedron in D^n, with $m > 1$. Then P is a generalized polytope in D^n if and only if P has finitely many sides and P has finite volume in $\langle P \rangle$.*

Proof: This follows immediately from Theorems 6.3.25 and 6.3.26. □

Theorem 6.4.9. *An m-dimensional generalized polytope P in D^n has at least $m + 1$ generalized vertices.*

Proof: By Theorem 6.4.7, we have that $\overline{P}$ is a polytope in E^n. By Theorem 6.4.2, we have that $\overline{P}$ has at least $m+1$ vertices. Now by the argument in the proof of Theorem 6.3.25, the vertices of $\overline{P}$ are the generalized vertices of P. Therefore P has at least $m + 1$ generalized vertices. □

Definition: A *generalized m-simplex* in D^n is an m-dimensional generalized polytope in D^n with exactly $m + 1$ generalized vertices.

Note that a generalized 0-simplex is just a point. A generalized 1-simplex is either a geodesic segment or a geodesic ray or a geodesic.

Theorem 6.4.10. *A convex polyhedron in D^n is a generalized m-simplex in D^n if and only if its closure in E^n is an m-simplex in E^n.*

Proof: Suppose that P is a generalized m-simplex. By Theorem 6.4.7, we have that $\overline{P}$ is a polytope in E^n. By the argument in the proof of Theorem 6.3.25, the vertices of $\overline{P}$ are the generalized vertices of P. Therefore $\overline{P}$ has exactly $m + 1$ vertices. Thus $\overline{P}$ is an m-simplex in E^n.

Conversely, suppose that $\overline{P}$ is an m-simplex in E^n. Then P is a polytope in D^n by Theorem 6.4.7. By the argument in the proof of Theorem 6.3.25, the vertices of $\overline{P}$ are the generalized vertices of P. Therefore P has exactly $m + 1$ generalized vertices. Thus P is a generalized m-simplex. □

Theorem 6.4.11. *An m-dimensional generalized polytope P in D^n, with $m > 1$, has at least $m + 1$ sides.*

Proof: By Theorem 6.4.7, we have that $\overline{P}$ is a polytope in E^n. By Theorem 6.3.20, the sides of $\overline{P}$ are the closures of the sides of P. Now by 6.4.3, we have that $\overline{P}$ has at least $m + 1$ sides. Therefore P has at least $m + 1$ sides. □

Theorem 6.4.12. *An m-dimensional generalized polytope P in D^n, with $m > 1$, is a generalized m-simplex if and only if P has exactly $m+1$ sides.*

Proof: By Theorem 6.4.7, we have that $\overline{P}$ is a polytope in E^n. By Theorem 6.4.10, we have that P is a generalized m-simplex if and only if $\overline{P}$ is an m-simplex in E^n. By Theorem 6.3.20, the sides of $\overline{P}$ are the closures of the sides of P. Therefore P is a generalized m-simplex if and only P has exactly $m + 1$ sides by Theorem 6.4.4. □

Definition: An *ideal polytope* in D^n is a generalized polytope in D^n all of whose generalized vertices are ideal.

Definition: An *ideal m-simplex* in D^n is a generalized m-simplex in D^n all of whose generalized vertices are ideal.

Example: Let $v_0, \ldots, v_m$ be $m + 1$ affinely independent vectors in S^{n-1}, with $m > 0$. Then their convex hull is a Euclidean m-simplex Δ inscribed in S^{n-1}. Therefore Δ minus its vertices is an ideal m-simplex in D^n by Theorem 6.4.10.

Theorem 6.4.13. *Let P be a generalized polytope in D^n that is not a geodesic of D^n. Then the group of symmetries of P in $\langle P \rangle$ is finite.*

Proof: Let Γ be the group of symmetries of P in $\langle P \rangle$. Then Γ permutes the generalized vertices of P. Let g be an element of Γ that fixes all the generalized vertices of P. We claim that $g = 1$. The proof is by induction on $m = \dim P$. This is certainly true if $m = 0$, so assume that $m > 0$, and the claim is true for all $(m-1)$-dimensional generalized polytopes in D^n that are not geodesics. Let v be a generalized vertex of P. Then P has a side S that is not incident with v, since P is the convex hull of its generalized vertices and P is not a geodesic. If S is a geodesic of D^n, then $g = 1$, since g fixes the endpoints of S and v. If S is not a geodesic, then by the induction hypothesis, g is the identity on $\langle S \rangle$. Therefore $g = 1$ by Theorem 4.3.6. Hence Γ injects into the group of permutations of the generalized vertices of P. Therefore Γ is finite. □

Regular Polytopes

Let P be an m-dimensional polytope in X. A *flag* of P is a sequence $(F_0, F_1, \ldots, F_m)$ of faces of P such that $\dim F_i = i$ for each i and F_i is a side of F_{i+1} for each $i < m$. Let $\mathcal{F}$ be the set of all flags of P and let Γ be the group of symmetries of P in $\langle P \rangle$. Then Γ acts on $\mathcal{F}$ by

$$g(F_0, F_1, \ldots, F_m) = (gF_0, gF_1, \ldots, gF_m).$$

Definition: A *regular polytope* in X is a polytope P in X whose group of symmetries in $\langle P \rangle$ acts transitively on the set of its flags.

Theorem 6.4.14. *Let P be a regular polytope in X. Then P is inscribed in a sphere of $\langle P \rangle$ centered at the centroid of P.*

Proof: Let Γ be the group of symmetries of P. Then Γ acts transitively on the vertices $v_1, \ldots, v_k$ of P. Now each element of Γ fixes the centroid c of P by Theorem 6.4.6. Therefore

$$d(c, v_1) = d(c, v_i) \quad \text{for each } i.$$

Hence P is inscribed in the sphere of $\langle P \rangle$ centered at c of radius $d(c, v_1)$. □

The regular polytopes in X are completely classified. First, we consider the classification of Euclidean regular polytopes.

(1) A 1-dimensional, Euclidean, regular polytope is a line segment.

(2) A 2-dimensional, Euclidean, regular polytope is a regular polygon.

(3) A 3-dimensional, Euclidean, regular polytope is a regular solid. Up to similarity, there only five regular solids, the regular tetrahedron, hexahedron, octahedron, dodecahedron, and icosahedron.

(4) There are up to similarity only six 4-dimensional, Euclidean, regular polytopes. They are called the 5-cell, 8-cell, 16-cell, 24-cell, 120-cell, and 600-cell. A k-cell has k sides.

(5) For $n \geq 5$, there are up to similarity only three n-dimensional, Euclidean, regular polytopes, the regular n-simplex with $n+1$ sides, the n-cube with $2n$ sides, and its dual with 2^n sides.

The classification of regular polytopes in S^n and H^n is essentially the same as the classification of regular polytopes in E^n. The only difference is that in S^n and H^n regular polytopes of the same combinatorial type come in different nonsimilar sizes.

Theorem 6.4.15. *Let P be a polytope in S^n. Then P is regular, with centroid e_{n+1}, if and only if the gnomonic projection of P into E^n is regular with centroid* 0.

Proof: We may assume that $\langle P\rangle = S^n$. Suppose that P is regular with centroid e_{n+1}. Let A be a symmetry of P. Then A is an element of $\mathrm{O}(n+1)$ that fixes e_{n+1}. Hence, the restriction of A to E^n is an element $\overline{A}$ of $\mathrm{O}(n)$. The gnomonic projection of S^n_+ onto E^n is given by

$$\phi(x) = \overline{x}/x_{n+1}, \quad \text{where } \overline{x} = (x_1, \ldots, x_n).$$

Observe that

$$\phi(Ax) = \overline{Ax}/(Ax)_{n+1} = \overline{A}\overline{x}/x_{n+1} = \overline{A}\phi(x).$$

Therefore, we have

$$\overline{A}\phi(P) = \phi(AP) = \phi(P).$$

Hence $\overline{A}$ is a symmetry of $\phi(P)$. Therefore $\phi(P)$ is regular in E^n. Let $v_1, \ldots, v_k$ be the vertices of P. Then we have

$$v_1 + \cdots + v_k = |v_1 + \cdots + v_k|e_{n+1}.$$

Therefore, we have

$$\overline{v}_1 + \cdots + \overline{v}_k = 0.$$

Observe that

$$\cos\theta(v_i, e_{n+1}) = v_i \cdot e_{n+1} = (v_i)_{n+1}.$$

Therefore, we have

$$(v_1)_{n+1} = (v_i)_{n+1} \quad \text{for all } i.$$

Hence

$$\frac{(\overline{v}_1/(v_1)_{n+1}) + \cdots + (\overline{v}_k/(v_k)_{n+1})}{k} = \frac{\overline{v}_1 + \cdots + \overline{v}_k}{k(v_1)_{n+1}} = 0.$$

Thus, the centroid of $\phi(P)$ is 0.

Conversely, suppose that $\phi(P)$ is regular with centroid 0. Let A be a symmetry of $\phi(P)$. Then A is an element of $\mathrm{O}(n)$. Let $\hat{A}$ be the element of $\mathrm{O}(n+1)$ that extends A and fixes e_{n+1}. Then we have $A\phi = \phi\hat{A}$. Hence, we have

$$\hat{A}P = \hat{A}\phi^{-1}\phi(P) = \phi^{-1}A\phi(P) = \phi^{-1}\phi(P) = P.$$

Hence $\hat{A}$ is a symmetry of P. Therefore P is regular.

Now since the symmetries of P of the form $\hat{A}$ fix e_{n+1} and act transitively on the vertices of P, we conclude as before that

$$(v_i)_{n+1} = (v_1)_{n+1} \quad \text{for all } i.$$

Therefore

$$\frac{\overline{v}_1 + \cdots + \overline{v}_k}{k(v_1)_{n+1}} = \frac{(\overline{v}_1/(v_1)_{n+1}) + \cdots + (\overline{v}_k/(v_k)_{n+1})}{k} = 0.$$

Hence, we have

$$\overline{v}_1 + \cdots + \overline{v}_k = 0.$$

Therefore, we have

$$v_1 + \cdots + v_k = |v_1 + \cdots + v_k|e_{n+1}.$$

Thus, the centroid of P is e_{n+1}. $\square$

Theorem 6.4.16. *Let P be a polytope in D^n. Then P is regular with centroid* 0 *if and only if P is regular in E^n with centroid* 0.

Proof: The proof is the same as the proof of Theorem 6.4.15 with S^n replaced by H^n. □

Regular Ideal Polytopes

Let P be an ideal polytope in D^n. A *flag* of P is defined as before except that vertices are now ideal.

Definition: A *regular ideal polytope* in D^n is an ideal polytope P in D^n whose group of symmetries in $\langle P \rangle$ acts transitively on the set of its flags.

Theorem 6.4.17. *An ideal polytope P in D^n is regular if and only if P is congruent to an ideal polytope in D^n whose closure in E^n is a regular polytope in E^n.*

Proof: We may assume that $\langle P \rangle = D^n$ and $n > 1$. Let Γ be the group of symmetries of P. Then Γ is finite by Theorem 6.4.13. Hence Γ fixes a point of D^n by Theorems 5.5.1 and 5.5.2. By conjugating Γ, we may assume that Γ fixes 0. Then every symmetry of P is a symmetry of $\overline{P}$. Therefore, if P is regular, then $\overline{P}$ is regular.

Conversely, suppose that $\overline{P}$ is regular. Then the centroid of $\overline{P}$ is 0, since $\overline{P}$ is inscribed in S^{n-1}. See Exercise 6.4.6. Hence, every symmetry of $\overline{P}$ is a symmetry of P. Therefore P is regular. □

Exercise 6.4

1. Prove that a subset S of E^n is an m-simplex if and only if S is the convex hull of an affinely independent subset $\{v_0, \ldots, v_m\}$ of E^n.
2. An *edge* of a convex polyhedron P in X is a 1-face of P. Prove that an m-dimensional polytope in X, with $m > 1$, has at least $m(m+1)/2$ edges and at least $m(m+1)/2$ ridges.
3. Prove that an m-simplex in X has $\binom{m+1}{k+1}$ k-faces for each $k = 0, \ldots, m$.
4. Let P be a polytope in X. Prove that the centroid of P is in P°.
5. Prove that the centroid of a regular polytope P in X is the only point of $\langle P \rangle$ fixed by all the symmetries of P in $\langle P \rangle$.
6. Let P be an n-dimensional polytope in X that is inscribed in a sphere S of X. Prove that S is unique.
7. Prove that the group of symmetries of an n-simplex in X is isomorphic to the group of permutations of its vertices.

§6.5. Fundamental Domains

Let Γ be a group acting on a metric space X. The *orbit space* of the action of Γ on X is defined to be the set of Γ-orbits

$$X/\Gamma = \{\Gamma x : x \in X\}$$

topologized with the quotient topology from X. The quotient map will be denoted by

$$\pi : X \to X/\Gamma.$$

Recall that if A and B are subsets of X, then the *distance* from A to B in X is defined to be

$$\text{dist}(A, B) = \inf\{d(x, y) : x \in A \quad \text{and} \quad y \in B\}.$$

The *orbit space distance function*

$$d_\Gamma : X/\Gamma \times X/\Gamma \to \mathbb{R}$$

is defined by the formula

$$d_\Gamma(\Gamma x, \Gamma y) = \text{dist}(\Gamma x, \Gamma y). \tag{6.5.1}$$

If d_Γ is a metric on X/Γ, then d_Γ is called the *orbit space metric* on X/Γ.

Theorem 6.5.1. *Let Γ be a group of isometries of a metric space X. Then d_Γ is a metric on X/Γ if and only if each Γ-orbit is a closed subset of X.*

Proof: Let x, y be in X and let g, h be in Γ. Then

$$d(gx, hy) = d(x, g^{-1}hy).$$

Therefore

$$\text{dist}(\Gamma x, \Gamma y) = \text{dist}(x, \Gamma y).$$

Suppose that d_Γ is a metric and $\Gamma x \neq \Gamma y$. Then

$$\text{dist}(x, \Gamma y) = d_\Gamma(\Gamma x, \Gamma y) > 0.$$

Let $r = \text{dist}(x, \Gamma y)$. Then $B(x, r) \subset X - \Gamma y$. Hence $X - \Gamma y$ is open and therefore Γy is closed. Thus, each Γ-orbit is a closed subset of X.

Conversely, suppose that each Γ-orbit is a closed subset of X. If x, y are in X and $\Gamma x \neq \Gamma y$, then

$$d_\Gamma(\Gamma x, \Gamma y) = \text{dist}(x, \Gamma y) > 0.$$

Thus d_Γ is nondegenerate.

Now let x, y, z be in X and let g, h be in Γ. Then

$$\begin{aligned} d(x, gy) + d(y, hz) &= d(x, gy) + d(gy, ghz) \\ &\geq d(x, ghz) \\ &\geq \text{dist}(x, \Gamma z). \end{aligned}$$

Therefore

$$\text{dist}(x, \Gamma z) \leq \text{dist}(x, \Gamma y) + \text{dist}(y, \Gamma z).$$

Hence d_Γ satisfies the triangle inequality. Thus d_Γ is a metric on X/Γ. $\square$

Corollary 1. *If Γ is a discontinuous group of isometries of a metric space X, then d_Γ is a metric on X/Γ.*

Proof: By Theorem 5.3.4, each Γ-orbit is a closed subset of X. □

Theorem 6.5.2. *Let Γ be a group of isometries of a metric space X such that d_Γ is a metric on X/Γ. Then the metric topology on X/Γ, determined by d_Γ, is the quotient topology; if $\pi : X \to X/\Gamma$ is the quotient map, then for each x in X and $r > 0$, we have*

$$\pi(B(x,r)) = B(\pi(x),r).$$

Proof: Let x be in X and suppose that $r > 0$. Then clearly

$$\pi(B(x,r)) \subset B(\pi(x),r).$$

To see the reversed inclusion, suppose that y is in X and

$$d_\Gamma(\Gamma x, \Gamma y) < r.$$

Then we have

$$\mathrm{dist}(x, \Gamma y) < r.$$

Consequently, there is a g in Γ such that $d(x,gy) < r$. Moreover, we have $\pi(gy) = \Gamma y$. Thus, we have

$$\pi(B(x,r)) = B(\pi(x),r).$$

Hence π is open and continuous with respect to the metric topology on X/Γ.

Let U be an open subset of X/Γ with respect the quotient topology. Then $\pi^{-1}(U)$ is open in X. Therefore $U = \pi(\pi^{-1}(U))$ is open in the metric topology on X/Γ. Let x be in X and suppose that $r > 0$. Then

$$\pi^{-1}(B(\pi(x),r)) = \bigcup_{g\in\Gamma} B(gx,r).$$

Therefore $B(\pi(x),r)$ is open in the quotient topology on X/Γ. Thus, the metric topology on X/Γ determined by d_Γ is the quotient topology. □

Fundamental Regions

Definition: A subset R of a metric space X is a *fundamental region* for a group Γ of isometries of X if and only if

(1) the set R is open in X;

(2) the members of $\{gR : g \in \Gamma\}$ are mutually disjoint; and

(3) $X = \cup\{g\overline{R} : g \in \Gamma\}$.

Definition: A subset D of a metric space X is a *fundamental domain* for a group Γ of isometries of X if and only if D is a connected fundamental region for Γ.

Theorem 6.5.3. *If a group Γ of isometries of a metric space X has a fundamental region, then Γ is a discrete subgroup of* $\mathrm{I}(X)$.

Proof: Let x be a point of a fundamental region R for a group of isometries Γ of a metric space X. Then the evaluation map

$$\varepsilon : \Gamma \to \Gamma x,$$

defined by $\varepsilon(g) = gx$, is a continuous. Now the point x is open in Γx, since $R \cap \Gamma x = \{x\}$. Moreover, the stabilizer Γ_x is trivial. Hence $1 = \varepsilon^{-1}(x)$ is open in Γ. Therefore Γ is discrete by the proof of Lemma 1 of §5.3. □

Theorem 6.5.4. *If R is a fundamental region for a group Γ of isometries of a metric space X, then for each $g \neq 1$ in Γ, we have*

$$\overline{R} \cap g\overline{R} \subset \partial R.$$

Proof: Let x be a point of $\overline{R} \cap g\overline{R}$ and let r be a positive real number. Then $B(x, r)$ contains a point of R, since x is in $\overline{R}$, and a point of gR, since x is in $g\overline{R}$. As R and gR are disjoint, $B(x, r)$ meets R and $X - R$. Hence x is in ∂R. Thus ∂R contains $\overline{R} \cap g\overline{R}$ for each $g \neq 1$ in Γ. □

Definition: A fundamental region R for a discrete group Γ of isometries of $X = S^n, E^n$, or H^n is *proper* if and only if $\mathrm{Vol}(\partial R) = 0$, that is, ∂R is a null set in X.

Example 1. Let α be the antipodal map of S^n. Then $\Gamma = \{1, \alpha\}$ is a discrete subgroup of $\mathrm{I}(S^n)$ and any open hemisphere of S^n is a proper fundamental domain for Γ. The orbit space S^n/Γ is elliptic n-space P^n.

Example 2. Let τ_i be the translation of E^n by e_i for $i = 1, \ldots, n$. Then $\{\tau_1, \ldots, \tau_n\}$ generates a discrete subgroup Γ of $\mathrm{I}(E^n)$. A proper fundamental domain for Γ is the open unit n-cube $(0, 1)^n$ in E^n. The orbit space E^n/Γ is similar to the n-torus $(S^1)^n$.

Example 3. Let ρ be the reflection of H^n in a hyperplane P. Then $\Gamma = \{1, \rho\}$ is a discrete subgroup of $\mathrm{I}(H^n)$. Either one of the two open half-spaces of H^n bounded by P is a proper fundamental domain for Γ. The orbit space H^n/Γ is isometric to a closed half-space of H^n.

Theorem 6.5.5. *If Γ is a discrete group of isometries of $X = S^n, E^n$, or H^n, then all the proper fundamental regions for Γ have the same volume.*

Proof: Let R and S be proper fundamental regions for Γ. Observe that

$$X - \bigcup_{g\in\Gamma} gS \subset \bigcup_{g\in\Gamma} g\partial S.$$

Therefore, we have

$$\mathrm{Vol}\Big(X - \bigcup_{g\in\Gamma} gS\Big) = 0.$$

Hence

$$\begin{aligned}\mathrm{Vol}(R) &= \mathrm{Vol}\big(R \cap \big(\textstyle\bigcup_{g\in\Gamma} gS\big)\big)\\ &= \mathrm{Vol}\big(\textstyle\bigcup_{g\in\Gamma} R \cap gS\big)\\ &= \sum_{g\in\Gamma} \mathrm{Vol}(R \cap gS)\\ &= \sum_{g\in\Gamma} \mathrm{Vol}(g^{-1}R \cap S) \quad = \quad \mathrm{Vol}(S).\end{aligned}$$

□

Theorem 6.5.6. *If R is a fundamental region for a group Γ of isometries of a metric space X and g is an element of Γ fixing a point of X, then g is conjugate in Γ to an element h such that h fixes a point of ∂R.*

Proof: This is certainly true if $g = 1$, so assume that $g \neq 1$. Let x be a fixed point of g. Then there is a point y of $\overline{R}$ and an element f of Γ such that $fx = y$. Let $h = fgf^{-1}$. Then h fixes y. As R and hR are disjoint, y is in ∂R. □

Corollary 2. *Let R be a fundamental region for a discrete group Γ of isometries of E^n or H^n. If g is an elliptic element of Γ, then g is conjugate in Γ to an element h such that h fixes a point of ∂R.*

Proof: Every elliptic element of Γ has a fixed point. □

Lemma 1. *If Γ is a discrete group of isometries of H^n such that H^n/Γ is compact, then there is an $\ell > 0$ such that $d(x, hx) \geq \ell$ for all x in H^n and all nonelliptic h in Γ.*

Proof: Let x be an arbitrary point of H^n and set

$$r(x) = \frac{1}{2}\mathrm{dist}(x, \Gamma x - \{x\}).$$

Then any two open balls in

$$\{B(gx, r(x)) : g \in \Gamma\}$$

are either the same or are disjoint. Let $\pi : H^n \to H^n/\Gamma$ be the quotient map. As H^n/Γ is compact, the open cover

$$\{B(\pi(y), r(y)) : y \in H^n\}$$

has a Lebesgue number $\ell > 0$. Hence, there is a y in H^n such that $B(\pi(y), r(y))$ contains $B(\pi(x), \ell)$. Consequently $\bigcup_{g\in\Gamma} B(gy, r(y))$ contains $B(x, \ell)$. As $B(x, \ell)$ is connected, there is a g in Γ such that $B(gy, r(y))$ contains $B(x, \ell)$. By replacing y with gy, we may assume that $g = 1$.

Now let h be an arbitrary nonelliptic element of Γ. As $B(y, r(y))$ and $B(hy, r(y))$ are disjoint, $B(x, \ell)$ and $B(hx, \ell)$ are disjoint. Therefore $d(x, hx) \geq \ell$. □

Theorem 6.5.7. *If Γ is a discrete group of isometries of H^n such that H^n/Γ is compact, then every element of Γ is either elliptic or hyperbolic.*

Proof: On the contrary, suppose that Γ has a parabolic element f. We pass to the upper half-space model U^n. Then we may assume, without loss of generality, that $f(\infty) = \infty$. Then f is the Poincaré extension of a Euclidean isometry of E^{n-1}. By Theorem 4.6.1, we have for each $t > 0$,

$$\begin{aligned} \cosh d(te_n, f(te_n)) &= 1 + \frac{|te_n - f(te_n)|}{2t^2} \\ &= 1 + \frac{|e_n - f(e_n)|}{2t^2}. \end{aligned}$$

Hence

$$\lim_{t\to\infty} \cosh d(te_n, f(te_n)) = 1.$$

Therefore

$$\lim_{t\to\infty} d(te_n, f(te_n)) = 0.$$

But this contradicts Lemma 1. □

Corollary 3. *If Γ is a discrete group of isometries of H^n with a parabolic element, then every fundamental region for Γ is unbounded.*

Proof: Let R be a fundamental region for Γ. If R were bounded, then $\overline{R}$ would be compact; but the quotient map $\pi : H^n \to H^n/\Gamma$ maps $\overline{R}$ onto H^n/Γ, and so H^n/Γ would be compact contrary to Theorem 6.5.7. □

Locally Finite Fundamental Regions

Definition: A fundamental region R for a group Γ of isometries of a metric space X is *locally finite* if and only if $\{g\overline{R} : g \in \Gamma\}$ is a locally finite family of subsets of X.

Example: Every fundamental region of a discrete group Γ of isometries of S^n is locally finite, since Γ is finite.

Let R be a fundamental region for a discontinuous group Γ of isometries of a metric space X, and let $\overline{R}/\Gamma$ be the collection of disjoint subsets of $\overline{R}$,

$$\{\Gamma x \cap \overline{R} : x \in \overline{R}\},$$

topologized with the quotient topology. At times, it will be useful to adopt $\overline{R}/\Gamma$ as a geometric model for X/Γ. The importance of local finiteness in this scheme is underscored by the next theorem.

Theorem 6.5.8. *If R is a fundamental region for a discontinuous group Γ of isometries of a metric space X, then the inclusion $\iota : \overline{R} \to X$ induces a continuous bijection $\kappa : \overline{R}/\Gamma \to X/\Gamma$, and κ is a homeomorphism if and only if R is locally finite.*

Proof: The map κ is defined by $\kappa(\Gamma x \cap \overline{R}) = \Gamma x$. If x, y are in $\overline{R}$ and $\Gamma x = \Gamma y$, then we have

$$\Gamma x \cap \overline{R} = \Gamma y \cap \overline{R}.$$

Therefore κ is injective. As $\overline{R}$ contains a fundamental set, κ is subjective. Let $\eta : \overline{R} \to \overline{R}/\Gamma$ be the quotient map. Then we have a commutative diagram

$$\begin{array}{ccc} \overline{R} & \xrightarrow{\iota} & X \\ \eta \downarrow & & \downarrow \pi \\ \overline{R}/\Gamma & \xrightarrow{\kappa} & X/\Gamma. \end{array}$$

This implies that κ is continuous. Thus κ is a continuous bijection.

Now assume that R is locally finite. To prove that κ is a homeomorphism, it suffices to show that κ is an open map. Let U be an open subset of $\overline{R}/\Gamma$. As η is continuous and surjective, there is an open subset V of X such that $\eta^{-1}(U) = \overline{R} \cap V$ and $\eta(\overline{R} \cap V) = U$. Let

$$W = \bigcup_{g \in \Gamma} g(\overline{R} \cap V).$$

Then we have

$$\begin{aligned} \pi(W) &= \pi(\overline{R} \cap V) \\ &= \pi\iota(\overline{R} \cap V) \\ &= \kappa\eta(\overline{R} \cap V) \quad = \quad \kappa(U). \end{aligned}$$

In order to prove that $\kappa(U)$ is open, it suffices to prove that W is open in X, since π is an open map.

Let w be in W. We need to show that W contains an open ball B centered at w. As W is Γ-invariant, we may assume that w is $\overline{R} \cap V$. As R is locally finite, there is an open ball B centered at w that meets only finitely many Γ-images of $\overline{R}$, say $g_1\overline{R}, \ldots, g_m\overline{R}$. Then we have

$$B \subset g_1\overline{R} \cup \cdots \cup g_m\overline{R}.$$

If $g_i\overline{R}$ does not contain w, then $B - g_i\overline{R}$ is an open neighborhood of w, and so we may shrink B to avoid $g_i\overline{R}$. Thus, we may assume that each $g_i\overline{R}$ contains w. Then $g_i^{-1}w$ is in $\overline{R}$ for each i. As $\eta(g_i^{-1}w) = \eta(w)$, we have that $g_i^{-1}w$ is in $\eta^{-1}(U) = \overline{R} \cap V$. Hence w is in g_iV for each i. By shrinking B still further, we may assume that

$$B \subset g_1V \cap \cdots \cap g_mV.$$

Consequently $B \subset W$, since if x is in B, then x is in both $g_i\overline{R}$ and g_iV for some i, and so x is in $g_i(\overline{R} \cap V)$, which is contained in W. Therefore W is open and κ is an open map. Thus κ is a homeomorphism.

Conversely, suppose that κ is a homeomorphism and on the contrary there is a point y of X at which R is not locally finite. Then there is a sequence $\{x_i\}_{i=1}^{\infty}$ of points in R and a sequence $\{g_i\}_{i=1}^{\infty}$ of distinct elements of Γ such that $g_i x_i \to y$. As gR is open and disjoint from every other Γ-image of R, the point y is not in any gR. Let

$$K = \{x_1, x_2, \ldots\}.$$

As $K \subset R$, we have that $\pi(y)$ is not in $\pi(K)$.

We claim that K is closed in X. Let x be in $X - K$. Now $\Gamma y - \{x\}$ is a closed subset of X by Theorem 5.3.4. Therefore

$$\mathrm{dist}(x, \Gamma y - \{x\}) > 0.$$

Now let

$$r = \frac{1}{2}\mathrm{dist}(x, \Gamma y - \{x\}).$$

As the g_i are distinct, x is equal to at most finitely many $g_i^{-1}y$, since Γ_y is finite. Thus $d(x, g_i^{-1}y) \geq 2r$ for large enough i. As $g_i x_i \to y$, we have that $d(g_i x_i, y) < r$ for large enough i. Hence, for large enough i, we have

$$2r \leq d(x, g_i^{-1}y) \leq d(x, x_i) + d(x_i, g_i^{-1}y)$$

and

$$r < 2r - d(g_i x_i, y) \leq d(x, x_i).$$

Thus $B(x, r)$ contains only finitely many points of K, and so there is an open ball centered at x avoiding K. Thus $X - K$ is open and so K is closed.

As $K \subset R$, we have that $\eta^{-1}(\eta(K)) = K$, and so $\eta(K)$ is closed in $\overline{R}/\Gamma$. Therefore $\kappa\eta(K) = \pi(K)$ is closed in X/Γ, since κ is a homeomorphism. As π is continuous, we have $\pi(g_i x_i) \to \pi(y)$, that is, $\pi(x_i) \to \pi(y)$. As $\pi(K)$ is closed, $\pi(y)$ is in $\pi(K)$, which is a contradiction. Thus R is locally finite. □

Theorem 6.5.9. *Let x be a boundary point of a locally finite fundamental region R for a group Γ of isometries of a metric space X. Then $\partial R \cap \Gamma x$ is finite and there is an $r > 0$ such that if $N(\overline{R}, r)$ is the r-neighborhood of $\overline{R}$ in X, then*

$$N(\overline{R}, r) \cap \Gamma x = \partial R \cap \Gamma x.$$

Proof: As R is locally finite, there is an $r > 0$ such that $B(x, r)$ meets only finitely many Γ-images of $\overline{R}$, say $g_1^{-1}\overline{R}, \ldots, g_m^{-1}\overline{R}$. By shrinking r, if necessary, we may assume that x is in each $g_i^{-1}\overline{R}$. Suppose that gx is also in ∂R. Then x is in $g^{-1}\overline{R}$ and so $g = g_i$ for some i. Hence

$$\partial R \cap \Gamma x \subset \{g_1 x, \ldots, g_m x\}.$$

Moreover, for each i, there is a y_i in ∂R such that $x = g_i^{-1} y_i$. Therefore

$$\partial R \cap \Gamma x = \{g_1 x, \ldots, g_m x\}.$$

Next, suppose that $d(gx, y) < r$ with y in $\overline{R}$. Then $d(x, g^{-1}y) < r$. Hence g is in $\{g_1, \ldots, g_m\}$ and so gx is in ∂R. Thus

$$N(\overline{R}, r) \cap \Gamma x = \partial R \cap \Gamma x.$$ □

Theorem 6.5.10. *Let R be a fundamental region for a discontinuous group Γ of isometries of a locally compact metric space X such that X/Γ is compact. Then R is locally finite if and only if $\overline{R}$ is compact.*

Proof: Suppose that $\overline{R}$ is compact. Then the map $\kappa : \overline{R}/\Gamma \to X/\Gamma$ is a continuous bijection from a compact space to a Hausdorff space and so is a homeomorphism. Therefore R is locally finite by Theorem 6.5.8.

Conversely, suppose that R is locally finite and on the contrary $\overline{R}$ is not compact. Then $\overline{R}$ is not countably compact, since $\overline{R}$ is a metric space. Hence, there is an infinite sequence $\{x_i\}$ in $\overline{R}$ that has no convergent subsequence. As X/Γ is compact, $\{\pi(x_i)\}$ has a convergent subsequence. By passing to this subsequence, we may assume that $\{\pi(x_i)\}$ converges in X/Γ. As the quotient map π maps $\overline{R}$ onto X/Γ, there is a point x of $\overline{R}$ such that $\pi(x_i) \to \pi(x)$. As π maps R homeomorphically onto $\pi(R)$, the point x must be in ∂R. By Theorem 6.5.9, there is an $r > 0$ such that

$$N(\overline{R}, r) \cap \Gamma x = \partial R \cap \Gamma x.$$

Moreover, there are only finitely many elements $g_1, \ldots, g_m$ of Γ such that

$$\partial R \cap \Gamma x = \{g_1 x, \ldots, g_m x\}.$$

By shrinking r, if necessary, we may assume that $\overline{B}(g_i x, r)$ is compact for each $i = 1, \ldots, m$. As $\pi(x_i) \to \pi(x)$, there is a $k > 0$ such that

$$\mathrm{dist}(\Gamma x_i, \Gamma x) < r$$

for all $i \geq k$. Hence, there is a h_i in Γ for each $i \geq k$ such that

$$d(x_i, h_i x) < r.$$

Now since

$$N(\overline{R}, r) \cap \Gamma x = \partial R \cap \Gamma x,$$

we have $h_i x = g_j x$ for some $j = 1, \ldots, m$. Hence x_i is in the compact set

$$\overline{B}(g_1 x, r) \cup \cdots \cup \overline{B}(g_m x, r)$$

for all $i \geq k$. But this implies that $\{x_i\}$ has a convergent subsequence, which is a contradiction. Thus $\overline{R}$ is compact. □

Theorem 6.5.11. *If R is a locally finite fundamental region for a group Γ of isometries of a connected metric space X, then Γ is generated by the set*

$$\Psi = \{g \in \Gamma : \overline{R} \cap g\overline{R} \textit{ is nonempty}\}.$$

Proof: Let H be the subgroup of Γ generated by Ψ, and let x be in X. Then there is a g in Γ such that gx is in $\overline{R}$. Suppose that h is another element of Γ such that hx is in $\overline{R}$. Then gx is in $\overline{R} \cap gh^{-1}\overline{R}$ and so gh^{-1} is in Ψ. Hence $\mathrm{H}g = \mathrm{H}h$. This implies that there is a well-defined function $\phi : X \to \Gamma/\mathrm{H}$, defined by $\phi(x) = \mathrm{H}g$, where gx is in $\overline{R}$.

As R is locally finite, there is an open ball B centered at x that meets only finitely many Γ-images of $\overline{R}$, say $g_1\overline{R}, \ldots, g_m\overline{R}$. We may assume that each $g_i\overline{R}$ contains x. Then we have

$$B \subset g_1\overline{R} \cup \cdots \cup g_m\overline{R}.$$

If y is in B, then y is in $g_i\overline{R}$ for some i, and we have

$$\phi(y) = \mathrm{H}g_i^{-1} = \phi(x).$$

Thus ϕ is constant on B. Hence, the fibers of the map ϕ are open. As X is connected, ϕ is constant.

Let g be in Γ, let u be in R, and let v be in $g^{-1}R$. Then

$$\mathrm{H} = \phi(u) = \phi(v) = \mathrm{H}g$$

and so g is in H. This shows that $\mathrm{H} = \Gamma$. Thus Ψ generates Γ. □

Rigid Metric Spaces

Definition: A metric space X is *rigid* if and only if the only similarity of X that fixes each point of a nonempty open subset of X is the identity map of X.

Theorem 6.5.12. *If X is a geodesically connected and geodesically complete metric space, then X is rigid.*

Proof: Let ϕ be a similarity of X that fixes each point of a nonempty open subset W of X. Then the scale factor of ϕ is one, and so ϕ is an isometry of X. Let w be a point of W and let x be an arbitrary point of X not equal to w. Then there is a geodesic line $\lambda : \mathbb{R} \to X$ whose image contains w and x. Observe that

$$\phi\lambda : \mathbb{R} \to X$$

is also a geodesic line and $\phi\lambda$ agrees with λ on the open set $\lambda^{-1}(W)$. As every geodesic arc in X extends to a unique geodesic line, we deduce that $\phi\lambda = \lambda$. Therefore $\phi(x) = x$. Hence $\phi = 1$. Thus X is rigid. □

Example: It follows from Theorem 6.5.12 that S^n, E^n, and H^n are rigid metric spaces.

Definition: A subset F of a metric space X is a *fundamental set* for a group Γ of isometries of X if and only if F contains exactly one point from each Γ-orbit in X.

Theorem 6.5.13. *An open subset R of a rigid metric space X is a fundamental region for a group Γ of isometries of X if and only if there is a fundamental set F for Γ such that $R \subset F \subset \overline{R}$.*

Proof: Suppose that R is a fundamental region for Γ. Then the members of $\{gR : g \in \Gamma\}$ are mutually disjoint. Therefore R contains at most one element from each Γ-orbit in X. Now since

$$X = \cup\{g\overline{R} : g \in \Gamma\},$$

there is a fundamental set F for Γ such that $R \subset F \subset \overline{R}$ by the axiom of choice.

Conversely, suppose there is a fundamental set F for the group Γ such that $R \subset F \subset \overline{R}$, and suppose that g, h are elements of Γ such that $gR \cap hR$ is nonempty. Then there are points x, y of R such that $gx = hy$. Hence $h^{-1}gx = y$. As x and y are in F, we deduce that $h^{-1}gx = x$. Therefore $h^{-1}g$ fixes each point of $R \cap g^{-1}hR$. As X is rigid, $h^{-1}g = 1$. Hence $g = h$. Thus, the members of $\{gR : g \in \Gamma\}$ are mutually disjoint.

Now as $F \subset \overline{R}$, we have

$$X = \bigcup_{g \in \Gamma} gF = \bigcup_{g \in \Gamma} g\overline{R}.$$

Thus R is a fundamental region for Γ. □

If R is a fundamental region for a group Γ of isometries of a metric space X, then the stabilizer of every point of R is trivial. We next consider an example of a discontinuous group of isometries of a metric space X such that every point of X is fixed by some $g \neq 1$ in Γ. Hence, this group does not have a fundamental region.

Example: Let X be the union of the x-axis and y-axis of E^2 and let

$$\Gamma = \{1, \rho, \sigma, \alpha\},$$

where ρ and σ are the reflections in the x-axis and y-axis, respectively, and α is the antipodal map. Then Γ is a discontinuous group of isometries of X, since Γ is finite. Observe that every point of X is fixed by a nonidentity element of Γ. Hence Γ has no fundamental region. Moreover X is not rigid.

Theorem 6.5.14. *Let Γ be a discontinuous group of isometries of a rigid metric space X. Then there is a point x of X whose stabilizer Γ_x is trivial.*

Proof: Since Γ is discontinuous, the stabilizer of each point of X is finite. Let x be a point of X such that the order of the stabilizer subgroup Γ_x is as small as possible. Let s be half the distance from x to $\Gamma x - \{x\}$. Then for each g in Γ, we have that $B(x, s)$ meets $B(gx, s)$ if and only if $gx = x$. Hence, for each point y in $B(y, s)$, we have that $\Gamma_y \subset \Gamma_x$ and therefore $\Gamma_y = \Gamma_x$ because of the minimality of the order of Γ_x. Hence, every point of $B(x, s)$ is fixed by every element of Γ_x. Therefore $\Gamma_x = \{1\}$, since X is rigid. □

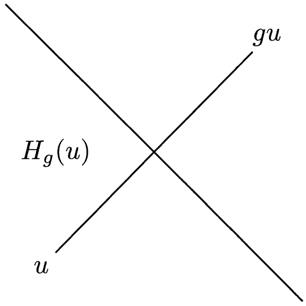


Figure 6.5.1. The half-space $H_g(u)$

Dirichlet Domains

Let Γ be a discontinuous group of isometries of a metric space X, and let u be a point of X whose stabilizer Γ_u is trivial. For each $g \neq 1$ in Γ, define

$$H_g(u) = \{x \in X : d(x,u) < d(x,gu)\}.$$

Observe that the set $H_g(u)$ is open in X. Moreover, if $X = S^n, E^n$, or H^n, then $H_g(u)$ is the open half-space of X containing u whose boundary is the perpendicular bisector of every geodesic segment joining u to gu. See Figure 6.5.1. The *Dirichlet domain* $D(u)$ for Γ, with *center* u, is either X if Γ is trivial or

$$D(u) = \cap\{H_g(u) : g \neq 1 \text{ in } \Gamma\}$$

if Γ is nontrivial.

Theorem 6.5.15. *Let $D(u)$ be the Dirichlet domain, with center u, for a discontinuous group Γ of isometries of a metric space X such that*

(1) *X is geodesically connected;*

(2) *X is geodesically complete;*

(3) *X is finitely compact.*

Then $D(u)$ is a locally finite fundamental domain for Γ.

Proof: This is clear if Γ is trivial, so assume that Γ is nontrivial. Let $r > 0$. Then $C(u,r)$ is compact. Hence $C(u,r)$ contains only finitely many points of an orbit Γx, since Γ is discontinuous. Let $K_g = X - H_g(u)$ for each $g \neq 1$ in Γ. Then K_g is closed in X. We next show that $\{K_g : g \neq 1 \text{ in } \Gamma\}$ is a locally finite family of sets in X. Suppose that $B(u,r)$ meets K_g in a point x. Then we have

$$\begin{aligned} d(u,gu) &\leq d(u,x) + d(x,gu) \\ &\leq d(u,x) + d(x,u) \\ &< 2r. \end{aligned}$$

Hence $B(u, 2r)$ contains gu. As $B(u, 2r)$ contains only finitely many points of Γu, the ball $B(u, r)$ meets only finitely many of the sets K_g. Therefore $\{K_g : g \neq 1 \text{ in } \Gamma\}$ is a locally finite family of closed sets in X. Hence

$$X - D(u) = \cup\{K_g : g \neq 1 \text{ in } \Gamma\}$$

is a closed set. Thus $D(u)$ is open.

From each orbit Γx, choose a point nearest to u and let F be the set of chosen points. Then F is a fundamental set for Γ. If x is in $D(u)$ and $g \neq 1$ in Γ, then

$$d(x, u) < d(x, gu) = d(g^{-1}x, u)$$

and so x is the unique nearest point of the orbit Γx to u. Thus $D(u) \subset F$.

Let x be an arbitrary point of F not equal to u and let $g \neq 1$ be in Γ. Then $d(x, u) \leq d(x, gu)$, since otherwise we would have

$$d(x, u) > d(x, gu) = d(g^{-1}x, u),$$

contrary to the assumption that x is in F. Let $[u, x]$ be a geodesic segment in X joining u to x. Let y be a point of the open segment (u, x). Then

$$\begin{aligned} d(y, u) &= d(x, u) - d(x, y) \\ &\leq d(x, gu) - d(x, y) \leq d(y, gu) \end{aligned}$$

with equality only if

$$d(x, gu) = d(x, y) + d(y, gu).$$

Suppose that we have equality. Let $[x, y]$ be the geodesic segment in $[x, u]$ joining x to y and let $[y, gu]$ be a geodesic segment in X joining y to gu. By Theorem 1.4.3, we have that $[x, y] \cup [y, gu]$ is a geodesic segment $[x, gu]$ in X joining x to gu. Now $[x, u]$ and $[x, gu]$ both extend $[x, y]$ and have the same length. Therefore $[x, u] = [x, gu]$, since X is geodesically complete. Hence $u = gu$, which is a contradiction. Therefore, we must have

$$d(y, u) < d(y, gu).$$

Hence y is in $H_g(u)$ for all $g \neq 1$ in Γ. Therefore y is in $D(u)$. Hence $[u, x) \subset D(u)$. Therefore x is in $\overline{D}(u)$. Hence $F \subset \overline{D}(u)$. Thus $D(u)$ is a fundamental region for Γ by Theorems 6.5.12 and 6.5.13. Moreover, if x is in $D(u)$, then $[u, x] \subset D(u)$, and so $D(u)$ is connected.

It remains only to show that $D(u)$ is locally finite. Suppose $r > 0$ and $B(u, r)$ meets $g\overline{D}(u)$. Then there is some x in $D(u)$ such that $d(u, gx) < r$. Moreover

$$\begin{aligned} d(u, gu) &\leq d(u, gx) + d(gx, gu) \\ &< r + d(x, u) \\ &\leq r + d(x, g^{-1}u) \\ &= r + d(gx, u) \\ &< 2r. \end{aligned}$$

But this is possible for only finitely many g. Thus $D(u)$ is locally finite. □

Theorem 6.5.16. *Let $D(u)$ be the Dirichlet domain, with center u, for a discontinuous group Γ of isometries of a metric space X such that*

(1) *X is geodesically connected;*

(2) *X is geodesically complete;*

(3) *X is finitely compact.*

Then

$$\overline{D}(u) = \{x \in X : x \textit{ is a nearest point of } \Gamma x \textit{ to } u\}.$$

Proof: This is clear if Γ is trivial, so assume that Γ is nontrivial. For each $g \neq 1$ in Γ, define

$$L_g = \{x \in X : d(x,u) \leq d(x,gu)\}.$$

Then L_g is a closed subset of X containing H_g. Now since

$$L_g = \{x \in X : d(x,u) \leq d(g^{-1}x,u)\},$$

we have

$$\cap\{L_g : g \neq 1 \text{ in } \Gamma\} = \{x \in X : x \text{ is a nearest point of } \Gamma x \text{ to } u\}.$$

Moreover, since

$$D(u) \quad = \quad \cap\{H_g(u) : g \neq 1 \text{ in } \Gamma\},$$

we have that

$$\overline{D}(u) \quad \subset \quad \cap\{L_g : g \neq 1 \text{ in } \Gamma\}.$$

Now suppose that x is a nearest point of Γx to u. Then we can choose a fundamental set F for Γ containing x such that each point of F is a nearest point in its orbit to u. From the proof of Theorem 6.5.15, we have that $F \subset \overline{D}(u)$. Thus x is in $\overline{D}(u)$. Therefore

$$\overline{D}(u) = \{x \in X : x \text{ is a nearest point of } \Gamma x \text{ to } u\}. \qquad \square$$

Exercise 6.5

1. Let R be a fundamental region for a group Γ of isometries of a metric space X and let $\hat{R}$ be the topological interior of $\overline{R}$. Prove that $\hat{R}$ is the largest fundamental region for Γ containing R.

2. Let R be a fundamental region for a group Γ of isometries of a metric space X. Prove that R is locally finite if and only if $\{gR : g \in \Gamma\}$ is a locally finite family of subsets of X.

3. Let Γ be a discontinuous group of isometries of a metric space X with a fundamental region R such that $\overline{R}$ is compact. Prove that

   (1) Γ is finitely generated, and

   (2) Γ has only finitely many conjugacy classes of elements with fixed points.

4. Let Γ be the subgroup of $\mathrm{I}(\mathbb{C})$ generated by $f : z \mapsto z + 1$ and $g : z \mapsto z + i$. Find a fundamental domain for Γ that is not locally finite.

5. Let Γ be a discontinuous group of isometries of a metric space X that has a fundamental region. Prove that the set of points of X that are not fixed by any $g \neq 1$ in Γ is an open dense subset of X.

6. Prove that the set $H_g(u)$ used in the definition of a Dirichlet domain is open.

7. Let $D(u)$ be a Dirichlet domain, with center u, for a group Γ as in Theorem 6.5.16. Prove that if x is in $\partial D(u)$, then $\partial D(u) \cap \Gamma x$ is a finite set of points that are all equidistant from u.

§6.6. Convex Fundamental Polyhedra

Throughout this section, $X = S^n, E^n$, or H^n with $n > 0$. Let Γ be a discrete group of isometries of X. By Theorem 6.5.14, there is a point u of X whose stabilizer Γ_u is trivial. Let $D(u)$ be the Dirichlet domain for Γ with center u. Then $D(u)$ is convex, since by definition $D(u)$ is either X or the intersection of open half-spaces of X. By Theorem 6.5.15, we have that $D(u)$ is a locally finite fundamental domain for Γ. Hence Γ has a convex, locally finite, fundamental domain.

Lemma 1. *If D is a convex, locally finite, fundamental domain for a discrete group Γ of isometries of X, then for each point x of ∂D, there is a $g \neq 1$ in Γ such that x is of $\overline{D} \cap g\overline{D}$.*

Proof: As D is locally finite, there is an $r > 0$ such that $B(x, r)$ meets only finitely many Γ-images of $\overline{D}$, say $g_1\overline{D}, \ldots, g_m\overline{D}$ with $g_1 = 1$. By shrinking r, if necessary, we may assume that x is in each $g_i\overline{D}$. As D is convex, $\partial D = \partial\overline{D}$. Therefore $B(x, r)$ contains a point not in $\overline{D}$. Hence $m > 1$. Thus, there is a $g \neq 1$ in D such that x is in $g\overline{D}$. □

Theorem 6.6.1. *If D is a convex, locally finite, fundamental domain for a discrete group Γ of isometries of X, then $\overline{D}$ is a convex polyhedron.*

Proof: Since D is convex in X, we have that $\overline{D}$ is closed and convex in X. Let $\mathcal{S}$ be the set of sides of D. We need to show that $\mathcal{S}$ is locally finite. Let x be an arbitrary point of X. If x is in D, then D is a neighborhood of x that meets no side of D. If x is in $X - \overline{D}$, then $X - \overline{D}$ is a neighborhood of x that meets no side of D. Hence, we may assume that x is in ∂D. As D is locally finite, there is an $r > 0$ such that $B(x, r)$ meets only finitely many Γ-images of $\overline{D}$, say $g_0\overline{D}, \ldots, g_m\overline{D}$ with $g_0 = 1$. By shrinking r, if necessary, we may assume that x is in each $g_i\overline{D}$. Now for each $i > 0$, we have that $\overline{D} \cap g_i\overline{D}$ is a nonempty convex subset of ∂D. By Theorem 6.2.6(1), there is a side S_i of D containing $\overline{D} \cap g_i\overline{D}$. By Lemma 1, we have

$$B(x,r) \cap \partial D \subset \bigcup_{i=1}^{m} (\overline{D} \cap g_i\overline{D}).$$

Therefore

$$B(x,r) \cap \partial D \subset S_1 \cup \cdots \cup S_m.$$

Now suppose that S is a side of D meeting $B(x,r)$. Then $B(x,r)$ meets S°, since $\overline{S^\circ} = S$. By Theorem 6.2.6(3), we have that $S = S_i$ for some i. Thus $B(x,r)$ meets only finitely many sides of D. Hence $\mathcal{S}$ is locally finite. Thus $\overline{D}$ is a convex polyhedron. □

Corollary 1. *Every convex, locally finite, fundamental domain for a discrete a group Γ of isometries of X is proper.*

Proof: Let D be a convex, locally finite, fundamental domain for Γ. Then the sides of D form a locally finite family of null sets in X. Hence ∂D is the union of a countable number of null sets and so is a null set. Thus D is proper. □

Fundamental Polyhedra

Definition: A *convex fundamental polyhedron* for a discrete group Γ of isometries of X is a convex polyhedron P in X whose interior is a locally finite fundamental domain for Γ.

Let Γ be a discrete group of isometries of X. By Theorem 6.6.1, the closure $\overline{D}$ of any convex, locally finite, fundamental domain D for Γ is a convex fundamental polyhedron for Γ. In particular, the closure $\overline{D}(u)$ of any Dirichlet domain $D(u)$ for Γ is a convex fundamental polyhedron for Γ, called the *Dirichlet polyhedron* for Γ with center u.

Example: Let $\Gamma = \mathrm{PSL}(2,\mathbb{Z})$ and regard Γ as a subgroup of $\mathrm{I}(U^2)$. Then Γ is discrete. Let T be the generalized hyperbolic triangle with vertices $\pm\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and ∞. See Figure 6.6.1. Then T is the Dirichlet polygon for Γ with center ti for any $t > 1$.

Let Γ be a discrete group of isometries of X and let u be a point of X whose stabilizer Γ_u is trivial. For each $g \neq 1$ in Γ, define

$$P_g(u) = \{x \in X : d(x,u) = d(x,gu)\}.$$

Then $P_g(u)$ is the unique hyperplane of X that bisects and is orthogonal to every geodesic segment in X joining u to gu.

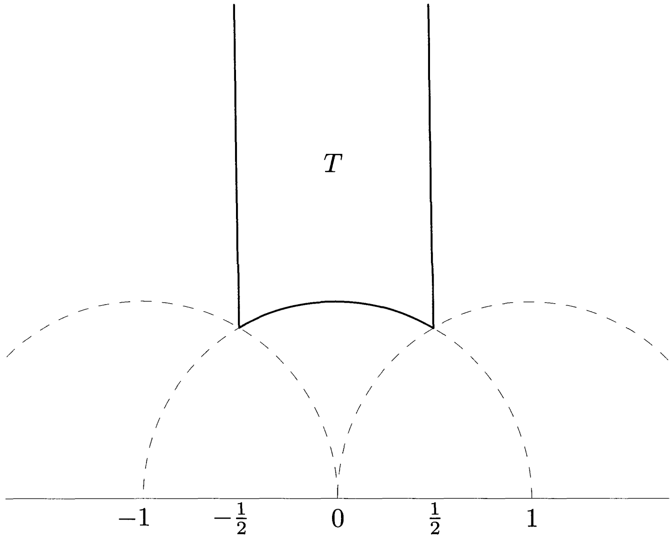


Figure 6.6.1. A Dirichlet polygon T for $\mathrm{PSL}(2,\mathbb{Z})$

Theorem 6.6.2. *Let S be a side of a Dirichlet domain $D(u)$, with center u, for a discrete group Γ of isometries of X. Then there is a unique element $g \neq 1$ of Γ that satisfies one (or all) of the following three properties:*

(1) $\langle S\rangle = P_g(u)$;

(2) $S = \overline{D}(u) \cap g\overline{D}(u)$;

(3) $g^{-1}S$ *is a side of* $D(u)$.

Proof: (1) Since

$$\partial D(u) \subset \cup\{P_g(u) : g \neq 1 \text{ in } \Gamma\},$$

we have that

$$S \subset \cup\{P_g(u) : g \neq 1 \text{ in } \Gamma\}.$$

Therefore

$$S = \cup\{S \cap P_g(u) : g \neq 1 \text{ in } \Gamma\}.$$

Now $S \cap P_g(u)$ is a closed convex subset of X for each $g \neq 1$ in Γ. As Γ is countable, we must have

$$\dim(S \cap P_g(u)) = n - 1$$

for some g; otherwise, the $(n-1)$-dimensional volume of S would be zero. Now since

$$\dim(S \cap P_g(u)) = n - 1$$

we have that $\langle S\rangle = P_g(u)$.

Let g, h be elements of Γ such that

$$P_g(u) = \langle S \rangle = P_h(u).$$

Since $P_g(u)$ is the perpendicular bisector of a geodesic segment from u to gu, we have that $gu = hu$. But u is fixed only by the identity element of Γ, and so $g = h$. Thus, there is a unique element g of Γ such that $\langle S \rangle = P_g(u)$.

(2) By (1) there is a unique element $g \neq 1$ of Γ such that $S \subset P_g(u)$. Let x be an arbitrary point of S. Then $d(x, u) = d(x, gu)$. By Theorem 6.5.16, we have that x is a nearest point of Γx to u. Now

$$d(g^{-1}x, u) = d(x, gu) = d(x, u).$$

Therefore $g^{-1}x$ is also a nearest point of Γx to u. Hence $g^{-1}x$ is in $\overline{D}(u)$ by Theorem 6.5.16. Therefore $g^{-1}S \subset \overline{D}(u)$. Hence

$$S \subset \overline{D}(u) \cap g\overline{D}(u).$$

But $\overline{D}(u) \cap g\overline{D}(u)$ is a convex subset of $\partial D(u)$. Therefore

$$S = \overline{D}(u) \cap g\overline{D}(u),$$

since S is a maximal convex subset of $\partial D(u)$.

Suppose that h is another nonidentity element of Γ such that

$$S = \overline{D}(u) \cap h\overline{D}(u).$$

Let x be an arbitrary point of S. Then $h^{-1}x$ is in $\overline{D}(u)$ and so

$$d(x, u) = d(h^{-1}x, u) = d(x, hu).$$

Hence x is in $P_h(u)$. Therefore $S \subset P_h(u)$. Hence $g = h$ by the uniqueness of g in (1). Thus, there is a unique $g \neq 1$ in Γ such that

$$S = \overline{D}(u) \cap g\overline{D}(u).$$

(3) By (2), there is unique element $g \neq 1$ of Γ such that

$$S = \overline{D}(u) \cap g\overline{D}(u).$$

Then we have

$$g^{-1}S = g^{-1}\overline{D}(u) \cap \overline{D}(u).$$

Therefore $g^{-1}S \subset \partial D(u)$. Hence, there is a side T of $\overline{D}(u)$ containing $g^{-1}S$. By (1) there is a unique element $h \neq 1$ of Γ such that

$$T = \overline{D}(u) \cap h\overline{D}(u).$$

Hence, we have

$$g^{-1}S \subset \overline{D}(u) \cap h\overline{D}(u)$$

and so we have

$$S \subset g\overline{D}(u) \cap gh\overline{D}(u).$$

Thus, we have

$$S \subset \overline{D}(u) \cap gh\overline{D}(u).$$

Suppose that $gh \neq 1$. We shall derive a contradiction. Since S is a maximal convex subset of $\partial D(u)$, we have

$$S = \overline{D}(u) \cap gh\overline{D}(u).$$

Then $gh = g$ by (2), and so $h = 1$, which is a contradiction. It follows that $gh = 1$ and so $h = g^{-1}$. Thus $g^{-1}S = T$.

Suppose that f is another nonidentity element of Γ such that $f^{-1}S$ is a side of $D(u)$. Then we have

$$f^{-1}S = \overline{D}(u) \cap f^{-1}\overline{D}(u)$$

and so we have

$$S = \overline{D}(u) \cap f\overline{D}(u).$$

Hence $f = g$ by (2). Thus, there is a unique element $g \neq 1$ of Γ such that $g^{-1}S$ is a side of $D(u)$. □

Definition: A convex fundamental polyhedron P for Γ is *exact* if and only if for each side S of P there is an element g of Γ such that $S = P \cap gP$.

It follows from Theorem 6.6.2(2) that every Dirichlet polyhedron for a discrete group is exact. Figure 6.6.2 illustrates an inexact, convex, fundamental polygon P for $\mathrm{PSL}(2,\mathbb{Z})$. The polygon P is inexact, since the two bounded sides of P are neither congruent nor left invariant by an element of $\mathrm{PSL}(2,\mathbb{Z})$. See Theorem 6.6.3.

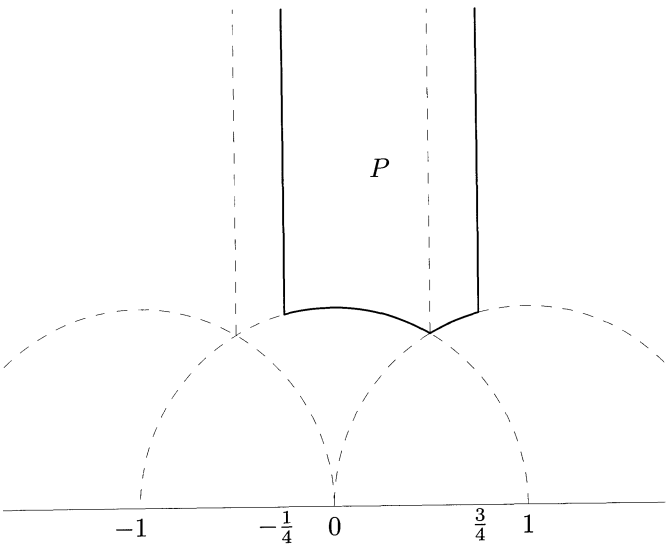


Figure 6.6.2. An inexact, convex, fundamental polygon P for $\mathrm{PSL}(2,\mathbb{Z})$

Theorem 6.6.3. *If S is a side of an exact, convex, fundamental polyhedron P for a discrete group Γ of isometries of X, then there is a unique element $g \neq 1$ of Γ such that*

$$S = P \cap gP;$$

moreover, $g^{-1}S$ is a side of P.

Proof: Since P is exact, there is an element g of Γ such that $S = P \cap gP$. Clearly $g \neq 1$. If $h \neq 1$ is another element of Γ such that $S = P \cap hP$, then gP° and hP° overlap; therefore $gP^\circ = hP^\circ$ and so $g = h$. Thus, there is a unique element $g \neq 1$ of Γ such that $S = P \cap gP$. The proof that $g^{-1}S$ is a side of P is the same as the proof of Theorem 6.6.2(3). □

Exercise 6.6

1. Let u, v be distinct points of X and let
$$P = \{x \in X : d(x, u) = d(x, v)\}.$$
Prove that P is the unique hyperplane of X that bisects and is orthogonal to every geodesic segment in X joining u to v.
2. Let Γ be the subgroup of $\mathrm{I}(\mathbb{C})$ generated by the translations of $\mathbb{C}$ by 1 and $\frac{1}{2} + \frac{\sqrt{3}}{2}i$. Determine the Dirichlet polygon of Γ with center 0 in $\mathbb{C}$.
3. Let T be the generalized hyperbolic triangle in Figure 6.6.1. Prove that T is the Dirichlet polygon for $\mathrm{PSL}(2, \mathbb{Z})$ with center ti for any $t > 1$.

§6.7. Tessellations

Throughout this section, $X = S^n, E^n$, or H^n with $n > 0$.

Definition: A *tessellation* of X is a collection $\mathcal{P}$ of n-dimensional convex polyhedra in X such that

(1) the interiors of the polyhedra in $\mathcal{P}$ are mutually disjoint;

(2) the union of the polyhedra in $\mathcal{P}$ is X; and

(3) the collection $\mathcal{P}$ is locally finite.

Definition: A tessellation $\mathcal{P}$ of X is *exact* if and only if each side S of a polyhedron P in $\mathcal{P}$ is a side of exactly two polyhedrons P and Q in $\mathcal{P}$.

An example of an exact tessellation is the grid pattern tessellation of E^2 by congruent squares. An example of an inexact tessellation is the familiar brick pattern tessellation of E^2 by congruent rectangles.

Definition: A *regular tessellation* of X is an exact tessellation of X consisting of congruent regular polytopes.

The three regular tessellations of the plane, by equilateral triangles, squares, and regular hexagons, have been known since antiquity. The five regular tessellations of the sphere induced by the five regular solids have been known since the Middle Ages. We are interested in tessellations of X by congruent polyhedra because of the following theorem.

Theorem 6.7.1. *Let P be an n-dimensional convex polyhedron in X and let Γ be a group of isometries of X. Then Γ is discrete and P is an (exact) convex fundamental polyhedron for Γ if and only if*

$$\mathcal{P} = \{gP : g \in \Gamma\}$$

is an (exact) tessellation of X.

Proof: Suppose that Γ is discrete and P is a convex fundamental polyhedron for Γ. Then P° is a locally finite fundamental domain for Γ. Hence, we have that

(1) the members of $\{gP^\circ : g \in \Gamma\}$ are mutually disjoint;

(2) $X = \cup\{gP : g \in \Gamma\}$; and

(3) the collection $\mathcal{P}$ is locally finite.

Thus $\mathcal{P}$ is a tessellation of X.

Now assume that P is exact. Let S be a side of P. Then there is a unique element of $g \neq 1$ of Γ such that $S = P \cap gP$; moreover $g^{-1}S$ is a side of P. Hence S is a side of gP. Therefore S is a side of exactly two polyhedrons P and gP of $\mathcal{P}$. As $\mathcal{P}$ is Γ-equivariant, the same is true for any side of any polyhedron in $\mathcal{P}$. Thus $\mathcal{P}$ is exact.

Conversely, suppose that $\mathcal{P}$ is a tessellation of X. Then

(1) the members of $\{gP^\circ : g \in \Gamma\}$ are mutually disjoint;

(2) $X = \cup\{gP : g \in \Gamma\}$; and

(3) the collection $\mathcal{P}$ is locally finite.

Hence P° is a a locally finite fundamental domain for Γ. Therefore Γ is discrete by Theorem 6.5.3 and P is a convex fundamental polyhedron for the group Γ.

Now assume that $\mathcal{P}$ is exact. Then for each side S of P, there is a g in Γ such that S is a side of gP. Hence $S \subset P \cap gP$. Since $P \cap gP \subset \partial P$ and S is a maximal convex subset of ∂P, we have that $S = P \cap gP$. Thus P is exact. $\square$

Definition: A collection $\mathcal{P}$ of n-dimension convex polyhedra in X is said to be *connected* if and only if for each pair P, Q in $\mathcal{P}$ there is a finite sequence $P_1, \ldots, P_m$ in $\mathcal{P}$ such that $P = P_1, P_m = Q$, and P_{i-1} and P_i share a common side for each $i > 1$.

Theorem 6.7.2. *Every exact tessellation of X is connected.*

Proof: The proof is by induction on the dimension n of X. The theorem is obviously true when $n = 1$, so assume that $n > 1$ and the theorem is true in dimension $n-1$. Let $\mathcal{P}$ be an exact tessellation of X and let P be a polyhedron in $\mathcal{P}$. Let U be the union of all the polyhedra Q in $\mathcal{P}$ for which there is a finite sequence $P_1, \ldots, P_m$ in $\mathcal{P}$ such that $P = P_1, P_m = Q$, and P_{i-1} and P_i share a common side for each $i > 1$. Then U is closed in X, since $\mathcal{P}$ is locally finite.

We now show that U is open in X. Let x be a point of U. Choose r such that $0 < r < \pi/2$ and $C(x, r)$ meets only the polyhedra of $\mathcal{P}$ containing x. Let Q be a polyhedron in $\mathcal{P}$ containing x. Then r is less than the distance from x to any side of Q not containing x. By Theorem 6.3.19, the set $Q \cap S(x,r)$ is an $(n-1)$-dimensional convex polyhedron in $S(x,r)$; moreover, if $\mathcal{S}(x)$ is the set of sides of Q containing x, then $\{T \cap S(x,r) : T \in \mathcal{S}(x)\}$ is the set of sides of $Q \cap S(x,r)$. Therefore $\mathcal{P}$ restricts to an exact tessellation $\mathcal{T}$ of $S(x,r)$. By the induction hypothesis, $\mathcal{T}$ is connected. Consequently, each polyhedron in $\mathcal{P}$ containing x is contained in U. Therefore U contains $B(x,r)$. Thus U is both open and closed in X. As X is connected, $U = X$. Thus $\mathcal{P}$ is connected. □

Theorem 6.7.3. *Let P be an exact, convex, fundamental polyhedron for a discrete group Γ of isometries of X. Then Γ is generated by the set*

$$\Phi = \{g \in \Gamma : P \cap gP \text{ is a side of } P\}.$$

Proof: By Theorem 6.7.1, we have that $\mathcal{P} = \{gP : g \in \Gamma\}$ is an exact tessellation of X. By Theorem 6.7.2, the tessellation $\mathcal{P}$ is connected. Let g be an arbitrary element of Γ. Then there is a finite sequence of elements $g_1, \ldots, g_m$ of Γ with $P = g_1P$, $g_mP = gP$, and $g_{i-1}P$ and g_iP share a common side for each $i > 1$. This implies that $g_1 = 1$, $g_m = g$, and P and $g_{i-1}^{-1}g_iP$ share a common side for each $i > 1$. We may assume that $g_{i-1} \neq g_i$ for each $i > 1$. Then $g_{i-1}^{-1}g_i$ is in Φ for each $i > 1$. As $g = g_1(g_1^{-1}g_2)\cdots(g_{m-1}^{-1}g_m)$, we have that Φ generates Γ. □

Theorem 6.7.4. *If a discrete group Γ of isometries of X has a finite-sided, exact, convex, fundamental polyhedron P, then Γ is finitely generated.*

Proof: By Theorem 6.6.3, the set of sides $\mathcal{S}$ of P is in one-to-one correspondence with the set $\Phi = \{g \in \Gamma : P \cap gP \in \mathcal{S}\}$. Therefore Φ is finite and so Γ is finitely generated by Theorem 6.7.3. □

Side-Pairing

Let S be a side of an exact, convex, fundamental polyhedron P for a discrete group Γ of isometries of X. By Theorem 6.6.3, there is a unique element g_S of Γ such that

$$S = P \cap g_S(P). \tag{6.7.1}$$

Furthermore $S' = g_S^{-1}(S)$ is a side of P. The side S' is said to be *paired to* the side S by the element g_S of Γ. As

$$S' = P \cap g_S^{-1}(P),$$

we have that $g_{S'} = g_S^{-1}$. Therefore S is paired to S' by g_S^{-1} and $S'' = S$. The Γ-*side-pairing* of P is defined to be the set

$$\Phi = \{g_S : S \text{ is a side of } P\}.$$

The elements of Φ are called the *side-pairing transformations* of P.

Two points x, x' of P are said to be *paired* by Φ, written $x \simeq x'$, if and only if there is a side S of P such that x is in S, x' is in S', and $g_S(x') = x$. If $g_S(x') = x$, then $g_{S'}(x) = x'$. Therefore $x \simeq x'$ if and only if $x' \simeq x$. Two points x, y of P are said to be *related* by Φ, written $x \sim y$, if either $x = y$ or there is a finite sequence $x_1, \ldots, x_m$ of points of P such that

$$x = x_1 \simeq x_2 \simeq \cdots \simeq x_m = y.$$

Being related by Φ is obviously an equivalence relation on the set P. The equivalence classes of P are called the *cycles* of Φ. If x is in P, we denote the cycle of Φ containing x by $[x]$.

Theorem 6.7.5. *If P is an exact, convex, fundamental polyhedron for a discrete group Γ of isometries of X, then for each point x of P,*

(1) *the cycle $[x]$ is finite, and*

(2) *$[x] = P \cap \Gamma x$.*

Proof: (1) It is clear from the definition of a cycle that $[x] \subset P \cap \Gamma x$. Hence $[x]$ is finite by Theorem 6.5.9.

(2) Let y be in $P \cap \Gamma x$. Then there is an f in Γ such that $y = fx$. Hence x is in $f^{-1}P$. As P is locally finite, there is an $r > 0$ such that $B(x, r)$ meets only finitely many Γ-images of P, say $g_1 P, \ldots, g_m P$. By shrinking r, we may assume that x is in $g_i P$ for each i. By shrinking r still further, we may assume that $r < \pi/2$ and r is less than the distance from x to any side of $g_i P$ not containing x. Now for each i, the set $g_i P \cap S(x, r)$ is an $(n-1)$-dimensional convex polyhedron in the sphere $S(x, r)$ by Theorem 6.3.19. Moreover

$$\mathcal{T} = \{g_i P \cap S(x, r) : i = 1, \ldots, m\}$$

is an exact tessellation of $S(x, r)$. By Theorem 6.7.2, the tessellation $\mathcal{T}$ is connected. Hence, there are elements $f_1, \ldots, f_\ell$ of Γ such that x is in $f_i^{-1}P$

for each i, and $P = f_1^{-1}P$, $f^{-1}P = f_\ell^{-1}P$, and $f_{i-1}^{-1}P$ and $f_i^{-1}P$ share a common side for each $i > 1$. This implies that $f_1 = 1$, $f_\ell = f$, and P and $f_{i-1}f_i^{-1}P$ share a common side S_i for each $i > 1$. We may assume that $i > 1$ and $f_{i-1} \neq f_i$ for each $i > 1$. Then $f_{i-1}f_i^{-1} = g_{S_i}$ for each $i > 1$. Let $x_1 = x$ and $x_i = f_i x$ for each $i > 1$. As x is in $f_i^{-1}P$, we have that $f_i x$ is in P. Hence x_i is in P for each i. Now

$$g_{S_i}(x_i) = f_{i-1}f_i^{-1}(x_i) = f_{i-1}x = x_{i-1}.$$

Hence x_{i-1} is in $P \cap g_{S_i}(P)$. Therefore x_{i-1} is in S_i and x_i is in S_i' for each $i > 1$. Hence, we have

$$x = x_1 \simeq x_2 \simeq \cdots \simeq x_\ell = y.$$

Therefore $x \sim y$. Thus $[x] = P \cap \Gamma x$. □

Dihedral Angles

Let P be an n-dimensional convex polyhedron in X. Sides S and T of P are said to be *adjacent* if and only if either $X = S^1$ and S, T are the sides of P or $n > 1$ and $S \cap T$ is a side of both S and T. In particular, the one side of a semicircle in S^1 is adjacent to itself.

Let S and T be sides of P. We now define the *dihedral angle* $\theta(S,T)$ of P between S and T. First of all, if $S = T$, then $\theta(S,T)$ is defined to be π. If S and T are distinct, nonadjacent sides of P, then $\theta(S,T)$ is defined to be 0. Now assume that S and T are adjacent. If $X = S^1$, then $\theta(S,T)$ is defined to be the angle between the endpoints of P.

Next assume that $n > 1$. Then the hyperplanes $\langle S\rangle$ and $\langle T\rangle$ subdivide X into four regions, one of which contains P; moreover,

$$\langle S\rangle \cap \langle T\rangle = \langle S \cap T\rangle.$$

Let x be any point in $S \cap T$ and let $\lambda, \mu : \mathbb{R} \to X$ be geodesic lines such that

(1) $\lambda(0) = x = \mu(0)$;

(2) λ and μ are normal to $\langle S\rangle$ and $\langle T\rangle$, respectively; and

(3) $\lambda'(0)$ and $\mu'(0)$ are directed away from the respective half-spaces of X containing P.

Let α be the angle between λ and μ at the point x. Clearly α does not depend on the choice of x. *The dihedral angle* of P between S and T is defined to be the angle

$$\theta(S,T) = \pi - \alpha. \tag{6.7.2}$$

See Figure 6.7.1. Note that as $0 < \alpha < \pi$, we have

$$0 < \theta(S,T) < \pi.$$

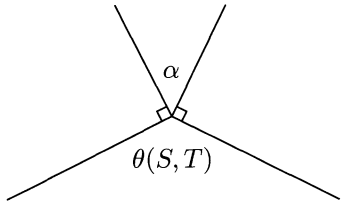


Figure 6.7.1. The dihedral angle $\theta(S,T)$ between adjacent sides

In general, we have that

$$0 \leq \theta(S,T) \leq \pi.$$

The dihedral angle $\theta(S,T)$ is said to be *proper* if and only if

$$0 < \theta(S,T) < \pi.$$

Note that $\theta(S,T)$ is proper if and only if S and T are distinct adjacent sides of P.

Cycles of Polyhedra

Definition: A *cycle of polyhedra* in X is a finite set

$$\mathcal{C} = \{P_0, \ldots, P_{m-1}\}$$

of n-dimensional convex polyhedra in X such that for each i (mod m),

(1) there are adjacent sides S_i and S_{i+1} of P_i such that $P_i \cap P_{i+1} = S_{i+1}$;

(2) $\sum\limits_{i=0}^{m-1} \theta(S_i, S_{i+1}) = 2\pi$; and

(3) if $n > 1$, then $R = \bigcap\limits_{i=0}^{m-1} P_i$ is a side of S_i for each i.

See Figure 6.7.2. Note that a collection $\mathcal{C}$ of geodesic segments in S^1 is a cycle if and only if $\mathcal{C}$ is a tessellation of S^1.

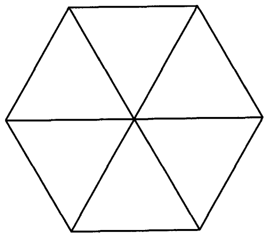


Figure 6.7.2. A cycle of equilateral triangles in E^2

Theorem 6.7.6. *Let R be a ridge of a polyhedron P in an exact tessellation $\mathcal{P}$ of X. Then the set of all polyhedra in $\mathcal{P}$ containing R forms a cycle whose intersection is R.*

Proof: Let S be one of the two sides of P containing R. We inductively define sequences

$$P_0, P_1, \ldots \quad \text{and} \quad S_0, S_1, \ldots$$

such that for each i,

(1) P_i is in $\mathcal{P}$ and S_i is a side of P_i;

(2) $P_0 = P$ and $S_0 = S$;

(3) R is a side of S_i;

(4) S_i and S_{i+1} are adjacent sides of P_i; and

(5) $P_i \cap P_{i+1} = S_{i+1}$.

Now R is contained in only finitely many polyhedra in $\mathcal{P}$, since $\mathcal{P}$ is locally finite. Hence, the sequence $\{P_i\}$ involves only finitely many distinct polyhedra. Evidently, the terms $P_0, P_1, \ldots, P_{k-1}$ are distinct if

$$\sum_{i=0}^{k-1} \theta(S_i, S_{i+1}) \leq 2\pi.$$

Hence, the first repetition of the sequence occurs at the first polyhedron P_m such that

$$\sum_{i=0}^{m} \theta(S_i, S_{i+1}) > 2\pi.$$

Clearly P_m intersects the interior of P_0 and so $P_m = P_0$. Hence $S_m = S_0$ and

$$\sum_{i=0}^{m-1} \theta(S_i, S_{i+1}) = 2\pi.$$

Now as

$$R = S_i \cap S_{i+1} \quad \text{for each } i,$$

we have that

$$R = \bigcap_{i=0}^{m-1} P_i.$$

Therefore $\{P_0, \ldots, P_{m-1}\}$ is a cycle of polyhedra whose intersection is R.

Let Q be any polyhedron in $\mathcal{P}$ containing R. Then clearly Q meets the interior of $\cup_{i=0}^{m-1} P_i$. This implies that Q meets the interior of P_i for some i, whence $Q = P_i$. Thus $\{P_0, \ldots, P_{m-1}\}$ is the set of polyhedra in $\mathcal{P}$ containing R. $\square$

Cycle Relations

Let P be an exact, convex, fundamental polyhedron for a discrete group Γ of isometries of X. We next consider certain relations in Γ that can be derived from the ridges and sides of P.

Let R be a side of a side S of P. Define a sequence $\{S_i\}_{i=1}^\infty$ of sides of P inductively as follows:

(1) Let $S_1 = S$.

(2) Let S_2 be the side of P adjacent to S_1' such that $g_{S_1}(S_1' \cap S_2) = R$.

(3) Let S_{i+1} be the side of P adjacent to S_i' such that

$$g_{S_i}(S_i' \cap S_{i+1}) = S_{i-1}' \cap S_i \quad \text{for each } i > 1.$$

We call $\{S_i\}_{i=1}^\infty$ the *sequence of sides* of P determined by R and S.

Theorem 6.7.7. *Let R be a side of a side S of an exact, convex, fundamental polyhedron P for a discrete group Γ of isometries of X, and let $\{S_i\}_{i=1}^\infty$ be the sequence of sides of P determined by R and S. Then there is a least positive integer ℓ and a positive integer k such that*

(1) *$S_{i+\ell} = S_i$ for each i,*

(2) *$\sum_{i=1}^{\ell} \theta(S_i', S_{i+1}) = 2\pi/k$, and*

(3) *the element $g_{S_1} g_{S_2} \cdots g_{S_\ell}$ has order k.*

Proof: Define a sequence $\{g_i\}_{i=0}^\infty$ of elements of Γ by $g_0 = 1$ and

$$g_i = g_{S_1} g_{S_2} \cdots g_{S_i} \quad \text{for each } i > 0.$$

We now prove that $\{g_i P\}_{i=0}^\infty$ forms a cycle of polyhedra in X. As S_i' and S_{i+1} are adjacent sides of P for each i, we have that $g_i S_i'$ and $g_i S_{i+1}$ are adjacent sides of $g_i P$ for each i; moreover,

$$g_i P \cap g_{i+1} P = g_i(P \cap g_{S_{i+1}} P) = g_i S_{i+1}$$

and $g_i S_{i+1} = g_{i+1} S_{i+1}'$ for each $i > 1$.

Now for each $i > 0$, we have

$$\begin{aligned} g_i S_{i+1} \cap g_{i+1} S_{i+2} &= g_{i+1} S_{i+1}' \cap g_{i+1} S_{i+2} \\ &= g_{i+1}(S_{i+1}' \cap S_{i+2}) \\ &= g_i(S_i' \cap S_{i+1}) \\ &= g_{i-1} S_i \cap g_i S_{i+1}. \end{aligned}$$

Therefore, we have

$$\bigcap_{i=0}^{\infty} g_i P = S_1 \cap g_{S_1}(S_2) = R.$$

By Theorem 6.7.6, there is an integer $m > 0$ such that $\{g_i P\}_{i=0}^{m-1}$ is a cycle of polyhedra. Hence $g_{i+m} P = g_i P$ for each i, and so $g_{i+m} = g_i$ for each i.

Now since

$$\begin{aligned} g_{i-1}S_{i+m} &= g_{i+m-1}S_{i+m} \\ &= g_{i+m-1}P \cap g_{i+m}P \\ &= g_{i-1}P \cap g_iP \quad = \quad g_{i-1}S_i, \end{aligned}$$

we find that $S_{i+m} = S_i$ for each i.

Let ℓ be the least positive integer such that $S_{i+\ell} = S_i$ for each i. Then $k = m/\ell$ is a positive integer. As

$$\sum_{i=1}^{m} \theta(g_iS_i', g_iS_{i+1}) = 2\pi,$$

we have that

$$k\sum_{i=1}^{\ell} \theta(S_i', S_{i+1}) = 2\pi.$$

Moreover, as $g_m = 1$, we have that $g_\ell^k = 1$, and since $g_j \neq 1$ for $1 < j < m$, we deduce that k is the order of g_ℓ. □

Let R be a side of a side S of an exact, convex, fundamental polyhedron P for a discrete group Γ of isometries of X, and let $\{S_i\}_{i=1}^{\infty}$ be the sequence of sides of P determined by R and S. By Theorem 6.7.7, there is a least positive integer ℓ such that $S_{i+\ell} = S_i$ for each i. The finite sequence $\{S_i\}_{i=1}^{\ell}$ is called the *cycle of sides* of P determined by R and S. The element $g_{S_1}g_{S_2}\cdots g_{S_\ell}$ of Γ is called the *cycle transformation* of the cycle of sides $\{S_i\}_{i=1}^{\ell}$. By Theorem 6.7.7, the cycle transformation $g_{S_1}g_{S_2}\cdots g_{S_\ell}$ has finite order k. The relation

$$(g_{S_1}g_{S_2}\cdots g_{S_\ell})^k = 1 \qquad (6.7.3)$$

in Γ is called the *cycle relation* of Γ determined by the cycle of sides $\{S_i\}_{i=1}^{\ell}$. For each side S of P, the relation

$$g_S g_{S'} = 1 \qquad (6.7.4)$$

is called the *side-pairing relation* determined by the side S.

Remark: The cycle relations together with the side-pairing relations form a complete set of relations for the generators

$$\Phi = \{g_S : S \text{ is a side of } P\}$$

of the group Γ; that is, any relation among the generators Φ can be derived from these relations. For a proof, see §13.5.

Example: Let L, S, R be the three sides occurring left to right in the Dirichlet polygon T for $\mathrm{PSL}(2,\mathbb{Z})$ in Figure 6.6.1. Then

$$g_R(z) = z + 1 \quad \text{and} \quad g_S(z) = -1/z.$$

Hence $R' = L$, $S' = S$, and $L' = R$. Observe that $\{S, R\}$ is a cycle of sides of T whose cycle transformation g_Sg_R has order three. Moreover g_S has order two. The relations $(g_Sg_R)^3 = 1$ and $g_S^2 = 1$ form a complete set of relations for the generators $\{g_S, g_R\}$ of $\mathrm{PSL}(2,\mathbb{Z})$.

Exercise 6.7

1. Let S be a side of an exact, convex, fundamental polyhedron P for Γ. Show that $S' = S$ if and only if g_S has order two in Γ.

2. Let $\{S_i\}_{i=1}^{\ell}$ be a cycle of sides of an exact, convex, fundamental polyhedron P for Γ. Show that the cycle transformation $g_{S_1} \cdots g_{S_\ell}$ leaves $S'_\ell \cap S_1$ invariant.

3. Furthermore, if $X = E^n$ or H^n, with $n > 1$, prove that $g_{S_1} \cdots g_{S_\ell}$ fixes a point of $S'_\ell \cap S_1$.

4. Let Γ be the discrete group of isometries of E^2 generated by the translations of E^2 by e_1 and e_2. Then $P = [0,1]^2$ is an exact, convex, fundamental polygon for Γ. Find all the cycles of sides of P and the corresponding cycle relations of Γ.

5. Let P be an exact, convex, fundamental polyhedron for Γ with only finitely many sides. Prove that P has only finitely many cycles of sides.

6. Let R be a ridge of an exact, convex, fundamental polyhedron P for Γ and let S and T be the two sides of P such that $R = S \cap T$. Let $\{S_i\}_{i=1}^{\ell}$ be the cycle of sides of P determined by R and S. Show that $\{S'_{\ell-i}\}_{i=0}^{\ell-1}$ is the cycle of sides P determined by R and T. Conclude that the pair consisting of the cycle transformation $g_{S_1} \cdots g_{S_\ell}$ and its inverse depends only on R.

7. Let R be a side of a side S of an exact, convex, fundamental polyhedron P for Γ and let R' be the side of S' such that $g_S(R') = R$. Let $\{S_i\}_{i=1}^{\ell}$ be the cycle of sides of P determined by R and S. Show that $\{S_2, \ldots, S_\ell, S_1\}$ is the cycle of sides of P determined by R' and S_2. Conclude that the cycle transformation of $\{S_2, \ldots, S_\ell, S_1\}$ determined by R' and S_2 is conjugate in Γ to the cycle transformation of $\{S_i\}_{i=1}^{\ell}$ determined by R and S.

§6.8. Historical Notes

§6.1. All the essential material in §6.1 appeared in Beltrami's 1868 papers *Saggio di interpetrazione della geometria non-euclidea* [38] and *Teoria fondamentale degli spazii di curvatura costante* [39]. See also Klein's 1871-73 paper *Ueber die sogenannte Nicht-Euklidische Geometrie* [224], [227].

§6.2. Convex curves and surfaces were defined by Archimedes in his third century B.C. treatise *On the sphere and cylinder* [23]. Convex sets in Euclidean n-space were first studied systematically by Minkowski. See, for example, his 1911 treatise *Theorie der konvexen Körper, insbesondere Begründung ihres Oberflächenbegriffs* [295]. The Euclidean versions of Theorems 6.2.1-6.2.3 were proved by Steinitz in his 1913-16 paper *Bedingt konvergente Reihen und konvexe Systeme* [380], [381], [382]. For a survey of convexity theory, see Berger's 1990 article *Convexity* [42]. References for the theory of convex sets are Grünbaum's 1967 text *Convex Polytopes* [172] and Brøndsted's 1983 text *An Introduction to Convex Polytopes* [59].

§6.3. Convex polyhedra in hyperbolic 3-space were defined by Poincaré in his 1881 note *Sur les groupes kleinéens* [329]. General polyhedra in Euclidean n-space were studied by Klee in his 1959 paper *Some characterizations of convex polyhedra* [223]. General polyhedra in hyperbolic n-space were considered by Andreev in his 1970 paper *Intersection of plane boundaries of a polytope with acute angles* [15]. Theorem 6.3.26 appeared in Vinberg's 1967 paper *Discrete groups generated by reflections in Lobacevskii spaces* [397].

§6.4. Euclidean polygons and the regular solids were thoroughly studied in Euclid's *Elements* [118]. General 3-dimensional Euclidean polytopes were first studied by Descartes in his seventeenth century manuscript *De solidorum elementis* [105], which was not published until 1860. General 3-dimensional Euclidean polytopes were studied by Euler in his 1758 paper *Elementa doctrinae solidorum* [121]. In particular, Euler introduced the concept of an *edge* of a polyhedron in this paper. Polytopes in Euclidean n-space and spherical n-space were first studied by Schläfli in his 1852 treatise *Theorie der vielfachen Kontinuität* [362], which was published posthumously in 1901. In particular, Schläfli classified all the regular Euclidean and spherical polytopes in this treatise. The most important results of Schläfli's treatise were published in his 1855 paper *Réduction d'une intégrale multiple, qui comprend l'arc de cercle et l'aire du triangle sphérique comme cas particuliers* [359] and in his 1858-60 paper *On the multiple integral* $\int dxdy \cdots dz$ [360], [361]. Convex polytopes in hyperbolic n-space were considered by Dehn in his 1905 paper *Die Eulersche Formel im Zusammenhang mit dem Inhalt in der Nicht-Euklidischen Geometrie* [101]. For a characterization of 3-dimensional hyperbolic polytopes, see Hodgson, Rivin, and Smith's 1992 paper *A characterization of convex hyperbolic polyhedra and of convex polyhedra inscribed in the sphere* [197] and Hodgson and Rivin's 1993 paper *A characterization of compact convex polyhedra in hyperbolic 3-space* [196]. References for the theory of convex polytopes are Grünbaum's 1967 text [172], Coxeter's 1973 treatise *Regular Polytopes* [92], and Brøndsted's 1983 text [59].

§6.5. The concept of a *fundamental region* arose in the theory of lattices. For example, Gauss spoke of an elementary parallelogram of a plane lattice in his 1831 review [149] of a treatise on quadratic forms. The concept of a fundamental region for a Fuchsian group was introduced by Poincaré in his 1881 note *Sur les fonctions fuchsiennes* [327]. See also Klein's 1883 paper *Neue Beiträge zur Riemannschen Funktionentheorie* [233]. Theorem 6.5.5 was essentially proved by Siegel in his 1943 paper *Discontinuous groups* [375]. Moreover, the concept of a *locally finite fundamental region* was introduced by Siegel in this paper. The 2-dimensional version of Theorem 6.5.7 was proved by Klein in his 1883 paper [233]. Theorem 6.5.8 appeared in Beardon's 1974 paper *Fundamental domains for Kleinian groups* [33]. Theorem 6.5.11 was essentially proved by Siegel in his 1943 paper [375]. The *Dirichlet domain* of a plane lattice was introduced by Dirichlet in his 1850 paper *Über die Reduction der positiven quadratischen Formen* [107]. Theorem 6.5.15 appeared in Busemann's 1948 paper *Spaces with non-positive curvature* [62]. For the theory of fundamental regions of Fuchsian groups, see Beardon's 1983 text *The Geometry of Discrete Groups* [34].

§6.6. According to Klein's historical study *Development of Mathematics in the 19th Century* [238], Gauss determined the fundamental polygon for the elliptic modular group in Figure 6.6.1. This fundamental polygon was described by Dedekind in his 1877 paper *Schreiben an Herrn Borchardt über die Theorie der elliptischen Modulfunktionen* [100]. The term *fundamental polygon* was introduced by Klein for certain subgroups of the elliptic modular group in his 1879 paper *Ueber die Transformation der elliptischen Functionen* [231]. The notion of a fundamental polygon was extended to all Fuchsian groups by Poincaré in his 1881 note [327]. See also Dyck's 1882 paper *Gruppentheoretische Studien* [111]. Fundamental polyhedra for Kleinian groups were introduced by Poincaré in his 1881 note [329]. The 2-dimensional version of Theorem 6.6.1 was proved by Beardon in his 1983 text [34]. Theorem 6.6.1 for dimension $n > 2$ seems to be new.

§6.7. The three regular tessellations of the plane were discovered by the Pythagoreans according to Heath's 1921 treatise *A History of Greek Mathematics* [186]. The five regular tessellations of the sphere were described by Abu l-Wafa in the 10th century according to a manuscript reported by Woepcke in his 1855 article *Recherches sur l'histoire des sciences mathématics chez les orientaux, d'après des traités inédits arabes et persans* [415]. For the classification of the regular tessellations of S^n, E^n, and H^n, see Coxeter's 1973 treatise *Regular Polytopes* [92] and Coxeter's 1956 paper *Regular honeycombs in hyperbolic space* [90]. The general notion of a tessellation of the hyperbolic plane generated by a fundamental polygon appeared in Poincaré's 1881 note [327]. The concepts of *side-pairing transformation* and *cycle of vertices* determined by a fundamental polygon for a Fuchsian group were introduced by Poincaré in his 1881 note *Sur les fonctions fuchsiennes* [328]. See also his 1882 paper *Théorie des groupes fuchsiens* [330]. Tessellations of hyperbolic space generated by a fundamental polyhedron were considered by Poincaré in his 1883 *Mémoire sur les groupes kleinéens* [332].

CHAPTER 7

Classical Discrete Groups

In this chapter, we study classical discrete groups of isometries of S^n, E^n, and H^n. We begin with the theory of discrete reflection groups. In Section 7.4, we study the theory of crystallographic groups. The chapter ends with a proof of Selberg's lemma.

§7.1. Reflection Groups

Throughout this section, $X = S^n, E^n$, or H^n with $n > 0$.

Lemma 1. *Let x be a point inside a horosphere Σ of H^n. Then the shortest distance from x to Σ is along the unique hyperbolic line passing through x Lorentz orthogonal to Σ.*

Proof: We pass to the conformal ball model B^n of hyperbolic space and move x to the origin. Then the shortest distance from 0 to Σ is obviously along the unique diameter of B^n orthogonal to Σ. See Figure 7.1.1. □

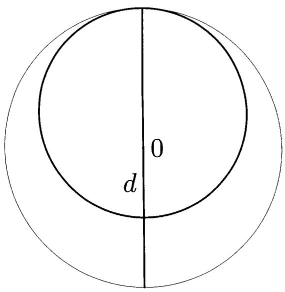


Figure 7.1.1. The shortest distance d from the origin to a horocycle of B^2

Let S be a side of an n-dimensional convex polyhedron P in X. The reflection of X in the side S of P is the reflection of X in the hyperplane $\langle S\rangle$ spanned by S.

Theorem 7.1.1. *Let G be the group generated by the reflections of X in the sides of a finite-sided, n-dimensional, convex polyhedron P in X of finite volume. Then*

$$X = \cup\{gP : g \in G\}.$$

Proof: The proof is by induction on the dimension n. The theorem is obviously true when $n = 1$, so assume that $n > 1$ and the theorem is true in dimension $n-1$. Let x be a point of P and let $G(x)$ be the subgroup of G generated by all the reflections of X in the sides of P that contain x. Let $r(x)$ be a real number such that $0 < r(x) < \pi/2$ and the ball $C(x, r(x))$ meets only the sides of P containing x. By Theorem 6.3.19, the set $P \cap S(x, r(x))$ is an $(n-1)$-dimensional, convex polyhedron in the sphere $S(x, r(x))$. From the induction hypothesis, we have

$$S(x, r(x)) = \cup\big\{g(P \cap S(x, r(x))) : g \in G(x)\big\}.$$

Now since P is convex, we deduce that

$$B(x, r(x)) \subset \cup\{gP : g \in G(x)\}.$$

By Theorems 6.3.25 and 6.3.26, the polyhedron P has only finitely many ideal vertices, say $v_1, \ldots, v_m$. For each i, let B_i be a horoball based at v_i such that $\overline{B}_i$ meets only the sides of P incident with v_i. For each i, let G_i be the subgroup of G generated by all the reflections of X in the sides of P that are incident with v_i. By Theorem 6.3.23, the set $P \cap \partial B_i$ is a compact, Euclidean, $(n-1)$-dimensional, convex polyhedron in the horosphere ∂B_i. We deduce from the induction hypothesis that

$$B_i \subset \cup\{gP : g \in G_i\}.$$

By Lemma 1, there is a horoball B_i' based at v_i such that $B_i' \subset B_i$ and $\mathrm{dist}(B_i', \partial B_i) = 1$ for each i. Set

$$P_0 = P - \bigcup_{i=1}^{m} B_i'.$$

Then P_0 is compact by Theorem 6.3.26. Let $\ell > 0$ be a Lebesgue number for the open cover $\{B(x, r(x)) : x \in P_0\}$ of P_0 such that $\ell < 1$. Let

$$U = \cup\{gP : g \in G\}.$$

We claim that $N(P, \ell) \subset U$. Observe that $N(P_0, \ell) \subset U$. Let x be a point of $P \cap B_i'$. Then we have

$$B(x, \ell) \subset B_i \subset U.$$

Hence $N(B_i', \ell) \subset U$ for each i. Therefore $N(P, \ell) \subset U$ as claimed. Now as U is G-equivariant, we deduce that $N(gP, \ell) \subset U$ for each g in G. Therefore $N(U, \ell) \subset U$, and so $U = X$. □

Let P be an exact, convex, fundamental polyhedron for a discrete group Γ of isometries of X. Then for each side S of P, there is a unique element g_S of Γ such that

$$S = P \cap g_S(P).$$

The group Γ is defined to be a *discrete reflection group*, with respect to the polyhedron P, if and only if g_S is the reflection of X in the hyperplane $\langle S \rangle$ for each side S of P.

Definition: An angle α is a *submultiple* of an angle β if and only if there is a positive integer k such that $\alpha = \beta/k$ or $\alpha = \beta/\infty = 0$.

Theorem 7.1.2. *Let Γ be a discrete reflection group with respect to the polyhedron P. Then all the dihedral angles of P are submultiples of π; moreover, if g_S and g_T are the reflections in adjacent sides S and T of P, and $\theta(S,T) = \pi/k$, then $g_S g_T$ has order k in Γ.*

Proof: Let S, T be adjacent sides of P. Then $\{S, T\}$ is a cycle of sides of P. By Theorem 6.7.7, there is a positive integer k such that

$$2\theta(S,T) = 2\pi/k$$

and the element $g_S g_T$ has order k in Γ. $\square$

Theorem 7.1.3. *Let P be a finite-sided, n-dimensional, convex polyhedron in X of finite volume all of whose dihedral angles are submultiples of π. Then the group Γ generated by the reflections of X in the sides of P is a discrete reflection group with respect to the polyhedron P.*

Proof: (1) The proof is by induction on n. The theorem is obviously true when $n = 1$, so assume that $n > 1$ and the theorem is true in dimension $n-1$. The idea of the proof is to construct a topological space $\tilde{X}$ for which the theorem is obviously true, and then to show that $\tilde{X}$ is homeomorphic to X by a covering space argument.

(2) Let $\Gamma \times P$ be the cartesian product of Γ and P. We topologize $\Gamma \times P$ by giving Γ the discrete topology and $\Gamma \times P$ the product topology. Then $\Gamma \times P$ is the topological sum of the subspaces

$$\{\{g\} \times P : g \in \Gamma\}.$$

Moreover, the mapping $(g, x) \mapsto gx$ is a homeomorphism of $\{g\} \times P$ onto gP for each g in Γ.

(3) Let $\mathcal{S}$ be the set of sides of P and for each S in $\mathcal{S}$, let g_S be the reflection of X in the side S of P. Let $\Phi = \{g_S : S \in \mathcal{S}\}$. Two points (g, x) and (h, y) of $\Gamma \times P$ are said to be *paired* by Φ, written $(g, x) \simeq (h, y)$, if and only if $g^{-1}h$ is in Φ and $gx = hy$. Suppose that $(g, x) \simeq (h, y)$. Then there is a side S of P such that $g^{-1}h = g_S$. As $g_S^{-1} = g_S$, we have that $(h, y) \simeq (g, x)$. Furthermore x is in $P \cap g_S(P) = S$, and so $x = g_S x = y$.

Two points (g, x) and (h, y) of $\Gamma \times P$ are said to be *related* by Φ, written $(g, x) \sim (h, y)$, if and only if there is a finite sequence, $(g_0, x_0), \ldots, (g_k, x_k)$, of points of $\Gamma \times P$ such that $(g, x) = (g_0, x_0)$, $(g_k, x_k) = (h, y)$, and

$$(g_{i-1}, x_{i-1}) \simeq (g_i, x_i) \quad \text{for } i = 1, \ldots, k.$$

Being related by Φ is obviously an equivalence relation on $\Gamma \times P$; moreover, if $(g, x) \sim (h, y)$, then $x = y$. Let $[g, x]$ be the equivalence class of (g, x) and let $\tilde{X}$ be the quotient space of $\Gamma \times P$ of equivalence classes.

(4) If $(g, x) \simeq (h, x)$, then obviously $(fg, x) \simeq (fh, x)$ for each f in Γ. Hence Γ acts on $\tilde{X}$ by $f[g, x] = [fg, x]$. For a subset A of P, set

$$[A] = \{[1, x] : x \in A\}.$$

Then if g is in Γ, we have

$$g[A] = \{[g, x] : x \in A\}.$$

If (g, x) is in $\Gamma \times P^\circ$, then $[g, x] = \{(g, x)\}$. Consequently, the members of $\{g[P^\circ] : g \in \Gamma\}$ are mutually disjoint in $\tilde{X}$.

(5) We now show that $\tilde{X}$ is connected. Let $\eta : \Gamma \times P \to \tilde{X}$ be the quotient map. As η maps $\{g\} \times P$ onto $g[P]$, we have that $g[P]$ is connected. In view of the fact that

$$\tilde{X} = \cup\{g[P] : g \in \Gamma\},$$

it suffices to show that for any g in Γ, there is a finite sequence $g_0, \ldots, g_k$ in Γ such that $[P] = g_0[P]$, $g_k[P] = g[P]$, and $g_{i-1}[P]$ and $g_i[P]$ intersect for each $i > 0$. As Γ is generated by the elements of Φ, there are sides S_i of P such that $g = g_{S_1} \cdots g_{S_k}$. Let $g_0 = 1$ and $g_i = g_{S_1} \cdots g_{S_i}$ for $i = 1, \ldots, k$. Now as

$$S_i = P \cap g_{S_i}(P),$$

we have that

$$[S_i] \subset [P] \cap g_{S_i}[P].$$

Therefore, we have

$$g_{i-1}[S_i] \subset g_{i-1}[P] \cap g_i[P].$$

Thus $\tilde{X}$ is connected.

(6) Let x be a point of P, let $\mathcal{S}(x)$ be the set of all the sides of P containing x, and let $\Gamma(x)$ be the subgroup of Γ generated by the elements of $\{g_S : S \in \mathcal{S}(x)\}$. We now show that $\Gamma(x)$ is finite. Let r be a real number such that $0 < r < \pi/2$ and r is less than the distance from x to any side of P not containing x. By Theorem 6.3.19, we have that $P \cap S(x, r)$ is an $(n-1)$-dimensional convex polyhedron in the sphere $S(x, r)$ and

$$\{T \cap S(x, r) : T \in \mathcal{S}(x)\}$$

is the set of sides of $P \cap S(x, r)$. Clearly $P \cap S(x, r)$ is compact and all the dihedral angles of $P \cap S(x, r)$ are submultiples of π. By the induction hypothesis, $\Gamma(x)$ restricts to a discrete reflection group with respect to $P \cap S(x, r)$. Hence $\Gamma(x)$ is finite, since $S(x, r)$ is compact.

(7) We next show that

$$[1,x] = \{(g,x) : g \in \Gamma(x)\}.$$

Let (g,x) be in $[1,x]$. Then there is a sequence $g_0,\ldots,g_k$ in Γ such that $(1,x) = (g_0,x)$, $(g_k,x) = (g,x)$, and $(g_{i-1},x) \simeq (g_i,x)$ for all $i > 0$. Hence $g_i x = x$ for all i and there is a side S_i in $\mathcal{S}(x)$ such that $g_i = g_{i-1}g_{S_i}$ for $i = 1,\ldots,k$. Therefore $g = g_{S_1}\cdots g_{S_k}$. Thus g is in $\Gamma(x)$. Consequently

$$[1,x] \subset \{(g,x) : g \in \Gamma(x)\}.$$

Now let g be an element of $\Gamma(x)$. Since $\Gamma(x)$ is generated by the set $\{g_S : S \in \mathcal{S}(x)\}$, there are sides S_i in $\mathcal{S}(x)$ such that $g = g_{S_1}\cdots g_{S_k}$. Let $g_0 = 1$ and $g_i = g_{S_1}\cdots g_{S_i}$ for $i = 1,\ldots,k$. Then g_i is in $\Gamma(x)$ for all i. As $g_{i-1}^{-1}g_i = g_{S_i}$, we have that $(g_{i-1},x) \simeq (g_i,x)$ for all $i > 0$. Hence $(1,x) \sim (g,x)$. Thus

$$[1,x] = \{(g,x) : g \in \Gamma(x)\}.$$

(8) For each point x of P and real number r as in (6), define

$$\tilde{B}(x,r) = \bigcup_{g\in\Gamma(x)} g[P \cap B(x,r)].$$

Suppose that g is in $\Gamma(x)$ and y is $P \cap B(x,r)$. Then $\mathcal{S}(y) \subset \mathcal{S}(x)$, and so $\Gamma(y) \subset \Gamma(x)$. As

$$[1,y] = \{(h,y) : h \in \Gamma(y)\}.$$

we have that

$$[g,y] = \{(gh,y) : h \in \Gamma(y)\}.$$

Consequently

$$\eta^{-1}(\tilde{B}(x,r)) = \bigcup_{g\in\Gamma(x)} \{g\} \times (P \cap B(x,r)).$$

Hence $\tilde{B}(x,r)$ is an open neighborhood of $[1,x]$ in $\tilde{X}$; moreover $\tilde{B}(x,r)$ intersects $g[P]$ if and only if g is in $\Gamma(x)$.

(9) Let $\kappa : \tilde{X} \to X$ be the map defined by $\kappa[g,x] = gx$. We now show that κ maps $\tilde{B}(x,r)$ onto $B(x,r)$. By Theorem 6.7.1, we have that

$$\{gP \cap S(x,r) : g \in \Gamma(x)\}$$

is a tessellation of $S(x,r)$. Consequently, the members of

$$\{gP^\circ \cap B(x,r) : g \in \Gamma(x)\}$$

are mutually disjoint and

$$B(x,r) = \bigcup_{g\in\Gamma(x)} \big(gP \cap B(x,r)\big).$$

Now as κ maps $g[P \cap B(x,r)]$ onto $gP \cap B(x,r)$ for each g in $\Gamma(x)$, we have that κ maps $\tilde{B}(x,r)$ onto $B(x,r)$.

(10) We now show that κ maps $\tilde{B}(x,r)$ injectively into $B(x,r)$. Let g,h be in $\Gamma(x)$, let y,z be in $P \cap B(x,r)$, and suppose that $\kappa[g,y] = \kappa[h,z]$.

Then $gy = hz$. Hence P and $g^{-1}hP$ intersect at $y = g^{-1}hz$. As y is in $P \cap B(x,r)$, we have that $\Gamma(y) \subset \Gamma(x)$. Now there is an $s > 0$ such that

$$B(y,s) \subset B(x,r),$$

and

$$B(y,s) = \bigcup_{f \in \Gamma(y)} \big(fP \cap B(y,s)\big).$$

Hence $g^{-1}hP \cap B(y,s)$ intersects $fP^\circ \cap B(y,s)$ for some f in $\Gamma(y)$. But the members of

$$\{fP^\circ \cap B(x,r) : f \in \Gamma(x)\}$$

are mutually disjoint. Therefore $g^{-1}h = f$ for some f in $\Gamma(y)$. Hence

$$y = f^{-1}y = h^{-1}gy = z$$

and

$$[g,y] = g[1,y] = g[g^{-1}h,y] = [h,y] = [h,z].$$

Thus κ maps $\tilde{B}(x,r)$ bijectively onto $B(x,r)$.

(11) We now show that κ maps $\tilde{B}(x,r)$ homeomorphically onto $B(x,r)$. Let g be in $\Gamma(x)$. As $\kappa\eta$ maps $\{g\} \times P \cap B(x,r)$ homeomorphically onto $gP \cap B(x,r)$, we have that κ maps $g[P \cap B(x,r)]$ homeomorphically onto $gP \cap B(x,r)$. Now since

$$B(x,r) = \bigcup_{g \in \Gamma(x)} \big(gP \cap B(x,r)\big),$$

and each set $gP \cap B(x,r)$ is closed in $B(x,r)$, and $\Gamma(x)$ is finite, we deduce that κ maps $\tilde{B}(x,r)$ homeomorphically onto $B(x,r)$.

(12) Now let g be an element of Γ. Then left multiplication by g is a homeomorphism of $\tilde{X}$, since left multiplication by g is a homeomorphism of $\Gamma \times P$. Hence $g\tilde{B}(x,r)$ is an open neighborhood of $[g,x]$ in $\tilde{X}$. As $\kappa(g\tilde{B}(x,r)) = g\kappa(\tilde{B}(x,r))$, we have that κ maps $g\tilde{B}(x,r)$ homeomorphically onto $B(gx,r)$. Thus κ is a local homeomorphism.

(13) We now show that $\tilde{X}$ is Hausdorff. Let

$$\begin{aligned} [g,x] &= \{(g_1,x),\ldots,(g_k,x)\}, \\ [h,y] &= \{(h_1,y),\ldots,(h_\ell,y)\} \end{aligned}$$

be distinct points of $\tilde{X}$. Then they are disjoint subsets of $\Gamma \times P$. Now choose r as before so that κ maps $\tilde{B}(x,r)$ homeomorphically onto $B(x,r)$ and κ maps $\tilde{B}(y,r)$ homeomorphically onto $B(y,r)$. We may choose r small enough so that the sets

$$\begin{aligned} \eta^{-1}(g\tilde{B}(x,r)) &= \bigcup_{i=1}^{k}\{g_i\} \times (P \cap B(x,r)), \\ \eta^{-1}(h\tilde{B}(y,r)) &= \bigcup_{j=1}^{\ell}\{h_j\} \times (P \cap B(y,r)) \end{aligned}$$

are disjoint in $\Gamma \times P$, since if $g_i \neq h_j$, then $\{g_i\} \times P$ and $\{h_j\} \times P$ are disjoint; while if $x \neq y$, we can choose r small enough so that $B(x,r)$ and $B(y,r)$

are disjoint. Therefore $g\tilde{B}(x,r)$ and $h\tilde{B}(y,r)$ are disjoint neighborhoods of $[g,x]$ and $[h,y]$, respectively, in $\tilde{X}$. Thus $\tilde{X}$ is Hausdorff.

(14) Let v be an ideal vertex of P, let $\mathcal{S}(v)$ be the set of all the sides of P incident with v, and let $\Gamma(v)$ be the subgroup of Γ generated by the set $\{g_S : S \in \mathcal{S}(v)\}$. Let B be a horoball based at v such that $\overline{B}$ meets only the sides in $\mathcal{S}(v)$. Then $P \cap \partial B$ is an $(n-1)$-dimensional, Euclidean, convex polyhedron in the horosphere ∂B and

$$\{S \cap \partial B : S \in \mathcal{S}(v)\}$$

is the set of sides of $P \cap \partial B$. Clearly $P \cap \partial B$ is compact and all the dihedral angles of $P \cap \partial B$ are submultiples of π. By the induction hypothesis, $\Gamma(v)$ restricts to a discrete reflection group with respect to $P \cap \partial B$.

(15) Define

$$\tilde{B} = \bigcup_{g \in \Gamma(v)} g[P \cap B].$$

By the same argument as in (8), we have

$$\eta^{-1}(\tilde{B}) = \bigcup_{g \in \Gamma(v)} \{g\} \times (P \cap B).$$

Hence $\tilde{B}$ is an open subset of $\tilde{X}$, and $\tilde{B}$ intersects $g[P]$ if and only if g is in $\Gamma(v)$. By the same arguments as in (9) and (10), κ maps $\tilde{B}$ bijectively onto B. As κ is an open map, κ maps $\tilde{B}$ homeomorphically onto B.

(16) Let $v_1, \ldots, v_m$ be the ideal vertices of P and for each i, let B_i be a horoball based at v_i such that $\overline{B}_i$ meets only the sides of P incident with v_i. Let B_i' be the horoball based at v_i such that $B_i' \subset B_i$ and $\mathrm{dist}(B_i', \partial B_i) = 1$. Now set

$$P_0 = P - \bigcup_{i=1}^{m} B_i'.$$

Then P_0 is compact. Let x be a point of P. Choose $r(x) > 0$ as before so that κ maps $\tilde{B}(x, r(x))$ homeomorphically onto $B(x, r(x))$. As P_0 is compact, the open covering $\{B(x, r(x)) : x \in P_0\}$ of P_0 has a Lebesgue number ℓ such that $0 < \ell < 1$. If x is in P_0, let y be a point of P_0 such that $B(x,\ell) \subset B(y, r(y))$, and let $\tilde{B}(x)$ be the subset of $\tilde{B}(y, r(y))$ that is mapped onto $B(x,\ell)$ by κ. If x is in B_i', let $\tilde{B}(x)$ be the subset of $\tilde{B}_i$ that is mapped onto $B(x,\ell)$ by κ. Then $\tilde{B}(x)$ is an open neighborhood of $[1,x]$ in $\tilde{X}$ that is mapped homeomorphically onto $B(x,\ell)$ by κ. Moreover, if g is in Γ, then $g\tilde{B}(x)$ is an open neighborhood of $[g,x]$ in $\tilde{X}$ that is mapped homeomorphically onto $B(gx,\ell)$ by κ. Thus, if y is in the image of κ, then $B(y,\ell)$ is in the image of κ. Therefore κ is surjective.

(17) Next, let $\alpha : [a,b] \to X$ be a geodesic arc from y to z such that $|\alpha| < \ell$ and suppose that $\kappa[g,x] = y$. We now show that α lifts to a unique curve $\tilde{\alpha} : [a,b] \to \tilde{X}$ such that $\tilde{\alpha}(a) = [g,x]$. Now as κ maps $g\tilde{B}(x)$ homeomorphically onto $B(gx,\ell)$, the map α lifts to a curve $\tilde{\alpha} : [a,b] \to \tilde{X}$ such that $\tilde{\alpha}(a) = [g,x]$ and $\tilde{\alpha}([a,b]) \subset g\tilde{B}(x)$. Suppose that $\hat{\alpha} : [a,b] \to \tilde{X}$ is a different lift of α starting at $[g,x]$. Then $\hat{\alpha}^{-1}(g\tilde{B}(x))$ is a proper open neighborhood of a in $[a,b]$, since $\hat{\alpha}$ is continuous and not equal to $\tilde{\alpha}$. Let t

be the first point of $[a, b]$ not in this neighborhood. Then $\tilde{\alpha}(t) \neq \hat{\alpha}(t)$. As $\tilde{X}$ is Hausdorff, there are disjoint open neighborhoods U and V of $\tilde{\alpha}(t)$ and $\hat{\alpha}(t)$, respectively. Choose $s < t$ in the open neighborhood $\tilde{\alpha}^{-1}(U) \cap \hat{\alpha}^{-1}(V)$ of t. Then $\hat{\alpha}(s)$ is in $g\tilde{B}(x)$ and so must be equal to $\tilde{\alpha}(s)$. As U and V are disjoint, we have a contradiction. Therefore, the lift $\tilde{\alpha}$ is unique.

(18) We now show that $\kappa : \tilde{X} \to X$ is a covering projection. Let z be a point of X. We will show that $B(z, \ell)$ is evenly covered by κ. Since κ is surjective, there is a point $[g, x]$ of $\tilde{X}$ such that $\kappa[g, x] = z$. Then κ maps the open neighborhood $g\tilde{B}(x)$ of $[g, x]$ in $\tilde{X}$ homeomorphically onto $B(z, \ell)$. Next, suppose that $[h, y] \neq [g, x]$ and $\kappa[h, y] = z$. We claim that $g\tilde{B}(x)$ and $h\tilde{B}(y)$ are disjoint. On the contrary, suppose that $[f, w]$ is in $g\tilde{B}(x) \cap h\tilde{B}(y)$. Let $\alpha : [a, b] \to X$ be a geodesic arc from z to fw. As fw is in $B(z, \ell)$, we have that $|\alpha| < \ell$. Hence α lifts to unique curves $\tilde{\alpha}_1, \tilde{\alpha}_2 : [a, b] \to \tilde{X}$ starting at $[g, x]$ and $[h, y]$, respectively. Both $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ end at $[f, w]$, since $[f, w]$ is the only point in $g\tilde{B}(x)$ and in $h\tilde{B}(y)$ that is mapped to fw by κ. By the uniqueness of the lift of α^{-1} starting at $[f, w]$, we have that $[g, x] = [h, y]$, which is a contradiction. Hence $g\tilde{B}(x)$ and $h\tilde{B}(y)$ are disjoint, and so $B(z, \ell)$ is evenly covered by κ. Thus κ is a covering projection.

(19) Now $\kappa : \tilde{X} \to X$ is a homeomorphism, since X is simply connected and $\tilde{X}$ is connected. Therefore, the members of $\{gP^\circ : g \in \Gamma\}$ are mutually disjoint, since the members of $\{g[P^\circ] : g \in \Gamma\}$ are mutually disjoint; and

$$X = \cup\{gP : g \in \Gamma\},$$

since we have

$$\tilde{X} = \cup\{g[P] : g \in \Gamma\}.$$

(20) We now show that

$$\mathcal{P} = \{gP : g \in \Gamma\}$$

is locally finite. Let y be an arbitrary point of X. Then there is a unique element $[f, x]$ of $\tilde{X}$ such that $\kappa[f, x] = y$. Let r be such that $0 < r < \pi/2$ and r is less than the distance from any side of P not containing x. Then the open neighborhood $f\tilde{B}(x, r)$ of $[f, x]$ intersects $g[P]$ if and only if $f^{-1}g$ is in $\Gamma(x)$. Hence, the set

$$\kappa(f\tilde{B}(x, r)) = B(fx, r) = B(y, r)$$

intersects gP if and only if $f^{-1}g$ is in $\Gamma(x)$. As $\Gamma(x)$ is finite, we have that $B(y, r)$ meets only finitely many members of $\mathcal{P}$. Thus $\mathcal{P}$ is locally finite.

(21) If gS is any side of gP, then gS is also a side of gg_SP, and since

$$gP \cap gg_SP = gS,$$

we have that gP and gg_SP are the only polyhedra of $\mathcal{P}$ containing gS as a side. Thus $\mathcal{P}$ is an exact tessellation of X. Therefore Γ is discrete and P is an exact, convex, fundamental polyhedron for Γ by Theorem 6.7.1. Thus Γ is a discrete reflection group with respect to the polyhedron P. □

Example 1. Let

$$P = \{x \in S^n : x_i \geq 0 \text{ for } i = 1, \ldots, n+1\}.$$

Then P is a regular n-simplex in S^n whose dihedral angle is $\pi/2$. Therefore, the group Γ generated by the reflections in the sides of P is a discrete reflection group with respect to P by Theorem 7.1.3. Obviously, the tessellation $\{gP : g \in \Gamma\}$ of S^n contains 2^{n+1} simplices, and so Γ has order 2^{n+1}. It is worth noting that the vertices of the regular tessellation $\{gP : g \in \Gamma\}$ of S^n are the vertices of an $(n+1)$-dimensional, Euclidean, regular, polytope inscribed in S^n.

Example 2. Let P be an n-cube in E^n. Then P is a regular polytope in E^n whose dihedral angle is $\pi/2$. Therefore, the group Γ generated by the reflections in the sides of P is a discrete reflection group with respect to P by Theorem 7.1.3.

Example 3. Form a cycle of hyperbolic triangles by reflecting in the sides of a $30° - 45°$ hyperbolic right triangle, always keeping the vertex at the $30°$ angle fixed. As $30° = 360°/12$, there are 12 triangles in this cycle, and their union is a hyperbolic regular hexagon P whose dihedral angle is $90°$. See Figure 7.1.2. Let Γ be the group generated by the reflections in the sides of P. Then Γ is a discrete reflection group with respect to P by Theorem 7.1.3.

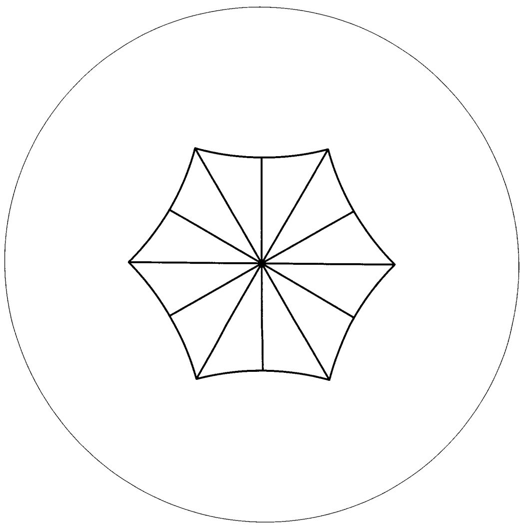


Figure 7.1.2. A cycle of twelve 30° - 45° hyperbolic right triangles

Example 4. Let $D(r)$ be a regular dodecahedron inscribed on the sphere $S(0,r)$ in E^3 with $0 < r < 1$. Then $D(r)$ is a hyperbolic regular dodecahedron in the projective disk model D^3 of hyperbolic 3-space. Let $\theta(r)$ be the hyperbolic dihedral angle of $D(r)$. When r is small, $\theta(r)$ is approximately equal to but less than the value of the dihedral angle of a Euclidean regular dodecahedron $\theta(0)$, which is approximately $116.6°$. As r increases to 1, the angle $\theta(r)$ decreases continuously to its limiting value $\theta(1)$, the dihedral angle of a regular ideal dodecahedron in D^3. See Figure 7.1.3.

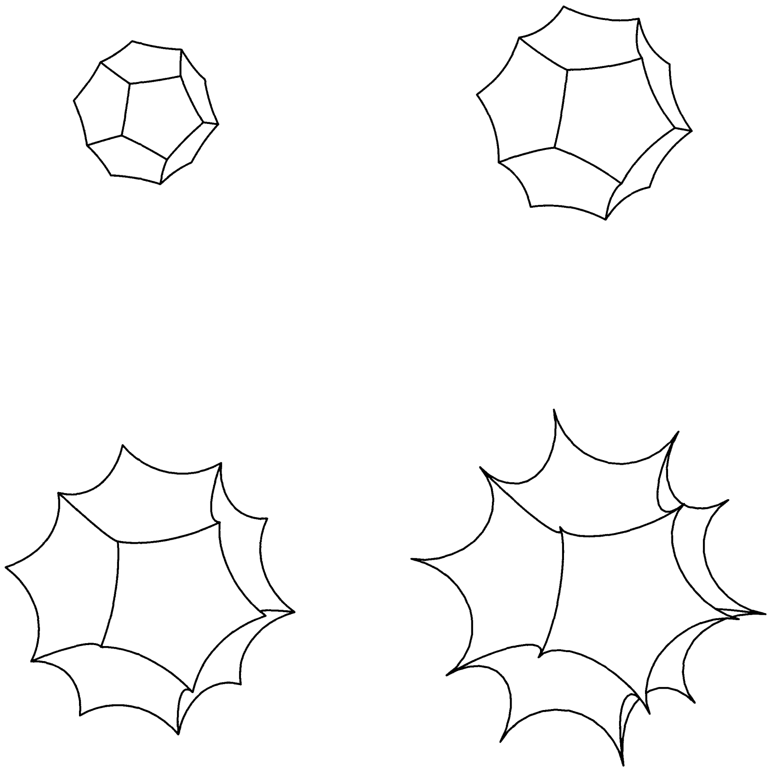


Figure 7.1.3. Four views of an expanding, hyperbolic, regular, dodecahedron centered at the origin in the conformal ball model of hyperbolic 3-space

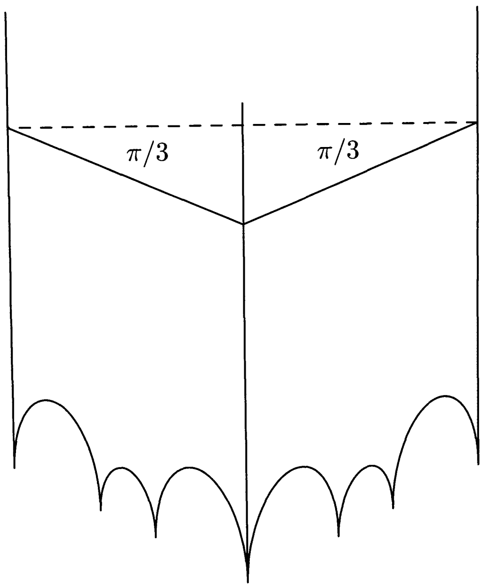


Figure 7.1.4. A regular ideal dodecahedron in U^3 with a vertex at ∞

To find the value of $\theta(1)$, we consider a regular ideal dodecahedron in the upper half-space model U^3 of hyperbolic 3-space with an ideal vertex at ∞. Since the dodecahedron is regular, the link of the ideal vertex at ∞ is an equilateral triangle. Therefore $\theta(1) = 60°$. See Figure 7.1.4.

Now as $\theta(r)$ is a continuous function of r, taking values in the interval $[\theta(1),\ \theta(0)]$, there is a unique value of r such that $\theta(r) = 90°$. Let $P = D(r)$ for this r. Then P is a hyperbolic regular dodecahedron whose dihedral angle is $\pi/2$. Let Γ be the group generated by the reflections in the sides of P. Then Γ is a discrete reflection group with respect to P by Theorem 7.1.3.

Example 5. By the previous discussion, a regular ideal dodecahedron P in H^3 has dihedral angle $\pi/3$. Let Γ be the group generated by the reflections in the sides of P. Then Γ is a discrete reflection group with respect to P by Theorem 7.1.3.

Example 6. The 24 points $\pm e_i$, for $i = 1, 2, 3, 4$, and $(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})$ of S^3 are the vertices of a regular 24-cell in E^4. Let P be the corresponding regular ideal 24-cell in B^4. The link of an ideal vertex of P is a cube. Therefore, the dihedral angle of P is $\pi/2$. Let Γ be the group generated by the reflections in the sides of P. Then Γ is a discrete reflection group with respect to P by Theorem 7.1.3.

Let Γ be a discrete reflection group with respect to a polyhedron P. Then all the dihedral angles of P are submultiples of π by Theorem 7.1.2. Let $\{S_i\}$ be the sides of P and for each pair of indices i, j, let $k_{ij} = \pi/\theta(S_i, S_j)$. Let F be the group freely generated by the symbols $\{S_i\}$ and let g_{S_i} be the reflection of X in the hyperplane $\langle S_i \rangle$. Then the map $\phi : F \to \Gamma$, defined by $\phi(S_i) = g_{S_i}$, is an epimorphism. By Theorem 7.1.2, the kernel of ϕ contains the words $(S_iS_j)^{k_{ij}}$ whenever k_{ij} is finite.

Let G be the quotient of F by the normal closure of the words

$$\left\{(S_iS_j)^{k_{ij}} : k_{ij} \text{ is finite}\right\}.$$

Then ϕ induces an epimorphism $\psi : G \to \Gamma$. We shall prove that ψ is an isomorphism when P has finitely many sides and finite volume. This fact is usually expressed by saying that

$$\left(S_i; (S_iS_j)^{k_{ij}}\right)$$

is a group presentation for Γ under the mapping $S_i \mapsto g_{S_i}$. Here it is understood that $(S_iS_j)^{k_{ij}}$ is to be deleted when $k_{ij} = \infty$.

Theorem 7.1.4. *Let Γ be a discrete reflection group with respect to a finite-sided polyhedron P in X of finite volume. Let $\{S_i\}$ be the set of sides of P and for each pair of indices i, j, let $k_{ij} = \pi/\theta(S_i, S_j)$. Then*

$$\left(S_i; (S_iS_j)^{k_{ij}}\right)$$

is a group presentation for Γ under the mapping $S_i \mapsto g_{S_i}$.

Proof: The proof follows the same outline as the proof of Theorem 7.1.3, and so only the necessary alterations will be given. The start of the induction requires proof. If $n = 1$, then P is a geodesic segment and Γ is obviously a dihedral group of order $2k_{12}$. It is then an exercise in group presentations to show that Γ has the presentation

$$\left(S_1, S_2; S_1^2, S_2^2, (S_1S_2)^{k_{12}}\right).$$

The main alteration in the proof of Theorem 7.1.3 is to replace Γ by G in the construction of the covering space $\tilde{X}$. Everything goes through as before except where the induction hypothesis is used in steps (6) and (14). Here one draws the additional conclusion that $\Gamma(x)$ has the presentation

$$\left(S_i \in \mathcal{S}(x); (S_iS_j)^{k_{ij}}\right).$$

Since the subgroup $G(x)$ of G generated by the set $\{S_i : S_i \in \mathcal{S}(x)\}$ satisfies the same relations and maps onto $\Gamma(x)$, we deduce that $G(x)$ has the same presentation. In particular, the mapping $S_i \mapsto g_{S_i}$ induces an isomorphism from $G(x)$ onto $\Gamma(x)$. Now everything goes through as before. The final conclusion is that the mapping $S_i \mapsto g_{S_i}$ induces an isomorphism from G to Γ. $\square$

Coxeter Groups

Definition: A *Coxeter group* G is an abstract group defined by a group presentation of the form $(S_i; (S_iS_j)^{k_{ij}})$, where

(1) the indices i, j vary over some countable indexing set $\mathcal{I}$;

(2) the exponent k_{ij} is either a positive integer or ∞ for each i, j;

(3) $k_{ij} = k_{ji}$;

(4) $k_{ii} = 1$ for each i;

(5) $k_{ij} > 1$ if $i \neq j$; and

(6) if $k_{ij} = \infty$, then the relator $(S_iS_j)^{k_{ij}}$ is deleted.

Note that if $i \neq j$, then the relator $(S_jS_i)^{k_{ji}}$ is derivable from the relators S_i^2, S_j^2, and $(S_iS_j)^{k_{ij}}$; and therefore only one of the relators $(S_iS_j)^{k_{ij}}$ and $(S_jS_i)^{k_{ji}}$ is required and the other may be deleted.

Let $G = (S_i, i \in \mathcal{I}; (S_iS_j)^{k_{ij}})$ be a Coxeter group. The *Coxeter graph* of G is the labeled graph with vertices $\mathcal{I}$ and edges

$$\{(i, j) : k_{ij} > 2\}.$$

Each edge (i, j) is labeled by k_{ij}. For simplicity, the edges with $k_{ij} = 3$ are usually not labeled in a representation of a Coxeter graph.

Example 7. The Coxeter group $G = (S_1; S_1^2)$ is a cyclic group of order two. Its Coxeter graph is a single vertex.

Example 8. The Coxeter group $G(k) = (S_1, S_2; S_1^2, S_2^2, (S_1S_2)^k)$ is a dihedral group of order $2k$. Its Coxeter graph, when $k > 2$, is a single edge with the label k.

Let Γ be a discrete reflection group with respect to a finite-sided polyhedron P of finite volume. Let $\{S_i\}$ be the set of sides of P, and for each pair of indices i, j, let $k_{ij} = \pi/\theta(S_i, S_j)$. Then the Coxeter group

$$G = \left(S_i; (S_iS_j)^{k_{ij}}\right)$$

is isomorphic to Γ by Theorem 7.1.4. Thus Γ is a Coxeter group.

Example 9. Let Γ be the group generated by the reflections in the sides of a rectangle P in E^2. By Theorem 7.1.4, the group Γ has the presentation

$$\left(S_1, S_2, S_3, S_4; S_i^2, (S_iS_{i+1})^2 \quad i \bmod 4\right).$$

The Coxeter graph of Γ consists of two disjoint edges labeled by ∞.

A Coxeter group G is said to be *irreducible* or *reducible* according as its Coxeter graph is connected or disconnected. We leave it as an exercise to show that a reducible Coxeter group is the direct product of the irreducible Coxeter groups represented by the connected components of its graph. For example, the discrete reflection group in Example 9 is the direct product of the two infinite dihedral groups $(S_1, S_3; S_1^2, S_3^2)$ and $(S_2, S_4; S_2^2, S_4^2)$. This is not surprising, since a rectangle in E^2 is the cartesian product of two line segments. In general, the geometric basis for the direct product decomposition of a reducible discrete reflection group is the fact that orthogonal reflections commute.

Exercise 7.1

1. Let Γ be a discrete reflection group with respect to a polyhedron P. Prove that P is the Dirichlet polyhedron for Γ with center any point of P°.
2. Let Γ be a discrete reflection group with respect to a polyhedron P. Prove that X/Γ is isometric to P.
3. Let Γ be the group generated by two reflections of E^1 about the endpoints of a line segment. Show that Γ has the presentation $(S, T; S^2, T^2)$.
4. Let k be a positive integer or ∞. Prove that the element ST generates a cyclic normal subgroup of order k and index 2 in the dihedral group
$$G(k) = \big(S, T; S^2, T^2, (ST)^k\big).$$
Interpret this fact geometrically in terms of reflections of S^1 or E^1.
5. Prove that a reducible Coxeter group G is the direct product of the irreducible Coxeter groups represented by the connected components of the Coxeter graph of G.
6. Prove that the group Γ in Example 1 is an elementary 2-group of rank $n+1$.
7. Show that the Coxeter graph of the group Γ in Example 3 is connected.
8. Let P be an n-dimensional convex polyhedron in S^n all of whose dihedral angles are submultiples of π. Prove that P has at most $n+1$ sides.
9. Let P be an n-dimensional convex polyhedron in E^n all of whose dihedral angles are submultiples of π. Prove that P has at most $2n$ sides.
10. Let Γ be a discrete reflection group with respect to a finite-sided polyhedron P in X of finite volume and let S and T be distinct nonadjacent sides of P. Prove that the element $g_S g_T$ has infinite order in Γ.
11. Prove that Theorem 7.1.1 is still true without the hypothesis that P has finite volume.
12. Prove that Theorem 7.1.3 is still true without the hypothesis that P has finite volume.
13. Prove that Theorem 7.1.4 is still true without the hypothesis that P has finite volume.

§7.2. Simplex Reflection Groups

Throughout this section, $X = S^n, E^n$, or H^n with $n > 0$. Let Δ be an n-simplex in X all of whose dihedral angles are submultiples of π. By Theorem 7.1.3, the group Γ generated by the reflections of X in the sides of Δ is a discrete group of isometries of X. The group Γ is called an *n-simplex reflection group.*

We shall also include the case of a 0-simplex Δ in S^0. We regard the antipodal map α of S^0 to be a reflection of S^0. Since $\{\Delta, \alpha(\Delta)\}$ is a tessellation of S^0, we also call the group Γ generated by α, a *0-simplex reflection group.* The Coxeter graph of Γ is defined to be a single vertex.

Assume that $n = 1$. Then Δ is a geodesic segment in X. Clearly Γ is a dihedral group of order $2k$, with $k > 1$, where π/k is the angle of Δ. The Coxeter graph of Γ is either two vertices if $k = 2$ or an edge labeled by k if $k > 2$. If $X = S^1$, then k is finite, whereas if $X = E^1$ or H^1, then $k = \infty$.

Assume that $n = 2$. Then there are integers a, b, c, with $2 \leq a \leq b \leq c$, such that Δ is a triangle $T(a, b, c)$ in X whose angles are π/a, π/b, π/c. Note that $T(a, b, c)$ is determined up to similarity in X by the integers a, b, c. The group Γ generated by the reflections in the sides of $T(a, b, c)$ is denoted by $G(a, b, c)$. Let $G_0(a, b, c)$ be the subgroup of $G(a, b, c)$ of orientation preserving isometries. Then $G_0(a, b, c)$ has index two in $G(a, b, c)$. The group $G_0(a, b, c)$ is called a *triangle group*, whereas $G(a, b, c)$ is called a *triangle reflection group.*

Spherical Triangle Reflection Groups

Assume that $X = S^2$. By Theorem 2.5.1, we have

$$\frac{\pi}{a} + \frac{\pi}{b} + \frac{\pi}{c} > \pi.$$

Hence, the integers a, b, c satisfy the inequality

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} > 1.$$

There are an infinite number of solutions of the form

$$\frac{1}{2} + \frac{1}{2} + \frac{1}{c} > 1,$$

and only three more solutions,

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{3} > 1, \quad \frac{1}{2} + \frac{1}{3} + \frac{1}{4} > 1, \quad \frac{1}{2} + \frac{1}{3} + \frac{1}{5} > 1.$$

The Coxeter graph of the group $G(2, 2, 2)$ consists of three vertices, and so $G(2, 2, 2)$ is an elementary 2-group of order 8. The Coxeter graph of $G(2, 2, c)$, for $c > 2$, is the disjoint union of a vertex and an edge labeled by c. Hence $G(2, 2, c)$ is the direct product of a group of order 2 and a dihedral group of order $2c$. Thus $G(2, 2, c)$ has order $4c$. The tessellation of

S^2 generated by reflecting in the sides of $T(2,2,5)$ is illustrated in Figure 7.2.1(a).

By Theorem 2.5.5, the area of $T(2,3,3)$ is

$$\frac{\pi}{2} + \frac{\pi}{3} + \frac{\pi}{3} - \pi = \frac{\pi}{6}.$$

As the area of S^2 is 4π, the tessellation

$$\{gT(2,3,3) : g \in G(2,3,3)\}$$

contains 24 triangles, and so $G(2,3,3)$ has order 24. The tessellation can be partitioned into 4 cycles, each consisting of 6 triangles cycling about a $60°$ vertex. The union of each of these cycles is a spherical equilateral triangle. See Figure 7.2.1(b). This gives a regular tessellation of S^2 by 4 equilateral triangles. It is clear from the geometry of these two tessellations that $G(2,3,3)$ is the group of symmetries of the regular tetrahedron inscribed in S^2 with its vertices at the corners of the 4 equilateral triangles. Consequently $G(2,3,3)$ is a symmetric group on four letters. The triangle group $G_0(2,3,3)$ is an alternating group on four letters called the *tetrahedral group*. The Coxeter graph of $G(2,3,3)$ is

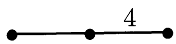
The area of $T(2,3,4)$ is $\pi/12$. Therefore, the tessellation

$$\{gT(2,3,4) : g \in G(2,3,4)\}$$

contains 48 triangles, and so $G(2,3,4)$ has order 48. The tessellation can be partitioned into 6 cycles, each consisting of 8 triangles cycling about a $45°$ vertex. The union of each of these cycles is a spherical regular quadrilateral. See Figure 7.2.1(c). This gives a regular tessellation of S^2 by 6 quadrilaterals. It is clear from the geometry of these two tessellations that $G(2,3,4)$ is the group of symmetries of the cube inscribed in S^2 with its vertices at the corners of the 6 quadrilaterals. The above tessellation of S^2 by 48 triangles can also be partitioned into 8 cycles, each consisting of 6 triangles cycling about a $60°$ vertex. The union of each of these cycles is a spherical equilateral triangle. See Figure 7.2.1(c). This gives a regular tessellation of S^2 by 8 equilateral triangles. It is clear from the geometry of these two tessellations that $G(2,3,4)$ is the group of symmetries of the regular octahedron inscribed in S^2 with its vertices at the corners of the 8 equilateral triangles. Now since a regular octahedron is antipodally symmetric, we have

$$G(2,3,4) = \{\pm 1\} \times G_0(2,3,4).$$

The triangle group $G_0(2,3,4)$ is a symmetric group on four letters called the *octahedral group*. The Coxeter graph of $G(2,3,4)$ is



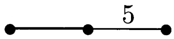
The area of $T(2,3,5)$ is $\pi/30$. Therefore, the tessellation

$$\{gT(2,3,5) : g \in G(2,3,5)\}$$

contains 120 triangles, and so $G(2,3,5)$ has order 120. The tessellation can be partitioned into 12 cycles, each consisting of 10 triangles cycling about a $36°$ vertex. The union of each of these cycles is a spherical regular pentagon. See Figure 7.2.1(d). This gives a regular tessellation of S^2 by 12 pentagons. It is clear from the geometry of these two tessellations that $G(2,3,5)$ is the group of symmetries of the regular dodecahedron inscribed in S^2 with its vertices at the corners of the 12 pentagons. The above tessellation of S^2 by 120 triangles can also be partitioned into 20 cycles, each consisting of 6 triangles cycling about a $60°$ vertex. The union of each of these cycles is a spherical equilateral triangle. See Figure 7.2.1(d). This gives a regular tessellation of S^2 by 20 equilateral triangles. It is clear from the geometry of these two tessellations that $G(2,3,5)$ is the group of symmetries of the regular icosahedron inscribed in S^2 with its vertices at the corners of the 20 equilateral triangles. Now since a regular icosahedron is antipodally symmetric, we have

$$G(2,3,5) = \{\pm 1\} \times G_0(2,3,5).$$

The triangle group $G_0(2,3,5)$ is an alternating group on five letters called the *icosahedral group*. The Coxeter graph of $G(2,3,5)$ is



Euclidean Triangle Reflection Groups

Now assume that $X = E^2$. Then we have

$$\frac{\pi}{a} + \frac{\pi}{b} + \frac{\pi}{c} = \pi.$$

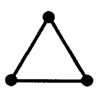
Hence, the integers a, b, c satisfy the equation

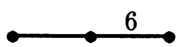
$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = 1.$$

There are exactly three solutions,

$$\frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1, \quad \frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 1, \quad \frac{1}{2} + \frac{1}{3} + \frac{1}{6} = 1.$$

Note that $T(3,3,3)$ is an equilateral triangle, $T(2,4,4)$ is an isosceles right triangle, and $T(2,3,6)$ is a $30°$-$60°$ right triangle. The tessellation of E^2 generated by reflecting in the sides of $T(a,b,c)$ in each of the three cases is illustrated in Figure 7.2.2. The Coxeter graphs of the groups $G(3,3,3), G(2,4,4)$, and $G(2,3,6)$ are, respectively,





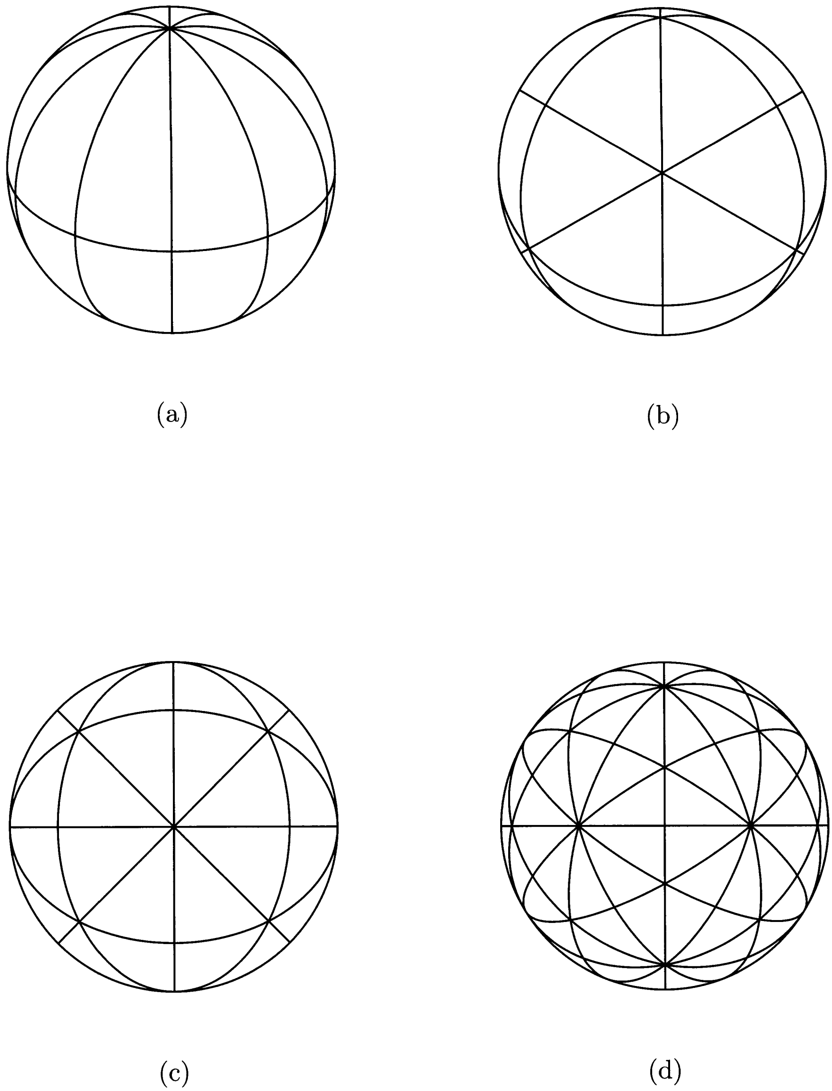


Figure 7.2.1. Tessellations of S^2 obtained by reflecting in the sides of a triangle

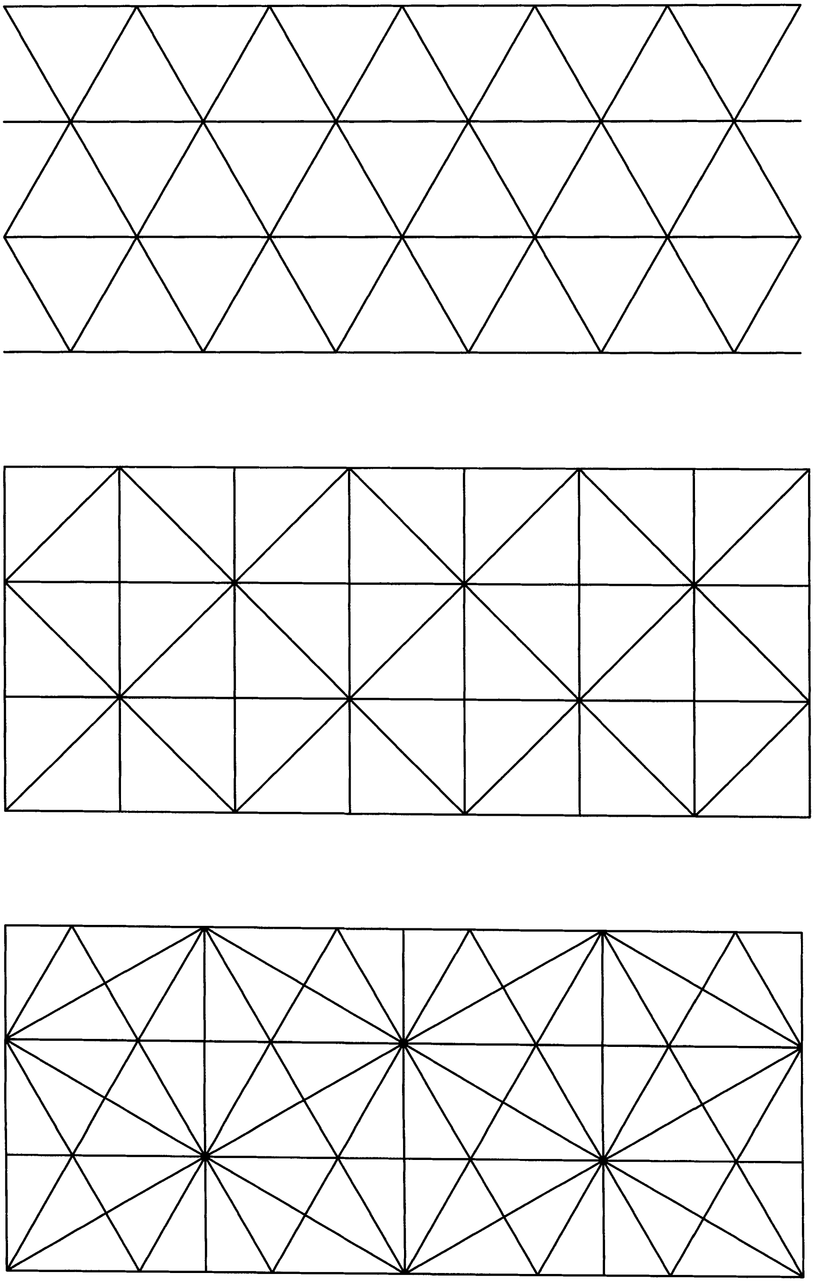


Figure 7.2.2. Tessellations of E^2 obtained by reflecting in the sides of a triangle

Hyperbolic Triangle Reflection Groups

Now assume that $X = H^2$. By Theorem 3.5.1, we have

$$\frac{\pi}{a} + \frac{\pi}{b} + \frac{\pi}{c} < \pi.$$

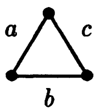
Hence, the integers a, b, c satisfy the inequality

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} < 1.$$

There are an infinite number of solutions to this inequality. Each solution determines a hyperbolic triangle $T(a, b, c)$ and a corresponding reflection group $G(a, b, c)$. Of all these triangles, $T(2, 3, 7)$ has the least area, $\pi/42$.

The Coxeter graph of a hyperbolic reflection group $G(a, b, c)$ is either

or



according as $a = 2$ or $a > 2$. Figure 7.2.3 illustrates the tessellation of B^2 generated by reflecting in the sides of $T(2, 4, 6)$. Note that this tessellation is the underlying geometry of Escher's circle print in Figure 1.2.5.

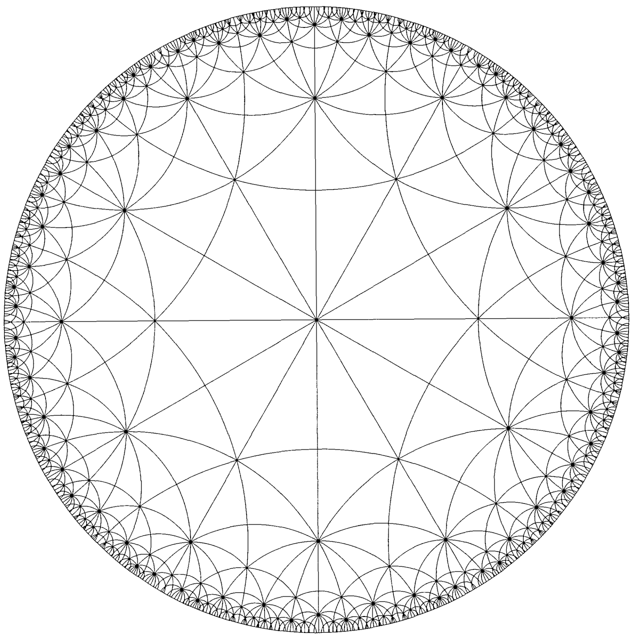


Figure 7.2.3. Tessellation of B^2 obtained by reflecting in the sides of $T(2, 4, 6)$

Theorem 7.2.1. *Let a, b, c, a', b', c' be integers such that*

$$2 \leq a \leq b \leq c \quad \text{and} \quad 2 \leq a' \leq b' \leq c'.$$

Then the triangle reflection groups $G(a, b, c)$ and $G(a', b', c')$ are isomorphic if and only if $(a, b, c) = (a', b', c')$.

Proof: Suppose that $G(a, b, c)$ and $G(a', b', c')$ are isomorphic. Assume first that $G(a, b, c)$ is finite. Then $G(a, b, c)$ and $G(a', b', c')$ are isomorphic spherical triangle reflections groups. From the description of all the spherical triangle reflection groups, we deduce that $(a, b, c) = (a', b', c')$. Thus, we may assume that $G(a, b, c)$ is infinite. Then $G(a, b, c)$ is either a Euclidean or hyperbolic triangle reflection group. In either case, every element of finite order in $G(a, b, c)$ is elliptic.

By Theorem 6.5.6, every element of finite order in $G(a, b, c)$ is conjugate in $G(a, b, c)$ to an element that fixes a point on the boundary of the triangle $T(a, b, c)$. Let x, y, z be the vertices of $T(a, b, c)$ corresponding to the angles $\pi/a, \pi/b, \pi/c$. In view of the fact that

$$\{gT(a, b, c) : g \in G(a, b, c)\}$$

is a tessellation of X, the stabilizer subgroup of each side of $T(a, b, c)$ is the group of order two generated by the reflection in the corresponding side of $T(a, b, c)$. Furthermore, the stabilizer subgroup at the vertex x, y, or z is a dihedral group of order $2a, 2b$, or $2c$, respectively.

Let v be an arbitrary vertex of $T(a, b, c)$ and let G_v be the stabilizer subgroup at v. Then

$$\{gT(a, b, c) : g \in G_v\}$$

forms a cycle of triangles around the vertex v. Consequently, no two vertices of $T(a, b, c)$ are in the same orbit. Therefore, two elements in $G_x \cup G_y \cup G_z$ are conjugate in $G(a, b, c)$ if and only if they are conjugate in the same stabilizer G_v, since $gG_vg^{-1} = G_{gv}$. Hence, the integers $\{2, a, b, c\}$ are characterized by $G(a, b, c)$ as the orders of the maximal finite cyclic subgroups of $G(a, b, c)$. As this set is invariant under isomorphism, we have that $\{2, a, b, c\} = \{2, a', b', c'\}$. Therefore $(a, b, c) = (a', b', c')$. □

Barycentric Subdivision

Let P be an n-dimensional polytope in X. The *barycentric subdivision* of P is the subdivision of P into n-simplices whose vertices can be ordered $\{v_0, \ldots, v_n\}$ so that v_k is the centroid of a k-face F_k of P for each k, and F_k is a side of F_{k+1} for each $k = 0, \ldots, n-1$. In particular, all the simplices of the barycentric subdivision of P share the centriod of P as a common vertex, and the side of such a simplex opposite the centroid of P is part of the barycentric subdivision of a side of P. For example, Figure 7.1.2 illustrates the barycentric subdivision of a regular hexagon in B^2.

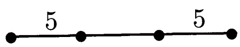
Tetrahedron Reflection Groups

We now consider some examples of tetrahedron reflection groups determined by regular tessellations of S^3, E^3, and H^3.

Example 1. Let P be a regular Euclidean 4-simplex inscribed in S^3. Then radial projection of ∂P onto S^3 gives a regular tessellation of S^3 by five tetrahedra. Now since three of these tetrahedra meet along each edge, their dihedral angle is $2\pi/3$. Let T be one of these tetrahedra. Then barycentric subdivision divides T into 24 congruent tetrahedra. Let Δ be one of these tetrahedra. Then the dihedral angles of Δ are all submultiples of π as indicated in Figure 7.2.4. Therefore, the group Γ generated by reflecting in the sides of Δ is a discrete reflection group with respect to Δ by Theorem 7.1.3. It is clear from the geometry of Δ and T that Γ is the group of symmetries of P. Therefore Γ is a symmetric group on five letters, and so Γ has order $5! = 120$. The Coxeter graph of Γ is

Example 2. Let P be a cube in E^3. The dihedral angle of P is $\pi/2$. Observe that barycentric subdivision divides P into 48 congruent tetrahedra. Let Δ be one of these tetrahedra. Then the dihedral angles of Δ are all submultiples of π as indicated in Figure 7.2.5. Therefore, the group Γ generated by reflecting in the sides of Δ is a discrete reflection group with respect to Δ by Theorem 7.1.3. It is worth noting that Γ is the group of symmetries of the regular tessellation of E^3 by cubes obtained by reflecting in the sides of P. The Coxeter graph of Γ is

Example 3. By the argument in Example 4 of §7.1, there is a hyperbolic regular dodecahedron P whose dihedral angle is $2\pi/5$. Observe that barycentric subdivision divides P into 120 congruent tetrahedra. Let Δ be one of these tetrahedra. Then the dihedral angles of Δ are all submultiples of π as indicated in Figure 7.2.6. Therefore, the group Γ generated by reflecting in the sides of Δ is a discrete reflection group with respect to Δ by Theorem 7.1.3. It is worth noting that Γ is the group of symmetries of the regular tessellation of H^3 by dodecahedra obtained by reflecting in the sides of P. The Coxeter graph of Γ is



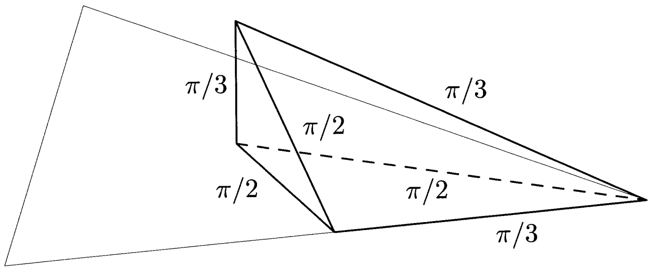


Figure 7.2.4. A spherical tetrahedron with dihedral angles submultiples of π

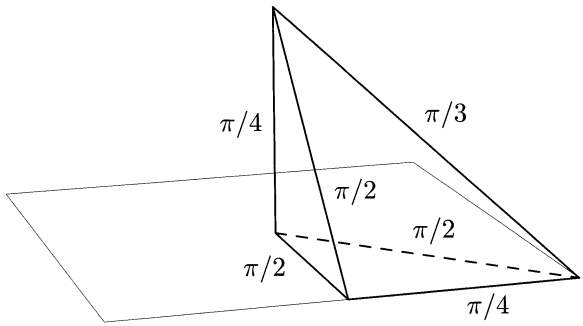


Figure 7.2.5. A Euclidean tetrahedron with dihedral angles submultiples of π

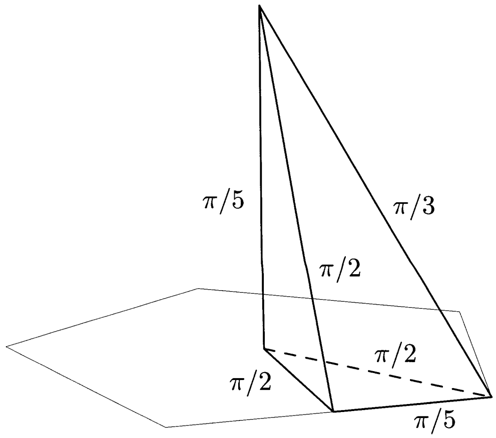


Figure 7.2.6. A hyperbolic tetrahedron with dihedral angles submultiples of π

Bilinear Forms

We now review some of the elementary theory of bilinear forms. Recall that a *bilinear form* on a real vector space V is a function from $V \times V$ to $\mathbb{R}$, denoted by $(v, w) \mapsto \langle v, w \rangle$, such that for all v, w in V,

(1) $\langle v, \ \rangle$ and $\langle \ , w \rangle$ are linear functions from V to $\mathbb{R}$ (bilinearity);

(2) $\langle v, w \rangle = \langle w, v \rangle$ (symmetry).

Moreover, $\langle \ , \ \rangle$ is said to be *nondegenerate* if and only if

(3) if $v \neq 0$, then there is a $w \neq 0$ such that $\langle v, w \rangle \neq 0$ (nondegeneracy).

Remark: A nondegenerate bilinear form on V is the same as an inner product on V.

A bilinear form $\langle \ , \ \rangle$ on V is said to be *positive semidefinite* if and only if

(4) $\langle v, v \rangle \geq 0$ for all v in V.

Finally, a bilinear form $\langle \ , \ \rangle$ on V is said to be *positive definite* if and only if

(5) $\langle v, v \rangle > 0$ for all nonzero v in V.

Now suppose that $\langle \ , \ \rangle$ is a bilinear form on $\mathbb{R}^n$. The *matrix* A of $\langle \ , \ \rangle$ is the real $n \times n$ matrix (a_{ij}) defined by

$$a_{ij} = \langle e_i, e_j \rangle.$$

Observe that A is a symmetric matrix. We say that A is *positive definite, positive semidefinite*, or *nondegenerate* according as $\langle \ , \ \rangle$ has the same property. By the Gram-Schmidt process, there is a basis $u_1, \ldots, u_n$ of $\mathbb{R}^n$ such that

$$\begin{aligned} \langle u_i, u_j \rangle &= 0 \quad \text{if } i \neq j, \\ \langle u_i, u_i \rangle &= \begin{cases} 1 & \text{if } 1 \leq i \leq p, \\ -1 & \text{if } p+1 \leq i \leq q, \\ 0 & \text{if } q+1 \leq i \leq n, \end{cases} \end{aligned}$$

where p, q are integers such that $0 \leq p \leq q \leq n$. Note that A is positive (semi) definite if and only if $p = n$ $(p = q)$, and A is nondegenerate if and only if $q = n$. Furthermore q is equal to the rank of A. The pair $(p, q-p)$ is called the *type* of A.

Given any real symmetric $n \times n$ matrix A, we define the *bilinear form* of A on $\mathbb{R}^n$ by the formula

$$\langle x, y \rangle = x \cdot Ay.$$

Clearly, A is the matrix of the bilinear form of A.

The *null space* of a bilinear form $\langle \ , \ \rangle$ on $\mathbb{R}^n$ is the set

$$\{y \in \mathbb{R}^n : \langle x, y \rangle = 0 \text{ for all } x \text{ in } \mathbb{R}^n\}.$$

Clearly, the null space of the bilinear form of a matrix A is the null space of A.

Definition: The *Gram matrix* of an n-simplex Δ in X whose sides are $S_1, \ldots, S_{n+1}$ is the $(n+1)\times(n+1)$ matrix whose ijth entry is $-\cos\theta(S_i, S_j)$.

Theorem 7.2.2. *Let θ_{ij}, for $i, j = 1, \ldots, n+1$, be real numbers such that*

(1) *$\theta_{ij} = \theta_{ji}$ for all i, j,*

(2) *$\theta_{ii} = \pi$ for each i, and*

(3) *θ_{ij} is in the interval $(0, \pi/2]$ if $i \neq j$.*

Let A be the $(n+1)\times(n+1)$ matrix whose ijth entry is $-\cos\theta_{ij}$ and let A_i be the $n \times n$ matrix obtained from A by deleting the ith row and ith column. Then there is an n-simplex Δ in either S^n, E^n, or H^n whose Gram matrix is A if and only if A_i is positive definite for each $i = 1, \ldots, n+1$. Furthermore Δ is

(1) *spherical if and only if A is positive definite,*

(2) *Euclidean if and only if A is of type $(n, 0)$,*

(3) *hyperbolic if and only if A is of type $(n, 1)$.*

Proof: (1) Suppose that Δ is an n-simplex in S^n with sides $S_1, \ldots, S_{n+1}$ such that $\theta(S_i, S_j) = \theta_{ij}$ for all $i, j = 1, \ldots, n+1$. Let $\langle S_i\rangle$ be the hyperplane of S^n containing S_i and let V_i be the n-dimensional vector subspace of $\mathbb{R}^{n+1}$ such that

$$\langle S_i\rangle = V_i \cap S^n.$$

Let H_i be the half-space of $\mathbb{R}^{n+1}$ bounded by V_i and containing Δ. Then

$$\Delta = \left(\bigcap_{i=1}^{n+1} H_i\right) \cap S^n.$$

Let v_i be the unit normal of V_i directed into H_i. Then

$$H_i = \{x \in \mathbb{R}^{n+1} : x \cdot v_i \geq 0\}.$$

Let B be the $(n+1)\times(n+1)$ matrix whose jth column vector is v_j. Then the orthogonal complement of the column space of B is the set

$$\{x \in \mathbb{R}^{n+1} : x \cdot v_i = 0 \ \text{ for } i = 1, \ldots, n+1\}.$$

But this set is $\cap_{i=1}^{n+1} V_i = \{0\}$. Therefore $v_1, \ldots, v_{n+1}$ form a basis of $\mathbb{R}^{n+1}$. Thus B is nonsingular.

Next, define a positive definite inner product in $\mathbb{R}^{n+1}$ by the formula

$$\langle x, y\rangle = Bx \cdot By.$$

Then for each i, j, we have

$$\begin{aligned}\langle e_i, e_j\rangle &= Be_i \cdot Be_j \\ &= v_i \cdot v_j \\ &= \cos(\pi - \theta_{ij}) \ = \ -\cos\theta_{ij}.\end{aligned}$$

Therefore A is the matrix of this inner product, and so A is positive definite. Furthermore, A_i is positive definite for each $i = 1, \ldots, n+1$.

Conversely, suppose that A is positive definite. Then there is an orthonormal basis $u_1, \ldots, u_{n+1}$ of $\mathbb{R}^{n+1}$ with respect to the inner product of A. Let B be the $(n+1) \times (n+1)$ matrix whose jth column vector is u_j. Then $B^t A B = I$. Let $C = B^{-1}$. Then $A = C^t C$. Let v_j be the jth column vector of C. Then $v_1, \ldots, v_{n+1}$ form a basis of $\mathbb{R}^{n+1}$ and $A = (v_i \cdot v_j)$. Let

$$Q = \{y \in \mathbb{R}^{n+1} : y_i \geq 0 \quad \text{for } i = 1, \ldots, n+1\}.$$

Then the set Q is an $(n+1)$-dimensional convex polyhedron in E^{n+1} with $n+1$ sides and one vertex at the origin.

Now let

$$H_i = \{x \in \mathbb{R}^{n+1} : v_i \cdot x \geq 0\}$$

and

$$V_i = \{x \in \mathbb{R}^{n+1} : v_i \cdot x = 0\}.$$

Observe that

$$C^t \left(\bigcap_{i=1}^{n+1} H_i \right) \subset Q.$$

Let y be an arbitrary vector in Q. Set $x = B^t y$. Then $C^t x = y$. Hence $v_i \cdot x \geq 0$ for all i, and so x is in $\cap_{i=1}^{n+1} H_i$. Therefore

$$C^t \left(\bigcap_{i=1}^{n+1} H_i \right) = Q.$$

Hence $\cap_{i=1}^{n+1} H_i$ is an $(n+1)$-dimensional convex polyhedron in E^{n+1} with $n+1$ sides

$$V_i \cap \left(\bigcap_{j=1}^{n+1} H_j \right) \quad \text{for } i = 1, \ldots, n+1$$

and exactly one vertex at the origin. Therefore

$$\Delta = \left(\bigcap_{i=1}^{n+1} H_i \right) \cap S^n$$

is an n-dimensional convex polyhedron in S^n with sides

$$S_i = V_i \cap \left(\bigcap_{j=1}^{n+1} H_j \right) \cap S^n \quad \text{for } i = 1, \ldots, n+1.$$

Moreover Δ is contained in an open hemisphere of S^n by induction on n and Theorem 6.3.8. Therefore Δ is an n-simplex in S^n by Theorem 6.4.4. Furthermore, for all i, j, we have

$$\theta(S_i, S_j) = \pi - \theta(v_i, v_j) = \pi - (\pi - \theta_{ij}) = \theta_{ij}.$$

(2) Suppose that Δ is an n-simplex in E^n with sides $S_1, \ldots, S_{n+1}$ such that $\theta(S_i, S_j) = \theta_{ij}$ for all $i, j = 1, \ldots, n+1$. Let P_i be the hyperplane of $\mathbb{R}^n$ containing S_i and let H_i be the half-space of $\mathbb{R}^n$ bounded by P_i and containing Δ. Then

$$\Delta = \bigcap_{i=1}^{n+1} H_i.$$

Let v_i be the unit normal of P_i directed into H_i. Then for all i, j, we have

$$v_i \cdot v_j = \cos(\pi - \theta_{ij}) = -\cos\theta_{ij}.$$

By translating Δ, if necessary, we may assume that the vertex of Δ opposite the side S_j is the origin. Then the set

$$\left(\bigcap_{\substack{i=1 \\ i\neq j}}^{n+1} H_i\right) \cap S^{n-1}$$

is an $(n-1)$-simplex in S^{n-1}. By the argument in (1), we have that $v_1, \ldots, \hat{v}_j, \ldots, v_{n+1}$ form a basis of $\mathbb{R}^n$ and A_j is positive definite; moreover, this is true for each $j = 1, \ldots, n+1$.

Let B be the $n \times (n+1)$ matrix whose jth column is v_j for $j = 1, \ldots, n+1$. Define a bilinear form on $\mathbb{R}^{n+1}$ by the formula

$$\langle x, y\rangle = Bx \cdot By.$$

Then the matrix of this form is A. Moreover, the null space of this form is the null space of B. As the rank of B is n, the null space of B is 1-dimensional. Therefore, the null space of the bilinear form of A is 1-dimensional. Hence A is of type $(n, 0)$.

Conversely, suppose that A is of type $(n, 0)$ and A_i is positive definite for each $i = 1, \ldots, n+1$. Consider the bilinear form of A defined by

$$\langle x, y\rangle = x \cdot Ay.$$

Clearly, the null space of the form is the orthogonal complement in $\mathbb{R}^{n+1}$ of the column space of A. Let x be a nonzero vector in the null space of A. Then each component x_i of x is nonzero, since A_i is positive definite for each $i = 1, \ldots, n+1$. Define y and z in $\mathbb{R}^{n+1}$ by

$$y_i = \begin{cases} x_i & \text{if } x_i > 0 \\ 0 & \text{if } x_i < 0, \end{cases}$$

$$z_i = \begin{cases} x_i & \text{if } x_i < 0 \\ 0 & \text{if } x_i > 0. \end{cases}$$

Then $x = y + z$. As $\langle x, y\rangle = 0$, we have

$$\langle y, y\rangle + \langle y, z\rangle = 0.$$

Now observe that

$$\langle y, z\rangle = \sum_{i\neq j} a_{ij} y_i z_j.$$

As $a_{ij} \leq 0$ and $y_i z_j \leq 0$ for each $i \neq j$, we have that $\langle y, z\rangle \geq 0$. Therefore $\langle y, y\rangle = 0$, since A is positive semidefinite. If $z \neq 0$, then some component of y is zero, and so $y = 0$, since A_i is positive definite for each i. Hence, either $y = 0$ or $z = 0$. Thus, all the components of x have the same sign.

Now as A is of type $(n, 0)$, there is a nonsingular $(n+1) \times (n+1)$ matrix C such that

$$A = C^t \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & 0 \end{pmatrix} C.$$

Let v_j be the jth column vector of C and let $\overline{v}_j$ be the vector in $\mathbb{R}^n$ obtained by dropping the last coordinate of v_j. Then $A = (\overline{v}_i \cdot \overline{v}_j)$.

Let $\overline{C}$ be the $n \times n$ matrix whose jth column vector is $\overline{v}_j$. Then

$$\overline{C}e_i \cdot \overline{C}e_j = \overline{v}_i \cdot \overline{v}_j.$$

Hence, the restriction of the bilinear form of A to $\mathbb{R}^n$ is given by

$$\langle x, y \rangle = \overline{C}x \cdot \overline{C}y.$$

As A_{n+1} is positive definite, the matrix $\overline{C}$ must be nonsingular. Therefore $\overline{v}_1, \ldots, \overline{v}_n$ form a basis of $\mathbb{R}^n$. Furthermore

$$\overline{v}_i \cdot \overline{v}_i = -\cos \pi = 1 \quad \text{for each } i = 1, \ldots, n+1.$$

Now let

$$H_i = \{x \in \mathbb{R}^n : \overline{v}_i \cdot x \geq 0\}$$

and

$$V_i = \{x \in \mathbb{R}^n : \overline{v}_i \cdot x = 0\}.$$

Let B be the $(n+1) \times n$ matrix whose ith row is $\overline{v}_i$. As $BB^t = A$, the column space of B is the column space of A. Suppose that x is in $\cap_{i=1}^{n+1} H_i$. Then $\overline{v}_i \cdot x \geq 0$ for each $i = 1, \ldots, n+1$. Hence, each component of Bx is nonnegative. Let y be a nonzero vector in the null space of A. Then y is orthogonal to the column space of A. Hence $(Bx) \cdot y = 0$. As all the components of y have the same sign, we deduce that $Bx = 0$. Therefore x is in $\cap_{i=1}^{n+1} V_i = \{0\}$. Thus $\cap_{i=1}^{n+1} H_i = \{0\}$.

Let

$$H_0 = \{x \in \mathbb{R}^n : \overline{v}_{n+1} \cdot x \geq -1\}$$

and let

$$\Delta = \bigcap_{i=0}^{n} H_i.$$

By applying the linear isomorphism $\overline{C}^t$, we may assume, without loss of generality, that $\overline{v}_i = e_i$ for each $i = 1, \ldots, n$. Then for each $j = 1, \ldots, n$,

$$\overline{v}_{n+1} \cdot e_j = \overline{v}_{n+1} \cdot \overline{v}_j = -\cos \theta_{n+1,j} \leq 0.$$

Now if $\overline{v}_{n+1} \cdot e_j = 0$, then e_j is in $\cap_{i=1}^{n+1} H_i = \{0\}$, which is a contradiction. Hence, all the coordinates of $\overline{v}_{n+1}$ are negative. Therefore Δ is the n-simplex bounded by the n coordinate hyperplanes of $\mathbb{R}^n$ and the hyperplane

$$V_0 = \{x \in \mathbb{R}^n : (-\overline{v}_{n+1}) \cdot x = 1\}.$$

Returning to the general case, we find that the $n+1$ sides of the n-simplex Δ are

$$S_i = V_i \cap \left(\bigcap_{j=0}^{n} H_j\right) \quad \text{for } i = 0, \ldots, n,$$

and for all i, j mod $(n+1)$, we have $\theta(S_i, S_j) = \theta_{ij}$.

(3) Suppose that Δ is an n-simplex in H^n with sides $S_1, \ldots, S_{n+1}$ such that $\theta(S_i, S_j) = \theta_{ij}$ for all $i, j = 1, \ldots, n+1$. Let $\langle S_i \rangle$ be the hyperplane of H^n containing S_i and let V_i be the n-dimensional, time-like, vector subspace of $\mathbb{R}^{n,1}$ such that

$$\langle S_i \rangle = V_i \cap H^n.$$

Let H_i be the half-space of $\mathbb{R}^{n+1}$ bounded by V_i and containing Δ. Then

$$\Delta = \left(\bigcap_{i=1}^{n+1} H_i\right) \cap H^n.$$

Let v_i be the unit Lorentz normal of V_i directed into H_i. Then

$$H_i = \{x \in \mathbb{R}^{n,1} : x \circ v_i \geq 0\}.$$

Let B be the $(n+1) \times (n+1)$ matrix whose jth column vector is v_j. Then the Lorentz orthogonal complement of the column space of B is the set

$$\{x \in \mathbb{R}^{n,1} : x \circ v_i = 0 \quad \text{for} \quad i = 1, \ldots, n+1\}.$$

But this set is $\cap_{i=1}^{n+1} V_i = \{0\}$. Therefore $v_1, \ldots, v_{n+1}$ form a basis of $\mathbb{R}^{n+1}$. Thus B is nonsingular.

Next, define a bilinear form on $\mathbb{R}^{n+1}$ of type $(n, 1)$ by the formula

$$\langle x, y \rangle = Bx \circ By.$$

Then for all i, j, we have

$$\begin{aligned} \langle e_i, e_j \rangle &= Be_i \circ Be_j \\ &= v_i \circ v_j \\ &= \cos(\pi - \theta_{ij}) \quad = \quad -\cos\theta_{ij}. \end{aligned}$$

Hence A is the matrix of this form, and so A is of type $(n, 1)$.

Let u_k be the vertex of Δ opposite the side S_k and let r_k be half the distance from u_k to S_k in H^n. Then the set

$$\Delta' = S(u_k, r_k) \cap \Delta$$

is a spherical $(n-1)$-simplex with sides

$$S_i' = S_i \cap S(u_k, r_k) \quad \text{for } i \neq k.$$

Furthermore, we have

$$\theta(S_i', S_j') = \theta(S_i, S_j) \quad \text{for } i, j \neq k.$$

Therefore A_k is positive definite by (1) for each $k = 1, \ldots, n+1$.

Conversely, suppose that A is of type $(n,1)$ and A_i is positive definite for each $i = 1,\ldots,n+1$. Let J be the diagonal $(n+1)\times(n+1)$ matrix with diagonal entries $1,\ldots,1,-1$. Then there is a nonsingular $(n+1)\times(n+1)$ matrix C such that $A = C^tJC$. Let v_j be the jth column vector of C. Then $v_1,\ldots,v_{n+1}$ form a basis of $\mathbb{R}^{n+1}$ and $A = (v_i \circ v_j)$. Let

$$Q = \{y \in \mathbb{R}^{n+1} : y_i \geq 0 \quad \text{for } i = 1,\ldots,n+1\}.$$

Then the set Q is an $(n+1)$-dimensional convex polyhedron in E^{n+1} with $n+1$ sides and one vertex at the origin.

Now let

$$H_i = \{x \in \mathbb{R}^{n,1} : v_i \circ x \geq 0\}$$

and

$$V_i = \{x \in \mathbb{R}^{n,1} : v_i \circ x = 0\}.$$

Then H_i is a half-space of $\mathbb{R}^{n+1}$ bounded by the n-dimensional vector subspace V_i of $\mathbb{R}^{n+1}$. As before, we have

$$C^tJ\left(\bigcap_{i=1}^{n+1} H_i\right) = Q.$$

Therefore $\cap_{i=1}^{n+1} H_i$ is an $(n+1)$-dimensional convex polyhedron in E^{n+1} with $n+1$ sides

$$V_i \cap \left(\bigcap_{j=1}^{n+1} H_j\right) \quad \text{for } i = 1,\ldots,n+1$$

and exactly one vertex at the origin. As

$$v_i \circ v_i = -\cos\pi = 1,$$

we have that v_i is space-like, and so V_i is time-like. Therefore, the set

$$\Delta = \left(\bigcap_{i=1}^{n+1} H_i\right) \cap H^n$$

is an n-dimensional convex polyhedron in H^n with sides

$$S_i = V_i \cap \left(\bigcap_{j=1}^{n+1} H_j\right) \cap H^n \quad \text{for } i = 1,\ldots,n+1.$$

Furthermore, for all i,j, we have

$$\begin{aligned}\theta(S_i,S_j) &= \pi - \theta(v_i,v_j)\\ &= \pi - (\pi - \theta_{ij}) = \theta_{ij}.\end{aligned}$$

It remains only to show that Δ is compact. Define a bilinear form on $\mathbb{R}^{n+1}$ by the formula

$$\langle x,y\rangle = Cx \circ Cy.$$

Then the matrix of this form is A. As A_j is positive definite, this form is positive definite on the vector subspace $\langle e_1,\ldots,\hat{e}_j,\ldots,e_{n+1}\rangle$. Hence, the Lorentzian inner product on $\mathbb{R}^{n,1}$ is positive definite on the vector subspace

$$W_j = \langle v_1,\ldots,\hat{v}_j,\ldots,v_{n+1}\rangle.$$

Therefore W_j is space-like. Let

$$L_j = \bigcap_{\substack{i=1 \\ i \neq j}}^{n+1} V_i.$$

Then L_j is the 1-dimensional vector subspace of $\mathbb{R}^{n+1}$ spanned by the 1-dimensional edge of $\cap_{i=1}^{n+1} H_i$ that is opposite the side

$$V_j \cap \left(\bigcap_{i=1}^{n+1} H_i \right).$$

Observe that

$$L_j = \{x \in \mathbb{R}^{n,1} : x \circ v_i = 0 \ \text{ for all } \ i \neq j\}.$$

Hence L_j is the Lorentz orthogonal complement of W_j. Consequently L_j is time-like. Hence $L_j \cap H^n$ is a vertex of Δ. Thus Δ has vertices $L_j \cap H^n$ for $j = 1, \ldots, n+1$. Therefore Δ is compact by induction on n and Theorem 6.3.7. Thus Δ is an n-simplex in H^n. This completes the proof of (3).

In order to complete the proof, we need to prove that if A_i is positive definite for each $i = 1, \ldots, n+1$, then A is either positive definite or of type $(n, 0)$ or $(n, 1)$. This is left as an exercise for the reader. □

Classification of Simplex Reflection Groups

Let Γ be the group generated by the reflections of X in the sides of an n-simplex Δ all of whose dihedral angles are submultiples of π. Let v be a vertex of Δ and let Γ_v be the subgroup of Γ consisting of the elements of Γ fixing v. Then Γ_v is a spherical $(n-1)$-simplex reflection group. Moreover, the subgraph of the Coxeter graph of Γ, obtained by deleting the vertex corresponding to the side of Δ opposite v and its adjoining edges, is the Coxeter graph of Γ_v. By induction, every subgraph of the Coxeter graph of Γ obtained by deleting vertices and their adjoining edges is the Coxeter graph of a spherical simplex reflection group.

The group Γ is said to be *irreducible* if and only if its Coxeter graph is connected. Suppose that Γ is irreducible. Then we can delete vertices and their adjoining edges from the Coxeter graph of Γ so that after each deletion we obtain a connected subgraph. Now the only labels on the irreducible spherical triangle reflection groups are 3, 4, and 5. Therefore, if $n > 2$, the Coxeter graph of Γ has only 3, 4, and 5 as possible labels. Hence, there are only finitely many possible Coxeter graphs of n-simplex reflection groups for each $n > 2$. In view of Theorem 7.2.2, it is straightforward to list all the possible Coxeter graphs of n-simplex reflections groups for a given n. Spherical and Euclidean n-simplex reflection groups exist in all dimensions n; however, hyperbolic n-simplex reflection groups exist only for dimensions $n \leq 4$. Figures 7.2.7-7.2.9 illustrate the Coxeter graphs of all the irreducible, simplex, reflection groups.

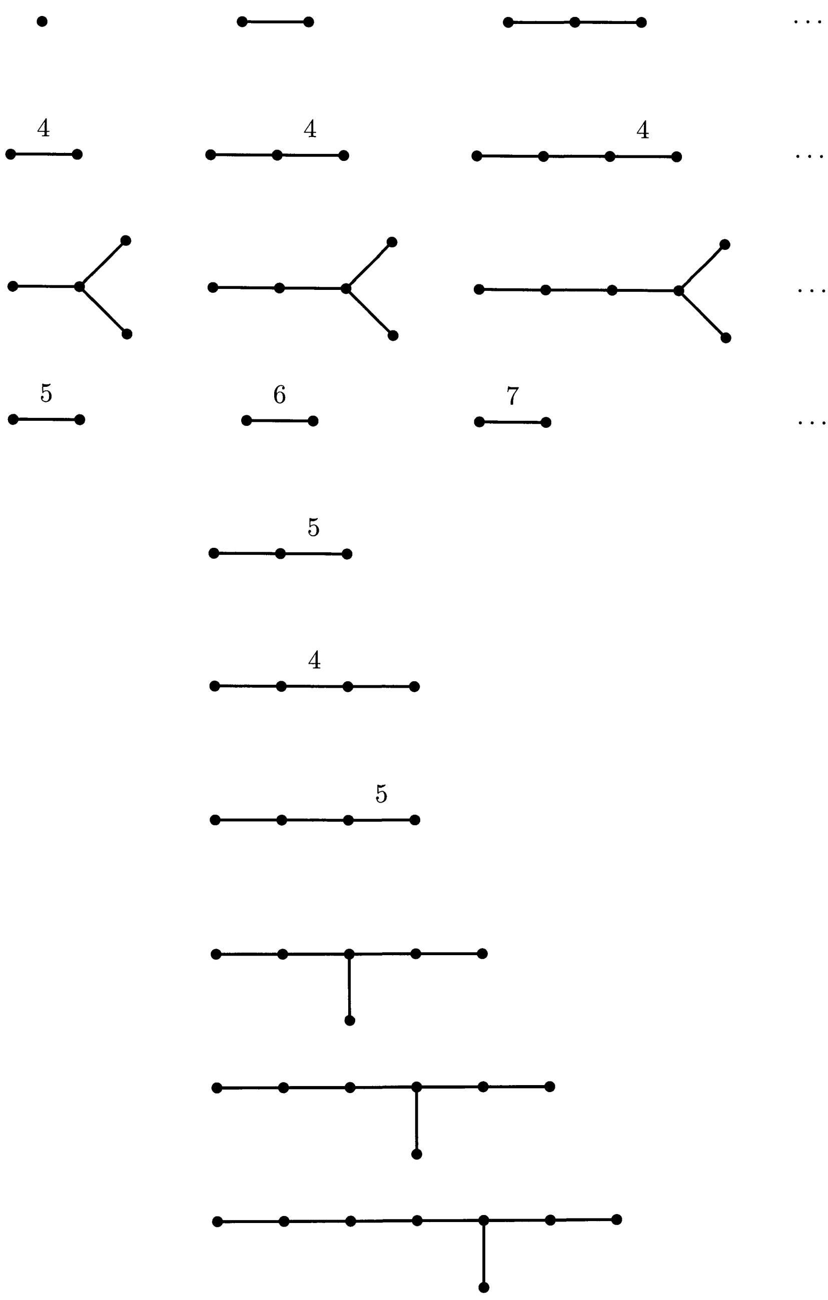


Figure 7.2.7. The irreducible, spherical, simplex, reflection groups

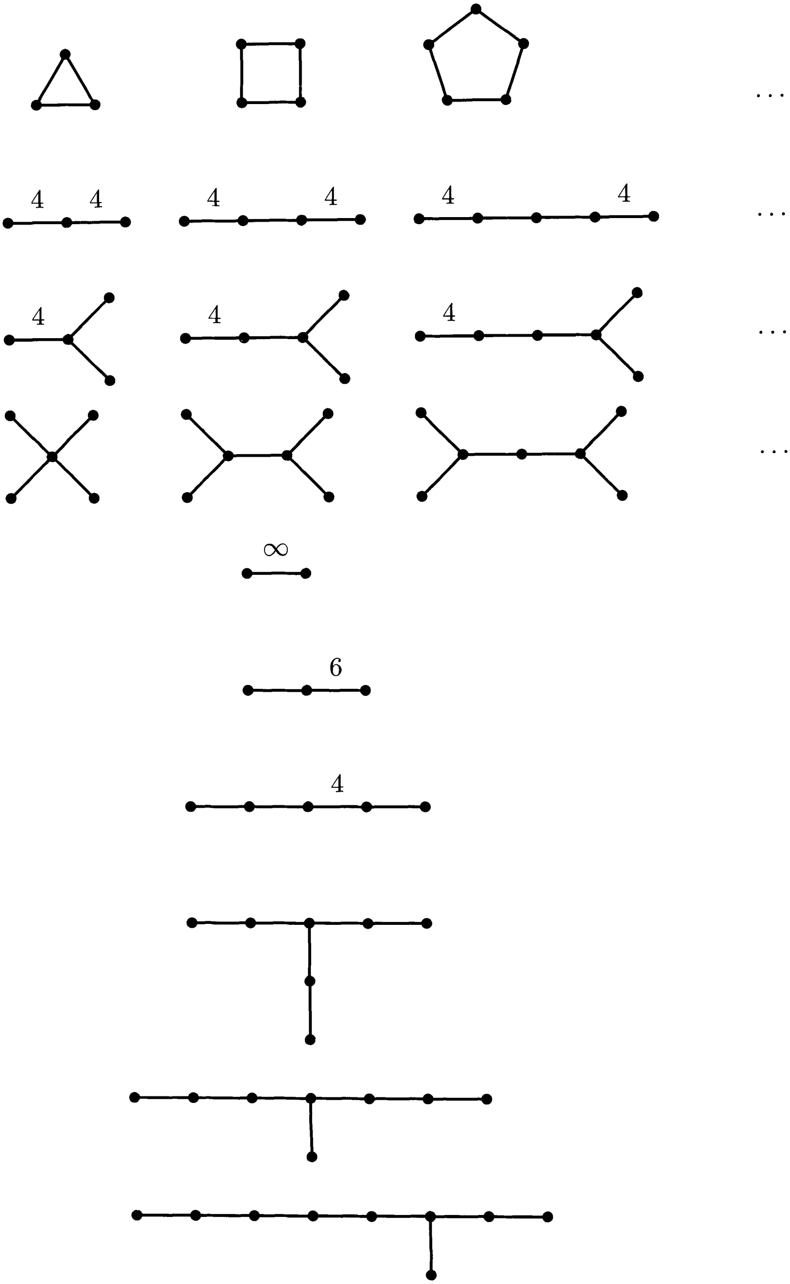


Figure 7.2.8. The Euclidean, simplex, reflection groups

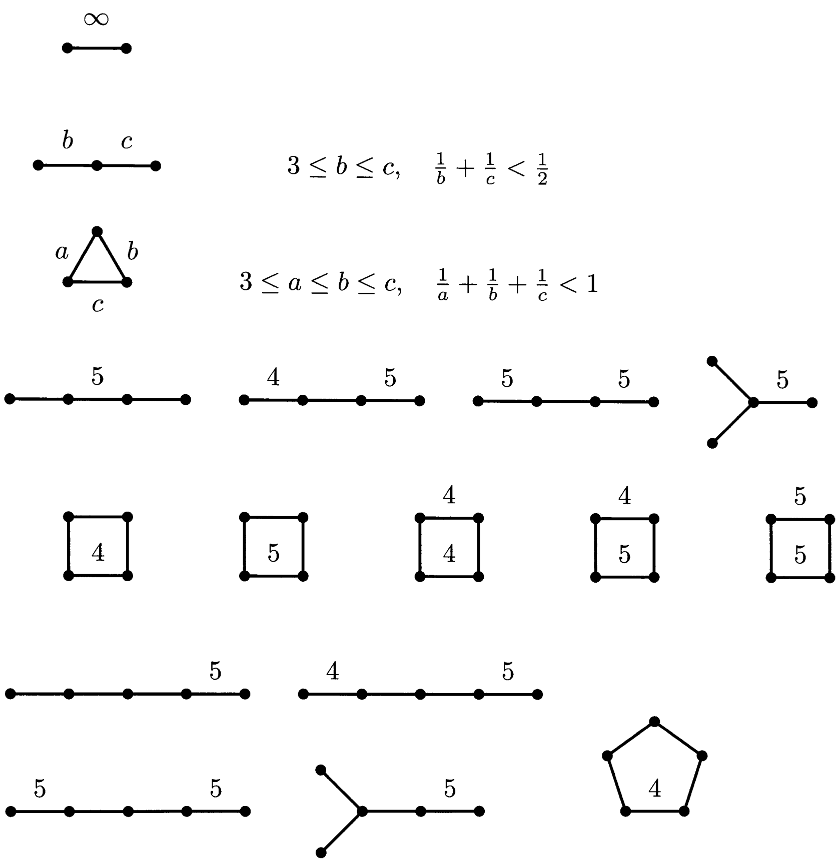


Figure 7.2.9. The hyperbolic, simplex, reflection groups

Exercise 7.2

1. Prove that $G_0(2, 3, 4)$ is a symmetric group on four letters and $G_0(2, 3, 5)$ is an alternating group on five letters.

2. Prove that $T(2, 3, 7)$ is the triangle of least hyperbolic area among all the hyperbolic triangles $T(a, b, c)$.

3. Prove that $G(2, 4, 6)$ contains the group Γ in Example 3 of §7.1 as a subgroup of index 12.

4. Let a, b, c, a', b', c' be integers such that $2 \leq a \leq b \leq c$ and $2 \leq a' \leq b' \leq c'$. Prove that the triangle groups $G_0(a, b, c)$ and $G_0(a', b', c')$ are isomorphic if and only if $(a, b, c) = (a', b', c')$.

5. Let f, g, h be the reflections in the sides of $T(a, b, c)$ opposite the angles $\pi/a, \pi/b, \pi/c$. Prove that $G_0(a, b, c)$ has the group presentation
$$\left(u, v; u^c, v^b, (uv)^a\right),$$
where $u \mapsto gf$ and $v \mapsto fh$.

6. Prove that the group of symmetries of an $(n + 1)$-dimensional, Euclidean, regular polytope inscribed in S^n is isomorphic to a spherical, n-simplex, reflection group.

7. Prove the regular tessellations of S^n correspond under radial projection to the $(n + 1)$-dimensional, Euclidean, regular polytopes inscribed in S^n.

8. Prove that the group of symmetries of a regular tessellation of X is an n-simplex reflection group.

9. Let A be as in Theorem 7.2.2 and suppose that A_i is positive definite for each $i = 1, \ldots, n+1$. Prove that A is either positive definite or of type $(n, 0)$ or $(n, 1)$ according as $\det A$ is positive, zero, or negative.

10. Prove that every Euclidean or hyperbolic simplex reflection group is irreducible.

11. Prove that every hyperbolic n-simplex reflection group is nonelementary when $n > 1$.

§7.3. Generalized Simplex Reflection Groups

Let Δ be a generalized n-simplex in H^n all of whose dihedral angles are submultiples of π. Then the group Γ generated by the reflections of H^n in the sides of Δ is a discrete group of isometries of H^n by Theorem 7.1.3. The group Γ is called a (generalized) simplex reflection group. Figure 7.3.1 illustrates the Coxeter graphs of the hyperbolic, noncompact triangle, reflection groups. Figure 7.3.2 illustrates the tessellation of B^2 obtained by reflecting in the sides of an ideal triangle.

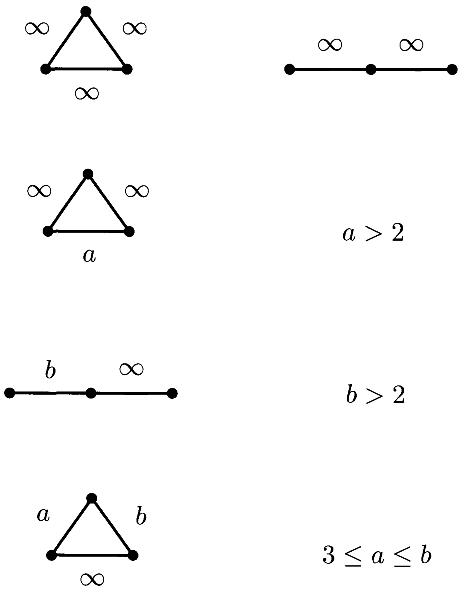


Figure 7.3.1. The hyperbolic, noncompact triangle, reflection groups

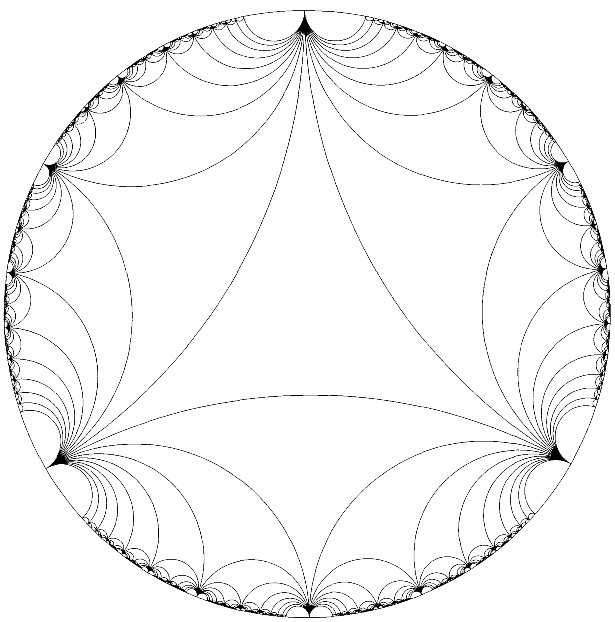


Figure 7.3.2. Tessellation of B^2 obtained by reflecting an ideal triangle

Example: Let Γ be the subgroup of $\mathrm{PO}(2,1)$ of all the matrices with integral entries. Then Γ is a discrete subgroup of $\mathrm{PO}(2,1)$, since Γ is a subgroup of the discrete group $\mathrm{GL}(3,\mathbb{Z})$. We now show that Γ is a discrete reflection group with respect to a triangle $T(2,4,\infty)$ in H^2. Clearly Γ acts on the set $S = H^2 \cap \mathbb{Z}^3$. Observe that the point $e_3 = (0,0,1)$ is in S. The stabilizer of e_3 in Γ is isomorphic to $\mathrm{O}(2) \cap \mathrm{GL}(2,\mathbb{Z})$. Hence Γ is a dihedral group of order eight generated by the $90°$ rotation about the z-axis and the reflection in the xz-plane.

Observe that the points of $S - \{e_3\}$ nearest to e_3 are the four points $(\pm 2, \pm 2, 3)$. Let A be the Lorentzian matrix that represents the unique reflection of H^2 that maps e_3 to $(2,2,3)$. Then $A = A^{-1} = (JAJ)^t$. Therefore A is of the form

$$\begin{pmatrix} a & b & 2 \\ b & c & 2 \\ -2 & -2 & 3 \end{pmatrix}.$$

From the information that the columns of A form a Lorentz orthonormal basis of $\mathbb{R}^{2,1}$ and $\det A = -1$, we deduce that

$$A = \begin{pmatrix} -1 & -2 & 2 \\ -2 & -1 & 2 \\ -2 & -2 & 3 \end{pmatrix}.$$

Therefore A is in Γ. Observe that A fixes the plane $z = x + y$. Hence A fixes the hyperbolic line of H^2 given by the conditions

$$z = x + y, \quad x^2 + y^2 - z^2 = -1, \quad z > 0.$$

Substituting the first equation into the second, we see that A fixes the hyperbolic line of H^2 given by the equation $xy = 1/2$.

Next, observe that the reflections

$$(x,y,z) \mapsto (x,-y,z) \quad \text{and} \quad (x,y,z) \mapsto (y,x,z)$$

fix the hyperbolic lines $y = 0$ and $x = y$, respectively, of H^2. Let T be the triangle in H^2 defined by the inequalities

$$xy \leq 1/2, \quad y \geq 0, \quad x \geq y.$$

Then clearly $T = T(2,4,\infty)$. See Figure 7.3.3. Let Γ_1 be the subgroup of Γ generated by the matrices representing the reflections in the sides of T. Then Γ_1 is a discrete reflection group with respect to T.

Let g be an element of Γ. Then there is an f in Γ_1 such that fge_3 is in T. Clearly e_3 is the only point of S contained in T. Therefore $fge_3 = e_3$. Thus fg is in the stabilizer of e_3 in Γ. As the stabilizer of e_3 in Γ is a subgroup of Γ_1, we have that g is in Γ_1. Therefore $\Gamma = \Gamma_1$. Thus Γ is a triangle reflection group with respect to $T(2,4,\infty)$. A nice consequence of this fact is that the set S of integral points of H^2 is the set of hyperbolic centers of all the ideal squares of the tessellation of H^2 in Figure 7.3.4.

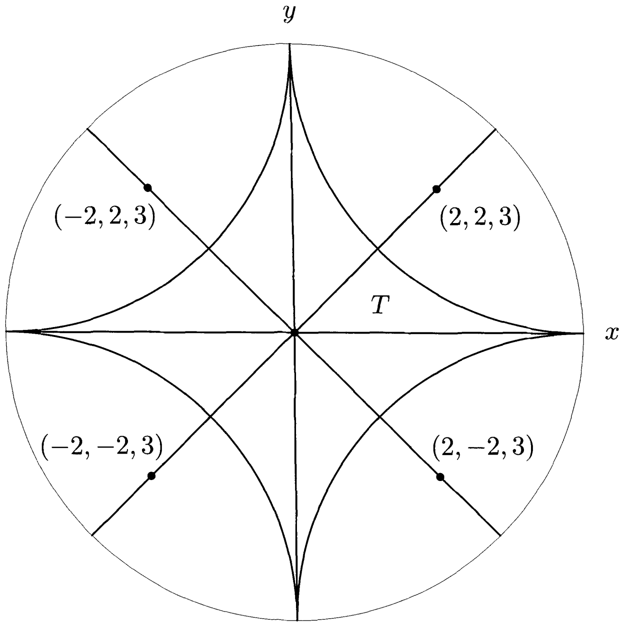


Figure 7.3.3. A triangle $T(2,4,\infty)$ in H^2

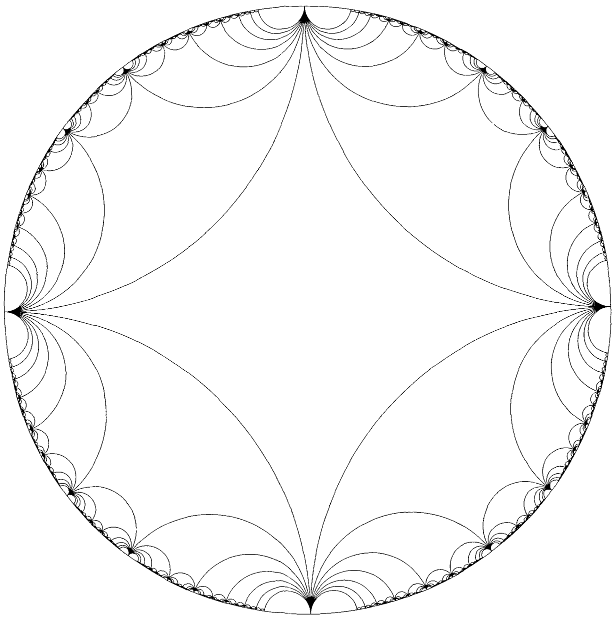


Figure 7.3.4. Tessellation of the unit disk by ideal squares

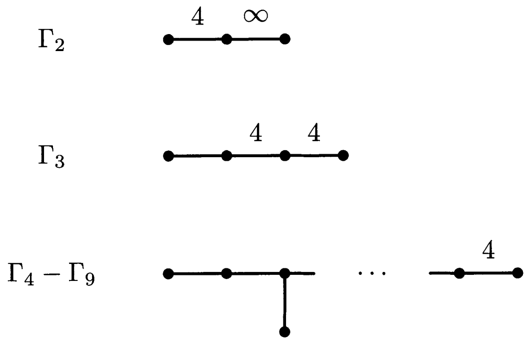


Figure 7.3.5. Coxeter graphs of the groups Γ_n for $n = 2, \ldots, 9$

Let Γ_n be the subgroup of $\mathrm{PO}(n,1)$ consisting of all the matrices with integral entries. Then Γ_n is a discrete subgroup of $\mathrm{PO}(n,1)$, since Γ_n is a subgroup of the discrete group $\mathrm{GL}(n+1,\mathbb{Z})$. The group Γ_n is a hyperbolic, noncompact n-simplex, reflection group for $n = 2, 3, \ldots, 9$. The Coxeter graphs of these groups are listed in Figure 7.3.5.

Definition: The *Gram matrix* of a generalized n-simplex Δ in H^n, with sides $S_1, \ldots, S_{n+1}$, is the $(n+1) \times (n+1)$ matrix whose ijth entry is $-\cos\theta(S_i, S_j)$.

Theorem 7.3.1. *Let θ_{ij}, for $i, j = 1, \ldots, n+1$, be real numbers such that*

(1) $\theta_{ij} = \theta_{ji}$ *for all* i, j,

(2) $\theta_{ii} = \pi$ *for each* i, *and*

(3) θ_{ij} *is in the interval* $[0, \pi/2]$ *if* $i \neq j$.

Let A be the $(n+1) \times (n+1)$ matrix whose ijth entry is $-\cos\theta_{ij}$ and let A_i be the $n \times n$ matrix obtained from A by deleting the ith row and ith column. Then there is a noncompact generalized n-simplex Δ in H^n whose Gram matrix is A if and only if

(1) *every column of A has more than one nonzero entry;*

(2) *the matrix A_i is the Gram matrix of either a spherical or Euclidean $(n-1)$-simplex for each $i = 1, \ldots, n+1$; and*

(3) *the matrix A_i is the Gram matrix of a Euclidean $(n-1)$-simplex for some i.*

Proof: Suppose that Δ is a noncompact generalized n-simplex in H^n with sides $S_1, \ldots, S_{n+1}$ such that $\theta(S_i, S_j) = \theta_{ij}$ for all $i, j = 1, \ldots, n+1$. If the vertex of Δ opposite the side S_k is finite, then A_k is the Gram matrix of a spherical $(n-1)$-simplex by the same argument as in the proof of Theorem 7.2.2(3).

Suppose that the vertex of Δ opposite the side S_k is ideal. We pass to the upper half-space model U^n. Then we may assume, without loss of generality, that the ideal vertex of Δ opposite the side S_k is ∞. Let B be a horoball based at ∞ such that $\overline{B}$ does not meet S_k. Then $\Delta' = \partial B \cap \Delta$ is a Euclidean $(n-1)$-simplex with sides $S_i' = S_i \cap \partial B$ for $i \neq k$ by Theorem 6.3.23. Clearly, we have

$$\theta(S_i', S_j') = \theta(S_i, S_j) \quad \text{for } i, j \neq k.$$

Therefore A_k is the Gram matrix of the Euclidean $(n-1)$-simplex Δ'. By our hypothesis, Δ has at least one ideal vertex. Hence A_i is the Gram matrix of a Euclidean $(n-1)$-simplex for some i.

Let $v_1, \ldots, v_{n+1}$ in $\mathbb{R}^{n+1}$ be defined as in the proof of Theorem 7.2.2(3). Then for each i, j, we have

$$v_i \circ v_j = -\cos\theta_{ij}.$$

Let C be the $(n+1)\times(n+1)$ matrix whose jth column vector is v_j. Define a bilinear form on $\mathbb{R}^{n+1}$ by the formula

$$\langle x, y\rangle = Cx \circ Cy.$$

Then A is the matrix of this form. As A_j is positive semidefinite, this form is positive semidefinite on the vector subspace $\langle e_1, \ldots, \hat{e}_j, \ldots, e_{n+1}\rangle$. Hence, the Lorentzian inner product on $\mathbb{R}^{n,1}$ is positive semidefinite on the vector subspace

$$W_j = \langle v_1, \ldots, \hat{v}_j, \ldots, v_{n+1}\rangle.$$

Therefore W_j is either space-like or light-like.

On the contrary, suppose that the jth column of A has only one nonzero entry, namely, $-\cos\theta_{jj} = 1$. Then v_j is Lorentz orthogonal to W_j. Therefore v_j is either time-like or light-like. But $v_j \circ v_j = 1$, and so we have a contradiction. Thus, every column of A must have at least two nonzero entries. Thus A satisfies (1)-(3).

Conversely, suppose that A satisfies (1)-(3). Then A_i is the Gram matrix of a Euclidean $(n-1)$-simplex for some i. By reindexing, if necessary, we may assume that A_{n+1} is the Gram matrix of a Euclidean $(n-1)$-simplex. Then $\mathbb{R}^n$ has a basis $\{u_1, \ldots, u_n\}$ such that $\langle u_i, u_j\rangle = 0$ if $i \neq j$, and $\langle u_i, u_i\rangle = 1$ for $i = 1, \ldots, n-1$, and $\langle u_n, u_n\rangle = 0$. Now the matrix of the bilinear form of A with respect to the basis $\{u_1, \ldots, u_n, e_{n+1}\}$ is

$$B = \begin{pmatrix} 1 & & 0 & & * \\ & \ddots & & & \vdots \\ 0 & & 1 & & * \\ & & & 0 & b \\ * & \cdots & * & b & 1 \end{pmatrix},$$

where $b = \langle u_n, e_{n+1}\rangle$. Write $u_n = (c_1, \ldots, c_n)$ as a vector in $\mathbb{R}^n$. Then by the argument in the proof of Theorem 7.2.2(2), all the components c_i of u_n

have the same sign. Hence

$$b = \sum_{i=1}^{n} c_i \langle e_i, e_{n+1} \rangle \neq 0,$$

since $\langle e_i, e_{n+1} \rangle \leq 0$ for all $i < n+1$ with inequality for some $i < n+1$. By expanding the determinant of B along the $(n+1)$st column, we find that

$$\det B = -b^2 < 0.$$

Hence, the rank of B, and therefore of A, is $n+1$. As the bilinear form of A is positive definite on the $(n-1)$-dimensional vector subspace $\langle u_1, \ldots, u_{n-1} \rangle$, the matrix A must be of type $(n, 1)$.

Define Δ as in the proof of Theorem 7.2.2(3). Then the same argument there proves that Δ is an n-dimensional convex polyhedron in H^n with sides $S_1, \ldots, S_{n+1}$ such that $\theta(S_i, S_j) = \theta_{ij}$ for all i, j. Moreover, if A_j is positive definite, then the n sides $S_1, \ldots, \hat{S}_j, \ldots, S_{n+1}$ intersect at a vertex of Δ.

Suppose that A_j is of type $(n, 0)$. Let W_j be the n-dimensional vector subspace of $\mathbb{R}^{n+1}$ defined as in the proof of Theorem 7.2.2(3). By the same argument there, the Lorentzian inner product on $\mathbb{R}^{n,1}$ is of type $(n, 0)$ on W_j. Therefore W_j is light-like. Hence, the Lorentz orthogonal complement L_j of W_j is light-like. By the same argument as in the proof of Theorem 7.2.2(3), we deduce that L_j represents an ideal vertex of Δ opposite the side S_j. Thus Δ has $n+1$ generalized vertices and at least one ideal vertex. Therefore Δ is a noncompact generalized n-simplex in H^n. □

It follows from Theorem 7.3.1 and the fact that the Coxeter graphs of Euclidean simplex reflection groups are connected that a Coxeter graph is the graph of a hyperbolic, noncompact n-simplex, reflection group if and only if it has the following properties:

(1) The number of vertices is $n+1$.

(2) The graph is connected.

(3) Any subgraph obtained by deleted a vertex and its adjoining edges is the Coxeter graph of either a spherical or Euclidean $(n-1)$-simplex reflection group.

(4) Some subgraph obtained by deleting a vertex and its adjoining edges is the Coxeter graph of a Euclidean $(n-1)$-simplex reflection group.

For each dimension $n \geq 3$, there are only finitely many such graphs, and such graphs exist only for $n \leq 9$. Figure 7.3.6 illustrates the Coxeter graphs of all the hyperbolic, noncompact tetrahedron, reflection groups. The number of Coxeter graphs of hyperbolic, noncompact n-simplex, reflection groups for $n = 4, \ldots, 9$ is $9, 12, 3, 4, 4, 3$, respectively.

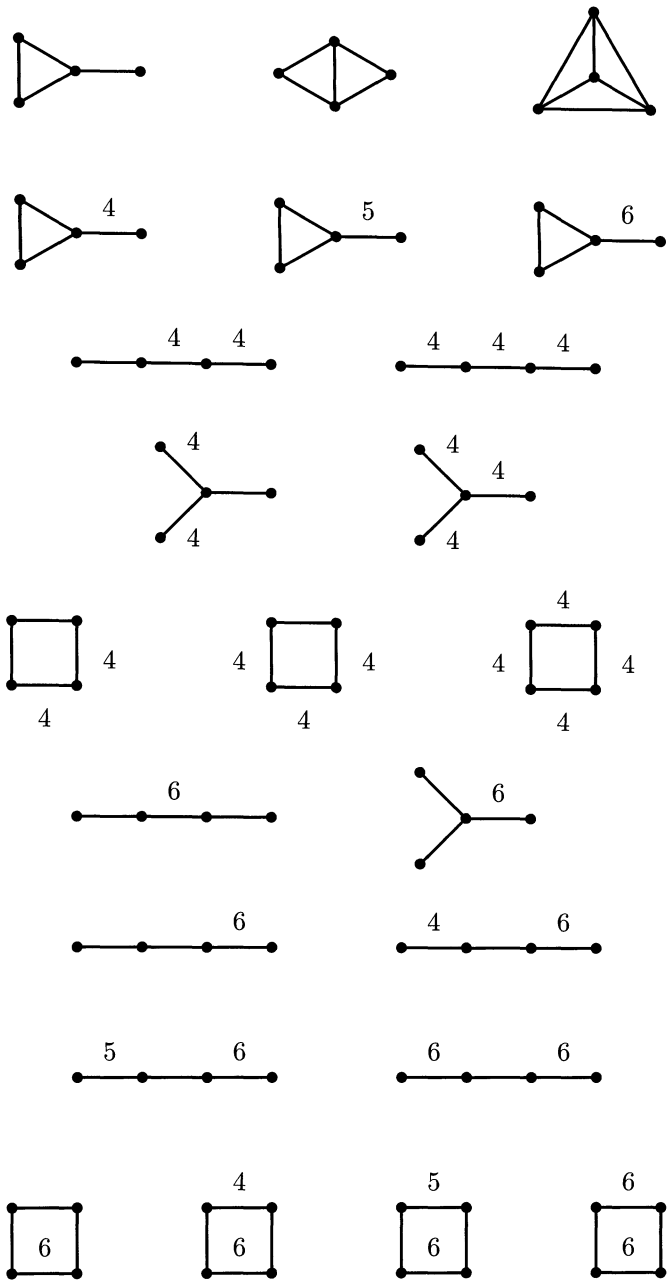


Figure 7.3.6. The hyperbolic, noncompact tetrahedron, reflection groups

Exercise 7.3

1. Prove that $\mathrm{PSL}(2, \mathbb{Z})$ is isomorphic to the subgroup of orientation preserving isometries of a reflection group with respect to a triangle $T(2, 3, \infty)$.
2. Prove that Γ_3 is a hyperbolic, noncompact tetrahedron, reflection group.
3. Construct the Coxeter graphs of all the hyperbolic, noncompact 4-simplex, reflection groups.
4. Prove that each label of the Coxeter graph of a hyperbolic, noncompact n-simplex, reflection group, with $n \geq 4$, is at most 4.
5. Prove that the dimension n of a hyperbolic, noncompact n-simplex, reflection group is at most 9.

§7.4. Crystallographic Groups

In this section, we study the theory of crystallographic groups.

Definition: An n-dimensional *crystallographic group* is a discrete group Γ of isometries of E^n such that E^n/Γ is compact.

Examples of crystallographic groups are the Euclidean, simplex, reflection groups in Figure 7.2.8.

Theorem 7.4.1. *Let Γ be a discrete group of isometries of E^n. Then the following are equivalent:*

(1) *The group Γ is crystallographic.*

(2) *Every convex fundamental polyhedron for Γ is compact.*

(3) *The group Γ has a compact Dirichlet polyhedron.*

Proof: (1) implies (2) by Theorem 6.5.10. Clearly (2) implies (3), and (3) implies (1). □

Let P be a convex fundamental polyhedron for an n-dimensional crystallographic group Γ. Then P is compact by Theorem 7.4.1. Therefore P is bounded and has only finitely many sides. We regard P to be a model for an n-dimensional crystal, and the tessellation $\{gP : g \in \Gamma\}$ of E^n to be a model for a crystalline structure.

The study of crystalline structures is called *crystallography.* By the end of the nineteenth century, crystallographers had classified 1-, 2-, and 3-dimensional crystallographic groups. For each of these dimensions, it was determined that there is only a finite number of essentially different kinds of crystallographic groups. This led Hilbert to ask, in problem 18 on his celebrated list of problems, if there is only a finite number of essentially different kinds of crystallographic groups in each dimension. This problem was answered affirmatively by L. Bieberbach in 1910. Bieberbach proved that there are only finitely many isomorphism classes of n-dimensional crystallographic groups for each n. In this section, we shall prove Bieberbach's theorem.

Lemma 1. *If* H *is a subgroup of finite index of a discrete group* Γ *of isometries of* $X = E^n$ *or* H^n, *then* X/Γ *is compact if and only if* X/H *is compact.*

Proof: Suppose that X/H is compact. Define a function

$$\phi : X/\mathrm{H} \to X/\Gamma$$

by $\phi(\mathrm{H}x) = \Gamma x$. Let $\pi : X \to X/\Gamma$ and $\eta : X \to X/\mathrm{H}$ be the quotient maps. Then $\pi = \phi\eta$. Therefore ϕ is continuous. As ϕ is surjective, X/Γ is compact.

Conversely, suppose that X/Γ is compact. Let D be a Dirichlet domain for Γ. Then D is a locally finite fundamental domain for Γ. Therefore $\overline{D}$ is compact by Theorem 6.5.10. Let $g_1\mathrm{H}, \ldots, g_m\mathrm{H}$ be the cosets of H in Γ and define

$$K = g_1^{-1}\overline{D} \cup \cdots \cup g_m^{-1}\overline{D}.$$

Then K is a compact subset of X. Let x be a point of X. Then there is a g in Γ such that gx is in $\overline{D}$; moreover, there is an index i such that $g = g_i h$ for some h in H. Hence hx is in $g_i^{-1}\overline{D}$. Thus $\mathrm{H}x$ is in $\eta(K)$. This shows that $X/\mathrm{H} = \eta(K)$ and therefore X/H is compact. □

Theorem 7.4.2. *Let* Γ *be a discrete group of isometries of* E^n. *Then* Γ *is crystallographic if and only if the subgroup* T *of translations of* Γ *is of finite index and has rank* n.

Proof: Suppose that Γ is crystallographic. By Theorem 5.4.3, the group Γ has an abelian subgroup H of finite index containing T; moreover, H is also crystallographic by Lemma 1. By Theorem 5.4.4, there is an m-plane P of E^n on which H acts by translation. Since points at a distance d from P stay at a distance d from P under the action of H, the orbit space E^n/H is unbounded if $m < n$. As E^n/H is compact, we must have $m = n$. Therefore H is a lattice subgroup of $\mathrm{I}(E^n)$. Hence $\mathrm{H} = \mathrm{T}$, and T is of finite index in Γ and has rank n.

Conversely, suppose that the subgroup T of translations of Γ is of finite index and has rank n. By Theorem 5.3.2, there is a basis $v_1, \ldots, v_n$ of $\mathbb{R}^n$ such that T is the group generated by the translations of E^n by $v_1, \ldots, v_n$. Clearly, the parallelepiped P spanned by $v_1, \ldots, v_n$ is a convex fundamental polyhedron for T. As P is compact, E^n/T is also compact. Therefore E^n/Γ is compact by Lemma 1. □

Let Γ be an n-dimensional crystallographic group and let $\mathrm{T} = \mathrm{T}(\Gamma)$ be its group of translations. Then T is a free abelian group of rank n and has finite index in Γ by Theorem 7.4.2. Furthermore, by Theorem 5.4.4, the subgroup T of Γ is characterized as the unique maximal free abelian subgroup of Γ. Consequently, the rank n of T is an isomorphism invariant of Γ. Therefore, the dimension n of Γ is an isomorphism invariant of Γ.

Let $\eta : \Gamma \to \mathrm{O}(n)$ be the natural projection defined by $\eta(a + A) = A$. The image Π of η is called the *point group* of Γ. As T is the kernel of η, we have an exact sequence of groups

$$1 \to \mathrm{T} \to \Gamma \to \Pi \to 1. \tag{7.4.1}$$

Therefore T is a normal subgroup of Γ and Π is a finite group. Furthermore, conjugation in Γ induces a left action of Π on T that makes T into a Π-module. Let $L = L(\Gamma)$ be the lattice subgroup of $\mathbb{R}^n$ corresponding to T. If $a + A$ is in Γ and b is in L, then

$$(a + A)(b + I)(a + A)^{-1} = Ab + I. \tag{7.4.2}$$

Hence Π acts on L by left matrix multiplication. By Theorem 5.4.4, the group T is a maximal abelian subgroup of Γ. Hence Π acts effectively on T and therefore on L. Consequently, we have a faithful representation of Π into $\mathrm{Aut}(L)$ given by $A \mapsto \phi_A$ where $\phi_A(x) = Ax$. As L is isomorphic to $\mathbb{Z}^n$, we have an exact sequence of groups

$$0 \to \mathbb{Z}^n \to \Gamma \to Q \to 1, \tag{7.4.3}$$

where Q is a finite subgroup of $\mathrm{GL}(n, \mathbb{Z})$ and the left action of Q on $\mathbb{Z}^n$ induced by conjugation in Γ is the natural action of Q on $\mathbb{Z}^n$. The standard method of proving that there are only finitely many isomorphism classes of n-dimensional crystallographic groups is to prove that there are only finitely many isomorphism classes of group extensions of the form (7.4.3). We shall take a different, more geometric, approach which exploits the geometry of lattices in $\mathbb{R}^n$.

Lemma 2. *Let $B(a, r)$ be the open ball in E^n with center a and radius r. Then there is a positive constant c_n, depending only on n, such that*

$$\mathrm{Vol}(B(a, r)) = c_n r^n.$$

Proof: Without loss we may assume that $a = 0$. Integrating with respect to spherical coordinates, we have

$$\begin{aligned} \mathrm{Vol}(B(0, r)) &= \int_0^{2\pi}\int_0^{\pi}\cdots\int_0^r \rho^{n-1}\sin^{n-2}\theta_1\cdots\sin\theta_{n-2}\,d\rho\, d\theta_1\cdots d\theta_{n-1} \\ &= \frac{r}{n}^n \mathrm{Vol}(S^{n-1}). \end{aligned}$$

Hence, the desired constant is

$$c_n = \frac{1}{n}\mathrm{Vol}(S^{n-1}). \qquad \square$$

Definition: A lattice L in $\mathbb{R}^n$ is *full scale* if and only if all the nonzero vectors of L have norm at least 1.

Lemma 3. *Let L be a full scale lattice in $\mathbb{R}^n$ and for each $r \geq 0$, let $N(r)$ be the number of vectors in L whose norm is at most r. Then*

$$N(r) \leq (2r+1)^n.$$

Proof: Since L is full scale, the distance between any two distinct vectors in L is at least 1. Consequently, the open balls of radius $\frac{1}{2}$ centered at the $N(r)$ vectors of L, whose norm is at most r, are pairwise disjoint and are all contained in the ball of radius $r + \frac{1}{2}$ centered at the origin. Comparing the volumes, we deduce from Lemma 2 that

$$N(r)\left(\tfrac{1}{2}\right)^n \leq \left(r + \tfrac{1}{2}\right)^n. \qquad \square$$

Lemma 4. *Let $\{v_1, \ldots, v_n\}$ be a basis for $\mathbb{R}^n$. Then for each x in $\mathbb{R}^n$, there are integers $k_1, \ldots, k_n$ such that*

$$\left| x - \sum_{i=1}^n k_i v_i \right| \leq \tfrac{1}{2}\big(|v_1| + \cdots + |v_n|\big).$$

Proof: Let x be in $\mathbb{R}^n$. Then there are real numbers $t_1, \ldots, t_n$ such that $x = \sum_{i=1}^n t_i v_i$. Let k_i be an integer nearest to t_i in $\mathbb{R}$. Then we have

$$\begin{aligned} \left| x - \sum_{i=1}^n k_i v_i \right| &= \left| \sum_{i=1}^n (t_i - k_i) v_i \right| \\ &\leq \sum_{i=1}^n |(t_i - k_i) v_i| \\ &\leq \tfrac{1}{2}\big(|v_1| + \cdots + |v_n|\big). \end{aligned} \qquad \square$$

Lemma 5. *Let V be a vector subspace of $\mathbb{R}^n$ spanned by m linearly independent unit vectors $v_1, \ldots, v_m$ in a full scale lattice L in $\mathbb{R}^n$. If a vector u in L is not in V, then its $V^\perp$-component w has norm*

$$|w| > (m+3)^{-n}.$$

Proof: On the contrary, let u be a vector in L whose $V^\perp$-component w satisfies

$$0 < |w| \leq (m+3)^{-n}.$$

Now let

$$k = (m+3)^n.$$

Then $k|w| \leq 1$. Hence, the vectors $0, u, 2u, \ldots, ku$ are at a distance at most 1 from V. By Lemma 4, we may add suitable integral linear combinations of $v_1, \ldots, v_m$ to each of these vectors to obtain $k+1$ new distinct vectors in

L whose $V^\perp$-components have not changed but whose V-components have norm at most $m/2$. These $k+1$ vectors of L have norm less than

$$r = (m/2) + 1.$$

By Lemma 3, we have

$$k + 1 \leq N(r) \leq (2r+1)^n = (m+3)^n,$$

which is a contradiction. Therefore

$$|w| > (m+3)^{-n}. \qquad \square$$

Definition: An n-dimensional crystallographic group Γ is *normalized* if and only if its lattice $L(\Gamma)$ is full scale and contains n linearly independent unit vectors.

Lemma 6. *Let Γ be an n-dimensional crystallographic group. Then Γ is isomorphic to a normalized n-dimensional crystallographic group.*

Proof: By changing scale, we may assume that a shortest nonzero vector in $L(\Gamma)$ is a unit vector. Now assume by induction that $L(\Gamma)$ is full scale and contains $m < n$ linearly independent unit vectors $v_1, \ldots, v_m$. We shall find an n-dimensional crystallographic group Γ' isomorphic to Γ such that $L(\Gamma')$ is full scale and contains $m+1$ linearly independent unit vectors.

Let V be the vector subspace of $\mathbb{R}^n$ spanned by $v_1, \ldots, v_m$. Assume first that the action of the point group Π of Γ on $L(\Gamma)$ does not leave V invariant. Then there is an element A of Π and an index i such that Av_i is not in V. Let $v_{m+1} = Av_i$. Then $v_1, \ldots, v_{m+1}$ are $m+1$ linearly independent unit vectors in $L(\Gamma)$. Therefore Γ is the desired group.

Now assume that Π leaves V invariant. Then Π also leaves $V^\perp$ invariant. For each $t > 0$, define a linear automorphism α_t of $\mathbb{R}^n$ by the formula

$$\alpha_t(u) = v + tw,$$

where $u = v + w$ with v in V and w in $V^\perp$. Let $a + A$ be in Γ. As A leaves V and $V^\perp$ invariant, we have

$$\alpha_t(a+A)\alpha_t^{-1} = \alpha_t(a) + A.$$

Hence, for each $t > 0$, the group

$$\Gamma_t = \alpha_t \Gamma \alpha_t^{-1}$$

is a subgroup of $\mathrm{I}(E^n)$. As

$$\mathrm{T}(\Gamma_t) = \alpha_t \mathrm{T}(\Gamma) \alpha_t^{-1}$$

and $\mathrm{T}(\Gamma_t)$ is of finite index in Γ_t for each $t > 0$, we have that Γ_t is an n-dimensional crystallographic group for each $t > 0$. Moreover, we have

$$L(\Gamma_t) = \alpha_t(L(\Gamma)).$$

Let u be an arbitrary vector in $L(\Gamma) - V$ and write $u = v + w$ with v in V and w in $V^\perp$. Then for t such that

$$0 < t \leq |w|^{-1}(m+3)^{-n},$$

the vector $v + tw$ is in $L(\Gamma_t) - V$ and

$$|tw| \leq (m+3)^{-n}.$$

By Lemma 5, the lattice $L(\Gamma_t)$ cannot be full scale. Let

$$s = \inf\{t : L(\Gamma_t) \text{ is full scale}\}.$$

Then $0 < s \leq 1$. As $|\alpha_t(u)| \geq 1$ for all $t > s$, we have that $|\alpha_s(u)| \geq 1$, since $|\alpha_t(u)|$ is a continuous function of t. Therefore $L(\Gamma_s)$ is full scale.

Let u_0 be a shortest vector in $L(\Gamma_s) - V$. We claim that u_0 is a unit vector. On the contrary, suppose that $|u_0| > 1$. By replacing Γ by Γ_s, we may assume that $s = 1$. Write $u_0 = v_0 + w_0$ with v_0 in V and w_0 in $V^\perp$. As $|u|^2 \geq |u_0|^2$, we have

$$|v|^2 + |w|^2 \geq |v_0|^2 + |w_0|^2.$$

Let $t = |u_0|^{-1}$. Then

$$\begin{aligned}
|\alpha_t(u)|^2 &= |v + tw|^2 \\
&= |v|^2 + t^2|w|^2 \\
&\geq |v|^2 + t^2(|v_0|^2 + |w_0|^2 - |v|^2) \\
&= |v|^2(1 - t^2) + t^2|u_0|^2 \\
&\geq t^2|u_0|^2 \\
&= 1.
\end{aligned}$$

Therefore $L(\Gamma_t)$ is full scale contrary to the minimality of s. Thus, we have that $v_{m+1} = u_0$ is a unit vector. Hence $v_1, \ldots, v_{m+1}$ are $m+1$ linearly independent unit vectors in $L(\Gamma_s)$. Therefore Γ_s is the desired group. This completes the induction. Thus Γ is isomorphic to a normalized n-dimensional crystallographic group. □

Theorem 7.4.3. (Bieberbach's theorem) *For each dimension n, there are only finitely many isomorphism classes of n-dimensional crystallographic groups.*

Proof: Fix a positive integer n. By Lemma 6, it suffices to show that there are only finitely many isomorphism classes of normalized n-dimensional crystallographic groups. Let Γ be such a group. Then $L(\Gamma)$ contains n linearly independent unit vectors $w_1, \ldots, w_n$. For each i, let $\omega_i = w_i + I$ be the corresponding translation in Γ, and let H be the subgroup of $\mathrm{T}(\Gamma)$ generated by $\omega_1, \ldots, \omega_n$. Then H is a free abelian group of rank n and therefore has finite index in $\mathrm{T}(\Gamma)$. By Theorem 7.4.2, the group $\mathrm{T}(\Gamma)$ has finite index in Γ. Hence H is of finite index in Γ.

By Lemma 4, we may choose for each coset $\mathrm{H}\omega$ of H in Γ a representative $\omega = w + A$ whose translation vector w has norm $|w| \leq n/2$. Let $\omega_{n+1}, \ldots, \omega_m$ be the chosen coset representatives. Then every element ϕ of Γ can be expressed uniquely in the form

$$\phi = (a_1 w_1 + \cdots + a_n w_n + I)\omega_p,$$

where $a_1, \ldots, a_n$ and p are integers with $n + 1 \leq p \leq m$. We shall call this expression the *normal form* for ϕ.

Since every element of Γ has a unique normal form, there are for each $i, j = 1, \ldots, m$, unique integers c_{ijk} and $f(i, j) > n$ such that

$$\omega_i \omega_j = (c_{ij1} w_1 + \cdots + c_{ijn} w_n + I)\omega_{f(i,j)}.$$

The integers c_{ijk} and $f(i, j)$ completely determine Γ, since one can find the normal form of a product of elements ϕ, ψ of Γ given the normal forms for ϕ, ψ and $\omega_i\omega_j$ for each $i, j = 1, \ldots, m$. To see this, let

$$\begin{aligned} \phi &= (a_1 w_1 + \cdots + a_n w_n + I)\omega_p, \\ \psi &= (b_1 w_1 + \cdots + b_n w_n + I)\omega_q \end{aligned}$$

be the normal forms for ϕ and ψ. Then

$$\phi\psi = (a_1 w_1 + \cdots + a_n w_n + I)\omega_p(\omega_1^{b_1} \cdots \omega_n^{b_n})\omega_q.$$

To find the normal form for $\phi\psi$, it suffices to find the normal form of $\omega_p(\omega_1^{b_1} \cdots \omega_n^{b_n})\omega_q$. If $b_1 > 0$, we replace $\omega_p\omega_1$ by its normal form. This has the effect of lowering b_1 to $b_1 - 1$. If $b_1 < 0$, we replace $\omega_p\omega_1^{-1}$ by its normal form

$$\omega_p\omega_1^{-1} = (d_1 w_1 + \cdots + d_n w_n + I)\omega_i.$$

Observe that

$$\omega_i\omega_1 = (-d_1 w_1 - \cdots - d_n w_n + I)\omega_p.$$

Hence i is the unique integer such that $p = f(i, 1)$; moreover $d_k = -c_{i1k}$ for each $k = 1, \ldots, n$. Thus, we can raise b_1 to $b_1 + 1$. It is clear that by repeated application of these two steps we can find the normal form of $\phi\psi$.

Even more is true. The integers c_{ijk} and $f(i, j)$ determine Γ up to isomorphism, in the sense that if Γ' is another normalized n-dimensional crystallographic group with the same set of integers, then Γ and Γ' are isomorphic. To see this, let $w_1', \ldots, w_n'$ be the corresponding unit vectors of $L(\Gamma')$ and let $\omega_{n+1}', \ldots, \omega_m'$ be the corresponding coset representatives. Then the function $\xi : \Gamma \to \Gamma'$, defined by

$$\xi((a_1 w_1 + \cdots + a_n w_n + I)\omega_p) = (a_1 w_1' + \cdots + a_n w_n' + I)\omega_p',$$

is an isomorphism, since ξ is obviously a bijection, and the same algorithm determines the normal form for a product in each group. Thus, to show that there are only finitely many isomorphism classes of normalized n-dimensional crystallographic groups, it suffices to show that the absolute values of the integers c_{ijk} and m have an upper bound depending only on the dimension n.

Now the elements ω_i, ω_j and $\omega_{f(i,j)}$ have translation vectors of length at most $n/2$. Consequently, the translation vector of

$$c_{ij1}w_1 + \cdots + c_{ijn}w_n + I = \omega_i\omega_j\omega_{f(i,j)}^{-1}$$

has length at most $3n/2$. Let v_k be the component of w_k perpendicular to the hyperplane spanned by $w_1, \ldots, w_{k-1}, w_{k+1}, \ldots, w_n$. Then

$$|c_{ijk}v_k| \leq 3n/2.$$

By Lemma 5, we have that

$$|v_k| > (n+2)^{-n}.$$

Hence, for each i, j, k, we have

$$|c_{ijk}| \leq \frac{3n}{2}(n+2)^n.$$

We next find an upper bound for m. First of all, we have

$$m - n = [\Gamma : \mathrm{H}] = [\Gamma : \mathrm{T}(\Gamma)][\mathrm{T}(\Gamma) : \mathrm{H}].$$

Now the translations among the representatives $\omega_{n+1}, \ldots, \omega_m$ form a complete set of coset representatives for H in $\mathrm{T}(\Gamma)$. Each translation vector w_i has norm at most $n/2$ and, by Lemma 3, is one of at most $(n+1)^n$ vectors in $L(\Gamma)$. Hence

$$[\mathrm{T}(\Gamma) : \mathrm{H}] \leq (n+1)^n.$$

Next, observe that

$$[\Gamma : \mathrm{T}(\Gamma)] = |\Pi|,$$

where Π is the point group of Γ. Let A be in Π. Then A is uniquely determined by its images Aw_i for $i = 1, \ldots, n$. By Lemma 3, the vector Aw_i is one of at most 3^n different unit vectors in $L(\Gamma)$. Hence A is one of at most $(3^n)^n$ different matrices in $\mathrm{O}(n)$. Hence

$$[\Gamma : \mathrm{T}(\Gamma)] \leq (3^n)^n.$$

Thus, we have

$$m \leq n + (3^n)^n(n+1)^n.$$

□

Remark: The exact number of isomorphism classes of n-dimensional crystallographic groups for $n = 1, 2, 3, 4$ is 2, 17, 219, 4783, respectively.

The Splitting Group

Let Γ be an n-dimensional crystallographic group and let m be the order of the point group Π of Γ. Let Γ^* be the subgroup of $\mathrm{I}(E^n)$ generated by $\mathrm{T}(\Gamma)^{\frac{1}{m}}$ and Γ. Then Γ^* has the same point group Π. Therefore

$$\begin{aligned}
[\Gamma^* : \Gamma] &= [\mathrm{T}(\Gamma)^{\frac{1}{m}} : \mathrm{T}(\Gamma)] \\
&= [\tfrac{1}{m}L(\Gamma) : L(\Gamma)] \\
&= [(\tfrac{1}{m}\mathbb{Z})^n : \mathbb{Z}^n] \\
&= m^n.
\end{aligned}$$

Hence Γ^* is also an n-dimensional crystallographic group with

$$L(\Gamma^*) = \tfrac{1}{m}L(\Gamma). \tag{7.4.4}$$

The group Γ^* is called the *splitting group* of Γ.

Lemma 7. *If Γ^* is the splitting group of Γ, then the following exact sequence splits*

$$1 \to \mathrm{T}(\Gamma^*) \to \Gamma^* \to \Pi \to 1.$$

Proof: Let $\eta : \Gamma^* \to \Pi$ be the natural projection. For each A in Π, choose ϕ_A in Γ such that $\eta(\phi_A) = A$. Then for each A, B in Π, there is an element $\tau(A, B)$ of $\mathrm{T}(\Gamma)$ such that

$$\phi_A\phi_B = \tau(A, B)\phi_{AB}.$$

Let $\phi_A = a_A + A$ for each A. Then

$$\phi_A\phi_B = a_A + Aa_B + AB.$$

Hence, we have

$$\tau(A, B) = a_A + Aa_B - a_{AB} + I.$$

Define a function $f : \Pi \times \Pi \to L(\Gamma)$ by the formula

$$f(A, B) = a_A + Aa_B - a_{AB}.$$

Taking the sum of both sides of the last equation, as B ranges over all the elements of Π, gives

$$\sum_{B\in\Pi} f(A, B) = ma_A + A\sum_{B\in\Pi} a_B - \sum_{B\in\Pi} a_B.$$

Define $\sigma : \Gamma \to \Gamma^*$ by

$$\sigma(A) = -\frac{1}{m}\sum_{C\in\Pi} f(A, C) + a_A + A.$$

Let $s = \sum_{C\in\Pi} a_C$. Then

$$\sigma(A) = -\frac{1}{m}(A - I)s + A.$$

Observe that

$$\begin{aligned}
\sigma(AB) &= -\frac{1}{m}(AB - I)s + AB \\
&= -\frac{1}{m}(A - I)s - \frac{1}{m}(AB - A)s + AB \\
&= \sigma(A)\sigma(B).
\end{aligned}$$

Therefore σ is a homomorphism such that $\eta\sigma$ is the identity on Π. □

Theorem 7.4.4. *Let $\xi : \Gamma_1 \to \Gamma_2$ be an isomorphism of n-dimensional crystallographic groups. Then there is an affine bijection α of $\mathbb{R}^n$ such that for each ϕ in Γ_1, we have*

$$\xi(\phi) = \alpha\phi\alpha^{-1}.$$

Proof: Since the subgroup of translations of a crystallographic group is characterized as the unique maximal free abelian subgroup, we have

$$\xi(\mathrm{T}(\Gamma_1)) = \mathrm{T}(\Gamma_2).$$

Hence ξ induces an isomorphism $\overline{\xi} : \Pi_1 \to \Pi_2$ between the point groups of Γ_1 and Γ_2. For each A in Π_1, choose ϕ_A in Γ_1 such that $\eta_1(\phi_A) = A$ where $\eta_1 : \Gamma_1 \to \Pi_1$ is the natural projection. Then $\{\phi_A : A \in \Pi_1\}$ is a set of coset representatives for $\mathrm{T}(\Gamma_1^*)$ in Γ_1^*. Let τ be an arbitrary element of $\mathrm{T}(\Gamma_1^*)$ and let m be the order of Π_1 and Π_2. Define $\xi^* : \Gamma_1^* \to \Gamma_2^*$ by

$$\xi^*(\tau\phi_A) = [\xi(\tau^m)]^{\frac{1}{m}}\xi(\phi_A).$$

Then ξ^* is an isomorphism, since ξ^* maps $\mathrm{T}(\Gamma_1^*)$ isomorphically onto $\mathrm{T}(\Gamma_2^*)$, and ξ^* agrees with the isomorphism $\overline{\xi}$. Moreover ξ^* extends ξ.

By Lemma 7, the exact sequence

$$1 \to \mathrm{T}(\Gamma_i^*) \to \Gamma_i^* \to \Pi_i \to 1$$

splits for each $i = 1, 2$. Let $\sigma_i : \Pi_i \to \Gamma_i^*$ be a splitting homomorphism. The finite group $\sigma_i(\Pi_i)$ has a fixed point in E^n. By a change of origin, we may assume that $\sigma_i(\Pi_i)$ fixes the origin. Then $\sigma_i(\Pi_i) = \Pi_i$ for $i = 1, 2$. Hence, every element of Γ_i^* is of the form τA with τ in $\mathrm{T}(\Gamma_i^*)$ and A in Π_i. Let $v_1, \ldots, v_n$ generate $L(\Gamma_1)$ and define $w_1, \ldots, w_n$ by

$$w_j + I = \xi(v_j + I) \quad \text{for} \quad j = 1, \ldots, n.$$

Then $w_1, \ldots, w_n$ generate $L(\Gamma_2)$. Hence, there is a unique linear automorphism α of $\mathbb{R}^n$ such that $\alpha(v_j) = w_j$ for $j = 1, \ldots, n$.

Let A be in Π_1 and let a be in $L(\Gamma_1^*)$. Then

$$A(a + I)A^{-1} = Aa + I.$$

Hence, we have

$$\xi^*(A(a + I)A^{-1}) = \xi^*(Aa + I).$$

Therefore

$$\xi^*(A)(\alpha(a) + I)\xi^*(A)^{-1} = \alpha Aa + I$$

and so we have

$$\xi^*(A)\alpha(a) + I = \alpha Aa + I.$$

Hence, we have

$$\xi^*(A)\alpha = \alpha A.$$

Thus, we have

$$\xi^*(A) = \alpha A\alpha^{-I}.$$

Hence, we have

$$\begin{aligned} \xi^*(\tau A) &= \xi^*(\tau)\xi^*(A) \\ &= (\alpha\tau\alpha^{-1})(\alpha A\alpha^{-1}) \\ &= \alpha(\tau A)\alpha^{-1}. \end{aligned}$$

□

Bieberbach Groups

Definition: An n-dimensional *Bieberbach group* is a group G for which there is an exact sequence of groups

$$0 \longrightarrow \mathbb{Z}^n \xrightarrow{\iota} G \xrightarrow{\eta} Q \longrightarrow 1 \tag{7.4.5}$$

such that Q is a finite subgroup of $\mathrm{GL}(n,\mathbb{Z})$ and the left action of Q on $\mathbb{Z}^n$ induced by conjugation in G is the natural action of Q on $\mathbb{Z}^n$.

For example, any n-dimensional crystallographic group is an n-dimensional Bieberbach group. We shall algebraically characterize crystallographic groups by showing that every n-dimensional Bieberbach group is isomorphic to an n-dimensional crystallographic group.

Lemma 8. *Let G be an n-dimensional Bieberbach group and let Q be a finite subgroup of* $\mathrm{GL}(n,\mathbb{Z})$ *as in the exact sequence* 7.4.5. *Then G can be embedded as a subgroup of finite index in the semidirect product $\mathbb{Z}^n \rtimes Q$.*

Proof: For each q in Q, choose an element x_q of G such that $\eta(x_q) = q$ and $x_1 = 1$. Then for each q, r in Q, there is a unique element $f(q,r)$ of $\mathbb{Z}^n$ such that

$$x_q x_r = \iota f(q,r) x_{qr}.$$

The function $f : Q \times Q \to \mathbb{Z}^n$ completely determines G, since if a, b are in $\mathbb{Z}^n$, then

$$(\iota(a)x_q)(\iota(b)x_r) = \iota(a + qb + f(q,r))x_{qr}.$$

The associativity of the group operation in G gives rise to the following cocycle identity for f. For each q, r, s in Q, we have

$$f(q,r) + f(qr,s) = qf(r,s) + f(q,rs).$$

We next construct a new n-dimensional Bieberbach group G^* from G and f. Let $G^* = \mathbb{Z}^n \times Q$ as a set and let $m = |Q|$. Define a multiplication in G^* by the formula

$$(a,q)(b,r) = (a + qb + mf(q,r), qr).$$

It is straightforward to check that G^* is a group with this multiplication. Let $\kappa : \mathbb{Z}^n \to G^*$ and $\pi : G^* \to Q$ be the natural injection and projection. Then we have an exact sequence

$$0 \longrightarrow \mathbb{Z}^n \xrightarrow{\kappa} G^* \xrightarrow{\pi} Q \longrightarrow 1.$$

Moreover, we have

$$(0,q)(a,1)(0,q)^{-1} = (qa,1).$$

Therefore G^* is an n-dimensional Bieberbach group.

Next, we show that π has a right inverse. Define $\sigma : Q \to G^*$ by

$$\sigma(q) = \Big(-\sum_{s\in Q} f(q,s), q\Big).$$

Taking the sum of both sides of the cocycle identity for f gives

$$mf(q,r) + \sum_{s\in Q} f(qr,s) = q\sum_{s\in Q} f(r,s) + \sum_{s\in Q} f(q,s).$$

Hence

$$\begin{aligned}\sigma(qr) &= \Big(-\sum_{s\in Q} f(qr,s), qr\Big)\\ &= \Big(-\sum_{s\in Q} f(q,s) - q\sum_{s\in Q} f(r,s) + mf(q,r), qr\Big)\\ &= \sigma(q)\sigma(r).\end{aligned}$$

Thus σ is a homomorphism such that $\pi\sigma$ is the identity on Q.

Next, define a function

$$\xi : \mathbb{Z}^n \rtimes Q \to G^*$$

by the formula

$$\xi(a,q) = \kappa(a)\sigma(q).$$

Then ξ is an isomorphism. Hence, it suffices to show that G can be embedded in G^* as a subgroup of finite index.

Define $\varepsilon : G \to G^*$ by

$$\varepsilon(\iota(a)x_q) = (ma, q).$$

Then we have

$$\begin{aligned}\varepsilon(\iota(a)x_q\iota(b)x_r) &= \varepsilon(\iota(a+qb+f(q,r))x_{qr})\\ &= (m(a+qb+f(q,r)), qr)\\ &= (ma+q(mb)+mf(q,r)), qr)\\ &= (ma,q)(mb,r)\\ &= \varepsilon(\iota(a)x_q)\varepsilon(\iota(b)x_r).\end{aligned}$$

Thus ε is a homomorphism. Clearly ε is a monomorphism and

$$[G^* : \varepsilon(G)] = [\mathbb{Z}^n : (m\mathbb{Z})^n] = m^n. \qquad \square$$

Lemma 9. *Let Q be a finite subgroup of* $\mathrm{GL}(n,\mathbb{R})$ *(resp.* $\mathrm{GL}(n,\mathbb{C})$*). Then Q is conjugate in* $\mathrm{GL}(n,\mathbb{R})$ *(resp.* $\mathrm{GL}(n,\mathbb{C})$*) to a finite subgroup of* $\mathrm{O}(n)$ *(resp.* $\mathrm{U}(n)$*).*

Proof: Define an inner product on $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) by the formula

$$\langle x, y\rangle = \sum_{q\in Q} qx * qy.$$

This product is obviously bilinear, Hermitian symmetric, and nondegenerate; moreover, for each q in Q, we have

$$\langle qx, qy\rangle = \langle x, y\rangle.$$

By the Gram-Schmidt process, we construct an orthonormal basis $v_1, \ldots, v_n$ for $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) with respect to this inner product. Define A in $\mathrm{GL}(n,\mathbb{R})$ (resp. $\mathrm{GL}(n,\mathbb{C})$) by $Ae_i = v_i$ for $i = 1, \ldots, n$. Then

$$\begin{aligned}\langle Ax, Ay\rangle &= \left\langle A\sum_{j=1}^{n} x_i e_i, A\sum_{j=1}^{n} y_j e_j\right\rangle \\ &= \left\langle \sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j\right\rangle \\ &= \sum_{i=1}^{n} x_i \overline{y}_i \\ &= x * y.\end{aligned}$$

If q is in Q and x, y are in $\mathbb{R}^n$ (resp. $\mathbb{C}^n$), then

$$\begin{aligned}A^{-1}qAx * A^{-1}qAy &= \langle qAx, qAy\rangle \\ &= \langle Ax, Ay\rangle \\ &= x * y.\end{aligned}$$

Thus $A^{-1}qA$ is an orthogonal (resp. unitary) transformation. Hence $A^{-1}QA$ is a finite subgroup of $\mathrm{O}(n)$ (resp. $\mathrm{U}(n)$). □

Theorem 7.4.5. *Let G be an n-dimensional Bieberbach group. Then G is isomorphic to an n-dimensional crystallographic group.*

Proof: As every subgroup of finite index of an n-dimensional crystallographic group is again an n-dimensional crystallographic group, we may assume, by Lemma 8, that G is a semidirect product $\mathbb{Z}^n \rtimes Q$, where Q is a finite subgroup of $\mathrm{GL}(n,\mathbb{Z})$. By Lemma 9, there is a matrix A in $\mathrm{GL}(n,\mathbb{R})$ such that AQA^{-1} is a subgroup of $\mathrm{O}(n)$. The group $L = A(\mathbb{Z}^n)$ is a lattice in $\mathbb{R}^n$ and $\Pi = AQA^{-1}$ acts naturally on L. The function

$$\alpha : \mathbb{Z}^n \rtimes Q \longrightarrow L \rtimes \Pi$$

defined by the formula

$$\alpha(a, q) = (Aa, AqA^{-1})$$

is obviously an isomorphism. Now define a function

$$\beta : L \rtimes \Pi \longrightarrow \mathrm{I}(E^n)$$

by the formula

$$\beta(a, A) = a + A.$$

Then β is clearly a monomorphism. Let $T = \beta(L)$. Then T is generated by n linearly independent translations. Therefore T is a discrete subgroup of $\mathrm{I}(E^n)$. As T is of finite index in $\Gamma = \mathrm{Im}\,\beta$, we have that Γ is an n-dimensional crystallographic group. Thus G is isomorphic to an n-dimensional crystallographic group. □

Exercise 7.4

1. Prove that a discrete group Γ of isometries of E^n is crystallographic if and only if the translation direction vectors of its parabolic elements span $\mathbb{R}^n$. See Exercise 5.4.6.
2. Let Γ be a crystallographic group. Prove that an element $a + A$ of Γ is a translation if and only if $|A - I| < 1/2$.
3. Verify that G^* in the proof of Lemma 8 is a group.
4. Derive the cocycle identity for f in the proof of Lemma 8.
5. Prove that the group G^* in the proof of Lemma 8 is isomorphic to the splitting group of G when G is crystallographic.

§7.5. Torsion-Free Linear Groups

In this section, we prove Selberg's lemma. In order to prove this lemma, we need to review some commutative ring theory.

Integral Domains

In this section, all rings are commutative with identity.

Definition: A ring A is an *integral domain* if and only if $0 \neq 1$ in A and whenever $ab = 0$ in A, then either $a = 0$ or $b = 0$.

Clearly, any subring of a field in an integral domain. Let S be a subset of an integral domain A. Then S is said to be *multiplicatively closed* if and only if 1 is in S and S is closed under multiplication. Suppose that S is multiplicatively closed. Define an equivalence relation on $A \times S$ by

$$(a, s) \cong (b, t) \text{ if and only if } at = bs.$$

Let a/s be the equivalence class of (a, s) and let $S^{-1}A$ be the set of equivalence classes. Then $S^{-1}A$ is a ring with fractional addition and multiplication. The ring $S^{-1}A$ is called the *ring of fractions* of A with respect to the multiplicatively closed set S.

Observe that the mapping a $\mapsto a/1$ is a ring monomorphism of A into $S^{-1}A$. Hence, we may regard A as a subring of $S^{-1}A$. Note that $S^{-1}A$ is also an integral domain. If $S = A - \{0\}$, then $S^{-1}A$ is a field, called the *field of fractions* of A. Thus, any integral domain is a subring of a field.

Definition: An ideal P of a ring A is *prime* if and only if A/P is an integral domain.

An ideal M of a ring A is said to be *maximal* if and only if M is proper ($M \neq A$) and A contains no ideals between M and A. Any maximal ideal M of a ring A is prime, because A/M is a field. By Zorn's Lemma, any proper ideal I of a ring A is contained in a maximal ideal of A.

Let P be a prime ideal of an integral domain A. Then $S = A - P$ is a multiplicatively closed subset of A. The ring $A_P = S^{-1}A$ is called the *localization* of A at P.

Definition: A ring A is *local* if and only if A has a unique maximal ideal.

Lemma 1. *If M is a proper ideal of a ring A such that every element of $A - M$ is a unit of A, then A is a local ring with M its maximal ideal.*

Proof: Let I be a proper ideal of A. Then every element of I is a nonunit. Hence $I \subset M$, and so M is the only maximal ideal of A. □

Theorem 7.5.1. *If P is a prime ideal of an integral domain A, then A_P is a local ring.*

Proof: Let $S = A - P$. Then $M = \{a/s : a \in P \text{ and } s \in S\}$ is a proper ideal of A_P. If b/t is in $A_P - M$, then b is in S, and so b/t is a unit of A_P. Therefore A_P is a local ring with M its maximal ideal by Lemma 1. □

Integrality

Let A be a subring of a ring B. An element b of B is said to be *integral* over A if and only if b is a root of a monic polynomial with coefficients in A, that is, there are elements $a_1, \ldots, a_n$ of A such that

$$b^n + a_1 b^{n-1} + \cdots + a_n = 0. \tag{7.5.1}$$

Clearly, every element of A is integral over A.

Let $b_1, \ldots, b_m$ be elements of B and let $A[b_1, \ldots, b_m]$ be the subring of B generated by A and $b_1, \ldots, b_m$. Note that every element of the ring $A[b_1, \ldots, b_m]$ can be expressed as a polynomial in $b_1, \ldots, b_m$ with coefficients in A. If $B = A[b_1, \ldots, b_m]$, we say that B is *finitely generated* over A, and $b_1, \ldots, b_m$ are *generators* of B over A.

Theorem 7.5.2. *Let A be a subring of an integral domain B and let b be an element of B. Then the following are equivalent:*

(1) *The element b is integral over A.*

(2) *The ring $A[b]$ is a finitely generated A-module.*

(3) *The ring $A[b]$ is contained in subring C of B such that C is a finitely generated A-module.*

Proof: Assume that (1) holds. From Formula 7.5.1, we have

$$b^{n+i} = -(a_1 b^{n+i-1} + \cdots + a_n b^i) \quad \text{for all } i \geq 0.$$

Hence, by induction, all positive powers of b are in the A-module generated by $1, b, \ldots, b^{n-1}$. Thus $A[b]$ is generated, as an A-module, by $1, b, \ldots, b^{n-1}$. Thus (1) implies (2).

To see that (2) implies (3), let $C = A[b]$.

Assume that (3) holds. Let $c_1, \ldots, c_n$ be generators of C as an A-module. Then there are coefficients a_{ij} in A such that for each $i = 1, \ldots, n$,

$$bc_i = \sum_{j=1}^{n} a_{ij} c_j.$$

Then we have that

$$\sum_{j=1}^{n} (\delta_{ij} b - a_{ij}) c_j = 0.$$

By multiplying on the left by the adjoint of the matrix $(\delta_{ij} b - a_{ij})$, we deduce that

$$\det(\delta_{ij} b - a_{ij}) c_j = 0 \quad \text{for } j = 1, \ldots, n.$$

Therefore, we have

$$\det(\delta_{ij} b - a_{ij}) = 0.$$

Expanding out the determinant gives a equation of the form (7.5.1). Hence b is integral over A. Thus (3) implies (1). □

Corollary 1. *If A is a subring of an integral domain B, and $b_1, \ldots, b_m$ are elements of B, each integral over A, then the ring $A[b_1, \ldots, b_m]$ is a finitely generated A-module.*

Proof: The proof is by induction on m. The case $m = 1$ follows from Theorem 7.5.2. Let $A_i = A[b_1, \ldots, b_i]$ and assume that A_{m-1} is a finitely generated A-module. Then $A_m = A_{m-1}[b_m]$ is a finitely generated A_{m-1}-module by Theorem 7.5.2. Thus A_m is a finitely generated A-module. □

Corollary 2. *If A is a subring of an integral domain B, then the set C of all elements of B that are integral over A is a subring of B containing A.*

Proof: Let c, d be in C. Then $A[c, d]$ is a finitely generated A-module by Corollary 1. Hence $c + d$ and cd are integral over A by Theorem 7.5.2. Thus C is a subring of B. □

Let A be a subring of an integral domain B. The subring C of B of all elements of B that are integral over A is called the *integral closure* of A in B. If $C = A$, then A is said to be *integrally closed* in B. If $C = B$, then B is said to be *integral* over A.

Lemma 2. *Let A be a subring of an integral domain B such that B is integral over A.*

(1) *If Q is a prime ideal of B, and $P = A \cap Q$, then B/Q is integral over A/P.*

(2) *If S is a multiplicatively closed subset of A, then $S^{-1}B$ is integral over $S^{-1}A$.*

Proof: Let b be in B. Then there are elements $a_1, \ldots, a_n$ of A such that

$$b^n + a_1 b^{n-1} + \cdots + a_n = 0.$$

Upon reducing mod Q, we find that $b + Q$ is integral over A/P.

(2) Let b/s be in $S^{-1}B$. Then dividing the last equation by s^n gives

$$(b/s)^n + (a_1/s)(b/s)^{n-1} + \cdots + (a_n/s^n) = 0.$$

Thus b/s is integral over $S^{-1}A$. □

Lemma 3. *Let A be a subring of an integral domain B such that B is integral over A. Then A is a field if and only if B is a field.*

Proof: Suppose that A is a field and b is a nonzero element of B. Then there are coefficients $a_1, \ldots, a_n$ in A such that

$$b^n + a_1 b^{n-1} + \cdots + a_n = 0,$$

and n is as small as possible. As B is an integral domain, we have that $a_n \neq 0$. Hence

$$b^{-1} = -a_n^{-1}(b^{n-1} + a_1 b^{n-2} + \cdots + a_{n-1})$$

exists in B, and so B is a field.

Conversely, suppose that B is a field and a is a nonzero element of A. Then a^{-1} exists in B and so is integral over A. Hence, there are coefficients $a_1, \ldots, a_n$ in A such that

$$a^{-n} + a_1 a^{-n+1} + \cdots + a_n = 0.$$

Then we have

$$a^{-1} = -(a_1 + a_2 a + \cdots + a_n a^{n-1})$$

is an element of A, and so A is a field. □

Lemma 4. *Let A be a subring of an integral domain B such that B is integral over A, let Q be a prime ideal of B, and let $P = A \cap Q$. Then P is maximal in A if and only if Q is maximal in B.*

Proof: By Lemma 2(1), we have that B/Q is integral over A/P. As Q is prime, we have that B/Q is an integral domain. Therefore A/P is a field if and only if B/Q is a field by Lemma 3. □

Theorem 7.5.3. *Let A be a subring of an integral domain B such that B is integral over A, and let P be a prime ideal of A. Then there is a prime ideal Q of B such that $A \cap Q = P$.*

Proof: Let $B_P = (A - P)^{-1}B$. Then B_P is integral over A_P by Lemma 2(2). Consider the commutative diagram of natural injections

$$\begin{array}{ccc} A & \longrightarrow & B \\ \alpha \downarrow & & \downarrow \beta \\ A_P & \longrightarrow & B_P. \end{array}$$

Let N be a maximal ideal of B_P. Then $M = A_P \cap N$ is maximal in A_P by Lemma 4. Hence M is the unique maximal ideal of the local ring A_P. Let $Q = \beta^{-1}(N)$. Then Q is a prime ideal of B such that

$$A \cap Q = \alpha^{-1}(M) = P.$$ □

Valuation Rings

Definition: A subring B of a field F is a *valuation ring* of F if and only if for each nonzero element x of F, either x is in B or x^{-1} is in B.

Theorem 7.5.4. *If B is a valuation ring of a field F, then*

(1) *B is a local ring; and*

(2) *B is integrally closed in F.*

Proof: (1) Let M be the set of nonunits of B. If x is in M and b in B, then bx is in M, otherwise $(bx)^{-1}$ would be in B, and therefore the element $x^{-1} = b(bx)^{-1}$ would be in B, which is not the case. Now let x, y be nonzero elements of M. Then either xy^{-1} is in B or $x^{-1}y$ is in B. If xy^{-1} is in B, then $x + y = (1 + xy^{-1})y$ is in M, and likewise if $x^{-1}y$ is in B. Hence M is an ideal of B and therefore B is a local ring by Lemma 1.

(2) Let x in F be integral over B. Then there are coefficients $b_1, \ldots, b_n$ in B such that

$$x^n + b_1x^{n-1} + \cdots + b_n = 0.$$

If x is in B, then we are done, otherwise x^{-1} is in B and so

$$x = -(b_1 + b_2x^{-1} \cdots + b_nx^{1-n})$$

is in B. Thus B is integrally closed in F. □

Let F be a field and let K be an algebraically closed field. Let Σ be the set of all pairs (A, α), where A is a subring of F and $\alpha : A \to K$ is a homomorphism. Define a partial ordering on Σ by the rule

$$(A, \alpha) \leq (B, \beta) \text{ if and only if } A \subset B \text{ and } \beta \mid A = \alpha.$$

By Zorn's Lemma, the set Σ has a maximal element.

Theorem 7.5.5. *Let (B, β) be a maximal element of Σ. Then B is a valuation ring of F.*

Proof: We first show that B is a local ring with $M = \ker \beta$ its maximal ideal. The ring $\beta(B)$ is an integral domain, since it is a subring of the field K. Therefore M is prime. We extend β to a homomorphism $\gamma : B_M \to K$ by setting

$$\gamma(b/s) = \beta(b)/\beta(s)$$

for all b in B and s in $B - M$, which is allowable, since $\beta(s) \neq 0$. As the pair (B, β) is maximal, we have that $B = B_M$. Therefore, every element of $B - M$ is a unit, and so B is a local ring and M is its maximal ideal by Lemma 1.

Now let x be a nonzero element of F and let $M[x]$ be the ideal of $B[x]$ of all polynomials in x with coefficients in M. We now show that either $M[x] \neq B[x]$ or $M[x^{-1}] \neq B[x^{-1}]$. On the contrary, suppose that $M[x] = B[x]$ and $M[x^{-1}] = B[x^{-1}]$. Then there are coefficients $a_0, \ldots, a_m$ and $b_0, \ldots, b_n$ in M such that

$$a_0 + a_1x + \cdots + a_mx^m = 1,$$
$$b_0 + b_1x^{-1} + \cdots + b_nx^{-n} = 1$$

and m and n are as small as possible. By replacing x by x^{-1}, if necessary, we may assume that $m \geq n$. Multiplying the second equation by x^n gives

$$(1 - b_0)x^n = b_1x^{n-1} + \cdots + b_n.$$

As b_0 is in M, we have that $1 - b_0$ is in $B - M$ and so is a unit of B. Therefore, we can write

$$x^n = c_1x^{n-1} + \cdots + c_n$$

with c_i in M. Hence, we can replace x^m by $c_1x^{m-1} + \cdots + c_nx^{m-n}$ in the first equation. This contradicts the minimality of m. Thus, either $M[x] \neq B[x]$ or $M[x^{-1}] \neq B[x^{-1}]$.

We now show that either x is in B or x^{-1} is in B. Let $B' = B[x]$. By replacing x by x^{-1}, if necessary, we may assume that $M[x] \neq B'$. Then $M[x]$ is contained in a maximal ideal M' of B'; and $B \cap M' = M$, since $B \cap M'$ is a proper ideal of B containing M. Hence, the inclusion of B into B' induces an embedding of the field $k = B/M$ into the field $k' = B'/M'$. Moreover $k' = k[\overline{x}]$ where $\overline{x} = x + M'$. Hence, if $\overline{x} \neq 0$, there are coefficients $c_0, \ldots, c_n$ in k such that

$$\overline{x}^{-1} \quad = \quad c_0 + c_1\overline{x} + \cdots + c_n\overline{x}^n.$$

Hence, we have

$$0 \quad = \quad -1 + c_0\overline{x} + \cdots + c_n\overline{x}^{n+1}.$$

Therefore $\overline{x}$ is algebraic over k.

Now the homomorphism $\beta : B \to K$ induces an embedding $\overline{\beta} : k \to K$ because $M = \ker \beta$. Let $p(t)$ be the irreducible polynomial for $\overline{x}$ over k.

As K is algebraically closed, the polynomial $(\overline{\beta}p)(t)$ has a root r in K. We extend $\overline{\beta}$ to a homomorphism $\overline{\beta}' : k' \to K$ as follows: Let y be in k'. Then there is a polynomial $f(t)$ over k such that $y = f(\overline{x})$. Define

$$\overline{\beta}'(y) = (\overline{\beta}f)(r).$$

Then $\overline{\beta}'$ is well defined, since if $g(t)$ is another polynomial over k such that $y = g(\overline{x})$, then $(g-f)(\overline{x}) = 0$, and so $p(t)$ divides $(g-f)(t)$, whence $(\overline{\beta}p)(t)$ divides $(\overline{\beta}(g-f))(t)$ and so

$$(\overline{\beta}g)(r) = (\overline{\beta}f)(r).$$

Clearly $\overline{\beta}'$ is a ring homomorphism extending $\overline{\beta}$. Composing $\overline{\beta}'$ with the natural projection $B' \to k'$ gives a homomorphism $\beta' : B' \to K$ extending β. As (B, β) is maximal, $B = B'$, and so x is in B. Thus B is a valuation ring of F. □

Corollary 3. *If A is a subring of a field F, then the integral closure C of A in F is the intersection of all the valuation rings of F containing A.*

Proof: Let B be a valuation ring of F containing A. Then B is integrally closed in F by Theorem 7.5.4. Hence, any element of F that is integral over A is an element of B. Therefore $C \subset B$.

Now let x be an element of $F - C$ and let $A' = A[x^{-1}]$. Then x is not in A', since otherwise there would be coefficients $a_0, \ldots, a_n$ in A such that

$$x = a_0 + a_1x^{-1} + \cdots + a_nx^{-n}$$

and so we would have

$$x^{n+1} - a_0x^n - \cdots - a_n = 0$$

and therefore x would be in C, which is not the case. Hence x^{-1} is a nonunit of A' and so is contained in a maximal ideal M of A'. Let $\overline{k}$ be the algebraic closure of the field $k = A'/M$ and let $\alpha : A' \to \overline{k}$ be the composition of the natural projection $A' \to k$ followed by the inclusion $k \to \overline{k}$. Then α can be extended to a homomorphism $\beta : B \to \overline{k}$ where B is a valuation ring of F containing A' by Theorem 7.5.5. Then x^{-1} is also a nonunit in B, since $\beta(x^{-1}) = 0$. Therefore x is not in B. Hence C is the intersection of all the valuation rings of F containing A. □

Lemma 5. *Every algebraically closed field is infinite.*

Proof: Let K be an algebraically closed field and on the contrary, suppose that K is finite. Let p be the characteristic of K. Then K is a finite dimensional vector space over the field of order p. Hence K has p^n elements for some positive integer n. Therefore, the group K^* of units of K has order $p^n - 1$. Let q be a prime not dividing $p^n - 1$. Then the polynomial $(t^q - 1)/(t - 1)$ has no root in K, since the order of every element of K^* divides $p^n - 1$. Thus K is not algebraically closed, which is a contradiction. Hence K must be infinite. □

Theorem 7.5.6. *Let A be a subring of an integral domain B such that B is finitely generated over A, and let b be a nonzero element of B. Then there exists a nonzero element a of A with the property that any homomorphism α of A into an algebraically closed field K, such that $\alpha(a) \neq 0$, can be extended to a homomorphism $\beta : B \to K$ such that $\beta(b) \neq 0$.*

Proof: By induction on the number of generators of B over A, we reduce immediately to the case where B is generated over A by a single element x. Assume first that x is not algebraic over A, that is, no nonzero polynomial with coefficients in A has x as a root. As $B = A[x]$, there are coefficients $a_0, \ldots, a_n$ in A, with $a_0 \neq 0$, such that

$$b = a_0x^n + a_1x^{n-1} + \cdots + a_n.$$

Set $a = a_0$ and let

$$\alpha : A \to K$$

be a homomorphism such that $\alpha(a) \neq 0$. Now the nonzero polynomial

$$\alpha(a_0)t^n + \alpha(a_1)t^{n-1} + \cdots + \alpha(a_n)$$

has at most n roots in K; therefore, there is an element y of K such that

$$\alpha(a_0)y^n + \alpha(a_1)y^{n-1} + \cdots + \alpha(a_n) \neq 0,$$

since K is infinite by Lemma 5. Extend $\alpha : A \to K$ to a homomorphism

$$\beta : B \to K$$

by setting $\beta(x) = y$. Then $\beta(b) \neq 0$, as required.

Assume next that x is algebraic over A. Then x is integral over the field F of fractions of A. As b is in $F[x]$, we have that b is integral over F by Theorem 7.5.2. Hence b is algebraic over A, and therefore b^{-1} is algebraic over A. Hence, there are coefficients $c_0, \ldots, c_m$ and $d_0, \ldots, d_n$ in A, with $c_0d_0 \neq 0$, such that

$$c_0x^m + c_1x^{m-1} + \cdots + c_m = 0,$$
$$d_0b^{-n} + d_1b^{1-n} + \cdots + d_n = 0.$$

Set $a = c_0d_0$ and let $\alpha : A \to K$ be a homomorphism such that $\alpha(a) \neq 0$. Then α can be extended first to a homomorphism

$$\alpha' : A[a^{-1}] \to K$$

by setting

$$\alpha'(a^{-1}) = \alpha(a)^{-1},$$

and then to a homomorphism $\gamma : C \to K$, where C is a valuation ring of the field of fractions of B, by Theorem 7.5.5. As $a = c_0d_0$, we have that x is integral over $A[a^{-1}]$. Therefore x is in C by Corollary 3, and so C contains B. Likewise, since $a = c_0d_0$, we have that b^{-1} is integral over $A[a^{-1}]$. Therefore b^{-1} is in C, and so b is a unit in C. Hence $\gamma(b) \neq 0$. Now take $\beta : B \to K$ to be the restriction of γ to B. □

Selberg's Lemma

Let A be a subring of $\mathbb{C}$. Then A is said to be *finitely generated* if and only if A is finitely generated over $\mathbb{Z}$, that is, there are a finite number of elements $a_1, \ldots, a_m$ of A, called the *generators* of A, such that every element of A can be expressed as a polynomial in $a_1, \ldots, a_m$ with coefficients in $\mathbb{Z}$.

Theorem 7.5.7. *Let A be a finitely generated subring of $\mathbb{C}$. Then every subgroup of $\mathrm{GL}(n, A)$ has a torsion-free normal subgroup of finite index.*

Proof: For each prime p in $\mathbb{Z}$, let α_p be the composite

$$\mathbb{Z} \xrightarrow{\text{proj}} \mathbb{Z}_p \xrightarrow{\text{inj}} \overline{\mathbb{Z}}_p,$$

where $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ and $\overline{\mathbb{Z}}_p$ is the algebraic closure of $\mathbb{Z}_p$. By Theorem 7.5.6, there is a nonzero integer m with the property that for any prime p not dividing m, the homomorphism $\alpha_p : \mathbb{Z} \to \overline{\mathbb{Z}}_p$ can be extended to a homomorphism $\beta_p : A \to \overline{\mathbb{Z}}_p$. As $\beta_p(1) = 1$, the kernel of β_p is a proper ideal of A. Let M_p be a maximal ideal of A containing $\ker \beta_p$. Then

$$p\mathbb{Z} = \mathbb{Z} \cap \ker \beta_p \subset \mathbb{Z} \cap M_p.$$

As $p\mathbb{Z}$ is a maximal ideal of $\mathbb{Z}$, we have that $\mathbb{Z} \cap M_p = p\mathbb{Z}$. Therefore A/M_p is a field of characteristic p.

Now $\beta_p : A \to \overline{\mathbb{Z}}_p$ induces an embedding of $A/\ker \beta_p$ into $\overline{\mathbb{Z}}_p$. As $\overline{\mathbb{Z}}_p$ is an algebraic extension of $\mathbb{Z}_p$, we have that $A/\ker \beta_p$ is algebraic over $\mathbb{Z}_p$. Therefore A/M_p is an algebraic extension of $\mathbb{Z}_p$. As A is finitely generated over $\mathbb{Z}$, we have that A/M_p is finitely generated over $\mathbb{Z}_p$. Therefore A/M_p is a finite extension of $\mathbb{Z}_p$ by Corollary 1. Hence A/M_p is a finite field.

Let $\mathrm{GL}_n(A, M_p)$ be the kernel of the natural projection from $\mathrm{GL}_n(A)$ into $\mathrm{GL}_n(A/M_p)$. Then $\mathrm{GL}_n(A, M_p)$ is a normal subgroup of $\mathrm{GL}_n(A)$ of finite index, since $\mathrm{GL}_n(A/M_p)$ is a finite group. Let Γ be an arbitrary subgroup of $\mathrm{GL}_n(A)$ and set

$$\Gamma_p = \Gamma \cap \mathrm{GL}_n(A, M_p).$$

Then Γ_p is a normal subgroup of Γ of finite index.

Let p, q be distinct primes not dividing m and set

$$\Gamma_{p,q} = \Gamma_p \cap \Gamma_q.$$

Then $\Gamma_{p,q}$ is a normal subgroup of Γ of finite index. We now prove that $\Gamma_{p,q}$ is torsion-free by contradiction. Let g be an element of $\Gamma_{p,q}$ of finite order $r > 1$. We may assume, without loss of generality, that r is prime. As $g^r = I$, each eigenvalue of g is an rth root of unity. By Lemma 9 of §7.4, we have that g is conjugate in $\mathrm{GL}(n, \mathbb{C})$ to a unitary matrix. Hence g is conjugate to a diagonal matrix. Now since the order of g is r, at least one eigenvalue of g is a primitive rth root of unity ω.

Let $B = A[\omega]$. By Theorem 7.5.3, there is a prime ideal Q_p of B such that $A \cap Q_p = M_p$. Let $\phi(t)$ be the characteristic polynomial of g. As g is in $\mathrm{GL}_n(A, M_p)$, we have

$$\phi(t) \equiv (t-1)^n \bmod M_p[t].$$

Therefore, we have

$$\phi(\omega) \equiv (\omega - 1)^n \bmod Q_p.$$

As $\phi(\omega) = 0$, we have that $\omega - 1$ is in Q_p, since B/Q_p is an integral domain. Hence, there is a nonzero element x of Q_p such that $\omega = 1 + x$. Observe that

$$1 = (1+x)^r = 1 + rx + \frac{r(r-1)}{2}x^2 + \cdots + x^r.$$

Therefore, there is a y in Q_p such that

$$1 = 1 + x(r+y).$$

Thus $x(r+y) = 0$ and so $r + y = 0$. Hence r is in $\mathbb{Z} \cap Q_p = p\mathbb{Z}$. As r is prime, we have that $r = p$. Likewise $r = q$, and we have a contradiction. Thus $\Gamma_{p,q}$ is torsion-free. □

Corollary 4. (Selberg's lemma) *Every finitely generated subgroup Γ of $\mathrm{GL}(n, \mathbb{C})$ has a torsion-free normal subgroup of finite index.*

Proof: Let Γ be the group generated by $g_1, \ldots, g_m$ and let A be the subring of $\mathbb{C}$ generated by all the entries of the matrices $g_1^{\pm 1}, \ldots, g_m^{\pm 1}$. Then Γ is a subgroup of $\mathrm{GL}(n, A)$ and so has a torsion-free normal subgroup of finite index by Theorem 7.5.7. □

Corollary 5. *Every finitely generated subgroup of $\mathrm{I}(H^n)$ has a torsion-free normal subgroup of finite index.*

Proof: The group $\mathrm{PO}(n, 1)$ is a subgroup of $\mathrm{GL}(n+1, \mathbb{C})$. □

Exercise 7.5

1. Let Γ be a group with a torsion-free subgroup of finite index. Prove that Γ has a torsion-free normal subgroup of finite index.
2. Γ be a group with a torsion-free subgroup of finite index. Prove that there is an upper bound on the set of finite orders of elements of Γ.
3. Let A be a finitely generated subring of $\mathbb{C}$. Prove that every subgroup of $\mathrm{PSL}(2, A)$ has a torsion-free normal subgroup of finite index.
4. Prove that every finitely generated subgroup of $\mathrm{PSL}(2, \mathbb{C})$ has a torsion-free normal subgroup of finite index.
5. Prove that every finitely generated subgroup Γ of $\mathrm{GL}(n, \mathbb{C})$ is residually finite, that is, for each $g \neq 1$ in Γ, there is normal subgroup Γ_g of Γ of finite index such that g is in $\Gamma - \Gamma_g$. Conclude that every finitely generated group of hyperbolic isometries is residually finite.

§7.6. Historical Notes

§7.1. Theorems 7.1.2 and 7.1.3 for 2- and 3-dimensional hyperbolic polyhedra appeared in Poincaré's 1883 *Mémoire sur les groupes kleinéens* [332]. Theorems 7.1.3 and 7.1.4 for spherical and Euclidean n-simplices appeared in Coxeter's 1932 paper *The polytopes with regular-prismatic vertex figures II* [86]. See also Witt's 1941 paper *Spiegelungsgruppen und Aufzählung halbeinfacher Liescher Ringe* [414]. Theorems 7.1.1 and 7.1.3 for compact polyhedra were proved by Aleksandrov in his 1954 Russian paper *On the filling of space by polyhedra* [12] and in general by Seifert in his 1975 paper *Komplexe mit Seitenzuordnung* [371]. Coxeter groups were introduced by Coxeter in his 1935 paper *The complete enumeration of finite groups of the form* $R_i^2 = (R_iR_j)^{k_{ij}} = 1$ [88].

§7.2. The spherical, Euclidean, and hyperbolic triangle reflection groups were determined by Schwarz in his 1873 paper *Ueber diejenigen Fälle, in welchen die Gaussische hypergeometrische Reihe eine algebraische Function ihres vierten Elementes darstellt* [369]. Hyperbolic, tetrahedron, reflection groups were considered by Dyck in his 1883 paper *Über die durch Gruppen linearer Transformationen gegebenen regulären Gebietseintheilungen des Raumes* [112]. The spherical, tetrahedron, reflection groups were determined by Goursat in his 1889 paper *Sur les substitutions orthogonales et les divisions régulières de l'espace* [155]. The spherical and Euclidean, n-simplex, reflection groups were enumerated by Coxeter in his 1931 note *Groups whose fundamental regions are simplexes* [85]. See also Coxeter's 1934 paper *Discrete groups generated by reflections* [87]. The hyperbolic, tetrahedron, reflection groups were described by Coxeter and Whitrow in their 1950 paper *World-structure and non-Euclidean honeycombs* [96]. The hyperbolic, compact n-simplex, reflection groups were enumerated by Lannér in his 1950 thesis *On complexes with transitive groups of automorphisms* [253]. Theorem 7.2.2 for spherical and Euclidean n-simplices appeared in Coxeter's 1932 paper [86]. See also Witt's 1941 paper [414]. Theorem 7.2.2 for hyperbolic n-simplices appeared in Vinberg's 1967 paper *Discrete groups generated by reflections in Lobacevskii spaces* [397].

§7.3. Theorem 7.3.1 and Figure 7.3.5 appeared in Vinberg's 1967 paper [397]. The hyperbolic, noncompact n-simplex, reflection groups were enumerated by Chein in his 1969 paper *Recherche des graphes des matrices de Coxeter hyperboliques d'ordre* ≤ 10 [79]. For a survey of hyperbolic reflection groups, see Vinberg's 1985 paper *Hyperbolic reflection groups* [398]. References for reflection groups are Bourbaki's 1968 treatise *Groupes et Algèbres de Lie* [54], Coxeter's 1973 treatise *Regular Polytopes* [92], and Humphreys' 1990 treatise *Reflection Groups and Coxeter Groups* [201]. A complete list of the Coxeter graphs of the hyperbolic, noncompact n-simplex, reflection groups can be found in Humphreys' 1990 treatise [201]. For the history of reflection groups, see the historical notes in Bourbaki's 1968 treatise [54] and in Coxeter's 1973 treatise [92].

§7.4. Theorem 7.4.1 appeared in Auslander's 1965 paper *An account of the theory of crystallographic groups* [26]. Theorems 7.4.2 and 7.4.3 were proved by Bieberbach in his 1911 paper *Über die Bewegungsgruppen der Euklidischen Räume I* [46]. Our proof of Theorem 7.4.3 was given by Buser in his 1985 paper *A geometric proof of Bieberbach's theorems on crystallographic groups* [65]. Theorem 7.4.4 was proved by Bieberbach in his 1912 paper *Über die Bewegungsgruppen der Euklidischen Räume II* [47]. A description of the 2-dimensional crystallographic groups can be found in Coxeter and Moser's 1980 treatise *Generators and Relations for Discrete Groups* [95]. For the history and classification of crystallographic groups, see the 1978 treatise *Crystallographic Groups of Four-Dimensional Space* of Brown, Bülow, Neubüser, Wondratschek, and Zassenhaus [61]. Lemma 9 was proved by Moore in his 1898 paper *An universal invariant for finite groups of linear substitutions* [304] and by Loewy in his 1898 paper *Ueber bilineare Formen mit conjugirt imaginären Variabeln* [269]. Theorem 7.4.5 appeared in Zassenhaus' 1948 paper *Über einen Algorithmus zur Bestimmung der Raumgruppen* [422]. As a reference for crystallographic groups, see Farkas' 1981 article *Crystallographic groups and their mathematics* [130].

§7.5. The material on integrality and valuation rings is basic commutative ring theory which was adapted from Chapter 5 of Atiyah and Macdonald's 1969 text *Introduction to Commutative Algebra* [25]. Selberg's lemma was proved by Selberg in his 1960 paper *On discontinuous groups in higher-dimensional symmetric spaces* [372]. For another proof of Selberg's lemma, see Alperin's 1987 paper *An elementary account of Selberg's lemma* [14].

CHAPTER 8

Geometric Manifolds

In this chapter, we lay down the foundation for the theory of hyperbolic manifolds. We begin with the notion of a geometric space. Examples of geometric spaces are S^n, E^n, and H^n. In Sections 8.2 and 8.3, we study manifolds locally modeled on a geometric space X via a group G of similarities of X. Such a manifold is called an (X,G)-manifold. In Section 8.4, we study the relationship between the fundamental group of an (X,G)-manifold and its (X,G)-structure. In Section 8.5, we study the role of metric completeness in the theory of (X,G)-manifolds. In particular, we prove that if M is a complete (X,G)-manifold, with X simply connected, then there is a discrete subgroup Γ of G of isometries acting freely on X such that M is isometric to X/Γ. The chapter ends with a discussion of the role of curvature in the theory of spherical, Euclidean, and hyperbolic manifolds.

§8.1. Geometric Spaces

We begin our study of geometric manifolds with the definition of a topological manifold without boundary.

Definition: An *n-manifold* (without boundary) is a Hausdorff space M that is locally homeomorphic to E^n, that is, for each point u of M, there is an open neighborhood U of u in M such that U is homeomorphic to an open subset of E^n.

Example: Euclidean n-space E^n is an n-manifold.

Definition: A *closed* manifold is a compact manifold (without boundary).

Example: Spherical n-space S^n is a closed n-manifold.

Definition: An *open* manifold is a manifold (without boundary) all of whose connected components are noncompact.

Example: Hyperbolic n-space H^n is an open n-manifold.

Definition: An *n-manifold-with-boundary* is a Hausdorff space M that is locally homeomorphic to $\overline{U}^n = \{x \in E^n : x_n \geq 0\}$.

Example: Closed upper half-space $\overline{U}^n$ is n-manifold-with-boundary.

Let M be an n-manifold-with-boundary and let M° be the set of points of M that have an open neighborhood homeomorphic to an open subset of U^n. Then M° is an open subset of M called the *interior* of M. The interior M° of M is an n-manifold. Let $\partial M = M - M^\circ$. Then ∂M is a closed subset of M called the *boundary* of M. The boundary ∂M of M is an $(n-1)$-manifold. A manifold-with-boundary is often called a manifold; however, in this book, a manifold will mean a manifold without boundary.

Definition: An n-dimensional *geometric space* is a metric space X satisfying the following axioms:

(1) The metric space X is geodesically connected; that is, each pair of distinct points of X are joined by a geodesic segment in X.

(2) The metric space X is geodesically complete; that is, each geodesic arc $\alpha : [a, b] \to X$ extends to a unique geodesic line $\lambda : \mathbb{R} \to X$.

(3) There is a continuous function $\varepsilon : E^n \to X$ and a $k > 0$ such that ε maps $B(0, k)$ homeomorphically onto $B(\varepsilon(0), k)$; for each point u of S^{n-1}, the map $\lambda : \mathbb{R} \to X$, defined by $\lambda(t) = \varepsilon(tu)$, is a geodesic line such that λ restricts to a geodesic arc on the interval $[-k, k]$;

(4) The metric space X is homogeneous.

One should compare Axioms 1-4 with Euclid's Postulates 1-4 in §1.1. Note that Axioms 3 and 4 imply that X is an n-manifold.

Example 1. Euclidean n-space E^n is an n-dimensional geometric space.

Example 2. Spherical n-space S^n is an n-dimensional geometric space. Define $\varepsilon : E^n \to S^n$ by $\varepsilon(0) = e_{n+1}$ and

$$\varepsilon(x) = (\cos|x|)e_{n+1} + (\sin|x|)\frac{x}{|x|} \quad \text{for } x \neq 0.$$

Then ε satisfies Axiom 3 with $k = \pi/2$.

Example 3. Hyperbolic n-space H^n is an n-dimensional geometric space. Define $\varepsilon : E^n \to H^n$ by $\varepsilon(0) = e_{n+1}$ and

$$\varepsilon(x) = (\cosh|x|)e_{n+1} + (\sinh|x|)\frac{x}{|x|} \quad \text{for } x \neq 0.$$

Then ε satisfies Axiom 3 for all $k > 0$.

Theorem 8.1.1. *Let X be an n-dimensional geometric space and suppose that $\varepsilon : E^n \to X$ is a function satisfying Axiom 3. Then for each geodesic line $\lambda : \mathbb{R} \to X$ such that $\lambda(0) = \varepsilon(0)$, there is a point u of S^{n-1} such that $\lambda(t) = \varepsilon(tu)$ for all t.*

Proof: Let $\lambda : \mathbb{R} \to X$ be a geodesic line such that $\lambda(0) = \varepsilon(0)$. Then there is a $c > 0$ such that the restriction of λ to $[0, c]$ is a geodesic arc. Let k be the constant in Axiom 3 and choose $b > 0$ but less than both c and k. Then $\lambda(b)$ is in $B(\varepsilon(0), k)$. Hence, there is a point u of S^{n-1} such that $\varepsilon(bu) = \lambda(b)$. Define $\alpha : [0, c] \to X$ by

$$\alpha(t) = \begin{cases} \varepsilon(tu), & 0 \leq t \leq b, \\ \lambda(t), & b \leq t \leq c. \end{cases}$$

Then α is the composite of two geodesic arcs. Hence α is a geodesic arc by Theorem 1.4.3, since

$$d(\lambda(0), \lambda(b)) + d(\lambda(b), \lambda(c)) = d(\lambda(0), \lambda(c)).$$

By Axiom 2, the arc α extends to a unique geodesic line $\mu : \mathbb{R} \to X$. Now λ and μ both extend the restriction of λ to $[b, c]$. Therefore $\lambda = \mu$. Hence $\lambda(t) = \varepsilon(tu)$ for $0 \leq t \leq b$. Furthermore $\lambda(t) = \varepsilon(tu)$ for all t, since λ is the unique geodesic line extending the restriction of λ to $[0, b]$. □

Theorem 8.1.2. *Let $B(x, r)$ be an open ball in a geometric space X and let $\overline{B}(x, r)$ be its topological closure in X. Then*

$$\overline{B}(x, r) = \{y \in X : d(x, y) \leq r\} = C(x, r).$$

Furthermore the closed ball $C(x, r)$ is compact.

Proof: In general, in a metric space, $\overline{B}(x, r) \subset C(x, r)$. As every point of the set $\{y \in X : d(x, r) = r\}$ is joined to x by a geodesic segment in $\overline{B}(x, r)$ by Axiom 1, we also have the reverse inclusion. Thus, we have

$$\overline{B}(x, r) = C(x, r).$$

Let $\varepsilon : E^n \to X$ be a function satisfying Axiom 3 with $\varepsilon(0) = x$. As ε is continuous, $\varepsilon(\overline{B}(0, r)) \subset \overline{B}(x, r)$. Let y be an arbitrary point of $C(x, r)$. By Axiom 1, there is a geodesic arc $\alpha : [0, \ell] \to X$ from x to y. By Axiom 2, the arc α extends to a geodesic line $\lambda : \mathbb{R} \to X$. By Theorem 8.1.1, there is a point u of S^{n-1} such that $\lambda(t) = \varepsilon(tu)$ for all t. Hence $y = \varepsilon(\ell u)$, where $\ell = d(x, y) \leq r$. Therefore y is in $\varepsilon(C(0, r))$. Hence $\varepsilon(C(0, r)) = C(x, r)$. As $C(0, r)$ is compact and ε is continuous, $C(x, r)$ is compact. □

Free Group Actions

Let Γ be a discrete group of isometries of an n-dimensional geometric space X. Then Γ is discontinuous by Theorems 5.3.5 and 8.1.2. Hence X/Γ is a metric space by Theorems 5.3.4 and 6.5.1. We next consider a sufficient condition on the action of Γ on X so that X/Γ is an n-manifold.

Definition: A group Γ acting on a set X acts *freely* on X if and only if for each x in X, the stabilizer subgroup $\Gamma_x = \{g \in \Gamma : gx = x\}$ is trivial.

Example: The group $\{\pm 1\}$ acts freely on S^n.

Definition: A function $\xi : X \to Y$ between metric spaces is a *local isometry* if and only if for each point x of X, there is an $r > 0$ such that ξ maps $B(x, r)$ isometrically onto $B(\xi(x), r)$.

Theorem 8.1.3. *Let Γ be a group of isometries of a metric space X such that Γ acts freely and discontinuously on X. Then the quotient map*

$$\pi : X \to X/\Gamma$$

is a local isometry and a covering projection. Furthermore, if X is connected, then Γ is the group of covering transformations of π.

Proof: Let x be an arbitrary point of X. Then we have

$$\pi(B(x,r)) = B(\pi(x), r)$$

for each $r > 0$ by Theorem 6.5.2. Hence π is an open map. Now as Γ acts freely on X, the map $g \mapsto gx$ is a bijection from Γ onto Γx. The set $\Gamma x - \{x\}$ is closed by Theorem 5.3.4. Hence, we have

$$\mathrm{dist}(x, \Gamma x - \{x\}) > 0.$$

Now set

$$s = \frac{1}{2}\mathrm{dist}(x, \Gamma x - \{x\})$$

and let y, z be arbitrary points of $B(x, s/2)$. Then $d(y, z) < s$. Let $g \neq 1$ be in Γ. Then

$$d(x, gx) \leq d(x, y) + d(y, gz) + d(gz, gx).$$

Hence, we have

$$\begin{aligned} d(y, gz) &\geq d(x, gx) - d(x, y) - d(z, x) \\ &\geq 2s - s/2 - s/2 \quad = \quad s. \end{aligned}$$

Therefore

$$d_\Gamma(\pi(y), \pi(z)) = \mathrm{dist}(\Gamma y, \Gamma z) = d(y, z).$$

Thus π maps $B(x, s/2)$ isometrically onto $B(\pi(x), s/2)$, and so π is a local isometry.

Now let g, h be in Γ and suppose that $B(gx, s)$ and $B(hx, s)$ overlap. Then $B(x, s)$ and $B(g^{-1}hx, s)$ overlap. Consequently

$$d(x, g^{-1}hx) < 2s.$$

Because of the choice of s, we have that $g^{-1}h = 1$ and so $g = h$. Thus, the open balls $\{B(gx, s) : g \in \Gamma\}$ are mutually disjoint in X. The orbit space metric d_Γ on X/Γ is the distance function between Γ-orbits in X. Therefore $\pi^{-1}(B(\pi(x), s))$ is the s-neighborhood of Γx in X. Hence, we have

$$\pi^{-1}(B(\pi(x), s)) = \bigcup_{g\in\Gamma} B(gx, s).$$

As each $h \neq 1$ in Γ moves $B(gx, s)$ off itself, no two points of $B(gx, s)$ are in the same Γ-orbit. Therefore π maps $B(gx, s)$ bijectively onto $B(\pi(x), s)$. Furthermore, since π is an open map, π maps $B(gx, s)$ homeomorphically onto $B(\pi(x), s)$ for each g in Γ. Hence $B(\pi(x), s)$ is evenly covered by π. Thus π is a covering projection.

If g is in Γ, then $\pi g = \pi$, and so g is a covering transformation of π. Now assume that X is connected. Choose a base point x_0 of X. Let $\tau : X \to X$ be a covering transformation of π. Then $\pi\tau = \pi$. Hence $\pi\tau(x_0) = \pi(x_0)$, and so there is an element g of Γ such that $\tau(x_0) = gx_0$. Now g and τ are both lifts of $\pi : X \to X/\Gamma$ with respect to π that agree at one point. Therefore $\tau = g$ by the unique lifting property of covering projections. Thus Γ is the group of covering transformations of π. □

X-Space-Forms

Let Γ be a discrete group of isometries of an n-dimensional geometric space X such that Γ acts freely on X. Then the orbit space X/Γ is called an *X-space-form*. By Theorem 8.1.3, an X-space-form is an n-manifold.

Choose a base point x_0 of X. Let $\alpha : [0, 1] \to X/\Gamma$ be a loop based at the point Γx_0. Lift α to a curve $\tilde{\alpha} : [0, 1] \to X$ starting at x_0. Then

$$\pi\tilde{\alpha}(1) = \alpha(1) = \Gamma x_0.$$

Now since Γ acts freely on X, there is a unique element g_α of Γ such that $\tilde{\alpha}(1) = g_\alpha x_0$. By the covering homotopy theorem, the element g_α depends only on the homotopy class $[\alpha]$ in the fundamental group $\pi_1(X/\Gamma, \Gamma x_0)$. Hence, we may define a function

$$\eta : \pi_1(X/\Gamma) \to \Gamma$$

by the formula $\eta([\alpha]) = g_\alpha$.

Theorem 8.1.4. *Let X be a simply connected geometric space and let X/Γ be an X-space-form. Then $\eta : \pi_1(X/\Gamma) \to \Gamma$ is an isomorphism.*

Proof: Let $\alpha, \beta : [0,1] \to X/\Gamma$ be loops based at Γx_0 and let $\tilde{\alpha}, \tilde{\beta} : [0,1] \to X$ be lifts starting at x_0. Then the curve $\tilde{\alpha}(g_\alpha \tilde{\beta}) : [0,1] \to X$ lifts $\alpha\beta$ and starts at x_0. Observe that

$$\tilde{\alpha} g_\alpha \tilde{\beta}(1) = g_\alpha g_\beta x_0.$$

Therefore

$$\eta([\alpha][\beta]) = \eta([\alpha\beta]) = g_\alpha g_\beta = \eta([\alpha])\eta([\beta]).$$

Thus η is a homomorphism.

Let g be an arbitrary element of Γ. As X is geodesically connected, there is a curve $\gamma : [0,1] \to X$ from x_0 to gx_0. Then $\pi\gamma : [0,1] \to X/\Gamma$ is a loop based at Γx_0 whose lift starting at x_0 is γ. Hence $\eta[\pi\gamma] = g$. Thus η is surjective. To see that η is injective, assume that $\eta([\alpha]) = 1$. Then $\tilde{\alpha}$ is a loop in X. As X is simply connected, $[\tilde{\alpha}] = 1$ and so

$$[\alpha] = \pi_*[\tilde{\alpha}] = 1.$$

Hence η is injective. Thus η is an isomorphism. □

Theorem 8.1.5. *Let X be a simply connected geometric space. Then two X-space-forms X/Γ and X/H are isometric if and only if Γ and H are conjugate in the group $\mathrm{I}(X)$ of isometries of X.*

Proof: Let ϕ be an element of $\mathrm{I}(X)$ such that $\mathrm{H} = \phi\Gamma\phi^{-1}$. Then for each g in Γ and x in X, we have

$$\phi g x = (\phi g \phi^{-1})\phi x.$$

Hence $\phi g x$ is in the same H-orbit as ϕx. Thus ϕ induces a homeomorphism

$$\overline{\phi} : X/\Gamma \to X/\mathrm{H}$$

defined by $\overline{\phi}(\Gamma x) = \mathrm{H}\phi x$. If x and y are in X, then

$$\begin{aligned} d_{\mathrm{H}}(\overline{\phi}(\Gamma x), \overline{\phi}(\Gamma y)) &= d_{\mathrm{H}}(\mathrm{H}\phi x, \mathrm{H}\phi y) \\ &= d_{\mathrm{H}}(\phi\phi^{-1}\mathrm{H}\phi x, \phi\phi^{-1}\mathrm{H}\phi y) \\ &= d_{\mathrm{H}}(\phi\Gamma x, \phi\Gamma y) \\ &= d_\Gamma(\Gamma x, \Gamma y). \end{aligned}$$

Thus $\overline{\phi}$ is an isometry.

Conversely, suppose that $\xi : X/\Gamma \to X/\mathrm{H}$ is an isometry. By Theorem 8.1.3, the quotient maps $\pi : X \to X/\Gamma$ and $\eta : X \to X/\mathrm{H}$ are covering projections. Since X is simply connected, ξ lifts to a homeomorphism $\tilde{\xi}$ such that the following diagram commutes:

$$\begin{array}{ccc} X & \xrightarrow{\tilde{\xi}} & X \\ \pi \downarrow & & \downarrow \eta \\ X/\Gamma & \xrightarrow{\xi} & X/\mathrm{H}. \end{array}$$

As π, ξ, and η are local isometries, $\tilde{\xi}$ is also a local isometry.

Let x, y be distinct points of X. As X is geodesically connected, there is a geodesic arc $\alpha : [0, \ell] \to X$ from x to y. Since $\tilde{\xi}$ is a local isometry, the curve $\tilde{\xi}\alpha$ is rectifiable and

$$|\tilde{\xi}\alpha| = |\alpha| = \ell = d(x, y).$$

Therefore, we have

$$d(\tilde{\xi}(x), \tilde{\xi}(y)) \leq d(x, y).$$

Likewise, we have

$$d(\tilde{\xi}^{-1}(x), \tilde{\xi}^{-1}(y)) \leq d(x, y).$$

Hence, we have

$$\begin{array}{rcl} d(x, y) & = & d(\tilde{\xi}^{-1}\tilde{\xi}(x), \tilde{\xi}^{-1}\tilde{\xi}(y)) \\ & \leq & d(\tilde{\xi}(x), \tilde{\xi}(y)). \end{array}$$

Therefore, we have

$$d(\tilde{\xi}(x), \tilde{\xi}(y)) = d(x, y).$$

Thus $\tilde{\xi}$ is an isometry of X.

Let g be an arbitrary element of Γ. Then we have

$$\begin{array}{rcl} \eta\tilde{\xi}g\tilde{\xi}^{-1} & = & \xi\pi g\tilde{\xi}^{-1} \\ & = & \xi\pi\tilde{\xi}^{-1} \\ & = & \eta\tilde{\xi}\tilde{\xi}^{-1} \\ & = & \eta. \end{array}$$

Hence $\tilde{\xi}g\tilde{\xi}^{-1}$ is a covering transformation of η. Therefore $\tilde{\xi}g\tilde{\xi}^{-1}$ is in H by Theorem 8.1.3. Thus H contains $\tilde{\xi}\Gamma\tilde{\xi}^{-1}$. By reversing the roles of Γ and H, we have that Γ contains $\tilde{\xi}^{-1}\mathrm{H}\tilde{\xi}$. Hence $\tilde{\xi}\Gamma\tilde{\xi}^{-1} = \mathrm{H}$. Thus Γ and H are conjugate in $\mathrm{I}(X)$. □

Exercise 8.1

1. Prove that elliptic n-space P^n is an n-dimensional geometric space.
2. Prove that the n-torus $T^n = E^n/\mathbb{Z}^n$ is an n-dimensional geometric space.
3. A metric space X is said to be *locally geodesically convex* if and only if for each x in X, there is an $r > 0$ such that any two distinct points in $B(x, r)$ are joined by a unique geodesic segment in X. Prove that every geometric space is locally geodesically convex.
4. Prove that every X-space-form is geodesically connected.
5. Let X/Γ be an X-space-form and let $N(\Gamma)$ be the normalizer of Γ in $\mathrm{I}(X)$. Prove that $\mathrm{I}(X/\Gamma)$ is isomorphic to $N(\Gamma)/\Gamma$.

§8.2. Clifford-Klein Space-Forms

Let $X = S^n, E^n$, or H^n. Then an X-space-form is called a *Clifford-Klein space-form*. Thus, a Clifford-Klein space-form is an orbit space X/Γ where Γ is a discrete group of isometries of X acting freely on X. A Clifford-Klein space-form X/Γ is also called a *spherical, Euclidean*, or *hyperbolic space-form* according as $X = S^n, E^n$, or H^n.

Theorem 8.2.1. *A discrete group Γ of isometries of $X = E^n$ or H^n acts freely on X if and only if Γ is torsion-free.*

Proof: As Γ is discontinuous, the stabilizer Γ_x is finite for each x in X. Hence, if Γ is torsion-free, then $\Gamma_x = \{1\}$ for each x in X, and so Γ acts freely on X. Conversely, suppose that Γ acts freely on X. Then every nonidentity element of Γ is either parabolic or hyperbolic, and so every nonidentity element of Γ has infinite order. Thus Γ is torsion-free. $\square$

Definition: The *volume* of a Clifford-Klein space-form X/Γ is the volume of any proper fundamental region R of Γ in X.

Note that the volume of a Clifford-Klein space-form X/Γ is well defined, since all the proper fundamental regions of Γ have the same volume by Theorem 6.5.5.

Theorem 8.2.2. *If X/Γ and X/H are two isometric Clifford-Klein space-forms, then*

$$\mathrm{Vol}(X/\Gamma) = \mathrm{Vol}(X/\mathrm{H}).$$

Proof: By Theorem 8.1.5, there is an isometry ϕ of X such that $\mathrm{H} = \phi\Gamma\phi^{-1}$. Let R be a proper fundamental region for Γ. We now show that $\phi(R)$ is a proper fundamental region for H. First of all, $\phi(R)$ is an open set, since R is open. Let F be a fundamental set for Γ such that $R \subset F \subset \overline{R}$. As $\mathrm{H}\phi x = \phi\Gamma x$ for each x in X, we have that $\phi(F)$ is a fundamental set for H. Moreover

$$\phi(R) \subset \phi(F) \subset \overline{\phi(R)}.$$

Furthermore

$$\mathrm{Vol}(\partial(\phi(R))) = \mathrm{Vol}(\phi(\partial R)) = \mathrm{Vol}(\partial R) = 0.$$

Therefore $\phi(R)$ is a proper fundamental region for H by Theorem 6.5.13. Finally

$$\mathrm{Vol}(X/\Gamma) = \mathrm{Vol}(R) = \mathrm{Vol}(\phi(R)) = \mathrm{Vol}(X/\mathrm{H}).$$ $\square$

Definition: A Clifford-Klein space-form X/Γ is *orientable* if and only if every element of Γ is orientation preserving.

Spherical Space-Forms

It follows from Theorem 8.1.3 that every spherical space-form S^n/Γ is finitely covered by S^n. Hence, every spherical space form is a closed n-manifold with a finite fundamental group when $n > 1$.

Example 1. Clearly, the group $\{\pm 1\}$ acts freely on S^n. The space-form $S^n/\{\pm 1\}$ is elliptic n-space P^n.

Theorem 8.2.3. *Spherical n-space S^n and elliptic n-space P^n are the only spherical space-forms of even dimension n.*

Proof: Let $M = S^n/\Gamma$ be a space-form of even dimension n and let A be a nonidentity element of Γ. Then A is an odd dimensional orthogonal matrix. By Theorem 5.4.2, we deduce that ± 1 is an eigenvalue of A. Hence 1 is an eigenvalue of A^2. Therefore A^2 fixes a point of S^n. As Γ acts freely on S^n, we must have that $A^2 = I$. Consequently, all the rotation angles of A are π. Hence A is conjugate in $\mathrm{O}(n+1)$ to $-I$. As $-I$ commutes with every matrix in $\mathrm{O}(n+1)$, we have $A = -I$. Thus $M = P^n$. □

Theorem 8.2.4. *Every spherical space-form S^n/Γ of odd dimension n is orientable.*

Proof: Let $M = S^n/\Gamma$ be a space-form of odd dimension n and let A be a nonidentity element of Γ. Then A is an even dimensional orthogonal matrix. As Γ acts freely on S^n, the matrix A has no eigenvalue equal to 1. By Theorem 5.4.2, we deduce that A has an even number of eigenvalues equal to -1. Hence A is a rotation. Consequently, every element of Γ preserves an orientation of S^n and therefore M is orientable. □

Example 2. Identify S^3 with the unit sphere in $\mathbb{C}^2$ given by

$$\{(z, w) \in \mathbb{C}^2 : |z|^2 + |w|^2 = 1\}.$$

Let p and q be positive coprime integers. Then the matrix

$$\begin{pmatrix} e^{2\pi i/p} & 0 \\ 0 & e^{2\pi i q/p} \end{pmatrix}$$

is unitary and has order p. Let Γ be the finite cyclic subgroup of $\mathrm{U}(2)$ generated by this matrix. Then Γ acts freely on S^3 as a group of isometries. The space-form

$$L(p, q) = S^3/\Gamma$$

is called the *(p, q)-lens space*. It is known that two lens spaces $L(p, q)$ and $L(p', q')$ are homeomorphic if and only if $p = p'$ and either $q \equiv \pm q'$ (mod p) or $qq' \equiv \pm 1$ (mod p). In particular, $L(5, 1)$ and $L(5, 2)$ have isomorphic fundamental groups but are not homeomorphic. Thus, the homeomorphism type of a spherical space-form is not determined, in general, by the isomorphism type of its fundamental group.

Euclidean Space-Forms

Let E^n/Γ be a Euclidean space-form. Then Γ is a torsion-free discrete group of isometries of E^n. By the characterization of discrete Euclidean groups in §5.4, the group Γ is a finite extension of a finitely generated free abelian group of rank at most n.

Example 3. Let Γ be a lattice subgroup of $\mathrm{I}(E^n)$. Then Γ is a torsion-free discrete subgroup of $\mathrm{I}(E^n)$. The space-form E^n/Γ is called a *Euclidean n-torus.*

Theorem 8.2.5. *Every compact, n-dimensional, Euclidean space-form is finitely covered by a Euclidean n-torus.*

Proof: Let E^n/Γ be a compact Euclidean space-form. By Theorem 7.4.2, the subgroup T of translations of Γ is of finite index and of rank n; moreover, T is a normal subgroup of Γ. Now the action of Γ on E^n induces an action of Γ/T on E^n/T such that if g is in Γ and x is in E^n, then

$$(\mathrm{T}g)(\mathrm{T}x) = \mathrm{T}gx.$$

The group Γ/T acts as a group of isometries of E^n/T, since

$$\begin{aligned} d_{\mathrm{T}}(\mathrm{T}g\mathrm{T}x, \mathrm{T}g\mathrm{T}y) &= d_{\mathrm{T}}(\mathrm{T}gx, \mathrm{T}gy) \\ &= d_{\mathrm{T}}(g\mathrm{T}x, g\mathrm{T}y) \\ &= d_{\mathrm{T}}(\mathrm{T}x, \mathrm{T}y). \end{aligned}$$

Furthermore Γ/T acts discontinuously on E^n/T, since it is finite.

Next, we show that Γ/T acts freely on E^n/T. Suppose that

$$(\mathrm{T}g)(\mathrm{T}x) = \mathrm{T}x.$$

Then $\mathrm{T}gx = \mathrm{T}x$. Hence $gx = hx$ for some h in T. Therefore $h^{-1}gx = x$. As Γ acts freely on E^n, we have that $h^{-1}g = 1$. Therefore $g = h$, and so g is in T. Thus Γ/T acts freely on E^n/T.

By Theorem 8.1.3, the quotient map

$$\pi : E^n/\mathrm{T} \to (E^n/\mathrm{T})/(\Gamma/\mathrm{T})$$

is a covering projection. Clearly $(E^n/\mathrm{T})/(\Gamma/\mathrm{T})$ is isometric to E^n/Γ. Thus E^n/Γ is finitely covered by the Euclidean n-torus E^n/T. □

Corollary 1. *If E^n/Γ is a compact Euclidean space-form, then Γ is a torsion-free finite extension of a free abelian group of rank n.*

Example 4. Let τ_i be the translation of E^2 by e_i, for $i = 1, 2$, and let ρ be the reflection of E^2 in the line $y = 1/2$. Let Γ be the group generated by $\rho\tau_1$ and τ_2. Then Γ is a torsion-free discrete subgroup of $\mathrm{I}(E^2)$. The space-form E^2/Γ is a Klein bottle that is double covered by the Euclidean torus E^2/T, where T is generated by τ_1^2 and τ_2.

Two Euclidean space-forms E^n/Γ and E^n/H are said to be *affinely equivalent* if and only if there is a homeomorphism $\phi : E^n/\Gamma \to E^n/\mathrm{H}$ induced by an affine bijection of $\mathbb{R}^n$. By Theorem 7.4.4, two compact Euclidean space-forms have isomorphic fundamental groups if and only if they are affinely equivalent. Moreover, there are only finitely many isomorphism classes of n-dimensional crystallographic groups by Theorem 7.4.3. Therefore, there are only finitely many affine equivalence classes of n-dimensional, compact, Euclidean, space-forms. The exact number of affine equivalence classes of n-dimensional, compact, Euclidean, space-forms for $n = 1, 2, 3, 4$ is $1, 2, 10, 74$, respectively.

Hyperbolic Space-Forms

Our main goal is to understand the geometry and topology of hyperbolic space-forms. We begin by studying the elementary hyperbolic space-forms.

Definition: A hyperbolic space-form H^n/Γ is *elementary* if and only if Γ is an elementary subgroup of $\mathrm{I}(H^n)$.

The *type* of an elementary space-form H^n/Γ is defined to be the elementary type of Γ. By the characterization of elementary discrete subgroups of $\mathrm{I}(H^n)$ in §5.5, a space-form H^n/Γ is elementary if and only if Γ contains an abelian subgroup of finite index.

Let H^n/Γ be an elementary space-form. Assume first that Γ is of elliptic type. Then Γ is finite by Theorem 5.5.2, but Γ is torsion-free by Theorem 8.2.1. Therefore Γ is trivial. Thus, the only n-dimensional, elementary, hyperbolic space-form of elliptic type is H^n.

Next, assume that Γ is of parabolic type. We now pass to the upper half-space model and consider Γ to be a subgroup of $\mathrm{I}(U^n)$. By Theorem 8.1.5, we may assume that Γ fixes ∞. Then Γ corresponds under Poincaré extension to an infinite discrete subgroup of $\mathrm{I}(E^{n-1})$ by Theorem 5.5.5. As Γ acts trivially on the second factor of the cartesian product

$$U^n = E^{n-1} \times \mathbb{R}_+,$$

we deduce that U^n/Γ is homeomorphic to $(E^{n-1}/\Gamma) \times \mathbb{R}_+$. As Γ is torsion-free, E^{n-1}/Γ is a Euclidean space-form. The next theorem says that the similarity type of E^{n-1}/Γ is a complete isometric invariant of U^n/Γ.

Theorem 8.2.6. *Let U^n/Γ and U^n/H be two elementary space-forms of parabolic type such that both Γ and H fix ∞. Then U^n/Γ and U^n/H are isometric if and only if E^{n-1}/Γ and E^{n-1}/H are similar.*

Proof: By Theorem 8.1.5, the space-forms U^n/Γ and U^n/H are isometric if and only if Γ and H are conjugate in $\mathrm{I}(U^n)$. As Γ and H both fix ∞, they are conjugate in $\mathrm{I}(U^n)$ if and only if they are conjugate in the subgroup of

$\mathrm{I}(U^n)$ that fixes ∞. The group $\mathrm{S}(E^{n-1})$ of similarities of E^{n-1} corresponds under Poincaré extension to the subgroup of $\mathrm{I}(U^n)$ that fixes ∞. Thus Γ and H are conjugate in $\mathrm{I}(U^n)$ if and only if they are conjugate in $\mathrm{S}(E^{n-1})$. The same argument as in the proof of Theorem 8.1.5 shows that Γ and H are conjugate in $\mathrm{S}(E^{n-1})$ if and only if E^{n-1}/Γ and E^{n-1}/H are similar. Thus U^n/Γ and U^n/H are isometric if and only if E^{n-1}/Γ and E^{n-1}/H are similar. □

Now assume that Γ is of hyperbolic type. From the description of an elementary discrete group of hyperbolic type in §5.5, we have that Γ is an infinite cyclic group generated by a hyperbolic element of $\mathrm{I}(U^n)$. By Theorem 8.1.5, we may assume that Γ is generated by a Möbius transformation ϕ of U^n defined by $\phi(x) = kAx$ with $k > 1$ and A an orthogonal transformation of E^n that fixes the n-axis. A fundamental domain for Γ is the two-sided region

$$\{x \in U^n : 1 < x_n < k\}.$$

Let $\mathrm{K} = \{k^m : m \in \mathbb{Z}\}$. The two sides of the fundamental domain of Γ are paired by ϕ. Consequently U^n/Γ is a $(n-1)$-dimensional vector bundle over the circle $\mathbb{R}_+/\mathrm{K}$.

Next observe that the geodesic segment $[e_n, ke_n]$ in U^n projects to a simple closed curve ω in U^n/Γ, called the *fundamental cycle* of U^n/Γ. The *length* of ω is defined to be $\log k$, which is the hyperbolic length of $[e_n, ke_n]$. The *torsion angles* of U^n/Γ are defined to the angles of rotation of A.

Theorem 8.2.7. *Two elementary space-forms U^n/Γ_1 and U^n/Γ_2 of hyperbolic type are isometric if and only if they have the same fundamental cycle length and torsion angles.*

Proof: By Theorem 8.1.5, the space-forms U^n/Γ_1 and U^n/Γ_2 are isometric if and only if Γ_1 and Γ_2 are conjugate in $\mathrm{I}(U^n)$. Hence, we may assume that Γ_i is generated by a Möbius transformation ϕ_i of U^n, given by $\phi_i = k_iA_i$, with $k_i > 1$ and A_i an orthogonal transformation of E^n that fixes the n-axis for $i = 1, 2$.

Now suppose that Γ_1 and Γ_2 are conjugate in $\mathrm{I}(U^n)$. Then there is a Möbius transformation ψ of U^n such that $\phi_1 = \psi\phi_2^{\pm 1}\psi^{-1}$. As the fixed points of $\psi\phi_2^{\pm 1}\psi^{-1}$ are $\psi\{0, \infty\}$, we deduce that ψ leaves the set $\{0, \infty\}$ invariant. Assume first that ψ fixes both 0 and ∞. Then there is a $\ell > 0$ and B in $\mathrm{O}(n)$ that fixes e_n such that $\psi = \ell B$. This implies that

$$\psi\phi_2^{\pm 1}\psi^{-1} = B\phi_2^{\pm 1}B^{-1}.$$

Hence, we have

$$k_1A_1 = k_2^{\pm 1}BA_2^{\pm 1}B^{-1}.$$

As $k_1, k_2 > 1$, we have that $k_1 = k_2$ and $A_1 = BA_2B^{-1}$. Therefore U^n/Γ_1 and U^n/Γ_2 have the same fundamental cycle length and torsion angles.

Now assume that ψ switches 0 and ∞. Then we may assume, by the first case, that $\psi(x) = x/|x|^2$. Then $\psi\phi_2^{\pm 1}\psi^{-1} = k_2^{\mp 1}A_2^{\pm}$. Hence, we have that $k_1 A_1 = k_2^{\mp 1}A_2^{\pm}$. As $k_1, k_2 > 1$, we have that $k_1 = k_2$ and $A_1 = A_2^{-1}$. Therefore U^n/Γ_1 and U^n/Γ_2 have the same fundamental cycle length and torsion angles.

Conversely, suppose that U^n/Γ_1 and U^n/Γ_2 have the same fundamental cycle length and torsion angles. Then $k_1 = k_2$, and A_1 and A_2 are conjugate in $\mathrm{O}(n)$ by an orthogonal transformation that fixes e_n. Therefore ϕ_1 and ϕ_2 are conjugate in $\mathrm{I}(U^n)$. Thus Γ_1 and Γ_2 are conjugate in $\mathrm{I}(U^n)$ if and only if they have the same fundamental cycle length and torsion angles. □

Exercise 8.2

1. Show that $E^1/2\pi\mathbb{Z}$ is isometric to S^1.

2. Prove that the lens spaces $L(p,q)$ and $L(p,q')$ are isometric if and only if $q \equiv \pm q' \pmod{p}$ or $qq' \equiv \pm 1 \pmod{p}$.

3. Show that the volume of a spherical space-form S^n/Γ is given by the formula
$$\mathrm{Vol}(S^n/\Gamma) = \mathrm{Vol}(S^n)/|\Gamma|.$$

4. Show that the Klein bottle group Γ of Example 4 is a torsion-free discrete subgroup of $\mathrm{I}(E^2)$.

5. Let E^n/Γ be an noncompact Euclidean space-form such that Γ is nontrivial and the subgroup T of translations of Γ is of finite index in Γ. Prove that E^n/Γ is finitely covered by a Euclidean space-form isometric to $T^m \times E^{n-m}$, where T^m is a Euclidean m-torus with $0 < m < n$.

6. Let E^n/Γ and E^n/H be Euclidean n-tori with rectangular fundamental polyhedra P and Q, respectively. Prove that E^n/Γ and E^n/H are isometric if and only if P and Q are congruent in E^n.

7. Let E^n/Γ and E^n/H be Euclidean n-tori with rectangular fundamental polyhedra P and Q, respectively. Prove that E^n/Γ and E^n/H are similar if and only if P and Q are similar in E^n.

8. Prove that two Euclidean space-forms E^n/Γ and E^n/H are similar if and only if Γ and H are conjugate in $\mathrm{S}(E^n)$.

9. Let E^n/Γ and E^n/H be compact Euclidean space-forms and let $\mathrm{A}(\mathbb{R}^n)$ be the group of affine bijections of $\mathbb{R}^n$. Prove that the following are equivalent:

   (1) E^n/Γ and E^n/H are affinely equivalent;

   (2) Γ and H are conjugate in $\mathrm{A}(\mathbb{R}^n)$;

   (3) Γ and H are isomorphic.

10. Prove that every elementary hyperbolic space-form has infinite volume.

§8.3. (X, G)-Manifolds

Let G a group of similarities of an n-dimensional geometric space X and let M be an n-manifold. An (X, G)-*atlas* for M is defined to be a family of functions

$$\Phi = \{\phi_i : U_i \to X\}_{i \in \mathcal{I}},$$

called *charts*, satisfying the following conditions:

(1) The set U_i, called a *coordinate neighborhood*, is an open connected subset of M for each i.

(2) The chart ϕ_i maps the coordinate neighborhood U_i homeomorphically onto an open subset of X for each i.

(3) The coordinate neighborhoods $\{U_i\}_{i \in \mathcal{I}}$ cover M.

(4) If U_i and U_j overlap, then the function

$$\phi_j \phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j),$$

called a *coordinate change*, agrees in a neighborhood of each point of its domain with an element of G. See Figure 8.3.1.

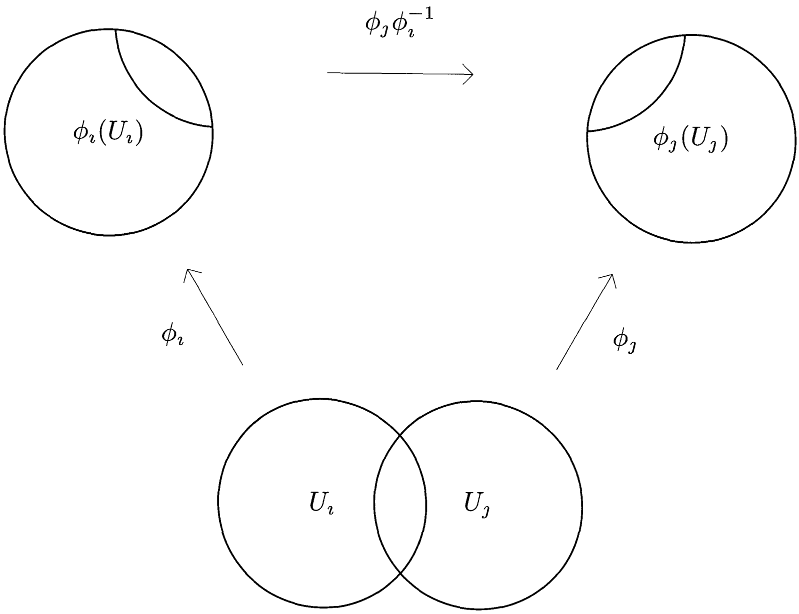


Figure 8.3.1. A coordinate change

Theorem 8.3.1. *Let Φ be an (X, G)-atlas for M. Then there is a unique maximal (X, G)-atlas for M containing Φ.*

Proof: Let $\Phi = \{\phi_i : U_i \to X\}$ and let $\overline{\Phi}$ be the set of all functions $\phi : U \to X$ such that

(1) the set U is an open connected subset of M;

(2) the function ϕ maps U homeomorphically onto an open subset of X;

(3) the function
$$\phi\phi_i^{-1} : \phi_i(U_i \cap U) \to \phi(U_i \cap U)$$
agrees in a neighborhood of each point of its domain with an element of G for each i.

Clearly $\overline{\Phi}$ contains Φ. Suppose that $\phi : U \to X$ and $\psi : V \to X$ are in $\overline{\Phi}$. Then for each i, we have that
$$\psi\phi^{-1} : \phi(U \cap V \cap U_i) \to \psi(U \cap V \cap U_i)$$
is the composite $\psi\phi_i^{-1}\phi_i\phi^{-1}$, and therefore it agrees in a neighborhood of each point of its domain with an element of G. As $\{U_i\}$ is an open cover of M, we have that $\psi\phi^{-1} : \phi(U \cap V) \to \psi(U \cap V)$ agrees in a neighborhood of each point of its domain with an element of G. Thus $\overline{\Phi}$ is an (X, G)-atlas for M. Clearly $\overline{\Phi}$ contains every (X, G)-atlas for M containing Φ, and so $\overline{\Phi}$ is the unique maximal (X, G)-atlas for M containing Φ. □

Definition: An (X, G)-*structure* for an n-manifold M is a maximal (X, G)-atlas for M.

Definition: An (X, G)-*manifold* M is an n-manifold M together with an (X, G)-structure for M.

Let M be an (X, G)-manifold. A *chart* for M is an element $\phi : U \to X$ of the (X, G)-structure of M. If u is a point of M, then a *chart* for (M, u) is a chart $\phi : U \to X$ for M such that u is in U.

Example 1. An $(S^n, \mathrm{I}(S^n))$-structure on a manifold is called a *spherical structure*, and an $(S^n, \mathrm{I}(S^n))$-manifold is called a *spherical n-manifold.*

Example 2. A $(E^n, \mathrm{I}(E^n))$-structure on a manifold is called a *Euclidean structure*, and a $(E^n, \mathrm{I}(E^n))$-manifold is called a *Euclidean n-manifold.*

Example 3. An $(H^n, \mathrm{I}(H^n))$-structure on a manifold is called a *hyperbolic structure*, and an $(H^n, \mathrm{I}(H^n))$-manifold is called a *hyperbolic n-manifold.*

Example 4. A $(E^n, \mathrm{S}(E^n))$-structure on a manifold is called a *Euclidean similarity structure*, and a $(E^n, \mathrm{S}(E^n))$-manifold is called a *Euclidean similarity n-manifold.*

X-Space-Forms

Let Γ be a discrete group of isometries of an n-dimensional geometric space X such that Γ acts freely on X. Then the quotient map $\pi : X \to X/\Gamma$ is a local isometry. Hence X/Γ is an n-manifold. For each x in X, choose $r(x) > 0$ so that π maps $B(x, r(x))$ isometrically onto $B(\pi(x), r(x))$. Let $U_x = B(\pi(x), r(x))$ and let $\phi_x : U_x \to X$ be the inverse of the restriction of π to $B(x, r(x))$. Then $\{U_x\}_{x \in X}$ is an open cover of X/Γ and ϕ_x maps U_x homeomorphically onto $B(x, r(x))$ for each x in X. Furthermore U_x is connected for each x in X, since $B(x, r(x))$ is connected.

Let x, y be points of X such that U_x and U_y overlap and consider the function

$$\phi_y\phi_x^{-1} : \phi_x(U_x \cap U_y) \to \phi_y(U_x \cap U_y).$$

Let z be an arbitrary point of $\phi_x(U_x \cap U_y)$ and set $w = \phi_y\phi_x^{-1}(z)$. Then $\pi(z) = \pi(w)$. Hence, there is a g in Γ such that $gz = w$. As g is continuous at z, there is an $\epsilon > 0$ such that $\phi_y(U_x \cap U_y)$ contains $gB(z, \epsilon)$. By shrinking $B(z, \epsilon)$, if necessary, we may assume that $\phi_x(U_x \cap U_y)$ contains $B(z, \epsilon)$. As $\pi g = \pi$, the map $\phi_y^{-1}g$ agrees with ϕ_x^{-1} on $B(z, \epsilon)$. Thus $\phi_y\phi_x^{-1}$ agrees with g on $B(z, \epsilon)$. This shows that $\{\phi_x : U_x \to X\}_{x \in X}$ is an (X, Γ)-atlas for X/Γ. By Theorem 8.3.1, this atlas determines an (X, Γ)-structure on X/Γ, called the *induced (X, Γ)-structure.* Thus X/Γ together with the induced (X, Γ)-structure is an (X, Γ)-manifold.

Let G be a subgroup of $\mathrm{S}(X)$ containing Γ. Clearly, an (X, Γ)-atlas for X/Γ is also an (X, G)-atlas for X/Γ; therefore, the induced (X, Γ)-structure on X/Γ determines an (X, G)-structure on X/Γ, called the *induced (X, G)-structure.* In particular, X/Γ, with the induced $(X, \mathrm{I}(X))$-structure, is an $(X, \mathrm{I}(X))$-manifold. Thus, every X-space-form is an $(X, \mathrm{I}(X))$-manifold.

Theorem 8.3.2. *Let X be a geodesically connected and geodesically complete metric space. If g and h are similarities of X that agree on a nonempty open subset of X, then $g = h$.*

Proof: The metric space X is rigid by Theorem 6.5.12. □

Theorem 8.3.3. *Let $\phi_j\phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$ be a coordinate change of an (X, G)-manifold M. Then $\phi_j\phi_i^{-1}$ agrees with an element of G on each connected component of its domain.*

Proof: Let C be a connected component of $\phi_i(U_i \cap U_j)$. Suppose that w and x are in C. Then there are open subsets $W_1, \ldots, W_m$ of C such that w is in W_1, the sets W_k and W_{k+1} overlap for $k = 1, \ldots, m-1$, the set W_m contains x, and $\phi_j\phi_i^{-1}$ agrees with an element g_k of G on W_k. As g_k and g_{k+1} agree on the nonempty open set $W_k \cap W_{k+1}$, we have that $g_k = g_{k+1}$ by Theorem 8.3.2. Therefore, all the g_k are the same. Thus $\phi_j\phi_i^{-1}$ agrees with g_1 at x and therefore on C. □

Metric (X, G)-Manifolds

Definition: A *metric* (X, G)*-manifold* is a connected (X, G)-manifold M such that G is a group of isometries of X.

Let $\gamma : [a, b] \to M$ be a curve in a metric (X, G)-manifold M. We now define the X*-length* of γ. Assume first that $\gamma([a, b])$ is contained in a coordinate neighborhood U. Let $\phi : U \to X$ be a chart for M. The X*-length* of γ is defined to be

$$\|\gamma\| = |\phi\gamma|.$$

The X-length of γ does not depend on the choice of the chart ϕ, since if $\psi : V \to X$ is another chart for M such that V contains $\gamma([a, b])$, then there is an isometry g in G that agrees with $\psi\phi^{-1}$ on $\phi\gamma([a, b])$ by Theorem 8.3.3 and therefore

$$|\phi\gamma| = |g\phi\gamma| = |\psi\phi^{-1}\phi\gamma| = |\psi\gamma|.$$

Now assume that $\gamma : [a, b] \to M$ is an arbitrary curve. As $\gamma([a, b])$ is compact, there is a partition

$$a = t_0 < t_1 < \cdots < t_m = b$$

of $[a, b]$ such that $\gamma([t_{i-1}, t_i])$ is contained in a coordinate neighborhood U_i for each $i = 1, \ldots, m$. Let γ_{t_{i-1},t_i} be the restriction of γ to $[t_{i-1}, t_i]$. The X*-length* of γ is defined to be

$$\|\gamma\| = \sum_{i=1}^{m} \|\gamma_{t_{i-1},t_i}\|.$$

The X-length of γ does not depend on the choice of the partition $\{t_i\}$, since if

$$a = s_0 < s_1 < \cdots < s_\ell = b$$

is another partition such that $\gamma([s_{i-1}, s_i])$ is contained in a coordinate neighborhood V_i, then there is a third partition

$$a = r_0 < r_1 < \cdots < r_k = b$$

such that $\{r_i\} = \{s_i\} \cup \{t_i\}$, and therefore

$$\sum_{i=1}^{m} \|\gamma_{t_{i-1},t_i}\| = \sum_{i=1}^{k} \|\gamma_{r_{i-1},r_i}\| = \sum_{i=1}^{\ell} \|\gamma_{s_{i-1},s_i}\|.$$

Definition: A curve γ in a metric (X, G)-manifold M is X*-rectifiable* if and only if $\|\gamma\| < \infty$.

Lemma 1. *Any two points in a metric (X, G)-manifold M can be joined by an X-rectifiable curve in M.*

Proof: Define a relation on M by $u \sim v$ if and only if u and v are joined by an X-rectifiable curve in M. It is easy to see that this is an equivalence relation on M. Let $[u]$ be an equivalence class and suppose that v is in $[u]$. Let $\psi : V \to X$ be a chart for (M, v). Then there is an $r > 0$ such that $\psi(V)$ contains $B(\psi(v), r)$. Let x be an arbitrary point in $B(\psi(v), r)$. As X is geodesically connected, there is a geodesic arc $\alpha : [a, b] \to X$ from $\psi(v)$ to x. Clearly $B(\psi(v), r)$ contains $\alpha([a, b])$. Hence $\psi^{-1}\alpha : [a, b] \to M$ is an X-rectifiable curve from v to $\psi^{-1}(x)$. This shows that $[u]$ contains the open set $\psi^{-1}(B(\psi(v), r))$. Thus $[u]$ is open in M. As M is connected, $[u]$ must be all of M. Thus, any two points of M can be joined by an X-rectifiable curve. □

Theorem 8.3.4. *Let M be a metric (X, G)-manifold. Then the function $d : M \times M \to \mathbb{R}$, defined by*

$$d(u, v) = \inf_{\gamma} \|\gamma\|,$$

where γ varies over all X-rectifiable curves from u to v, is a metric on M.

Proof: By Lemma 1, the function d is well defined. Clearly d is nonnegative and $d(u, u) = 0$ for all u in M. To see that d is nondegenerate, let u, v be distinct points of M. Since M is Hausdorff, there is a chart $\phi : U \to X$ for (M, u) such that v is not in U. Choose $r > 0$ such that $\phi(U)$ contains $C(\phi(u), r)$. By Theorem 8.1.2, the sphere

$$S(\phi(u), r) = \{x \in X : d(\phi(u), x) = r\}$$

is compact. Hence, the set

$$T = \phi^{-1}(S(\phi(u), r))$$

is closed in M, since M is Hausdorff.

Let $\gamma : [a, b] \to M$ be an arbitrary X-rectifiable curve from u to v. Since $\gamma([a, b])$ is connected and contains both u and v, it must meet T. Hence, there is a first point c in the open interval (a, b) such that $\gamma(c)$ is in T. Let $\gamma_{a,c}$ be the restriction of γ to $[a, c]$. Then the image of $\gamma_{a,c}$ is contained in $\phi^{-1}(C(\phi(u), r))$. Consequently, we have

$$\|\gamma\| \geq \|\gamma_{a,c}\| = |\phi\gamma_{a,c}| \geq d_X(\phi(u), \phi\gamma(c)) = r.$$

Therefore, we have

$$d(u, v) \geq r > 0.$$

Thus d is nondegenerate.

If $\gamma : [a, b] \to M$ is an X-rectifiable curve from u to v, then

$$\gamma^{-1} : [a, b] \to M$$

is an X-rectifiable curve from v to u, and $\|\gamma^{-1}\| = \|\gamma\|$. Consequently d is symmetric.

If $\alpha : [a, b] \to M$ is an X-rectifiable curve from u to v, and $\beta : [b, c] \to M$ is an X-rectifiable curve from v to w, then $\alpha\beta : [a, c] \to M$ is an X-rectifiable curve from u to w, and

$$\|\alpha\beta\| = \|\alpha\| + \|\beta\|.$$

This implies the triangle inequality

$$d(u, w) \leq d(u, v) + d(v, w).$$

Thus d is a metric on M. □

Let M be a metric (X, G)-manifold. Then the metric d in Theorem 8.3.4 is called the *induced metric* on M. Henceforth, we shall assume that a metric (X, G)-manifold is a metric space with the induced metric.

Theorem 8.3.5. *Let $\phi : U \to X$ be a chart for a metric (X, G)-manifold M, let x be a point of $\phi(U)$, and let $r > 0$ be such that $\phi(U)$ contains $B(x, r)$. Then ϕ^{-1} maps $B(x, r)$ homeomorphically onto $B(\phi^{-1}(x), r)$.*

Proof: Clearly ϕ^{-1} maps $B(x, r)$ into $B(\phi^{-1}(x), r)$. Let v be an arbitrary point of $B(\phi^{-1}(x), r)$. Then there is an X-rectifiable curve $\gamma : [a, b] \to M$ from $\phi^{-1}(x)$ to v such that $\|\gamma\| < r$. Suppose that v is not in $\phi^{-1}(B(x, r))$. We shall derive a contradiction. Let $s = (\|\gamma\| + r)/2$. Since $\gamma([a, b])$ is connected and contains both $\phi^{-1}(x)$ and v, it must meet $\phi^{-1}(S(x, s))$. Hence, there is a first point c in (a, b) such that $\gamma(c)$ is in $\phi^{-1}(S(x, s))$. Let $\gamma_{a,c}$ be the restriction of γ to $[a, c]$. Then the image of $\gamma_{a,c}$ is contained in $\phi^{-1}(C(x, s))$. Consequently

$$\|\gamma\| \geq \|\gamma_{a,c}\| = |\phi\gamma_{a,c}| \geq s,$$

which is a contradiction. Thus ϕ^{-1} maps $B(x, r)$ onto $B(\phi^{-1}(x), r)$. □

Corollary 1. *If M is a metric (X, G)-manifold, then the topology of M is the metric topology determined by the induced metric.*

Theorem 8.3.6. *Let $\phi : U \to X$ be a chart for a metric (X, G)-manifold M, let x be a point of $\phi(U)$, and let $r > 0$ be such that $\phi(U)$ contains $B(x, r)$. Then ϕ^{-1}maps $B(x, r/2)$ isometrically onto $B(\phi^{-1}(x), r/2)$; therefore ϕ is a local isometry.*

Proof: By Theorem 8.3.5, the function ϕ^{-1} maps $B(x, r/2)$ bijectively onto $B(\phi^{-1}(x), r/2)$. Hence, we only need to show that ϕ^{-1} preserves distances on $B(x, r/2)$. Let y, z be distinct points of $B(x, r/2)$. As X is geodesically connected, there is a geodesic arc $\alpha : [0, \ell] \to X$ from y to z. By the triangle inequality, $d_X(y, z) < r$. Hence, every point in $\alpha([0, \ell])$ is at most a distance $r/2$ from either y or z. Therefore $B(x, r)$ contains $\alpha([0, \ell])$. Hence

$$d(\phi^{-1}(y), \phi^{-1}(z)) \leq \|\phi^{-1}\alpha\| = |\alpha| = d_X(y, z).$$

Now let $\gamma : [a, b] \to M$ be any X-rectifiable curve from $\phi^{-1}(y)$ to $\phi^{-1}(z)$. Assume first that U contains $\gamma([a, b])$. Then

$$\|\gamma\| = |\phi\gamma| \geq d_X(y, z).$$

Now assume that U does not contain $\gamma([a, b])$. Set

$$s = \max\{d_X(x, y), d_X(x, z)\} + (r/2).$$

Then $s < r$. Hence, there is a first point c in (a, b) such that $\gamma(c)$ is in $\phi^{-1}(S(x, s))$, and there is a last point d in (a, b) such that $\gamma(d)$ is in $\phi^{-1}(S(x, s))$. Let $\gamma_{a,c}$ be the restriction of γ to $[a, c]$ and let $\gamma_{d,b}$ be the restriction of γ to $[d, b]$. Then

$$\begin{aligned}
\|\gamma\| &\geq \|\gamma_{a,c}\| + \|\gamma_{d,b}\| \\
&= |\phi\gamma_{a,c}| + |\phi\gamma_{d,b}| \\
&\geq d_X(y, \phi\gamma(c)) + d_X(\phi\gamma(d), z) \\
&\geq r/2 + r/2 \\
&> d_X(y, z).
\end{aligned}$$

Thus, in general, we have

$$\|\gamma\| \geq d_X(y, z).$$

Hence, we have

$$d(\phi^{-1}(y), \phi^{-1}(z)) \geq d_X(y, z).$$

Since we have already established the reverse inequality, we have that ϕ^{-1} maps $B(x, r/2)$ isometrically onto $B(\phi^{-1}(x), r/2)$. □

Example: The unit circle S^1 in $\mathbb{C}$ is a Euclidean 1-manifold. The complex argument mapping

$$\arg : S^1 - \{-1\} \to \mathbb{R}$$

is a chart for S^1 whose image is the open interval $(-\pi, \pi)$. Observe that $(-\pi/2, \pi/2)$ is the largest open interval centered at the origin that is mapped isometrically onto its image by $\arg^{-1}$. This example shows why the radius r is halved in Theorem 8.3.6.

Exercise 8.3

1. Prove Corollary 1.
2. Let $\gamma : [a, b] \to M$ be a curve in a metric (X, G)-manifold. Prove that the X-length of γ is the same as the length of γ with respect to the induced metric.
3. Let X/Γ be an X-space-form. Show that the induced metric on X/Γ is the orbit space metric d_Γ.
4. Prove that every metric (X, G)-manifold is locally geodesically convex.
5. Prove that any two points of a metric (X, G)-manifold M can be joined by a piecewise geodesic curve in M.

§8.4. Developing

Let $\phi : U \to X$ be a chart for an (X, G)-manifold M and let $\gamma : [a, b] \to M$ be a curve whose initial point $\gamma(a)$ is in U. Then there is a partition

$$a = t_0 < t_1 < \cdots < t_m = b$$

and a set $\{\phi_i : U_i \to X\}_{i=1}^m$ of charts for M such that $\phi_1 = \phi$ and U_i contains $\gamma([t_{i-1}, t_i])$ for each $i = 1, \ldots, m$. Let g_i be the element of G that agrees with $\phi_i\phi_{i+1}^{-1}$ on the connected component of $U_i \cap U_{i+1}$ containing $\gamma(t_i)$. Let γ_i be the restriction of γ to the interval $[t_{i-1}, t_i]$. Then $\phi_i\gamma_i$ and $g_i\phi_{i+1}\gamma_{i+1}$ are curves in X and

$$g_i\phi_{i+1}\gamma(t_i) = \phi_i\phi_{i+1}^{-1}\phi_{i+1}\gamma(t_i) = \phi_i\gamma(t_i).$$

Thus $g_i\phi_{i+1}\gamma_{i+1}$ begins where $\phi_i\gamma_i$ ends, and so we can define a curve $\hat{\gamma} : [a, b] \to X$ by the formula

$$\hat{\gamma} = (\phi_1\gamma_1)(g_1\phi_2\gamma_2)(g_1g_2\phi_3\gamma_3)\cdots(g_1\cdots g_{m-1}\phi_m\gamma_m).$$

We claim that $\hat{\gamma}$ does not depend on the choice of the charts $\{\phi_i\}$ once a partition of $[a, b]$ has been fixed. Suppose that $\{\psi_i : V_i \to X\}_{i=1}^m$ is another set of charts for M such that $\psi_1 = \phi$ and V_i contains $\gamma([t_{i-1}, t_i])$ for each $i = 1, \ldots, m$. Let h_i be the element of G that agrees with $\psi_i\psi_{i+1}^{-1}$ on the component of $V_i \cap V_{i+1}$ containing $\gamma(t_i)$. As $U_i \cap V_i$ contains $\gamma([t_{i-1}, t_i])$, it is enough to show that

$$g_1 \cdots g_{i-1}\phi_i = h_1 \cdots h_{i-1}\psi_i$$

on the component of $U_i \cap V_i$ containing $\gamma([t_{i-1}, t_i])$ for each i. This is true by hypothesis for $i = 1$. We proceed by induction. Suppose that it is true for $i - 1$. Let f_i be the element of G that agrees with $\psi_i\phi_i^{-1}$ on the component of $U_i \cap V_i$ containing $\gamma([t_{i-1}t_i])$. On the one hand, f_i agrees with

$$\psi_i(\psi_{i-1}^{-1}h_{i-2}^{-1}\cdots h_1^{-1})(g_1\cdots g_{i-2}\phi_{i-1})\phi_i^{-1}$$

on the component of $\phi_i(U_{i-1} \cap V_{i-1} \cap U_i \cap V_i)$ containing $\gamma(t_{i-1})$. On the other hand, $(h_{i-1}^{-1}\cdots h_1^{-1})(g_1\cdots g_{i-1})$ agrees with

$$(\psi_i\psi_{i-1}^{-1})(h_{i-2}^{-1}\cdots h_1^{-1})(g_1\cdots g_{i-2})(\phi_{i-1}\phi_i^{-1})$$

on the component of $\phi_i(U_{i-1} \cap V_{i-1} \cap U_i \cap V_i)$ containing $\gamma(t_{i-1})$. Therefore

$$f_i = (h_{i-1}^{-1}\cdots h_1^{-1})(g_1\cdots g_{i-1})$$

by Theorem 8.3.2. Hence

$$\begin{aligned}
(g_1\cdots g_{i-1})\phi_i &= (h_1\cdots h_{i-1})(h_{i-1}^{-1}\cdots h_1^{-1})(g_1\cdots g_{i-1})\phi_i \\
&= (h_1\cdots h_{i-1})f_i\phi_i \\
&= (h_1\cdots h_{i-1})\psi_i
\end{aligned}$$

on the component of $U_i \cap V_i$ containing $\gamma([t_{i-1}, t_i])$. This completes the induction.

Next, we show that $\hat{\gamma}$ does not depend on the partition of $[a,b]$. Let $\{s_i\}_{i=1}^{\ell}$ be another partition with charts $\{\psi_i : V_i \to X\}_{i=1}^{\ell}$. Then $\{r_i\} = \{s_i\} \cup \{t_i\}$ is a partition of $[a,b]$ containing both partitions. Since the charts $\{\phi_i\}$ and $\{\psi_i\}$ can both be used in turn for the partition $\{r_i\}$, we deduce that all three partitions determine the same curve $\hat{\gamma}$. The curve $\hat{\gamma} : [a,b] \to X$ is called the *continuation* of $\phi\gamma_1$ along γ.

Theorem 8.4.1. *Let $\phi : U \to X$ be a chart for an (X,G)-manifold M, let $\alpha, \beta : [a,b] \to M$ be curves with the same initial point in U and the same terminal point in M, and let $\hat{\alpha}, \hat{\beta}$ be the continuations of $\phi\alpha_1, \phi\beta_1$ along α, β, respectively. If α and β are homotopic by a homotopy that keeps their endpoints fixed, then $\hat{\alpha}$ and $\hat{\beta}$ have the same endpoints, and they are homotopic by a homotopy that keeps their endpoints fixed.*

Proof: This is clear if α and β differ only along a subinterval (c,d) such that $\alpha([c,d])$ and $\beta([c,d])$ are contained in a simply connected coordinate neighborhood U. In the general case, let $H : [a,b]^2 \to M$ be a homotopy from α to β that keeps the endpoints fixed. As $[a,b]$ is compact, there is a partition $a = t_0 < t_1 < \cdots < t_m = b$ such that $H([t_{i-1}, t_i] \times [t_{j-1}, t_j])$ is contained in a simply connected coordinate neighborhood U_{ij} for each $i, j = 1, \ldots, m$. Let α_{ij} be the curve in M defined by applying H to the curve in $[a,b]^2$ illustrated in Figure 8.4.1(a), and let β_{ij} be the curve in M defined by applying H to the curve in $[a,b]^2$ illustrated in Figure 8.4.1(b). Then by the first remark, $\hat{\alpha}_{ij}$ and $\hat{\beta}_{ij}$ have the same endpoints and are homotopic by a homotopy keeping their endpoints fixed. By composing all these homotopies starting at the lower right-hand corner of $[a,b]^2$, proceeding right to left along each row of rectangles $[t_{i-1}, t_i] \times [t_{j-1}, t_j]$, and ending at the top left-hand corner of $[a,b]^2$, we find that $\hat{\alpha}$ and $\hat{\beta}$ are homotopic by a homotopy keeping their endpoints fixed. $\square$

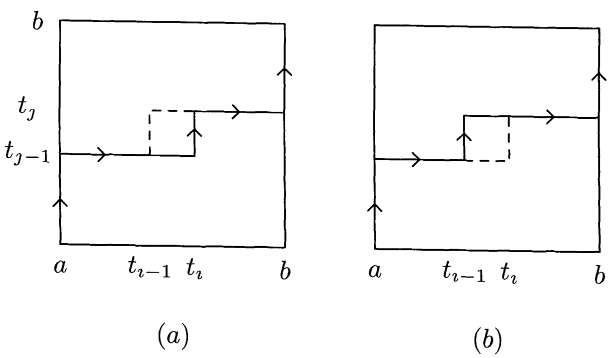


Figure 8.4.1. Alternate routes from (a,a) to (b,b) in the square $[a,b]^2$

(X, G)-Maps

Definition: A function $\xi : M \to N$ between (X, G)-manifolds is an (X, G)-*map* if and only if ξ is continuous and for each chart $\phi : U \to X$ for M and chart $\psi : V \to X$ for N such that U and $\xi^{-1}(V)$ overlap, the function

$$\psi\xi\phi^{-1} : \phi(U \cap \xi^{-1}(V)) \to \psi(\xi(U) \cap V)$$

agrees in a neighborhood of each point of its domain with an element of G.

Theorem 8.4.2. *A function $\xi : M \to N$ between (X, G)-manifolds is an (X, G)-map if and only if for each point u of M, there is a chart $\phi : U \to X$ for (M, u) such that ξ maps U homeomorphically onto an open subset of N and $\phi\xi^{-1} : \xi(U) \to X$ is a chart for N.*

Proof: Suppose that $\xi : M \to N$ is an (X, G)-map and u is an arbitrary point of M. Let $\psi : V \to X$ be a chart for $(N, \xi(u))$. Since ξ is continuous at u, there is a chart $\phi : U \to X$ for (M, u) such that $\xi(U) \subset V$. Then

$$\psi\xi\phi^{-1} : \phi(U) \to \psi\xi(U)$$

agrees with an element g of G, since $\phi(U)$ is connected. Hence ξ maps U homeomorphically onto an open subset of N, and $\phi\xi^{-1} : \xi(U) \to X$ agrees with $g^{-1}\psi : V \to X$. Therefore $\phi\xi^{-1}$ is a chart for N.

Conversely, suppose that for each point u of M, there is a chart $\phi : U \to X$ for (M, u) such that ξ maps U homeomorphically onto an open subset of N, and $\phi\xi^{-1} : \xi(U) \to X$ is a chart for N. Then ξ is continuous. Let $\chi : W \to X$ and $\psi : V \to X$ be charts for M and N, respectively, such that W and $\xi^{-1}(V)$ overlap, and let u be an arbitrary point of the set $W \cap \xi^{-1}(V)$. Then there is a chart $\phi : U \to X$ for (M, u) such that ξ maps U homeomorphically onto an open subset of N and $\phi\xi^{-1} : \xi(U) \to X$ is a chart for N. Observe that in a neighborhood of u, the function

$$\psi\xi\chi^{-1} : \chi(W \cap \xi^{-1}(V)) \to \psi(\xi(W) \cap V)$$

agrees with $(\psi\xi\phi^{-1})(\phi\chi^{-1})$. As $\phi\chi^{-1}$ and $\psi\xi\phi^{-1}$ are coordinate changes for M and N, respectively, $\psi\xi\chi^{-1}$ agrees in a neighborhood of u with an element of G. Thus ξ is an (X, G)-map. □

Theorem 8.4.3. *Let $\phi : U \to X$ be a chart for a simply connected (X, G)-manifold M. Then there is a unique (X, G)-map $\hat{\phi} : M \to X$ extending the chart ϕ.*

Proof: Fix a point u in U and let v be an arbitrary point of M. Then there is a curve $\alpha : [a, b] \to M$ from u to v. Let $\hat{\alpha} : [a, b] \to X$ be the continuation of $\phi\alpha_1$ along α. Then $\hat{\alpha}(b)$ does not depend on the choice of α by Theorem 8.4.1, since M is simply connected. Hence, we may define a function $\hat{\phi} : M \to X$ by $\hat{\phi}(v) = \hat{\alpha}(b)$.

Let $\psi : V \to X$ be a chart for (M, v) such that $\psi = \phi$ if v is in U. Then there is a partition

$$a = t_0 < t_1 < \cdots < t_m = b$$

and a set of charts $\{\phi_i : U_i \to X\}_{i=1}^m$ for M such that $\phi_1 = \phi$, and U_i contains $\alpha([t_{i-1}, t_i])$ for each $i = 1, \ldots, m$, and $\phi_m = \psi$. Let α_i be the restriction of α to $[t_{i-1}, t_i]$ and let g_i be the element of G that agrees with $\phi_i\phi_{i+1}^{-1}$ on the connected component of $U_i \cap U_{i+1}$ containing $\alpha(t_i)$. Then

$$\hat{\alpha} = (\phi_1\alpha_1)(g_1\phi_2\alpha_2)\cdots(g_1 \cdots g_{m-1}\phi_m\alpha_m).$$

Let $\beta : [b, c] \to V$ be a curve from v to w and let $g = g_1 \cdots g_{m-1}$. Then $\widehat{\alpha\beta} = \hat{\alpha} g \psi \beta$. Hence $\hat{\phi}(w) = \widehat{\alpha\beta}(c) = g\psi(w)$. Therefore $\hat{\phi}(w) = g\psi(w)$ for all w in V. Hence $\hat{\phi}$ maps V homeomorphically onto the open subset $g\psi(V)$ of X and $\psi\hat{\phi}^{-1} : \hat{\phi}(V) \to X$ is the restriction of g^{-1}. Thus $\hat{\phi}$ is an (X, G)-map by Theorem 8.4.2; moreover, $\hat{\phi}$ extends ϕ.

Now let $\xi : M \to X$ be any (X, G)-map extending ϕ. Without loss of generality, we may assume that the set of charts $\{\phi_i : U_i \to X\}_{i=1}^m$ for M has the property that

$$\phi_i\xi^{-1} : \xi(U_i) \to X$$

is a chart for X. Then $\phi_i\xi^{-1}$ extends to an element h_i^{-1} of G. Hence $\xi(w) = h_i\phi_i(w)$ for all w in U_i. As $\xi(w) = \phi(w)$ for all w in U, we have that $h_1\phi = \phi$ and so $h_1 = 1$. We proceed by induction. Suppose that $h_{i-1} = g_1 \cdots g_{i-2}$. Then for each w in U_{i-1}, we have

$$\begin{aligned} \xi(w) &= h_{i-1}\phi_{i-1}(w) \\ &= g_1 \cdots g_{i-2}\phi_{i-1}(w) \\ &= \hat{\phi}(w). \end{aligned}$$

Hence

$$h_i\phi_i(w) = \xi(w) = \hat{\phi}(w) = g_1 \cdots g_{i-1}\phi_i(w)$$

for all w in $U_{i-1} \cap U_i$. Therefore $h_i = g_1 \cdots g_{i-1}$. Hence, by induction, we have that

$$\xi(v) = h_m\phi_m(v) = g\phi_m(v) = \hat{\phi}(v).$$

Therefore $\xi = \hat{\phi}$. Thus $\hat{\phi}$ is unique. □

Theorem 8.4.4. *Let M be a simply connected (X, G)-manifold. If $\xi_1, \xi_2 : M \to X$ are (X, G)-maps, then there is a unique element g of G such that $\xi_2 = g\xi_1$.*

Proof: Let $\phi : U \to X$ be a chart for M such that $\phi\xi_i^{-1} : \xi_i(U) \to X$ is a chart for X for $i = 1, 2$. By Theorem 8.3.3, there is an element g_i of G extending $\phi\xi_i^{-1} : \xi_i(U) \to X$. As $g_i\xi_i$ is an (X, G)-map extending ϕ for $i = 1, 2$, we have that $g_1\xi_1 = g_2\xi_2$ by the uniqueness of $\hat{\phi}$. Let $g = g_2^{-1}g_1$. Then $\xi_2 = g\xi_1$. If h is another element of G such that $\xi_2 = h\xi_1$, then $g\xi_1 = h\xi_1$ whence $g = h$ by Theorem 8.3.2. Thus g is unique. □

The Developing Map

Let M be a connected (X, G)-manifold and let $\kappa : \tilde{M} \to M$ be a universal covering projection. Then $\tilde{M}$ is simply connected. Let $\{\phi_i : U_i \to X\}$ be an (X, G)-atlas for M such that U_i is simply connected for each i. Then the set U_i is evenly covered by κ for each i. Let $\{U_{ij}\}$ by the set of sheets over U_i and let $\kappa_{ij} : U_{ij} \to U_i$ be the restriction of κ. Define $\phi_{ij} : U_{ij} \to X$ by $\phi_{ij} = \phi_i \kappa_{ij}$. Then ϕ_{ij} maps U_{ij} homeomorphically onto the open set $\phi(U_i)$ in X. Suppose that U_{ij} and $U_{k\ell}$ overlap. Then U_i and U_k overlap. Consider the function

$$\phi_{ij}\phi_{k\ell}^{-1} : \phi_{k\ell}(U_{ij} \cap U_{k\ell}) \to \phi_{ij}(U_{ij} \cap U_{k\ell}).$$

If x is in $\phi_{k\ell}(U_{ij} \cap U_{k\ell})$, then

$$\phi_{ij}\phi_{k\ell}^{-1}(x) = \phi_i \kappa_{ij} \kappa_{k\ell}^{-1}(x) \phi_k^{-1} = \phi_i \phi_k^{-1}(x).$$

Hence $\phi_{ij}\phi_{k\ell}^{-1}$ agrees in a neighborhood of each point of its domain with an element of G. Therefore $\{\phi_{ij} : U_{ij} \to X\}$ is an (X, G)-atlas for $\tilde{M}$. We shall assume that $\tilde{M}$ is an (X, G)-manifold with the (X, G)-structure determined by this (X, G)-atlas.

Observe that κ maps the coordinate neighborhood U_{ij} homeomorphically onto U_i, and $\phi_{ij}\kappa^{-1} : \kappa(U_{ij}) \to X$ is the chart $\phi_i : U_i \to X$ for M. Thus κ is an (X, G)-map by Theorem 8.4.2.

Let $\tau : \tilde{M} \to \tilde{M}$ be a covering transformation of κ and let $\tilde{u}$ be an arbitrary point of $\tilde{M}$. Then there is an i such that $\kappa(\tilde{u})$ is in U_i. Hence, there is a j such that $\tilde{u}$ is in U_{ij}. As τ permutes the sheets over U_i, there is a k such that $\tau(U_{ij}) = U_{ik}$. Observe that $\phi_{ij}\tau^{-1} : \tau(U_{ij}) \to X$ is the chart $\phi_{ik} : U_{ik} \to X$. Therefore τ is an (X, G)-map.

Let $\phi : U \to X$ be a chart for $\tilde{M}$. Then ϕ extends to a unique (X, G)-map $\delta : \tilde{M} \to X$ by Theorem 8.4.3. The map

$$\delta : \tilde{M} \to X$$

is called the *developing map* for M determined by the chart ϕ. By Theorem 8.4.4, any two developing maps for M differ only by composition with an element of G. Thus, the developing map δ is unique up to composition with an element of G.

Holonomy

Choose a base point u of M and a base point $\tilde{u}$ of $\tilde{M}$ such that $\kappa(\tilde{u}) = u$. Let $\alpha : [0, 1] \to M$ be a loop based at u. Then α lifts to a unique curve $\tilde{\alpha}$ in $\tilde{M}$ starting at $\tilde{u}$. Let $\tilde{v}$ be the endpoint of $\tilde{\alpha}$. Then there is a unique covering transformation τ_α of κ such that $\tau_\alpha(\tilde{u}) = \tilde{v}$. The covering transformation τ_α depends only on the homotopy class of α in the fundamental group $\pi_1(M, u)$ by Theorem 8.4.1. Let $\beta : [0, 1] \to M$ be another loop based at u. Then $\widetilde{\alpha\beta} = (\tilde{\alpha})(\tau_\alpha\tilde{\beta})$ and so $\tau_{\alpha\beta} = \tau_\alpha\tau_\beta$.

Let $\delta : \tilde{M} \to X$ be a developing map for M. As $\delta\tau_\alpha : \tilde{M} \to X$ is an (X,G)-map, there is a unique element g_α of G such that $\delta\tau_\alpha = g_\alpha\delta$. Define

$$\eta : \pi_1(M,u) \to G$$

by $\eta([\alpha]) = g_\alpha$. Then η is well defined, since g_α depends only on the homotopy class of α. Observe that

$$\delta\tau_{\alpha\beta} = \delta\tau_\alpha\tau_\beta = g_\alpha\delta\tau_\beta = g_\alpha g_\beta\delta.$$

Hence

$$\eta([\alpha][\beta]) = \eta([\alpha\beta]) = g_\alpha g_\beta = \eta([\alpha])\eta([\beta]).$$

Thus η is a homomorphism. The homomorphism $\eta : \pi_1(M) \to G$ is called the *holonomy* of M determined by the developing map δ.

Note, if $\delta' : \tilde{M} \to X$ is another developing map for M, then there is a g in G such that $\delta' = g\delta$, and therefore

$$\delta'\tau_\alpha = g\delta\tau_\alpha = gg_\alpha\delta = gg_\alpha g^{-1}\delta'.$$

Hence, the holonomy η' of M determined by δ' differs from the holonomy of M determined by δ by conjugation by g.

Theorem 8.4.5. *Let M be a connected (X,G)-manifold and let H be a subgroup of G. Then the (X,G)-structure of M contains an (X,H)-structure for M if and only if H contains the image of a holonomy $\eta : \pi_1(M) \to G$ for M.*

Proof: Suppose that the (X,G)-structure of M contains an (X,H)-structure. Then H contains the image of any holonomy for M defined in terms of the (X,H)-structure for M. Conversely, suppose that H contains the image of a holonomy $\eta : \pi_1(M) \to G$ for M. Let $\delta : \tilde{M} \to X$ be the developing map that determines η, and let $\{\phi_i : U_i \to X\}$ be an (X,G)-atlas for M such that U_i is evenly covered by the covering projection $\kappa : \tilde{M} \to M$ for each i. Let $\{U_{ij}\}$ be the set of sheets over U_i and let $\kappa_{ij} : U_{ij} \to U_i$ be the restriction of κ. Define $\phi_{ij} : U_{ij} \to X$ by $\phi_{ij} = \phi_i\kappa_{ij}$. Then $\{\phi_{ij} : U_{ij} \to X\}$ is an (X,G)-atlas for $\tilde{M}$. Hence δ maps U_{ij} homeomorphically onto an open subset of X for each i and j.

For each i, choose a sheet U_{ij} over U_i and define $\psi_i : U_i \to X$ by setting $\psi_i = \delta\kappa_{ij}^{-1}$. Then ψ_i maps U_i homeomorphically onto an open subset of X for each i. Assume that U_i and U_k overlap and consider the function

$$\psi_k\psi_i^{-1} : \psi_i(U_i \cap U_k) \to \psi_k(U_i \cap U_k).$$

Then for some j and ℓ, we have

$$\psi_k\psi_i^{-1}(x) = \delta\kappa_{k\ell}^{-1}\kappa_{ij}\delta^{-1}(x)$$

for each x in $\psi_i(U_i \cap U_k)$. Hence $\psi_k\psi_i^{-1}$ agrees in a neighborhood of each point of its domain with $\delta\tau\delta^{-1}$ for some covering transformation τ of κ. By hypothesis, $\delta\tau\delta^{-1}$ agrees with an element of H. Hence $\{\psi_i : U_i \to X\}$ is an (X,H)-atlas for M.

Now as $\phi_{ij} : U_{ij} \to X$ is a chart for $\tilde{M}$, we have that $\phi_{ij}\delta^{-1} : \delta(U_{ij}) \to X$ is the restriction of an element of G. Since

$$\phi_i\psi_i^{-1} = \phi_i\kappa_{ij}\delta^{-1} = \phi_{ij}\delta^{-1},$$

we have that $\phi_i\psi_i^{-1}$ is the restriction of an element of G. This implies that $\{\psi_i\}$ is contained in the (X,G)-structure of M. Consequently, the (X,H)-structure on M determined by $\{\psi_i\}$ is contained in the (X,G)-structure of M. Thus, the (X,G)-structure of M contains an (X,H)-structure. □

Definition: An (X,G)-manifold M is *orientable* if and only if the (X,G)-structure of M contains an (X,G_0)-structure for M, where G_0 is the group of orientation preserving elements of G.

By Theorem 8.4.5, a connected (X,G)-manifold M is orientable if and only if the image of a holonomy $\eta : \pi_1(X) \to G$ for M consists of orientation preserving elements of G.

Exercise 8.4

1. Prove that an (X,G)-map is a local homeomorphism.
2. Prove that a composition of (X,G)-maps is an (X,G)-map.
3. Let X be a geometric space and let G be a subgroup of $\mathrm{S}(X)$. Prove that a function $\xi : X \to X$ is an (X,G)-map if and only if ξ is in G.
4. Let M be an (X,G)-manifold and let $\kappa : \tilde{M} \to M$ be a covering projection. Prove that $\tilde{M}$ has a unique (X,G)-structure so that κ is an (X,G)-map.
5. Let M and N be (X,G)-manifolds, let $\kappa : \tilde{M} \to M$ be a covering projection, and let $\xi : M \to N$ and $\tilde{\xi} : \tilde{M} \to N$ be functions such that $\tilde{\xi} = \xi\kappa$. Prove that ξ is an (X,G)-map if and only if $\tilde{\xi}$ is an (X,G)-map.
6. Prove that an (X,G)-map $\xi : M \to N$ between metric (X,G)-manifolds is a local isometry.
7. Let U be a nonempty open connected subset of $X = S^n, E^n$, or H^n, and let $\phi : U \to X$ be a distance preserving function. Prove that ϕ extends to a unique isometry of X.
8. Let $X = S^n, E^n$, or H^n, and let $\xi : M \to N$ be a function between metric $(X, \mathrm{I}(X))$-manifolds. Prove that ξ is an $(X, \mathrm{I}(X))$-map if and only if ξ is a local isometry.
9. Let M be a connected (X,G)-manifold and let H be a normal subgroup of G. Prove that the (X,G)-structure of M contains an (X,H)-structure if and only if H contains the image of every holonomy for M.
10. Let M be a connected (X,G)-manifold and let H be a normal subgroup of G. Suppose that the (X,G)-structure of M contains an (X,H)-structure for M. Prove that the set of (X,H)-structures for M contained in the (X,G)-structure of M is in one-to-one correspondence with G/H.

§8.5. Completeness

In this section, we study the role of various forms of completeness in the theory of (X, G)-manifolds. We begin with the most elementary form of completeness.

Metric Completeness

Definition: An infinite sequence $\{x_i\}_{i=1}^\infty$ in a metric space X is a *Cauchy sequence* if and only if for each $\epsilon > 0$, there is a positive integer k such that $d(x_i, x_j) < \epsilon$ for all $i, j \geq k$.

Lemma 1. *Let $\{x_i\}_{i=1}^\infty$ be a Cauchy sequence in a metric space X. Then $\{x_i\}$ converges in X if and only if $\{x_i\}$ has a limit point in X.*

Proof: Let y be a limit point of $\{x_i\}$ in X. We shall prove that $\{x_i\}$ converges to y. Let $\epsilon > 0$. As $\{x_i\}$ is a Cauchy sequence, there is an integer k such that for all $i, j \geq k$, we have

$$d(x_i, x_j) < \epsilon/2.$$

As y is a limit point of $\{x_i\}$, there is an integer $\ell \geq k$ such that

$$d(x_\ell, y) < \epsilon/2.$$

Hence, for all $i \geq k$, we have

$$d(x_i, y) \leq d(x_i, x_\ell) + d(x_\ell, y) < \epsilon.$$

Thus $x_i \to y$ in X. □

Definition: A metric space X is *complete* if and only if every Cauchy sequence in X converges in X.

Theorem 8.5.1. *Let X be a metric space and suppose there is an $\epsilon > 0$ such that $\overline{B}(x, \epsilon)$ is compact for all x in X. Then X is complete.*

Proof: Let $\{x_i\}$ be a Cauchy sequence in X. Then there is a positive integer k such that $d(x_i, x_j) < \epsilon$ for all $i, j \geq k$. Hence $B(x_k, \epsilon)$ contains x_i for all $i \geq k$. As $\overline{B}(x_k, \epsilon)$ is compact, the sequence $\{x_i\}$ has a limit point in $\overline{B}(x_k, \epsilon)$. Hence $\{x_i\}$ converges by Lemma 1. Thus X is complete. □

Theorem 8.5.2. *Let Γ be a group of isometries of a finitely compact metric space all of whose Γ-orbits are closed. Then X/Γ is a complete metric space.*

Proof: Let $B(x, r)$ be an open ball in X. Then the quotient map $\pi : X \to X/\Gamma$ maps $B(x, r)$ onto $B(\pi(x), r)$ by Theorem 6.5.2. As $\overline{B}(x, r)$ is compact, we have

$$\pi(\overline{B}(x, r)) = \overline{B}(\pi(x), r).$$

Hence $\overline{B}(\pi(x), r)$ is compact. Thus X/Γ is complete by Theorem 8.5.1. □

Theorem 8.5.3. *Let Γ be a group of isometries of a metric space X such that each Γ-orbit is a closed discrete subset of X. If X/Γ is complete, then X is complete.*

Proof: Let $\{x_i\}$ be a Cauchy sequence in X. Then $\{\Gamma x_i\}$ is a Cauchy sequence in X/Γ, since

$$\text{dist}(\Gamma x_i, \Gamma x_j) \leq d(x_i, x_j).$$

Hence $\{\Gamma x_i\}$ converges to an orbit Γy. Set

$$s = \frac{1}{2}\text{dist}(y, \Gamma y - \{y\}).$$

Then $s > 0$, since Γy is a closed discrete subset of X. Now for all g in Γ, we have that

$$s = \frac{1}{2}\text{dist}(gy, \Gamma y - \{gy\}).$$

As $\{x_i\}$ is a Cauchy sequence, there is an integer k such that $d(x_i, x_j) < s/2$ for all $i, j, \geq k$. Suppose that $0 < \epsilon \leq s/2$. As $\Gamma x_i \to \Gamma y$, there is an integer $\ell \geq k$ and an element g_i of Γ such that $d(x_i, g_i y) < \epsilon$ for all $i \geq \ell$. Hence, if $i \geq \ell$, then

$$d(x_k, g_i y) \leq d(x_k, x_i) + d(x_i, g_i y) < s.$$

But $B(x_k, s)$ contains at most one point of Γy. Therefore, there is an element g of Γ such that $g_i y = gy$ for all $i \geq \ell$. Moreover $d(x_i, gy) < \epsilon$ for all $i \geq \ell$. Therefore $x_i \to gy$. Thus X is complete. □

Theorem 8.5.4. *Let X be a complete metric space and let $\xi : X \to X$ be a similarity that is not an isometry. Then ξ has a unique fixed point in X.*

Proof: By replacing ξ by ξ^{-1}, if necessary, we may assume that the scale factor k of ξ is less than one. Let x be any point of X. Define a sequence $\{x_m\}_{m=1}^{\infty}$ in X by $x_m = \xi^m(x)$ for each m. Then for $m < n$, we have

$$\begin{aligned}
d(x_m, x_n) &= d(\xi^m(x), \xi^n(x)). \\
&\leq \sum_{\ell=m}^{n-1} d(\xi^\ell(x), \xi^{\ell+1}(x)) \\
&= (k^m + k^{m+1} + \cdots + k^{n-1}) d(x, \xi(x)) \\
&= \left(\frac{k^m - k^n}{1-k}\right) d(x, \xi(x)) \\
&< k^m \left(\frac{d(x, \xi(x))}{1-k}\right).
\end{aligned}$$

Consequently $\{x_m\}$ is a Cauchy sequence in X. Therefore, the sequence $\{x_m\}$ converges to a point y in X. As ξ is continuous, the sequence $\{\xi(x_m)\}$ converges to $\xi(y)$. But $\xi(x_m) = x_{m+1}$. Hence $\{x_m\}$ and $\{\xi(x_m)\}$ converge to the same point, and so $\xi(y) = y$. Thus y is a fixed point of ξ in X.

Now let z be another fixed point of ξ. Then

$$d(y,z) = d(\xi(y), \xi(z)) = kd(y,z).$$

Hence $d(y,z) = 0$ and so $y = z$. Thus y is the unique fixed point of ξ. □

Geodesic Completeness

We next consider the role of geodesic completeness in the theory of metric (X, G)-manifolds. Recall that a metric space X is geodesically complete if and only if each geodesic arc $\alpha : [a,b] \to X$ extends to a unique geodesic line $\lambda : \mathbb{R} \to X$.

Theorem 8.5.5. *If M is a geodesically complete metric (X, G)-manifold, then M is geodesically connected.*

Proof: Let u, v be points of M, with $d(u,v) = \ell > 0$, and let $\phi : U \to X$ be a chart for (M, u). Choose $r > 0$ so that $\phi(U)$ contains $B(\phi(u), 2r)$. Then ϕ maps $B(u,r)$ isometrically onto $B(\phi(u), r)$ by Theorem 8.3.6.

Assume first that v is in $B(u,r)$. Then $\phi(v)$ is in $B(\phi(u), r)$ and

$$d(\phi(u), \phi(v)) = d(u,v) = \ell.$$

As X is geodesically connected, there is a geodesic arc $\alpha : [0,\ell] \to X$ from $\phi(u)$ to $\phi(v)$. Observe that

$$|\alpha| = \ell = d(u,v) < r.$$

Therefore $B(\phi(u), r)$ contains the image of α. Hence $\phi^{-1}\alpha : [0,\ell] \to M$ is a geodesic arc from u to v.

Now assume that v is not in $B(u,r)$. Let S be a sphere $S(u,\epsilon)$ in M with $\epsilon < r$. Then the function $\delta : S \to \mathbb{R}$, defined by $\delta(z) = d(z,v)$, is continuous. As S is compact, there is a point w on S at which δ attains its minimum value. Since w is in $B(u,r)$, there is a geodesic arc $\beta : [0,b] \to M$ from u to w. Moreover β extends to a unique geodesic line $\lambda : \mathbb{R} \to M$, since M is geodesically complete.

We claim that $\lambda(\ell) = v$. To prove this result, we shall prove that $d(\lambda(t), v) = \ell - t$ for all t in $[b, \ell]$. First of all, since every curve from u to v must intersect S, we have

$$\begin{aligned} d(u,v) &\geq \operatorname{dist}(u,S) + \operatorname{dist}(S,v) \\ &= d(u,w) + d(w,v) \\ &\geq d(u,v). \end{aligned}$$

Hence, we have

$$d(\lambda(b), v) = d(w,v) = \ell - b.$$

Now let s be the supremum of all t in $[b,\ell]$ such that $d(\lambda(t), v) = \ell - t$. Then $d(\lambda(s), v) = \ell - s$ by the continuity of $d(\lambda(t), v)$ as a function of t. Let $\lambda_{0,s} : [0,s] \to M$ be the restriction of λ. As

$$d(u,v) \leq d(u, \lambda(s)) + d(\lambda(s), v),$$

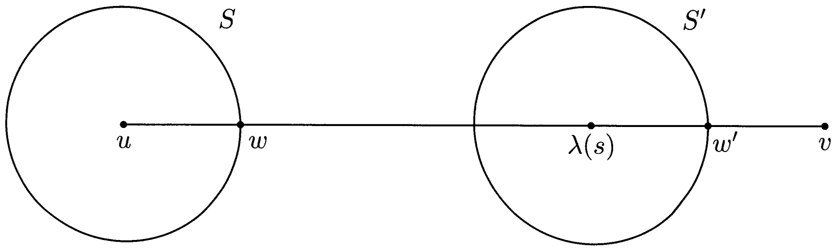


Figure 8.5.1. A geodesic segment joining u to v

we have that

$$\ell \leq d(\lambda(0), \lambda(s)) + \ell - s.$$

Hence, we have

$$\|\lambda_{0,s}\| = s \leq d(\lambda(0), \lambda(s)).$$

Therefore $\|\lambda_{0,s}\| = d(\lambda(0), \lambda(s))$. Consequently $\lambda_{0,s}$ is a geodesic arc. Suppose that $s < \ell$. We shall derive a contradiction.

Let $\psi : V \to X$ be a chart for $(M, \lambda(s))$. Choose $r' > 0$ so that $\psi(V)$ contains $B(\psi\lambda(s), 2r')$. Let S' be a sphere $S(\lambda(s), \epsilon')$ with

$$\epsilon' < \min\{r', \ell - s\}$$

and let w' be a point on S' nearest to v. See Figure 8.5.1. Now since

$$d(\lambda(s), v) = \ell - s \quad \text{and} \quad \epsilon' < \ell - s,$$

we have that v is not in the closed ball $C(\lambda(s), \epsilon')$. Therefore

$$\begin{aligned} d(\lambda(s), v) &\geq \operatorname{dist}(\lambda(s), S') + \operatorname{dist}(S', v) \\ &= d(\lambda(s), w') + d(w', v) \quad \geq \quad d(\lambda(s), v). \end{aligned}$$

Hence $d(\lambda(s), v) = \epsilon' + d(w', v)$ and so $d(w', v) = (\ell - s) - \epsilon'$. Therefore

$$\begin{aligned} d(u, w') &\geq d(u, v) - d(w', v) \\ &= \ell - (\ell - s - \epsilon') \\ &= s + \epsilon' \\ &= d(u, \lambda(s)) + d(\lambda(s), w') \quad \geq \quad d(u, w'). \end{aligned}$$

Let $\gamma : [0, s + \epsilon'] \to M$ be the composite of $\lambda_{0,s}$ and a geodesic arc from $\lambda(s)$ to w'. Then γ is a geodesic arc by Theorem 1.4.3, since

$$d(u, w') = d(u, \lambda(s)) + d(\lambda(s), w').$$

As M is geodesically complete, the arc γ extends to a unique geodesic line $\mu : \mathbb{R} \to M$. But μ also extends $\lambda_{0,s}$. Therefore $\mu = \lambda$. Hence λ agrees with γ, and so $\lambda(s+\epsilon') = w'$. Therefore

$$d(\lambda(s+\epsilon'), v) = \ell - (s+\epsilon').$$

But this contradicts the supremacy of s. Therefore $s = \ell$. Hence $\lambda(\ell) = v$ and $\lambda_{0,\ell}$ is a geodesic arc in M from u to v. Thus M is geodesically connected. □

Lemma 2. *Let X be a geometric space. Then there is a $k > 0$ such that if $\lambda : \mathbb{R} \to X$ is a geodesic line, then λ restricts to a geodesic arc on the interval $[-k, k]$.*

Proof: Let k be as in Axiom 3 for a geometric space. Then k has the desired property by Axioms 3 and 4 and Theorem 8.1.1. □

Theorem 8.5.6. *Let M be a metric (X, G)-manifold and let $\xi : M \to X$ be a local isometry. Then M is geodesically complete if and only if ξ is a covering projection.*

Proof: Suppose that ξ is a covering projection. Let $\alpha : [a, b] \to M$ be a geodesic arc in M. As ξ is a local isometry, $\xi\alpha : [a, b] \to X$ is a geodesic curve. Consequently, $\xi\alpha$ extends to a unique geodesic line $\lambda : \mathbb{R} \to X$. Since ξ is a covering projection, λ lifts to a geodesic line $\mu : \mathbb{R} \to M$ such that $\mu(a) = \alpha(a)$. By unique path lifting, μ extends α. Now let $\mu' : \mathbb{R} \to M$ be another geodesic line extending α. Then $\xi\mu' : \mathbb{R} \to X$ is a geodesic line extending $\xi\alpha$. Therefore $\xi\mu' = \lambda$. By the unique lifting property of covering projections, $\mu' = \mu$. Hence μ is the unique geodesic line in M extending α. Thus M is geodesically complete.

Conversely, suppose that M is a geodesically complete. We first show that geodesic arcs in X can be lifted with respect to ξ. Let $\alpha : [a, b] \to X$ be a geodesic arc and let u be a point of M such that $\xi(u) = \alpha(a)$. Since ξ is a local isometry, there is a geodesic arc $\beta : [a, c] \to M$ such that $\beta(a) = u$, $c < b$, and $\xi\beta$ is the restriction $\alpha_{a,c}$ of α to $[a, c]$. As M is geodesically complete, β extends to a unique geodesic line $\mu : \mathbb{R} \to M$. Since ξ is a local isometry, $\xi\mu : \mathbb{R} \to X$ is a geodesic line extending $\alpha_{a,c}$. Hence $\xi\mu : \mathbb{R} \to X$ is the unique geodesic line extending α. Let $\tilde{\alpha} : [a, b] \to M$ be the restriction of μ. Then $\tilde{\alpha}(a) = u$ and $\xi\tilde{\alpha} = \alpha$. Thus, geodesic arcs can be lifted with respect to ξ.

Next, we show that ξ is surjective. Let x be a point in the image of ξ and let y be any other point of X. As X is geodesically connected, there is a geodesic arc $\alpha : [0, \ell] \to X$ from x to y. As x is in the image of ξ, we can lift α to a curve $\tilde{\alpha} : [0, \ell] \to M$ with respect to ξ. Then

$$\xi\tilde{\alpha}(\ell) = \alpha(\ell) = y.$$

Hence y is in the image of ξ. Thus ξ is surjective.

Now let $B(x,r)$ be an arbitrary open ball in X. We next show that

$$\xi^{-1}(B(x,r)) = \bigcup_{u \in \xi^{-1}(x)} B(u,r).$$

As ξ is a local isometry, we have

$$\xi(B(u,r)) \subset B(x,r)$$

for each u in $\xi^{-1}(x)$. Therefore

$$\bigcup_{u \in \xi^{-1}(x)} B(u,r) \subset \xi^{-1}(B(x,r)).$$

Now let v be an arbitrary point in $\xi^{-1}(B(x,r))$. Then $\xi(v)$ is in $B(x,r)$. Let $\alpha : [0,\ell] \to X$ be a geodesic arc from $\xi(v)$ to x, and let $\tilde{\alpha} : [0,\ell] \to M$ be a lift of α with respect to ξ such that $\tilde{\alpha}(0) = v$. Then

$$\xi\tilde{\alpha}(\ell) = \alpha(\ell) = x.$$

Thus $\tilde{\alpha}(\ell)$ is in $\xi^{-1}(x)$. Moreover

$$\|\tilde{\alpha}\| = |\alpha| = d(x, \xi(v)) < r.$$

Therefore v is in $B(\tilde{\alpha}(\ell), r)$. This shows that

$$\xi^{-1}(B(x,r)) \subset \bigcup_{u \in \xi^{-1}(x)} B(u,r).$$

Since we have already established the reverse inclusion, we have

$$\xi^{-1}(B(x,r)) = \bigcup_{u \in \xi^{-1}(x)} B(u,r).$$

Let u be in $\xi^{-1}(x)$. We next show that ξ maps $B(u,r)$ onto $B(x,r)$. Let y be an arbitrary point of $B(x,r)$ other than x. Then there is a geodesic arc $\alpha : [0,\ell] \to X$ from x to y. Moreover, there is a lift $\tilde{\alpha} : [0,\ell] \to M$ with respect to ξ such that $\tilde{\alpha}(0) = u$. Then $\xi\tilde{\alpha}(\ell) = \alpha(\ell) = y$. Furthermore

$$\|\tilde{\alpha}\| = |\alpha| = d(x,y) < r.$$

Therefore $\tilde{\alpha}(\ell)$ is in $B(u,r)$. This shows that ξ maps $B(u,r)$ onto $B(x,r)$.

By Lemma 2, there is a $k > 0$ such that if $\lambda : \mathbb{R} \to X$ is a geodesic line, then λ restricts to a geodesic arc on $[-k,k]$. Let u be in $\xi^{-1}(x)$. We next show that ξ maps $B(u,k)$ bijectively onto $B(x,k)$. We have already established that ξ maps $B(u,k)$ onto $B(x,k)$. On the contrary, suppose that v, w are distinct points of $B(u,k)$ such that $\xi(v) = \xi(w)$. By Theorem 8.5.5, there is a geodesic arc $\alpha : [-b,b] \to M$ from v to w. As the endpoints of α are in $B(u,k)$, we have

$$2b = d(v,w) \ \leq \ d(v,u) + d(u,w) \ < \ 2k.$$

Therefore $0 < b < k$. As M is geodesically complete, α extends to a geodesic line $\mu : \mathbb{R} \to M$. Because of the choice of k, the geodesic line $\xi\mu : \mathbb{R} \to X$ restricts to a geodesic arc on $[-k,k]$. Therefore $\xi\alpha : [-b,b] \to X$ is a geodesic arc from $\xi(v)$ to $\xi(w)$, which is a contradiction. Hence ξ maps $B(u,k)$ bijectively onto $B(x,k)$.

By the triangle inequality, the sets $\{B(u,k/2) : u \in \xi^{-1}(x)\}$ are pairwise disjoint. Now since ξ maps $B(u,k/2)$ homeomorphically onto $B(x,k/2)$ for each u in $\xi^{-1}(x)$ and

$$\xi^{-1}(B(x,k/2)) = \bigcup_{u\in\xi^{-1}(x)} B(u,k/2),$$

the set $B(x,k/2)$ is evenly covered by ξ. Thus ξ is a covering projection. □

Complete (X,G)-Manifolds

Let $\delta : \tilde{M} \to X$ be a developing map for a connected (X,G)-manifold M. Let $\{U_i\}$ be the collection of all the open connected sets U_i of $\tilde{M}$ such that δ maps U_i homeomorphically into X, and let $\phi_i : U_i \to X$ be the restriction of δ. Then $\{\phi_i\}$ is an $(X,\{1\})$-structure for $\tilde{M}$, and $\{\phi_i\}$ is contained in the (X,G)-structure on $\tilde{M}$, since δ is an (X,G)-map. We shall regard the universal covering space $\tilde{M}$ to be an $(X,\{1\})$-manifold with the $(X,\{1\})$-structure $\{\phi_i\}$. Then δ is also a developing map for the $(X,\{1\})$-manifold $\tilde{M}$, since $\delta : \tilde{M} \to X$ is the unique $(X,\{1\})$-map extending $\phi_i : U_i \to X$. Note that the $(X,\{1\})$-structure on $\tilde{M}$ is unique up to multiplication by an element of G. Therefore, the induced metric on $\tilde{M}$ is unique up to multiplication by a scale factor of an element of G.

Definition: An (X,G)-manifold M is *complete* if and only if the universal covering space of each connected component of M is a complete metric space.

Theorem 8.5.7. *Let M be a metric (X,G)-manifold. Then the following are equivalent:*

1. *M is complete;*
2. *M is geodesically complete;*
3. *M is a complete metric space.*

Proof: Suppose that M is complete. Then $\tilde{M}$ is a complete metric space. We now show that $\tilde{M}$ is geodesically complete. Let $\alpha : [a,b] \to \tilde{M}$ be a geodesic arc and let $\delta : \tilde{M} \to X$ be a developing map for M. Then $\delta\alpha : [a,b] \to X$ is a geodesic curve. Hence, there is a unique geodesic line $\lambda : \mathbb{R} \to X$ extending $\delta\alpha$. Let I be the largest interval in $\mathbb{R}$ containing $[a,b]$ for which there is a map $\mu : I \to \tilde{M}$ lifting λ with respect to δ. Then I is open, since δ is a local homeomorphism. On the contrary, suppose that I is not all of $\mathbb{R}$. Then there is a sequence of real numbers $\{t_i\}$ in I converging to an endpoint c of I. As δ is a local isometry, μ is locally a geodesic arc. Therefore, μ does not increase distances. Hence $\{\mu(t_i)\}$ is a Cauchy sequence in $\tilde{M}$. As $\tilde{M}$ is a complete metric space, $\{\mu(t_i)\}$ converges to a point $\tilde{u}$ in $\tilde{M}$. Now extend μ to a function $\overline{\mu} : I \cup \{c\} \to \tilde{M}$ by setting

$\overline{\mu}(c) = \tilde{u}$. Then $\overline{\mu}$ is continuous, since the point $\tilde{u}$ does not depend on the choice of the sequence $\{t_i\}$ converging to the point c. Observe that

$$\delta\overline{\mu}(c) = \lim_{i\to\infty} \delta\mu(t_i) = \lim_{i\to\infty} \lambda(t_i) = \lambda(c).$$

Hence $\overline{\mu} : I \cup \{c\} \to \tilde{M}$ further lifts λ. But this contradicts the maximality of I. Thus I is all of $\mathbb{R}$ and $\mu : \mathbb{R} \to \tilde{M}$ is a geodesic line extending α.

Let $\mu' : \mathbb{R} \to \tilde{M}$ be another geodesic line extending α. As δ is a local isometry, $\delta\mu' : \mathbb{R} \to X$ is a geodesic line extending $\delta\alpha$. Hence $\delta\mu' = \lambda = \delta\mu$. Therefore $\mu' = \mu$, since δ is a local homeomorphism. Hence μ is the unique geodesic line extending α. Thus $\tilde{M}$ is geodesically complete. Therefore M is geodesically complete, since the universal covering projection $\kappa : \tilde{M} \to M$ is a local isometry. Thus (1) implies (2).

Now assume that M is geodesically complete. Then $\tilde{M}$ is geodesically complete, since the universal covering projection $\kappa : \tilde{M} \to M$ is a local isometry. Therefore $\delta : \tilde{M} \to X$ is a covering projection by Theorem 8.5.6. Furthermore, the proof of Theorem 8.5.6 shows that there is an $r > 0$ such that $B(x, 2r)$ is evenly covered by δ for all x in X. Let $\tilde{u}$ be a point of $\tilde{M}$. From the proof of Theorem 8.5.6, we have that δ maps $B(\tilde{u}, r)$ onto $B(\delta(\tilde{u}), r)$. As δ is continuous, we have

$$\delta(\overline{B}(\tilde{u}, r)) \subset \overline{B}(\delta(\tilde{u}), r).$$

By a geodesic arc lifting argument, δ maps $\overline{B}(\tilde{u}, r)$ onto $\overline{B}(\delta(\tilde{u}), r)$. Now as δ maps $\overline{B}(\tilde{u}, r)$ homeomorphically onto $\overline{B}(\delta(\tilde{u}), r)$, we have that $\overline{B}(\tilde{u}, r)$ is compact for each point $\tilde{u}$ of $\tilde{M}$. By the same argument, the covering projection $\kappa : \tilde{M} \to M$ maps $\overline{B}(\tilde{u}, r)$ onto $\overline{B}(\kappa(\tilde{u}), r)$. Therefore $\overline{B}(u, r)$ is compact for each point u of M. Hence M is a complete metric space by Theorem 8.5.1. Thus (2) implies (3).

Now assume that M is a complete metric space. Let Γ be the group of covering transformations of the universal covering $\kappa : \tilde{M} \to M$. Then Γ is a group of isometries of $\tilde{M}$ and κ induces a homeomorphism $\overline{\kappa} : \tilde{M}/\Gamma \to M$. Moreover $\overline{\kappa}$ is a local isometry, since κ and the quotient map $\pi : \tilde{M} \to \tilde{M}/\Gamma$ are local isometries. Now the homeomorphism $\overline{\kappa} : \tilde{M}/\Gamma \to M$ induces an (X, G)-manifold structure on $\tilde{M}/\Gamma$. We claim that the orbit space metric d_Γ on $\tilde{M}/\Gamma$ is the same as the induced (X, G)-manifold metric d on $\tilde{M}/\Gamma$. First of all, d_Γ and d agree locally, since $\overline{\kappa} : \tilde{M}/\Gamma \to M$ is a local isometry; moreover $d_\Gamma \leq d$, since arc length with respect to d_Γ is the same as X-length. Finally, $d_\Gamma = d$, since π preserves X-length. Therefore $\overline{\kappa} : \tilde{M}/\Gamma \to M$ is an isometry. Hence $\tilde{M}/\Gamma$ is a complete metric space. Therefore $\tilde{M}$ is a complete metric space by Theorem 8.5.3. Thus (3) implies (1). □

Definition: An (X, G)-structure Φ for a manifold M is *complete* if and only if M, with the (X, G)-structure Φ, is a complete (X, G)-manifold.

Theorem 8.5.8. *Let M be an (X, G)-manifold and let G_1 be the group of isometries in G. Then M is complete if and only if the (X, G)-structure of M contains a complete (X, G_1)-structure for M.*

Proof: Without loss of generality, we may assume that M is connected. Suppose that M is complete. Then the universal covering space $\tilde{M}$ is a complete metric space. Let $\tau : \tilde{M} \to \tilde{M}$ be a nonidentity covering transformation of the universal covering projection $\kappa : \tilde{M} \to M$. Then τ is an (X, G)-map. Hence τ is locally a similarity. Moreover, as $\tilde{M}$ is connected, all the local scale factors of τ have the same value k. Let $\gamma : [a, b] \to \tilde{M}$ be an X-rectifiable curve from u to v. Then $\|\tau\gamma\| = k\|\gamma\|$. Hence, we have that

$$d(\tau(u), \tau(v)) \leq kd(u, v).$$

Likewise, we have

$$d(\tau^{-1}(u), \tau^{-1}(v)) \leq k^{-1}d(u, v).$$

Observe that

$$kd(u, v) = kd(\tau^{-1}(\tau(u)), \tau^{-1}(\tau(v))) \leq d(\tau(u), \tau(v)).$$

Therefore, we have

$$d(\tau(u), \tau(v)) = kd(u, v).$$

Thus τ is a similarity. Since τ has no fixed points, τ is an isometry by Theorem 8.5.4.

Let $\eta : \pi_1(M) \to G$ be the holonomy determined by δ. Then η is defined by $\eta([\alpha]) = g_\alpha$ where $\delta\tau_\alpha = g_\alpha\delta$ and τ_α is a certain covering transformation of κ. As δ and τ_α are local isometries, g_α is an isometry of X. Hence, the image of η is contained in the group G_1 of isometries in G. Therefore, the (X, G)-structure Φ of M contains an (X, G_1)-structure Φ_1 for M by Theorem 8.4.5. Moreover Φ_1 is complete, since $\tilde{M}$ is a complete metric space.

Conversely, suppose that the (X, G)-structure Φ of M contains a complete (X, G_1)-structure Φ_1 for M. Then $\tilde{M}$ is a complete metric space. Therefore M is a complete (X, G)-manifold. □

Definition: A function $\xi : M \to N$ between (X, G)-manifolds is an (X, G)-*equivalence* if and only if ξ is a bijective (X, G)-map.

Clearly, the inverse of an (X, G)-equivalence is also an (X, G)-equivalence. Two (X, G)-manifolds M and N are said to be (X, G)-*equivalent* if and only if there is an (X, G)-equivalence $\xi : M \to N$. Note that an (X, G)-equivalence $\xi : M \to N$ between metric (X, G)-manifolds is an isometry.

Theorem 8.5.9. *Let G be a group of similarities of a simply connected geometric space X and let M be a complete connected (X, G)-manifold. Let $\delta : \tilde{M} \to X$ be a developing map for M and let $\eta : \pi_1(M) \to G$ be the holonomy of M determined by δ. Then δ is an $(X, \{1\})$-equivalence, η maps $\pi_1(M)$ isomorphically onto a freely acting discrete group Γ of isometries of X, and δ induces an (X, G)-equivalence from M to X/Γ.*

Proof: First of all, $\tilde{M}$ is geodesically complete by Theorem 8.5.7. Hence, the developing map $\delta : \tilde{M} \to X$ is a covering projection by Theorem 8.5.6. Therefore δ is a homeomorphism, since X is simply connected. Hence δ is an $(X, \{1\})$-equivalence and so is an isometry. Now $\pi_1(M)$ corresponds to the group of covering transformations of the universal covering $\kappa : \tilde{M} \to M$ which corresponds via δ to the image of η. Therefore η maps $\pi_1(M)$ isomorphically onto a freely acting discrete group Γ of isometries of X. Moreover δ induces a homeomorphism $\overline{\delta}$ such that the following diagram commutes:

$$\begin{array}{ccc} \tilde{M} & \xrightarrow{\delta} & X \\ \kappa \downarrow & & \downarrow \pi \\ M & \xrightarrow{\overline{\delta}} & X/\Gamma, \end{array}$$

where π is the quotient map. As κ, δ, and π are (X, G)-maps, $\overline{\delta}$ is an (X, G)-map. Hence $\overline{\delta}$ is an (X, G)-equivalence. □

Theorem 8.5.10. *Let M be a metric $(X, \mathrm{I}(X))$-manifold with X simply connected. Then the following are equivalent:*

(1) *The manifold M is complete.*

(2) *There is an $\epsilon > 0$ such that each closed ϵ-ball in M is compact.*

(3) *All the closed balls in M are compact.*

(4) *There is a sequence $\{M_i\}_{i=1}^{\infty}$ of compact subsets of the manifold M such that $M = \cup_{i=1}^{\infty} M_i$ and $N(M_i, 1) \subset M_{i+1}$ for each i.*

Proof: Assume that M is complete. Then M is isometric to an X-space-form X/Γ by Theorem 8.5.9. Now all the closed balls in X are compact by Theorem 8.1.2. Hence, all the closed balls in X/Γ are compact by Theorem 6.5.2. Therefore, all the closed balls in M are compact. Thus (1) implies (3). As (3) implies (2), and (2) implies (1) by Theorem 8.5.1, we have that (1)-(3) are equivalent.

Now assume that all the closed balls in M are compact. Let u be a point of M. For each integer $i > 0$, let $M_i = C(u, i)$. Then $M = \cup_{i=1}^{\infty} M_i$ and

$$N(M_i, 1) \subset M_{i+1}$$

for each i. Thus (3) implies (4).

Now assume that (4) holds. Let $\{u_i\}$ be a Cauchy sequence in M. Then there is an integer k such $d(u_i, u_j) < 1$ for all $i, j \geq k$. As $M = \cup_{i=1}^{\infty} M_i$, there is an integer ℓ such that

$$\{u_1, \ldots, u_k\} \subset M_\ell.$$

Then the set $M_{\ell+1}$ contains the entire sequence $\{u_i\}$, since

$$N(M_\ell, 1) \subset M_{\ell+1}.$$

As $M_{\ell+1}$ is compact, the sequence $\{u_i\}$ converges. Therefore M is complete. Hence (4) implies (1). Thus (1)-(4) are equivalent. □

Exercise 8.5

1. Prove that every locally compact, homogeneous, metric space X is complete.
2. Let X be a connected n-manifold with a complete metric. Prove that a function $\xi : X \to X$ is an isometry if and only if it preserves distances.
   Hint: Use invariance of domain.
3. Prove that a local isometry $\xi : M \to N$ between metric (X, G)-manifolds is an isometry if and only if it is a bijection.
4. Let M be a metric (X, G)-manifold and let $\kappa : \tilde{M} \to M$ be a covering projection with $\tilde{M}$ connected. Prove that M is geodesically complete if and only if $\tilde{M}$ is geodesically complete.
5. Prove that a local isometry $\xi : M \to N$ between geodesically complete metric (X, G)-manifolds is a covering projection.
6. Prove that a connected (X, G)-manifold M is complete if and only if every (or some) developing map $\delta : \tilde{M} \to X$ for M is a covering projection.
7. Let X be a simply connected geometric space. Prove that a function $\xi : X \to X$ is an isometry if and only if it is a local isometry.
8. Prove that the universal covering space $\tilde{X}$ of a geometric space X is also a geometric space.
9. Let M be an $(X, \mathrm{I}(X))$-manifold and let $\tilde{X}$ be the universal covering space of X. Prove that the $(X, \mathrm{I}(X))$-structure of M lifts to an $(\tilde{X}, \mathrm{I}(\tilde{X}))$-structure for M.
10. Let M be a complete connected $(X, \mathrm{I}(X))$-manifold and let $\tilde{X}$ be the universal covering space of X. Prove that M is $(\tilde{X}, \mathrm{I}(\tilde{X}))$-equivalent to an $\tilde{X}$-space-form.

§8.6. Curvature

In this section, we briefly describe the role of curvature in the theory of spherical, Euclidean, and hyperbolic manifolds. We assume that the reader is familiar with the basic theory of Riemannian manifolds. In particular, every connected Riemannian manifold has a natural metric space structure.

Theorem 8.6.1. *A connected Riemannian n-manifold X is an n-dimensional geometric space if and only if X is homogeneous.*

Proof: Suppose that X is homogeneous. Then X is a complete metric space by Theorem 8.5.1. Hence X is geodesically connected and geodesically complete by the Hopf-Rinow-Whitehead Theorem. The exponential map at any point of X determines a function $\varepsilon : E^n \to X$ satisfying Axiom 3 for a geometric space. □

Remark: It is a theorem of Berestovskii that an n-dimensional geometric space X has a Riemannian metric compatible with its topology such that every isometry of X is an isometry of the Riemannian metric.

Definition: An n-dimensional *geometry* is a simply connected, homogeneous, Riemannian n-manifold X for which there is at least one X-space-form of finite volume.

Euclidean 1-dimensional geometry E^1 is the only 1-dimensional geometry up to isometry. If $n > 1$, then S^n, E^n, and H^n are examples of nonsimilar n-dimensional geometries. These geometries are characterized as the geometries of constant curvature because of the following theorem.

Theorem 8.6.2. *Let X be a Riemannian n-manifold such that X is*

(1) *connected,*

(2) *complete,*

(3) *simply connected, and*

(4) *of constant sectional curvature.*

Then X is similar to either S^n, E^n, or H^n.

Remark: One should compare conditions 1-4 in Theorem 8.6.2 with Euclid's Postulates 1-4 in §1.1.

Corollary 1. *If X is a 2-dimensional geometry, then X is similar to either S^2, E^2, or H^2.*

Proof: As X is homogeneous, X is of constant curvature. □

Two n-dimensional geometries X and Y are said to be *equivalent* if and only if there is a diffeomorphism $\phi : X \to Y$ such that ϕ induces an isomorphism $\phi_* : \mathrm{I}(X) \to \mathrm{I}(Y)$ defined by

$$\phi_*(g) = \phi g \phi^{-1}.$$

It is a theorem of Thurston that there are, up to equivalence, exactly eight 3-dimensional geometries.

We end the chapter with the definition of a geometric manifold.

Definition: A *geometric n-manifold* is an $(X, \mathrm{S}(X))$-manifold, where $\mathrm{S}(X)$ is the group of similarities of an n-dimensional geometry X.

Spherical, Euclidean, and hyperbolic manifolds are examples of geometric manifolds.

§8.7. Historical Notes

§8.1. The concept of an *n-dimensional manifold* was introduced by Riemann in his 1854 lecture *Über die Hypothesen, welche der Geometrie zu Grunde liegen* [349]. For a discussion, see Scholz's 1992 article *Riemann's vision of a new approach to geometry* [366], and for the early history of manifolds, see Scholz's 1980 thesis *Geschichte des Mannigfaltigkeitsbegriffs von Riemann bis Poincaré* [363]. The concept of a *geometric space* was introduced here as a metric space generalization of a homogeneous Riemannian manifold. Theorem 8.1.3 for Clifford-Klein space-forms appeared in Hopf's 1926 paper *Zum Clifford-Kleinschen Raumproblem* [198]. The fundamental group was introduced by Poincaré in his 1895 memoir *Analysis situs* [336]. In particular, Theorem 8.1.4 for Clifford-Klein space-forms was described in this paper. Theorem 8.1.5 for closed geometric surfaces was essentially proved by Poincaré in his 1885 paper *Sur un théorème de M. Fuchs* [334].

§8.2. The elliptic plane was introduced by Cayley in his 1859 paper *A sixth memoir upon quantics* [76]. In 1873, Clifford described a Euclidean torus embedded in elliptic 3-space in his paper *Preliminary Sketch of Biquaternions* [82]. Closed hyperbolic surfaces were constructed by Poincaré in his 1882 paper *Théorie des groupes fuchsiens* [330]. In 1890, Klein proposed the problem of determining all the closed spherical, Euclidean, and hyperbolic manifolds in his paper *Zur Nicht-Euklidischen Geometrie* [234]. Killing recognized that a closed spherical, Euclidean, or hyperbolic manifold can be represented as an orbit space of a discontinuous group of isometries acting freely in his 1891 paper *Ueber die Clifford-Klein'schen Raumformen* [222]. In particular, Killing introduced the term *Clifford-Klein space-form* in this paper. For the historical context of Killing's work, see Hawkins' 1980 article *Non-Euclidean geometry and Weierstrassian mathematics* [184]. Theorem 8.2.3 appeared in Killing's 1891 paper [222]. Theorem 8.2.4 appeared in Hopf's 1926 paper [198]. The lens spaces $L(5,1)$ and $L(5,2)$ were shown to be nonhomeomorphic by Alexander in his 1919 paper *Note on two three-dimensional manifolds with the same group* [13]. For the classification of lens spaces, see Brody's 1960 paper *The topological classification of the lens spaces* [58], and for the classification of spherical space-forms, see Wolf's 1984 treatise *Spaces of Constant Curvature* [416]. Theorem 8.2.5 appeared in Auslander and Kuranishi's 1957 paper *On the holonomy group of locally Euclidean spaces* [27]. The Euclidean plane-forms were described by Klein in his 1928 treatise *Vorlesungen über nicht-euklidische Geometrie* [237]. The 3-dimensional Euclidean space-forms were enumerated by Nowacki in his 1934 paper *Die euklidischen, dreidimensionalen, geschlossenen und offenen Raumformen* [322]. See also Hantzsche and Wendt's 1935 paper *Dreidimensionale euklidische Raumformen* [178]. References for Euclidean space-forms are Wolf's 1984 treatise [416] and Charlap's 1986 text *Bieberbach Groups and Flat Manifolds* [77].

§8.3. The concept of an (X, G)-manifold originated in the notion of a locally homogeneous Riemannian manifold introduced by Cartan in his 1926 paper *L'application des espaces de Riemann et l'analysis situs* [69]. The concept of an (X, G)-manifold was introduced by Veblen and Whitehead in their 1931 paper *A set of axioms for differential geometry* [395]. For further development of the theory of (X, G)-manifolds, see Goldman's 1988 paper *Geometric structures on manifolds and varieties of representations* [154].

§8.4. The concept of the developing map originated in the notion of a developable surface introduced by Euler in his 1772 paper *De solidis quorum superficiem in planum explicare licet* [125]. For commentary, see Cajori's 1929 article *Generalizations in geometry as seen in the history of developable surfaces* [66]. Theorem 8.4.1 appeared in Ehresmann's 1936 paper *Sur les espaces localement homogènes* [115]. The developing map and holonomy homomorphism for locally homogeneous Riemannian manifolds were described by Cartan in his 1926 paper [69].

§8.5. The concept of metric completeness was introduced by Fréchet in his 1906 paper *Sur quelques points du calcul fonctionnel* [137]. For the history of metric completeness, see Dugac's 1984 article *Histoire des espaces complets* [109]. Theorem 8.5.4 for the Euclidean plane was proved by Euler in his 1795 paper *De centro similitudinis* [129]. Theorems 8.5.5, 8.5.7, and 8.5.10 for Riemannian surfaces were proved by Hopf and Rinow in their 1931 paper *Ueber den Begriff der vollständigen differentialgeometrischen Fläche* [199] and were extended to Riemannian n-manifolds by Whitehead in his 1935 paper *On the covering of a complete space by the geodesics through a point* [410]. See also Cohn-Vossen's 1935 paper *Existenz kürzester Wege* [83]. Theorem 8.5.9 for spherical, Euclidean, or hyperbolic manifolds was proved by Hopf in his 1926 paper [198] and was extended to locally homogeneous Riemannian manifolds by Whitehead in his 1932 paper *Locally homogeneous spaces in differential geometry* [409].

§8.6. Berestovskii proved his theorem on geometric spaces in his 1982 paper *Homogeneous Busemann G-spaces* [41]. The notion of an n-dimensional geometry originated in Riemann's concept of a manifold of constant curvature which he introduced in his 1854 lecture [349]. For a discussion, see von Helmholtz's 1876 paper *On the origin and significance of geometrical axioms* [399]. The notion of an n-dimensional geometry was developed by Killing, Lie, and Cartan in their work on Lie groups. For a discussion, see Cartan's 1936 article *Le rôle de la théorie des groupes de Lie dans l'évolution de la géométrie moderne* [70]. Theorem 8.6.2 appeared in Riemann's 1854 lecture [349]. For a proof, see Vol. II of Spivak's 1979 treatise *Differential Geometry* [378]. Thurston's theorem on 3-dimensional geometries appeared in his 1982 article *Three dimensional manifolds, Kleinian groups, and hyperbolic geometry* [390]. For a discussion, see Scott's 1984 survey *The geometries of 3-manifolds* [370]. The 4-dimensional geometries are described in Wall's 1985 paper *Geometries and geometric structures in real dimension 4 and complex dimension 2* [400].

CHAPTER 9

Geometric Surfaces

In this chapter, we study the geometry of geometric surfaces. The chapter begins with a review of the topology of compact surfaces. In Section 9.2, a geometric method for constructing spherical, Euclidean, and hyperbolic surfaces is given. The fundamental relationship between the Euler characteristic of a closed geometric surface and its area is derived in Section 9.3. In Section 9.4, the set of similarity equivalence classes of Euclidean or hyperbolic structures on a closed surface is shown to have a natural topology. The geometry of closed geometric surfaces is studied in Sections 9.5 and 9.6. The chapter ends with a study of the geometry of complete hyperbolic surfaces of finite area.

§9.1. Compact Surfaces

A *surface* is a connected 2-dimensional manifold. A compact surface is called a *closed surface.*

Definition: A *triangulation* of a closed surface M consists of a finite family of functions

$$\{\phi_i : \Delta^2 \to M\}_{i=1}^m$$

with the following properties:

(1) The function ϕ_i maps the standard 2-simplex Δ^2 homeomorphically onto a subset T_i of M, called a *triangle.* The *vertices* and *edges* of T_i are the images of the vertices and edges of Δ^2 under ϕ_i.

(2) The surface M is the union of the triangles $T_1, \ldots, T_m$.

(3) If $i \neq j$, then the intersection of T_i and T_j is either empty, a common vertex of each triangle, or a common edge of each triangle.

Figure 7.2.1 illustrates four different triangulations of S^2. It is a fundamental theorem of the topology of surfaces that every closed surface has a triangulation. Given a triangulation of a closed surface M, let v be the number of vertices, e the number of edges, and t the number of triangles. The *Euler characteristic* of M is the integer

$$\chi(M) = v - e + t. \tag{9.1.1}$$

It is a basic theorem of algebraic topology that $\chi(M)$ does not depend on the choice of the triangulation. More generally, if M is a cell complex with a 0-cells, b 1-cells, and c 2-cells, then

$$\chi(M) = a - b + c. \tag{9.1.2}$$

If M_1 and M_2 are surfaces, then we can form a new surface $M_1\#M_2$, called the *connected sum* of M_1 and M_2, as follows: Let $\phi_i : \Delta^2 \to M_i$, for $i = 1, 2$, be a function that maps Δ^2 homeomorphically into M_i and set

$$M_i' = M - \phi_i(\operatorname{Int}\Delta^2)$$

for $i = 1, 2$. Then $M_1\#M_2$ is defined to be the quotient space of the disjoint union $M_1' \coprod M_2'$ obtained by identifying $\phi_1(x)$ with $\phi_2(x)$ for each x in $\partial\Delta^2$. The topological type of $M_1\#M_2$ does not depend on the choice of the functions ϕ_1 and ϕ_2. Evidently, if M_1 and M_2 are closed, then

$$\chi(M_1\#M_2) = \chi(M_1) + \chi(M_2) - 2, \tag{9.1.3}$$

since we can choose ϕ_1 and ϕ_2 to be part of triangulations of M_1 and M_2.

Starting from the fact that closed surfaces can be triangulated, it is not difficult to classify all closed surfaces up to homeomorphism. The classification of closed surfaces is summarized in the following theorem.

Theorem 9.1.1. *A closed surface is homeomorphic to either a sphere, a connected sum of tori, or a connected sum of projective planes.*

Orientability

Let $\{\phi_i : \Delta^2 \to M\}_{i=1}^m$ be a triangulation of a closed surface M. Orient the standard 2-simplex Δ^2 with the positive orientation from E^2. Then ϕ_i orients the triangle $T_i = \phi_i(\Delta^2)$ for each i. In particular, ϕ_i orients each of the three edges of T_i. A triangulation of M is said to be *oriented* if and only if each edge of the triangulation receives opposite orientations from the two adjacent triangles of which it is an edge. See Figure 9.1.1.

Let ρ be the reflection of Δ^2 in the line $y = x$. Then ρ reverses the orientation of Δ^2. A triangulation $\{\phi_i : \Delta^2 \to M\}_{i=1}^m$ for M is said to be *orientable* if and only if an oriented triangulation of M can be obtained from $\{\phi_i\}_{i=1}^m$ by replacing each ϕ_i by ϕ_i or $\phi_i\rho$. The surface M is said to be *orientable* if and only if it has an orientable triangulation. It is a basic theorem of algebraic topology that a closed surface M is orientable

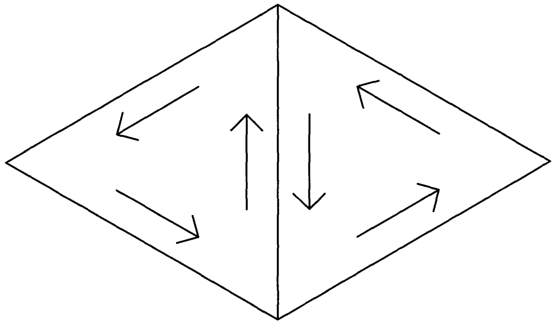


Figure 9.1.1. Adjacent oriented triangles with compatible orientations

if and only if every triangulation of M is orientable. Furthermore, a closed surface is orientable if and only if it is either a sphere or a connected sum of tori.

A connected sum of n tori is called a closed orientable surface of *genus* n. A 2-sphere is also called a closed orientable surface of *genus zero.* The relationship between the Euler characteristic of a closed orientable surface M and its genus is given by the formula

$$\chi(M) = 2(1 - \text{genus}(M)). \tag{9.1.4}$$

A connected sum of n projective planes is called a closed nonorientable surface of *genus* n. A closed nonorientable surface of genus two is also called a *Klein bottle.* The relationship between the Euler characteristic of a closed nonorientable surface M and its genus is given by the formula

$$\chi(M) = 2 - \text{genus}(M). \tag{9.1.5}$$

The next theorem states that the Euler characteristic and orientability form a complete set of topological invariants for the classification of closed surfaces.

Theorem 9.1.2. *Two closed surfaces are homeomorphic if and only if they have the same Euler characteristic and both are orientable or both are nonorientable.*

Surfaces-with-boundary

A *surface-with-boundary* is a connected 2-manifold-with-boundary. Let M be a compact surface-with-boundary. The boundary ∂M of M is a disjoint union of a finite number of topological circles. Let M^* be the closed surface obtained from M by gluing a disk along its boundary to each boundary circle of M. We now state the classification theorem for compact surfaces-with-boundary.

Theorem 9.1.3. *Two compact surfaces-with-boundary M_1 and M_2 are homeomorphic if and only if they both have the same number of boundary components and the closed surfaces M_1^* and M_2^*, obtained from M_1 and M_2 by gluing a disk to each boundary component, are homeomorphic.*

Triangulations and the *Euler characteristic* of a compact surface-with-boundary M are defined in the same way as for closed surfaces. If M has m boundary components, then the relationship between the Euler characteristics of M and M^* is given by the formula

$$\chi(M^*) = \chi(M) + m. \tag{9.1.6}$$

A compact surface-with-boundary M is said to be *orientable* if and only if the closed surface M^* is orientable. The next theorem follows from Theorems 9.1.2 and 9.1.3.

Theorem 9.1.4. *Two compact surfaces-with-boundary are homeomorphic if and only if they have the same number of boundary components, the same Euler characteristic, and both are orientable or both are nonorientable.*

§9.2. Gluing Surfaces

In this section, we construct spherical, Euclidean, and hyperbolic surfaces by gluing together convex polygons in $X = S^2, E^2$, or H^2 along their sides.

Let $\mathcal{P}$ be a finite family of disjoint convex polygons in X and let G be a group of isometries of X.

Definition: A *G-side-pairing* for $\mathcal{P}$ is a subset of G,

$$\Phi = \{g_S : S \in \mathcal{S}\},$$

indexed by the collection $\mathcal{S}$ of all the sides of the polygons in $\mathcal{P}$ such that for each side S in $\mathcal{S}$,

(1) there is a side S' in $\mathcal{S}$ such that $g_S(S') = S$;

(2) the isometries g_S and $g_{S'}$ satisfy the relation $g_{S'} = g_S^{-1}$; and

(3) if S is a side of P in $\mathcal{P}$ and S' is a side of P' in $\mathcal{P}$, then
$$P \cap g_S(P') = S.$$

It follows from (1) that S' is uniquely determined by S. The side S' is said to be *paired to* the side S by Φ. From (2), we deduce that $S'' = S$. Thus, the mapping $S \mapsto S'$ is an involution of the set $\mathcal{S}$. It follows from (3) that $g_S \neq 1$ for all S.

Let $\Phi = \{g_S : S \in \mathcal{S}\}$ be a G-side-pairing for $\mathcal{P}$ and set

$$\Pi = \bigcup_{P \in \mathcal{P}} P.$$

Two points x, x' of Π are said to be *paired* by Φ, written $x \simeq x'$, if and only if there is a side S in $\mathcal{S}$ such that x is in S, and x' is in S', and $g_S(x') = x$. If $g_S(x') = x$, then $g_{S'}(x) = x'$. Therefore $x \simeq x'$ if and only if $x' \simeq x$.

Two points x, y of Π are said to be *related* by Φ, written $x \sim y$, if and only if either $x = y$ or there is a finite sequence $x_1, \ldots, x_m$ of points of Π such that

$$x = x_1 \simeq x_2 \simeq \cdots \simeq x_m = y.$$

Being related by Φ is obviously an equivalence relation on the set Π. The equivalence classes of Π are called the *cycles* of Φ. If x is in Π, we denote the cycle of Φ containing x by $[x]$.

Let

$$[x] = \{x_1, \ldots, x_m\}$$

be a finite cycle of Φ. Let P_i be the polygon in $\mathcal{P}$ containing the point x_i and let θ_i be the angle subtended by P_i at the point x_i for each $i = 1, \ldots, m$. The *angle sum* of $[x]$ is defined to be the real number

$$\theta[x] = \theta_1 + \cdots + \theta_m.$$

Definition: A G-side-pairing Φ for $\mathcal{P}$ is *proper* if and only if each cycle of Φ is finite and has angle sum 2π.

Example 1. Let P be a closed hemisphere in S^2. Pair ∂P to itself by the antipodal map α of S^2. Then each point x in P° forms a cycle whose angle sum is 2π, and each pair of antipodal points x, x' in ∂P form a cycle whose angle sum is 2π. Therefore, this $\{I, \alpha\}$-side-pairing is proper.

Example 2. Let P be a rectangle in E^2. Pair the opposite sides of P by translations. Then each point x in P° forms a cycle whose angle sum is 2π. See Figure 9.2.1(a). Each pair of points x, x' directly across from each other in the interior of opposite sides forms a cycle whose angle sum is 2π. See Figure 9.2.1(b). Finally, the four vertices x_1, x_2, x_3, x_4 of P form a cycle whose angle sum is 2π. See Figure 9.2.1(c). Therefore, this $\mathrm{T}(E^2)$-side-pairing is proper.

Example 3. Let P be an exact fundamental polygon for a discrete group Γ of isometries of X acting freely on X. For each side S of P, there is a unique element g_S of Γ such that $P \cap g_S P = S$. Then

$$\Phi = \{g_S : S \text{ is a side of } P\}$$

is a proper Γ-side-pairing by Theorems 6.7.5 and 6.7.7.

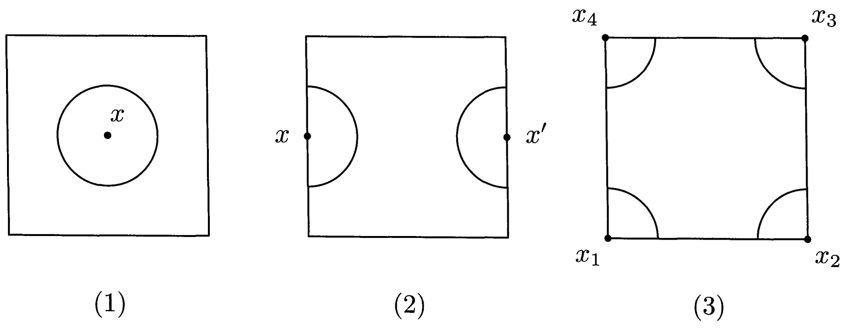


Figure 9.2.1. Cycles in a rectangle

Theorem 9.2.1. *If $\Phi = \{g_S : S \in \mathcal{S}\}$ is a proper G-side-pairing for $\mathcal{P}$, then for each side S in $\mathcal{S}$,*

1. *the isometry g_S fixes no point of S'; and*
2. *the sides S and S' are equal if and only if S is a great circle of S^2 and g_S is the antipodal map of S^2.*

Proof: (1) On the contrary, suppose that g_S fixes a point x of S'. Assume first that x is in the interior of S'. Then $[x] = \{x\}$ and $\theta[x] = \pi$, which is a contradiction. Assume now that x is an endpoint of S'. Then x is an endpoint of exactly one other side T in $\mathcal{S}$. As $g_S(S') = S$, we have that x is in S, and so either $S = S'$ or $S = T$. If $S = S'$, then g_S would fix S pointwise, contrary to the first case; therefore $S = T$. Then $[x] = \{x\}$ and $\theta[x] < \pi$, which is a contradiction. Thus g_S fixes no point of S'.

(2) If S is a great circle and g_S is the antipodal map of S^2, then

$$S' = g_S^{-1}(S) = S.$$

Conversely, suppose that $S' = S$. As $g_{S'} = g_S^{-1}$, we have that g_S has order two. Let x be a point of S. Then $x' = g_S(x)$ is also a point of S. If x and x' were not antipodal points, then g_S would fix the midpoint of the geodesic segment joining x to x' in S contrary to (1). Therefore x and x' are antipodal points of S^2. Hence S is invariant under the antipodal map of S^2, and so S must be a great circle. Hence, the polygon P in $\mathcal{P}$ containing S is a hemisphere. As g_S is the antipodal map on S and $P \cap g_S(P) = S$, we have that g_S is the antipodal map of S^2. □

Let Φ be a proper G-side-pairing for $\mathcal{P}$. Then Π is the topological sum of the polygons in $\mathcal{P}$, since $\mathcal{P}$ is a finite family of disjoint closed subsets of X. Let M be the quotient space of Π of cycles of Φ. The space M is said to be obtained by gluing together the polygons in $\mathcal{P}$ by Φ. We next prove the gluing theorem for geometric surfaces.

Theorem 9.2.2. *Let G be a group of isometries of X and let M be a space obtained by gluing together a finite family $\mathcal{P}$ of disjoint convex polygons in X by a proper G-side-pairing Φ. Then M is a 2-manifold with an (X,G)-structure such that the natural injection of P° into M is an (X,G)-map for each P in $\mathcal{P}$.*

Proof: Without loss of generality, we may assume that each polygon in $\mathcal{P}$ has at least one side. Let $\pi : \Pi \to M$ be the quotient map and let x be a point of Π. We now construct an open neighborhood $U(x,r)$ of $\pi(x)$ in M and a homeomorphism

$$\phi_x : U(x,r) \to B(x,r)$$

for all sufficiently small values of r.

Let P be the polygon in $\mathcal{P}$ containing x. There are three cases to consider. Either (1) x is in P°, or (2) x is in the interior of a side S of P, or (3) x is a vertex of P. See Figure 9.2.1. If x is in P°, then $[x] = \{x\}$. If x is in the interior of a side of P, then $[x] = \{x, x'\}$, with $x \neq x'$, since Φ is proper. If x is a vertex of P, then x is the endpoint of exactly two sides of P, and so x is paired to exactly two other points of Π, since Φ is proper. In this case, each element of $[x]$ is paired to exactly two other elements of $[x]$. Thus, in all three cases, the cycle $[x]$ can be ordered

$$[x] = \{x_1, x_2, \ldots, x_m\}$$

so that

$$x = x_1 \simeq x_2 \simeq \cdots \simeq x_m \simeq x.$$

Moreover, if $m > 1$, then there is a unique side S_i in $\mathcal{S}$ such that $g_{S_i}(x_{i+1}) = x_i$ for $i = 1, \ldots, m-1$, and $g_{S_m}(x_1) = x_m$.

Let $g_1 = 1$ and $g_i = g_{S_1} \cdots g_{S_{i-1}}$ for $i = 2, \ldots, m$. Then $g_i x_i = x$ for each i. Let P_i be the polygon in $\mathcal{P}$ containing the point x_i for each i. Let r be a positive real number such that r is less than one-third the distance from x_i to x_j for each $i \neq j$ and from x_i to any side of P_i not containing x_i for each i. Then the sets $P_i \cap B(x_i, r)$, for $i = 1, \ldots, m$, are disjoint.

Let θ_i be the angle subtended by P_i at the point x_i. Then $P_i \cap B(x_i, r)$ is a sector of the open disk $B(x_i, r)$ whose angular measure is θ_i. Hence

$$g_i(P_i \cap B(x_i, r)) = g_i P_i \cap B(x, r)$$

is a sector of the open disk $B(x,r)$ whose angular measure is θ_i. If $m = 1$, then

$$B(x,r) \;=\; P \cap B(x,r) = g_1 P_1 \cap B(x,r).$$

If $m = 2$, then

$$\begin{aligned} B(x,r) &= (P \cap B(x,r)) \cup (g_{S_1} P_2 \cap B(x,r)) \\ &= (g_1 P_1 \cap B(x,r)) \cup (g_2 P_2 \cap B(x,r)). \end{aligned}$$

Now assume that $m > 2$. Observe that the polygons P_i and $g_{S_i}(P_{i+1})$ lie on opposite sides of their common side S_i, and so the polygons $g_i P_i$ and

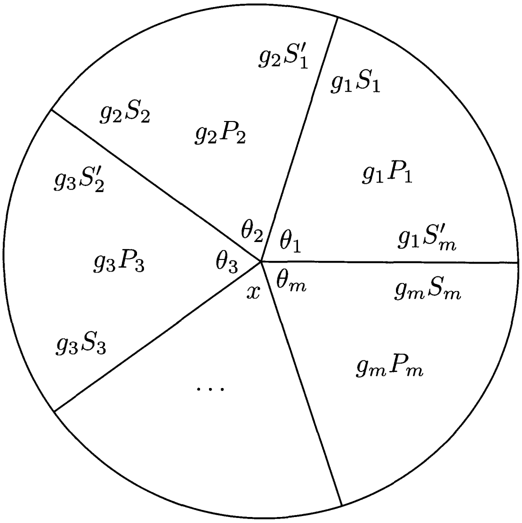


Figure 9.2.2. The partition of $B(x,r)$ into sectors by a proper side-pairing

$g_{i+1}P_{i+1}$ lie on opposite sides of their common side g_iS_i for $i = 1,\dots,m-1$. As $S_i = g_{S_i}(S_i')$ for $i = 1,\dots,m$, we have that $g_iS_i = g_{i+1}S_i'$ for $i = 1,\dots,m-1$. Now S_i and S_{i-1}' are the two sides of P_i whose endpoint is x_i for $i = 2,\dots,m$, and so g_iS_i and $g_iS_{i-1}' = g_{i-1}S_{i-1}$ are the two sides of g_iP_i whose endpoint is x for $i = 2,\dots,m$. Therefore, the sectors $g_iP_i \cap B(x,r)$, for $i = 1,\dots,m$, occur in sequential order rotating about the point x. See Figure 9.2.2. Since $\theta[x] = 2\pi$, we have

$$B(x,r) = \bigcup_{i=1}^{m} \big(g_iP_i \cap B(x,r)\big).$$

The polygons P_m and $g_{S_m}(P)$ lie on opposite sides of their common side S_m, and so the polygons $g_{S_m}^{-1}(P_m)$ and P lie on opposite sides of their common side S_m'. Now as S_1 and S_m' are the two sides of P whose endpoint is x, we deduce that

$$g_mP_m = g_{S_m}^{-1}P_m.$$

Therefore $g_m = g_{S_m}^{-1}$. Hence, we have the cycle relation $g_{S_1}\cdots g_{S_m} = 1$.

In all three cases, let

$$U(x,r) = \pi\Big(\bigcup_{i=1}^{m} P_i \cap B(x_i,r)\Big).$$

Now as the set

$$\pi^{-1}(U(x,r)) = \bigcup_{i=1}^{m} P_i \cap B(x_i,r)$$

is open in Π, we have that $U(x,r)$ is an open subset of M.

Define a function

$$\psi_x : \bigcup_{i=1}^{m} P_i \cap B(x_i, r) \to B(x, r)$$

by $\psi_x(z) = g_i z$ if z is in $P_i \cap B(x_i, r)$. Then ψ_x induces a continuous function

$$\phi_x : U(x, r) \to B(x, r).$$

The function ϕ_x is a bijection with a continuous inverse defined by

$$\phi_x^{-1}(z) = \pi(g_i^{-1} z) \quad \text{if } z \text{ is in } g_i P_i \cap B(x, r).$$

Hence ϕ_x is a homeomorphism.

Next, we show that M is Hausdorff. Let x and y be points of Π such that $\pi(x)$ and $\pi(y)$ are distinct points of M. Let $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_n\}$ be the cycles of Φ containing x and y, respectively. Then $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_n\}$ are disjoint subsets of Π. Let P_i be the polygon in $\mathcal{P}$ containing x_i for $i = 1, \ldots, m$, and let Q_j be the polygon in $\mathcal{P}$ containing y_j for $j = 1, \ldots, n$. Then we can choose radii r and s as before so that

$$\pi\Big(\bigcup_{i=1}^{m} P_i \cap B(x_i, r)\Big) = U(x, r)$$

and

$$\pi\Big(\bigcup_{j=1}^{n} Q_j \cap B(y_j, s)\Big) = U(y, s).$$

Moreover, we can choose r and s small enough so that

$$\bigcup_{i=1}^{m} P_i \cap B(x_i, r) \quad \text{and} \quad \bigcup_{j=1}^{n} Q_j \cap B(y_j, s)$$

are disjoint subsets of Π. As

$$\bigcup_{i=1}^{m} P_i \cap B(x_i, r) = \pi^{-1}(U(x, r))$$

and

$$\bigcup_{j=1}^{n} Q_j \cap B(y_j, s) = \pi^{-1}(U(y, s)),$$

we deduce that $U(x, r)$ and $U(y, r)$ are disjoint open neighborhoods of $\pi(x)$ and $\pi(y)$ in M. Thus M is Hausdorff, and therefore M is a 2-manifold.

Next, we show that

$$\{\phi_x : U(x, r) \to B(x, r)\}$$

is an (X, G)-atlas for M. By construction, $U(x, r)$ is an open connected subset of M and ϕ_x is a homeomorphism. Moreover $U(x, r)$ is defined for each point $\pi(x)$ of M and sufficiently small radius r. Hence $\{U(x, r)\}$ is an open cover of M. It remains only to show that if $U(x, r)$ and $U(y, s)$ overlap, then the coordinate change

$$\phi_y \phi_x^{-1} : \phi_x\big(U(x, r) \cap U(y, s)\big) \to \phi_y\big(U(x, r) \cap U(y, s)\big)$$

agrees in a neighborhood of each point of its domain with an element of G.

As before, we have

$$\pi^{-1}(U(x,r)) = \bigcup_{i=1}^{m} P_i \cap B(x_i, r),$$
$$\pi^{-1}(U(y,s)) = \bigcup_{j=1}^{n} Q_j \cap B(y_j, s).$$

By reversing the roles of x and y, if necessary, we may assume that $m \leq n$. If $m > 1$, let S_i be the side of P_i containing x_i as before, and if $n > 1$, let T_j be the side of Q_j containing y_j as before. Let $g_1, \ldots, g_m$ and $h_1, \ldots, h_n$ be the elements of G constructed as before for x and y. Because of the $1/3$ bounds on r and s, there is only one index j, say ℓ, such that the set

$$P \cap B(x,r) \cap Q_j \cap B(y_j, s)$$

is nonempty. We shall prove that the coordinate change $\phi_y \phi_x^{-1}$ is the restriction of the element h_ℓ of G.

Assume first that $m = 1$. Then x is in P° and

$$\pi^{-1}(U(x,r)) = B(x,r).$$

Therefore

$$\begin{aligned} & U(x,r) \cap U(y,s) \\ = \quad & \pi(B(x,r)) \cap \pi\Big(\bigcup_{j=1}^{n} Q_j \cap B(y_j,s)\Big) \\ = \quad & \pi\big(B(x,r) \cap \bigcup_{j=1}^{n} Q_j \cap B(y_j,s)\big) \\ = \quad & \pi\big(B(x,r) \cap B(y_\ell,s)\big). \end{aligned}$$

Hence

$$\phi_x\big(U(x,r) \cap U(y,s)\big) = B(x,r) \cap B(y_\ell, s)$$

and

$$\phi_y\big(U(x,r) \cap U(y,s)\big) = h_\ell\big(B(x,r) \cap B(y_\ell, s)\big).$$

Therefore, the coordinate change

$$\phi_y \phi_x^{-1} : B(x,r) \cap B(y_\ell, s) \to h_\ell\big(B(x,r) \cap B(y_\ell, s)\big)$$

is the restriction of h_ℓ.

Assume next that $m = 2$. Then x is in the interior of a side S of P and x' is in the interior of a side S' of P' and the set

$$P' \cap B(x',r) \cap Q_j \cap B(y_j, s)$$

is nonempty only for $j = \ell - 1$ or $\ell + 1 \pmod{n}$. By reversing the ordering of $y_1, \ldots, y_n$, if necessary, we may assume that this intersection is nonempty only for $j = \ell + 1$. Then $P = Q_\ell$, $P' = Q_{\ell+1}$, $S = T_\ell$, and

$$\begin{aligned} & U(x,r) \cap U(y,s) \\ = \quad & \pi\big[\big(P \cap B(x,r)\big) \cup \big(P' \cap B(x',r)\big)\big] \cap \pi\big[\bigcup_{j=1}^{n} Q_j \cap B(y_j,s)\big] \\ = \quad & \pi\big[\bigcup_{j=1}^{n} P \cap B(x,r) \cap Q_j \cap B(y_j,s) \cup \bigcup_{j=1}^{n} P' \cap B(x',r) \cap Q_j \cap B(y_j,s)\big] \\ = \quad & \pi\big[\big(P \cap B(x,r) \cap B(y_\ell,s)\big) \cup \big(P' \cap B(x',r) \cap B(y_{\ell+1},s)\big)\big]. \end{aligned}$$

Hence

$$\begin{aligned}
&\phi_x\big(U(x,r)\cap U(y,s)\big)\\
&\quad= \big(P\cap B(x,r)\cap B(y_\ell,s)\big)\cup g_S\big(P'\cap B(x',r)\cap B(y_{\ell+1},s)\big)\\
&\quad= \big(P\cap B(x,r)\cap B(y_\ell,s)\big)\cup \big(g_S(P')\cap B(x,r)\cap B(y_\ell,s)\big)\\
&\quad= B(x,r)\cap B(y_\ell,s)
\end{aligned}$$

and

$$\begin{aligned}
&\phi_y\big(U(x,r)\cap U(y,s)\big)\\
&\quad= h_\ell\big(P\cap B(x,r)\cap B(y_\ell,s)\big)\cup h_{\ell+1}\big(P'\cap B(x',r)\cap B(y_{\ell+1},s)\big)\\
&\quad= h_\ell\big[\big(P\cap B(x,r)\cap B(y_\ell,s)\big)\cup g_S\big(P'\cap B(x',r)\cap B(y_{\ell+1},s)\big)\big]\\
&\quad= h_\ell\big[\big(P\cap B(x,r)\cap B(y_\ell,s)\big)\cup \big(g_S(P')\cap B(x,r)\cap B(y_\ell s)\big)\big]\\
&\quad= h_\ell\big(B(x,r)\cap B(y_\ell,s)\big).
\end{aligned}$$

Now on the set

$$P\cap B(x,r)\cap B(y_\ell,s),$$

the map $\phi_y\phi_x^{-1}$ is the restriction of h_ℓ, and on the set

$$g_S\big(P'\cap B(x',r)\cap B(y_{\ell+1},s)\big),$$

the map $\phi_y\phi_x^{-1}$ is the restriction of $h_{\ell+1}g_S^{-1}=h_\ell$. Hence, the coordinate change

$$\phi_y\phi_x^{-1}: B(x,r)\cap B(y_\ell,s)\to h_\ell\big(B(x,r)\cap B(y_\ell,s)\big)$$

is the restriction of h_ℓ.

Assume now that $m>2$. Then both x and y are vertices. As $U(x,r)$ and $U(y,s)$ overlap, $\pi(x)=\pi(y)$ because of the bounds on r and s. Hence $x=y_\ell$. Let $t=\min\{r,s\}$. Then

$$\begin{aligned}
U(x,r)\cap U(y,s) &= U(x,t),\\
\phi_x(U(x,t)) &= B(x,t),\\
\phi_y(U(x,t)) &= B(y,t).
\end{aligned}$$

Now either

$$x_\imath = y_{\ell+\imath-1}\ (\mathrm{mod}\ m)$$

or

$$x_\imath = y_{\ell-\imath-1}\ (\mathrm{mod}\ m).$$

By reversing the ordering of $y_1,\ldots,y_m$, if necessary, we may assume that the former holds. Then

$$P_\imath = Q_{\ell+\imath-1}\ (\mathrm{mod}\ m)$$

and

$$S_\imath = T_{\ell+\imath-1}\ (\mathrm{mod}\ m).$$

Now observe that

$$\begin{aligned} g_\imath &= g_{S_1} \cdots g_{S_{\imath-1}} \\ &= g_{T_\ell} \cdots g_{T_{\ell+\imath-2}} \\ &= h_\ell^{-1} h_{\ell+\imath-1} \pmod m \end{aligned}$$

and so we have

$$h_{\ell+\imath-1} = h_\ell g_\imath \pmod m.$$

Now as

$$B(x,t) = \bigcup_{\imath=1}^{m} g_\imath P_\imath \cap B(x,t),$$

the map $\phi_y \phi_x^{-1}$ is the restriction of

$$h_{\ell+\imath-1} g_\imath^{-1} = (h_\ell g_\imath) g_\imath^{-1} = h_\ell$$

on the set $g_\imath P_\imath \cap B(x,t)$ for each $i = 1, \ldots, m$. Hence, the coordinate change

$$\phi_y \phi_x^{-1} : B(x,t) \to B(y,t)$$

is the restriction of h_ℓ. Thus, in all three cases, $\phi_y \phi_x^{-1}$ agrees with an element of G. This completes the proof that $\{\phi_x\}$ is an (X,G)-atlas for M.

Let P be a polygon in $\mathcal{P}$ and let $\iota : P^\circ \to M$ be the natural injection of P° into M. Then for each point x in P° and chart $\phi_x : U(x,r) \to B(x,r)$, the map

$$\iota^{-1} : \iota B(x,r) \to B(x,r)$$

is ϕ_x. Therefore ι is an (X,G)-map by Theorem 8.4.2. Thus, the (X,G)-structure of M has the property that the natural injection of P° into M is an (X,G)-map for each P in $\mathcal{P}$. □

Example 4. Let n be an integer greater than one. Then

$$\frac{\pi}{2n} + \frac{\pi}{4n} + \frac{\pi}{4n} = \frac{\pi}{n} < \pi.$$

Hence, there is a hyperbolic triangle of the form $\triangle\left(\frac{\pi}{2n}, \frac{\pi}{4n}, \frac{\pi}{4n}\right)$ by Theorem 3.5.9. Now reflecting in the sides of $\triangle$, keeping the vertex whose angle is $\pi/2n$ fixed, generates a cycle of $4n$ hyperbolic triangles whose union is a regular hyperbolic $4n$-gon P whose dihedral angle is $\pi/2n$. We position P in B^2 so that its center is the origin. See Figure 9.2.3.

Now label the sides of P in positive order by the symbols

$$S_1, T_1, S_1', T_1', \ldots, S_n, T_n, S_n', T_n'$$

as in Figure 9.2.3. The side $S_\imath'$ is paired to the side $S_\imath$ by first reflecting in the straight line passing through the origin and the center of the side labeled $T_\imath$, and then reflecting in the side of P labeled $S_\imath$. The side $T_\imath'$ is paired to the side $T_\imath$ by first reflecting in the straight line passing through the origin and the center of the side labeled $S_\imath'$, and then reflecting in the side of P labeled $T_\imath$. The $4n$ vertices of P form a cycle whose angle sum is 2π. Therefore, this side-pairing is proper.

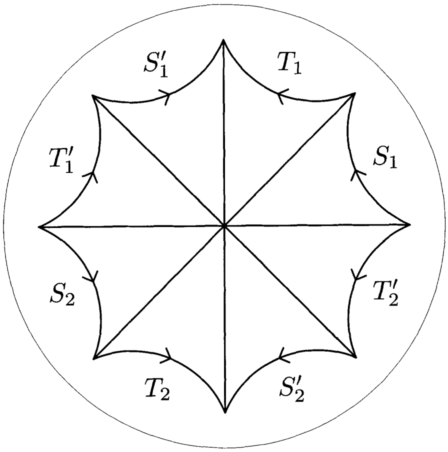


Figure 9.2.3. A regular hyperbolic octagon

Let M be the space obtained from P by gluing together its sides by this side-pairing. Then M is a closed surface with a $(B^2, \mathrm{I}_0(B^2))$-structure by Theorem 9.2.2. It is evident from the gluing pattern of P that M is a connected sum of n tori. Thus M is a closed orientable surface of genus $n > 1$.

Example 5. Let n be an integer greater than two. Then

$$\frac{\pi}{n} + \frac{\pi}{2n} + \frac{\pi}{2n} = \frac{2\pi}{n} < \pi.$$

Hence, there is a hyperbolic triangle of the form $\triangle(\frac{\pi}{n}, \frac{\pi}{2n}, \frac{\pi}{2n})$ by Theorem 3.5.9. Now reflecting in the sides of $\triangle$, keeping the vertex whose angle is π/n fixed, generates a cycle of $2n$ hyperbolic triangles whose union is a regular hyperbolic $2n$-gon Q whose dihedral angle is π/n. We position Q in B^2 so that its center is the origin.

We now divide the sides of Q into pairs of consecutive sides. Each of these pairs of consecutive sides of Q are paired by a rotation about the origin followed by the reflection in the corresponding side of Q. The $2n$ vertices of Q form a cycle whose angle sum is 2π. Therefore, this side-pairing is proper.

Let M be the space obtained from Q by gluing together its sides by this side-pairing. Then M is a closed surface with a $(B^2, \mathrm{I}(B^2))$-structure by Theorem 9.2.2. It is evident from the gluing pattern of Q that M is a connected sum of n projective planes. Thus M is a closed nonorientable surface of genus $n > 2$.

The Generalized Gluing Theorem

In later applications, we shall need a more general version of Theorem 9.2.2. The first step towards this generalized gluing theorem is to generalize the notion of a convex polygon so as to allow vertices in the interior of a side.

Definition: An *abstract convex polygon* P in X is a convex polygon P in X together with a collection $\mathcal{E}$ of subsets of ∂P, called the *edges* of P, such that

(1) each edge of P is a closed, 1-dimensional, convex subset of ∂P;

(2) two edges of P meet only along their boundaries;

(3) the union of the edges of P is ∂P;

(4) the collection $\mathcal{E}$ is a locally finite family of subsets of X.

By Theorem 6.2.6, a convex polygon P in X, together with the collection $\mathcal{S}$ of its sides, is an abstract convex polygon. Note that, in general, an edge of an abstract convex polygon P may or may not be equal to the side of P containing it. The *vertices* of an abstract convex polygon P are defined to be the endpoints of the edges of P. A vertex of an abstract convex polygon P may be in the interior of a side of P.

We next generalize the notion of a disjoint set of convex polygons so as to allow the possibility that the polygons may live in different copies of X.

Definition: A *disjoint set of abstract convex polygons* of X is a set of functions

$$\Xi = \{\xi_P : P \in \mathcal{P}\}$$

indexed by a set $\mathcal{P}$ such that

(1) the function $\xi_P : X \to X_P$ is a similarity for each P in $\mathcal{P}$;

(2) the index P is an abstract convex polygon in X_P for each P in $\mathcal{P}$;

(3) the polygons in $\mathcal{P}$ are mutually disjoint.

Let Ξ be a disjoint set of abstract convex polygons of X and let G be a group of similarities of X.

Definition: A *G-edge-pairing* for Ξ is a set of functions

$$\Phi = \{\phi_E : E \in \mathcal{E}\}$$

indexed by the collection $\mathcal{E}$ of all the edges of the polygons in $\mathcal{P}$ such that for each edge E of a polygon P in $\mathcal{P}$,

(1) there is a polygon P' in $\mathcal{P}$ such that the function $\phi_E : X_{P'} \to X_P$ is a similarity;

(2) the similarity $\xi_P^{-1}\phi_E\xi_{P'}$ is in G;

(3) there is an edge E' of P' such that $\phi_E(E') = E$;

(4) the similarities ϕ_E and $\phi_{E'}$ satisfy the relation $\phi_{E'} = \phi_E^{-1}$;

(5) the polygons P and $\phi_E(P')$ are situated so that $P \cap \phi_E(P') = E$.

Let Φ be a G-edge-pairing for Ξ. Then the pairing of edge points by elements of Φ generates an equivalence relation on the set $\Pi = \cup_{P\in\mathcal{P}} P$. The equivalence classes are called the *cycles* of Φ, and Φ is said to be *proper* if and only if every cycle of Φ is finite and has angle sum 2π. Topologize Π with the direct sum topology and let M be the quotient space of Π of cycles of Φ. The space M is said to be obtained by gluing together the polygons of Ξ by Φ.

The proof of the next theorem follows the same outline as the proof of Theorem 9.2.2 and is therefore left to the reader.

Theorem 9.2.3. *Let G be a group of similarities of X and let M be a space obtained by gluing together a disjoint set Ξ of abstract convex polygons of X by a proper G-edge-pairing Φ. Then M is a 2-manifold with an (X,G)-structure such that the natural injection of P° into M is an (X,G)-map for each polygon P of Ξ.*

Exercise 9.2

1. With the same definitions as in the proof of Theorem 9.2.2, prove that for each index i, there is at most one index j such that the following set is nonempty:
$$P_i \cap B(x_i, r) \cap Q_j \cap B(y_j, s).$$

2. Show that the same gluing pattern on the sides of a square in E^2, as in Example 5, yields a Euclidean structure on the Klein bottle.

3. Let P be a convex fundamental polygon for a discrete group Γ of isometries of X and let $\mathcal{E}$ be the collection of all 1-dimensional convex subsets of ∂P of the form $P \cap gP$ for some g in Γ. Prove that P together with $\mathcal{E}$ is an abstract convex polygon in X.

4. Let P be as in Exercise 3. For each edge E of P, let g_E be the element of Γ such that $P \cap g_E(P) = E$. Prove that $\Phi = \{g_E : E \in \mathcal{E}\}$ is a Γ-edge-pairing for P.

5. Prove Theorem 9.2.3.

§9.3. The Gauss-Bonnet Theorem

We next prove the Gauss-Bonnet Theorem for closed geometric surfaces.

Theorem 9.3.1. *If $\kappa = 1, 0$, or -1 is the curvature of a closed spherical, Euclidean, or hyperbolic surface M, then*

$$\kappa\text{Area}(M) = 2\pi\chi(M).$$

Proof: As M is compact, M is complete. By Theorem 8.5.9, we may assume that M is a space-form X/Γ. Let P be an exact fundamental polygon for Γ. Then P is compact by Theorem 6.5.10.

If P has no sides, then $P = S^2 = M$ and

$$\text{Area}(M) = 4\pi = 2\pi\chi(M).$$

If P has one side, then P is a closed hemisphere of S^2, and so $M = P^2$ by Theorem 9.2.1(2), and

$$\text{Area}(M) = 2\pi = 2\pi\chi(M).$$

If P has two sides, then P is a lune of S^2, but any side-pairing of a lune is not proper. Therefore, we may assume that P has at least three sides. Then the 2nd barycentric subdivision of P subdivides P into triangles and projects to a triangulation of M so that each triangle of the subdivision of P is mapped homeomorphically onto a triangle of the triangulation.

Let $\triangle_1, \ldots, \triangle_t$ be the triangles of the 2nd barycentric subdivision of P. Then $e = 3t/2$ is the number of edges of the triangulation of M. Let v be the number of vertices of the triangulation of M. Then

$$\chi(M) = v - e + t = v - \tfrac{1}{2}t.$$

Suppose that $\kappa = 1$ or -1. Then by Theorems 2.5.5 and 3.5.5, we have

$$\begin{aligned}
\kappa\text{Area}(M) &= \kappa\text{Area}(P) \\
&= \kappa\sum_{i=1}^{t}\text{Area}(\triangle_i(\alpha_i, \beta_i, \gamma_i)) \\
&= \sum_{i=1}^{t}(\alpha_i + \beta_i + \gamma_i - \pi) \\
&= 2\pi v - t\pi \\
&= 2\pi(v - \tfrac{1}{2}t) \quad = \quad 2\pi\chi(M).
\end{aligned}$$

Now suppose that $\kappa = 0$. Then we have

$$2\pi v = \sum_{i=1}^{t}(\alpha_i + \beta_i + \gamma_i) = t\pi.$$

Hence, we have

$$\chi(M) = (v - \tfrac{1}{2}t) = 0.$$

Thus, we have

$$\kappa\text{Area}(M) = 2\pi\chi(M). \qquad \square$$

Theorem 9.3.2. *If M is a closed surface, then M has*

(1) *a spherical structure if and only if $\chi(M) > 0$,*

(2) *a Euclidean structure if and only if $\chi(M) = 0$,*

(3) *a hyperbolic structure if and only if $\chi(M) < 0$.*

Proof: (1) If $\chi(M) > 0$, then M is either a sphere or projective plane by Theorem 9.1.1, both of which have a spherical structure. Conversely, if M has a spherical structure, then $\chi(M) > 0$ by Theorem 9.3.1.

(2) If $\chi(M) = 0$, then M is either a torus or a Klein bottle by Theorem 9.1.1, both of which have a Euclidean structure. Conversely, if M has a Euclidean structure, then $\chi(M) = 0$ by Theorem 9.3.1.

(3) If $\chi(M) < 0$, then M is either a closed orientable surface of genus n, with $n > 1$, or a closed nonorientable surface of genus n, with $n > 2$, both of which have a hyperbolic structure by the constructions in Examples 4 and 5 in §9.2. Conversely, if M has a hyperbolic structure, then $\chi(M) < 0$ by Theorem 9.3.1. □

Exercise 9.3

1. Let T be a triangle in S^2, E^2, or H^2. Prove that the centroid of T is the intersection of the three geodesic segments joining a vertex of T to the midpoint of the opposite side of T.

2. Let P be a compact convex polygon in $X = S^2, E^2$, or H^2 with n sides and $n \geq 3$. Prove that the 2nd barycentric subdivision of P divides P into $12n$ triangles.

3. With the same definitions as in the proof of Theorem 9.3.1, prove that each triangle of the barycentric subdivision of P is mapped homeomorphically onto the image in M by the quotient map from X to M.

4. With the same definitions as in the proof of Theorem 9.3.1, prove that the 2nd barycentric subdivision of P projects to a triangulation of M.

§9.4. Moduli Spaces

Let M be a closed surface such that $\chi(M) \leq 0$. By Theorem 9.3.2, the surface M has a Euclidean or hyperbolic structure according as $\chi(M) = 0$ or $\chi(M) < 0$. In this section, we show that the set of similarity equivalence classes of Euclidean or hyperbolic structures on M has a natural topology.

If $\chi(M) = 0$, let $\mathcal{E}(M)$ be the *set of Euclidean structures* for M, and if $\chi(M) < 0$, let $\mathcal{H}(M)$ be the *set of hyperbolic structures* for M. Let $X = E^2$ or H^2 according as $\chi(M) = 0$ or $\chi(M) < 0$, and let $\mathcal{S}(M)$ be the set of *complete* $(X, \mathrm{S}(X))$*-structures* for M. We begin by studying the relationship between $\mathcal{S}(M)$ and $\mathcal{E}(M)$ or $\mathcal{H}(M)$. First of all, if $\chi(M) < 0$, then $\mathcal{S}(M) = \mathcal{H}(M)$, since $\mathrm{S}(H^2) = \mathrm{I}(H^2)$ and every hyperbolic structure for M is complete because M is compact. Thus, we may assume that $\chi(M) = 0$.

Define a left action of $\mathrm{S}(E^2)$ on $\mathcal{E}(M)$ as follows: If $\xi : E^2 \to E^2$ is a similarity and

$$\Phi = \{\phi_i : U_i \to E^2\}$$

is a Euclidean structure for M, define $\xi\Phi$ to be the Euclidean structure for M given by

$$\xi\Phi = \{\xi\phi_i : U_i \to E^2\}.$$

Clearly, $\mathrm{I}(E^2)$ acts trivially on $\mathcal{E}(M)$. Hence, the action of $\mathrm{S}(E^2)$ on $\mathcal{E}(M)$ induces an action of $\mathrm{S}(E^2)/\mathrm{I}(E^2)$ on $\mathcal{E}(M)$. The group $\mathrm{S}(E^2)/\mathrm{I}(E^2)$ is isomorphic to $\mathbb{R}_+$. Consequently, there is a corresponding action of $\mathbb{R}_+$ on $\mathcal{E}(M)$ defined as follows: If $k > 0$ and $\Phi = \{\phi_i : U_i \to E^2\}$ is in $\mathcal{E}(M)$, then

$$k\Phi = \{k\phi_i : U_i \to E^2\}.$$

Clearly, this action of $\mathbb{R}_+$ on $\mathcal{E}(M)$ is effective. Furthermore, we see that two elements of $\mathcal{E}(M)$ are in the same $\mathrm{S}(E^2)$-orbit if and only if they differ by a change of scale.

Given a Euclidean structure Φ for M, let $\hat{\Phi}$ be the unique complete $(E^2, \mathrm{S}(E^2))$-structure for M containing Φ.

Lemma 1. *If Φ is a Euclidean structure for M, then $\hat{\Phi}$ is the disjoint union of the Euclidean structures $\{k\Phi : k > 0\}$.*

Proof: Clearly, the Euclidean structures $\{k\Phi : k > 0\}$ are disjoint and

$$\cup\{k\Phi : k > 0\} \subset \hat{\Phi}.$$

Let $\phi : U \to E^2$ be an arbitrary chart in $\hat{\Phi}$. We shall prove that ϕ is in $k\Phi$ for some $k > 0$. Define a function $f : U \to \mathbb{R}_+$ as follows: For each point u of U, choose a chart $\phi_i : U_i \to E^2$ of Φ such that u is in U_i. Then $\phi\phi_i^{-1}$ agrees with an element g of $\mathrm{S}(E^2)$ in a neighborhood of u. Define $f(u)$ to be the scale factor of g. Observe that $f(u)$ does not depend on the choice of the chart ϕ_i, since if $\phi_j : U_j \to E^2$ is another chart in Φ such that u is in U_j, then

$$\phi\phi_j^{-1} = (\phi\phi_i^{-1})(\phi_i\phi_j^{-1})$$

in a neighborhood of u, and $\phi_i\phi_j^{-1}$ agrees with an isometry of E^2 in this neighborhood. It is clear from the definition of f that f is locally constant; therefore, f is constant, since U is connected.

Let k be the constant value of f. If $\phi_i : U \to E^2$ is a chart in Φ such that U and U_i overlap, then $k^{-1}\phi\phi_i^{-1}$ agrees with an element of $\mathrm{I}(E^2)$ in a neighborhood of each point of $\phi_i(U \cap U_i)$. Therefore $k^{-1}\phi$ is in Φ. Hence ϕ is in $k\Phi$. Thus

$$\hat{\Phi} = \dot{\cup}\{k\Phi : k > 0\}.$$

□

Theorem 9.4.1. *If M is a closed surface such that $\chi(M) = 0$, then the mapping $\Phi \mapsto \hat{\Phi}$ induces a bijection from $\mathrm{S}(E^2)\backslash\mathcal{E}(M)$ onto $\mathcal{S}(M)$.*

Proof: If ξ is an $\mathrm{S}(E^2)$ and Φ is in $\mathcal{E}(M)$, then $\widehat{\xi\Phi} = \hat{\Phi}$. Hence, the mapping $\Phi \mapsto \hat{\Phi}$ induces a function

$$\sigma : \mathrm{S}(E^2)\backslash\mathcal{E}(M) \to \mathcal{S}(M).$$

Suppose that Φ and Φ' are elements of $\mathcal{E}(M)$ such that $\hat{\Phi} = \hat{\Phi}'$. By Lemma 1, there is a $k > 0$ such that $\Phi' = k\Phi$. Hence Φ and Φ' are in the same $\mathrm{S}(E^2)$-orbit of $\mathcal{E}(M)$. Therefore σ is injective. Now let Ψ be an arbitrary element of $\mathcal{S}(M)$. By Theorem 8.5.8, we have that Ψ contains a Euclidean structure Φ for M. As $\hat{\Phi} = \Psi$, we have that σ is surjective. Thus σ is a bijection. □

Moduli Space

Two $(X, \mathrm{S}(X))$-structures Ψ and Ψ' for M are said to be *similar* if and only if (M, Ψ) and (M, Ψ') are $(X, \mathrm{S}(X))$-equivalent. Let $\mathcal{M}(M)$ be the set of similarity equivalence classes of complete $(X, \mathrm{S}(X))$-structures for M.

(1) If $\chi(M) = 0$, then $\mathcal{M}(M)$ is in one-to-one correspondence with the set of similarity classes of Euclidean structures for M by Theorem 9.4.1.

(2) If $\chi(M) < 0$, then $\mathcal{M}(M)$ is the set of isometry classes of hyperbolic structures for M.

The set $\mathcal{M}(M)$ is called the *moduli space* of Euclidean or hyperbolic structures for M.

We next study the relationship between $\mathcal{S}(M)$ and $\mathcal{M}(M)$. Let $\mathrm{Hom}(M)$ be the group of homeomorphisms of M. Define a right action of $\mathrm{Hom}(M)$ on $\mathcal{S}(M)$ as follows: If $h : M \to M$ is a homeomorphism and

$$\Psi = \{\psi_i : V_i \to X\}$$

is an element of $\mathcal{S}(M)$, define Ψh to be the element of $\mathcal{S}(M)$ given by

$$\Psi h = \{\psi_i h : h^{-1}(V_i) \to X\}.$$

Theorem 9.4.2. *If M is a closed surface such that $\chi(M) \leq 0$, then the natural projection from $\mathcal{S}(M)$ to $\mathcal{M}(M)$ induces a bijection from the set $\mathcal{S}(M)/\mathrm{Hom}(M)$ onto $\mathcal{M}(M)$.*

Proof: Let $h : M \to M$ be a homeomorphism and let

$$\Psi = \{\psi_i : V_i \to X\}$$

be an element of $\mathcal{S}(M)$. Then for each i and j, we have

$$(\psi_i h)(\psi_j h)^{-1} = \psi_i \psi_j^{-1}.$$

Hence h is an $(X, \mathrm{S}(X))$-map from $(M, \Psi h)$ to (M, Ψ). As h is a bijection, $(M, \Psi h)$ and (M, Ψ) are $(X, \mathrm{S}(X))$-equivalent. Hence, the natural projection from $\mathcal{S}(M)$ to $\mathcal{M}(M)$ induces a surjection

$$\mu : \mathcal{S}(M)/\mathrm{Hom}(M) \to \mathcal{M}(M).$$

Let Ψ and Ψ' be similar elements of $\mathcal{S}(M)$. Then there is an $(X, \mathrm{S}(X))$-equivalence $h : (M, \Psi') \to (M, \Psi)$. As h is a local homeomorphism and a bijection, h is a homeomorphism. If $\psi_i : V_i \to X$ and $\psi_j : V_j \to X$ are charts in Ψ and Ψ', respectively, then $\psi_i h \psi_j^{-1}$ agrees in a neighborhood of each point of its domain with an element of $\mathrm{S}(X)$. Therefore $\psi_i h$ is in Ψ'. Hence $\Psi h = \Psi'$. Thus Ψ and Ψ' are in the same $\mathrm{Hom}(M)$-orbit in $\mathcal{S}(M)$. Hence μ is injective. Thus μ is a bijection. □

Teichmüller Space

Let $\mathrm{Hom}_1(M)$ be the group of all homeomorphisms of M homotopic to the identity map of M. The *Teichmüller space* of Euclidean or hyperbolic structures for M is defined to be the set

$$\mathcal{T}(M) = \mathcal{S}(M)/\mathrm{Hom}_1(M).$$

The group $\mathrm{Hom}_1(M)$ is a normal subgroup of $\mathrm{Hom}(M)$. The quotient

$$\mathrm{Map}(M) = \mathrm{Hom}(M)/\mathrm{Hom}_1(M)$$

is called the full *mapping class group* of M. The action of $\mathrm{Hom}(M)$ on $\mathcal{S}(M)$ induces an action of $\mathrm{Map}(M)$ on $\mathcal{T}(M)$; moreover, the quotient map from $\mathcal{T}(M)$ to $\mathcal{M}(M)$ induces a bijection from $\mathcal{T}(M)/\mathrm{Map}(M)$ onto $\mathcal{M}(M)$.

The Dehn-Nielsen Theorem

Choose a base point u of M and let $h : M \to M$ be a homeomorphism. Then h induces an isomorphism

$$h_* : \pi_1(M, u) \to \pi_1(M, h(u)).$$

Let $\alpha : [0,1] \to M$ be a curve from u to $h(u)$. Then α determines a change of base point isomorphism

$$\alpha_* : \pi_1(M, h(u)) \to \pi_1(M, u)$$

defined by

$$\alpha_*([\gamma]) = [\alpha\gamma\alpha^{-1}].$$

The composite $\alpha_* h_*$ is an automorphism of $\pi_1(M) = \pi_1(M, u)$. Let $\beta : [0,1] \to M$ be another curve from u to $h(u)$. Then $\beta_* h_*$ is also an automorphism of $\pi_1(M)$. Moreover

$$\beta_* h_* = \beta_* \alpha_*^{-1} \alpha_* h_* = (\beta\alpha^{-1})_* \alpha_* h_*.$$

The automorphism $(\beta\alpha^{-1})_*$ of $\pi_1(M)$ is just conjugation by $[\beta\alpha^{-1}]$.

Let $\mathrm{Inn}(\pi_1(M))$ be the group of inner automorphisms of $\pi_1(M)$. Then the quotient group

$$\mathrm{Out}(\pi_1(M)) = \mathrm{Aut}(\pi_1(M))/\mathrm{Inn}(\pi_1(M))$$

is called the *outer automorphism group* of $\pi_1(M)$. Let $[h_*]$ be the coset $\alpha_* h_* \mathrm{Inn}(\pi_1(M))$ in $\mathrm{Out}(\pi_1(M))$. Then $[h_*]$ does not depend on the choice of the curve α. If h is homotopic to the identity map of M, then $\alpha_* h_*$ is an inner automorphism of $\pi_1(M)$, and so $[h_*] = 1$. Thus, the mapping $h \mapsto [h_*]$ induces a function

$$\nu : \mathrm{Map}(M) \to \mathrm{Out}(\pi_1(M)).$$

The next theorem is a basic theorem of surface theory.

Theorem 9.4.3. (The Dehn-Nielsen Theorem) *If M is a closed surface with $\chi(M) \leq 0$, then $\nu : \mathrm{Map}(M) \to \mathrm{Out}(\pi_1(M))$ is an isomorphism.*

Proof: We shall only prove that ν is a monomorphism. We begin by showing that ν is a homomorphism. Let $g, h : M \to M$ be homeomorphisms, let $\alpha : [0,1] \to M$ be a curve from the base point u to $h(u)$, and let $\beta : [0,1] \to M$ be a curve from u to $g(u)$. Then $\beta g\alpha : [0,1] \to M$ is a curve from u to $gh(u)$. Hence

$$\begin{aligned} \nu[gh] &= (\beta g\alpha)_*(gh)_*\mathrm{Inn}(\pi_1(M)) \\ &= (\beta_* g_* \alpha_* h_* \mathrm{Inn}(\pi_1(M)) \\ &= (\beta_* g_*)(\alpha_* h_*)\mathrm{Inn}(\pi_1(M)) \quad = \quad \nu[g]\nu[h]. \end{aligned}$$

Thus ν is a homomorphism.

Let $h : M \to M$ be a homeomorphism such that $\nu[h] = 1$ in $\mathrm{Out}(\pi_1(M))$ and let $\alpha : [0,1] \to M$ be a curve from u to $h(u)$. Then there is a loop $\gamma : [0,1] \to M$ based at u such that $\alpha_* h_* = \gamma_*$. Hence $h_* = (\alpha^{-1}\gamma)_*$. By replacing α by $\gamma^{-1}\alpha$, we may assume that $h_* = \alpha_*^{-1}$.

Now M has a cell structure with one 0-cell u, k 1-cells, and one 2-cell. Let $\gamma_i : [0,1] \to M$, for $i = 1, \ldots, k$, be characteristic maps for the 1-cells of M. Then

$$h\gamma_i \simeq \alpha^{-1}\gamma_i\alpha \simeq \gamma_i \quad \text{for each } i.$$

Hence, there are homotopies $H_i : [0,1]^2 \to M$ from $h\gamma_i$ to γ_i such that $H_i(0,t) = H_i(1,t)$ for all t and $H_i(0,t) = H_j(0,t)$ for all t and all i, j.

Let h_1 be the restriction of h to the 1-skeleton M^1 of M. Define a homotopy

$$H : M^1 \times [0,1] \to M$$

by $H(\gamma_i(s), t) = H_i(s,t)$. Then H is well defined and a homotopy of h_1 to the inclusion map of M^1 into M. As $\chi(M) \leq 0$, we have that $\pi_2(M) = 0$. Hence, we can extend H to a homotopy of h to the identity map of M. Therefore $[h] = 1$ in $\mathrm{Map}(M)$. Thus ν is a monomorphism. $\square$

Deformation Space

Let $\eta : \pi_1(M) \to \mathrm{I}(X)$ be a holonomy for M with respect to a complete $(X, \mathrm{S}(X))$-structure Ψ for M. The holonomy η depends on the choice of a developing map for M. If η' is another holonomy for M with respect to Ψ, then there is a similarity ξ of X such that

$$\eta'(c) = \xi\eta(c)\xi^{-1}$$

for each c in $\pi_1(M)$.

Let $[\eta]$ denote the orbit $\mathrm{S}(X)\eta$ under the left action of $\mathrm{S}(X)$ on the set of homomorphisms $\mathrm{Hom}(\pi_1(M), \mathrm{I}(X))$ by conjugation. Then $[\eta]$ does not depend on the choice of the developing map for M. Thus, the mapping $\Psi \mapsto [\eta]$ defines a function from $\mathcal{S}(M)$ into

$$\mathrm{S}(X)\backslash\mathrm{Hom}(\pi_1(M), \mathrm{I}(X)).$$

Now by Theorem 8.5.9, the holonomy η maps $\pi_1(M)$ isomorphically onto a discrete subgroup of $\mathrm{I}(X)$. A homomorphism in $\mathrm{Hom}(\pi_1(M), \mathrm{I}(X))$ mapping $\pi_1(M)$ isomorphically onto a discrete subgroup of $\mathrm{I}(X)$ is called a *discrete faithful representation* of $\pi_1(M)$ in $\mathrm{I}(X)$. Let $\mathrm{D}(\pi_1(M), \mathrm{I}(X))$ be the *set of discrete faithful representations* of $\pi_1(M)$ in $\mathrm{I}(X)$. Then $\mathrm{D}(\pi_1(M), \mathrm{I}(X))$ is invariant under the action of $\mathrm{S}(X)$.

The *deformation space* of M is defined to be the set

$$\mathcal{D}(M) = \mathrm{S}(X)\backslash\mathrm{D}(\pi_1(M), \mathrm{I}(X)).$$

Note that the mapping $\Psi \mapsto [\eta]$ defines a function from $\mathcal{S}(M)$ to $\mathcal{D}(M)$.

Let $h : M \to M$ be a homeomorphism and let $\delta : \tilde{M} \to X$ be the developing map for M that determines the holonomy η. Let $\kappa : \tilde{M} \to M$ be the universal covering projection and let $\tilde{h} : \tilde{M} \to \tilde{M}$ be a lift of h with respect to κ. Then $\delta\tilde{h} : \tilde{M} \to X$ is a developing map for the $(X, \mathrm{S}(X))$-structure Ψh for M. We now compute the holonomy for M determined by $\delta\tilde{h}$ in terms of η and h.

Choose a base point $\tilde{u}$ of $\tilde{M}$ such that $\kappa(\tilde{u}) = u$. Let $\alpha : [0, 1] \to M$ be a loop based at u. Then α lifts to a unique curve $\tilde{\alpha}$ in $\tilde{M}$ starting at $\tilde{u}$. Let $\tilde{v}$ be the endpoint of $\tilde{\alpha}$ and let τ_α be the unique covering transformation of κ such that $\tau_\alpha(\tilde{u}) = \tilde{v}$. Then there is a unique element g_α of $\mathrm{I}(X)$ such that

$$\delta\tau_\alpha = g_\alpha\delta.$$

The holonomy $\eta : \pi_1(M, u) \to \mathrm{I}(X)$ is defined by $\eta([\alpha]) = g_\alpha$.

Let $u' = h(u)$, $\tilde{u}' = \tilde{h}(\tilde{u})$, and $\eta' : \pi_1(M, u') \to \mathrm{I}(X)$ be the holonomy for M determined by δ. Then

$$\kappa\tilde{h}\tau_\alpha = h\kappa\tau_\alpha = h\kappa$$

and

$$\kappa\tau_{h\alpha}\tilde{h} = \kappa\tilde{h} = h\kappa.$$

Now as
$$\tilde{h}\tau_\alpha(\tilde{u}) = \tilde{h}(\tilde{v}) = \tau_{h\alpha}\tilde{h}(\tilde{u}),$$
we have that
$$\tilde{h}\tau_\alpha = \tau_{h\alpha}\tilde{h}.$$
Hence, we have
$$\delta\tilde{h}\tau_\alpha = \delta\tau_{h\alpha}\tilde{h} = g_{h\alpha}\delta\tilde{h}.$$
Thus, the holonomy for M determined by $\delta\tilde{h}$ is
$$\eta' h_* : \pi_1(M, u) \to \mathrm{I}(X).$$

Note that η' is defined relative to the base point $u' = h(u)$. We now switch the base point back to u. Let $\tilde{\gamma} : [0,1] \to \tilde{M}$ be a curve from $\tilde{u}$ to $\tilde{u}'$ and set $\gamma = \kappa\tilde{\gamma}$. Then $\gamma : [0,1] \to M$ is a curve from u to u'. Let $\beta : [0,1] \to M$ be a loop based at u' and let $\tilde{\beta} : [0,1] \to \tilde{M}$ the lift of β starting at $\tilde{u}'$. Then $\gamma\beta\gamma^{-1} : [0,1] \to M$ is a loop based at u and
$$\tilde{\gamma}\tilde{\beta}(\tau_\beta\tilde{\gamma}^{-1}) : [0,1] \to \tilde{M}$$
is the lift of $\gamma\beta\gamma^{-1}$ starting at $\tilde{u}$. Observe that
$$\tilde{\gamma}\tilde{\beta}(\tau_\beta\tilde{\gamma}^{-1})(1) = \tau_\beta(\tilde{u}).$$
Hence $\tau_{\gamma\beta\gamma^{-1}} = \tau_\beta$. Thus $\eta' = \eta\gamma_*$ where
$$\gamma_* : \pi_1(M, u') \to \pi_1(M, u)$$
is the change of base point isomorphism. Therefore, the holonomy for M determined by $\delta\tilde{h}$ is
$$\eta\gamma_* h_* : \pi_1(M, u) \to \mathrm{I}(X).$$

Now suppose that $h : M \to M$ is homotopic to the identity map of M. Then the automorphism
$$\gamma_* h_* : \pi_1(M) \to \pi_1(M)$$
is conjugation by an element b of $\pi_1(M)$. If c is in $\pi_1(M)$, then
$$\eta\gamma_* h_*(c) = \eta(bcb^{-1}) = \eta(b)\eta(c)\eta(b)^{-1}.$$
Therefore, we have that
$$\eta\gamma_* h_* = \eta(b) \cdot \eta.$$
Hence Ψ and Ψh determine the same element $[\eta]$ of $\mathcal{D}(M)$. Thus, the mapping $\Psi \mapsto [\eta]$ induces a function $\rho : \mathcal{T}(M) \to \mathcal{D}(M)$ defined by
$$\rho([\Psi]) = [\eta],$$
where $[\Psi] = \Psi\mathrm{Hom}_1(M)$.

Theorem 9.4.4. *If M is a closed surface such that $\chi(M) \leq 0$, then the function $\rho : \mathcal{T}(M) \to \mathcal{D}(M)$, defined by $\rho([\Psi]) = [\eta]$, where η is a holonomy for (M, Ψ), is a bijection.*

Proof: We first show that ρ is injective. Let Ψ_1 and Ψ_2 be complete $(X, \mathrm{S}(X))$-structures for M such that $\rho([\psi_1]) = \rho([\Psi_2])$. Let $\delta_i : \tilde{M} \to X$ be a developing map for (M, Ψ_i) and let

$$\eta_i : \pi_1(M, u) \to \mathrm{I}(X)$$

be the holonomy for M determined by δ_i for $i = 1, 2$. Then $\rho([\Psi_i]) = [\eta_i]$ for $i = 1, 2$. Therefore $[\eta_1] = [\eta_2]$. Hence, there is a similarity ξ of X such that $\eta_2 = \xi \cdot \eta_1$. Now $\xi\delta_1$ is also a developing map for (M, Ψ_1); moreover, $\xi\delta_1$ determines the holonomy $\xi \cdot \eta_1$. Hence, by replacing δ_1 with $\xi\delta_1$, we may assume that $\eta_1 = \eta_2$.

Let $\Gamma = \mathrm{Im}(\eta_i)$ for $i = 1, 2$. Then Γ acts freely and discontinuously on X by Theorem 8.5.9. Let $\overline{\delta}_i : M \to X/\Gamma$ be the map induced by δ_i for $i = 1, 2$. Then $\overline{\delta}_i$ is an $(X, \mathrm{S}(X))$-equivalence from (M, Ψ_i) to X/Γ for $i = 1, 2$. Let $h = \overline{\delta}_2^{-1}\overline{\delta}_1$. Then h is an $(X, \mathrm{S}(X))$-equivalence from (M, Ψ_1) to (M, Ψ_2). Therefore $\Psi_2 h = \Psi$ by Theorem 9.4.2.

Let $\Gamma x_i = \overline{\delta}_i(u)$ and let

$$\vartheta_i : \pi_1(X/\Gamma, \Gamma x_i) \to \Gamma$$

be the holonomy for X/Γ for $i = 1, 2$. Then η_i is the composite

$$\pi_1(M) \xrightarrow{(\overline{\delta}_i)_*} \pi_1(X/\Gamma) \xrightarrow{\vartheta_i} \Gamma.$$

Let $\tilde{\gamma} : [0, 1] \to X$ be a curve from x_1 to x_2 and set $\gamma = \pi\tilde{\gamma}$. Then $\gamma : [0, 1] \to X/\Gamma$ is a curve from Γx_1 to Γx_2 and $\vartheta_2 = \vartheta_1\gamma_*$. Hence

$$\begin{aligned}
(\overline{\delta}_2^{-1}\gamma^{-1})_* h_* &= (\overline{\delta}_2^{-1}\gamma^{-1})_*(\overline{\delta}_2^{-1})_*(\overline{\delta}_1)_* \\
&= (\overline{\delta}_2^{-1})_*\gamma_*^{-1}(\overline{\delta}_1)_* \\
&= \eta_2^{-1}\vartheta_2\gamma_*^{-1}\vartheta_1^{-1}\eta_1 \\
&= \eta_2^{-1}\eta_1 \\
&= 1.
\end{aligned}$$

Therefore h is homotopic to the identity map of M by Theorem 9.4.3. Hence $[\Psi_1] = [\Psi_2]$. Thus ρ is injective.

We now show that ρ is surjective. Let $\eta : \pi_1(M) \to \mathrm{I}(X)$ be a discrete faithful representation of $\pi_1(M)$ in $\mathrm{I}(X)$ and set $\Gamma = \mathrm{Im}(\eta)$. Since M has either a Euclidean or hyperbolic structure, $\pi_1(M)$ is torsion-free. Therefore Γ is a torsion-free discrete subgroup of $\mathrm{I}(X)$. Hence Γ acts freely and discontinuously on X, and so X/Γ is either a Euclidean or hyperbolic surface.

Let $\vartheta : \pi_1(X/\Gamma) \to \Gamma$ be the holonomy for X/Γ. Then $\vartheta^{-1}\eta : \pi(M) \to \pi_1(X/\Gamma)$ is an isomorphism. Consequently M and X/Γ are homeomorphic. By Theorem 9.4.3, there is a homeomorphism $h : M \to X/\Gamma$ such that

$$\alpha_* h_* = \vartheta^{-1}\eta\iota,$$

where α_* is a change of base point isomorphism and ι is an inner automorphism of $\pi_1(M)$.

Let $\Psi = \{\psi_i : V_i \to X\}$ be the $(X, \mathrm{S}(X))$-structure for X/Γ. Then

$$\Psi h = \{\psi_i h : h^{-1}(V_i) \to X\}$$

is a complete $(X, \mathrm{S}(X))$-structure for M. Lift h to a homeomorphism $\tilde{h} : \tilde{M} \to X$. Then $\tilde{h}$ is a developing map for $(M, \Psi h)$. The holonomy for M determined by $\tilde{h}$ is $\vartheta\beta_* h_*$ where β_* is a change of base point isomorphism. Therefore, we have

$$\begin{aligned} \rho[\Psi h]) &= [\vartheta\beta_* h_*] \\ &= [\vartheta\beta_*\alpha_*^{-1}\alpha_* h_*] \\ &= [\vartheta(\beta\alpha)_*^{-1}\alpha_* h_*] \\ &= [\vartheta\alpha_* h_*] \;=\; [\eta]. \end{aligned}$$

Hence ρ is surjective. Thus ρ is a bijection. □

The group $\mathrm{Aut}(\pi_1(M))$ acts on $\mathrm{D}(\pi_1(M), \mathrm{I}(X))$ on the right. Moreover, if ζ is an automorphism of $\pi_1(M)$ and η is in $\mathrm{D}(\pi_1(M), \mathrm{I}(X))$ and ξ is a similarity of X, then

$$(\xi \cdot \eta)\zeta = \xi \cdot (\eta\zeta).$$

Hence, the action of $\mathrm{Aut}(\pi_1(M))$ on $\mathrm{D}(\pi_1(M), \mathrm{I}(X))$ induces an action of $\mathrm{Aut}(\pi_1(M))$ on $\mathcal{D}(M)$. Let ι be an inner automorphism of $\pi_1(M)$. Then there is a b in $\pi_1(M)$ such that $\iota(c) = bcb^{-1}$ for all c in $\pi_1(M)$. If η is in $\mathrm{D}(\pi_1(M), \mathrm{I}(X))$, then

$$\eta\iota(c) = \eta(bcb^{-1}) = \eta(b)\eta(c)\eta(b)^{-1}.$$

Hence $\eta\iota = \eta(b) \cdot \eta$. Therefore $\mathrm{Inn}(\pi_1(M))$ acts trivially on $\mathcal{D}(M)$. Hence, the action of $\mathrm{Aut}(\pi_1(M))$ on $\mathcal{D}(M)$ induces an action of $\mathrm{Out}(\pi_1(M))$ on $\mathcal{D}(M)$. Let

$$\mathcal{O}(M) = \mathcal{D}(M)/\mathrm{Out}(\pi_1(M)).$$

Theorem 9.4.5. *If M is a closed surface such that $\chi(M) \leq 0$, then the function $\rho : \mathcal{T}(M) \to \mathcal{D}(M)$ induces a bijection $\overline{\rho} : \mathcal{M}(M) \to \mathcal{O}(M)$.*

Proof: Let Ψ be a complete $(X, \mathrm{S}(X))$-structure for M and let $h : M \to M$ be a homeomorphism. Let $\eta : \pi_1(M) \to \mathrm{I}(X)$ be a holonomy for (M, Ψ). Then there is a change of base point isomorphism γ_* such that $\eta\gamma_* h_* : \pi_1(M) \to \mathrm{I}(X)$ is the holonomy for Ψh. Hence

$$\begin{aligned} \rho([\Psi][h]) &= \rho([\Psi h]) \\ &= [\eta\gamma_* h_*] \\ &= [\eta][h_*] \\ &= \rho([\Psi])\nu([h]). \end{aligned}$$

By Theorems 9.4.3 and 9.4.4, we have that ρ induces a bijection from $\mathcal{T}(M)/\mathrm{Map}(M)$ onto $\mathcal{D}(M)/\mathrm{Out}(\pi_1(M))$. Thus ρ induces a bijection from $\mathcal{M}(M)$ onto $\mathcal{O}(M)$. □

We now define a topology for each of the sets $\mathcal{D}(M)$, $\mathcal{O}(M)$, $\mathcal{T}(M)$, and $\mathcal{M}(M)$. First, topologize $\pi_1(M)$ with the discrete topology and the set $\mathrm{C}(\pi_1(M), \mathrm{I}(X))$ of all functions from $\pi_1(M)$ to $\mathrm{I}(X)$ with the compact-open topology. Then $\mathrm{C}(\pi_1(M), \mathrm{I}(X))$ is the cartesian product $\mathrm{I}(X)^{\pi_1(M)}$ with the product topology.

Next, we topologize $\mathrm{D}(\pi_1(M), \mathrm{I}(X))$ with the subspace topology inherited from $\mathrm{C}(\pi_1(M), \mathrm{I}(X))$. Now we topologize $\mathcal{D}(M)$ and $\mathcal{O}(M)$ with the quotient topology inherited from $\mathrm{D}(\pi_1(M), \mathrm{I}(X))$ and $\mathcal{D}(M)$, respectively. Finally, we topologize $\mathcal{T}(M)$ and $\mathcal{M}(M)$ so that $\rho : \mathcal{T}(M) \to \mathcal{D}(M)$ and $\overline{\rho} : \mathcal{M}(M) \to \mathcal{O}(M)$ are homeomorphisms. Then $\mathcal{M}(M)$ has the quotient topology inherited from $\mathcal{T}(M)$.

Remark: It is a fundamental theorem of Teichmüller space theory that Teichmüller space $\mathcal{T}(M)$ is homeomorphic to a finite dimensional Euclidean space. Moreover $\mathcal{T}(M)$ has a finitely compact metric such that the mapping class group $\mathrm{Map}(M)$ acts discontinuously on $\mathcal{T}(M)$ by isometries. Therefore, the orbit space $\mathcal{T}(M)/\mathrm{Map}(M)$ has a complete metric. Now $\mathcal{T}(M)/\mathrm{Map}(M)$ is homeomorphic to $\mathcal{M}(M)$. Therefore, moduli space $\mathcal{M}(M)$ has a complete metric.

Exercise 9.4

1. Let Φ and Φ' be Euclidean structures for M. Prove that $\hat{\Phi}$ and $\hat{\Phi}'$ are similar if and only if (M, Φ) and (M, Φ') are similar metric spaces.
2. Let Φ and Φ' be hyperbolic structures for M. Prove that Φ and Φ' are similar if and only if (M, Φ) and (M, Φ') are isometric.
3. Let Φ and Φ' be hyperbolic structures for M. Prove that $[\Phi] = [\Phi']$ in $\mathcal{T}(M)$ if and only if there is an isometry from (M, Φ) to (M, Φ') that is homotopic to the identity map of M.
4. Let $h : M \to M$ be a homeomorphism of a surface M and let $\alpha : [0, 1] \to M$ be a curve from u to $h(u)$. Prove that if h is homotopic to the identity map of M, then $\alpha_* h_*$ is an inner automorphism of $\pi_1(M, u)$.
5. Let M be a closed surface. Prove that the natural action of $\mathrm{Hom}_1(M)$ on M is transitive.
6. Let u be a point of a surface M and let $h : M \to M$ be a homeomorphism. Prove that h is homotopic to a homeomorphism $g : M \to M$ such that $g(u) = u$.
7. Prove that Nielsen's homomorphism ν is surjective if M is a torus.
8. Prove that Nielsen's homomorphism ν is surjective if M is a Klein bottle.
9. Let M be a closed surface. Prove that $\mathrm{Aut}(\pi_1(M))$ is a countable group. Conclude that $\mathrm{Out}(\pi_1(M))$ is a countable group.
10. Prove that $\mathrm{C}(\pi_1(M), \mathrm{I}(X))$ is the cartesian product $\mathrm{I}(X)^{\pi_1(M)}$ with the product topology.

§9.5. Closed Euclidean Surfaces

In this section, we classify the Euclidean structures on the torus T^2. By definition, T^2 is the orbit space $E^2/\mathbb{Z}^2$. Therefore T^2 has a Euclidean structure as a Euclidean space-form. This Euclidean structure on T^2 is far from unique. We shall prove that T^2 has an uncountable number of nonsimilar Euclidean structures.

Theorem 9.5.1. *The deformation space $\mathcal{D}(T^2)$ is homeomorphic to the upper half-plane U^2; moreover, the right action of the group $\mathrm{Aut}(\pi_1(T^2))$ on $\mathcal{D}(T^2)$ corresponds to the right action of $\mathrm{GL}(2,\mathbb{Z})$ on U^2 given by*

$$z \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{cases} \frac{az+c}{bz+d} & \text{if } ad-bc=1, \\ \frac{a\overline{z}+c}{b\overline{z}+d} & \text{if } ad-bc=-1. \end{cases}$$

Proof: We shall identify $\pi_1(T^2)$ with $\mathbb{Z}^2$ and E^2 with $\mathbb{C}$. By Theorem 5.4.4, every homomorphism in $\mathrm{D}(\mathbb{Z}^2, \mathrm{I}(\mathbb{C}))$ maps $\mathbb{Z}^2$ into the subgroup $\mathrm{T}(\mathbb{C})$ of translations of $\mathbb{C}$. By Corollary 1 of Theorem 5.2.4, we may identify $\mathrm{T}(\mathbb{C})$ with $\mathbb{C}$.

We now show that $\mathrm{Hom}(\mathbb{Z}^2, \mathbb{C})$ is homeomorphic to $\mathbb{C}^2$. Define

$$h : \mathrm{Hom}(\mathbb{Z}^2, \mathbb{C}) \to \mathbb{C}^2$$

by the formula

$$h(\eta) = (\eta(1,0), \eta(0,1)).$$

As each component of h is an evaluation map, h is continuous. The map h is obviously an isomorphism of groups. To see that h^{-1} is continuous, we regard $\mathrm{Hom}(\mathbb{Z}^2, \mathbb{C})$ to be a subspace of the cartesian product $\mathbb{C}^{\mathbb{Z}^2}$. Now $h^{-1} : \mathbb{C}^2 \to \mathbb{C}^{\mathbb{Z}^2}$ is defined by

$$h^{-1}(z,w)(m,n) = mz + nw.$$

Hence, each component of h^{-1}, given by $(z,w) \mapsto mz + nw$, is continuous and so h^{-1} is continuous. Thus h is a homeomorphism.

Let ξ be a similarity of $\mathbb{C}$. Then there is a nonzero complex number u and a complex number v such that

$$\xi(z) = \begin{cases} uz + v & \text{if } \xi \text{ preserves orientation,} \\ u\overline{z} + v & \text{if } \xi \text{ reverses orientation.} \end{cases}$$

Let τ be the translation of $\mathbb{C}$ by w. If ξ preserves orientation, then

$$\begin{aligned} \xi\tau\xi^{-1}(z) &= \xi\tau(u^{-1}z - u^{-1}v) \\ &= \xi(u^{-1}z - u^{-1}v + w) \\ &= z + uw. \end{aligned}$$

If ξ reverses orientation, then

$$\begin{aligned} \xi\tau\xi^{-1}(z) &= \xi\tau(\overline{u}^{-1}\overline{z} - \overline{u}^{-1}\overline{v}) \\ &= \xi(\overline{u}^{-1}\overline{z} - \overline{u}^{-1}\overline{v} + w) \\ &= z + u\overline{w}. \end{aligned}$$

Hence, the action of $S(\mathbb{C})$ on $T(\mathbb{C})$ by conjugation corresponds under the identification of $T(\mathbb{C})$ with $\mathbb{C}$ to multiplication by nonzero complex numbers of $\mathbb{C}$ possibly followed by complex conjugation. Moreover, the left action of $S(\mathbb{C})$ on $\mathrm{Hom}(\mathbb{Z}^2, \mathbb{C})$ corresponds under h to multiplication by nonzero complex numbers on $\mathbb{C}^2$ possibly followed by complex conjugation on $\mathbb{C}^2$.

By Theorem 5.3.2, a homomorphism $\eta : \mathbb{Z}^2 \to \mathbb{C}$ maps $\mathbb{Z}^2$ isomorphically onto a discrete subgroup of $\mathbb{C}$ if and only if $\eta(1,0)$ and $\eta(0,1)$ are linearly independent over $\mathbb{R}$. Hence $D(\mathbb{Z}^2, \mathbb{C})$ corresponds under h to the subset D of $\mathbb{C}^2$ of all pairs (z,w) such that z, w are linearly independent over $\mathbb{R}$. Now define $f : D \to U^2$ by

$$f(z,w) = \begin{cases} z/w & \text{if } \mathrm{Im}(z/w) > 0, \\ \overline{z}/\overline{w} & \text{if } \mathrm{Im}(z/w) < 0. \end{cases}$$

Then f is continuous and induces a continuous bijection

$$g : S(\mathbb{C})\backslash D \to U^2.$$

As the mapping $z \mapsto (z,1)$ from U^2 to D is continuous, we see that g^{-1} is continuous. Therefore g is a homeomorphism. Thus $\mathcal{D}(T^2)$ is homeomorphic to U^2.

We identify $\mathrm{Aut}(\mathbb{Z}^2)$ with the group $GL(2,\mathbb{Z})$ so that a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in $GL(2,\mathbb{Z})$ represents the automorphism of $\mathbb{Z}^2$ that maps $(1,0)$ to (a,c) and $(0,1)$ to (b,d). Then the right action of $\mathrm{Aut}(\mathbb{Z}^2)$ on $\mathrm{Hom}(\mathbb{Z}^2,\mathbb{C})$ corresponds under the isomorphism

$$h : \mathrm{Hom}(\mathbb{Z}^2, \mathbb{C}) \to \mathbb{C}^2$$

to the right action of $GL(2,\mathbb{Z})$ on $\mathbb{C}^2$ given by

$$(z,w)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (az + cw, bz + dw).$$

Hence, the right action of $GL(2,\mathbb{Z})$ on $S(\mathbb{C})\backslash D$ corresponds under the homeomorphism

$$g : S(\mathbb{C})\backslash D \to U^2$$

to the right action of $GL(2,\mathbb{Z})$ on U^2 given by

$$z \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{cases} \frac{az+c}{bz+d} & \text{if } ad - bc = 1, \\ \frac{a\overline{z}+c}{b\overline{z}+d} & \text{if } ad - bc = -1. \end{cases}$$ □

Theorem 9.5.2. *The moduli space $\mathcal{M}(T^2)$ is homeomorphic to the hyperbolic triangle $\triangle(i, \frac{1}{2} + \frac{\sqrt{3}}{2}i, \infty)$ in U^2.*

Proof: If $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is in $GL(2,\mathbb{Z})$, then

$$z \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot z,$$

where $\mathrm{GL}(2,\mathbb{Z})$ acts on the left by hyperbolic isometries of U^2. Hence, the orbit space $U^2/\mathrm{GL}(2,\mathbb{Z})$ is the same as the orbit space $\mathrm{PGL}(2,\mathbb{Z})\backslash U^2$. Now the triangle $\triangle(i, \frac{1}{2}+\frac{\sqrt{3}}{2}i, \infty)$ is a fundamental polygon for $\mathrm{PGL}(2,\mathbb{Z})$; moreover, $\mathrm{PGL}(2,\mathbb{Z})$ is a triangle reflection group with respect to $\triangle$. Therefore $\mathrm{PGL}(2,\mathbb{Z})\backslash U^2$ is homeomorphic to $\triangle$ by Theorem 6.5.8. Now $\mathcal{O}(T^2)$ is homeomorphic to $U^2/\mathrm{GL}(2,\mathbb{Z})$ by Theorem 9.5.1. Hence $\mathcal{M}(T^2)$ is homeomorphic to the triangle $\triangle$. □

Let P be the unit square in $\mathbb{C}$ with vertices $0, 1, 1+i, i$. The *Klein bottle* K^2 is, by definition, the surface obtained by gluing the opposite sides of P by the translation τ_1, defined by $\tau_1(z) = z + 1$, and the glide-reflection ρ_1, defined by $\rho_1(z) = -\overline{z} + 1 + i$. This side-pairing of P is proper, and so K^2 has a Euclidean structure by Theorem 9.2.2.

We leave it as an exercise to show that τ_1 and ρ_1 generate a discrete subgroup Γ_1 of $\mathrm{I}(\mathbb{C})$ and P is a fundamental polygon for Γ_1. The group Γ_1 is called the *Klein bottle group.* The group Γ_1 is isomorphic to $\pi_1(K^2)$ by Theorems 6.5.8, 6.5.10, and 8.1.4. Like the torus T^2, the Klein bottle K^2 has an uncountable number of nonsimilar Euclidean structures. The proof of the next theorem is left as an exercise for the reader.

Theorem 9.5.3. *The deformation space $\mathcal{D}(K^2)$ is homeomorphic to U^1; moreover, $\mathrm{Out}(\pi_1(K^2))$ acts trivially on $\mathcal{D}(K^2)$ and therefore the moduli space $\mathcal{M}(K^2)$ is also homeomorphic to U^1.*

Exercise 9.5

1. Let P be the parallelogram in $\mathbb{C}$, with vertices $0, 1, z, w$ in positive order around P, and let M be the torus obtained from P by gluing the opposite sides of P by translations. Prove that the class of M in $\mathcal{T}(T^2)$ corresponds to the point w of U^2 under the composite of the bijections of Theorems 9.4.4 and 9.5.1.

2. Show that τ_1 and ρ_1^2 generate a discrete subgroup of $\mathrm{T}(\mathbb{C})$ of index two in the Klein bottle group Γ_1. Conclude that Γ_1 is a discrete subgroup of $\mathrm{I}(\mathbb{C})$.

3. Prove that the square P in $\mathbb{C}$, with vertices $0, 1, 1+i, i$, is a fundamental polygon for the Klein bottle group Γ_1.

4. Prove that a discrete subgroup Γ of $\mathrm{I}(\mathbb{C})$ is isomorphic to Γ_1 if and only if there are v, w in $\mathbb{C}$ such that v, w are linearly independent over $\mathbb{R}$ and Γ is generated by τ and ρ defined by $\tau(z) = z + w$ and $\rho(z) = -(w/\overline{w})\overline{z} + v$.

5. Prove that $\mathcal{D}(K^2)$ is homeomorphic to U^1.

6. Let P be the parallelogram in $\mathbb{C}$, with vertices $0, 1, z, w$ in positive order around P, and let M be the Klein bottle obtained from P by gluing the opposite sides $[0, w]$ and $[1, z]$ by a translation and $[0, 1]$ and $[w, z]$ by a glide-reflection. Prove that the class of M in $\mathcal{T}(K^2)$ corresponds to the

point $\mathrm{Im}(z)$ of U^1 under the composite of the bijections of Theorems 9.4.4 and 9.5.3.

7. Prove that τ_1 generates a characteristic subgroup of Γ_1 and that $\Gamma_1/\langle\tau_1\rangle$ is an infinite cyclic group generated by $\langle\tau_1\rangle\rho_1$.
8. Prove that $\mathrm{Out}(\Gamma_1)$ is a Klein four-group generated by the cosets $\mathrm{Inn}(\Gamma_1)\alpha$ and $\mathrm{Inn}(\Gamma_1)\beta$, where $\alpha(\tau_1) = \tau_1$ and $\alpha(\rho_1) = \tau_1\rho_1$, and $\beta(\tau_1) = \tau_1$ and $\beta(\rho_1) = \rho_1^{-1}$.
9. Prove that $\mathrm{Out}(\pi_1(K^2))$ acts trivially on $\mathcal{D}(K^2)$.
10. Let $\kappa : \mathcal{M}(K^2) \to \mathcal{M}(T^2)$ be the function defined by mapping the class of a Klein bottle to the class of its orientable double cover. Prove that κ is well defined and that κ is neither surjective nor injective.

§9.6. Closed Geodesics

In this section, we study the geometry of closed geodesics of hyperbolic surfaces.

Definition: A *period* of a geodesic line $\lambda : \mathbb{R} \to X$ is a positive real number p such that $\lambda(t+p) = \lambda(t)$ for all t in $\mathbb{R}$. A geodesic line λ is *periodic* if it has a period.

Theorem 9.6.1. *A periodic geodesic line $\lambda : \mathbb{R} \to X$ has a smallest period p_1 and every period of λ is a multiple of p_1.*

Proof: Let P be the set of all real numbers p such that $\lambda(t+p) = \lambda(t)$ for all t. Then P consists of all the periods of λ, their negatives, and zero. The set P is clearly a subgroup of $\mathbb{R}$. Now since λ is a geodesic line, there is an $s > 0$ such that λ restricted to the closed interval $[-s, s]$ is a geodesic arc. Therefore λ is injective on $[-s, s]$. If p is a nonzero element of P, then $\lambda(p) = \lambda(0)$, and so p cannot lie in the open interval $(-s, s)$. Therefore 0 is open in P, and so P is a discrete subgroup of $\mathbb{R}$. By Theorem 5.3.2, the group P is infinite cyclic. Let p_1 be the positive generator of P. Then p_1 is the smallest period of λ, and every period of λ is a multiple of p_1. □

Definition: A *closed geodesic* in a metric space X is the image of a periodic geodesic line $\lambda : \mathbb{R} \to X$.

Example: Let $M = H^n/\Gamma$ be a space-form and let $\pi : H^n \to H^n/\Gamma$ be the quotient map. Let h be a hyperbolic element of Γ with axis L in H^n, and let $\tilde{\lambda} : \mathbb{R} \to H^n$ be a geodesic line whose image is L. Then h acts on L as a translation by a distance $p = d(\tilde{\lambda}(0), h\tilde{\lambda}(0))$. Therefore $\lambda = \pi\tilde{\lambda} : \mathbb{R} \to M$ is a periodic geodesic line with period p. Hence, the set $C = \lambda(\mathbb{R})$ is a closed geodesic of M. Observe that

$$C = \lambda(\mathbb{R}) = \pi\tilde{\lambda}(\mathbb{R}) = \pi(L).$$

Therefore, the axis L of h projects onto the closed geodesic C of M.

Definition: An element h of a group Γ is *primitive* in Γ if and only if h has no roots in Γ, that is, if $h = g^m$, with g in Γ, then $m = \pm 1$.

Theorem 9.6.2. *Let C be a closed geodesic of a space-form $M = H^n/\Gamma$. Then there is a primitive hyperbolic element h of Γ whose axis projects onto C. Moreover, the axis of a hyperbolic element f of Γ projects onto C if and only if there is an element g of Γ and a nonzero integer k such that $f = gh^kg^{-1}$.*

Proof: Since C is a closed geodesic, there is a periodic geodesic line $\lambda : \mathbb{R} \to M$ whose image is C. Let $\tilde{\lambda} : \mathbb{R} \to H^n$ be a lift of λ with respect to $\pi : H^n \to H^n/\Gamma$. Then $\tilde{\lambda}$ maps $\mathbb{R}$ isometrically onto a hyperbolic line L of H^n. Let p be the smallest period of λ. Then $\pi\tilde{\lambda}(p) = \pi\tilde{\lambda}(0)$. Hence, there is a nonidentity element h of Γ such that $\tilde{\lambda}(p) = h\tilde{\lambda}(0)$. Now $h\tilde{\lambda} : \mathbb{R} \to H^n$ also lifts λ and agrees with $\hat{\lambda} : \mathbb{R} \to H^n$, defined by

$$\hat{\lambda}(t) = \tilde{\lambda}(t + p),$$

at $t = 0$. As $\hat{\lambda}$ also lifts λ, we have that $h\tilde{\lambda} = \hat{\lambda}$ by the unique lifting property of the covering projection $\pi : H^n \to H^n/\Gamma$. Therefore h leaves L invariant. Hence h is hyperbolic with axis L. Moreover h is primitive in Γ, since

$$p = d(\tilde{\lambda}(0), h\tilde{\lambda}(0))$$

is the smallest period of λ. Thus h is a primitive hyperbolic element of Γ whose axis projects onto C.

Let f be a hyperbolic element of Γ and suppose that g is an element of Γ and k is a nonzero integer such that $f = gh^kg^{-1}$. Then the axis of f is gL. Therefore, the axis of f projects onto C.

Conversely, suppose that the axis K of f projects onto C. Then there exists an element g of Γ such that $K = gL$. Now $g^{-1}fg$ is a hyperbolic element of Γ with axis L. Hence $g^{-1}fg$ acts as a translation on L by a signed distance, say q. Now $\pm q$ is a period of λ, and so there is a nonzero integer k such that $q = kp$ by Theorem 9.6.1. Hence $g^{-1}fgh^{-k}$ fixes each point of L. As Γ acts freely on H^n, we have that $g^{-1}fgh^{-k} = 1$. Therefore $g^{-1}fg = h^k$ and so $f = gh^kg^{-1}$. □

Theorem 9.6.3. *Let $M = H^n/\Gamma$ be a compact space-form. Then every nonidentity element of Γ is hyperbolic.*

Proof: Since Γ is discrete and M is compact, every element of Γ is either elliptic or hyperbolic by Theorem 6.5.7. Moreover, since Γ acts freely on H^n, an elliptic element of Γ must be the identity. □

Closed Curves

Let $M = H^n/\Gamma$ be a space-form. A closed curve $\gamma : [0,1] \to M$ is said to be *elliptic*, *parabolic*, or *hyperbolic* if and only if for a lift $\tilde{\gamma} : [0,1] \to H^n$, the element g of Γ such that $\tilde{\gamma}(1) = g\tilde{\gamma}(0)$ is elliptic, parabolic, or hyperbolic, respectively. This does not depend on the choice of the lift $\tilde{\gamma}$, since if $\hat{\gamma} : [0,1] \to H^n$ is another lift of γ, then $\hat{\gamma} = f\tilde{\gamma}$ for some f in Γ and so

$$\begin{aligned} fgf^{-1}\hat{\gamma}(0) &= fgf^{-1}f\tilde{\gamma}(0) \\ &= fg\tilde{\gamma}(0) \\ &= f\tilde{\gamma}(1) \quad = \quad \hat{\gamma}(1). \end{aligned}$$

Note that a closed curve $\gamma : [0,1] \to M$ is elliptic if and only if γ is null homotopic (nonessential). Hence, an essential closed curve $\gamma : [0,1] \to M$ is either parabolic or hyperbolic. If M is compact, then every essential closed curve $\gamma : [0,1] \to M$ is hyperbolic by Theorem 9.6.3.

Definition: Two closed curves $\alpha, \beta : [0,1] \to X$ are *freely homotopic* if and only if there is a homotopy $H : [0,1]^2 \to X$ from α to β such that $H(0,t) = H(1,t)$ for all t.

Theorem 9.6.4. *Let $\gamma : [0,1] \to M$ be a hyperbolic closed curve in a complete hyperbolic n-manifold M. Then there is a periodic geodesic line $\lambda : \mathbb{R} \to M$ that is unique up to composition with a translation in $\mathbb{R}$, and there is a unique period p of λ such that γ is freely homotopic to the closed curve $\lambda_p : [0,1] \to M$ defined by $\lambda_p(t) = \lambda(pt)$.*

Proof: Since any closed curve freely homotopic to γ is in the same connected component of M as γ, we may assume that M is connected. As M is complete, we may assume that M is a space-form H^n/Γ by Theorem 8.5.9. Let $\tilde{\gamma} : [0,1] \to H^n$ be a lift of γ with respect to the quotient map $\pi : H^n \to H^n/\Gamma$. As γ is hyperbolic, the element h of Γ such that $h\tilde{\gamma}(0) = \tilde{\gamma}(1)$ is hyperbolic.

Let L be the axis of h in H^n and let $\tilde{\lambda} : \mathbb{R} \to H^n$ be a geodesic line parameterizing L in the same direction that h translates L. Then $\lambda = \pi\tilde{\lambda}$ is a geodesic line in M. Let $p > 0$ be such that

$$h\tilde{\lambda}(t) = \tilde{\lambda}(t+p).$$

Applying π, we find that

$$\lambda(t) = \lambda(t+p).$$

Thus p is a period for λ.

Define a homotopy $\tilde{H} : [0,1]^2 \to H^n$ from $\tilde{\gamma}$ to $\tilde{\lambda}_p$ by the formula

$$\tilde{H}(s,t) = \frac{(1-t)\tilde{\gamma}(s) + t\tilde{\lambda}_p(s)}{|||(1-t)\tilde{\gamma}(s) + t\tilde{\lambda}_p(s)|||}.$$

Observe that

$$
\begin{aligned}
h\tilde{H}(0,t) &= \frac{h((1-t)\tilde{\gamma}(0)+t\tilde{\lambda}(0))}{|||(1-t)\tilde{\gamma}(0)+t\tilde{\lambda}(0)|||} \\
&= \frac{(1-t)h\tilde{\gamma}(0)+th\tilde{\lambda}(0)}{|||h((1-t)\tilde{\gamma}(0)+t\tilde{\lambda}(0))|||} \\
&= \frac{(1-t)\tilde{\gamma}(1)+t\tilde{\lambda}(p)}{|||(1-t)\tilde{\gamma}(1)+t\tilde{\lambda}(p)|||} \quad = \quad \tilde{H}(1,t).
\end{aligned}
$$

Let $H = \pi\tilde{H}$. Then $H(0,t) = H(1,t)$ for all t. Hence γ is freely homotopic to λ_p via H.

We now prove uniqueness. Let $\mu : \mathbb{R} \to M$ be a periodic geodesic line and let q be a period of μ such that γ is freely homotopic to μ_q. Let $G : [0,1]^2 \to M$ be a homotopy from γ to μ_q such that $G(0,t) = G(1,t)$ for all t, and let $\tilde{G} : [0,1]^2 \to H^n$ be a lift of G such that $\tilde{\gamma}(s) = \tilde{G}(s,0)$ for all s. As $h\tilde{\gamma}(0) = \tilde{\gamma}(1)$, we have

$$h\tilde{G}(0,t) = \tilde{G}(1,t)$$

for all t by unique path lifting.

Let $\tilde{\mu} : \mathbb{R} \to H^n$ be the lift of μ such that $\tilde{\mu}(0) = \tilde{G}(0,1)$. Then $\tilde{G}$ is a homotopy from $\tilde{\gamma}$ to $\tilde{\mu}_q$. Hence

$$h\tilde{\mu}(0) = h\tilde{G}(0,1) = \tilde{G}(1,1) = \tilde{\mu}(q).$$

Now for each integer k, we have that γ^k is freely homotopic to μ_{kq}, and the above argument shows that $h^k\tilde{\mu}(0) = \tilde{\mu}(kq)$. Hence, we have

$$h\tilde{\mu}((k-1)q) = \tilde{\mu}(kq).$$

Therefore h maps the line segment $[\tilde{\mu}((k-1)q), \tilde{\mu}(kq)]$ to the line segment $[\tilde{\mu}(kq), \tilde{\mu}((k+1)q)]$. Thus h leaves $\tilde{\mu}(\mathbb{R})$ invariant, and so $\tilde{\mu}(\mathbb{R}) = L$. As $h\tilde{\mu}(0) = \tilde{\mu}(q)$, we have $p = q$, and μ and λ differ by a translation of $\mathbb{R}$. □

Definition: A closed curve $\gamma : [a,b] \to X$ is *simple* if and only if γ is injective on the interval $[a,b)$. A closed geodesic in a metric space X, defined by a periodic line $\lambda : \mathbb{R} \to X$, with smallest period p, is *simple* if and only if the restriction of λ to the closed interval $[0,p]$ is a simple closed curve.

Theorem 9.6.5. *Let $\gamma : [0,1] \to M$ be a hyperbolic, simple, closed curve in a complete, orientable, hyperbolic surface M. Then there is a periodic geodesic line $\lambda : \mathbb{R} \to M$ that is unique up to composition with a translation in $\mathbb{R}$, and there is a unique period p of λ such that γ is freely homotopic to the closed curve $\lambda_p : [0,1] \to M$ defined by $\lambda_p(t) = \lambda(pt)$. Furthermore p is the smallest period of λ and λ_p is simple.*

Proof: All but the last sentence of the theorem follows from Theorem 9.6.4. As in the proof of Theorem 9.6.4, let $\tilde{\gamma} : [0,1] \to H^2$ be a lift of

γ with respect to the quotient map $\pi : H^2 \to H^2/\Gamma$, and let h be the hyperbolic element of Γ such that $h\tilde{\gamma}(0) = \tilde{\gamma}(1)$. Let $C = \gamma([0,1])$. Then C is homeomorphic to S^1. Let $\tilde{C}$ be the component of $\pi^{-1}(C)$ containing $\tilde{\gamma}(0)$. Then we have

$$\tilde{C} = \cup\{h^k\tilde{\gamma}([0,1]) : k \in \mathbb{Z}\}$$

by unique path lifting.

Since γ represents an element of infinite order in $\pi_1(M)$, the covering $\tilde{C}$ of C is universal, and so $\tilde{C}$ is homeomorphic to $\mathbb{R}$. Let L be the axis of h in H^2. We now pass to the projective disk model D^2. Because of the attractive-repulsive nature of the endpoints of L in $\overline{D}^2$ with respect to h, the closure of $\tilde{C}$ in $\overline{D}^2$ is the union of $\tilde{C}$ and the two endpoints of L. Therefore, the closure of $\tilde{C}$ in $\overline{D}^2$ is homeomorphic to a closed interval whose interior is $\tilde{C}$ and whose endpoints are those of L.

Let $\tilde{\lambda} : \mathbb{R} \to D^2$ be a geodesic line parameterizing L in the same direction that h translates L, and let $p > 0$ be such that

$$h\tilde{\lambda}(t) = \tilde{\lambda}(t + p).$$

Then $\lambda = \pi\tilde{\lambda}$ is a geodesic line with period p, and γ is freely homotopic to λ_p by the proof of Theorem 9.6.4.

Let q be the smallest period of λ. We now show that $\lambda_q : [0,1] \to M$ is simple. On the contrary, suppose that λ_q is not simple. Then λ_q must cross itself transversely. Hence, there is an element g of Γ and another lift $g\tilde{\lambda} : \mathbb{R} \to D^2$ of λ such that the hyperbolic line $gL = g\tilde{\lambda}(\mathbb{R})$ intersects L at one point. As the endpoints of $\tilde{C}$ and $g\tilde{C}$ link, $\tilde{C}$ and $g\tilde{C}$ must intersect. See Figure 9.6.1. But $\tilde{C}$ and $g\tilde{C}$ are distinct components of $\pi^{-1}(C)$ and so are disjoint, which is a contradiction. Thus λ_q is simple.

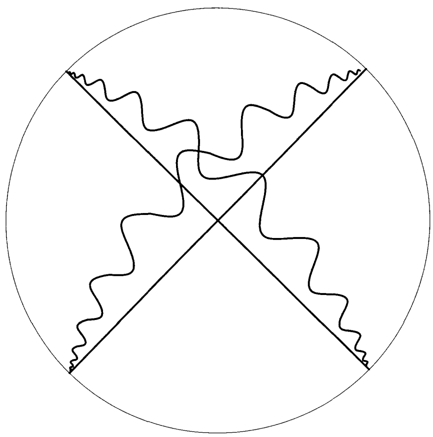


Figure 9.6.1. Lifts of two simple closed curves on a closed hyperbolic surface

Let $m = p/q$. Then $\lambda_p = \lambda_q^m$. Assume that $m > 1$. We shall derive a contradiction. Let g be the element of Γ such that $g\tilde{\lambda}(0) = \tilde{\lambda}(q)$. By unique path lifting, we have

$$\tilde{\lambda}_q g\tilde{\lambda}_q \cdots g^{m-1}\tilde{\lambda}_q = \tilde{\lambda}_p.$$

Therefore, we have

$$g^m\tilde{\lambda}(0) = g^{m-1}\tilde{\lambda}(q) = \tilde{\lambda}(p) = h\tilde{\lambda}(0).$$

Hence $h = g^m$. Consequently g has the same axis as h, and so g translates along L a distance q in the same direction as h.

Now, without loss of generality, we may assume that L is the line $(-e_2, e_2)$ of D^2. Then $\overline{\tilde{C}}$ divides $\overline{D}^2$ into two components, the left one that contains $-e_1$ and the right one that contains e_1. Observe that $g\tilde{C}$ is a component of $\pi^{-1}(C)$ different from $\tilde{C}$ and so must be in either the left or right component of $\overline{D}^2 - \overline{\tilde{C}}$. Say $g\tilde{C}$ is in the right component. Likewise $\overline{g\tilde{C}}$ divides $\overline{D}^2$ into two components, the left one that contains $-e_1$ and the right one that contains e_1. Moreover g maps the right component of $\overline{D}^2 - \overline{\tilde{C}}$ onto the right component of $\overline{D}^2 - \overline{g\tilde{C}}$ because g leaves invariant the right component of $S^1 - \{\pm e_2\}$. Hence $g^2\tilde{C}$ is in the right component of $\overline{D}^2 - \overline{\tilde{C}}$. By induction, we deduce that $g^m\tilde{C} = \tilde{C}$ is in the right component of $\overline{D}^2 - \overline{\tilde{C}}$, which is a contradiction. Therefore $m = 1$ and $p = q$. Thus γ is freely homotopic to the simple, closed, geodesic curve λ_p. □

Let $\gamma : [0,1] \to M$ be a hyperbolic, simple, closed curve in a complete orientable surface M. By Theorem 9.6.5, there is a periodic geodesic line $\lambda : \mathbb{R} \to M$, with smallest period p, that is unique up to composition with a translation in $\mathbb{R}$, such that γ is freely homotopic to $\lambda_p : [0,1] \to M$ defined by $\lambda_p(t) = \lambda(pt)$. Moreover λ_p is simple. The simple closed geodesic $\lambda(\mathbb{R})$ of M is said to *represent* the simple closed curve γ.

Definition: Two curves $\alpha, \beta : [0,1] \to X$ are *homotopically distinct* if and only if α is not freely homotopic to $\beta^{\pm 1}$.

Theorem 9.6.6. *Let $\alpha, \beta : [0,1] \to M$ be disjoint, homotopically distinct, hyperbolic, simple, closed curves in a complete, orientable, hyperbolic surface M. Then α and β are represented by disjoint, simple, closed geodesics of M.*

Proof: On the contrary, suppose that the simple closed geodesics representing α and β intersect. We may assume that M is a space-form H^2/Γ. Then there are lifts K and L of the geodesics in the universal cover H^2 that intersect. Now K and L do not coincide, since α and β are homotopically distinct. Therefore K and L intersect at one point.

Let $A = \alpha([0,1])$ and $B = \beta([0,1])$. Then there are lifts $\tilde{A}$ and $\tilde{B}$ of A and B, respectively, that have the same endpoints as K and L, respectively.

Consequently $\tilde{A}$ and $\tilde{B}$ must intersect. See Figure 9.6.1. Therefore A and B intersect, which is a contradiction. Thus, the simple closed geodesics representing α and β are disjoint. □

Theorem 9.6.7. *Let $\alpha, \beta : [0,1] \to M$ be homotopically distinct, hyperbolic, simple, closed curves in a complete, orientable, hyperbolic surface M whose images meet transversely at a single point. Then the simple closed geodesics of M, representing α and β, meet transversely at a single point.*

Proof: We may assume that M is a space-form H^2/Γ. Let $\pi : H^2 \to H^2/\Gamma$ be the quotient map. Let $A = \alpha([0,1])$, $B = \beta([0,1])$, and $\tilde{A}$ and $\tilde{B}$ be components of $\pi^{-1}(A)$ and $\pi^{-1}(B)$, respectively, such that $\tilde{A}$ and $\tilde{B}$ intersect. Let g and h be the hyperbolic elements of Γ that leave $\tilde{A}$ and $\tilde{B}$ invariant, respectively, and let K, L be the axis of g, h, respectively.

We now show that $\tilde{A}$ and $\tilde{B}$ meet transversely at a single point. As A and B meet transversely, $\tilde{A}$ and $\tilde{B}$ also meet transversely. Suppose that $\tilde{A}$ and $\tilde{B}$ meet at two points $\tilde{x}$ and $\tilde{y}$. Then $\pi(\tilde{x}) = x = \pi(\tilde{y})$. Hence, there exist nonzero integers k and ℓ such that $g^k\tilde{x} = \tilde{y} = h^\ell\tilde{x}$. Therefore $g^k = h^\ell$, and so $K = L$. Hence α and β or α and β^{-1} are homotopic by Theorem 9.6.5, which is a contradiction. Thus $\tilde{A}$ and $\tilde{B}$ meet transversely at a single point $\tilde{x}$. Therefore K and L meet at a single point $\tilde{z}$.

Next, we show that the geodesics $C = \pi(K)$ and $D = \pi(L)$, representing α and β, meet at a single point. Suppose that C and D meet at points z and w with $\pi(\tilde{z}) = z$. Let $\tilde{w}$ be a point of L such that $\pi(\tilde{w}) = w$. Then there is an element f of Γ such that fK meets L at a single point $\tilde{w}$. Consequently $f\tilde{A}$ meets $\tilde{B}$ at a point $\tilde{y}$. Then $\pi(\tilde{y}) = x$. As $\tilde{y}$ is in $\tilde{B}$, there is an integer m such that $\tilde{y} = h^m\tilde{x}$. Now since $f\tilde{A}$ and $h^m\tilde{A}$ meet at $\tilde{y}$, we have that $f\tilde{A} = h^m\tilde{A}$. Therefore $fK = h^mK$. As K and L meet at the point $\tilde{z}$, we have that h^mK and L meet at the point $h^m\tilde{z}$. Therefore $\tilde{w} = h^m\tilde{z}$. Hence $w = z$. Thus C and D meet transversely at a single point. □

Exercise 9.6

1. Let B^n/Γ be a space-form and let g and h be nonidentity elements of Γ with h hyperbolic. Prove that the following are equivalent:

   (1) The elements g and h are both hyperbolic with the same axis.

   (2) The elements g and h are both powers of the same element of Γ.

   (3) The elements g and h commute.

   (4) The elements g and h have the same fixed points in S^{n-1}.

   (5) The elements g and h have a common fixed point in S^{n-1}.

2. Let B^n/Γ be a compact space-form. Prove that every elementary subgroup of Γ is cyclic.

3. Let X be a geometric space and let $M = X/\Gamma$ be a space-form. Let $\lambda : \mathbb{R} \to M$ be a periodic geodesic line with smallest period p. Prove that there are only finitely many numbers t in the interval $[0, p]$ such that $\lambda(t) = \lambda(s)$ with $0 \leq s < t$. Conclude that a closed geodesic of M intersects itself only finitely many times.

4. Let $X = S^n, E^n$, or H^n, and let $M = X/\Gamma$ be a space-form. Let $\pi : X \to X/\Gamma$ be the quotient map. Prove that a closed geodesic C of M is simple if and only if $\pi^{-1}(C)$ is a disjoint union of geodesics of X.

5. Let $\gamma : [0,1] \to M$ be an essential closed curve in a complete Euclidean n-manifold M. Prove that there is a periodic geodesic line $\lambda : \mathbb{R} \to M$ and a unique period p of λ such that γ is freely homotopic to the closed curve $\lambda_p : [0,1] \to M$ defined by $\lambda_p(t) = \lambda(pt)$.

6. Let $\gamma : [0,1] \to M$ be an essential, simple, closed curve in a complete, orientable, Euclidean surface M. Prove that there is a periodic geodesic line $\lambda : \mathbb{R} \to M$ and a unique period p of λ such that γ is freely homotopic to the closed curve $\lambda_p : [0,1] \to M$ defined by $\lambda_p(t) = \lambda(pt)$. Furthermore p is the smallest period of λ and λ_p is simple.

7. Let γ and λ_p be as in Theorem 9.6.4. Prove that $|\lambda_p| \leq |\gamma|$. Conclude that λ_p has minimal length in its free homotopy class.

8. Prove that the infimum of the set of lengths of essential closed curves in a compact hyperbolic n-manifold M is positive.

9. Let X be a geometric space and let $M = X/\Gamma$ be a space-form. Let $\lambda, \mu : \mathbb{R} \to M$ be periodic geodesic lines such that $\lambda(\mathbb{R}) = \mu(\mathbb{R})$. Prove that there is an isometry ξ of $\mathbb{R}$ such that $\mu = \lambda\xi$. Conclude that the *length* of the closed geodesic $\lambda(\mathbb{R})$ is well defined to be the smallest period of λ.

10. Let X be a geometric space and let $M = X/\Gamma$ be a compact space-form. Prove that for each $\ell > 0$, there are only finitely many closed geodesics in M of length less than ℓ.

§9.7. Closed Hyperbolic Surfaces

In this section, we describe the Teichmüller space of a closed orientable surface of genus $n > 1$. The next theorem is a basic theorem of the topology of closed surfaces.

Theorem 9.7.1. *If M is a closed orientable surface of genus $n > 1$, then*

(1) *the maximum number of disjoint, homotopically distinct, essential, simple, closed curves in M is $3n - 3$; and*

(2) *the complement in M of a maximal number of disjoint, homotopically distinct, essential, simple, closed curves in M is the disjoint union of $2n - 2$ surfaces each homeomorphic to S^2 minus three disjoint closed disks.*

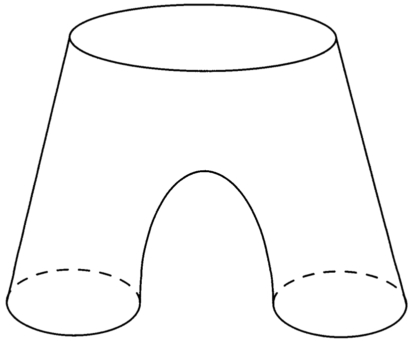


Figure 9.7.1. A pair of pants

Pairs of Pants

We shall call a space P homeomorphic to the complement in S^2 of three disjoint open disks a *pair of pants*. See Figure 9.7.1. A pair of pants is a compact orientable surface-with-boundary whose boundary consists of three disjoint topological circles. By Theorems 9.6.6 and 9.7.1, a closed, orientable, hyperbolic surface M of genus $n > 1$ can be subdivided by $3n - 3$ disjoint, simple, closed geodesics into the union of $2n - 2$ pairs of pants with the geodesics as their boundary circles. See Figure 9.7.2.

Let P be a pair of pants in a hyperbolic surface M such that each boundary circle of P is a simple closed geodesic of M. A *seam* of P is defined to be the image S of an injective geodesic curve $\sigma : [a, b] \to M$ such that the point $\sigma(a)$ is in a boundary circle A of P, the point $\sigma(t)$ is in the interior of P for $a < t < b$, the point $\sigma(b)$ is in another boundary circle B of P, and the geodesic section S is perpendicular to both A and B.

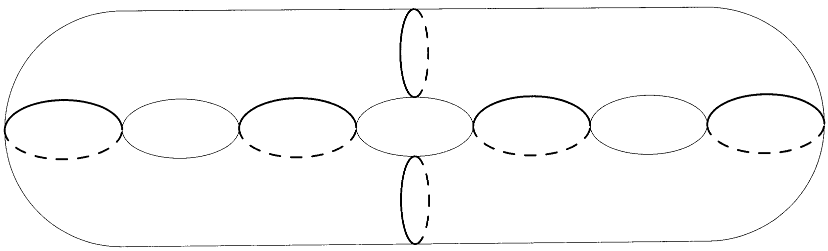


Figure 9.7.2. A maximal number of disjoint, homotopically distinct, essential, simple, closed curves on a closed orientable surface of genus three

Theorem 9.7.2. *Let P be a pair of pants in a hyperbolic surface M such that each boundary circle of P is a simple closed geodesic of M. Then any two boundary circles of P are joined by a unique seam of P. Moreover, the three seams of P are mutually disjoint.*

Proof: Let P' be a copy of P. For each point x of P, let x' be the corresponding point of P'. Let Q be the quotient space obtained from the disjoint union of P and P' by identifying x with x' for each point x of ∂P. We regard Q to be the union of P and P' with

$$\partial P = P \cap P' = \partial P'.$$

The space Q is a closed orientable surface of genus two called the *double* of P. See Figure 9.7.3.

Let A, B, C be the boundary circles of P. The hyperbolic structures on the interiors of P and P' extend to a hyperbolic structure on Q so that A, B, C are closed geodesics of Q. The hyperbolic surface Q is complete, since Q is compact.

Let $\alpha : [0, 1] \to P$ be a simple curve such that the point $\alpha(0)$ is in A, the point $\alpha(t)$ is in the interior of P for $0 < t < 1$, and the point $\alpha(1)$ is in B. Let α' be the corresponding simple curve in P'. Then $\alpha\alpha'^{-1}$ is an essential, simple, closed curve in Q. Hence $\alpha\alpha'^{-1}$ is freely homotopic to a simple closed curve δ whose image is a simple closed geodesic D in Q by Theorem 9.6.5. Now by Theorem 9.6.7, the geodesic D meets the geodesics A and B transversely in single points. Let $S = D \cap P$. Then S is a section of D contained in P joining A to B.

Let $\rho : Q \to Q$ be the map defined by $\rho(x) = x'$ and $\rho(x') = x$ for each point x of P. Then ρ is an isometry of Q. Observe that

$$\rho(\alpha\alpha'^{-1}) = \alpha'\alpha^{-1}.$$

Hence $\alpha'\alpha^{-1}$ is freely homotopic to $\rho\delta$, and ρD is the simple closed geodesic of Q that represents $\alpha'\alpha^{-1}$. Therefore $\rho D = D$ by Theorem 9.6.5. Consequently D is perpendicular to both A and B. Hence S is perpendicular to A and B. Thus S is a seam of P joining A to B.

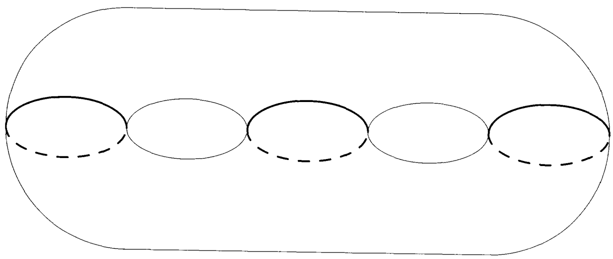


Figure 9.7.3. The double of a pair of pants

Now suppose that T is another geodesic section in P joining A to B that is perpendicular to A and B. Then $E = T \cup T'$ is a simple closed geodesic of Q. Let $\sigma, \tau : [0,1] \to P$ be simple curves starting in A whose images are S, T, respectively. Then σ is freely homotopic to τ by a homotopy keeping the endpoints on A and B. Hence $\sigma\sigma'^{-1}$ is freely homotopic to $\tau\tau'^{-1}$. Therefore $D = E$ by Theorem 9.6.5. Hence $S = T$. Thus, the seam S is unique.

Now suppose that T is the seam of P joining A to C. Let $\beta : [0,1] \to P$ be a simple curve such that the point $\beta(0)$ is in A, the point $\beta(t)$ is in the interior of P for $0 < t < 1$, the point $\beta(1)$ is in C, and the image of β is disjoint from the image of α. Then $\alpha\alpha'^{-1}$ and $\beta\beta'^{-1}$ are essential, homotopically distinct, disjoint, simple, closed curves in Q. Therefore, the simple closed geodesics representing them, D and $T \cup T'$, are disjoint by Theorem 9.6.6. Thus S and T are disjoint. □

Let P be a pair of pants in a hyperbolic surface M such that each boundary circle of P is a simple closed geodesic of M. If we split P apart along its seams, we find that P is the union of two subsets D_1 and D_2, meeting along the seams of P, each of which is homeomorphic to a disk. The boundary of each D_i is the union of six geodesic sections meeting only along their endpoints at right angles.

By replacing M with the double of P, we may assume that M is complete. Therefore, we may assume that M is a space-form H^2/Γ. Let $\pi : H^2 \to H^2/\Gamma$ be the quotient map and let H_i be a component of $\pi^{-1}(D_i)$ for $i = 1, 2$. As D_i is simply connected, π maps H_i homeomorphically onto D_i for $i = 1, 2$. The set H_i is a closed, connected, locally convex subset of H^2 and so is convex. Hence H_i is a convex hexagon in H^2 all of whose angles are right angles. Thus P can be obtained by gluing together two right-angled, convex, hyperbolic hexagons along alternate sides.

Theorem 9.7.3. *Let P be a pair of pants in a hyperbolic surface M such that each boundary circle of P is a simple closed geodesic of M. Let a, b, c be the lengths of the boundary circles of P and let H_1, H_2 be the right-angled, convex, hyperbolic hexagons obtained from P by splitting P along its seams. Then H_1 and H_2 are congruent with nonseam alternate sides of length $a/2, b/2, c/2$, respectively. Moreover P is determined, up to isometry, by the lengths a, b, c.*

Proof: As H_1 and H_2 have the same lengths for their seam alternate sides, H_1 and H_2 are congruent by Theorem 3.5.14. Hence H_1 and H_2 have the same lengths for their nonseam alternate sides. As these lengths add up to a, b, c, respectively, we find that the nonseam alternate sides of H_1 and H_2 have length $a/2, b/2, c/2$, respectively. As H_1 and H_2 are determined, up to congruence, by the lengths $a/2, b/2, c/2$, we deduce that P is determined, up to isometry, by the lengths a, b, c. □

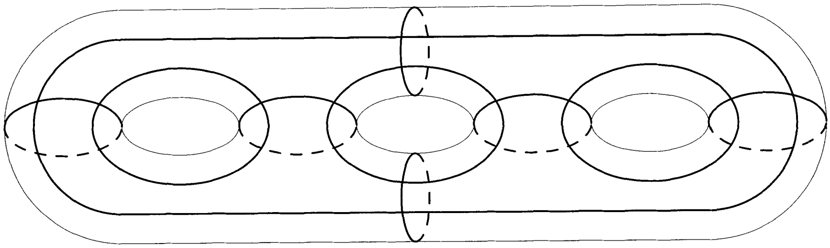


Figure 9.7.4. A marked, closed, orientable surface of genus three

Teichmüller Space

Let M be a closed orientable surface of genus $n > 1$. We mark M by choosing $3n - 3$ disjoint, homotopically distinct, essential, simple, closed curves $\alpha_i : [0, 1] \to M$, for $i = 1, \ldots, 3n - 3$, and $n + 1$ more disjoint, homotopically distinct, essential, simple, closed curves $\beta_j : [0, 1] \to M$, for $j = 1, \ldots, n + 1$, which together with the first set of curves divides M into closed disks as in Figure 9.7.4. Observe that the first set of curves divides M into pairs of pants and that the second set of curves forms a continuous set of topological seams for the pairs of pants.

Let Φ be a hyperbolic structure for M. By Theorem 9.6.6, the curves $\alpha_1, \ldots, \alpha_{3n-3}$ are represented by $3n - 3$ disjoint, simple, closed geodesics $A_1, \ldots, A_{3n-3}$ of (M, Φ). By Theorem 9.7.1, these geodesics divide M into $2n - 2$ pairs of pants. By Theorem 9.7.3, these pairs of pants are determined, up to isometry, by the lengths of their boundary circles. Let ℓ_i be the length of A_i for each $i = 1, \ldots, 3n - 3$.

In order to determine the isometry type of (M, Φ) from that of the pairs of pants, we need to measure the amount of twist with which the boundary circles of the pairs of pants are attached. We use the curves $\beta_1, \ldots, \beta_{n+1}$ to measure these twists. By Theorem 9.6.6, the curves $\beta_1, \ldots, \beta_{n+1}$ are represented by $n + 1$ disjoint, simple, closed geodesics $B_1, \ldots, B_{n+1}$. In the pairs of pants, these geodesics restrict to geodesic sections joining the boundary circles because of Theorem 9.6.7. Furthermore, in the pairs of pants, these geodesic sections are homotopic to the seams of the pairs of pants by homotopies keeping the endpoints on the curves $A_1, \ldots, A_{3n-3}$.

We orient M. This orients all the pairs of pants of M. Let P_i and Q_i be the pairs of pants of M with A_i as a boundary circle, and suppose that the orientation of A_i agrees with the orientation of P_i. Let $2a_i$ be the total radian measure that the above homotopies move, within P_i, the two endpoints on A_i. The number a_i measures the degree to which the two geodesic sections wrap around the two seams of P_i ending in A_i and

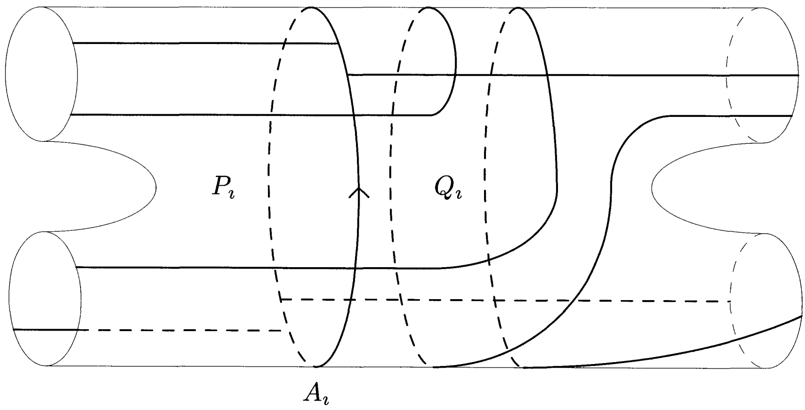


Figure 9.7.5. The four geodesic sections and seams ending in the geodesic A_i

is called the *winding degree* of (P_i, A_i). See Figure 9.7.5. The winding degree a_i does not depend on the choice of the homotopies. Let b_i be the winding degree of (Q_i, A_i). The real number $t_i = a_i - b_i$ is called the *twist coefficient* of A_i. The twist coefficient t_i measures the twist with which P_i and Q_i are attached at A_i relative to the given marking of M. Note that t_i is congruent modulo 2π to the angle that Q_i must rotate around A_i so that the corresponding seams of P_i and Q_i match up. See Figure 9.7.5.

Define a function

$$F : \mathcal{H}(M) \to \mathbb{R}^{6n-6}$$

by setting

$$F(\Phi) = (\log \ell_1, t_1, \log \ell_2, t_2, \ldots, \log \ell_{3n-3}, t_{3n-3}). \tag{9.7.1}$$

We shall call the components of $F(\Phi)$ the *length-twist coordinates* of the hyperbolic structure Φ for M.

Theorem 9.7.4. *Let M be a closed orientable surface of genus $n > 1$. Then the function $F : \mathcal{H}(M) \to \mathbb{R}^{6n-6}$ induces a bijection from $\mathcal{T}(M)$ to $\mathbb{R}^{6n-6}$.*

Proof: Let $h : M \to M$ be a homeomorphism that is homotopic to the identity map of M. Then h is an isometry from $(M, \Phi h)$ to (M, Φ). Consequently $h^{-1}A_i$ is a simple closed geodesic of $(M, \Phi h)$ for all i. As h^{-1} is homotopic to the identity map, $h^{-1}A_i$ is freely homotopic to A_i for each i. Hence, the curves $\alpha_1, \ldots, \alpha_{3n-3}$ are represented in $(M, \Phi h)$ by the geodesics $h^{-1}A_1, \ldots, h^{-1}A_{3n-3}$. Likewise, the curves $\beta_1, \ldots, \beta_{n+1}$ are represented in $(M, \Phi h)$ by the geodesics $h^{-1}B_1, \ldots, h^{-1}B_{n+1}$. As h^{-1} is an isometry, $h^{-1}A_i$ has the same length and twist coefficient as A_i for each i. Thus $F(\Phi h) = F(\Phi)$. Therefore F induces a function

$$\overline{F} : \mathcal{T}(M) \to \mathbb{R}^{6n-6}.$$

Next, we show that $\overline{F}$ is injective. Suppose that Φ and Φ' are hyperbolic structures for M such that $F(\Phi) = F(\Phi')$. Let $A_1, \ldots, A_{3n-3}$ be the simple closed geodesics in (M, Φ) representing $\alpha_1, \ldots, \alpha_{3n-3}$, and let $A'_1, \ldots, A'_{3n-3}$ be the simple closed geodesics in (M, Φ') representing $\alpha_1, \ldots, \alpha_{3n-3}$. Then A_i has the same length and twist coefficient as A'_i for each i. By Theorem 9.7.3, there is an isometry $h : (M, \Phi') \to (M, \Phi)$ mapping the geodesic A'_i onto the geodesic A_i for each i.

Let $B_1, \ldots, B_{n+1}$ be the simple closed geodesics in (M, Φ) representing $\beta_1, \ldots, \beta_{n+1}$, and let $B'_1, \ldots, B'_{n+1}$ be the simple closed geodesics in (M, Φ') representing $\beta_1, \ldots, \beta_{n+1}$. Now the sets $h(B'_1), \ldots, h(B'_{n+1})$ are simple closed geodesics in (M, Φ) that form a continuous set of topological seams for the pairs of pants of (M, Φ) and twist the same amount about the geodesics $A_1, \ldots, A_{3n-3}$ as the continuous set of topological seams $B_1, \ldots, B_{n+1}$. Consequently $h(B'_j)$ is freely homotopic to B_j for each j. Therefore $h(B'_j) = B_j$ for each j by Theorem 9.6.5.

Regard the geodesics $A'_1, \ldots, A'_{3n-3}$ and $B'_1, \ldots, B'_{n+1}$ as forming the 1-skeleton M^1 of a cell structure for M. Let h_1 be the restriction of h to M^1. Then we can construct a homotopy from h_1 to the inclusion map of M^1 into M, since A_i is freely homotopic to A'_i for each i and B_j is freely homotopic to B'_j by a homotopy consistent with the first set of homotopies for each j. Now since $\pi_2(M) = 0$, the homotopy of h_1 to the inclusion of M^1 into M can be extended to a homotopy of h to the identity map of M. As $\Phi' = \Phi h$, we have that $[\Phi'] = [\Phi]$ in $\mathcal{T}(M)$. Thus $\overline{F}$ is injective.

Next, we show that $\overline{F}$ is surjective. Let $(s_1, t_1, \ldots, s_{3n-3}, t_{3n-3})$ be a point of $\mathbb{R}^{6n-6}$ and set $\ell_i = e^{s_i}$ for $i = 1, \ldots, 3n-3$. By Theorem 3.5.14, there are $4n-4$ right-angled, convex, hyperbolic hexagons that can be glued together in pairs along alternate sides to give $2n-2$ pairs of pants whose $6n-6$ boundary circles have length $\ell_1, \ell_1, \ell_2, \ell_2, \ldots, \ell_{3n-3}, \ell_{3n-3}$, respectively, and which are in one-to-one correspondence with the $2n-2$ pairs of pants of M in such a way that the indexing of the lengths of the boundary circles of each of the pairs of pants corresponds to the indexing of the boundary circles of the corresponding pair of pants of M. Write $t_i = \theta_i + 2\pi k_i$, with $0 \leq \theta_i < 2\pi$ and k_i an integer. Let M' be the surface obtained by gluing together the $2n-2$ pairs of pants along the two boundary circles of length ℓ_i with a twist of θ_i for each i. By Theorem 9.2.3, the surface M' has a hyperbolic structure such that the circle C_i in M', obtained by gluing the two boundary circles of length ℓ_i, is a simple closed geodesic of length ℓ_i for each i. Furthermore, the one-to-one correspondence between the pairs of pants of M and M' extends to a homeomorphism $h : M \to M'$ mapping $\alpha_i([0,1])$ onto C_i for each i.

Let $\Phi = \{\phi_i : U_i \to H^2\}$ be the hyperbolic structure of M'. Then

$$\Phi h = \{\phi_i h : h^{-1}(U_i) \to H^2\}$$

is a hyperbolic structure for M such that h is an isometry from $(M, \Phi h)$ to (M', Φ). Let $A_i = \alpha_i([0,1])$ for each i. Then A_i is a simple closed geodesic of $(M, \Phi h)$ of length ℓ_i that represents α_i for each i. Moreover, the twist coefficient of A_i is congruent to θ_i modulo 2π. Hence, by replacing h with h composed with an appropriate number of Dehn twists about C_i for each i, we can assume that the twist coefficient of A_i is t_i for each i. Then

$$F(\Phi h) = (s_1, t_1, \ldots, s_{3n-3}, t_{3n-3}).$$

Hence $\overline{F}$ is surjective. Thus $\overline{F}$ is a bijection. □

Remark: It is a fundamental theorem of Teichmüller space theory that the bijection $\overline{F} : \mathcal{T}(M) \to \mathbb{R}^{6n-6}$ is a homeomorphism.

Corollary 1. *The moduli space $\mathcal{M}(M)$ of a closed orientable surface M of genus $n > 1$ is uncountable.*

Proof: Since $\pi_1(M)$ is finitely generated, the group $\mathrm{Out}(\pi_1(M))$ is countable. Therefore, the mapping class group $\mathrm{Map}(M)$ is countable, since the Nielsen homomorphism

$$\nu : \mathrm{Map}(M) \to \mathrm{Out}(\pi_1(M))$$

is injective. By Theorem 9.7.4, we have that $\mathcal{T}(M)$ is uncountable, and so the set $\mathcal{T}(M)/\mathrm{Map}(M)$ is uncountable. As there is a bijection from the set $\mathcal{T}(M)/\mathrm{Map}(M)$ to $\mathcal{M}(M)$, we have that $\mathcal{M}(M)$ is uncountable. □

Exercise 9.7

1. Prove Theorem 9.7.1.
2. Prove that the hyperbolic structure in the interior of a pair of pants extends to a unique hyperbolic structure on its double.
3. Let P be a pair of pants with boundary circles A, B, C and let $\alpha, \beta : [0,1] \to P$ be simple curves whose images are geodesic sections that begin in A, end in B, and are otherwise disjoint from A, B, C. Prove that α is freely homotopic to β by a homotopy that keeps the endpoints in A and B.
4. Let $\tilde{M}$ be a marked, closed, orientable surface of genus $n-1$ embedded in $\mathbb{R}^3$ so that the β_j curves all lie on the xy-plane, the α_i curves lie either on the xz-plane or on planes parallel to the yz-plane, and $\tilde{M}$ and its marking are invariant under a $180°$ rotation ϕ about the z-axis and the reflection ρ in the xy-plane. Let $\sigma = \rho\phi$ and let $\Gamma = \{I, \sigma\}$. Prove that $M = \tilde{M}/\Gamma$ is a closed nonorientable surface of genus n.
5. Let $\ell > 0$. Prove that $\tilde{M}$ in Exercise 4 has a hyperbolic structure $\tilde{\Phi}_\ell$ whose length-twist coordinates are $\log \ell, 0, \ldots, \log \ell, 0$, and such that ϕ and ρ are isometries. Conclude that $\tilde{\Phi}_\ell$ induces a hyperbolic structure Φ_ℓ on M.
6. Prove that the moduli space $\mathcal{M}(M)$ of a closed nonorientable surface M of genus $n > 2$ is uncountable.

§9.8. Hyperbolic Surfaces of Finite Area

In this section, we study the geometry of complete hyperbolic surfaces of finite area. We begin by determining the geometry of exact, convex, fundamental polygons of finite area.

Theorem 9.8.1. *Let P be an exact, convex, fundamental polygon of finite area for a discrete group Γ of isometries of H^2. Then P has only a finite number of sides and the sides of P can be cyclically ordered so that any two consecutive sides meet either in H^2 or at infinity.*

Proof: We pass to the projective disk model D^2. Let $\overline{P}$ be the closure of P in E^2 and suppose that $\overline{P}$ contains m points on S^1. Then $\overline{P}$ contains the convex hull $\overline{Q}$ of these m points. The set $Q = \overline{Q} \cap D^2$ is an ideal polygon with m sides. As Q can be subdivided into $m-2$ ideal triangles,

$$\text{Area}(Q) = (m-2)\pi.$$

As P contains Q and the area of P is finite, there must be an upper bound on the number of points of $\overline{P}$ on S^1. Thus $\overline{P}$ contains only finitely many points on S^1.

Let $\theta(v)$ be the angle subtended by P at a vertex v. Suppose that $v_1, \ldots, v_n$ are finite vertices of P and R is the convex hull of $v_1, \ldots, v_n$. Then R is a compact convex polygon with n sides. As R can be subdivided into $n-2$ triangles, we deduce that

$$\text{Area}(R) = (n-2)\pi - \sum_{i=1}^{n} \theta(v_i).$$

Therefore, we have

$$2\pi + \text{Area}(R) = \sum_{i=1}^{n} (\pi - \theta(v_i)).$$

Consequently

$$2\pi + \text{Area}(P) \geq \sum \{\pi - \theta(v) : v \text{ is a vertex of } P\}.$$

Hence, the sum $\sum_v (\pi - \theta(v))$ converges. Let

$$A = \{v : \theta(v) \leq 2\pi/3\}$$

and

$$B = \{v : \theta(v) > 2\pi/3\}.$$

Then A is a finite set, since the sum $\sum_v (\pi - \theta(v))$ converges.

Now the Γ-side-pairing of P induces an equivalence relation on the vertices of P whose equivalence classes are called cycles of vertices. Each cycle C of vertices is finite by Theorem 6.7.5 and corresponds to a cycle of sides of P, and so by Theorem 6.7.7, the angle sum

$$\theta(C) = \sum \{\theta(v) : v \in C\}$$

is a submultiple of 2π. Consequently, each cycle C of vertices contains at most two vertices from the set B and at least one vertex from the set A. Therefore, there are only finitely many cycles of vertices. As each cycle of vertices is finite, P has only finitely many vertices. This, together with the fact that $\overline{P} \cap S^1$ is finite, implies that P has only finitely many sides and the sides of P can be cyclically ordered so that any two consecutive sides meet either in D^2 or at an ideal vertex on the circle S^1 at infinity. □

We now determine the topology of a complete hyperbolic surface of finite area.

Theorem 9.8.2. *Let M be a complete hyperbolic surface of finite area. Then M is homeomorphic to a closed surface minus a finite number of points and*

$$\text{Area}(M) = -2\pi\chi(M).$$

Proof: Since M is complete, we may assume that M is a space-form H^2/Γ. Let P be an exact, convex, fundamental polygon for Γ. As

$$\text{Area}(P) = \text{Area}(H^2/\Gamma),$$

we have that P has finite area. By Theorem 9.8.1, the polygon P has only finitely many sides and the sides of P can be cyclically ordered so that any two consecutive sides meet either in H^2 or at infinity. We now pass to the projective disk model D^2. Let $\overline{P}$ be the closure of P in E^2. Then $\overline{P}$ is a compact convex polygon in E^2. By Theorem 6.5.8, the surface M is homeomorphic to the space P/Γ obtained from P by gluing together the sides of P paired by elements of Γ. This pairing extends to a side-pairing of $\overline{P}$. Let $\overline{P}/\Gamma$ be the space obtained from $\overline{P}$ by gluing together the sides of $\overline{P}$ paired by elements of Γ. Then $\overline{P}/\Gamma$ is a closed surface and P/Γ is homeomorphic to $\overline{P}/\Gamma$ minus the images of the ideal vertices of $\overline{P}$. Thus M is homeomorphic to a closed surface minus a finite number of points.

Now P/Γ is a cell complex, with some 0-cells removed, consisting of a 0-cells, b 1-cells, and one 2-cell. Let $v_1, \ldots, v_m$ be the finite vertices of P and let n be the number of sides of P. As P can be subdivided into $n-2$ generalized triangles, we deduce that

$$\begin{aligned}\text{Area}(P) &= (n-2)\pi - \sum_{i=1}^{m} \theta(v_i) \\ &= (2b-2)\pi - 2\pi a \\ &= -2\pi(a-b+1) \quad = \quad -2\pi\chi(P/\Gamma).\end{aligned}$$

Thus, we have that

$$\text{Area}(M) = -2\pi\chi(M).$$ □

Complete Gluing of Hyperbolic Surfaces

Let M be a hyperbolic 2-manifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, convex, finite-sided polygons in H^2 of finite area by a proper $I(H^2)$-side-pairing Φ. We shall determine necessary and sufficient conditions such that M is complete.

It will be more convenient for us to work in the conformal disk model B^2. Then the sides of each polygon in $\mathcal{P}$ can be cyclically ordered so that any two consecutive sides meet either in B^2 or at an ideal vertex on the circle S^1 at infinity. We may assume, without loss of generality, that no two polygons in $\mathcal{P}$ share an ideal vertex. Then the side-pairing Φ of the sides $\mathcal{S}$ of the polygons in $\mathcal{P}$ extends to a pairing of the ideal vertices of the polygons in $\mathcal{P}$. The pairing of the ideal vertices of the polygons in $\mathcal{P}$ generates an equivalence relation whose equivalence classes are called *cycles*. If v is an ideal vertex, we denote the cycle containing v by $[v]$.

Let v be an ideal vertex of a polygon P_v in $\mathcal{P}$. Then we can write

$$[v] = \{v_1, v_2, \ldots, v_m\}$$

with

$$v = v_1 \simeq v_2 \simeq \cdots \simeq v_m \simeq v.$$

Define sides $S_1, \ldots, S_m$ in $\mathcal{S}$ inductively as follows: Let S_1 be a side in $\mathcal{S}$ such that $g_{S_1}(v_2) = v_1$. Then v_1 is an ideal endpoint of S_1. Suppose that sides $S_1, \ldots, S_{j-1}$ have been defined so that v_i is an ideal endpoint of S_i and $g_{S_i}(v_{i+1}) = v_i$ for $i = 1, \ldots, j-1$. As $g_{S_{j-1}}(S'_{j-1}) = S_{j-1}$, we have that v_j is an ideal endpoint of S'_{j-1}. Let S_j be the other side in $\mathcal{S}$ whose ideal endpoint is v_j. Then $g_{S_j}(v_{j+1}) = v_j$ if $j < m$, and $g_{S_m}(v_1) = v_m$ if $j = m$. Thus $S_1, \ldots, S_m$ are defined. The sequence $\{S_i\}_{i=1}^m$ is called a *cycle of unbounded sides* corresponding to the cycle $[v]$ of ideal vertices.

Example 1. Let P be the ideal square in B^2 with vertices $\pm e_1$ and $\pm e_2$. Pair the opposite sides of P by first reflecting in the lines $y = \pm x$ and then reflecting in the corresponding side of P. This $I_0(B^2)$-side-pairing Φ is proper. The hyperbolic surface M obtained by gluing together the opposite sides of P by Φ is a once-punctured torus. Figure 9.8.1 illustrates the cycle of vertices of P and the corresponding cycle of unbounded sides.

Choose $\epsilon > 0$ so that the Euclidean ϵ-neighborhoods of the ideal vertices $v_1, \ldots, v_m$ are disjoint and meet only two sides in $\mathcal{S}$. Let P_i be the polygon in $\mathcal{P}$ containing the side S_i. Choose a point x_1 of S_1 so that the horocycle based at v_1 passing through x_1 is contained in $B(v_1, \epsilon)$. See Figure 9.8.2. The horocycle intersects P_1 in a horoarc α_1 that is perpendicular to the sides S'_m and S_1. Since $g_{S_1}^{-1}$ is continuous at v_1, we can choose x_1 closer to v_1, if necessary, so that the horocycle based at v_2 passing through the point $x'_1 = g_{S_1}^{-1}(x_1)$ is contained in $B(v_2, \epsilon)$. This horocycle intersects P_2 in a horoarc α_2 that is perpendicular to S'_1 and S_2. Let x_2 be the endpoint

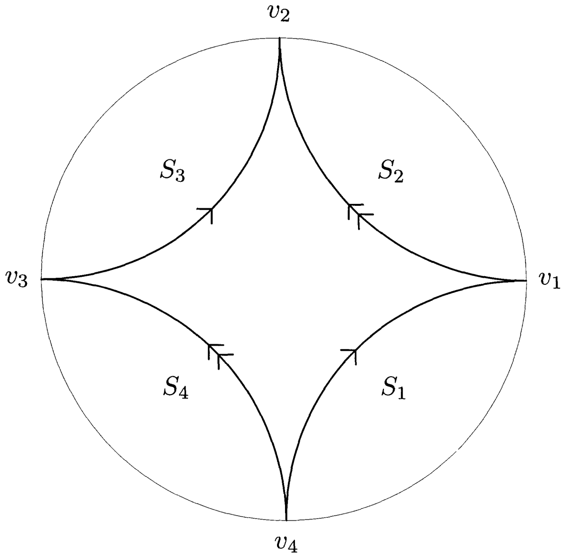


Figure 9.8.1. The cycle of sides of an ideal square with opposite sides paired

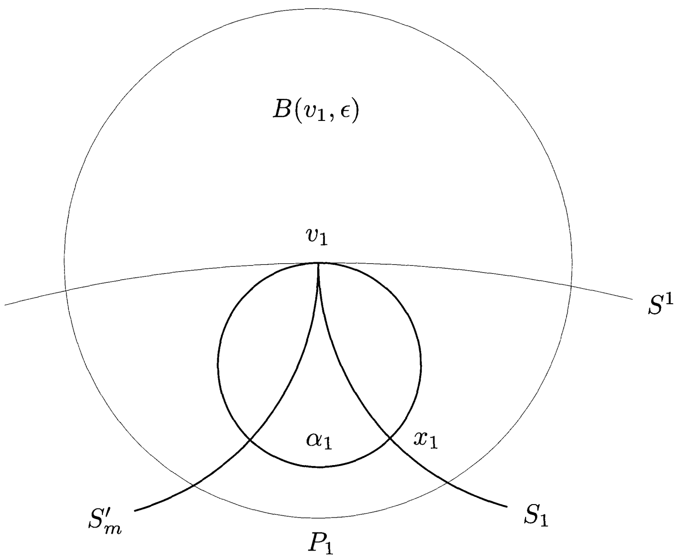


Figure 9.8.2. The horocycle based at v_1 passing through the point x_1

of α_2 in S_2. Continuing in this way, we construct a sequence of points $x_1, \ldots, x_m$ and horoarcs $\alpha_1, \ldots, \alpha_m$ such that x_i is an endpoint of α_i for $i = 1, \ldots, m$, and x'_{i-1} is an endpoint of α_i for $i = 2, \ldots, m$, and α_i is contained in $B(v_i, \epsilon)$ for $i = 1, \ldots, m$.

Let x'_0 be the endpoint of α_1 in S'_m. Define $d(v)$ to be $\pm d(x'_m, x'_0)$ with the sign positive if and only if x'_m is further away from v than x'_0. The real number $d(v)$ does not depend on the choice of x_1 because if $y_1, \ldots, y_m$ is another such sequence of points, then

$$d(x'_0, y'_0) = d(x_1, y_1) = d(x'_1, y'_1) = \cdots = d(x_m, y_m) = d(x'_m, y'_m)$$

and so

$$\begin{aligned} \pm d(x'_m, x'_0) &= \pm d(x'_m, y'_m) \pm d(y'_m, x'_0) \\ &= \pm d(x'_0, y'_0) \pm d(y'_m, x'_0) \\ &= \pm d(y'_m, y'_0). \end{aligned}$$

The real number $d(v)$ is called the *gluing invariant* of the ideal vertex v. For example, the gluing invariant of the ideal vertex v_1 in Figure 9.8.1 is zero.

Set $g_v = g_{S_1} \cdots g_{S_m}$. The element g_v is called the *cycle transformation* of the cycle of unbounded sides $\{S_i\}_{i=1}^m$. As $g_{S_i}(v_{i+1}) = v_i$ and $g_{S_m}(v_1) = v_m$, we have that g_v fixes v.

Theorem 9.8.3. *The gluing invariant $d(v)$ is zero if and only if the cycle transformation g_v is parabolic.*

Proof: Let f_i be the parabolic element of $\mathrm{I}(B^2)$ that fixes v_i and maps x_i to x'_{i-1} for $i = 1, \ldots, m$, and set $g_i = g_{S_i}$ for each i. As $g_i(v_{i+1}) = v_i$, $g_m(v_1) = v_m$, and $g_i(x'_i) = x_i$, we have that $f_1 g_1 \cdots f_m g_m$ fixes v and

$$f_1 g_1 \cdots f_m g_m(x'_m) = x'_0.$$

Suppose that $d(v) = 0$. Then $x'_m = x'_0$. Hence $f_1 g_1 \cdots f_m g_m$ fixes the side S'_m. Therefore $f_1 g_1 \cdots f_m g_m$ is either the reflection in S'_m or the identity map. Now g_i maps the side of S'_i containing P_{i+1} to the side of S_i not containing P_i for $i = 1, \ldots m$, and $P_{m+1} = P_1$; moreover, f_i maps the side of S_i not containing P_i to the side of S'_{i-1} containing P_i for $i = 1, \ldots, m$, and $S'_0 = S'_m$. Hence $f_1 g_1 \cdots f_m g_m$ maps the side of S'_m containing P_1 to the side of S'_m containing P_1. Therefore $f_1 g_1 \cdots f_m g_m$ must be the identity map. Now observe that

$$\begin{aligned} g_v^{-1} &= (f_1 g_1 \cdots f_m g_m)(g_m^{-1} \cdots g_1^{-1}) \\ &= \prod_{i=1}^{m} (g_1 \cdots g_{i-1} f_i g_{i-1}^{-1} \cdots g_1^{-1}). \end{aligned}$$

As each term of the above product is parabolic, with fixed point v, we have that g_v is parabolic with fixed point v.

Conversely, suppose that g_v is parabolic. Then from the last equation, we deduce that $f_1 g_1 \cdots f_m g_m$ is either parabolic, with fixed point v, or the identity map. As $f_1 g_1 \cdots f_m g_m$ leaves invariant the hyperbolic line containing S'_m, we have that $f_1 g_1 \cdots f_m g_m$ is the identity map. Therefore $x'_m = x'_0$ and so $d(v) = 0$. □

Theorem 9.8.4. *Let Γ_v be the group generated by the cycle transformation g_v. If g_v is parabolic, then there is a horodisk $B(v)$ based at v and an injective local isometry*

$$\iota : B(v)/\Gamma_v \to M$$

compatible with the projection of the polygon P_v to M.

Proof: We pass to the upper half-plane model U^2 and assume, without loss of generality, that $v = \infty$. Then g_v is a horizontal translation of U^2. Let $B(v)$ be the open horodisk based at v with the horoarc α_1 on its boundary. Then Γ_v acts freely and discontinuously on $B(v)$ as a group of isometries. Consequently $B(v)/\Gamma_v$ is a hyperbolic surface.

We now find a fundamental domain for Γ_v in $B(v)$. Define $g_1 = 1$ and $g_i = g_{S_1} \cdots g_{S_{i-1}}$ for $i = 2, \ldots, m$. As the polygons P_i and $g_{S_i}(P_{i+1})$ lie on opposite sides of their common side S_i for $i = 1, \ldots, m-1$, the polygons $g_i P_i$ and $g_{i+1} P_{i+1}$ lie on opposite sides of their common side $g_i S_i$ for $i = 1, \ldots, m-1$. Thus, the rectangular strips $g_i P_i \cap B(v)$ lie adjacent to each other in sequential order. See Figure 9.8.3. As g_v translates the side S'_m of $g_1 P_1$ onto the side $g_m S_m$ of $g_m P_m$, we see that the rectangular strip

$$\bigcup_{i=1}^{m} g_i P_i \cap B(v)$$

is the closure of a fundamental domain D for Γ_v in $B(v)$; moreover D is locally finite.

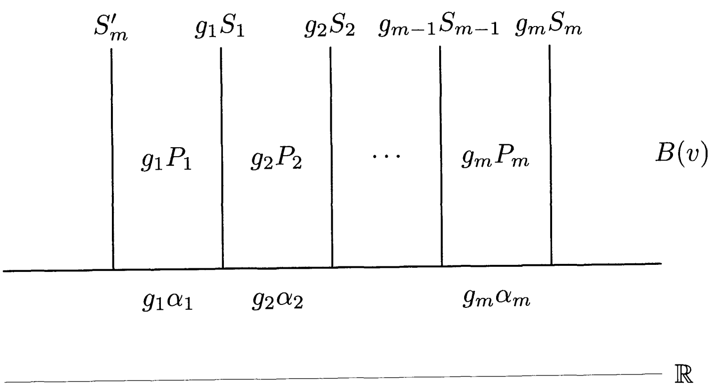


Figure 9.8.3. A fundamental domain for Γ_v in $B(v)$

By Theorem 6.5.8, the inclusion map of $\overline{D}$ into $B(v)$ induces a homeomorphism
$$\kappa : \overline{D}/\Gamma_v \to B(v)/\Gamma_v.$$

Let $\pi : \bigcup_{i=1}^{m} P_i \to M$ be the quotient map. Then we have a map $\psi : \overline{D} \to M$ defined by $\psi(z) = \pi g_i^{-1}(z)$ if z is in $g_i P_i \cap B(v)$. Clearly ψ induces an embedding
$$\phi : \overline{D}/\Gamma_v \to M.$$

Define $\iota : B(v)/\Gamma_v \to M$ by $\iota = \phi\kappa^{-1}$. It is clear from the gluing construction of the hyperbolic structure for M that ι is a local isometry. □

Lemma 1. *Let K and L be two vertical hyperbolic lines of U^2 and let α and β be two horizontal horoarcs joining K to L with β above α at a hyperbolic distance d. Then*
$$|\beta| = |\alpha|e^{-d}.$$

Proof: Let
$$\begin{aligned} K &= \{k + ti : t > 0\}, \\ L &= \{\ell + ti : t > 0\}, \\ \alpha(t) &= t + ai \quad \text{for } k \leq t \leq \ell, \\ \beta(t) &= t + bi \quad \text{for } k \leq t \leq \ell. \end{aligned}$$

Then we have
$$|\alpha| = \int_k^\ell \frac{|\alpha'(t)|}{\mathrm{Im}(\alpha(t))}dt = \int_k^\ell \frac{dt}{a} = \frac{(\ell - k)}{a}.$$
Likewise $|\beta| = (\ell - k)/b$. Now
$$|\alpha|/|\beta| = b/a = \exp(d_U(ai, bi)) = e^d.$$ □

Theorem 9.8.5. *Let M be a hyperbolic 2-manifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, convex, finite-sided polygons in H^2 of finite area by a proper $\mathrm{I}(H^2)$-side-pairing Φ. Then M is complete if and only if $d(v) = 0$ for each ideal vertex v of a polygon in $\mathcal{P}$.*

Proof: Without loss of generality, we may assume that M is connected. We pass to the conformal disk model B^2. Let v be an ideal vertex of a polygon in $\mathcal{P}$ and let $[v] = \{v_1, \ldots, v_m\}$ with $v = v_1$. Choose a sequence of points $x_1, \ldots, x_m$ of B^2 and a sequence of horoarcs $\alpha_1, \ldots, \alpha_m$ as before. Suppose that $d(v) < 0$. Then the images of these arcs in M appear as in Figure 9.8.4. By continuing along horoarcs, as indicated in Figure 9.8.4, we construct an infinite sequence of points $\{x_i\}_{i=1}^\infty$ of B^2 and an infinite sequence of horoarcs $\{\alpha_i\}_{i=1}^\infty$. Let α be the ray in M obtained by spiraling in along the images of the α_i. Then α has finite length, since the length

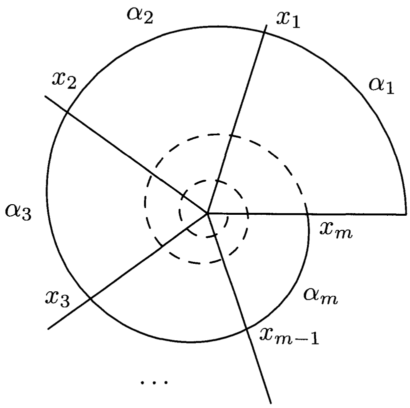


Figure 9.8.4. A sequence of horoarcs spiraling into a puncture of M

of each successive circuit around the puncture of M represented by v is reduced by a constant factor less than one because of Lemma 1. Consequently, the image of the sequence $\{x_i\}$ in M is a Cauchy sequence. As this sequence does not converge, M is incomplete. If $d(v) > 0$, we spiral around the puncture in the opposite direction and deduce that M is incomplete. Thus, if M is complete, then $d(v) = 0$ for each ideal vertex v.

Conversely, suppose that $d(v) = 0$ for each ideal vertex v. Then by Theorems 9.8.3 and 9.8.4, we can remove open horodisk neighborhoods of each ideal vertex to obtain a compact surface-with-boundary M_0 in M. For each $t > 0$, let M_t be the surface-with-boundary obtained by removing smaller horodisk neighborhoods bounded by horocycles at a distance t from the original ones. See Figure 9.8.5. Then M_t is compact for each $t > 0$ and $M = \bigcup_{t>0} M_t$.

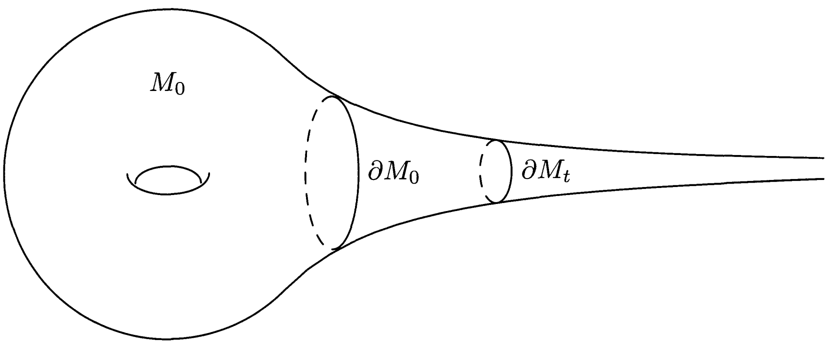


Figure 9.8.5. A complete hyperbolic surface M of finite area

Let x be a point of $M - M_t$. Then there is a $d > 0$ such that x is in ∂M_{t+d}. We claim that d is the distance in M from x to M_t. By the definition of M_{t+d}, we have that d is at most the distance in M from x to M_t. On the contrary, suppose that γ is a curve in M from x to a point y in M_t of length less than d. Then γ must cross ∂M_t, and so we may assume that y is in ∂M_t and the rest of γ lies in $M - M_t$. By Theorem 9.8.4, there is an injective local isometry

$$\iota : B(v)/\Gamma_v \to M$$

whose image is the component of $M - M_0$ containing x. Hence γ corresponds under ι to a curve in $B(v)/\Gamma_v$ of the same length. Let C_t be the horocycle in $B(v)$ at a distance t from $\partial B(v)$. Then $\iota^{-1}\gamma$ lifts to a curve $\tilde{\gamma}$ in $B(v)$ starting in C_{t+d} and ending in C_t. By Lemma 1 of §7.1, we have that $|\tilde{\gamma}| \geq d$, which is a contradiction. Thus d is the distance in M from x to M_t. Consequently M_{t+1} contains $N(M_t, 1)$ for each $t > 0$. Therefore M is complete by Theorem 8.5.10(4). □

Cusps

Let $B(\infty)$ be the open horodisk $\mathbb{R} \times (1, \infty)$ in U^2 and let f_c be the horizontal translation of U^2 by a Euclidean distance $c > 0$ in the positive direction. Let Γ_c be the infinite cyclic group generated by f_c. Then Γ_c acts freely and discontinuously on $B(\infty)$ as a group of isometries. Consequently $B(\infty)/\Gamma_c$ is a hyperbolic surface. The surface $B(\infty)/\Gamma_c$ is homeomorphic to $S^1 \times (1, \infty)$. Each horocycle $\mathbb{R} \times \{t\}$ in $B(\infty)$ projects to a horocircle in $B(\infty)/\Gamma_c$, corresponding to $S^1 \times \{t\}$ in $S^1 \times (1, \infty)$, whose length decreases exponentially with t because of Lemma 1. For this reason, a hyperbolic surface M, isometric to $B(\infty)/\Gamma_c$ for some $c > 0$, is called a *cusp* of circumference c.

The geometry of a cusp is easy to visualize because a cusp of circumference $c \leq \pi$ isometrically embeds in E^3. See Figure 1.1.5. The circumference of a cusp M is unique and an isometric invariant of M because it is the least upper bound of the lengths of the horocircles of M.

The *area* of a cusp M of circumference c is defined to be the area of the fundamental domain

$$D = (0, c) \times (1, \infty)$$

for Γ_c in $B(\infty)$. Hence, we have

$$\text{Area}(M) = \int_D \frac{dxdy}{y^2} = \int_1^\infty \int_0^c \frac{dxdy}{y^2} = c.$$

Thus, the area of a cusp M is equal to its circumference and is therefore finite even though M is unbounded.

We now determine the geometry of a complete hyperbolic surface of finite area.

Theorem 9.8.6. *Let M be a complete hyperbolic surface of finite area. Then there is a compact surface-with-boundary M_0 in M such that $M - M_0$ is the disjoint union of a finite number of cusps.*

Proof: Since M is complete, we may assume that M is a space-form H^2/Γ. Let P be an exact, convex, fundamental polygon for Γ. Then P has finite area and only finitely many sides. By Theorem 6.5.8, the inclusion map of P into H^2 induces a homeomorphism

$$\kappa : P/\Gamma \to H^2/\Gamma,$$

where P/Γ is the space obtained from P by gluing together the sides of P paired by the elements of a subset Φ of Γ. By Theorem 6.7.7, the $\mathrm{I}(H^2)$-side-pairing Φ is proper. Therefore P/Γ has a hyperbolic structure by Theorem 9.2.2. It is clear from the gluing construction of the hyperbolic structure for P/Γ that κ is a local isometry. Moreover, since P/Γ and H^2/Γ are both hyperbolic surfaces, κ is an isometry. Therefore P/Γ is complete.

We now pass to the conformal disk model B^2. Since P/Γ is complete, we can remove open horodisk neighborhoods of each ideal vertex of P to obtain a compact surface-with-boundary M_0 in M. Furthermore $M - M_0$ has a finite number of components, and for each component C of $M - M_0$ there is a ideal vertex v of P and an injective local isometry

$$\iota : B(v)/\Gamma_v \to M,$$

as in Theorem 9.8.4, mapping onto C. By replacing the horodisk neighborhood $B(v)$ of v by a smaller concentric horodisk, if necessary, we can arrange ι to map the cusp $B(v)/\Gamma_v$ isometrically onto C. Thus, we can choose M_0 so that each component of $M - M_0$ is a cusp. □

Discrete Groups

We now consider a general method for constructing a discrete torsion-free subgroup Γ of $\mathrm{I}(H^2)$ whose space-form H^2/Γ has finite area.

Theorem 9.8.7. *Let Φ be a proper $\mathrm{I}(H^2)$-side-pairing for a finite-sided convex polygon P in H^2 of finite area such that the gluing invariants of all the ideal vertices of P are zero. Then the group Γ generated by Φ is discrete and torsion-free, P is an exact, convex, fundamental polygon for Γ, and the inclusion map of P into H^2 induces an isometry from the hyperbolic surface M, obtained by gluing together the sides of P by Φ, to the space-form H^2/Γ.*

Proof: The quotient map $\pi : P \to M$ maps P° homeomorphically onto an open subset U of M. Let $\phi : U \to H^2$ be the inverse of π. From the construction of M, we have that ϕ is locally a chart for M. Therefore ϕ is a chart for M.

Let $\kappa : \tilde{M} \to M$ be a universal covering. As U is simply connected, $\phi : U \to H^2$ lifts to a chart $\tilde{\phi} : \tilde{U} \to H^2$ for $\tilde{M}$. Let $\delta : \tilde{M} \to H^2$ be the developing map determined by $\tilde{\phi}$. The hyperbolic surface M is complete by Theorem 9.8.5. Therefore δ is an isometry by Theorem 8.5.9. Let $\zeta = \kappa\delta^{-1}$. Then $\zeta : H^2 \to M$ is a covering projection extending π on P°. Moreover, by continuity, ζ extends π.

Let Γ be the group of covering transformations of ζ. By Theorem 8.5.9, we have that Γ is a torsion-free discrete group of isometries of H^2, and ζ induces an isometry from H^2/Γ to M. Now as U is simply connected, it is evenly covered by ζ. Hence, the members of $\{gP^\circ : g \in \Gamma\}$ are mutually disjoint. As $\pi(P) = M$, we have

$$X = \cup\{gP : g \in \Gamma\}.$$

Therefore P° is a fundamental domain for Γ.

Let g_S be an element of Φ. Choose a point y in the interior of the side S of P. Then there is an element y' in the interior of the side S' of P such that $g_S(y') = y$. Since $\pi(y') = y$, there is an element g of Γ such that $g(y') = y$. Since gS' does not extend into P°, we must have that gS' lies on the hyperbolic line extending S. Moreover, since pairs of points of S° equidistant from y are not identified by π, we have that g and g_S agree on S'. Furthermore, since gP lies on the opposite side of S from P, we deduce that $g = g_S$ by Theorem 4.3.6. Thus Γ contains Φ. Therefore P/Γ is a quotient of M.

Now by Theorem 6.5.8, the inclusion map of P into H^2 induces a continuous bijection from P/Γ to H^2/Γ. The composition of the induced maps

$$H^2/\Gamma \to M \to P/\Gamma \to H^2/\Gamma$$

restricts to the identity map of P° and so is the identity map by continuity. Therefore $M = P/\Gamma$.

Now since $\zeta : H^2 \to M$ induces an isometry from H^2/Γ to $M = P/\Gamma$, the inclusion map of P into H^2 induces an isometry from P/Γ to H^2/Γ. Therefore P is locally finite by Theorem 6.5.8. Hence P is an exact, convex, fundamental polygon for Γ. Finally Φ generates Γ by Theorem 6.7.3. □

Example 2. Let P be the ideal square in U^2 with vertices $-1, 0, 1, \infty$. See Figure 9.8.6. Pair the vertical sides of P by a horizontal translation and the sides incident with 0 by reflecting in the y-axis and then reflecting in the corresponding side of P. This $\mathrm{I}_0(U^2)$-side-pairing Φ is proper. The hyperbolic surface M obtained by gluing together the sides of P by Φ is a thrice-punctured sphere.

The complete hyperbolic structure of finite area on the thrice-punctured sphere is very special because the thrice-punctured sphere is the only surface that has a complete hyperbolic structure of finite area that is unique up to isometry.

Theorem 9.8.8. *The complete hyperbolic structure of finite area on the thrice-punctured sphere is unique up to isometry.*

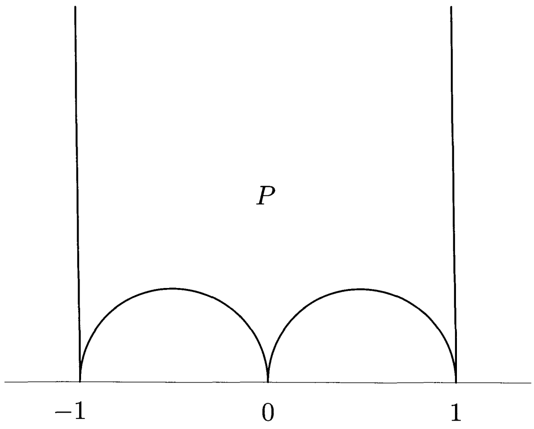


Figure 9.8.6. The ideal square P in U^2 with vertices $-1, 0, 1, \infty$

Proof: Let M be a thrice-punctured sphere with a complete hyperbolic structure of finite area. Then M is isometric to a space-form U^2/Γ of finite area. By Theorem 9.8.6, there is a compact surface-with-boundary M_0 in M such that $M - M_0$ is the disjoint union of three cusps. Therefore M_0 is a pair of pants. Consider the curves α, β, γ in M_0 shown in Figure 9.8.7. Observe that the simple closed curves $\alpha\beta^{-1}$, $\beta\gamma^{-1}$, and $\alpha\gamma^{-1}$ are freely homotopic to the boundary horocircles of M_0. Therefore, the elements of $\pi_1(M)$, represented by these curves, correspond to parabolic elements f, g, h of Γ_g. As $[\alpha\beta^{-1}]$ and $[\beta\gamma^{-1}]$ generate the free group $\pi_1(M)$ of rank two, f and g generate the free group Γ of rank two. Moreover, since

$$[\alpha\gamma^{-1}] = [\alpha\beta^{-1}][\beta\gamma^{-1}],$$

we have that $h = fg$.

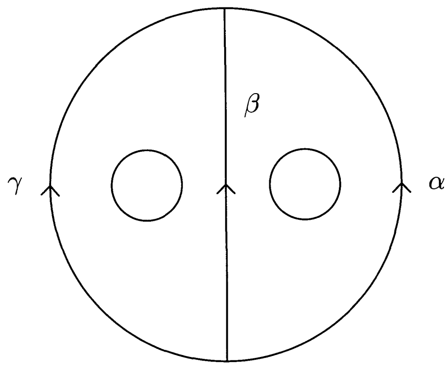


Figure 9.8.7. The pair of pants M_0 in a thrice-punctured sphere M

By conjugating Γ in $\mathrm{I}_0(U^2)$, we may assume that $f(z) = z + 2$. As g is parabolic, there are real numbers a, b, c, d, such that

$$g(z) = \frac{az+b}{cz+d}$$

with $ad - bc = 1$ and $a + d = 2$. Next, observe that

$$h(z) = \frac{(a+2c)z + (b+2d)}{cz+d}.$$

As h is parabolic, we have

$$a + 2c + d = \pm 2.$$

Hence $c = 0$ or -2. If $c = 0$, then f and h would commute, which is not the case, since Γ is a free group of rank two generated by f and h. Therefore $c = -2$.

The elements f and g do not commute, since they generate Γ. Therefore, the fixed point of g is on the real axis. By conjugating Γ by a horizontal translation of U^2, we may assume that the fixed point of g is 0. Then $b = 0$. Now since $ad - bc = 1$, we have that $ad = 1$. As $a + d = 2$, we deduce that $a = 1 = d$. Hence, we have

$$g(z) = \frac{z}{-2z+1}.$$

Now $g(1) = -1$. Hence g pairs the sides of the ideal square in Figure 9.8.6 incident with 0. Therefore Γ is the discrete group in Example 2. Thus, the complete hyperbolic structure of finite area on M is unique up to isometry. □

Exercise 9.8

1. Construct a hyperbolic convex polygon of finite area with infinitely many sides.
2. Let M be a surface obtained from a closed surface by removing a finite number of points. Prove that M has a complete hyperbolic structure of finite area if and only if $\chi(M) < 0$.
3. Let C be a cycle of m ideal vertices. Prove that C has $2m$ cycle transformations associated to its vertices and that all these transformations are conjugates of each other or their inverses. Conclude that if one of these transformations in parabolic, then they are all parabolic.
4. Prove that the horodisk $B(v)$ in Theorem 9.8.4 can be replaced by a smaller concentric horodisk so that ι maps the cusp $B(v)/\Gamma_v$ isometrically onto its image in M.
5. Prove that the group in Example 2 is the group of all linear fractional transformations $\gamma(z) = (az+b)/(cz+d)$ with a, b, c, d integers such that
$$ad - bc = 1 \quad \text{and} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod 2.$$
6. Prove that the once-punctured torus has an uncountable number of nonisometric complete hyperbolic structures of finite area.

§9.9. Historical Notes

§9.1. The Euler characteristic of the boundary of a convex polyhedron was essentially introduced by Euler in his 1758 paper *Elementa doctrinae solidorum* [121]. Euler proved that the Euler characteristic of the boundary of a convex polyhedron is two in his 1758 paper *Demonstratio nonnullarum insignium proprietatum quibus solida hedris planis inclusa sunt praedita* [122]. The Euler characteristic of a closed, orientable, polygonal surface was introduced by Lhuilier in his 1813 paper *Mémoire sur la polyédrométrie* [257]. In particular, Formula 9.1.4 appeared in this paper. A surface with a complex structure is called a *Riemann surface.* Closed Riemann surfaces were introduced and classified by Riemann in his 1857 paper *Theorie der Abel'schen Functionen* [348]. Closed orientable surfaces were classified by Möbius in his 1863 paper *Theorie der elementaren Verwandtschaft* [299]. The notion of orientability of a surface was introduced by Möbius in his 1865 paper *Ueber die Bestimmung des Inhaltes eines Polyëders* [300]. See also his paper *Zur Theorie der Polyëder und der Elementarverwandtschaft* [301], which was published posthumously in 1886. Formula 9.1.6 appeared in Jordan's 1866 paper *Recherches sur les polyèdres* [204]. Compact orientable surfaces-with-boundary were classified by Jordan in his 1866 paper *La déformation des surfaces* [203]. That the projective plane is nonorientable appeared in Klein's 1874 paper *Bemerkungen über den Zusammenhang der Flächen* [228]. See also Klein's 1876 paper *Ueber den Zusammenhang der Flächen* [230]. The Klein bottle was introduced by Klein in his 1882 treatise *Ueber Riemanns Theorie der algebraischen Functionen und ihrer Integrale* [232]. Theorems 9.1.2 and 9.1.4 appeared in Dyck's 1888 paper *Beiträge zur Analysis situs* [113]. For the early history of topology of surfaces, see Pont's 1974 treatise *La Topologie Algébrique des origines à Poincaré* [342] and Scholz's 1980 treatise *Geschichte des Mannigfaltigkeitsbegriffs von Riemann bis Poincaré* [363]. References for the topology of surfaces are Massey's 1967 text *Algebraic Topology: An Introduction* [283] and Moise's 1977 text *Geometric Topology in Dimensions 2 and 3* [302].

§9.2. In 1873, Clifford described a Euclidean torus embedded in elliptic 3-space in his paper *Preliminary sketch of biquaternions* [82]. In particular, he wrote, "The geometry of this surface is the same as that of a finite parallelogram whose opposite sides are regarded as identical." Closed hyperbolic surfaces were constructed by Poincaré in his 1882 paper *Théorie des groupes fuchsiens* [330] by gluing together the sides of hyperbolic convex polygons by proper side-pairings. As a reference for geometric surfaces, see Weeks' 1985 text *The Shape of Space* [406].

§9.3. The Gauss-Bonnet theorem for closed, orientable, Riemannian surfaces appeared in Dyck's 1888 paper [113] and was extended to nonorientable surfaces by Boy in his 1903 paper *Über die Curvatura integra und die Topologie geschlossener Flächen* [56]. Theorems 9.3.1 and 9.3.2 appeared in Weeks' 1985 text [406].

§9.4. The moduli space of a closed orientable surface M was introduced by Riemann in his 1857 paper [348] as the space of all conformal equivalence classes of Riemann surface structures on M. In particular, Riemann asserted that the moduli space of a closed orientable surface M of genus $n > 1$ can be parameterized by $3n - 3$ complex parameters that he called *moduli*. For a discussion, see Chap. V of Dieudonné's 1985 treatise *History of Algebraic Geometry* [106]. Klein asserted that every closed Riemann surface is conformally equivalent to either a spherical, Euclidean, or hyperbolic plane-form, that is unique up to orientation preserving similarity, in his 1883 paper *Neue Beiträge zur Riemann'schen Functionentheorie* [233]. Klein's assertion is called the *uniformization theorem*. The uniformization theorem was proved independently by Poincaré in his 1907 paper *Sur l'uniformisation des fonctions analytiques* [339] and by Koebe in his 1907 paper *Über die Uniformisierung beliebiger analytischen Kurven* [239]. For a discussion, see Abikoff's 1981 article *The uniformization theorem* [2]. It follows from the uniformization theorem that Riemann's moduli space of a closed orientable surface M of positive genus is equivalent to the moduli space of orientation preserving similarity classes of Euclidean or hyperbolic structures for M.

The Teichmüller space of a closed orientable surface appeared implicitly in Klein's 1883 paper [233] and in Poincaré's 1884 paper *Sur les groupes des équations linéaires* [333]. For a discussion, see §6.4 of Gray's 1986 treatise *Linear Differential Equations and Group Theory from Riemann to Poincaré* [160]. Teichmüller space was explicitly introduced by Teichmüller in his 1939 paper *Extremale quasikonforme Abbildungen und quadratische Differentiale* [387]. Theorem 9.4.3 for orientable surfaces was proved by Dehn and Nielsen and appeared in Nielsen's 1927 paper *Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen* [319]. Theorem 9.4.3 for nonorientable surfaces was proved by Mangler in his 1938 paper *Die Klassen von topologischen Abbildungen einer geschlossenen Flächen auf sich* [274]. The space of discrete faithful representations of a group appeared in Weil's 1960 paper *On discrete subgroups of Lie groups* [408].

§9.5. That the moduli space of the torus has complex dimension one appeared in Riemann's 1857 paper [348]. Theorems 9.5.1 and 9.5.2 appeared in Poincaré's 1884 paper [333].

§9.6. All the essential material in this section appeared in Poincaré's 1904 paper *Cinquième complément à l'analysis situs* [337].

§9.7. A closed orientable hyperbolic surface was implicitly decomposed into pairs of pants by Fricke and Klein in their 1897-1912 treatise *Vorlesungen über die Theorie der automorphen Functionen* [139]. Moreover, they implicitly showed that a pair of pants is the union of two congruent right-angled hyperbolic hexagons sewn together along seams. Instead of working with right-angled hexagons, they worked projectively with ultra-ideal triangles. An ultra-ideal triangle corresponds to a right-angled hexagon in the same way that the triangle $T(x, y, z)$ corresponds to the right-angled hexagon in Figure 3.5.10. Fricke and Klein also essentially proved that the Teichmüller space of a closed orientable surface of genus $n > 1$ is homeomorphic to $(6n - 6)$-dimensional Euclidean space. They expressed their coordinates in terms of the traces of the matrices in $\mathrm{SL}(2, \mathbb{R})$ that represent the transformations corresponding to the decomposition geodesics and certain other simple closed geodesics on a closed hyperbolic surface. Each trace determines the length of the corresponding simple closed geodesic. The twist coefficients of the decomposition geodesics were not clearly identified by Fricke and Klein. For discussions, see Keen's 1971-1973 paper *On Fricke moduli* [214], [215], Harvey's 1977 article *Spaces of discrete groups* [179], and Bers and Gardiner's 1986 paper *Fricke Spaces* [43].

An explicit decomposition of a closed, orientable, hyperbolic surface into right-angled hyperbolic hexagons was given by Löbell in his 1927 thesis *Die überall regulären unbegrenzten Flächen fester Krümmung* [265]. In particular, Löbell described the length coordinates and twist coordinates (modulo 2π) of a closed, orientable, hyperbolic surface. Löbell's decomposition and coordinates were described by Koebe in his 1928 paper *Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen. III* [243]. This decomposition was further studied by Nielsen and Fenchel in their 1959 manuscript *Discontinuous Groups of Non-Euclidean Motions* [321]. In particular, they implicitly unwound the twist coordinates. For a discussion, see Wolpert's 1982 paper *The Fenchel-Nielsen deformation* [417]. The length-twist coordinates of a closed, orientable, hyperbolic surface were explicitly described by Thurston in his 1979 lecture notes *The Geometry and Topology of 3-Manifolds* [389], by Douady in his 1979 exposé *L'espace de Teichmüller* [108], and by Abikoff in his 1980 treatise *The Real Analytic Theory of Teichmüller Space* [1]. For a characterization of a pair of pants in a hyperbolic surface, see Basmajian's 1990 paper *Constructing pairs of pants* [31].

§9.8. Theorem 9.8.1 was proved by Siegel in his 1945 paper *Some remarks on discontinuous groups* [376]. Theorem 9.8.2 was proved by Koebe in his 1928 paper [243]. The complete gluing of an open surface of finite area was considered by Poincaré in his 1884 paper [333]. For commentary, see Klein's 1891 paper *Ueber den Begriff des functionentheoretischen Fundamentalbereichs* [235]. Theorem 9.8.4 was essentially proved by Seifert in his 1975 paper *Komplexe mit Seitenzuordnung* [371]. Theorem 9.8.5 for a single polygon was proved by de Rham in his 1971 paper *Sur les polygones générateurs de groupes fuchsiens* [103] and by Maskit in his 1971 paper *On Poincaré's theorem for fundamental polygons* [281]. Theorem 9.8.5 was proved by Seifert in his 1975 paper [371]. Theorem 9.8.6 essentially appeared in Koebe's 1927 Preisschrift *Allgemeine Theorie der Riemannschen Mannigfaltigkeiten* [240]. See also his 1928 paper [243]. Theorem 9.8.7 appeared in de Rham's 1971 paper [103] and in Maskit's 1971 paper [281]. Theorem 9.8.8 is a consequence of the classification of all the complete hyperbolic structures on a thrice-punctured sphere given by Fricke and Klein in their 1897-1912 treatise [139].

CHAPTER 10

Hyperbolic 3-Manifolds

In this chapter, we construct some examples of hyperbolic 3-manifolds. We begin with a geometric method for constructing spherical, Euclidean, and hyperbolic 3-manifolds in Sections 10.1 and 10.2. Examples of complete hyperbolic 3-manifolds of finite volume are constructed in Section 10.3. The problem of computing the volume of a hyperbolic 3-manifold is taken up in Section 10.4. The chapter ends with a detailed study of hyperbolic Dehn surgery on the figure-eight knot complement.

§10.1. Gluing 3-Manifolds

In this section, we shall construct spherical, Euclidean, and hyperbolic 3-manifolds by gluing together convex polyhedra in $X = S^3, E^3$, or H^3 along their sides.

Let $\mathcal{P}$ be a finite family of disjoint convex polyhedra in X and let G be a group of isometries of X.

Definition: A *G-side-pairing* for $\mathcal{P}$ is a subset of G,

$$\Phi = \{g_S : S \in \mathcal{S}\},$$

indexed by the collection $\mathcal{S}$ of all the sides of the polyhedra in $\mathcal{P}$ such that for each side S in $\mathcal{S}$,

(1) there is a side S' in $\mathcal{S}$ such that $g_S(S') = S$;

(2) the isometries g_S and $g_{S'}$ satisfy the relation $g_{S'} = g_S^{-1}$; and

(3) if S is a side of P in $\mathcal{P}$ and S' is a side of P' in $\mathcal{P}$, then

$$P \cap g_S(P') = S.$$

It follows from (1) that S' is uniquely determined by S. The side S' is said to be *paired to* the side S by Φ. From (2), we deduce that $S'' = S$. The pairing of side points by elements of Φ generates an equivalence relation on the set $\Pi = \cup_{P\in\mathcal{P}} P$, and the equivalence classes are called the *cycles* of Φ.

The *solid angle* ω subtended by a polyhedron P in X at a point x of P is defined to be the real number

$$\omega = 4\pi\frac{\mathrm{Vol}(P\cap B(x,r))}{\mathrm{Vol}(B(x,r))},$$

where r is less than the distance from x to any side of P not containing x. It follows from Theorems 2.4.1 and 3.4.1 that ω does not depend on r.

Let $[x] = \{x_1,\dots,x_m\}$ be a finite cycle of Φ. Let P_i be the polyhedron in $\mathcal{P}$ containing the point x_i and let ω_i be the solid angle subtended by P_i at the point x_i for each $i = 1,\dots,m$. The *solid angle sum* of $[x]$ is defined to be the real number

$$\omega[x] = \omega_1 + \cdots + \omega_m.$$

If x is in the interior of a polyhedron in $\mathcal{P}$, then $[x] = \{x\}$ and $\omega[x] = 4\pi$. If x is in the interior of a side S of a polyhedron in $\mathcal{P}$, then $x' = g_S^{-1}(x)$ is in the interior of S' and $[x] = \{x, x'\}$; therefore $\omega[x] = 2\pi$ or 4π according as $x = x'$ or $x \neq x'$.

Now suppose that x is in the interior of an edge of a polyhedron in $\mathcal{P}$. Then every point of $[x]$ is in the interior of an edge of a polyhedron in $\mathcal{P}$, in which case $[x]$ is called an *edge cycle* of Φ. Let θ_i be the dihedral angle of P_i along the edge containing x_i for each i. The *dihedral angle sum* of the edge cycle $[x]$ is defined to be the real number

$$\theta[x] = \theta_1 + \cdots + \theta_m.$$

Note that $\omega_i = 2\theta_i$ for each i. Therefore $\omega[x] = 2\theta[x]$.

Definition: A G-side-pairing Φ for $\mathcal{P}$ is *proper* if and only if each cycle of Φ is finite and has solid angle sum 4π.

Theorem 10.1.1. *If G is a group of isometries of X and Φ is a proper G-side-pairing for a finite family $\mathcal{P}$ of disjoint convex polyhedra in X, then*

(1) *the isometry g_S fixes no point of S' for each S in $\mathcal{S}$;*

(2) *the sides S and S' are equal if and only if S is a great 2-sphere of S^3 and g_S is the antipodal map of S^3; and*

(3) *each edge cycle of Φ contains at most one point of an edge of a polyhedron in $\mathcal{P}$.*

Proof: (1) On the contrary, suppose that g_S fixes a point x of S'. Let $[x] = \{x_1,\dots,x_m\}$. Then $m \geq 2$, since Φ is proper. Let P_i be the polyhedron in $\mathcal{P}$ containing x_i for each i. Let r be a positive real number such that r is less than half the distance from x_i to x_j for each $i \neq j$ and from x_i to any side of P_i not containing x_i for each i. Then $P_i \cap S(x_i, r)$ is a polygon in the sphere $S(x_i, r)$ and the polygons $\{P_i \cap S(x_i, r)\}$ are disjoint. Now the side-pairing

Φ restricts to a proper $\mathrm{I}(S^2)$-side-pairing of the polygons $\{P_i \cap S(x_i, r)\}$. Let Σ be the space obtained by gluing together the polygons. Then Σ has a spherical structure by Theorem 9.2.3; moreover Σ is a 2-sphere, since Σ is compact, connected, and $\omega[x] = 4\pi$.

Let P be the polyhedron in $\mathcal{P}$ containing x. Then the side $S' \cap S(x,r)$ of $P \cap S(x,r)$ is paired to the side $S \cap S(x,r)$ of $P \cap S(x,r)$. Let y be a point of $S \cap S(x,r)$ and let $y' = g_S^{-1}(y)$. Then $y \neq y'$ by Theorem 9.2.1(1). As $P \cap S(x,r)$ is a convex polygon, there is a geodesic segment $[y, y']$ in $P \cap S(x,r)$ joining y to y'. As y is paired to y', the segment projects to a great circle of the sphere Σ, but this is a contradiction because the length of $[y, y']$ is at most half the length of a great circle of $S(x,r)$. Thus g_S fixes no point of S'.

(2) The proof of (2) is the same as the proof of Theorem 9.2.1(2).

(3) Suppose that $[x]$ is an edge cycle. Then the cycle $[x]$ can be ordered

$$[x] = \{x_1, x_2, \ldots, x_m\}$$

so that

$$x = x_1 \simeq x_2 \simeq \cdots \simeq x_m \simeq x.$$

Let E_i be the edge of the polyhedron in $\mathcal{P}$ containing x_i, and let k be the number of points of $[x]$ contained in E_1. Then E_i contains k points of $[x]$ for each i. Let y_i be the centroid of the points of $[x]$ in E_i for each i, and let $y = y_1$. Then we have

$$y = y_1 \simeq y_2 \simeq \cdots \simeq y_m \simeq y.$$

Moreover

$$d(x_1, y_1) = d(x_2, y_2) = \cdots = d(x_m, y_m).$$

Therefore $k = 1$ or 2. Now as

$$4\pi = \omega[x] = 2\theta[x] = 2k\theta[y] = k\omega[y] = 4k\pi,$$

we must have $k = 1$. □

Let Φ be a proper G-side-pairing for $\mathcal{P}$ and let M be the quotient space of Π of cycles of Φ. The space M is said to be obtained by gluing together the polyhedra in $\mathcal{P}$ by Φ.

Theorem 10.1.2. *Let G be a group of isometries of X and let M be a space obtained by gluing together a finite family $\mathcal{P}$ of disjoint convex polyhedra in X by a proper G-side-pairing Φ. Then M is a 3-manifold with an (X, G)-structure such that the natural injection of P° into M is an (X, G)-map for each P in $\mathcal{P}$.*

Proof: Without loss of generality, we may assume that each polyhedron in $\mathcal{P}$ has at least one side. Let x a point of Π and let $[x] = \{x_1, \ldots, x_m\}$. Let P_i be the polyhedron in $\mathcal{P}$ containing x_i for each i. If x_i in a side of P_i, then $m \geq 2$ by Theorem 10.1.1. Let $\delta(x)$ be the minimum distance from

x_i to x_j for each $i \neq j$ and from x_i to any side of P_i not containing x_i for each i.

Let r be a real number such that $0 < r < \delta(x)/2$. Then for each i, the set $P_i \cap S(x_i, r)$ is a polygon in the sphere $S(x_i, r)$, and the polygons $\{P_i \cap S(x_i, r)\}$ are disjoint. Now the side-pairing Φ restricts to a proper $\mathrm{I}(S^2)$-side-pairing of the polygons $\{P_i \cap S(x_i, r)\}$. Let $\Sigma(x, r)$ be the space obtained by gluing together the polygons. Then $\Sigma(x, r)$ has a spherical structure by Theorem 9.2.3. Now since $\Sigma(x, r)$ is compact, connected, and $\omega[x] = 4\pi$, we deduce that $\Sigma(x, r)$ is a 2-sphere.

Let $\pi : \Pi \to M$ be the quotient map. Then for each i, the restriction of π to the polygon $P_i \cap S(x_i, r)$ extends to an isometry

$$\xi_i : S(x_i, r) \to \Sigma(x, r).$$

Moreover, for each i, j, the isometry

$$\xi_j^{-1}\xi_i : S(x_i, r) \to S(x_j, r)$$

extends to a unique isometry g_{ij} of X, and $g_{ij}(x_i) = x_j$.

Suppose that the element g_S of Φ pairs the side $S' \cap S(x_i, r)$ of the polygon $P_i \cap S(x_i, r)$ to the side $S \cap S(x_j, r)$ of $P_j \cap S(x_j, r)$. Then $\xi_j^{-1}\xi_i$ agrees with g_S on the set $S' \cap S(x_i, r)$. Hence $\xi_j^{-1}\xi_i$ agrees with g_S on the great circle $\langle S' \rangle \cap S(x_i, r)$. Therefore g_{ij} agrees with g_S on the plane $\langle S' \rangle$. Now since g_{ij} and g_S both map $P_i \cap S(x_i, r)$ to the opposite side of the plane $\langle S \rangle$ from $P_j \cap S(x_j, r)$, we deduce that $g_{ij} = g_S$ by Theorem 4.3.6.

Now suppose that

$$x_i = x_{i_1} \simeq x_{i_2} \simeq \cdots \simeq x_{i_p} = x_j.$$

Then we have

$$\xi_j^{-1}\xi_i = (\xi_{i_p}^{-1}\xi_{i_{p-1}})(\xi_{i_{p-1}}^{-1}\xi_{i_{p-2}}) \cdots (\xi_{i_2}^{-1}\xi_{i_1}).$$

Hence, we have

$$g_{ij} = g_{i_{p-1}i_p} g_{i_{p-2}i_{p-1}} \cdots g_{i_1 i_2}.$$

Now the elements $g_{i_1 i_2}, \ldots, g_{i_{p-1}i_p}$ are in Φ by the previous argument. Therefore g_{ij} is in G for each i, j.

Define

$$U(x, r) = \bigcup_{i=1}^{m} \pi(P_i \cap B(x_i, r)).$$

As the set

$$\pi^{-1}(U(x, r)) = \bigcup_{i=1}^{m} P_i \cap B(x_i, r)$$

is open in Π, we have that $U(x, r)$ is an open subset of M.

Suppose that $x = x_k$ and define a function

$$\psi_x : \bigcup_{i=1}^{m} P_i \cap B(x_i, r) \to B(x, r)$$

by the rule

$$\psi_x(z) = g_{ik}(z) \ \text{ if } z \text{ is in } P_i \cap B(x_i, r).$$

Suppose that $g_S(x_i) = x_j$. Then $g_S = g_{ij}$. Let y be a point of $S \cap B(x_j, r)$ and let $y' = g_S^{-1}(y)$. Then y' is a point of $S' \cap B(x_i, r)$. As

$$\xi_k^{-1}\xi_i = (\xi_k^{-1}\xi_j)(\xi_j^{-1}\xi_i),$$

we have that $g_{ik} = g_{jk}g_{ij}$. Therefore

$$\psi_x(y) = g_{jk}(y) = g_{jk}g_S(y') = g_{ik}(y') = \psi_x(y').$$

Consequently ψ_x induces a continuous function

$$\phi_x : U(x, r) \to B(x, r).$$

For each t such that $0 < t < r$, the function ϕ_x restricts to the isometry

$$\xi_k^{-1} : \Sigma(x, t) \to S(x, t)$$

corresponding to t. Therefore ϕ_x is a bijection with a continuous inverse defined by

$$\phi_x^{-1}(z) = \pi g_{ik}^{-1}(z) \quad \text{if } z \text{ is in } g_{ik}(P_i \cap B(x_i, r)).$$

Hence ϕ_x is a homeomorphism. The same argument as in the proof of Theorem 9.2.2 shows that M is Hausdorff. Thus M is a 3-manifold.

Next, we show that

$$\big\{\phi_x : U(x, r) \to B(x, r) \;\big|\; x \text{ is in } \Pi \;\text{ and }\; r < \delta(x)/3\big\}$$

is an (X, G)-atlas for M. By construction, $U(x, r)$ is an open connected subset of M and ϕ_x is a homeomorphism. Moreover $U(x, r)$ is defined for each point $\pi(x)$ of M and sufficiently small radius r. Consequently $\{U(x, r)\}$ is an open cover of M.

Suppose that $U(x, r)$ and $U(y, s)$ overlap and $r < \delta(x)/3$ and $s < \delta(y)/3$. Let $F(x)$ be the face of the polyhedron in $\mathcal{P}$ that contains x in its interior. By reversing the roles of x and y, if necessary, we may assume that

$$\dim F(x) \geq \dim F(y).$$

As before, we have

$$\begin{aligned} \pi^{-1}(U(x, r)) &= \bigcup_{i=1}^{m} P_i \cap B(x_i, r), \\ \pi^{-1}(U(y, s)) &= \bigcup_{j=1}^{n} Q_j \cap B(y_j, s). \end{aligned}$$

Now for some i and j, the set $P_i \cap B(x_i, r)$ meets $Q_j \cap B(y_j, s)$. By reindexing, we may assume that $P_1 \cap B(x_1, r)$ meets $Q_1 \cap B(y_1, s)$. Then $P_1 = Q_1$ and $d(x_1, y_1) < r + s$. We claim that y_1 is in every side of P_1 that contains x_1. On the contrary, suppose that y_1 is not in a side of P_1 that contains x_1. Then $s < d(x_1, y_1)/3$. Therefore x_1 is in every side of P_1 that contains y_1, otherwise we would have the contradiction that $r < d(x_1, y_1)/3$. Hence $F(x_1)$ is a proper face of $F(y_1)$, which is a contradiction. Therefore y_1 is in every side of P_1 that contains x_1. This implies that for each i, the set $P_i \cap B(x_i, r)$ meets $Q_j \cap B(y_j, s)$ for some j.

We claim that the set $P_i \cap B(x_i, r)$ meets $Q_j \cap B(y_j, s)$ for only one index j. On the contrary, suppose that $P_i \cap B(x_i, r)$ meets $Q_j \cap B(y_j, s)$ and $Q_k \cap B(y_k, s)$. Then $P_i = Q_j = Q_k$. Now since y_j and y_k are in every side of P_i that contains x_i, we have that $F(y_j)$ and $F(y_k)$ are faces of $F(x_i)$. Moreover, $F(y_j)$ and $F(y_k)$ are distinct by Theorem 10.1.1. Therefore $F(y_j)$ and $F(y_k)$ are proper faces of $F(x_i)$. Consequently, we have

$$r < d(x_i, y_j)/3, \quad r < d(x_i, y_k)/3, \quad \text{and} \quad s < d(y_j, y_k)/3.$$

Now observe that

$$\begin{aligned} d(x_i, y_j) + d(x_i, y_k) &< (r + s) + (r + s) \\ &< d(x_i, y_j)/3 + d(x_i, y_k)/3 + 2d(y_j, y_k)/3 \\ &\leq d(x_i, y_j) + d(x_i, y_k), \end{aligned}$$

which is a contradiction. Therefore $P_i \cap B(x_i, r)$ meets $Q_j \cap B(y_j, s)$ for only one index j.

We claim that the set $Q_j \cap B(y_j, s)$ meets $P_i \cap B(x_i, r)$ for only one index i. On the contrary, suppose that $Q_j \cap B(y_j, s)$ meets $P_i \cap B(x_i, r)$ and $P_k \cap B(x_k, r)$. Then $P_i = Q_j = P_k$. Now since y_j is in every side of P_i that contains x_i or x_k, we have that $F(y_j)$ is a face of $F(x_i)$ and $F(x_k)$. Moreover $F(x_i)$ and $F(x_k)$ are distinct by Theorem 10.1.1. Therefore $F(y_j)$ is a proper face of $F(x_i)$ and $F(x_k)$. Consequently, we have

$$r < d(x_i, y_j)/3 < (r + s)/3.$$

Therefore $r < s/2$. As $s < \delta(y)/3$, we have that $r < \delta(y)/6$. Now observe that

$$d(x_i, y_j) < r + s < \delta(y)/2 \quad \text{and} \quad d(x_k, y_j) < r + s < \delta(y)/2.$$

From the construction of $U(y, r+s)$, we deduce that π maps $P_i \cap B(y_j, r+s)$ injectively into M. As x_i and x_k are in $P_i \cap B(y_j, r + s)$, we have a contradiction. Consequently, we can reindex $[y]$ so that $P_i \cap B(x_i, r)$ meets only $Q_i \cap B(y_i, s)$ for $i = 1, \ldots, m$. Then $P_i = Q_i$ for each i.

Let g_{ij} and h_{ij} be the elements of G constructed as before for x and y. Suppose that g_S pairs the side $S' \cap S(x_i, r)$ of $P_i \cap S(x_i, r)$ to the side $S \cap S(x_j, r)$ of $P_j \cap S(x_j, r)$. Then $g_S = g_{ij}$ and $g_S(x_i) = x_j$. Therefore x_i is in S'. Now since $P_i \cap B(x_i, r)$ meets $P_i \cap B(y_i, s)$, we have that y_i is also in S'. Now observe that $g_S(P_i \cap B(x_i, r))$ meets $g_S(P_i \cap B(y_i, s))$. Hence $P_j \cap B(x_j, r)$ meets $P_j \cap B(g_S y_i, s)$. Therefore $g_S y_i = y_j$. Hence $g_{ij} = h_{ij}$.

Now suppose that

$$x_i = x_{i_1} \simeq x_{i_2} \simeq \cdots \simeq x_{i_p} = x_j.$$

Then we deduce from the previous argument that

$$y_i = y_{i_1} \simeq y_{i_2} \simeq \cdots \simeq y_{i_p} = y_j$$

and

$$\begin{aligned} g_{ij} &= g_{i_{p-1}i_p} g_{i_{p-2}i_{p-1}} \cdots g_{i_1 i_2} \\ &= h_{i_{p-1}i_p} h_{i_{p-2}i_{p-1}} \cdots h_{i_1 i_2} \quad = \quad h_{ij}. \end{aligned}$$

Next, observe that

$$\begin{aligned}
&U(x,r)\cap U(y,s)\\
&= \pi\Big(\bigcup_{i=1}^{m} P_i\cap B(x_i,r)\Big)\cap\pi\Big(\bigcup_{j=1}^{n} Q_j\cap B(y_j,s)\Big)\\
&= \pi\Big(\Big[\bigcup_{i=1}^{m} P_i\cap B(x_i,r)\Big]\cap\Big[\bigcup_{j=1}^{n} Q_j\cap B(y_j,s)\Big]\Big)\\
&= \pi\Big(\bigcup_{i=1}^{m}\bigcup_{j=1}^{n}\big[P_i\cap B(x_i,r)\cap Q_j\cap B(y_j,s)\big]\Big)\\
&= \pi\Big(\bigcup_{i=1}^{m} P_i\cap B(x_i,r)\cap B(y_i,s)\Big).
\end{aligned}$$

Let $x = x_k$ and $y = y_\ell$. Then

$$\phi_x\big(U(x,r)\cap U(y,s)\big) = \bigcup_{i=1}^{m} g_{ik}\big(P_i\cap B(x_i,r)\cap B(y_i,s)\big)$$

and

$$\phi_y\big(U(x,r)\cap U(y,s)\big) = \bigcup_{i=1}^{m} h_{i\ell}\big(P_i\cap B(x_i,r)\cap B(y_i,s)\big).$$

Now on the set

$$g_{ik}\big(P_i\cap B(x_i,r)\cap B(y_i,s)\big),$$

the map $\phi_y\phi_x^{-1}$ is the restriction of

$$h_{i\ell}g_{ik}^{-1} = h_{i\ell}h_{ik}^{-1} = h_{i\ell}h_{ki} = h_{k\ell}$$

for each $i = 1,\ldots,m$. Therefore $\phi_y\phi_x^{-1}$ is the restriction of $h_{k\ell}$. Thus $\phi_y\phi_x^{-1}$ agrees with an element of G. This completes the proof that $\{\phi_x\}$ is an (X,G)-atlas for M.

The same argument as in the proof of Theorem 9.2.2 shows that the (X,G)-structure of M has the property that the natural injection map of P° into M is an (X,G)-map for each P in $\mathcal{P}$. □

The next theorem makes it much easier to apply Theorem 10.1.2.

Theorem 10.1.3. *Let G be a group of orientation preserving isometries of X and let $\Phi = \{g_S : S\in\mathcal{S}\}$ be a G-side-pairing for a finite family $\mathcal{P}$ of disjoint convex polyhedra in X. Then Φ is proper if and only if*

1. *each cycle of Φ is finite;*
2. *the isometry g_S fixes no point of S' for each S in $\mathcal{S}$; and*
3. *each edge cycle of Φ has dihedral angle sum 2π.*

Proof: Suppose that Φ is proper. Then every cycle of Φ is finite and has solid angle sum 4π; moreover, g_S fixes no point of S' for each S in $\mathcal{S}$ by Theorem 10.1.1. Let

$$[x] = \{x_1,\ldots,x_m\}$$

be an edge cycle of Φ. As $\omega[x] = 2\theta[x]$, we have that $\theta[x] = 2\pi$. Thus, every edge cycle of Φ has dihedral angle sum 2π.

Conversely, suppose that Φ satisfies (1)-(3). Then every cycle of Φ is finite by (1). Now let

$$[x] = \{x_1, \ldots, x_m\}$$

be a cycle of Φ. If x is in the interior of a polyhedron of $\mathcal{P}$, then $\omega[x] = 4\pi$. If x is in the interior of a side of a polyhedron of $\mathcal{P}$, then $\omega[x] = 4\pi$ by (2). If x is in the interior of an edge of a polyhedron of $\mathcal{P}$, then $[x]$ is an edge cycle, and we have by (3) that

$$\omega[x] = 2\theta[x] = 4\pi.$$

Now assume that x is a vertex of a polyhedron of $\mathcal{P}$. Then x_i is a vertex of a polyhedron P_i in $\mathcal{P}$ for each i. Let r be a positive real number such that r is less than half the distance from x_i to x_j for each $i \neq j$ and from x_i to any side of P_i not containing x_i for each i. Then $P_i \cap S(x_i, r)$ is a polygon in the sphere $S(x_i, r)$ and the polygons $\{P_i \cap S(x_i, r)\}$ are disjoint. Now the side-pairing Φ restricts to a proper side-pairing of the polygons $\{P_i \cap S(x_i, r)\}$. Hence, the space Σ obtained by gluing together the polygons has an orientable spherical structure by Theorem 9.2.3. Therefore Σ is a 2-sphere, since it is compact and connected. Hence $\omega[x] = 4\pi$. Thus Φ is proper. □

Example 1. Let P be a cube in E^3. Define a $\mathrm{T}(E^3)$-side-pairing Φ for P by pairing the opposite sides of P by translations. Then each edge cycle of Φ consists of four points. Therefore, each edge cycle of Φ has dihedral angle sum 2π. Hence Φ is proper by Theorem 10.1.3. Therefore, the space M obtained by gluing together the sides of P by Φ is a $\mathrm{T}(E^3)$-manifold by Theorem 10.1.2. The 3-manifold M is called the *cubical Euclidean 3-torus*.

Example 2. Let $D(r)$ be a regular spherical dodecahedron inscribed on the sphere $S(e_4, r)$ in S^3 with $0 < r \leq \pi/2$. Let $\theta(r)$ be the dihedral angle of $D(r)$. When r is small, $\theta(r)$ is approximately equal to but greater than the value of the dihedral angle of a Euclidean regular dodecahedron, which is approximately $116°, 34'$. As r increases, $\theta(r)$ increases continuously until it reaches $\theta(\pi/2)$, the dihedral angle of a regular dodecahedron in S^3 with vertices on S^2. As $\partial D(\pi/2) = S^2$, we have that $\theta(\pi/2) = 180°$. Now as $\theta(r)$ is a continuous function of r, taking values in the interval $(\theta(0),\ \theta(\pi/2)]$, there is a unique value of r such that $\theta(r) = 120°$. Let $P = D(r)$ for this value of r. Then P is a regular spherical dodecahedron all of whose proper dihedral angles are $2\pi/3$.

Define an $\mathrm{I}_0(S^3)$-side-pairing Φ for P by pairing the opposite sides of P with a twist of $\pi/5$. See Figure 10.1.1. Then each edge cycle of Φ consists of three points. Therefore, each edge cycle of Φ has dihedral angle sum 2π. Hence Φ is proper by Theorem 10.1.3. Therefore, the space M obtained by gluing together the sides of P by Φ is an orientable spherical 3-manifold by Theorem 10.1.2. The 3-manifold M is called the *Poincaré dodecahedral space*.

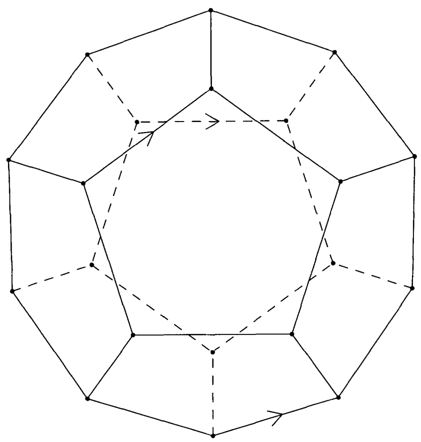


Figure 10.1.1. The gluing pattern for the Poincaré dodecahedral space

Example 3. By the argument in Example 4 of §7.1, there is a regular hyperbolic dodecahedron P in H^3 all of whose proper dihedral angles are $2\pi/5$. Define an $\mathrm{I}_0(H^3)$-side-pairing Φ for P by pairing the opposite sides of P with a twist of $3\pi/5$. See Figure 10.1.2. Then each edge cycle of Φ consists of five points. Therefore, each edge cycle of Φ has dihedral angle sum 2π. Hence Φ is proper by Theorem 10.1.3. Therefore, the space M obtained by gluing together the sides of P by Φ is a closed, orientable, hyperbolic 3-manifold by Theorem 10.1.2. The 3-manifold M is called the *Seifert-Weber dodecahedral space.*

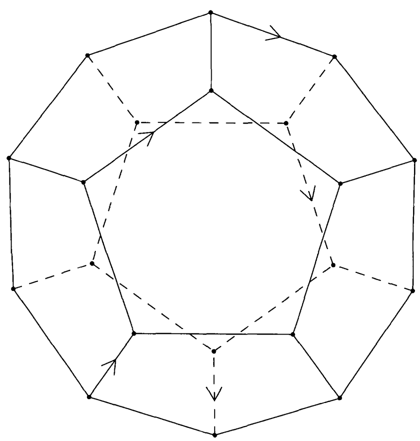


Figure 10.1.2. The gluing pattern for the Seifert-Weber dodecahedral space

Exercise 10.1

1. Prove that the Poincaré dodecahedral space has the same singular homology as the 3-sphere.
2. Prove that the fundamental group of the Poincaré dodecahedral space has order 120.
3. Compute the singular homology of the Seifert-Weber dodecahedral space.
4. Show that the Seifert-Weber dodecahedral space has a finite covering space whose first singular homology group is infinite.
5. Prove that there are an infinite number of nonisometric, closed, orientable, hyperbolic 3-manifolds.

§10.2. Complete Gluing of 3-Manifolds

Let M be a hyperbolic 3-manifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, convex, finite-sided polyhedra in H^3 of finite volume by a proper $\mathrm{I}(H^3)$-side-pairing Φ. In this section, we shall determine necessary and sufficient conditions such that M is complete.

It will be more convenient for us to work in the conformal ball model B^3. Then each polyhedron in $\mathcal{P}$ has only finitely many ideal vertices on the sphere S^2 at infinity by Theorems 6.3.25 and 6.3.26. We may assume, without loss of generality, that no two polyhedrons in $\mathcal{P}$ share an ideal vertex. Then the side-pairing Φ of the sides $\mathcal{S}$ of the polyhedra in $\mathcal{P}$ extends to a pairing of the ideal vertices of the polyhedra in $\mathcal{P}$, which, in turn, generates an equivalence relation on the set of all the ideal vertices of the polyhedra in $\mathcal{P}$. The equivalence classes are called *cycles*. The cycle containing an ideal vertex v is denoted by $[v]$. A cycle of ideal vertices is called a *cusp point* of the manifold M.

Let v be an ideal vertex of a polyhedron P_v in $\mathcal{P}$. Choose a horosphere Σ_v based at v that meets only the sides in $\mathcal{S}$ incident with v. The *link* of the ideal vertex v is defined to be the set

$$L(v) = P_v \cap \Sigma_v.$$

Note that $L(v)$ is a compact Euclidean polygon in the horosphere Σ_v, with respect to the natural Euclidean metric of Σ_v, whose similarity type does not depend on the choice of the horosphere Σ_v. For each cycle $[v]$ of ideal vertices, we shall assume that the horospheres $\{\Sigma_u : u \in [v]\}$ have been chosen small enough so that the links $\{L(u) : u \in [v]\}$ are disjoint. We next show that Φ determines a proper $\mathrm{S}(E^2)$-side-pairing of the polygons $\{L(u) : u \in [v]\}$.

Let g_S be an element of Φ and let u, u' be elements of $[v]$ such that $g_S(u') = u$. Then $\Sigma_{u'} \cap S'$ is a side of $L(u')$ and $\Sigma_u \cap S$ is a side of $L(u)$.

Now let

$$\overline{g}_S : \Sigma_{u'} \to g_S(\Sigma_{u'})$$

be the restriction of g_S. Then $\overline{g}_S$ is an isometry with respect to the natural Euclidean metrics of the horospheres $\Sigma_{u'}$ and $g_S(\Sigma_{u'})$. Observe that the line segment

$$g_S(\Sigma_{u'} \cap S') = g_S(\Sigma_{u'}) \cap S$$

is parallel to the line segment $\Sigma_u \cap S$ because $g_S(\Sigma_{u'})$ is concentric with Σ_u. Let

$$p_S : g_S(\Sigma_{u'}) \to \Sigma_u$$

be the radial projection of $g_S(\Sigma_{u'})$ onto Σ_u. Then p_S is a change of scale with respect to the natural Euclidean metrics of $g_S(\Sigma_{u'})$ and Σ_u. Define

$$h_S : \Sigma_{u'} \to \Sigma_u$$

by $h_S = p_S \overline{g}_S$. Then h_S is a similarity with respect to the natural Euclidean metrics of $\Sigma_{u'}$ and Σ_u. Moreover h_S maps the side $\Sigma_{u'} \cap S'$ of $L(u')$ onto the side $\Sigma_u \cap S$ of $L(u)$. Clearly $\{h_S\}$ is a proper $\mathrm{S}(E^2)$-side-pairing of the polygons $\{L(u)\}$. Here S ranges over the set of all the sides in $\mathcal{S}$ incident with the cycle $[v]$. We shall assume that the horospheres $\{\Sigma_u\}$ have been chosen so that $p_S = 1$ for the largest possible number of sides S.

Let $L[v]$ be the space obtained by gluing together the polygons $\{L(u)\}$ by $\{h_S\}$. Then $L[v]$ is a Euclidean similarity surface by Theorem 9.2.3. The surface $L[v]$ is called the *link* of the cusp point $[v]$ of the hyperbolic 3-manifold M obtained by gluing together the polyhedra in $\mathcal{P}$ by Φ. We now determine the topology of $L[v]$.

Theorem 10.2.1. *The link $L[v]$ of a cusp point $[v]$ of M is either a torus or a Klein bottle; moreover, if each element of Φ is orientation preserving, then $L[v]$ is a torus.*

Proof: By construction, $L[v]$ is a closed surface. By subdividing the polygons, if necessary, we may assume that all the polygons $\{L(u)\}$ are triangles. Let p, e, t be the number of vertices, edges, and triangles, respectively. Then we have $3t = 2e$, since each triangle has 3 edges and each edge bounds 2 triangles. Now the sum of all the angles of the triangles is πt on the one hand and $2\pi p$ on the other. Hence $t = 2p$. Therefore

$$\begin{aligned}\chi(L[v]) &= p - e + t \\ &= \frac{1}{2}t - \frac{3}{2}t + t = 0.\end{aligned}$$

Hence $L[v]$ is either a torus or a Klein bottle. If each element of Φ is orientation preserving, then each element of $\{h_S\}$ is orientation preserving, whence $L[v]$ is orientable; therefore $L[v]$ is a torus. □

Theorem 10.2.2. *The link $L[v]$ of a cusp point $[v]$ of M is complete if and only if links $\{L(u)\}$ for the ideal vertices in $[v]$ can be chosen so that Φ restricts to a side-pairing for $\{L(u)\}$.*

Proof: Suppose that Φ restricts to a side-pairing for $\{L(u)\}$. Then $h_S = \overline{g}_S$ for each S, and so $\{h_S\}$ is an $\mathrm{I}(E^2)$-side-pairing for $\{L(u)\}$. As $L[v]$ is compact, the $(E^2, \mathrm{I}(E^2))$-structure on $L[v]$ determined by $\{h_S\}$ is complete by Theorem 8.5.7. Hence $L[v]$ is a complete $(E^2, \mathrm{S}(E^2))$-surface by Theorem 8.5.8.

Conversely, suppose that $L[v]$ is complete. Let $\mathcal{G}$ be the abstract graph whose vertices are the elements of $[v]$ and whose edges are the sets $\{u, u'\}$ for which there is an element g_S of Φ such that $g_S(u') = u$. Then $\mathcal{G}$ is connected. Let $\mathcal{H}$ be the subgraph of $\mathcal{G}$ whose vertices are those of $\mathcal{G}$ and whose edges are the sets $\{u, u'\}$ for which there is an element g_S of Φ such that $g_S(u') = u$ and $p_S = 1$. We now show that $\mathcal{H}$ is connected. On the contrary, assume that $\mathcal{H}$ is disconnected. Then there is an edge $\{u, u'\}$ of $\mathcal{G}$ joining two components of $\mathcal{H}$. By rechoosing all the horospheres corresponding to one of these components by a uniform change of scale, we can add the edge $\{u, u'\}$ to $\mathcal{H}$. However, we assumed in the original choice of the horospheres that $\mathcal{H}$ has the largest possible number of edges. Thus $\mathcal{H}$ must be connected.

Now as $L[v]$ is complete, the $(E^2, \mathrm{S}(E^2))$-structure of $L[v]$ contains a $(E^2, \mathrm{I}(E^2))$-structure; moreover, since $\mathcal{H}$ is connected, we can choose the scale of the $(E^2, \mathrm{I}(E^2))$-structure on $L[v]$ so that the natural injection map of $L(u)^\circ$ into $L[v]$ is a local isometry for each u in $[v]$. Let g_S be an element of Φ such that $g_S(u') = u$. Then the restriction of h_S to the interior of the side $\Sigma_{u'} \cap S'$ of $L(u')$ is a local isometry because it factors through $L[v]$. Consequently h_S is an isometry and therefore $p_S = 1$. Thus Φ restricts to a side-pairing for $\{L(u)\}$. □

We now assume that $L[v]$ is complete. For greater clarity, we pass to the upper half-space model U^3 and assume, without loss of generality, that $v = \infty$. By Theorem 8.5.9, there is a group of isometries Γ_v of U^3 acting freely and discontinuously on Σ_v, and there is a $(E^2, \mathrm{I}(E^2))$-equivalence from Σ_v/Γ_v to $L[v]$ compatible with the projection from $L(v)$ to $L[v]$.

Let $B(v)$ be the open horoball based at v such that $\partial B(v) = \Sigma_v$. Then Γ_v acts freely and discontinuously on $B(v)$ as a group of isometries. Consequently $B(v)/\Gamma_v$ is a hyperbolic 3-manifold called a *solid horocusp*. It is clear from the gluing construction of M that we have the following 3-dimensional version of Theorem 9.8.4.

Theorem 10.2.3. *If the link $L[v]$ of a cusp point $[v]$ of M is complete, then there is an injective local isometry*

$$\iota : B(v)/\Gamma_v \to M$$

compatible with the projection of P_v to M.

We next consider the 3-dimensional version of Theorem 9.8.5.

Theorem 10.2.4. *Let M be a hyperbolic 3-manifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, convex, finite-sided polyhedra in H^3 of finite volume by a proper $\mathrm{I}(H^3)$-side-pairing Φ. Then M is complete if and only if $L[v]$ is complete for each cusp point $[v]$ of M.*

Proof: Suppose that $L[v]$ is incomplete for some ideal vertex v. By Theorem 10.2.2, there is a side S incident with $[v]$ such that $p_S \neq 1$. Let $\mathcal{H}$ be the graph in the proof of Theorem 10.2.2. Since $\mathcal{H}$ is connected, there are sides $S_1, \ldots, S_m$ incident with the cycle $[v]$ at ideal vertices $v_1, \ldots, v_m$, respectively, such that $g_{S_i}(v_{i+1}) = v_i$, and $g_{S_m}(v_1) = v_m$, and $p_{S_i} = 1$ for each $i = 1, \ldots, m-1$, and $S = S'_m$.

Let $L_i = L(v_i)$ for $i = 1, \ldots, m$. Choose a point x'_0 in the side $S \cap L_1$ of the polygon L_1. Let α_1 be a Euclidean geodesic arc in L_1 joining x'_0 to a point x_1 in the side $S_1 \cap L_1$ of L_1. We choose inductively a point x_i in the side $S_i \cap L_i$ of L_i and a Euclidean geodesic arc α_i in L_i joining x'_{i-1} to x_i for $i = 2, \ldots, m$ so that $p_S(x'_m) = x'_0$. If the point x'_m is closer to v_1 than x'_0, then the same argument as in the proof of Theorem 9.8.5 shows that the sequence $x_1, x_2, \ldots, x_m$ can be continued to a nonconvergent Cauchy sequence in M. If x'_0 is closer to v_1 than x'_m, then $x_m, x_{m-1}, \ldots, x_1$ can be continued to a nonconvergent Cauchy sequence in M. Thus M is incomplete.

Conversely, suppose that $L[v]$ is complete for each ideal vertex v. From Theorem 10.2.3, we deduce that there is a compact 3-manifold-with-boundary M_0 in M such that $M - M_0$ is the disjoint union of solid horocusps. The same argument as in the proof of Theorem 9.8.5 shows that M is complete. □

Exercise 10.2

1. Prove that the similarity type of the link of a cusp point $L[v]$ does not depend on the choice of the horospheres $\{\Sigma_u\}$.
2. Fill in the details of the proof of Theorem 10.2.3.
3. Prove that the horoball $B(v)$ in Theorem 10.2.3 can be replaced by a smaller concentric horoball so that ι maps the solid horocusp $B(v)/\Gamma_v$ isometrically onto its image in M.
4. Prove that a solid horocusp has finite volume.
5. Prove that the hypothesis of finite volume can be dropped from Theorem 10.2.4.
6. State and prove the 3-dimensional version of Theorem 9.8.7.

§10.3. Finite Volume Hyperbolic 3-Manifolds

In this section, we construct some examples of open, complete, hyperbolic 3-manifolds of finite volume obtained by gluing together a finite number of regular ideal polyhedra in H^3 along their sides. Each of these examples is homeomorphic to the complement of a knot or link in $\hat{E}^3$.

The Figure-Eight Knot Complement

Let T be a regular ideal tetrahedron in B^3. See Figure 10.3.1. Since the group of symmetries of T acts transitively on its edges, all the dihedral angles of T are the same. We now pass to the upper half-space model U^3 and position T with a vertex at ∞ as in Figure 10.3.2. Then a sufficiently high horizontal horosphere will intersect T in an equilateral Euclidean triangle. Therefore, all the dihedral angles of T are $\pi/3$.

Let T and T' be two disjoint regular ideal tetrahedrons in B^3. Label the sides and edges of T and T' as in Figure 10.3.3. Since a Möbius transformation of B^3 is determined by its action on the four vertices of T, the group of symmetries of T corresponds to the group of permutations of the vertices of T. Consequently, there is a unique orientation reversing isometry f_S of B^3 that maps T' onto T and side S' onto S in such a way as to preserve the gluing pattern between S' and S in Figure 10.3.3 for $S = A, B, C, D$.

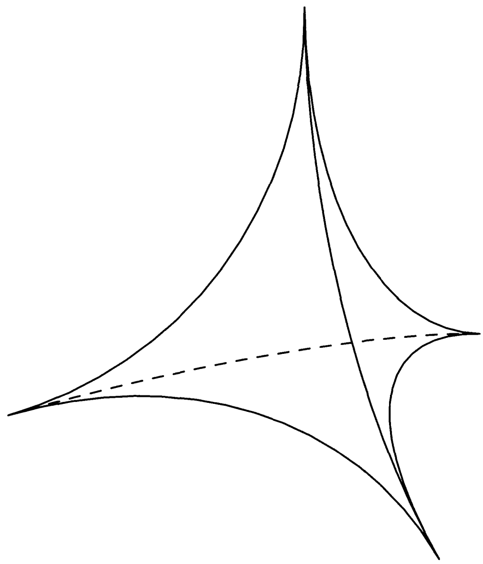


Figure 10.3.1. A regular ideal tetrahedron in B^3

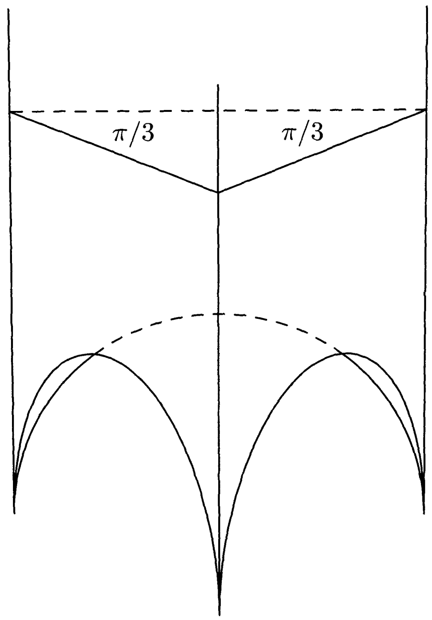


Figure 10.3.2. A regular ideal tetrahedron in U^3

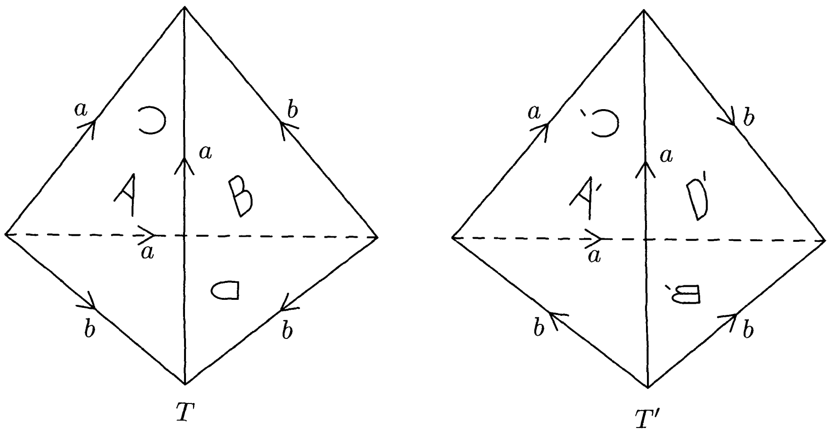


Figure 10.3.3. The gluing pattern for the figure-eight knot complement

Let g_S be the composite of f_S followed by the reflection in the side S. Then g_A, g_B, g_C, g_D and their inverses form an $\mathrm{I}_0(B^3)$-side-pairing Φ for $\{T, T'\}$. There are six points in each edge cycle of Φ. Hence, the dihedral angle sum of each edge cycle of Φ is 2π. Therefore Φ is a proper side-pairing.

Let M be the space obtained by gluing together T and T' by Φ. Then M is an orientable hyperbolic 3-manifold by Theorem 10.1.2. There is one cycle of ideal vertices. The link of the cusp point of M is a torus by Theorem 10.2.1. This can be seen directly in Figure 10.3.4.

Now choose disjoint horospheres based at the ideal vertices of T' that are invariant under the group of symmetries of T'. Then the isometries f_A, f_B, f_C, f_D will map these horospheres to horospheres based at the ideal vertices of T that are invariant under the group of symmetries of T. Consequently, these horospheres are paired by the elements of Φ. Therefore, the link of the cusp point of M is complete by Theorem 10.2.2. Thus M has one solid horocusp by Theorem 10.2.3. Finally M is complete by Theorem 10.2.4.

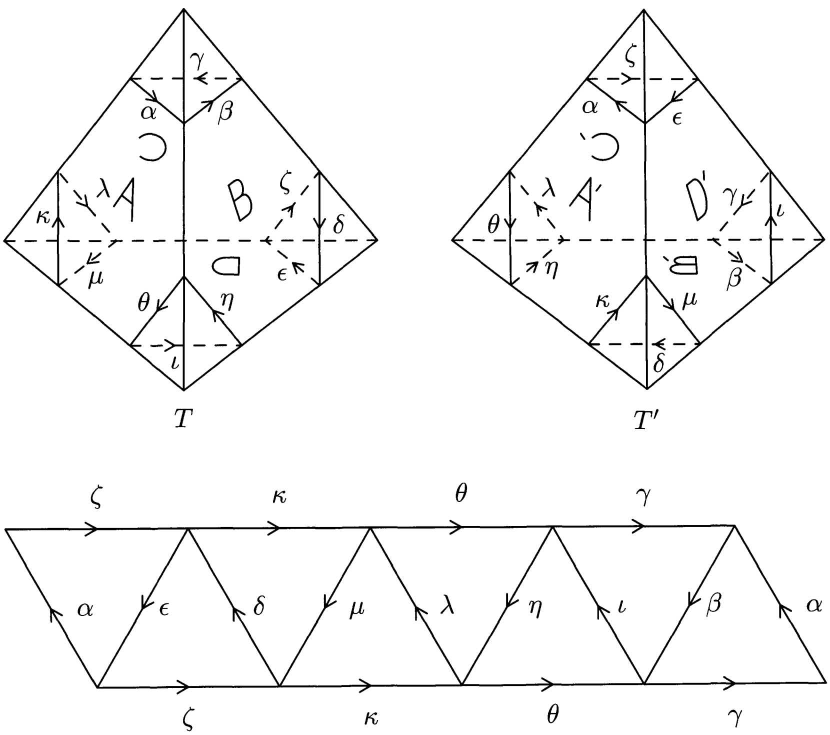


Figure 10.3.4. The link of the cusp point of the figure-eight knot complement

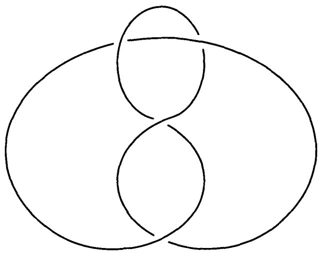


Figure 10.3.5. The figure-eight knot

Let K be a figure-eight knot in E^3. See Figure 10.3.5. We now show that M is homeomorphic to $\hat{E}^3 - K$. Drape the knot K over the top of the tetrahedron T and add directed arcs a, b to K as in Figure 10.3.6. These two arcs will correspond to the two edges a, b of M.

Now observe that the boundary of side A has the gluing pattern in Figure 10.3.7(a). The resulting quotient space is homeomorphic to a closed disk with two points removed as in Figure 10.3.7(b). This quotient space is homeomorphic to a disk with one interior point and part of its boundary removed as in Figures 10.3.7(c) and (d). The last disk spans the part of K that follows the contour of side A. Notice that the knot passes through the missing point of the interior of the disk in Figure 10.3.7(d).

Likewise, sides B, C, D of T give rise to disks that span the parts of K that follow the contours of sides B, C, D. See Figures 10.3.8-10.3.10. These four disks together with K form a 2-complex L whose 1-skeleton is the union of K and the arcs a, b. Let M^2 be the image of ∂T in M. From the compatibility of the gluing, we see that M^2 is homeomorphic to $L - K$.

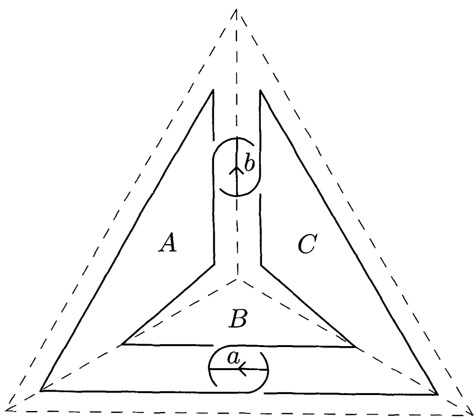


Figure 10.3.6. The figure-eight knot draped over the tetrahedron T

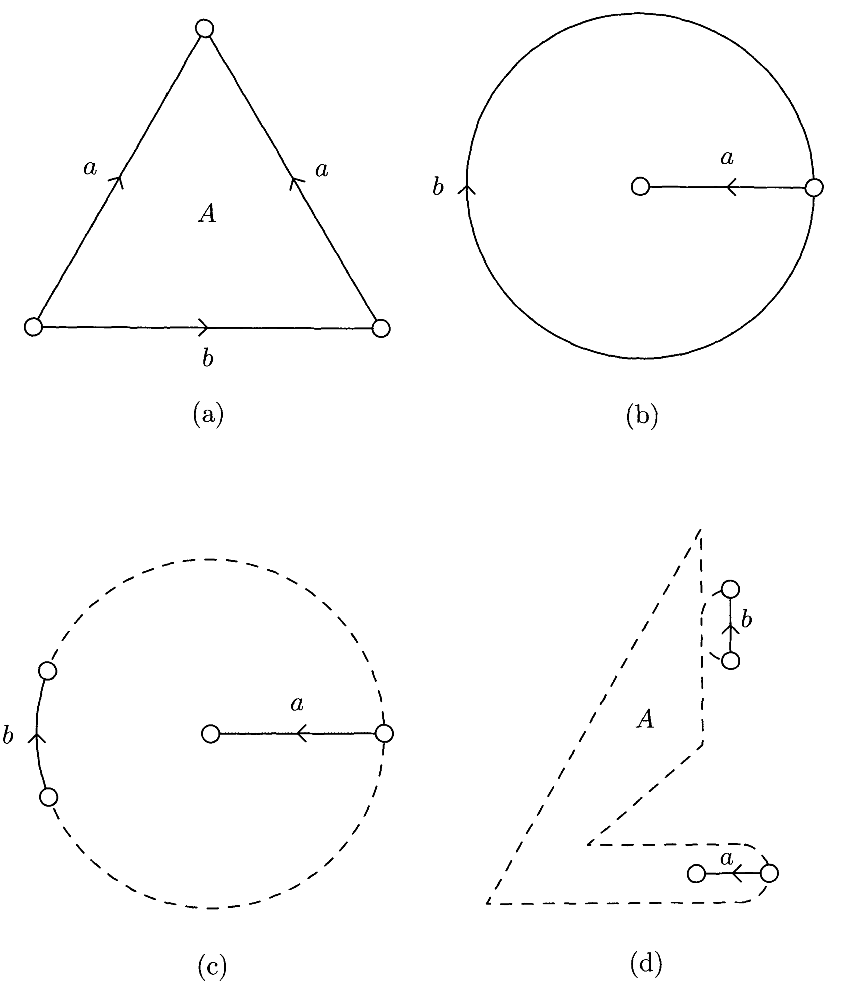


Figure 10.3.7. Side A deforming into a 2-cell of the complex L

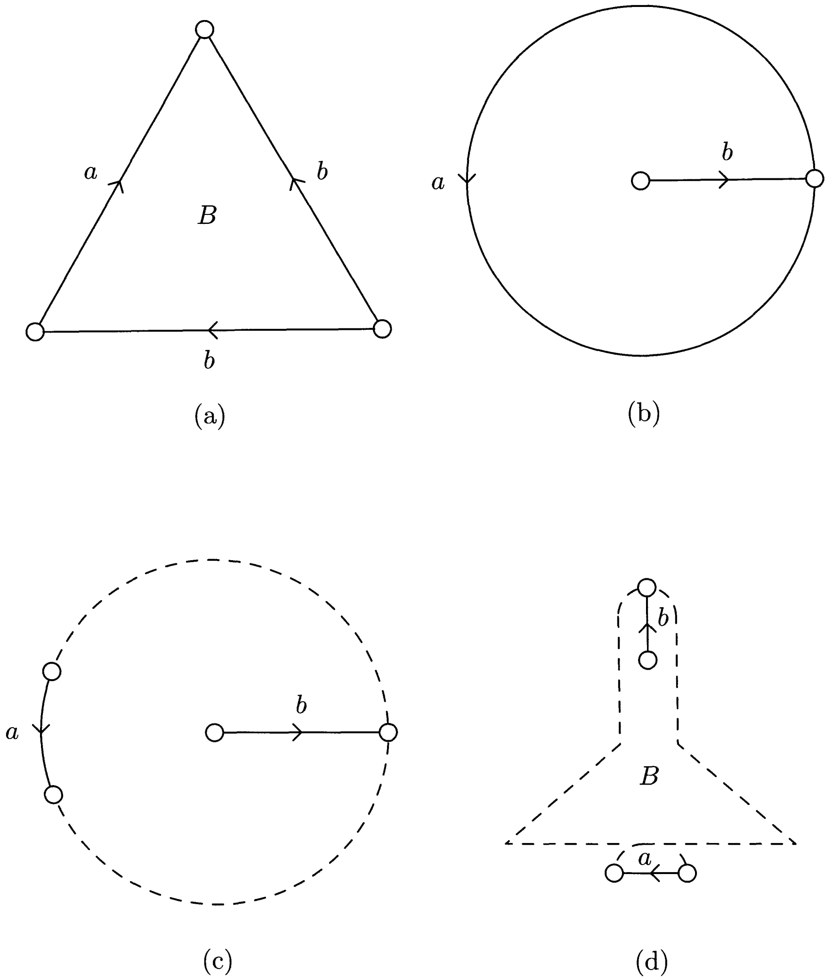


Figure 10.3.8. Side B deforming into a 2-cell of the complex L

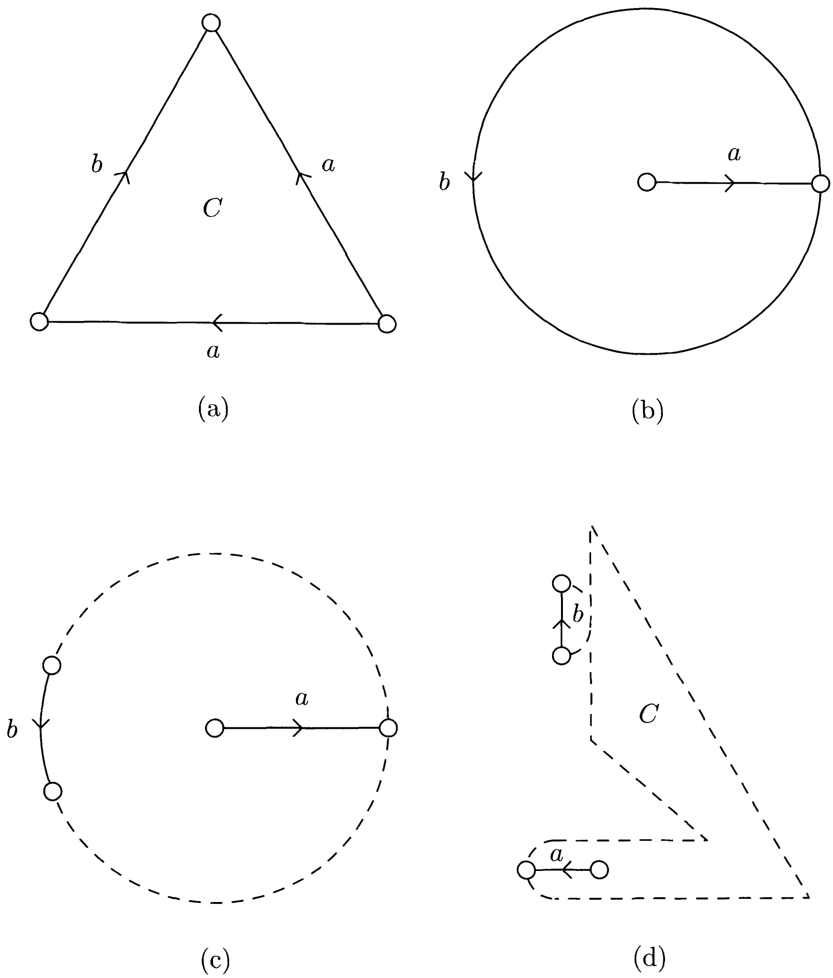


Figure 10.3.9. Side C deforming into a 2-cell of the complex L

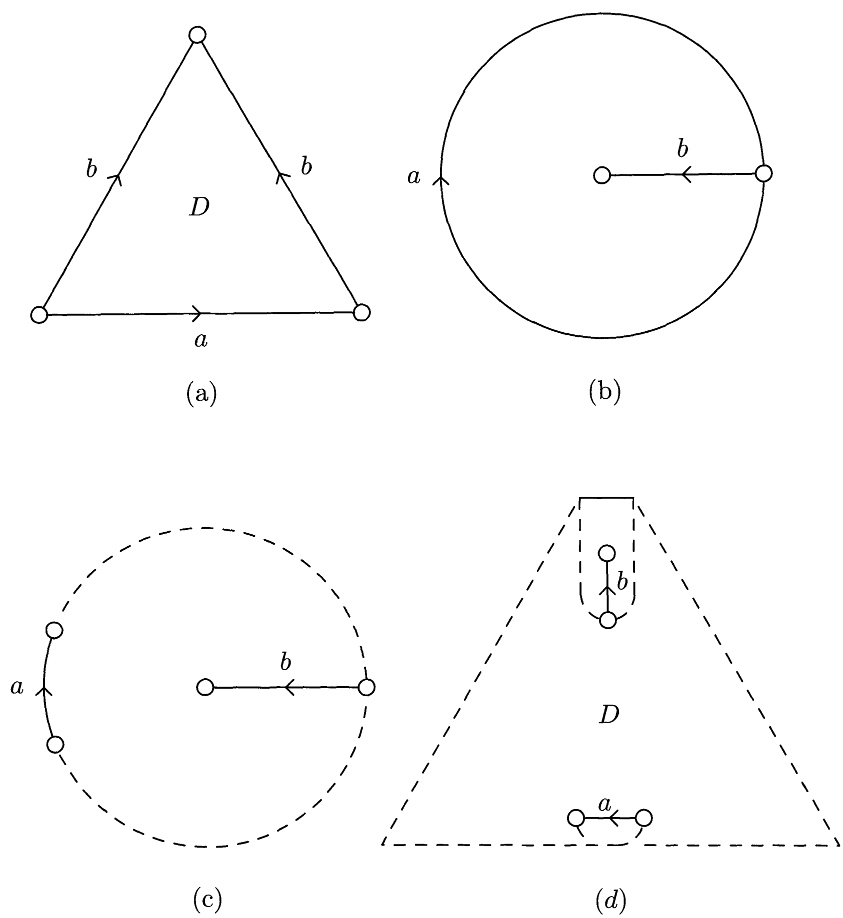


Figure 10.3.10. Side D deforming into a 2-cell of the complex L

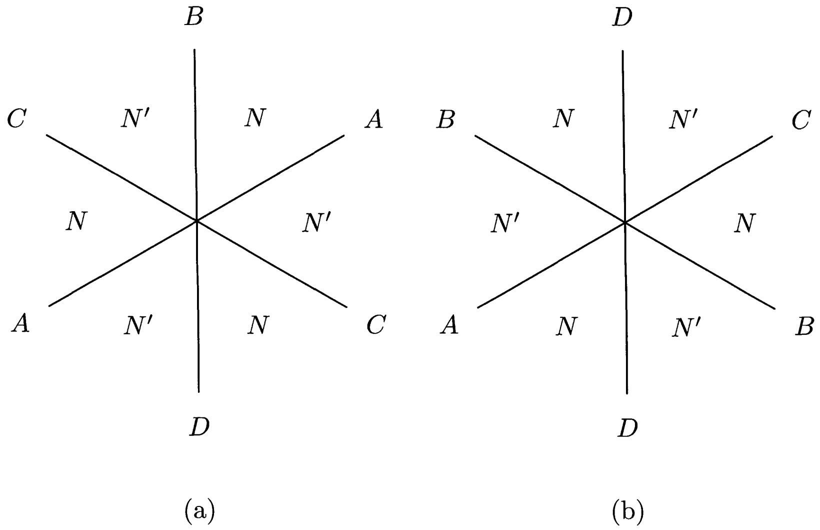


Figure 10.3.11. Cross sections normal to the arcs a and b pointing down

Observe that each of the arcs a, b meets all four of the 2-cells of L. By collapsing the arcs a, b to points, we see that L has the homotopy type of a 2-sphere. Consequently $\hat{E}^3 - L$ is the union of two open 3-balls. Now cut $\hat{E}^3 - K$ open along the interiors of the 2-cells of L and split apart the arcs a, b along their interiors to yield two connected 3-manifolds-with-boundary N and N' whose boundaries are 2-spheres minus four points with the same cell decomposition as the boundaries of T and T', respectively. Figure 10.3.11 illustrates cross sections of the subdivisions of $\hat{E}^3 - K$ normal to the arcs a and b. Notice that ∞ is in N. This explains the inside-out flip of the disks in Figures 10.3.7-10.3.10.

As the interiors of N and N' are open 3-balls, the manifolds N and N' are closed 3-balls minus four points on their boundaries. Consequently, there is a function

$$\phi : N \coprod N' \to T \coprod T'$$

that induces a homeomorphism from $\hat{E}^3 - K$ to M. Thus M is homeomorphic to the complement of a figure-eight knot in $\hat{E}^3$.

The Whitehead Link Complement

Let P be the regular ideal octahedron in B^3 with vertices $\pm e_1, \pm e_2, \pm e_3$. See Figure 10.3.12. By considering a regular ideal octahedron in U^3, with a vertex at ∞, as in Figure 10.3.13, we see that all the dihedral angles of P are $\pi/2$.

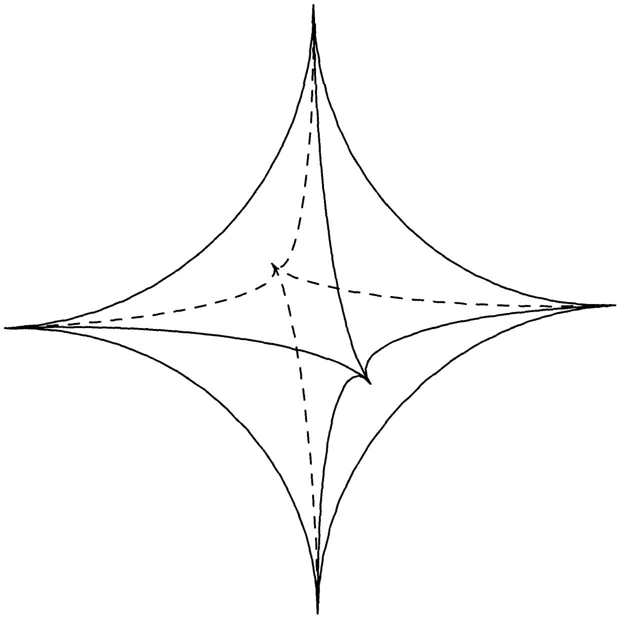


Figure 10.3.12. A regular ideal octahedron in B^3

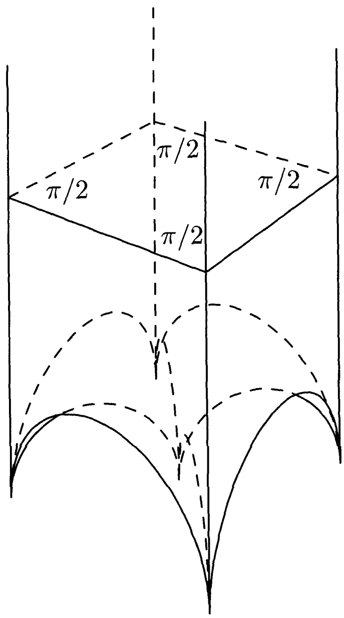


Figure 10.3.13. A regular ideal octahedron in U^3

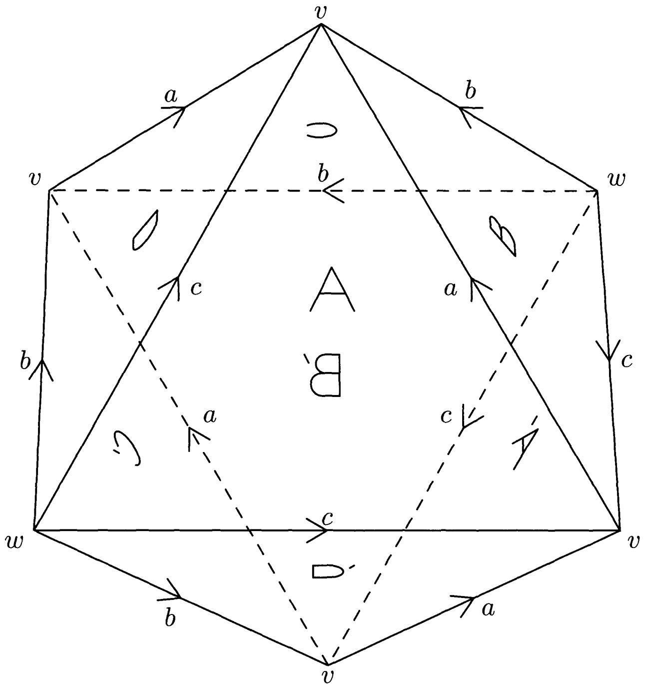


Figure 10.3.14. The gluing pattern for the Whitehead link complement

Now label the sides, edges, and vertices of P as in Figure 10.3.14. Let g_A be the Möbius transformation of B^3 that is the composite of the reflection in the plane of B^3 midway between the plane of side A and side A', then a $2\pi/3$ rotation in the plane of A about the center of A in the positive sense with respect to the outside of A, and then a reflection in the plane of A. Let g_B be defined as g_A except without the rotation. Let g_C be defined as g_A and let g_D be defined as g_B. Then g_A, g_B, g_C, g_D and their inverses form a $\mathrm{I}_0(B^3)$-side-pairing Φ for the polyhedron P. There are four points in each edge cycle of Φ. Hence, the dihedral angle sum of each edge cycle of Φ is 2π. Therefore Φ is a proper side-pairing.

Let M be the space obtained by gluing together the sides of P by Φ. Then M is an orientable hyperbolic 3-manifold by Theorem 10.1.2. There are two cycles of ideal vertices of P. The links of the cusp points of M are tori by Theorem 10.2.1. This can be seen directly in Figure 10.3.15. Each element g_S of Φ is the composite of an orthogonal transformation followed by the reflection in S. Consequently, disjoint horospheres based at the ideal vertices of P and equidistant from the origin are paired by the elements of Φ. Therefore, the links of the cusp points of M are complete by Theorem 10.2.2. Thus M has two disjoint solid horocusps by Theorem 10.2.3. Finally M is complete by Theorem 10.2.4.

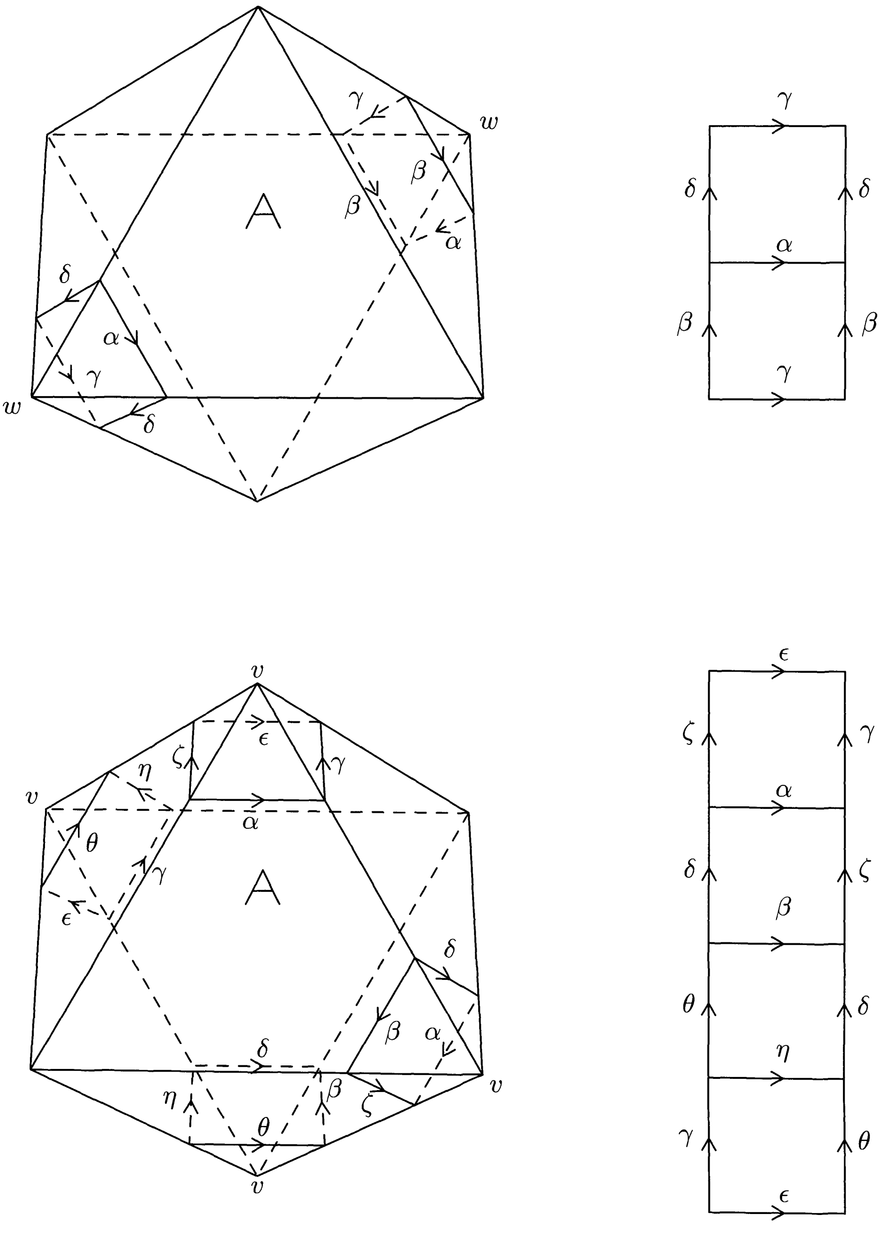


Figure 10.3.15. The links of the cusp points of the Whitehead link complement

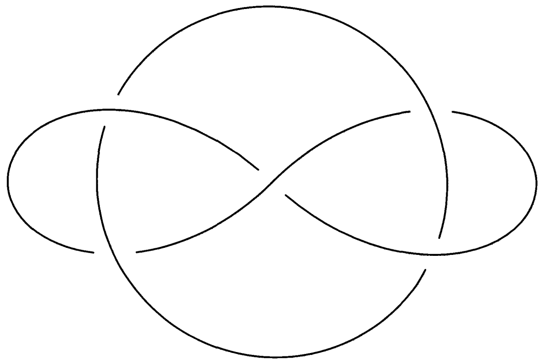


Figure 10.3.16. The Whitehead link

Let L be a Whitehead link in E^3. See Figure 10.3.16. We now show that M is homeomorphic to $\hat{E}^3 - L$. Drape the link L over the top pyramid of the regular octahedron and add three directed arcs a, b, c to L as in Figure 10.3.17. These three arcs will correspond to the three edges a, b, c of M.

Now observe that the boundary of side A of P has the gluing pattern in Figure 10.3.18(a). The resulting quotient space is homeomorphic to a closed disk with two points removed as in Figure 10.3.18(b). This quotient space is homeomorphic to a disk with one interior point and part of its boundary removed as in Figure 10.3.18(c). This last disk spans the right half of the component of L in Figure 10.3.17 that is in the shape of an infinity sign. Notice that the other component passes through the missing point of the interior of the disk in Figure 10.3.18(c).

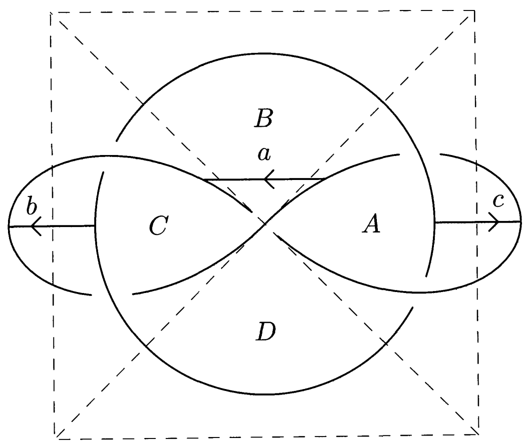


Figure 10.3.17. The Whitehead link draped over of a regular octahedron

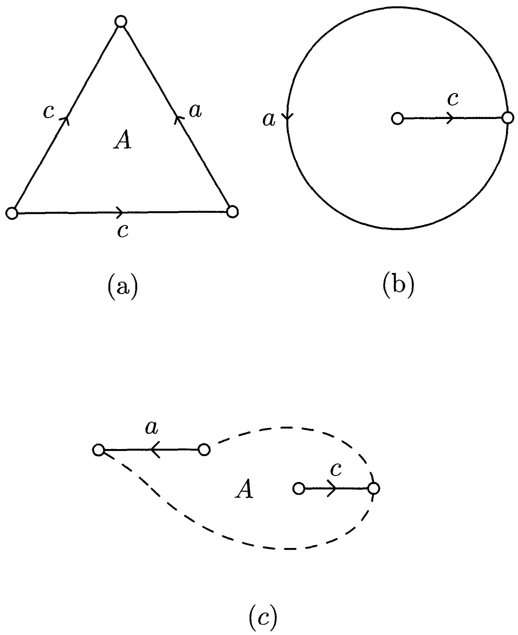


Figure 10.3.18. Side A deforming into a 2-cell of the complex K

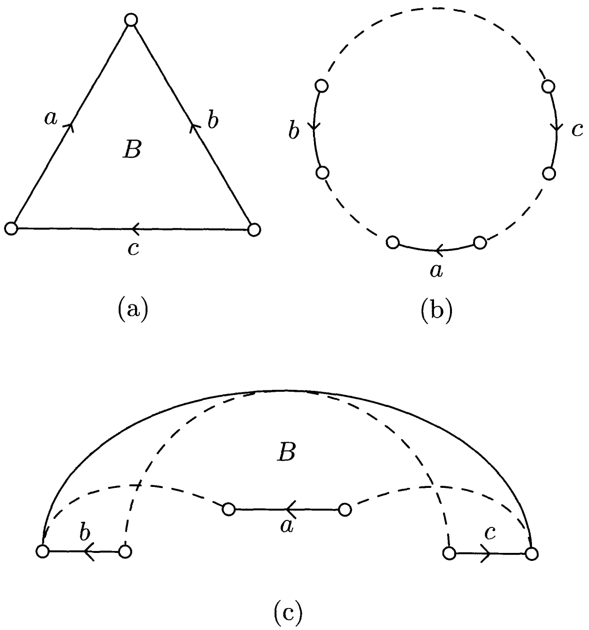


Figure 10.3.19. Side B deforming into a 2-cell of the complex K

Next, observe that the boundary of side B of P has the gluing pattern in Figure 10.3.19(a). The resulting quotient space is homeomorphic to a closed disk with part of the boundary removed as in Figure 10.3.19(b) and (c). The last disk spans the part of L in Figure 10.3.17 that follows the contour of side B. Likewise, the sides C and D of P give rise to disks that span the parts of L that follow the contours of sides C and D. These four disks together with L form a 2-complex K whose 1-skeleton is the union of L and the arcs a, b, c. Let M^2 be the image of ∂P in M. From the compatibility of the gluing, we see that M^2 is homeomorphic to $K - L$.

The 2-complex K is contractible because if we collapse the arcs a, b, c to points, we obtain a closed disk. Consequently $\hat{E}^3 - K$ is an open 3-ball. Now cut $\hat{E}^3 - L$ open along the interiors of the 2-cells of K and split apart the arcs a, b, c along their interiors to yield a 3-manifold-with-boundary N whose boundary is a 2-sphere minus six points with the same cell decomposition as ∂P. Now as the interior of N is an open 3-ball, N is a closed 3-ball minus six points on its boundary. Consequently, there is map $\phi : N \to P$ inducing a homeomorphism from $\hat{E}^3 - L$ to M. Thus M is homeomorphic to the complement of a Whitehead link in $\hat{E}^3$.

The Borromean Rings Complement

Let L be the Borromean rings in Figure 10.3.20. We now describe a hyperbolic structure for $\hat{E}^3 - L$.

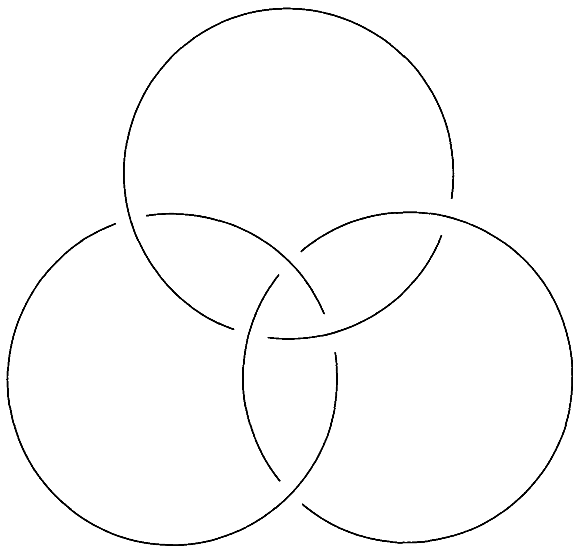


Figure 10.3.20. The Borromean rings

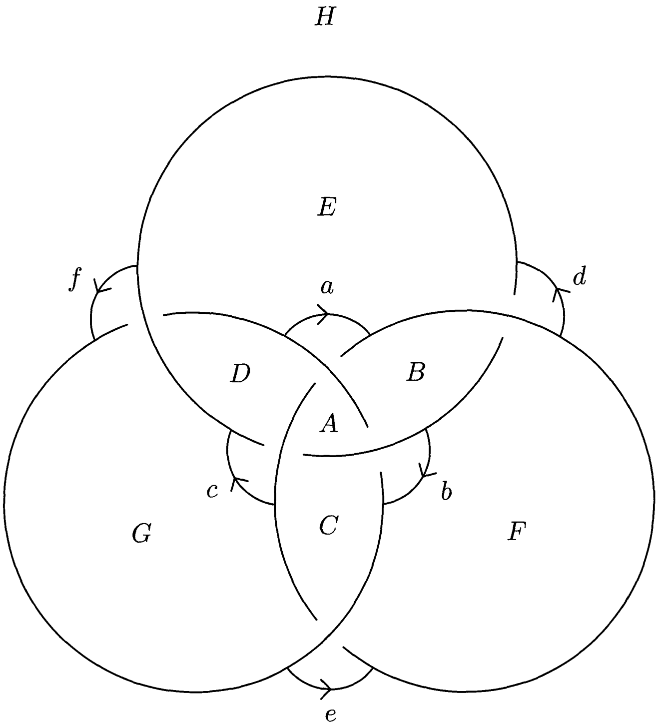


Figure 10.3.21. The 2-complex K

Adjoin six directed arcs $a, b, \ldots, f$ to L as in Figure 10.3.21. The union of L and these six arcs form the 1-skeleton of a 2-complex K whose 2-cells are disks corresponding to the eight regions $A, B, \ldots, H$ in Figure 10.3.21. Observe that each of the arcs $a, b, \ldots, f$ meets four of the 2-cells of K. By collapsing the arcs $a, b, \ldots, f$ to points, we see that K has the homotopy type of a 2-sphere. Consequently $\hat{E}^3 - K$ is the union of two open 3-balls.

Now cut $\hat{E}^3 - L$ open along the interiors of the 2-cells of K and split apart the arcs $a, b, \ldots, f$ along their interiors to yield two connected 3-manifolds-with-boundary N and N' whose boundaries are 2-spheres minus six points with the same cell decompositions as the boundaries of the octahedrons in Figure 10.3.22. As the interiors of N and N' are open 3-balls, N and N' are closed 3-balls minus six points on their boundaries. Consequently $\hat{E}^3 - L$ can be obtained by gluing together two regular ideal octahedrons along their sides by the side-pairing in Figure 10.3.22.

Notice that the paired sides are glued together with 120° rotations, alternating in direction from side to adjacent side. We leave it as an exercise to show that this side-pairing determines a complete hyperbolic structure for $\hat{E}^3 - L$.

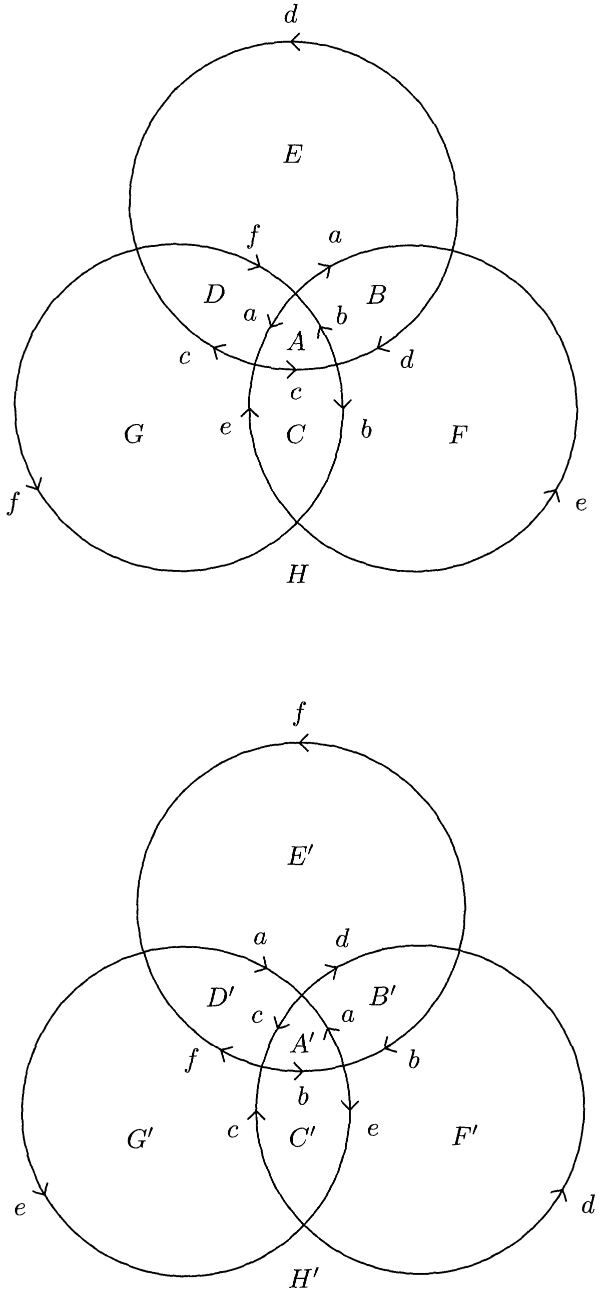


Figure 10.3.22. The gluing pattern for the Borromean rings complement

Exercise 10.3

1. Determine the class in $\mathcal{M}(T^2)$ of the link of the cusp point of the figure-eight knot complement.
2. Determine the classes in $\mathcal{M}(T^2)$ of the links of the cusp points of the Whitehead link complement.
3. Draw a picture of each of the 2-cells of the complex K in Figure 10.3.21.
4. Explain how the gluing pattern in Figure 10.3.22 is derived from the splitting of the complex in Figure 10.3.21.
5. Prove that the side-pairing of two regular ideal octahedrons described in Figure 10.3.22 induces a complete hyperbolic structure on the complement of the Borromean rings in $\hat{E}^3$.
6. Construct a complete hyperbolic manifold M by gluing together the sides of a regular ideal tetrahedron. The manifold M is called the *Gieseking manifold.*
7. Show that the link of the cusp point of the Gieseking manifold M is a Klein bottle. Conclude that M is nonorientable.
8. Show that the Gieseking manifold double covers the figure-eight knot complement.
9. Construct a complete, orientable, hyperbolic manifold M by gluing together two regular ideal tetrahedrons such that M is not homeomorphic to the figure-eight knot complement. The manifold M is called the *sister* of the figure-eight knot complement.
10. Show that the links of the cusp points of the figure-eight knot complement and its sister represent different classes in $\mathcal{M}(T^2)$.

§10.4. Hyperbolic Volume

In this section, we compute the volume of the hyperbolic 3-manifolds constructed in the last section. We begin by studying the geometry of ideal tetrahedra.

Ideal Tetrahedra

Let T be an ideal tetrahedron in H^3 and let Σ be a horosphere based at an ideal vertex v of T that does not meet the opposite side of T. Then

$$L(v) = \Sigma \cap T$$

is a Euclidean triangle, called the *link* of v in T. See Figure 10.4.1. Note that the orientation preserving similarity class of $L(v)$ does not depend on the choice of Σ.

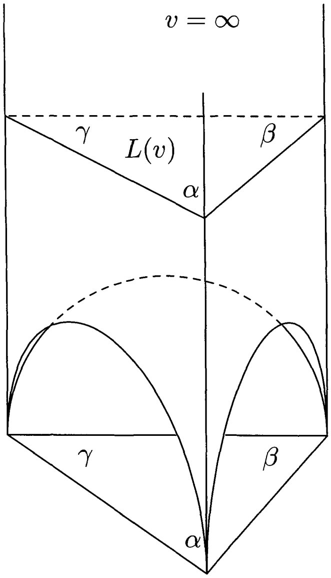


Figure 10.4.1. An ideal tetrahedron in U^3

Theorem 10.4.1. *The (orientation preserving) similarity class of the link $L(v)$ of a vertex v of an ideal tetrahedron T in H^3 determines T up to (orientation preserving) congruence.*

Proof: We pass to the upper half-space model U^3 and assume, without loss of generality, that $v = \infty$. Then the other three vertices of T form a triangle in E^2 that is in the orientation preserving similarity class of $L(v)$. See Figure 10.4.1. Now any (orientation preserving) similarity of E^2 extends to a unique (orientation preserving) isometry of U^3. Therefore, if T' is another ideal tetrahedron in U^3, with a vertex v' such that $L(v)$ is (directly) similar to $L(v')$, then T and T' are (directly) congruent. □

Theorem 10.4.2. *Let T be an ideal tetrahedron in H^3. Then T is determined, up to congruence, by the three dihedral angles α, β, γ of the edges incident to a vertex of T. Moreover, $\alpha + \beta + \gamma = \pi$ and the dihedral angles of opposite edges of T are equal. Furthermore, if α, β, γ are positive real numbers such that $\alpha + \beta + \gamma = \pi$, then there is an ideal tetrahedron in H^3 whose dihedral angles are α, β, γ.*

Proof: Let v be an ideal vertex of T. By Theorem 10.4.1, the congruence class of T is determined by the similarity class of $L(v)$, which, in turn, is determined by the dihedral angles α, β, γ of the edges of T incident to v.

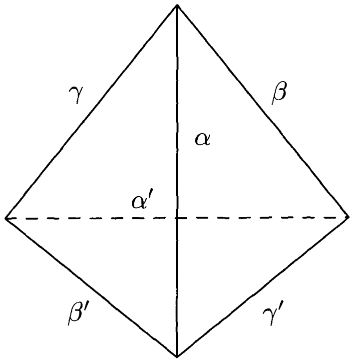


Figure 10.4.2. The dihedral angles of a tetrahedron

To see that the dihedral angles of the opposite sides of T are equal, label the dihedral angles of T as in Figure 10.4.2. Then we have the system of equations

$$\begin{cases} \alpha+\beta+\gamma=\pi, \\ \alpha+\beta'+\gamma'=\pi, \\ \alpha'+\beta'+\gamma=\pi, \\ \alpha'+\beta+\gamma'=\pi. \end{cases}$$

By adding the first two and the last two equations, we obtain the system

$$\begin{cases} 2\alpha+(\beta+\beta')+(\gamma+\gamma')=2\pi, \\ 2\alpha'+(\beta+\beta')+(\gamma+\gamma')=2\pi. \end{cases}$$

Therefore $\alpha=\alpha'$. The same argument shows that $\beta=\beta'$ and $\gamma=\gamma'$. The existence part of the theorem is left as an exercise for the reader. □

It follows from Theorems 10.4.1 and 10.4.2 that the orientation preserving similarity class of the link $L(v)$ of a vertex v of T does not depend on the choice of v. A simple geometric explanation of this fact is that the group of orientation preserving symmetries of T acts transitively on the set of vertices of T. See Exercise 10.4.3.

The Lobachevsky Function

We now study some of the properties of the *Lobachevsky function* $\text{Л}(\theta)$ defined by the formula

$$\text{Л}(\theta)=-\int_0^\theta \log|2\sin t|dt. \qquad (10.4.1)$$

Notice that the above integral is improper at all multiples of π. We will prove that $\text{Л}(\theta)$ is well defined and continuous for all θ. To begin with, we define $\text{Л}(0)=0$.

Let w be a complex number in the complement of the closed interval $[1, \infty)$. Then $1 - w$ is in the complement of the closed interval $(-\infty, 0]$. Define $\arg(1-w)$ to be the argument of $1-w$ in the interval $(-\pi, \pi)$. Then the formula

$$\log(1-w) = \log|1-w| + i\arg(1-w) \tag{10.4.2}$$

defines $\log(1-w)$ as an analytic function of w in the complement of the closed interval $[1, \infty)$. The relationship between $\log(1-w)$ and $\text{Л}(\theta)$ is revealed in the next lemma.

Lemma 1. *If* $0 < \theta < \pi$, *then*

$$\log(1 - e^{2i\theta}) = \log(2\sin\theta) + i(\theta - \pi/2).$$

Proof: Observe that

$$\begin{aligned} 1 - e^{2i\theta} &= 1 - (\cos 2\theta + i\sin 2\theta) \\ &= 1 - (\cos^2\theta - \sin^2\theta) - 2i\sin\theta\cos\theta \\ &= 2\sin^2\theta - 2i\sin\theta\cos\theta \\ &= 2\sin\theta(\sin\theta - i\cos\theta) \\ &= 2\sin\theta[\cos(\theta - \pi/2) + i\sin(\theta - \pi/2)]. \end{aligned}$$

The result now follows from Formula 10.4.2. □

Consider the function $\phi(w)$ defined by the formula

$$\phi(w) = \frac{-\log(1-w)}{w}. \tag{10.4.3}$$

The singularity at $w = 0$ is removable, since

$$\lim_{w \to 0} w\phi(w) = 0.$$

From the power series expansion

$$-\log(1-w) = \sum_{n=1}^{\infty} \frac{w^n}{n}, \quad \text{for } |w| < 1, \tag{10.4.4}$$

we find that

$$\phi(w) = \sum_{n=1}^{\infty} \frac{w^{n-1}}{n}, \quad \text{for } |w| < 1. \tag{10.4.5}$$

Thus $\phi(w)$ is analytic in the complement of the closed interval $[1, \infty)$.

The *dilogarithm function* $\psi(z)$ is defined as an analytic function of z on the complement of the closed interval $[1, \infty)$ by the formula

$$\psi(z) = \int_0^z \phi(w)dw. \tag{10.4.6}$$

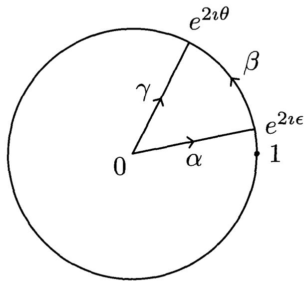


Figure 10.4.3. Curves α, β, γ in the unit disk

By integrating Formula 10.4.5, we find that

$$\psi(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2}, \qquad \text{for } |z| < 1.$$

Note that the above series converges uniformly on the closed disk $|z| \leq 1$. Now define

$$\psi(1) = \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Then $\psi(z)$ is continuous on the closed disk $|z| \leq 1$ and

$$\psi(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2}, \qquad \text{for } |z| \leq 1. \tag{10.4.7}$$

Let ϵ, θ be real numbers such that $0 < \epsilon < \theta < \pi$ and consider the curves α, β, γ in Figure 10.4.3. Since $\phi(w)$ is analytic in the complement of the closed interval $[1, \infty)$, we have

$$\int_\alpha \phi(w)dw + \int_\beta \phi(w)dw = \int_\gamma \phi(w)dw.$$

Hence, we have

$$\int_\beta \phi(w)dw = \psi(e^{2i\theta}) - \psi(e^{2i\epsilon}).$$

Let $w = e^{2i\theta}$. Then $dw/w = 2id\theta$. Hence, we have

$$\begin{aligned}
\int_\beta \phi(w)dw &= -\int_\beta \log(1-w)dw/w \\
&= -\int_\epsilon^\theta \log(1-e^{2it})2idt \\
&= -\int_\epsilon^\theta \big[\log(2\sin t) + i(t-\pi/2)\big]2idt \\
&= \big[t^2 - \pi t\big]_\epsilon^\theta - 2i\int_\epsilon^\theta \log(2\sin t)dt.
\end{aligned}$$

Thus

$$-2i\int_\epsilon^\theta \log(2\sin t)dt = \psi(e^{2i\theta}) - \psi(e^{2i\epsilon}) + \left[\pi t - t^2\right]_\epsilon^\theta.$$

Since ψ is continuous on the unit circle, we deduce that the improper integral

$$\int_0^\theta \log(2\sin t)dt = \lim_{\epsilon\to 0+}\int_\epsilon^\theta \log(2\sin t)dt$$

exists, and so $Л(\theta)$ is well defined for $0 < \theta < \pi$ and

$$2iЛ(\theta) = \psi(e^{2i\theta}) - \psi(1) + \pi\theta - \theta^2. \tag{10.4.8}$$

By letting $\theta \to \pi$, we find that $Л(\pi)$ exists and $Л(\pi) = 0$. Thus, Formula 10.4.8 holds for $0 \leq \theta \leq \pi$.

Theorem 10.4.3. *The function $Л(\theta)$ is well defined and continuous for all θ. Moreover, for all θ, the function $Л(\theta)$ satisfies the relations*

(1) $Л(\theta + \pi) = Л(\theta)$,

(2) $Л(-\theta) = -Л(\theta)$.

Proof: (1) As $Л(0) = 0 = Л(\pi)$ and $-\log|2\sin\theta|$ is periodic of period π, we deduce that $Л(\theta)$ is well defined for all θ, continuous, and periodic of period π. (2) As $-\log|2\sin\theta|$ is an even function, $Л(\theta)$ is an odd function. □

Theorem 10.4.4. *For each positive integer n, the function $Л(\theta)$ satisfies the identity*

$$Л(n\theta) = n\sum_{j=0}^{n-1} Л(\theta + j\pi/n).$$

Proof: Upon substituting $z = e^{2it}$ into the equation

$$z^n - 1 = \prod_{j=0}^{n-1}(z - e^{-2\pi ij/n}),$$

we obtain the equation

$$e^{2int} - 1 = \prod_{j=0}^{n-1} e^{2it}(1 - e^{-2it-2\pi ij/n}).$$

From the proof of Lemma 1, we have

$$|1 - e^{2i\theta}| = |2\sin\theta|.$$

Therefore, we have

$$|2\sin nt| = \prod_{j=0}^{n-1}|2\sin(t + j\pi/n)|.$$

Hence, we have

$$\int_0^\theta \log|2\sin nt|dt = \sum_{j=0}^{n-1}\int_0^\theta \log|2\sin(t+j\pi/n)|dt.$$

After changing variables, we have

$$\frac{1}{n}\int_0^{n\theta}\log|2\sin x|dx = \sum_{j=0}^{n-1}\int_{j\pi/n}^{\theta+j\pi/n}\log|2\sin x|dx.$$

Thus, we have

$$\frac{1}{n}Л(n\theta) = \sum_{j=0}^{n-1}Л(\theta+j\pi/n) - \sum_{j=0}^{n-1}Л(j\pi/n).$$

By Theorem 10.4.3, we have

$$Л((n-j)\pi/n) = Л(-j\pi/n) = -Л(j\pi/n).$$

Hence, we have

$$\sum_{j=0}^{n-1}Л(j\pi/n) = 0.$$

Thus, we have

$$\frac{1}{n}Л(n\theta) = \sum_{j=0}^{n-1}Л(\theta+j\pi/n).$$

□

By the fundamental theorem of calculus, we have

$$\begin{aligned}\frac{dЛ(\theta)}{d\theta} &= -\log|2\sin\theta|,\\ \frac{d^2Л(\theta)}{d\theta^2} &= -\cot\theta.\end{aligned}$$

Consequently, $Л(\theta)$ attains its maximum value at $\pi/6$ and its minimum value at $5\pi/6$. One can compute by numerical integration that

$$Л(\pi/6) = .5074708\ldots.$$

By Theorem 10.4.4, we have the equation

$$\frac{1}{2}Л(2\theta) = Л(\theta) + Л(\theta+\pi/2)$$

and therefore, by Theorem 10.4.3, we have

$$\frac{1}{2}Л(2\theta) = Л(\theta) - Л(\pi/2-\theta). \tag{10.4.9}$$

Substituting $\theta = \pi/6$ yields the equation

$$\frac{1}{2}Л(\pi/3) = Л(\pi/6) - Л(\pi/3).$$

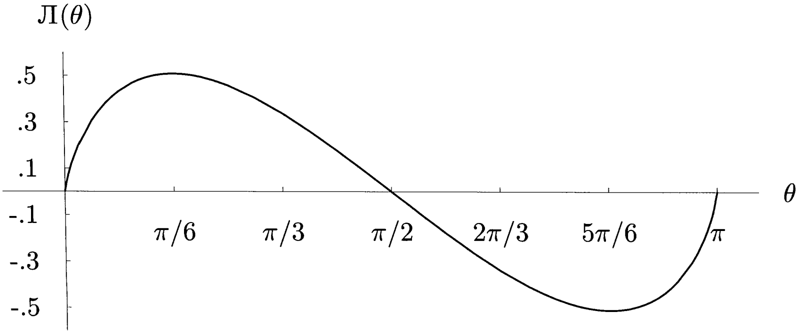


Figure 10.4.4. A graph of the Lobachevsky function

Thus, we have

$$Л(\pi/3) = \frac{2}{3}Л(\pi/6) = .3383138\ldots. \tag{10.4.10}$$

We now have enough information to sketch the graph of $Л(\theta)$. See Figure 10.4.4.

Volumes of Noncompact Tetrahedra

Consider a noncompact tetrahedron $S_{\alpha,\beta}$ in U^3 with three right dihedral angles and three other dihedral angles $\alpha, \pi/2 - \alpha, \beta$ as in Figure 10.4.5.

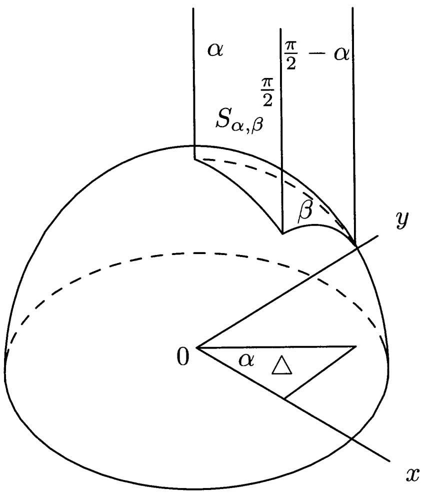


Figure 10.4.5. The tetrahedron $S_{\alpha,\beta}$

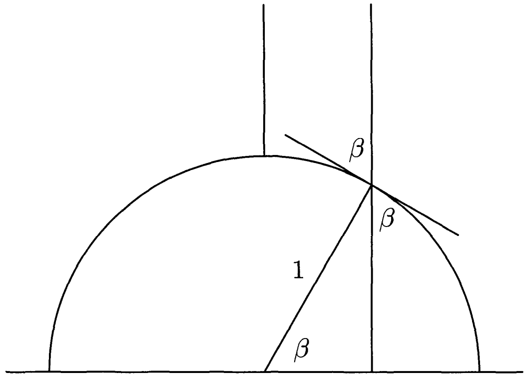


Figure 10.4.6. The side of $S_{\alpha,\beta}$ on the xz-plane

Theorem 10.4.5. *The volume of the tetrahedron $S_{\alpha,\beta}$ is given by*

$$\mathrm{Vol}(S_{\alpha,\beta}) = \frac{1}{4}\left[\text{Л}(\alpha+\beta) + \text{Л}(\alpha-\beta) + 2\text{Л}(\pi/2-\alpha)\right].$$

Proof: We may assume, without loss of generality, that the ideal vertex of $S_{\alpha,\beta}$ is at ∞ and that the base of $S_{\alpha,\beta}$ is on the unit sphere. The vertical projection of $S_{\alpha,\beta}$ to the xy-plane is a Euclidean right triangle $\triangle$. We may assume that $\triangle$ is situated as in Figure 10.4.5. From Figure 10.4.6 we deduce that the base of $\triangle$ has length $\cos\beta$. Thus $\triangle$ is the set of all points (x, y) satisfying the inequalities

$$\begin{aligned} 0 \leq x &\leq \cos\beta, \\ 0 \leq y &\leq x\tan\alpha. \end{aligned}$$

By Theorem 4.6.7, the element of hyperbolic volume of U^3 is $dxdydz/z^3$. Thus, the volume V of $S_{\alpha,\beta}$ is given by the formula

$$\begin{aligned} V &= \iint_{\triangle}\int_{\sqrt{1-x^2-y^2}}^{\infty} \frac{dxdydz}{z^3} \\ &= \iint_{\triangle} \frac{dxdy}{2(1-x^2-y^2)}. \end{aligned}$$

Now let $u = \sqrt{1-x^2}$. Then we have

$$\begin{aligned} V &= \int_0^{\cos\beta}\int_0^{x\tan\alpha} \frac{dydx}{2(u^2-y^2)} \\ &= \int_0^{\cos\beta} \frac{1}{4u}\log\left|\frac{u+x\tan\alpha}{u-x\tan\alpha}\right| dx \\ &= \int_0^{\cos\beta} \frac{1}{4u}\log\left|\frac{u\cos\alpha+x\sin\alpha}{u\cos\alpha-x\sin\alpha}\right| dx. \end{aligned}$$

Let $x = \cos\theta$. Then $u = \sin\theta$ and $dx/u = -d\theta$. Hence

$$\begin{aligned} V &= -\frac{1}{4}\int_{\pi/2}^{\beta} \log\left|\frac{2\sin(\theta+\alpha)}{2\sin(\theta-\alpha)}\right| d\theta \\ &= \frac{1}{4}\big[\text{Л}(\beta+\alpha) - \text{Л}(\pi/2+\alpha) - \text{Л}(\beta-\alpha) + \text{Л}(\pi/2-\alpha)\big] \\ &= \frac{1}{4}\big[\text{Л}(\alpha+\beta) + \text{Л}(\alpha-\beta) + 2\text{Л}(\pi/2-\alpha)\big]. \end{aligned}$$

□

Now suppose that the tetrahedron $S_{\alpha,\beta}$ has two ideal vertices. Then the vertex $(\cos\beta, \cos\beta\tan\alpha)$ of $\triangle$ is on the unit circle. Hence

$$\cos^2\beta + \cos^2\beta\tan^2\alpha = 1.$$

Thus $\tan^2\alpha = \tan^2\beta$ and so $\alpha = \beta$. By Theorem 10.4.5, we have

$$\text{Vol}(S_{\alpha,\alpha}) = \frac{1}{4}\big[\text{Л}(2\alpha) + 2\text{Л}(\pi/2-\alpha)\big].$$

By Formula 10.4.9, we have

$$2\text{Л}(\pi/2-\alpha) = 2\text{Л}(\alpha) - \text{Л}(2\alpha).$$

This implies the next result.

Corollary 1. *The volume of the tetrahedron $S_{\alpha,\alpha}$ is given by*

$$\text{Vol}(S_{\alpha,\alpha}) = \frac{1}{2}\text{Л}(\alpha).$$

Let $T_{\alpha,\beta,\gamma}$ be an ideal tetrahedron in U^3 with dihedral angles α, β, γ. We now compute the volume of $T_{\alpha,\beta,\gamma}$.

Theorem 10.4.6. *The volume of the ideal tetrahedron $T_{\alpha,\beta,\gamma}$ is given by*

$$\text{Vol}(T_{\alpha,\beta,\gamma}) = \text{Л}(\alpha) + \text{Л}(\beta) + \text{Л}(\gamma).$$

Proof: We may assume that one vertex of $T_{\alpha,\beta,\gamma}$ is at ∞ and that the base of $T_{\alpha,\beta,\gamma}$ is on the unit sphere. The vertical projection of $T_{\alpha,\beta,\gamma}$ to E^2 is a Euclidean triangle $\triangle$ with angles α, β, γ and vertices on the unit circle. There are three cases to consider. The origin is (1) in the interior of $\triangle$, (2) on a side of $\triangle$, or (3) in the exterior of $\triangle$.

(1) Suppose that the origin is in the interior of $\triangle$. Join the origin to the midpoints of the sides and the vertices of $\triangle$ by line segments. This subdivides $\triangle$ into six right triangles. Note that the pairs of triangles that share a perpendicular to a side of $\triangle$ are congruent. See Figure 10.4.7. Since an angle inscribed in a circle is measured by one half its intercepted arc, the angles around the origin are as indicated in Figure 10.4.7. Projecting this subdivision of $\triangle$ vertically upwards subdivides $T_{\alpha,\beta,\gamma}$ into six generalized tetrahedra of the form $S_{\theta,\theta}$ with $\theta = \alpha, \alpha, \beta, \beta, \gamma, \gamma$. See Figure 10.4.8. By Corollary 1, we have

$$\text{Vol}(T_{\alpha,\beta,\gamma}) = 2\left[\frac{1}{2}\text{Л}(\alpha) + \frac{1}{2}\text{Л}(\beta) + \frac{1}{2}\text{Л}(\gamma)\right]$$

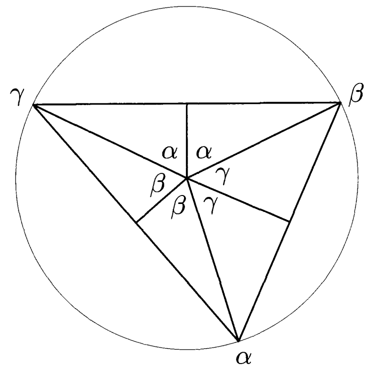


Figure 10.4.7. Subdivision of the triangle $\triangle$

(2) Now suppose that the origin is on a side of $\triangle$. Then $\triangle$ is inscribed in a semicircle. Hence, one of the angles of $\triangle$ is a right angle, say γ. Join the origin to the midpoints of the sides and vertices of $\triangle$ by line segments. This subdivides $\triangle$ into four right triangles. See Figure 10.4.9. The same argument as in case (1) shows that

$$\mathrm{Vol}(T_{\alpha,\beta,\pi/2}) = 2\left[\frac{1}{2}Л(\alpha) + \frac{1}{2}Л(\beta) + \frac{1}{2}Л(\pi/2)\right].$$

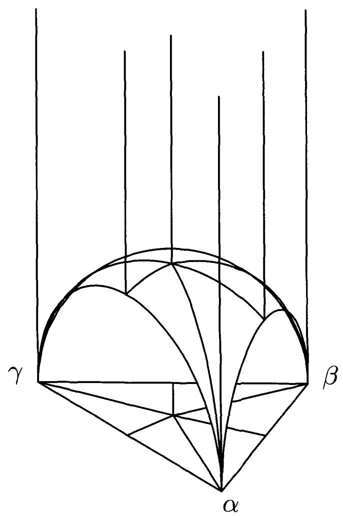


Figure 10.4.8. Subdivision of the tetrahedron $T_{\alpha,\beta,\gamma}$

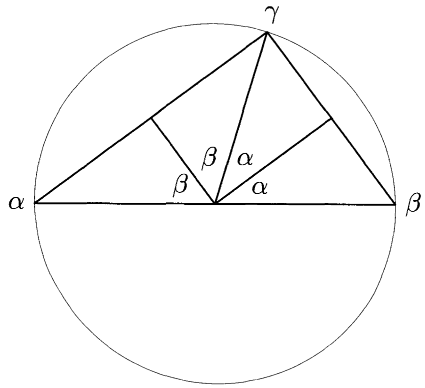


Figure 10.4.9. Subdivision of the triangle $\triangle$

(3) Now suppose that the origin is in the exterior of $\triangle$. Then one of the angles of $\triangle$ is obtuse, say γ. Join the origin to the midpoints of the sides and vertices of $\triangle$ by line segments. This expresses $\triangle$ as the union of four right triangles minus the union of two right triangles. See Figure 10.4.10. The same argument as in case (1) shows that

$$\mathrm{Vol}(T_{\alpha,\beta,\gamma}) = 2\left[\frac{1}{2}Л(\alpha) + \frac{1}{2}Л(\beta) - \frac{1}{2}Л(\pi-\gamma)\right]. \qquad \square$$

Example 1. The hyperbolic structure on the complement of the figure-eight knot constructed in the last section was obtained by gluing together two copies of $T_{\pi/3,\pi/3,\pi/3}$. Thus, its volume is $6Л(\pi/3) = 2.0298832\ldots$.

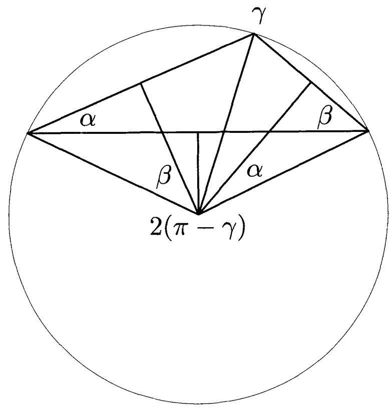


Figure 10.4.10. The triangle $\triangle$ expressed as the difference of right triangles

Theorem 10.4.7. *A tetrahedron of maximum volume in H^3 is a regular ideal tetrahedron.*

Proof: Since any tetrahedron in H^3 is contained in an ideal tetrahedron, it suffices to consider only ideal tetrahedra. Because of Theorem 10.4.6, we need to maximize the function

$$V(\alpha,\beta,\gamma) = Л(\alpha) + Л(\beta) + Л(\gamma)$$

subject to the constraints

$$\alpha,\beta,\gamma \geq 0 \quad \text{and} \quad \alpha+\beta+\gamma = \pi.$$

As V is continuous, it has a maximum value in the compact set $\alpha,\beta,\gamma \geq 0$ and $\alpha+\beta+\gamma = \pi$. Now $V(\alpha,\beta,\gamma) = 0$ if any one of α,β,γ is zero. Hence V attains its maximum value when $\alpha,\beta,\gamma > 0$. Let

$$f(\alpha,\beta,\gamma) = \alpha+\beta+\gamma.$$

Then by the Lagrange multiplier rule, there is a scalar λ such that

$$\mathrm{grad}(V) = \lambda\,\mathrm{grad}(f)$$

at any maximum point $(\alpha_0,\beta_0,\gamma_0)$. Then we have

$$Л'(\alpha_0) = Л'(\beta_0) = Л'(\gamma_0).$$

Therefore, we have

$$\sin\alpha_0 = \sin\beta_0 = \sin\gamma_0.$$

As $\alpha_0+\beta_0+\gamma_0 = \pi$, we deduce that $\alpha_0,\beta_0,\gamma_0 = \pi/3$. Thus, every ideal tetrahedron of maximum volume is regular. □

Let P be an ideal polyhedron in U^3 obtained by taking the cone to ∞ from an ideal n-gon on a hemispherical plane of U^3. Let $\alpha_1,\ldots,\alpha_n$ be the dihedral angles of P between its vertical sides and the base n-gon. We shall denote P by $P_{\alpha_1,\ldots,\alpha_n}$.

Theorem 10.4.8. *The polyhedron $P_{\alpha_1,\ldots,\alpha_n}$ has the following properties:*

(1) $\alpha_1+\alpha_2+\cdots+\alpha_n = \pi$,

(2) $\mathrm{Vol}(P_{\alpha_1,\ldots,\alpha_n}) = \sum_{i=1}^n Л(\alpha_i)$.

Proof: The proof is by induction on n. The case $n = 3$ follows from Theorems 10.4.2 and 10.4.6. Suppose that the theorem is true for $n-1$. By subdividing the base n-gon of $P_{\alpha_1,\ldots,\alpha_n}$ into an $(n-1)$-gon and a triangle, and taking the cone to ∞ on each polygon, we can subdivide $P_{\alpha_1,\ldots,\alpha_n}$ into the union of $P_{\alpha_1,\ldots,\alpha_{n-2},\beta}$ and $P_{\alpha_{n-1},\alpha_n,\pi-\beta}$. By the induction hypothesis, we have that

$$\alpha_1+\cdots+\alpha_{n-2}+\beta = \pi,$$
$$\alpha_{n-1}+\alpha_n+\pi-\beta = \pi.$$

Adding these two equations gives (1). Similarly, we have

$$\begin{aligned}\mathrm{Vol}(P_{\alpha_1,\ldots,\alpha_{n-2},\beta}) &= \left(\sum_{i=1}^{n-2} Л(\alpha_i)\right) + Л(\beta),\\ \mathrm{Vol}(P_{\alpha_{n-1},\alpha_n,\pi-\beta}) &= Л(\alpha_{n-1}) + Л(\alpha_n) + Л(\pi-\beta).\end{aligned}$$

Adding these two equations gives (2). □

Example 2. The hyperbolic structure on the complement of the Whitehead link constructed in the last section was obtained from a regular ideal octahedron, which can be subdivided into two copies of $P_{\pi/4,\pi/4,\pi/4,\pi/4}$. Therefore, its volume is

$$8Л(\pi/4) = 3.6638623\ldots.$$

Example 3. The hyperbolic structure on the complement of the Borromean rings constructed in the last section was obtained by gluing together two regular ideal octahedrons. Therefore, its volume is

$$16Л(\pi/4) = 7.3277247\ldots.$$

Exercise 10.4

1. Let L be the positive 3rd axis in U^3 and let r be a positive real number. Set
$$C(L,r) = \{x \in U^3 : d_U(x,L) = r\}.$$
Prove that $C(L,r)$ is a cone with axis L and cone point 0, and that the angle ϕ between L and $C(L,r)$ satisfies the equation $\sec\phi = \cosh r$.

2. Let K and L be two nonintersecting and nonasymptotic hyperbolic lines of B^3. Prove that there is a unique hyperbolic line M of B^3 perpendicular to both K and L.

3. Let T be an ideal tetrahedron in B^3 and let K, L, M be the perpendiculars to the opposite edges of T. Prove that the group Γ of orientation preserving symmetries of T contains the 180° rotations about K, L, M. Conclude that K, L, M meet at a common point and are pairwise orthogonal and that Γ acts transitively on the set of ideal vertices of T.

4. Deduce from Formula 10.4.8 that the function $Л(\theta)$ has the Fourier series expansion
$$Л(\theta) = \frac{1}{2}\sum_{n=1}^{\infty}\frac{\sin(2n\theta)}{n^2}.$$

5. Prove that the function $Л(\theta)$ has the series expansion
$$Л(\theta) = \theta - \theta\log(2\theta) + \sum_{n=1}^{\infty}\frac{B_n}{2n}\frac{(2\theta)^{2n+1}}{(2n+1)!} \quad \text{for } 0<\theta<\pi,$$
where $B_1 = 1/6,\ B_2 = 1/30,\ B_3 = 1/42, \ldots$ are Bernoulli numbers, by twice integrating the usual Laurent series expansion for the cotangent of θ.

6. Prove that the set of volumes of all the ideal tetrahedra in H^3 is the interval $(0, 3Л(\pi/3)]$.
7. Prove that a regular ideal hexahedron can be subdivided into five regular ideal tetrahedra.
8. Find the volume of a regular ideal dodecahedron.
9. Let P be a regular dodecahedron in H^3 whose dihedral angles are $2\pi/5$. Estimate the volume of P by finding the volumes of the inscribed and circumscribed balls about P.

§10.5. Hyperbolic Dehn Surgery

In this section, we construct hyperbolic structures for almost all the closed 3-manifolds obtained from $\hat{E}^3$ by performing Dehn surgery along the figure-eight knot. We begin by parameterizing Euclidean triangles.

Parameterization of Euclidean Triangles

Let $\triangle(u, v, w)$ be a Euclidean triangle in the complex plane $\mathbb{C}$ with vertices u, v, w labeled counterclockwise around $\triangle$. To each vertex of $\triangle$ we associate the ratio of the sides adjacent to the vertex in the following manner.

$$z(u) = \frac{w-u}{v-u}, \quad z(v) = \frac{u-v}{w-v}, \quad z(w) = \frac{v-w}{u-w}. \tag{10.5.1}$$

The complex numbers $z(u)$, $z(v)$, $z(w)$ are called the *vertex invariants* of the triangle $\triangle(u, v, w)$. See Figure 10.5.1.

Lemma 1. *The vertex invariants $z(u), z(v), z(w)$ depend only on the orientation preserving similarity class of the triangle $\triangle(u, v, w)$.*

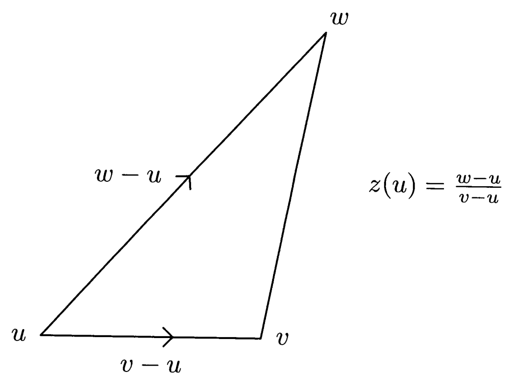


Figure 10.5.1. The vertex invariant $z(u)$ of the triangle $\triangle(u, v, w)$

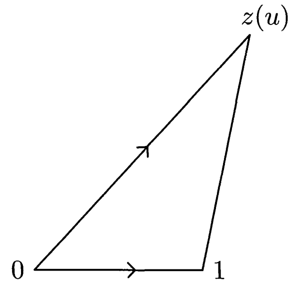


Figure 10.5.2. The triangle $\triangle(0, 1, z(u))$

Proof: An arbitrary orientation preserving similarity of $\mathbb{C}$ is of the form $x \mapsto ax + b$ with $a \neq 0$. Observe that

$$\begin{aligned} z(au+b) &= \frac{(aw+b)-(au+b)}{(av+b)-(au+b)} \\ &= \frac{a(w-u)}{a(v-u)} = z(u). \end{aligned}$$

□

Lemma 2. *Let $z(u)$ be a vertex invariant of a triangle $\triangle(u, v, w)$. Then*

(1) $\mathrm{Im}(z(u)) > 0$; *and*

(2) $\arg(z(u))$ *is the angle of* $\triangle(u, v, w)$ *at* u.

Proof: Define a similarity ϕ of $\mathbb{C}$ by

$$\phi(x) = \frac{x}{v-u} - \frac{u}{v-u}.$$

Then $\phi(u) = 0$, $\phi(v) = 1$, and $\phi(w) = z(u)$. As ϕ preserves orientation, the triangle $\triangle(0, 1, z(u))$ is labeled counterclockwise. See Figure 10.5.2. Hence $\mathrm{Im}(z(u)) > 0$, and $\arg(z(u))$ is the angle of $\triangle(u, v, w)$ at u. □

It is evident from Figure 10.5.2 that $z(u)$ determines the orientation preserving similarity class of $\triangle(u, v, w)$. Consequently $z(u)$ determines $z(v)$ and $z(w)$. By Lemma 1, we can calculate $z(v)$ and $z(w)$ from the triangle $\triangle(0, 1, z(u))$. This gives the relationships

$$z(v) = \frac{1}{1-z(u)}, \tag{10.5.2}$$

$$z(w) = \frac{z(u)-1}{z(u)}. \tag{10.5.3}$$

Example: For an equilateral triangle $\triangle(u, v, w)$, the vertex invariants $z(u)$, $z(v)$, $z(w)$ are all equal to $\frac{1}{2} + \frac{\sqrt{3}}{2}i$, since $\triangle(u, v, w)$ is directly similar to $\triangle(0, 1, \frac{1}{2} + \frac{\sqrt{3}}{2}i)$.

We now state precisely the parameterization of Euclidean triangles in $\mathbb{C}$ by their vertex invariants.

Theorem 10.5.1. *Let $\triangle(u, v, w)$ be a Euclidean triangle in $\mathbb{C}$, with vertices labeled counterclockwise and let $z_1 = z(u), z_2 = z(v), z_3 = z(w)$ be its vertex invariants. Then z_1, z_2, z_3 are in U^2 and satisfy the equations*

(1) $z_1 z_2 z_3 = -1$, *and*

(2) $1 - z_2 + z_1 z_2 = 0$.

Conversely, if z_1, z_2, z_3 are in U^2 and satisfy (1) *and* (2), *then there is a Euclidean triangle $\triangle$ in $\mathbb{C}$ that is unique up to orientation preserving similarity whose vertex invariants in counterclockwise order are z_1, z_2, z_3.*

Proof: By Formulas 10.5.2 and 10.5.3, we have

$$z_1 z_2 z_3 = z_1 \left(\frac{1}{1 - z_1}\right)\left(\frac{z_1 - 1}{z_1}\right) = -1.$$

As $z_2 = 1/(1 - z_1)$, we have $z_2 - z_1 z_2 = 1$.

Conversely, suppose that z_1, z_2, z_3 are in U^2 and satisfy equations (1) and (2). Then the vertex invariants of $\triangle(0, 1, z_1)$ are z_1, z_2, z_3. □

Parameterization of Ideal Tetrahedra

We now parameterize the ideal tetrahedra in H^3. Let v be a vertex of an ideal tetrahedron T in H^3. We label the edges of T, incident with v, with the corresponding vertex invariants z_1, z_2, z_3 of the link of v. Then opposite edges of T have the same label. The three parameters z_1, z_2, z_3 are indexed according to the right-hand rule with your thumb pointing towards a vertex of T. See Figure 10.5.3. The complex parameters z_1, z_2, z_3 are called the *edge invariants* of T.

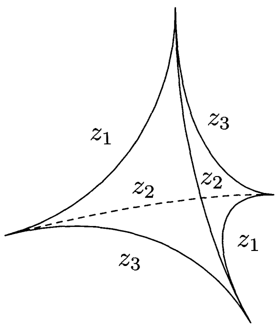


Figure 10.5.3. The edge invariants of an ideal tetrahedron

The next theorem follows immediately from Theorems 10.4.1 and 10.5.1.

Theorem 10.5.2. *Let z_1, z_2, z_3 be complex numbers in U^2 satisfying*

(1) $z_1 z_2 z_3 = -1$, *and*

(2) $1 - z_2 + z_1 z_2 = 0$.

Then there is a ideal tetrahedron T in H^3, unique up to orientation preserving congruence, whose edge invariants, in right-hand order, are z_1, z_2, z_3.

Gluing Consistency Conditions

Let Φ be an $\mathrm{I}_0(H^3)$-side-pairing for a finite family $\mathcal{T}$ of disjoint ideal tetrahedra in H^3. We now determine necessary and sufficient conditions on the edge invariants of the tetrahedra in $\mathcal{T}$ such that Φ is proper. The side-pairing Φ induces a pairing on the set $\mathcal{E}$ of edges of the tetrahedra in $\mathcal{T}$, which, in turn, generates an equivalence relation on $\mathcal{E}$. The equivalence classes of $\mathcal{E}$ are called *cycles of edges.*

Theorem 10.5.3. *Let Φ be an $\mathrm{I}_0(H^3)$-side-pairing for a finite family $\mathcal{T}$ of disjoint ideal tetrahedra in H^3. Then Φ is proper if and only if the invariants of each cycle of edges $\{E_1, \ldots E_m\}$ satisfy the equations*

(1) $z(E_1)z(E_2)\cdots z(E_m) = 1$, *and*

(2) $\arg z(E_1) + \arg z(E_2) + \cdots + \arg z(E_m) = 2\pi$,
where $0 < \arg z(E_i) < \pi$ *for each* i.

Proof: Let E_i be an edge of the side S_i of the tetrahedron T_i in $\mathcal{T}$. By reindexing, if necessary, we may assume that $g_{S_i}(E_{i+1}) = E_i$ for $i = 1, \ldots, m-1$ and $g_{S_m}(E_1) = E_m$. Define $g_1 = 1$ and $g_i = g_{S_1} \cdots g_{S_{i-1}}$ for $i = 2, \ldots, m+1$. Then $g_{m+1}(E_1) = E_1$. Orient T_i positively for each i. This orients each side of T_i. Now orient E_i positively with respect to S_i for each i. As g_{S_i} is orientation preserving, its restriction $g_{S_i} : S_i' \to S_i$ reverses orientation. As S_{i+1} and S_i' intersect along E_{i+1}, the edge E_{i+1} is oriented negatively with respect to S_i'. Therefore, the restriction $g_{S_i} : E_{i+1} \to E_i$ preserves orientation for $i = 1, \ldots, m-1$. Likewise $g_{S_m} : E_1 \to E_m$ preserves orientation. Hence g_{m+1} preserves the orientation of E_1. Thus, either g_{m+1} is the identity on E_1 or g_{m+1} acts as a nontrivial translation along E_1. In the latter case, Φ has an infinite cycle on E_1. Thus Φ has finite cycles if and only if g_{m+1} is the identity on E_1 for each cycle of edges $\{E_1, \ldots, E_m\}$.

The tetrahedrons T_i and $g_{S_i}(T_{i+1})$ lie on opposite side of their common side S_i and so the tetrahedrons $g_i T_i$ and $g_{i+1} T_{i+1}$ lie on opposite sides of their common side $g_i S_i$ for $i = 1, \ldots, m-1$. Now S_i and S_{i-1}' are the two sides of T_i intersecting along E_i and so $g_i S_i$ and $g_i S_{i-1}' = g_{i-1} S_{i-1}$

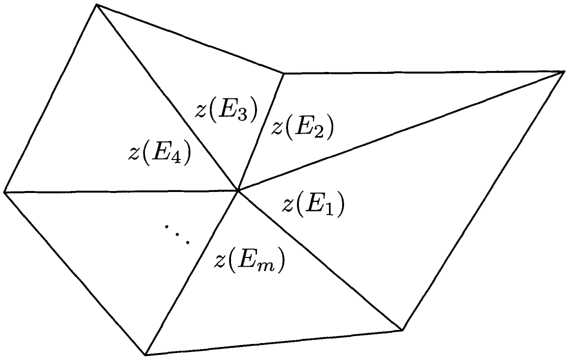


Figure 10.5.4. A cycle of Euclidean triangles

are the two sides of g_iT_i intersecting along E_1 for $i = 2, \ldots, m$. Therefore, the tetrahedra g_iT_i, for $i = 1, \ldots, m$, occur in sequential order rotating about the edge E_1 starting at the side S'_m of T_1 and ending at the side $g_mS_m = g_{m+1}S'_m$ of g_mT_m. Observe that $\{g_iT_i\}$ forms a cycle of tetrahedra around the edge E_1 if and only if the dihedral angle sum of the edges $E_1, \ldots, E_m$ is 2π and $g_{m+1} = 1$. Thus Φ is proper if and only if $\{g_iT_i\}$ forms a cycle of tetrahedra around E_1 for each cycle of edges $\{E_1, \ldots, E_m\}$.

By taking E_1 to be a vertical line of U^3, we see that $\{g_iT_i\}$ forms a cycle if and only if the orientation preserving similarity classes of Euclidean triangles determined by the invariants $z(E_1), \ldots, z(E_m)$ have representatives that form a cycle around a point of $\mathbb{C}$. See Figure 10.5.4. This will be the case if and only if

$$\arg z(E_1) + \arg z(E_2) + \cdots + \arg z(E_m) = 2\pi$$

and representatives can be chosen so that their sides match up correctly. As $|z(E_i)|$ is the ratio of the length of adjacent sides, the sides will match up correctly if and only if $|z(E_1)\cdots z(E_m)| = 1$. Thus Φ is proper if and only if the edge invariants of every cycle of edges satisfy equations (1) and (2). □

Hyperbolic Structures on the Figure-Eight Knot

Consider the gluing pattern on two parameterized ideal tetrahedrons T and T' in Figure 10.5.5 that gives the figure-eight knot complement. The gluing consistency equations for the two edge cycles are

$$z_1w_2z_2w_1z_2w_2 = 1 \quad \text{and} \quad z_1w_3z_3w_1z_3w_3 = 1,$$

or equivalently

$$z_1z_2^2w_1w_2^2 = 1 \quad \text{and} \quad z_1z_3^2w_1w_3^2 = 1.$$

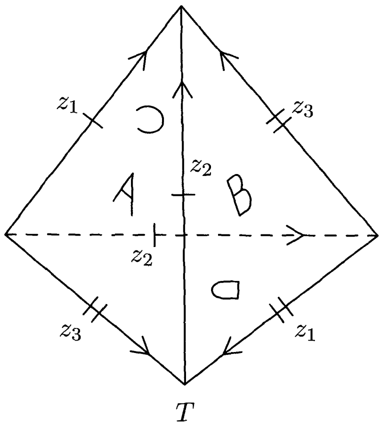


Figure 10.5.5. The gluing pattern for the figure-knot complement

As $z_1z_2z_3 = -1$ and $w_1w_2w_3 = -1$, the product of the two consistency equations is automatically satisfied

$$(z_1z_2z_3)^2(w_1w_2w_3)^2 = 1.$$

Thus, we need only consider one of the consistency equations, say

$$z_1z_2^2w_1w_2^2 = 1. \tag{10.5.4}$$

From Formulas 10.5.2 and 10.5.3, we have $z_2 = 1/(1 - z_1)$, and so $z_1z_2 = z_2 - 1$. Likewise $w_1w_2 = w_2 - 1$. Hence, upon substituting $z = z_2$ and $w = w_2$ into Formula 10.5.4, we have

$$z(z-1)w(w-1) = 1. \tag{10.5.5}$$

This gives the quadratic equation in z,

$$z^2 - z - (w(w-1))^{-1} = 0, \tag{10.5.6}$$

which has the solutions

$$z = \frac{1 \pm \sqrt{1 + 4(w(w-1))^{-1}}}{2}. \tag{10.5.7}$$

We want solutions such that $\mathrm{Im}(w) > 0$ and $\mathrm{Im}(z) > 0$. For each value of w, there is a unique solution for z, with $\mathrm{Im}(z) > 0$, provided the discriminant $1 + 4(w(w-1))^{-1}$ is not in the interval $[0, \infty)$.

Let $w = a + bi$ with a, b real and $b > 0$. Then

$$\begin{aligned} w(w-1) &= (a+bi)(a-1+bi) \\ &= (a(a-1) - b^2) + (b(a-1) + ab)i. \end{aligned}$$

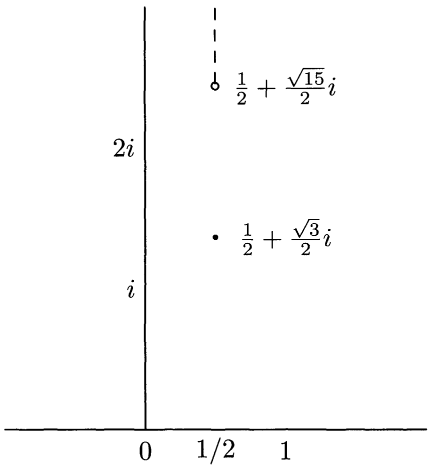


Figure 10.5.6. The solution space for w

Now suppose that $w(w-1)$ is real. Then

$$b(a-1) + ab = 0$$

and so $a = 1/2$. Thus

$$w(w-1) = -\frac{1}{4} - b^2.$$

Solving the inequality

$$1 + 4(w(w-1))^{-1} \geq 0$$

yields the inequality $b \geq \sqrt{15}/2$. Thus, the desired solutions correspond to the points in U^2 minus the ray $\{\frac{1}{2} + \frac{t}{2}i : t \geq \sqrt{15}\}$. See Figure 10.5.6.

We also need to satisfy the angle sum equations

$$\begin{array}{rcccl} s_1 & = & \arg(z_1) + 2\arg(z_2) + \arg(w_1) + 2\arg(w_2) & = & 2\pi, \\ s_2 & = & \arg(z_1) + 2\arg(z_3) + \arg(w_1) + 2\arg(w_3) & = & 2\pi. \end{array}$$

Now as

$$\arg(z_1) + \arg(z_2) + \arg(z_3) = \pi,$$

we have that

$$\arg(z_1) + 2\arg(z_2) < 2\pi.$$

Likewise, we have

$$\arg(w_1) + 2\arg(w_2) < 2\pi.$$

Therefore $s_1 < 4\pi$, and so $s_1 = 2\pi$. Likewise $s_2 = 2\pi$. The next theorem now follows from Theorem 10.5.3.

Theorem 10.5.4. *The hyperbolic structures on the figure-eight knot complement obtained by gluing together the parameterized ideal tetrahedrons T and T' according to the given pattern are parameterized by the points in the upper half w-plane minus the ray $\{\frac{1}{2}+\frac{t}{2}i : t \geq \sqrt{15}\}$. The parameterization is given by $w_2 = w$ and*

$$z_2 = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{w(w-1)}}.$$

The Uniqueness of the Complete Structure

Let M be the hyperbolic 3-manifold obtained by properly gluing together the ideal tetrahedrons T and T' according to the gluing pattern in Figure 10.5.5. We now show that $\frac{1}{2} + \frac{\sqrt{3}}{2}i$ is the only value of the parameter w for which M is complete.

Let L be the link of the cusp point of M. By Theorem 10.2.4, we have that M is complete if and only if L is complete. By Theorems 8.4.5, 8.5.8, and 8.5.9, we have that L is complete if and only if the holonomy

$$\eta : \pi_1(L) \to \mathrm{S}_0(\mathbb{C})$$

maps $\pi_1(L)$ isomorphically onto a freely acting discrete group of Euclidean isometries of $\mathbb{C}$. By Theorem 5.4.4, this is the case if and only if the image of η is a lattice group of translations of $\mathbb{C}$.

Now every element of $\mathrm{S}_0(\mathbb{C})$ is of the form $\phi(z) = \alpha z + \beta$ with α in $\mathbb{C}^*$ and β in $\mathbb{C}$; moreover, ϕ is a Euclidean translation if and only if $\alpha = 1$. Notice that the derivative of ϕ is $\phi'(z) = \alpha$, and so ϕ is a Euclidean translation if and only if $\phi'(z) = 1$.

We now compute the derivative of the holonomy of the similarity structure on L. Consider the pseudo-triangulation of L in Figure 10.5.7. After developing the triangulation of L onto $\mathbb{C}$, we can regard directed edges of the triangulation as vectors in $\mathbb{C}$. The ratio, as complex numbers, of any two vectors in the same triangle is known in terms of the vertex invariants. See Figure 10.5.1. This allows us to compute the derivative of the holonomy as a telescoping product of ratios.

Let x be the element of $\pi_1(L)$ represented by the base of the parallelogram in Figure 10.5.7. To compute $\eta'(x)$, we assign the value 1 to the base of triangle a and develop the triangulation of L onto $\mathbb{C}$ along x until we come to another copy of triangle a. See Figure 10.5.8(a). The values of the directed edges encountered along the way are given in terms of the vertex invariants by the equations

$$\frac{1}{v_1} = z_1, \ \frac{v_1}{v_2} = w_2, \ \ldots, \ \frac{v_{11}}{v_{12}} = z_3.$$

Therefore

$$\frac{1}{v_1}\frac{v_1}{v_2}\cdots\frac{v_{11}}{v_{12}} = z_1^2 z_2^2 z_3^4 w_1^2 w_2^2 = z_3^2 w_1^2 w_2^2.$$

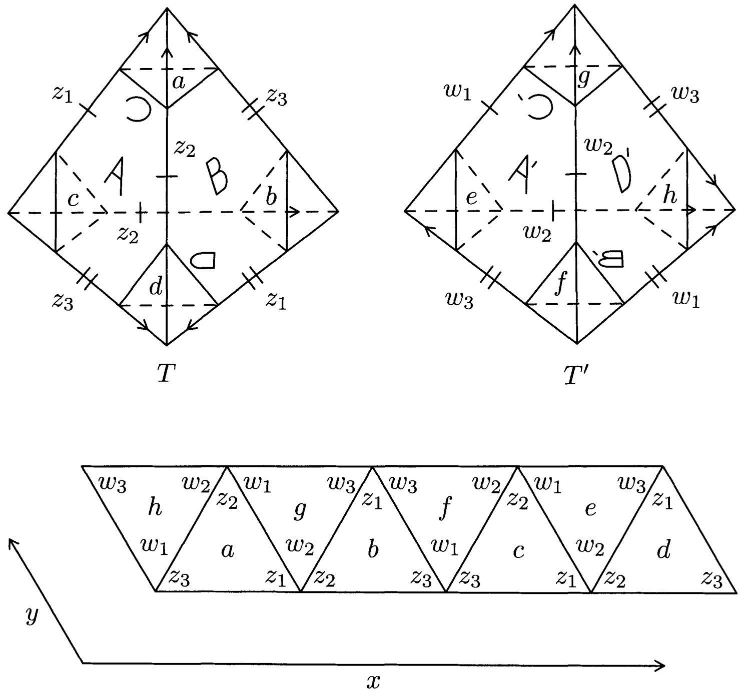


Figure 10.5.7. The link of the cusp point of the figure-eight knot complement

Hence, we have

$$1/v_{12} = z_3^2 w_1^2 w_2^2 = \left(\frac{w_1 w_2}{z_1 z_2}\right)^2 = \left(\frac{w-1}{z-1}\right)^2.$$

The value v_{12} of the base of the second triangle a is $\eta'(x)$. Thus

$$\eta'(x) = \left(\frac{z-1}{w-1}\right)^2. \tag{10.5.8}$$

Let y be the element of $\pi_1(L)$ represented by the left side of the parallelogram in Figure 10.5.7. From Figure 10.5.8(b), we compute

$$\eta'(y) = -z_3 w_1 w_3 = \frac{-1}{z_1 z_2 w_2} = \frac{1}{w(1-z)}.$$

From Formula 10.5.5, we have

$$\eta'(y) = z(1-w). \tag{10.5.9}$$

Now $\eta'(x) = 1$ if and only if $z = w$, and so $\eta'(x) = 1 = \eta'(y)$ if and only if $w(1-w) = 1$. Hence η' is trivial if and only if $w = \frac{1}{2} + \frac{\sqrt{3}}{2}i$. Thus M is complete if and only if $w = \frac{1}{2} + \frac{\sqrt{3}}{2}i$, that is, both T and T' are regular.

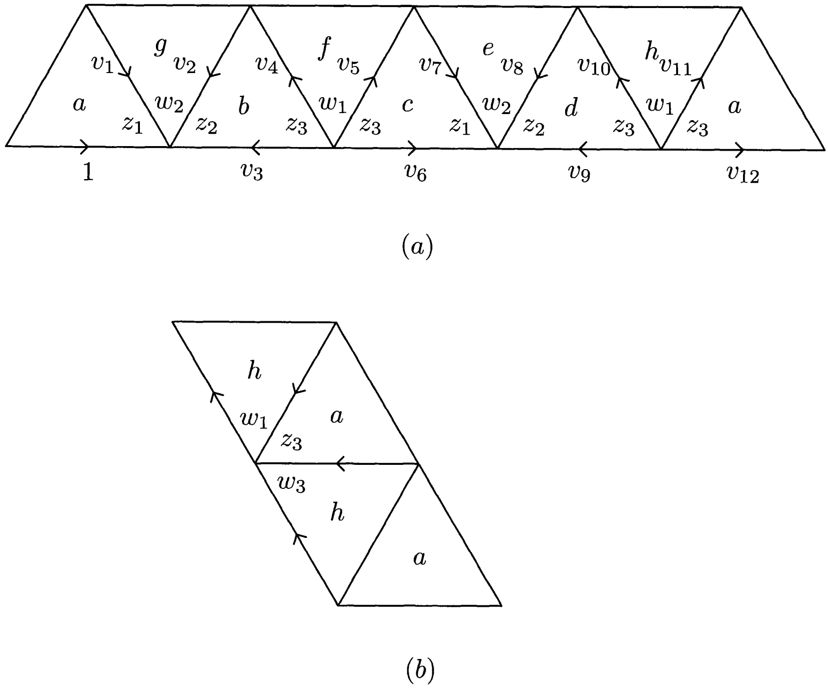


Figure 10.5.8. The developments of triangle a along x and y

The Metric Structure of the Link

We now assume that M is incomplete. Then the link L of the cusp point of M is incomplete. By Theorem 8.4.5, the image of the holonomy

$$\eta : \pi_1(L) \to \mathrm{S}_0(\mathbb{C})$$

contains an element ϕ that is not an isometry. Then $\phi(z) = \alpha z + \beta$ with $|\alpha| \neq 0, 1$. By composing the developing map $\delta : \tilde{L} \to \mathbb{C}$ with a translation of $\mathbb{C}$, we may assume that $\beta = 0$. Then ϕ fixes 0. As $\pi_1(L)$ is abelian, every element of $\mathrm{Im}(\eta)$ must also fix 0. Thus η maps into the subgroup $\mathrm{S}_0(\mathbb{C})_0$ of orientation preserving similarities of $\mathbb{C}$ that fix 0.

Every element of $\mathrm{S}_0(\mathbb{C})_0$ is of the form $z \mapsto kz$ for some nonzero complex number k. Hence, we may identify $\mathrm{S}_0(\mathbb{C})_0$ with the multiplicative group $\mathbb{C}^*$ of nonzero complex numbers. The exponential map $\exp : \mathbb{C} \to \mathbb{C}^*$ induces an isomorphism from the topological group $\mathbb{C}/2\pi i\mathbb{Z}$ to $\mathbb{C}^*$. Therefore exp induces a complete metric on $\mathbb{C}^*$ so that $\mathbb{C}/2\pi i\mathbb{Z}$ is isometric to $\mathbb{C}^*$ via exp. It is an exercise to show that $\mathbb{C}^*$ is a geometric space with $\mathrm{I}_0(\mathbb{C}^*) = \mathbb{C}^*$.

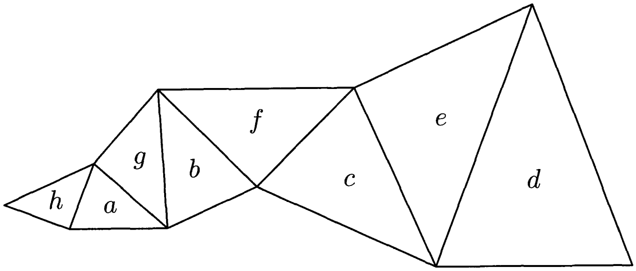


Figure 10.5.9. Triangles $\triangle'_a, \ldots, \triangle'_h$ for $w = \frac{1}{2} + \frac{1}{2}i$

We now show that the developing map $\delta : \tilde{L} \to \mathbb{C}$ maps into $\mathbb{C}^*$. Let $\triangle_i$, for $i = a, \ldots, h$, be the eight triangles in the triangulation of L. Lift these triangles to triangles $\tilde{\triangle}_i$, for $i = a, \ldots, h$, in $\tilde{L}$ that meet as in Figure 10.5.7. Let $\triangle'_i = \delta(\tilde{\triangle}_i)$ for $i = a, \ldots, h$. See Figure 10.5.9. Since $\tilde{L}$ is the union of the images of the triangles $\tilde{\triangle}_i$ under the covering transformations of the universal covering $\kappa : \tilde{L} \to L$, we have that $\delta(\tilde{L})$ is the union of the images of the triangles $\triangle'_i$ under the elements of $\mathrm{Im}(\eta)$. Since $\eta(y)$ does not fix a point in any of the triangles $\triangle'_i$, we see that 0 is not in any of the triangles $\triangle'_i$. As $\mathrm{Im}(\eta)$ leaves $\mathbb{C}^*$ invariant, we deduce that η maps into $\mathbb{C}^*$. Therefore L has the structure of a $(\mathbb{C}^*, \mathbb{C}^*)$-manifold by Theorem 8.4.5.

Now L is a complete $(\mathbb{C}^*, \mathbb{C}^*)$-manifold because L is compact. Hence $\tilde{L}$ is a complete $(\mathbb{C}^*, \mathbb{C}^*)$-manifold. Therefore $\delta : \tilde{L} \to \mathbb{C}^*$ is a universal covering by Theorem 8.5.6. The exponential map $\exp : \mathbb{C} \to \mathbb{C}^*$ is a universal covering of geometric spaces. We shall identify the group $\mathrm{T}(\mathbb{C})$ of translations of $\mathbb{C}$ with $\mathbb{C}$. Then the complete $(\mathbb{C}^*, \mathbb{C}^*)$-structure of L lifts to a complete $(\mathbb{C}, \mathbb{C})$-structure for L. Let $\tilde{\delta} : \tilde{L} \to \mathbb{C}$ be a lift of $\delta : \tilde{L} \to \mathbb{C}^*$ with respect to exp. Then $\tilde{\delta}$ is the developing map for L as a $(\mathbb{C}, \mathbb{C})$-manifold. Let $\tilde{\eta} : \pi_1(L) \to \mathbb{C}$ be the holonomy determined by $\tilde{\delta}$. Then $\eta = \exp\tilde{\eta}$ is the holonomy determined by δ.

Theorem 10.5.5. *The group* $\mathrm{Im}(\eta)$ *is a discrete subgroup of* $\mathbb{C}^*$ *and the map* $\delta : \tilde{L} \to \mathbb{C}^*$ *induces a* $(\mathbb{C}^*, \mathbb{C}^*)$*-equivalence from* L *to* $\mathbb{C}^*/\mathrm{Im}(\eta)$ *if and only if* $2\pi i$ *is in* $\mathrm{Im}(\tilde{\eta})$.

Proof: Since L is a complete $(\mathbb{C}, \mathbb{C})$-manifold, $\mathrm{Im}(\tilde{\eta})$ is a discrete subgroup of $\mathbb{C}$, and $\tilde{\delta} : \tilde{L} \to \mathbb{C}$ induces a $(\mathbb{C}, \mathbb{C})$-equivalence from L to $\mathbb{C}/\mathrm{Im}(\tilde{\eta})$ by Theorem 8.5.9. Observe that we have a commutative diagram of epimorphisms

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{\exp} & \mathbb{C}^* \\ \downarrow & & \downarrow \\ \mathbb{C}/\mathrm{Im}(\tilde{\eta}) & \xrightarrow{\overline{\exp}} & \mathbb{C}^*/\mathrm{Im}(\eta). \end{array}$$

Suppose that $\mathrm{Im}(\eta)$ is a discrete subgroup of $\mathbb{C}^*$ and δ induces a $(\mathbb{C}^*, \mathbb{C}^*)$-equivalence from L to $\mathbb{C}^*/\mathrm{Im}(\eta)$. Then $\overline{\exp}$ is an isomorphism. Now as $\exp(2\pi i) = 1$, we have that $2\pi i$ is in $\mathrm{Im}(\tilde{\eta})$.

Conversely, suppose that $2\pi i$ is in $\mathrm{Im}(\tilde{\eta})$. As $\eta = \exp \tilde{\eta}$, the kernel of η is nontrivial. Hence $\mathrm{Im}(\eta)$ is the direct sum of a finite cyclic group and an infinite cyclic group. Therefore, the infinite cyclic group generated by ϕ is of finite index in $\mathrm{Im}(\eta)$. As $\langle \phi \rangle$ is discrete, $\mathrm{Im}(\eta)$ is discrete. Since $2\pi i$ is in $\mathrm{Im}(\tilde{\eta})$, the map $\overline{\exp}$ is an isomorphism. As $\mathbb{C}/\mathrm{Im}(\tilde{\eta})$ is compact and $\mathbb{C}^*/\mathrm{Im}(\eta)$ is Hausdorff, $\overline{\exp}$ is a homeomorphism. Consequently δ induces a $(\mathbb{C}^*, \mathbb{C}^*)$-equivalence from L to $\mathbb{C}^*/\mathrm{Im}(\eta)$. □

Suppose that $\mathrm{Im}(\eta)$ is a discrete subgroup of $\mathbb{C}^*$ and $\delta : \tilde{L} \to \mathbb{C}^*$ induces a $(\mathbb{C}^*, \mathbb{C}^*)$-equivalence from L to $\mathbb{C}^*/\mathrm{Im}(\eta)$. Then $\tilde{\delta}^{-1} : \mathbb{C} \to \tilde{L}$ induces a covering projection from $\mathbb{C}^*$ to L that is a $(\mathbb{C}^*, \mathbb{C}^*)$-map. Consequently, the triangulation of L lifts to a triangulation of $\mathbb{C}^*$ by Euclidean triangles. Thus, the triangulation of L develops into an exact tessellation of $\mathbb{C}^*$ by Euclidean triangles. Figure 10.5.10 illustrates the exact tessellation of $\mathbb{C}^*$ when $\tilde{\eta}(y) = 2\pi i/10$.

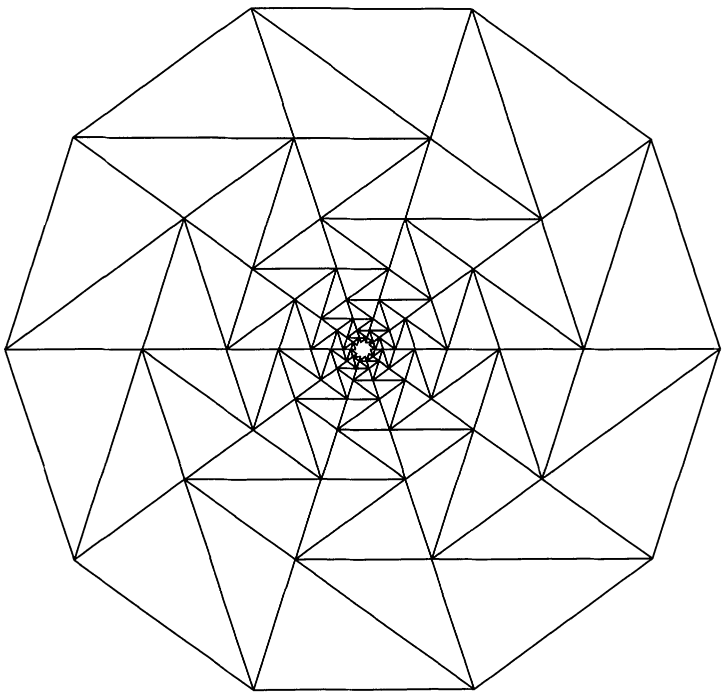


Figure 10.5.10. A tessellation of $\mathbb{C}^*$ by Euclidean triangles

Metric Completion

We now determine when the metric completion $\overline{M}$ of M is a hyperbolic 3-manifold. We shall identify the triangle $\triangle_i$ in L with a triangle in M that represents it, for each $i = a, \ldots, h$, such that these eight triangles in M meet as in Figure 10.5.9. Then we may identify the triangle $\tilde{\triangle}_i$ of $\tilde{L}$ with a triangle in the universal covering space $\tilde{M}$ that projects to the triangle $\triangle_i$ in M, for each $i = a, \ldots, h$, such that these eight triangles in $\tilde{M}$ meet as in Figure 10.5.9.

Regard $\mathbb{C}$ as the boundary of U^3 in $\mathbb{R}^3$. Let $\hat{\delta} : \tilde{M} \to U^3$ be the developing map for M that maps the triangle $\tilde{\triangle}_a$ onto a horizontal triangle directly above $\triangle'_a$. Let $\hat{\triangle}'_i = \hat{\delta}(\tilde{\triangle}_i)$ for $i = a, \ldots, h$. Then the triangles $\hat{\triangle}'_i$ lie on a horizontal horosphere of U^3 with $\hat{\triangle}'_i$ directly above $\triangle'_i$ for each i. Let $\hat{\eta} : \pi_1(M) \to \mathrm{I}_0(U^3)$ by the holonomy determined by $\hat{\delta}$. Then we have a commutative diagram

$$\begin{array}{ccc}
\pi_1(L) & \xrightarrow{\eta} & \mathrm{I}_0(\mathbb{C}^*) \\
i \downarrow & & \downarrow j \\
\pi_1(M) & \xrightarrow{\hat{\eta}} & \mathrm{I}_0(U^3),
\end{array}$$

where i and j are the injections induced by inclusion and Poincaré extension, respectively.

Let T_i be the tetrahedron in M corresponding to T or T' that contains the triangle $\triangle_i$ for $i = a, \ldots, h$. Then T_i lifts to a tetrahedron $\tilde{T}_i$ in $\tilde{M}$ containing $\tilde{\triangle}_i$. Let $T'_i = \hat{\delta}(\tilde{T}_i)$ for $i = a, \ldots, h$. Then T'_i is the ideal tetrahedron in U^3, with one vertex at ∞, directly above the triangle $\triangle'_i$.

Let C be a solid cone in U^3 centered about the 3rd axis, with its vertex at 0, such that the triangle $\hat{\triangle}'_i$ is outside of C for each $i = a, \ldots, h$. Then T'_i intersects ∂C in a triangle τ'_i directly above $\hat{\triangle}'_i$. See Figure 10.5.11. Let τ_i be the triangle in T_i corresponding to τ'_i. Since τ'_i is above $\hat{\triangle}'_i$, for $i = a, \ldots, h$, the triangles $\tau_a, \ldots, \tau_h$ meet only along their boundaries in M. Furthermore, since the image of $j\eta$ leaves ∂C invariant, the triangles $\tau_a, \ldots, \tau_h$ fit together to form a pseudo-triangulation of a torus S in M.

The torus S is the boundary of a closed neighborhood N in M of the cusp point of M. Let $\tilde{\tau}_i$ be the triangle in $\tilde{T}_i$ corresponding to τ_i for $i = a, \ldots, h$, and let $\tilde{N}$ be the component of the subspace of $\tilde{M}$ over N that contains $\tilde{\tau}_i$ for each i. As N deformation retracts onto S and $\pi_1(S)$ injects into $\pi_1(M)$, we have that $\pi_1(N)$ injects into $\pi_1(M)$. Hence $\tilde{N}$ is a universal covering space of N.

Let C_0 be C minus the 3rd axis. As the developing map $\delta : \tilde{L} \to \mathbb{C}^*$ is surjective, $\mathbb{C}^*$ is covered by the triangles $\tilde{\triangle}'_i$, for $i = a, \ldots, h$, and their images by elements of the image of the holonomy $\eta : \pi_1(L) \to \mathbb{C}^*$. Hence C_0 is covered by the tetrahedra T'_i, for $i = a, \ldots, h$, and their images by the elements of $j(\mathrm{Im}(\eta))$. Consequently $\hat{\delta}(\tilde{N}) = C_0$.

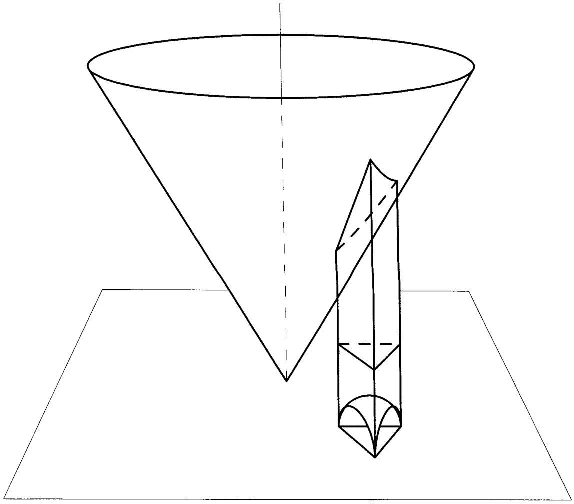


Figure 10.5.11. The triangles τ_i' (on the cone), $\hat{\triangle}_i'$, and $\triangle_i'$

Let U_0^3 be U^3 minus the 3rd axis. Then the universal covering

$$\exp : \mathbb{C} \to \mathbb{C}^*$$

extends to a universal covering

$$\widehat{\exp} : U^3 \to U_0^3.$$

The hyperbolic metric induced on U^3 by $\widehat{\exp}$ is not the Poincaré metric, so we shall denote U^3, with the induced metric, by $\tilde{U}_0^3$. Let $\tilde{C}_0$ be the subspace of $\tilde{U}_0^3$ over C_0. Then $\tilde{C}_0$ is a universal covering space of C_0.

Now since the developing map $\delta : \tilde{L} \to \mathbb{C}^*$ lifts to a homeomorphism $\tilde{\delta} : \tilde{L} \to \mathbb{C}$, the developing map $\hat{\delta} : \tilde{N} \to C_0$ lifts to a homeomorphism $\tilde{\hat{\delta}} : \tilde{N} \to \tilde{C}_0$. Let

$$\tilde{\jmath} : \mathrm{T}(\mathbb{C}) \to \mathrm{I}_0(\tilde{U}_0^3)$$

be the injection obtained by lifting $\jmath : \mathrm{I}_0(\mathbb{C}^*) \to \mathrm{I}_0(U^3)$. Since $\tilde{\delta} : \tilde{L} \to \mathbb{C}$ induces a $(\mathbb{C}, \mathbb{C})$-equivalence from L to $\mathbb{C}/\mathrm{Im}(\tilde{\eta})$, we conclude that the map $\tilde{\hat{\delta}} : \tilde{N} \to \tilde{C}_0$ induces an isometry from N to $\tilde{C}_0/\tilde{\jmath}(\mathrm{Im}(\tilde{\eta}))$.

Theorem 10.5.6. *Let M be an incomplete hyperbolic 3-manifold obtained by properly gluing together two ideal tetrahedrons according to the gluing pattern for the figure-eight knot complement. Then the metric completion $\overline{M}$ is a hyperbolic 3-manifold if and only if the holonomy $\tilde{\eta} : \pi_1(L) \to \mathbb{C}$ for the link L of the cusp point of M has the property that*

$$\mathrm{Im}(\tilde{\eta}) \cap i\,\mathbb{R} = 2\pi i\,\mathbb{Z}.$$

Proof: Suppose that
$$\mathrm{Im}(\tilde{\eta}) \cap i\,\mathbb{R} = 2\pi i\,\mathbb{Z}.$$

Let $\Gamma = j(\mathrm{Im}(\eta))$ and $\tilde{\Gamma} = \tilde{j}(\mathrm{Im}(\tilde{\eta}))$. As $\eta = \exp\tilde{\eta}$, the projection of $\tilde{C}_0$ onto C_0 induces an isometry from $\tilde{C}_0/\tilde{\Gamma}$ to C_0/Γ. Hence N is isometric to C_0/Γ. The metric completion of C_0 is C, since C is the closure of C_0 in the complete metric space U^3. The group Γ is generated by a hyperbolic transformation of U^3 whose axis is the core of C. Therefore Γ acts discontinuously on C. Hence C/Γ is a metric space homeomorphic to a solid torus. As C/Γ is compact, C/Γ is complete. Hence C/Γ is the metric completion of C_0/Γ, since C/Γ is the closure of C_0/Γ in C/Γ. Thus, the metric completion $\overline{N}$ of N is isometric to C/Γ.

Now observe that the hyperbolic structure of the interior of C_0/Γ extends to a hyperbolic structure on the interior of C/Γ. Hence, the hyperbolic structure of N° extends to a hyperbolic structure on $\overline{N}^\circ$. As $M - N^\circ$ is compact, the metric completion of M is $(M-N)\cup\overline{N}$, which is a hyperbolic 3-manifold.

Conversely, suppose that $\overline{M}$ is a hyperbolic 3-manifold. Let $\overline{\delta} : \widetilde{\overline{M}} \to U^3$ be the developing map for $\overline{M}$ that is consistent with the developing map $\delta : \tilde{M} \to U^3$ for M. Let $\overline{\eta} : \pi_1(\overline{M}) \to \mathrm{I}(U^3)$ be the holonomy determined by $\overline{\delta}$. Then we have a commutative diagram

$$\begin{array}{ccc}
\pi(L) & \stackrel{\eta}{\longrightarrow} & \mathrm{I}_0(\mathbb{C}^*) \\
i\downarrow & & \downarrow j \\
\pi_1(M) & \stackrel{\hat{\eta}}{\longrightarrow} & \mathrm{I}_0(U^3) \\
\downarrow & & \downarrow \\
\pi_1(\overline{M}) & \stackrel{\overline{\eta}}{\longrightarrow} & \mathrm{I}(U^3).
\end{array}$$

By Theorem 8.5.9, we have that $\mathrm{Im}(\overline{\eta})$ is a discrete torsion-free subgroup of $\mathrm{I}(U^3)$. Therefore $\Gamma = j(\mathrm{Im}(\eta))$ is a discrete torsion-free subgroup of $\mathrm{I}_0(U^3)$. As Γ fixes 0 and ∞, the group Γ is elementary of hyperbolic type. By Theorem 5.5.8, the group Γ contains an infinite cyclic subgroup of finite index generated by a hyperbolic transformation. Since Γ is torsion-free, Γ is an infinite cyclic group generated by a hyperbolic transformation. As $\eta = \exp\tilde{\eta}$, the image of $\tilde{\eta}$ is generated by an element in the kernel of exp and some other element not in $i\,\mathbb{R}$. Hence, there is a positive integer m such that
$$\mathrm{Im}(\tilde{\eta}) \cap i\,\mathbb{R} = m2\pi i\,\mathbb{Z}.$$

By Theorem 8.5.9, the map $\overline{\delta} : \widetilde{\overline{M}} \to U^3$ induces an isometry from $\overline{M}$ to $U^3/\mathrm{Im}(\overline{\eta})$. Consequently $\overline{\delta}$ induces an isometry from S to $\partial C/\Gamma$. This implies that $\delta : \tilde{L} \to \mathbb{C}^*$ induces a $(\mathbb{C}^*, \mathbb{C}^*)$-equivalence from L to $\mathbb{C}^*/\Gamma$. By Theorem 10.5.5, we have that $2\pi i$ is in $\mathrm{Im}(\tilde{\eta})$. Therefore $m = 1$. □

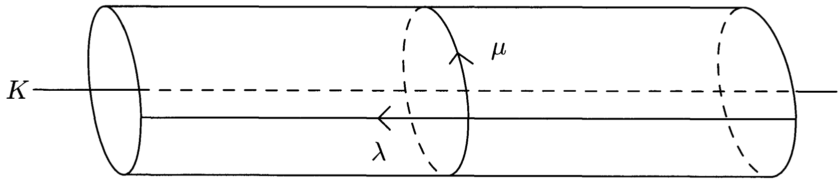


Figure 10.5.12. A meridian-longitude pair μ, λ for a knot K

The Dehn Surgery Invariant

Let K be a smooth knot in $\hat{E}^3$. A *meridian* of K is a simple closed curve μ on the surface of a tubular neighborhood N of K in $\hat{E}^3$ that bounds a disk in N. A meridian μ of K is unique up to isotopy; and so the element m of $\pi_1(\partial N)$ representing μ is unique up to sign. A *longitude* of K is an essential simple closed curve λ on ∂N that meets a meridian μ of K at only one point and is null homologous in $\hat{E}^3 - K$. A longitude λ of K is unique up to isotopy; and so the element ℓ of $\pi_1(\partial N)$ representing λ is unique up to sign. A meridian μ and longitude λ of K that meet at only one point are called a *meridian-longitude pair* of K and, by convention, are oriented by the right-hand rule with your thumb pointing in the direction of λ. See Figure 10.5.12. Finally, the pair m, ℓ generates $\pi_1(\partial N)$.

We now determine a meridian-longitude pair for the figure-eight knot K. From Figure 10.3.7, we see that the curve α in Figure 10.3.4 represents a meridian of K. Figure 10.5.13 illustrates α as it would appear in Figure 10.3.6. Let L be the link of the cusp point of M and assume first that L is complete. Starting on α, we follow a longitude on L, slightly above K in Figure 10.3.6, down through side A. The path of sides and regions encountered in Figure 10.3.6 is

$$AN'DNBN'ANCN'BNDN'CNA.$$

Hence, the longitude crosses the curves in Figure 10.3.4 in the order

$$\alpha, \epsilon, \delta, \kappa, \lambda, \eta, \iota, \gamma.$$

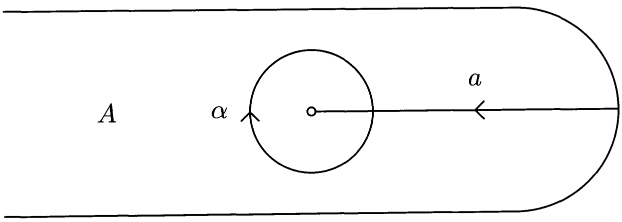


Figure 10.5.13. The meridian α of the figure-eight knot

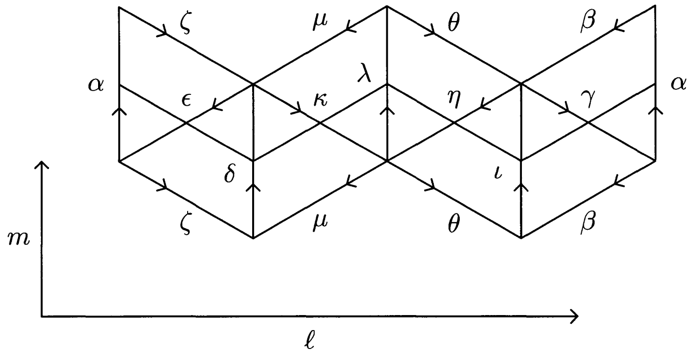


Figure 10.5.14. A representation of a meridian-longitude pair on L

Thus, the central zigzag path in Figure 10.5.14 represents a longitude for K. From Figures 10.5.7 and 10.5.14, we deduce that the meridian and longitude of K are represented by $m = y$ and $\ell = x + 2y$ in $\pi_1(L)$. From Formulas 10.5.8 and 10.5.9, we have

$$\eta(m) = z(1-w), \tag{10.5.10}$$
$$\eta(\ell) = z^2(1-z)^2. \tag{10.5.11}$$

Now assume that M is incomplete. The holonomy $\eta : \pi_1(L) \to \mathbb{C}^*$ lifts to a homomorphism $\tilde{\eta} : \pi_1(L) \to \mathbb{C}$ such that $\tilde{\eta}$ maps $\pi_1(L)$ isomorphically onto a lattice subgroup of $\mathbb{C}$. Therefore $\tilde{\eta}(m)$ and $\tilde{\eta}(\ell)$ form a basis for the real vector space $\mathbb{C}$. From Formulas 10.5.10 and 10.5.11, we have

$$\tilde{\eta}(m) = \log|z(1-w)| + i\arg(z(1-w)), \tag{10.5.12}$$
$$\tilde{\eta}(\ell) = 2\log|z(1-z)| + 2i\arg(z(1-z)). \tag{10.5.13}$$

Now $\arg(z(1-w))$ and $\arg(z(1-z))$ are continuous functions of w that approach 0 as $w \to \frac{1}{2} + \frac{\sqrt{3}}{2}i$. Hence, we have

$$\begin{aligned}\arg(z(1-w)) &= \arg(z) + \arg(1-w),\\ \arg(z(1-z)) &= \arg(z) + \arg(1-z),\end{aligned}$$

with

$$\begin{aligned}&0 < \arg(z) < \pi,\\ &-\pi < \arg(1-w) < 0,\\ &-\pi < \arg(1-z) < 0.\end{aligned}$$

Thus, we have

$$\begin{aligned}&-\pi < \arg(z(1-w)) < \pi,\\ &-\pi < \arg(z(1-z)) < \pi.\end{aligned}$$

Definition: If M is incomplete, the *Dehn surgery invariant* of M is the pair of real numbers (a, b) such that

$$a\tilde{\eta}(m) + b\tilde{\eta}(\ell) = 2\pi i. \tag{10.5.14}$$

If M is complete, the *Dehn surgery invariant* of M is ∞.

Let W be the solution space for w in Figure 10.5.6. Then the Dehn surgery invariant determines a map

$$d : W \to \hat{E}^2$$

such that $d(\frac{1}{2} + \frac{\sqrt{3}}{2}i) = \infty$. If $w \neq \frac{1}{2} + \frac{\sqrt{3}}{2}i$, then

$$d(w) = (a(w), b(w)), \tag{10.5.15}$$

where a and b satisfy the system of equations

$$a\log|z(1-w)| + 2b\log|z(1-z)| = 0, \tag{10.5.16}$$
$$a\arg(z(1-w)) + 2b\arg(z(1-z)) = 2\pi. \tag{10.5.17}$$

Theorem 10.5.7. *The Dehn surgery invariant map d is continuous.*

Proof: Let W_0 be W minus the point $\frac{1}{2} + \frac{\sqrt{3}}{2}i$. By Cramer's rule, a and b, satisfying Equations 10.5.16 and 10.5.17, are continuous functions of w on the set W_0. As both $\arg(z(1-w))$ and $\arg(z(1-z))$ approach 0 as $w \to \frac{1}{2} + \frac{\sqrt{3}}{2}i$, we deduce from Equation 10.5.17 that $(a(w), b(w)) \to \infty$ as $w \to \frac{1}{2} + \frac{\sqrt{3}}{2}i$. Hence d is continuous at the point $\frac{1}{2} + \frac{\sqrt{3}}{2}i$. □

Theorem 10.5.8. *Let M be an incomplete hyperbolic 3-manifold obtained by properly gluing together two ideal tetrahedrons according to the gluing pattern for the figure-eight knot complement. Then the metric completion $\overline{M}$ is a hyperbolic 3-manifold if and only if the Dehn surgery invariant of M is a pair (p, q) of coprime integers.*

Proof: By Theorem 10.5.6, the metric completion $\overline{M}$ is a hyperbolic 3-manifold if and only if

$$\mathrm{Im}(\tilde{\eta}) \cap i\,\mathbb{R} = 2\pi i\,\mathbb{Z}.$$

Now $\mathrm{Im}(\tilde{\eta}) \cap i\,\mathbb{R}$ is a subgroup of $\mathrm{Im}(\tilde{\eta})$ and therefore is a free abelian group of rank 0, 1, or 2. The last case is impossible since $\mathrm{Im}(\tilde{\eta}) \cap i\,\mathbb{R}$ would then be of finite index in $\mathrm{Im}(\tilde{\eta})$, and every subgroup of finite index of $\mathrm{Im}(\tilde{\eta})$ is generated by two linearly independent vectors of the real vector space $\mathbb{C}$. Hence $\mathrm{Im}(\tilde{\eta}) \cap i\,\mathbb{R}$ is a cyclic group. As $\mathrm{Im}(\tilde{\eta})$ is generated by $\tilde{\eta}(m)$ and $\tilde{\eta}(\ell)$, we have that

$$\mathrm{Im}(\tilde{\eta}) \cap i\,\mathbb{R} = 2\pi i\,\mathbb{Z}$$

if and only if there are coprime integers p, q such that

$$p\tilde{\eta}(m) + q\tilde{\eta}(\ell) = 2\pi i.$$

□

Dehn Surgery

Let N be a closed tubular neighborhood of the figure-eight knot K in E^3. Let p, q be coprime integers and let $M_{(p,q)}$ be the closed orientable 3-manifold obtained by gluing a solid torus V to $\hat{E}^3 - N^\circ$ along their boundaries by a homeomorphism that maps a meridian of V onto a simple closed curve in ∂N representing $m^p\ell^q$ in $\pi_1(\partial N)$. The 3-manifold $M_{(p,q)}$ is said to be obtained from $\hat{E}^3$ by (p,q)-*Dehn surgery* on K.

Theorem 10.5.9. *Let M be an incomplete hyperbolic 3-manifold, obtained by properly gluing together two ideal tetrahedrons according to the gluing pattern for the figure-eight knot K, whose Dehn surgery invariant is a pair (p,q) of coprime integers. Then the metric completion $\overline{M}$ is a hyperbolic 3-manifold homeomorphic to the 3-manifold $M_{(p,q)}$ obtained from $\hat{E}^3$ by (p,q)-Dehn surgery on K.*

Proof: By Theorem 10.5.8, the metric completion $\overline{M}$ is a hyperbolic 3-manifold. From the proof of Theorem 10.5.6, we have

$$\overline{M} = (M - N^\circ) \cup \overline{N},$$

where $\overline{N}$ is a solid torus isometric to C/Γ. The group $\Gamma = \jmath(\mathrm{Im}(\eta))$ is generated by a hyperbolic transformation $z \mapsto kz$, where $|k| > 1$. Let F be the frustrum in U^3 bounded by ∂C and the horospheres $x_3 = 1, |k|$. See Figure 10.5.15. Then F° is a fundamental domain for Γ in C, and $V = F/\Gamma$ is a solid torus that is glued to $M - N^\circ$ to give $\overline{M}$. Now $M - N^\circ$ is homeomorphic to the complement in $\hat{E}^3$ of a open tubular neighborhood of K. Therefore $\overline{M}$ is homeomorphic to a 3-manifold obtained from $\hat{E}^3$ by Dehn surgery on K. Observe that the bottom rim ρ of F in Figure 10.5.15 represents a meridian of V, and ρ corresponds to a rotation by 2π in Γ. As the Dehn surgery invariant of M is (p,q), the curve ρ represents the element $m^p\ell^q$ of $\pi_1(\partial N)$. Thus $\overline{M}$ is homeomorphic to $M_{(p,q)}$. □

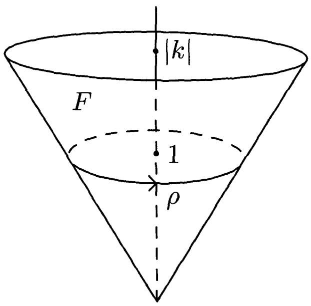


Figure 10.5.15. The frustrum F within the cone C

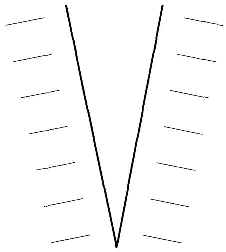


Figure 10.5.16. The compactification of W along the missing ray

Let $\hat{W}$ be the compactification of the solution space W obtained by adjoining to W the real axis, a copy of $\mathbb{R}$ along the ray

$$R = \{\tfrac{1}{2} + \tfrac{t}{2}i : t \geq \sqrt{15}\}$$

as indicated in Figure 10.5.16, and two more points $\pm\infty$, with $-\infty$ joining the left ends of the new lines together and $+\infty$ joining the right ends of the new lines together. Note that $\hat{W}$ is topologically a disk whose interior is W.

Let σ be the involution of W obtained by interchanging the solutions w and z of Equation 10.5.5,

$$z(z-1)w(w-1) = 1.$$

Then we deduce from Formulas 10.5.10 and 10.5.11 that

$$\begin{array}{rcccl} \sigma\eta(m) & = & \eta(m)^{-1} & = & \eta(-m), \\ \sigma\eta(\ell) & = & \eta(\ell)^{-1} & = & \eta(-\ell). \end{array}$$

Therefore, we deduce from Formula 10.5.14 that $d\sigma = -d$.

Lemma 3. *The involution σ of W extends to a continuous involution $\hat{\sigma}$ of $\hat{W}$.*

Proof: The function $\sigma : W \to W$ is defined by the formula

$$\sigma(w) = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{1}{w(w-1)}}.$$

Hence σ is analytic and therefore σ is continuous.

When w is near the interval $(-\infty, 0)$, we find that z is near the interval $(1, \infty)$. Hence σ extends continuously to $(-\infty, 0)$ by the formula

$$\hat{\sigma}(w) = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{w(w-1)}}.$$

We define $\hat{\sigma}(0) = +\infty$.

When w is near the interval $(0, 1/2]$, we find that z is near the right side of the ray R. Hence σ extends continuously to $(0, 1/2]$ by the formula

$$\hat{\sigma}(w) = \frac{1}{2} + i\sqrt{-\left(\frac{1}{4} + \frac{1}{w(w-1)}\right)},$$

where $\hat{\sigma}(w)$ is understood to lie in the right copy R_+ of the ray R.

When w is near the interval $[1/2, 1)$, we find that z is near the left side of R. Hence σ extends continuously to $[1/2, 1)$ by the formula

$$\hat{\sigma}(w) = \frac{1}{2} + i\sqrt{-\left(\frac{1}{4} + \frac{1}{w(w-1)}\right)},$$

where $\hat{\sigma}(w)$ is understood to lie in the left copy R_- of R. We define $\hat{\sigma}(1) = -\infty$.

When w is near the interval $(1, \infty)$, we find that z is near the interval $(-\infty, 0)$. Hence σ extends continuously to $(1, \infty)$ by the formula

$$\hat{\sigma}(w) = \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{w(w-1)}}.$$

We define $\hat{\sigma}(+\infty) = 0$.

When w is near the right side of R, we find that z is near the interval $(0, 1/2]$. Hence σ extends continuously to R_+ by the formula

$$\hat{\sigma}(w) = \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{w(w-1)}}.$$

When w is near the left side of R, we find that z is near the interval $[1/2, 1)$. Hence σ extends continuously to R_- by the formula

$$\hat{\sigma}(w) = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{w(w-1)}}.$$

Finally, we define $\hat{\sigma}(-\infty) = 1$. Then $\hat{\sigma}$ is a continuous involution of $\hat{W}$. □

Let τ be the involution of W defined by

$$\tau(w) = \overline{1 - w}.$$

Then $\tau(z) = \overline{1 - z}$, and we deduce from Formulas 10.5.10 and 10.5.11 that

$$\begin{aligned} \tau\eta(m) &= \overline{\eta(m)}^{-1}, \\ \tau\eta(\ell) &= \overline{\eta(\ell)}. \end{aligned}$$

Therefore, we deduce from Formulas 10.5.12-10.5.14 that

$$d\tau(w) = (a\tau(w), b\tau(w)) = (a(w), -b(w)).$$

Let $\rho : \hat{E}^2 \to \hat{E}^2$ be the reflection in the x-axis. Then $d\tau = \rho d$. Clearly τ extends to a continuous involution $\hat{\tau}$ of $\hat{W}$.

Lemma 4. *The Dehn surgery invariant map $d : W \to \hat{E}^2$ extends to a continuous function $\hat{d} : \hat{W} \to \hat{E}^2$.*

Proof: We begin by extending d to the open interval $(1, \infty)$. When w is near $(1, \infty)$, then z is near the interval $(-\infty, 0)$. Thus, for w in $(1, \infty)$, we define

$$z = \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{w(w-1)}},$$

$$\arg(w) = 0,\ \arg(1-w) = -\pi,\ \arg(z) = \pi,\ \arg(1-z) = 0.$$

Then $\arg(z(1-w)) = 0$ and $\arg(z(1-z)) = \pi$. From Equation 10.5.17, we find that $b(w) = 1$, and so from Equation 10.5.16, we have

$$a(w) = \frac{-2\log|z(1-z)|}{\log|z(1-w)|)}.$$

From Equation 10.5.5, we have

$$a(w) = \frac{-2\log(w(w-1))}{\log(w(1-z))}.$$

Define $\hat{d}$ on the interval $(1, \infty)$ by

$$\hat{d}(w) = (a(w), 1).$$

Then $\hat{d}$ is continuous on the set $W \cup (1, \infty)$.

Next, observe that

$$\begin{aligned} \frac{\log(w(w-1))}{\log(w(1-z))} &= \frac{\log(w) + \log(w-1)}{\log(w) + \log(1-z)} \\ &= \frac{1 + \frac{\log(w-1)}{\log(w)}}{1 + \frac{\log(1-z)}{\log(w)}} < 2 \end{aligned}$$

and that

$$\lim_{w\to\infty} \frac{\log(w(w-1))}{\log(w(1-z))} = 2.$$

Hence $a(w) > -4$ and $\lim\limits_{w\to\infty} a(w) = -4$. Now

$$a\big((1+\sqrt{5})/2\big) = 0$$

and $a(w) \leq 0$ for $w \geq (1+\sqrt{5})/2$. By continuity, we deduce that a maps the interval $[(1+\sqrt{5})/2, \infty)$ onto the interval $(-4, 0]$. Observe that

$$\begin{aligned} \hat{d}\big((1, (1+\sqrt{5})/2]\big) &= \hat{d}\hat{\sigma}\big((-\infty, (1-\sqrt{5})/2]\big) \\ &= \hat{d}\hat{\sigma}\hat{\tau}\big([(1+\sqrt{5})/2, \infty)\big) \\ &= -\rho\hat{d}\big([(1+\sqrt{5})/2, \infty)\big). \end{aligned}$$

Therefore a maps the interval $(1, (1+\sqrt{5})/2]$ onto the interval $[0, 4)$.

We now extend d to the right copy R_+ of the ray R. When w is near the right side of R, then z is near the interval $(0, 1/2]$. Thus, for w in R_+, we define

$$z = \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{w(w-1)}},$$

$\arg(z) = 0$, and $\arg(1 - z) = 0$. Then $\arg(z(1 - z)) = 0$ and

$$\arg(z(1 - w)) = \arg(1 - w).$$

From Equation 10.5.17, we find that

$$a(w) = \frac{2\pi}{\arg(1 - w)}.$$

As w varies from $\frac{1}{2} + \frac{\sqrt{15}}{2}i$ to $+\infty$ along R_+, the value of $a(w)$ increases from $-4.76679\ldots$ to -4. From Equation 10.5.16, we find that

$$b(w) = \frac{-a(w)\log|z(1-w)|}{2\log|z(1-z)|}.$$

From Equation 10.5.5, we have

$$\begin{aligned} b(w) &= \frac{-a(w)\log|w(1-z)|}{2\log|w(1-w)|} \\ &= \frac{-a(w)\log|w(1-z)|}{2\log|w\overline{w}|} \\ &= \frac{-a(w)\log|w(1-z)|}{4\log|w|} \\ &= -\frac{a(w)}{4}\left(1 + \frac{\log(1-z)}{\log|w|}\right). \end{aligned}$$

Hence, we have

$$b\left(\frac{1}{2} + \frac{\sqrt{15}}{2}i\right) = 0 \quad \text{and} \quad \lim_{w \to +\infty} b(w) = 1.$$

Define $\hat{d}$ on R_+ by

$$\hat{d}(w) = (a(w), b(w)).$$

Then $\hat{d}$ is continuous on the set $W \cup R_+$.

We next define $\hat{d}(+\infty) = (-4, 1)$ and show that $\hat{d}$ is continuous at $+\infty$. Suppose that w is in W with $|w|$ large and w is to the right of the ray R. Then $|z|$ is small. From the equation

$$|z|\ |z-1|\ |w|\ |w-1| = 1,$$

we deduce that

$$|z|\ |w|^2 \simeq 1.$$

Therefore, we have

$$\log|z| + 2\log|w| \simeq 0.$$

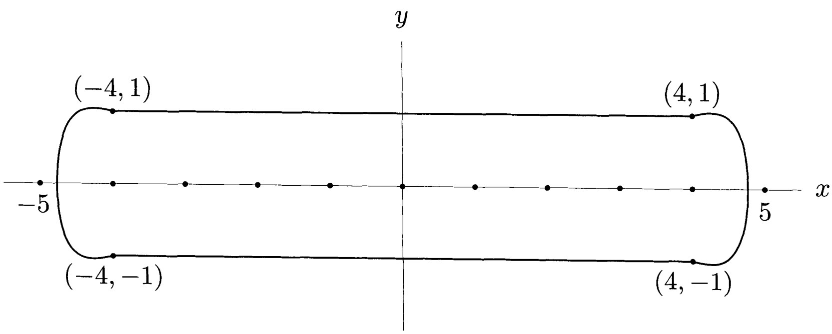


Figure 10.5.17. The image of the boundary of $\hat{W}$

From Equation 10.5.16, we find that $a + 4b \simeq 0$. From the equation

$$\arg(z) + \arg(z-1) + \arg(w) + \arg(w-1) = 2\pi,$$

we deduce that

$$\arg(z) \simeq \pi - 2\arg(w).$$

Therefore, we have

$$\begin{aligned}
&\arg(1-z) \simeq 0,\\
&\arg(1-w) \simeq \arg(w) - \pi,\\
&\arg(z(1-w)) \simeq -\arg(w),\\
&\arg(z(1-z)) \simeq \pi - 2\arg(w).
\end{aligned}$$

From Equation 10.5.17, we find that

$$-(a+4b)\arg(w) + 2b\pi \simeq 2\pi.$$

Therefore $(a, b) \simeq (-4, 1)$ with $(a, b) \to (-4, 1)$ as $w \to +\infty$. Thus $\hat{d}$ is continuous at the point $+\infty$.

Now, by symmetry, $d : W \to \hat{E}^2$ extends to a continuous function

$$\hat{d} : \hat{W} \to \hat{E}^2$$

such that $\hat{d}\hat{\sigma} = -\hat{d}$ and $\hat{d}\hat{\tau} = \rho\hat{d}$. Consequently $\hat{d}(\partial\hat{W})$ is a simple closed curve enclosing the origin that is symmetric with respect to the x and y axes. See Figure 10.5.17. □

Theorem 10.5.10. *Let p, q be coprime integers such that either $|p| > 4$ or $|q| > 1$, and let $M_{(p,q)}$ be the closed orientable 3-manifold obtained from $\hat{E}^3$ by (p,q)-Dehn surgery on the figure-eight knot. Then $M_{(p,q)}$ has a hyperbolic 3-manifold structure.*

Proof: Let C and D be the closed disks in $\hat{E}^2$ bounded by the simple closed curve $\hat{d}(\partial\hat{W})$ with $(0,0)$ in C. See Figure 10.5.17. Let

$$r : \hat{E}^2 - \{(0,0)\} \to D$$

be a retraction that retracts $C - \{(0,0)\}$ onto $\partial C = \partial D$. From Equation 10.5.17, we deduce that $(0,0)$ is not in the image of $\hat{d}$. Hence, the function

$$f : \hat{W} \to D$$

defined by $f = r\hat{d}$ is well defined and continuous.

We now prove that f is onto. On the contrary, suppose that f is not onto. Then f is homotopic to a map $g : \hat{W} \to \partial D$ such that f and g agree on $\partial\hat{W}$. Let $\partial f : \partial\hat{W} \to \partial D$ be the restriction of f. Then we have a commutative diagram of first homology groups and homomorphisms:

$$\begin{tikzcd} H_1(\partial\hat{W}) \arrow[r, "i_*"] \arrow[d, "(\partial f)_*"'] & H_1(\hat{W}) \arrow[dl, "g_*"] \\ H_1(\partial D) & \end{tikzcd}$$

As $H_1(\hat{W}) = 0$, we have that $(\partial f)_*$ is the zero homomorphism; but ∂f is a degree one map, which is a contradiction. Therefore f is onto.

Now since r retracts $C - \{(0,0)\}$ onto ∂D, we deduce that $D \subset \hat{d}(\hat{W})$. Therefore $D^\circ \subset d(W)$. The theorem now follows from Theorem 10.5.9, since (p,q) is in D°. □

Exercise 10.5

1. Prove that every Euclidean triangle in $\mathbb{C}$ is directly similar to a triangle whose vertices are $0, 1, z$, where z satisfies the inequalities $\mathrm{Im}(z) > 0$, $|z| \leq 1$, and $|z-1| \leq 1$.

2. Prove that $\mathbb{C}^*$ is a geometric space with $\mathrm{I}_0(\mathbb{C}^*) = \mathbb{C}^*$, where $\mathbb{C}^*$ acts on itself by multiplication.

3. Let $M_{(p,q)}$ be the hyperbolic 3-manifold obtained by (p,q)-Dehn surgery on the figure-eight knot and let M_∞ be the complete, hyperbolic, figure-eight knot complement. Prove that

$$\begin{aligned} \mathrm{Vol}(M_{(p,q)}) &< \mathrm{Vol}(M_\infty), \\ \lim_{(p,q)\to\infty} \mathrm{Vol}(M_{(p,q)}) &= \mathrm{Vol}(M_\infty). \end{aligned}$$

4. Prove that infinitely many nonisometric, closed, orientable, hyperbolic 3-manifolds can be obtained from the figure-eight knot by hyperbolic Dehn surgery.

5. Prove that the Seifert-Weber dodecahedral space cannot be obtained from the figure-eight knot by hyperbolic Dehn surgery.

§10.6. Historical Notes

§10.1. The concept of gluing together polyhedra to construct a 3-manifold was introduced by Poincaré in his 1895 paper *Analysis situs* [336]. In particular, Example 1 appeared in this paper. The first example of a closed hyperbolic 3-manifold was constructed by Löbell in his 1931 paper *Beispiele geschlossener dreidimensionaler Clifford-Kleinscher Räume negativer Krümmung* [268] by gluing together eight copies of a 14-sided, right-angled, hyperbolic polyhedron. For a description of Löbell's 3-manifold in terms of reflection groups, see Vesnin's 1987 paper *Three-dimensional hyperbolic manifolds of Löbell type* [396]. Examples 2 and 3 were given by Seifert and Weber in their 1933 paper *Die beiden Dodekaederräume* [405]. Moreover, Theorem 10.1.3 appeared in this paper. Other examples of closed hyperbolic 3-manifolds obtained by gluing together polyhedra can be found in Best's 1971 paper *On torsion-free discrete subgroups of* $\mathrm{PSL}(2,\mathbb{C})$ *with compact orbit space* [44], in Gucul's 1979 paper *On a series of compact 3-dimensional manifolds of constant negative curvature* [173], and in Molnár's 1989 paper *Two hyperbolic football manifolds* [303].

§10.2. Necessary and sufficient conditions for the complete gluing of a hyperbolic 3-manifold from a single polyhedron were given by Maskit in his 1971 paper *On Poincaré's theorem for fundamental polygons* [281]. Necessary and sufficient conditions for the complete gluing of a hyperbolic 3-manifold were given by Seifert in his 1975 paper *Komplexe mit Seitenzuordnung* [371]. The concept of the link of a cusp point of a hyperbolic 3-manifold was introduced by Thurston in his 1979 lecture notes *The Geometry and Topology of 3-Manifolds* [389], and all of the results of this section appeared in Thurston's notes.

§10.3. The first example of a complete hyperbolic 3-manifold of finite volume was constructed by Gieseking in his 1912 thesis *Analytische Untersuchungen über topologische Gruppen* [153] by gluing together the sides of a regular ideal tetrahedron. For a discription of the Gieseking manifold, see Adams' 1987 paper *The noncompact hyperbolic 3-manifold of minimal volume* [4]. The Gieseking manifold is nonorientable. Its orientable double cover is the figure-eight knot space. That the figure-eight knot space has a complete hyperbolic structure appeared in Riley's 1975 paper *A quadratic parabolic group* [350]. The construction of the complete hyperbolic structure on the figure-eight knot space by gluing together two regular ideal tetrahedrons was given by Thurston in his 1979 lecture notes [389]. The complements of the Whitehead link and the Borromean rings were first shown to have a complete hyperbolic structure by Riley. See Wielenberg's 1978 paper *The structure of certain subgroups of the Picard group* [412] and Riley's 1979 paper *An elliptical path from parabolic representations to hyperbolic structures* [351]. The construction of the complete hyperbolic structure on the Whitehead link and the Borromean rings by gluing together regular ideal octahedrons was given by Thurston in his 1979 lecture notes [389]. For examples of complete hyperbolic 3-manifolds obtained by gluing together ideal cubes or regular ideal dodecahedra, see Aitchison and Rubinstein's 1990 paper *An introduction to polyhedral metrics of non-positive curvature on 3-manifolds* [10] and their 1992 paper *Combinatorial cubings, cusps, and the dodecahedral knots* [11].

§10.4. Theorems 10.4.1 and 10.4.2 appeared in Thurston's 1979 lecture notes [389]. Clausen investigated the function $f(\phi) = 2Л(\phi/2)$ in his 1832 paper *Ueber die Function* $f(\phi) = \sin\phi + \frac{1}{2^2}\sin 2\phi + \frac{1}{3^2}\sin 3\phi + etc.$ [81]. In particular, Formula 10.4.9 appeared in this paper. Moreover, Theorem 10.4.3 is implicit in Clausen's Fourier series expansion of $f(\phi)$. The Lobachevsky function was originally defined to be minus the integral of $\log\cos\theta$ from 0 to θ by Lobachevsky in his 1836 Russian treatise *Application of imaginary geometry to certain integrals.* For a German translation with commentary, see *N. J. Lobatschefskijs Imaginäre Geometrie und Anwendung der imaginären Geometrie auf einige Integrale* [263]. The present Lobachevsky function was introduced by Milnor in Thurston's 1979 lecture notes [389]. All the results of this section appeared in Thurston's notes and in Milnor's 1982 paper *Hyperbolic geometry: the first 150 years* [290]. Theorem 10.4.5 was essentially proved by Lobachevsky in his 1836 treatise [263]. Theorem 10.4.7 appeared in Coxeter's 1935 paper *The functions of Schläfli and Lobatschefsky* [89] and was proved by Milnor in his 1982 paper [290]. For the computation of the volume of a compact hyperbolic tetrahedron, see Kellerhals' 1989 paper *On the volume of hyperbolic polyhedra* [217] and her 1991 paper *The dilogarithm and volumes of hyperbolic polytopes* [218].

Jørgensen and Thurston have proved that the set of all the volumes of complete hyperbolic 3-manifolds of finite volume is a well-ordered closed subset of the real line with all the volumes of open complete manifolds of finite volume as limit points from the left. In particular, there is a closed hyperbolic 3-manifold of minimum volume. Furthermore, volume is a finite-to-one function of complete hyperbolic 3-manifolds of finite volume. For a discussion, see Thurston's 1979 lecture notes [389] and Gromov's 1981 paper *Hyperbolic manifolds according to Thurston and Jørgensen* [168]. Wielenberg has constructed arbitrarily large finite sets of nonisometric, open, complete, hyperbolic 3-manifolds with the same finite volume in his 1980 paper *Hyperbolic 3-manifolds which share a fundamental polyhedron* [413], and Apanasov and Gutsul have constructed arbitrarily large finite sets of nonisometric, closed, hyperbolic 3-manifolds with the same volume in their 1992 paper *Greatly symmetric totally geodesic surfaces and closed hyperbolic 3-manifolds which share a fundamental polyhedron* [21]. For a lower bound on the volume of an open, complete, hyperbolic 3-manifold of finite volume, see Adams' 1988 paper *Volumes of N-cusped hyperbolic 3-manifolds* [5]. For a positive lower bound for the set of volumes of complete hyperbolic 3-manifolds, see Gehring and Martin's 1991 paper *Inequalities for Möbius transformations and discrete groups* [151]. See also Culler and Shalen's 1992 paper *Paradoxical decompositions, 2-generator Kleinian groups, and volumes of hyperbolic 3-manifolds* [98]. For an algebraic characterization of complete hyperbolic 3-manifolds of finite volume, see my 1987 paper *Euler characteristics of 3-manifold groups and discrete subgroups of* $\mathrm{SL}(2,\mathbb{C})$ [347].

§10.5. The similarity structures on the torus were considered by Kuiper in his 1950 paper *Compact spaces with a local structure determined by the group of similarity transformations in* E^n [247]. See also Fried's 1980 paper *Closed similarity manifolds* [140].

Hyperbolic Dehn surgery was introduced by Thurston in his 1979 lecture notes [389], and all the results of this section appeared in Thurston's notes. According to Thurston [389], he became interested in hyperbolic Dehn surgery because of Jørgensen's 1977 paper *Compact 3-manifolds of constant negative curvature fibering over the circle* [212]. Thurston has proved that most knot and link spaces have a complete hyperbolic structure and almost all Dehn surgeries on a hyperbolic knot or link space yield a hyperbolic 3-manifold. For details, see Thurston's 1979 lecture notes [389], his 1982 article *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry* [390], Morgan's 1984 paper *On Thurston's uniformization theorem for 3-dimensional manifolds* [305], McMullen's 1992 article *Riemann surfaces and the geometrization of 3-manifolds* [285], and Benedetti and Petronio's 1992 text *Lectures on Hyperbolic Geometry* [40].

For an analysis of the volumes of hyperbolic 3-manifolds obtained by Dehn surgery on a hyperbolic knot space, see Neumann and Zagier's 1985 paper *Volumes of hyperbolic 3-manifolds* [314]. For a computation of the volumes of closed, orientable, hyperbolic 3-manifolds of small complexity, see Matveev and Fomenko's 1988 paper *Constant energy surfaces of Hamiltonian systems, enumeration of 3-dimensional manifolds in increasing order of complexity, and computation of volumes of closed hyperbolic manifolds* [284]. Weeks has written a computer program called *SnapPea* that computes invariants of hyperbolic 3-manifolds. For a discussion, see Adams' 1990 review *SnapPea, The Weeks hyperbolic 3-manifold program* [6]. See also Weeks' 1993 paper *Convex hulls and isometries of cusped hyperbolic 3-manifolds* [407]. For a tabulation of hyperbolic knots and links and their invariants, see Adams, Hildebrand, and Weeks' 1991 paper *Hyperbolic invariants of knots and links* [8]. For an analysis of some of the complete hyperbolic 3-manifolds obtained by Dehn surgery on the Whitehead link complement, see Hodgson, Meyerhoff, and Weeks' 1992 paper *Surgeries on the Whitehead link yield geometrically similar manifolds* [195].

CHAPTER 11

Hyperbolic n-Manifolds

In this chapter, we take up the study of hyperbolic n-manifolds. We begin with a geometric method for constructing spherical, Euclidean, and hyperbolic n-manifolds. In Section 11.2, we prove Poincaré's fundamental polyhedron theorem for freely acting groups. In Section 11.3, we determine the simplices of maximum volume in hyperbolic n-space. In Section 11.4, we study the Gromov invariant of a closed, orientable, hyperbolic manifold. In Section 11.5, we study the measure homology of hyperbolic space-forms. In Section 11.6, we prove Mostow's rigidity theorem for closed, orientable, hyperbolic manifolds.

§11.1. Gluing n-Manifolds

In this section, we shall construct n-dimensional spherical, Euclidean, and hyperbolic manifolds by gluing together n-dimensional convex polyhedra. Let $X = S^n, E^n$, or H^n with $n > 0$.

Definition: An n-dimensional, *abstract, convex polyhedron* P in X is an n-dimensional convex polyhedron P in X together with a collection $\mathcal{F}$ of subsets of ∂P, called the *facets* of P, such that

(1) each facet of P is a closed, $(n-1)$-dimensional, convex subset of ∂P;

(2) two facets of P meet only along their boundaries;

(3) the union of the facets of P is ∂P;

(4) the collection $\mathcal{F}$ is locally finite in X.

By Theorem 6.2.6, an n-dimensional convex polyhedron P in X, together with the collection $\mathcal{S}$ of its sides, is an n-dimensional, abstract, convex polyhedron. Note that, in general, a facet of an abstract convex polyhedron

P may or may not be equal to the side of P containing it. It is an exercise to prove that every facet of an n-dimensional, abstract, convex polyhedron is an $(n-1)$-dimensional convex polyhedron.

Definition: A *disjoint set of n-dimensional, abstract, convex polyhedra* of X is a set of functions

$$\Xi = \{\xi_P : P \in \mathcal{P}\}$$

indexed by a set $\mathcal{P}$ such that

(1) the function $\xi_P : X \to X_P$ is a similarity for each P in $\mathcal{P}$;

(2) the index P is an n-dimensional abstract convex polyhedron in X_P for each P in $\mathcal{P}$; and

(3) the polyhedra in $\mathcal{P}$ are mutually disjoint.

Let Ξ be a disjoint set of n-dimensional, abstract, convex polyhedra of X and let G be a group of similarities of X.

Definition: A *G-facet-pairing* for Ξ is a set of functions

$$\Phi = \{\phi_F : F \in \mathcal{F}\}$$

indexed by the collection $\mathcal{F}$ of all the facets of the polyhedra in $\mathcal{P}$ such that for each facet F of a polyhedron P in $\mathcal{P}$,

(1) there is a polyhedron P' in $\mathcal{P}$ such that the function $\phi_F : X_{P'} \to X_P$ is a similarity;

(2) the similarity $g_F = \xi_P^{-1}\phi_F\xi_{P'}$ is in G;

(3) there is a facet F' of P' such that $\phi_F(F') = F$;

(4) the similarities ϕ_F and $\phi_{F'}$ satisfy the relation $\phi_{F'} = \phi_F^{-1}$;

(5) the polyhedrons P and $\phi_F(P')$ are situated so that $P \cap \phi_F(P') = F$.

Let Φ be a G-facet-pairing for Ξ. The pairing of facet points by elements of Φ generates an equivalence relation on the set $\Pi = \cup_{P\in\mathcal{P}} P$ whose equivalence classes are called the *cycles* of Φ. Topologize Π with the direct sum topology and let M be the quotient space of Π of cycles. The space M is said to be obtained by gluing together the polyhedra of Ξ by Φ.

The *normalized solid angle* $\hat{\omega}$ subtended by a polyhedron P in X at a point x of P is defined to be the real number

$$\hat{\omega} = \frac{\mathrm{Vol}(P \cap B(x,r))}{\mathrm{Vol}(B(x,r))},$$

where r is less than the distance from x to any side of P not containing x. It follows from Theorems 2.4.1 and 3.4.1 that $\hat{\omega}$ does not depend on r.

Let $[x] = \{x_1, \ldots, x_m\}$ be a finite cycle of Φ. Let P_i be the polyhedron in $\mathcal{P}$ containing the point x_i and let $\hat{\omega}_i$ be the normalized solid angle subtended by P_i at the point x_i for each $i = 1, \ldots, m$. The *normalized solid angle sum* of $[x]$ is defined to be

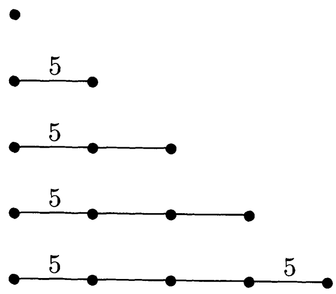
$$\hat{\omega}[x] = \hat{\omega}_1 + \cdots + \hat{\omega}_m.$$

Definition: A G-facet-pairing Φ for Ξ is *proper* if and only if each cycle of Φ is finite and has normalized solid angle sum 1.

The proof of the next theorem is by induction on n and follows the same outline as the proof of Theorem 10.1.2 and it is therefore left to the reader.

Theorem 11.1.1. *Let G be a group of similarities of X and let M be a space obtained by gluing together a disjoint set Ξ of n-dimensional, abstract, convex polyhedra of X by a proper G-facet-pairing Φ. Then M is an n-manifold with an (X, G)-structure such that the natural injection of P° into M is an (X, G)-map for each polyhedron P of Ξ.*

Example 1. We now consider an example of a closed hyperbolic 4-manifold obtained by gluing together the sides of a 4-dimensional, regular, convex polyhedron in H^4. For $n = 0, 1, 2, 3, 4$, let Γ_n be the discrete, n-simplex, reflection group whose Coxeter graph is, respectively,



For $n = 1, 2, 3$, the group Γ_n is a discrete group of isometries of S^n generated by the reflections of S^n in the sides of a spherical n-simplex Δ^n. The group Γ_4 is a discrete group of isometries of H^4 generated by the reflections of H^4 in the sides of a hyperbolic 4-simplex Δ^4. For $n = 1, 2, 3, 4$, let v_n be a vertex of Δ^n such that the subgroup of Γ_n fixing v_n is Γ_{n-1}. Then the images of Δ^n under Γ_{n-1} fit together at v_n to give the barycentric subdivision of a regular convex polyhedron P^n in S^n, if $n = 1, 2, 3$, or in H^4 if $n = 4$. The images of P^n under Γ_n form an exact tessellation of S^n, if $n = 1, 2, 3$, or of H^4 if $n = 4$, by congruent copies of P^n. The group of symmetries of this tessellation is Γ_n. The order of Γ_n, for $n = 0, 1, 2, 3, 4$, is $2, 10, 120, 14400, \infty$, respectively.

For $n = 1, 2, 3$, the convex hull of the set of vertices of this tessellation of S^n is a regular Euclidean convex polyhedron Q^{n+1} which is combinatorially equivalent to P^{n+1}. The set P^1 is an arc of twice the length of Δ^1 and so S^1 is tessellated by $10/2 = 5$ copies of it. Hence Q^2 is a regular pentagon. Therefore P^2 is a regular spherical pentagon and S^2 is tessellated by $120/10 = 12$ copies of it. Hence Q^3 is a regular dodecahedron. Therefore P^3 is a regular spherical dodecahedron and S^3 is tessellated by $14400/120 = 120$ copies of it. The 4-dimensional regular polyhedron Q^4 is called the 120-cell. Therefore P^4 is a regular hyperbolic 120-cell.

The polyhedron Q^4 has 120 sides, 720 ridges, 1200 edges, and 600 vertices. Each side of Q^4 is a regular dodecahedron and is parallel to its opposite side, $-S$. For each side S of P^4, let f_S be the reflection of H^4 that pairs S to its opposite side S' and let g_S be the composite of f_S followed by the reflection in the side S. Then $\{g_S\}$ is an $\mathrm{I}_0(H^4)$-side-pairing for P^4. We shall call $\Phi = \{g_S\}$ the *opposite side-pairing* of P^4.

Using known coordinates for the vertices of Q^4, one can check that each ridge cycle contains 5 points, each edge cycle contains 20 points, and all the vertices of P^4 belong to 1 cycle. Therefore Φ has finite cycles. Now the tessellation of H^4 by congruent copies of P^4 has the property that 5 copies of P^4 meet along a ridge, 20 copies of P^4 meet along an edge, and 600 copies of P^4 meet at a vertex. Consequently, the normalized solid angle subtended by P^4 at an interior ridge point is $1/5$, at an interior edge point is $1/20$, and at a vertex is $1/600$. Hence, each cycle has normalized solid angle sum 1. Thus Φ is proper.

Let M be the space obtained by gluing the sides of P^4 by the opposite side-pairing Φ. Then M is a closed, orientable, hyperbolic 4-manifold by Theorem 11.1.1. The manifold M is called the *Davis 120-cell space.*

Complete Gluing of n-Manifolds

We now consider gluing together polyhedra to form a complete manifold. We begin by proving a complete gluing theorem for Euclidean manifolds.

Theorem 11.1.2. *Let M be a Euclidean n-manifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, finite-sided, n-dimensional, convex polyhedra in E^n by a proper $\mathrm{I}(E^n)$-side-pairing Φ. Then M is complete.*

Proof: Without loss of generality, we may assume that M is connected. Then M is a metric space with the induced metric. We shall prove that M is complete by finding an $\epsilon > 0$ so that $\overline{B}(u, \epsilon)$ is compact for every u in M. It will then follow from Theorem 8.5.1 that M is complete.

Let Π be the union of the polyhedra in $\mathcal{P}$ and let $\pi : \Pi \to M$ be the quotient map. Let x be a point of Π and let $\{x_1, \ldots, x_m\}$ be the cycle of Φ containing x. Let P_i be the polyhedron in $\mathcal{P}$ containing x_i and let $r > 0$ be less than one-third the distance from x_i to any side of P_i not containing

x_i for each i. Then there is a chart

$$\phi_x : U(x,r) \to B(x,r)$$

for $(M, \pi(x))$. By Theorem 8.3.5, we have that ϕ_x^{-1} maps $B(x, r/2)$ homeomorphically onto $B(\pi(x), r/2)$. As $\overline{B}(x, r/2)$ is compact, we have

$$\phi_x^{-1}(\overline{B}(x, r/2)) = \overline{B}(\pi(x), r/2)$$

and therefore $\overline{B}(\pi(x), r/2)$ is compact.

Let Π^k be the union of all the k-faces of the polyhedra in $\mathcal{P}$ for each $k = 0, 1, \ldots, n$. Then Π^0 is a finite set. Let $r_0 > 0$ be less than one-sixth the distance from any point x of Π^0 to any side of a polyhedron in $\mathcal{P}$ not containing x. Then $\overline{B}(\pi(x), r_0)$ is compact for each x in Π^0. Now suppose that $r_k > 0$ and $\overline{B}(\pi(x), r_k)$ is compact for each x in Π^k. Let $r_{k+1} > 0$ be such that $r_{k+1} \leq r_k/2$ and for each $(k+1)$-face F of a polyhedron in $\mathcal{P}$, we have that r_{k+1} is less than one-sixth the distance from $F - N(\partial F, r_k/2)$ to any side of a polyhedron in $\mathcal{P}$ not containing F. Let x be a point of Π^{k+1}. Then there is a $(k+1)$-face F such that x is in F.

Assume first that x is in $N(\partial F, r_k/2)$. Then there is a point y of ∂F such that $|x - y| < r_k/2$. Hence $\pi(x)$ is in $B(\pi(y), r_k/2)$. By the triangle inequality, $B(\pi(x), r_{k+1}) \subset B(\pi(y), r_k)$. Therefore $\overline{B}(\pi(x), r_{k+1})$ is compact. Now assume that x is not in $N(\partial F, r_k/2)$. Let $\{x_1, \ldots, x_m\}$ be the cycle of x. Then there is a $(k+1)$-face F_i of a polyhedron in $\mathcal{P}$ such that x_i is in F_i° for each i. Moreover x_i is not in $N(\partial F_i, r_k/2)$ for each i because each element of Φ is an isometry. Therefore r_{k+1} is less than one-sixth the distance from x_i to any side of a polyhedron in $\mathcal{P}$ not containing x_i for each i. Hence $\overline{B}(\pi(x), r_{k+1})$ is compact. It follows by induction that $\overline{B}(\pi(x), r_n)$ is compact for all x in Π. □

Let M be a hyperbolic n-manifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, finite-sided, n-dimensional, convex polyhedra in B^n by a proper $\mathrm{M}(B^n)$-side-pairing Φ. We shall determine necessary and sufficient conditions such that M is complete. We may assume, without loss of generality, that no two polyhedrons in $\mathcal{P}$ meet at infinity. Then Φ extends to a side-pairing of the $(n-1)$-dimensional sides of the Euclidean closures of the polyhedra in $\mathcal{P}$, which, in turn, generates an equivalence relation on the union of the Euclidean closures of the polyhedra in $\mathcal{P}$. The equivalence classes are called *cycles*. We denote the cycle containing a point x by $[x]$.

Let P be a polyhedron in $\mathcal{P}$. A *cusp point* of P is a point c of $\overline{P} \cap S^{n-1}$ that is the intersection of the Euclidean closures of all the sides of P incident with c. The cycle of a cusp point of a polyhedron in $\mathcal{P}$ is called a *cusp point* of M. As each polyhedron in $\mathcal{P}$ has only finitely many cusp points, M has only finitely many cusp points.

Let c be a cusp point of a polyhedron in $\mathcal{P}$. Let b be a point in $[c]$ and let P_b be the polyhedron in $\mathcal{P}$ containing b in its Euclidean closure. The *link* of b is defined to be the $(n-1)$-dimensional, Euclidean, convex polyhedron $L(b)$ obtained by intersecting P_b with a horosphere Σ_b based at b that meets only the sides of P_b incident with b. We shall assume that the horospheres $\{\Sigma_b : b \in [c]\}$ have been chosen small enough so that the links of the points of $[c]$ are mutually disjoint. Then Φ determines a proper $S(E^{n-1})$-side-pairing for $\{L(b) : b \in [c]\}$ as in §10.2. Let $L[c]$ be the space obtained by gluing together the polyhedra $\{L(b)\}$ by this side-pairing. The space $L[c]$ is called the *link of the cusp point* $[c]$ of M.

Theorem 11.1.3. *The link $L[c]$ of a cusp point $[c]$ of M is a connected, Euclidean, similarity $(n-1)$-manifold.*

Proof: The space $L[c]$ is a $(E^{n-1}, S(E^{n-1}))$-manifold by Theorem 11.1.1. It follows directly from the definition of a cycle that $L[c]$ is connected. □

Theorem 11.1.4. *The link $L[c]$ of a cusp point $[c]$ of M is complete if and only if the links $\{L(b)\}$ for the points in $[c]$ can be chosen so that Φ restricts to a side-pairing for $\{L(b)\}$.*

Proof: If links for the points in $[c]$ can be chosen so that Φ restricts to a side-pairing for $\{L(b)\}$, then this side pairing for $\{L(b)\}$ is an $I(E^{n-1})$-side-pairing, and so $L[c]$ is complete by Theorem 11.1.2. The converse is proved by the same argument as in the proof of Theorem 10.2.2. □

Theorem 11.1.5. *If the link $L[c]$ of a cusp point $[c]$ of M is complete, then there is a horoball $B(c)$ based at the point c, a discrete subgroup Γ_c of* $M(B^n)$ *leaving $B(c)$ invariant, and an injective local isometry*

$$\iota : B(c)/\Gamma_c \to M$$

compatible with the projection of P_c to M.

Proof: The proof is the same as the proof of Theorem 10.2.3. □

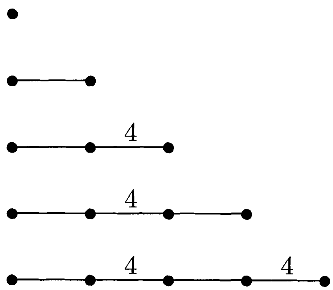
Theorem 11.1.6. *Let M be a hyperbolic n-manifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, finite-sided, n-dimensional, convex polyhedra in B^n by a proper* $M(B^n)$*-side-pairing Φ. Then M is complete if and only if $L[c]$ is complete for each cusp point $[c]$ of M.*

Proof: Without loss of generality, we may assume that M is connected. Suppose that $L[c]$ is incomplete for some cusp point $[c]$ of M. Then M is incomplete by the same argument as in the proof of Theorem 10.2.4. Conversely, suppose that $L[c]$ is complete for each cusp point $[c]$. Let M_0 be the manifold-with-boundary obtained from M by removing the image of the injective local isometry

$$\iota : B(c)/\Gamma_c \to M$$

of Theorem 11.1.5 for each cusp point $[c]$ of M. Then M_0 is complete by the same argument as in the proof of Theorem 11.1.2. Finally M is complete by the same argument as in the proof of Theorem 9.8.5. □

Example 2. We now consider an example of an open, complete, hyperbolic 4-manifold of finite volume obtained by gluing together the sides of a 4-dimensional, regular, ideal, convex polyhedron in H^4. For $n = 0, 1, 2, 3, 4$, let Γ_n be the discrete, n-simplex, reflection group whose Coxeter graph is, respectively,



For $n = 1, 2, 3$, the group Γ_n is a discrete group of isometries of S^n generated by the reflections of S^n in the sides of a spherical n-simplex Δ^n. The group Γ_4 is a discrete group of isometries of H^4 generated by the reflections of H^4 in the sides of a generalized hyperbolic 4-simplex Δ^4. For $n = 1, 2, 3, 4$, let v_n be a vertex of Δ^n such that the subgroup of Γ_n fixing v_n is Γ_{n-1}. Then the images of Δ^n under Γ_{n-1} fit together at v_n to give the barycentric subdivision of a regular convex polyhedron P^n in S^n, if $n = 1, 2, 3$, or in H^4 if $n = 4$. The images of P^n under Γ_n form an exact tessellation of S^n, if $n = 1, 2, 3$, or of H^4 if $n = 4$, by congruent copies of P^n. The group of symmetries of this tessellation is Γ_n. The order of Γ_n, for $n = 0, 1, 2, 3, 4$, is $2, 6, 48, 1152, \infty$, respectively.

For $n = 1, 2, 3$, the convex hull of the set of vertices of this tessellation of S^n is a regular Euclidean convex polyhedron Q^{n+1} that is combinatorially equivalent to P^{n+1}. The set P^1 is an arc of twice the length of Δ^1 and so S^1 is tessellated by $6/2 = 3$ copies of it. Hence Q^2 is an equilateral triangle. Therefore P^2 is a spherical equilateral triangle and S^2 is tessellated by $48/6 = 8$ copies of it. Hence Q^3 is a regular octahedron. Therefore P^3 is a regular spherical octahedron and S^3 is tessellated by $1152/48 = 24$ copies of it. The 4-dimensional regular polyhedron Q^4 is called the 24-cell. All the vertices of P^4 are ideal. Therefore P^4 is a regular, ideal, hyperbolic 24-cell.

The 24-cell Q^4 has 24 sides, 96 ridges, 96 edges, and 24 vertices. Each side S of Q^4 is a regular octahedron and is parallel to its opposite side, $-S$. We rotate Q^4 so that its vertices are $\pm e_i$, for $i = 1, 2, 3, 4$, and $(\pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2}, \pm\frac{1}{2})$. We pass to the projective model D^4 of hyperbolic space and rotate P^4 so that Q^4 and P^4 coincide. We now pair each side S of P^4 to its opposite side S' by an orientation reversing isometry g_S of D^4. For each of the eight sides of P^4 whose Euclidean centers are $(\pm\frac{1}{2}, 0, 0, \pm\frac{1}{2})$ and $(0, \pm\frac{1}{2}, \pm\frac{1}{2}, 0)$, let g_S be the composite of the antipodal map followed by the reflection in the side S. Now each side of P^4 has two vertices of the form $\pm e_i$ and $\pm e_j$ with $i \neq j$. For the remaining 16 sides of P^4, let g_S be the composition of the reflection of D^4 that pairs S to S' followed by the reflection of D^4 that transposes the vertices $\pm e_i$ and $\pm e_j$ of S, and then followed by the reflection in the side S. Then $\Phi = \{g_S\}$ is an $\mathrm{I}(D^4)$-side-pairing for P^4.

One can check that each ridge cycle contains 4 points and each edge cycle contains 8 points. Therefore Φ has finite cycles. Now the tessellation of D^4 by congruent copies of P^4 has the property that 4 copies of P^4 meet along a ridge and 8 copies of P^4 meet along an edge. Consequently, the normalized solid angle subtended by P^4 at an interior ridge point is $1/4$ and at an interior edge point is $1/8$. Hence, each cycle has normalized solid angle sum 1. Thus Φ is proper.

Let M be the space obtained by gluing the sides of P^4 by Φ. Then M is a hyperbolic 4-manifold by Theorem 11.1.1. The manifold M is noncompact and nonorientable but has finite volume. We shall call M the *hyperbolic 24-cell space.*

There are 6 cycles of ideal vertices of P^4. Each element g_S of Φ is the composite of a rotation about the origin followed by the reflection in S. Consequently, disjoint horospheres based at the ideal vertices of P^4 and equidistant from the origin are paired by the elements of Φ. Therefore, the links of the cusp points of M are complete by Theorem 11.1.4. Finally M is complete by Theorem 11.1.6.

Exercise 11.1

1. Prove that every facet of an n-dimensional, abstract, convex polyhedron is an $(n-1)$-dimensional convex polyhedron.

2. Let P be a convex fundamental polyhedron for a discrete group Γ of isometries of X and let $\mathcal{F}$ be the collection of $(n-1)$-dimensional convex subsets of ∂P of the form $P \cap gP$ for some g in Γ. Prove that P together with $\mathcal{F}$ is an abstract convex polyhedron in X.

3. For each facet F of P in Exercise 2, let g_F be the element of Γ such that $P \cap g_F(P) = F$. Prove that $\Phi = \{g_F : F \in \mathcal{F}\}$ is a Γ-facet-pairing for P.

4. Prove Theorem 11.1.1.

5. Let Γ be the group generated by the opposite side-pairing of the hyperbolic 120-cell P^4. Prove that Γ is a torsion-free subgroup of Γ_4 of index 14400.

6. Let P be a finite-sided convex polyhedron in E^n. Prove that for each $r > 0$, the set $P - N(\partial P, r)$ is either empty or a finite-sided convex polyhedron.

7. Let P and Q be disjoint, finite-sided, convex, polyhedrons in E^n. Prove that $\mathrm{dist}(P, Q) > 0$.

8. Explain why the argument in the proof of Theorem 11.1.2 breaks down in the hyperbolic case.

§11.2. Poincaré's Theorem

In this section, we prove Poincaré's fundamental polyhedron theorem for freely acting discrete groups of isometries of $X = S^n, E^n$, or H^n with $n > 1$. We begin by proving a weak version of Poincaré's theorem.

Theorem 11.2.1. *Let Φ be a proper $\mathrm{I}(X)$-side-pairing for an n-dimensional convex polyhedron P in X such that the $(X, \mathrm{I}(X))$-manifold M obtained by gluing together the sides of P by Φ is complete. Then the group Γ generated by Φ is discrete and acts freely, P is an exact, convex, fundamental polyhedron for Γ, and the inclusion of P into X induces an isometry from M to the space-form X/Γ.*

Proof: The quotient map $\pi : P \to M$ maps P° homeomorphically onto an open subset U of M. Let $\phi : U \to X$ be the inverse of π. From the construction of M, we have that ϕ is locally a chart for M. Therefore ϕ is a chart for M.

Let $\kappa : \tilde{M} \to M$ be a universal covering. As U is simply connected, $\phi : U \to X$ lifts to a chart $\tilde{\phi} : \tilde{U} \to X$ for $\tilde{M}$. Let $\delta : \tilde{M} \to X$ be the developing map determined by $\tilde{\phi}$. Then δ is an isometry by Theorem 8.5.9. Let $\zeta = \kappa\delta^{-1}$. Then $\zeta : X \to M$ is a covering projection extending π on P°. Moreover, by continuity, ζ extends π.

Let Γ be the group of covering transformations of ζ. By Theorem 8.5.9, we have that Γ is a freely acting discrete group of isometries of X and ζ induces an isometry from X/Γ to M. Now as U is simply connected, it is evenly covered by ζ. Hence, the members of $\{gP^\circ : g \in \Gamma\}$ are mutually disjoint. As $\pi(P) = M$, we have

$$X = \cup\{gP : g \in \Gamma\}.$$

Therefore P° is a fundamental domain for Γ.

Let g_S be an element of Φ. Choose a point y in the interior of the side S of P. Then there is an point y' in the interior of the side S' of P such that $g_S(y') = y$. Since $\pi(y') = y$, there is an element g of Γ such that $g(y') = y$. Since gS' does not extend into P°, we must have that gS' lies on the hyperplane $\langle S \rangle$.

Now since $\pi : P \to M$ maps a neighborhood of y in S injectively into M, we must have that g and g_S agree on a neighborhood of y' in S'. Hence $g = g_S$ on $\langle S' \rangle$. Furthermore, since gP lies on the opposite side of S from P, we deduce that $g = g_S$ by Theorem 4.3.6. Thus Γ contains Φ. Therefore P/Γ is a quotient of M.

Now by Theorem 6.5.8, the inclusion map of P into X induces a continuous bijection from P/Γ to X/Γ. The composition of the induced maps

$$X/\Gamma \to M \to P/\Gamma \to X/\Gamma$$

restricts to the identity map of P° and so is the identity map by continuity. Therefore $M = P/\Gamma$.

Now since $\zeta : X \to M$ induces an isometry from X/Γ to $M = P/\Gamma$, the inclusion map of P into X induces an isometry from P/Γ to X/Γ. Therefore P is locally finite by Theorem 6.5.8. Hence P is an exact, convex, fundamental polyhedron for Γ. Finally Φ generates Γ by Theorem 6.7.3. □

In order to apply Theorem 11.2.1, we need to know that the manifold M is complete. If $X = S^n$, then M is always complete, since M is compact. If $X = E^n$ and the polyhedron P is finite-sided, then M is complete by Theorem 11.1.2. If $X = H^n$ and P is finite-sided, then easily verifiable necessary and sufficient conditions for M to be complete are given by Theorems 11.1.4 and 11.1.6. If $X = H^n$ and P has infinitely many sides, then M may fail to be complete even though the conditions of Theorem 11.1.6 are satisfied. This phenomenon is exhibited by the next example.

Example 1. We now consider a proper side-pairing Φ for an infinite-sided hyperbolic polygon P, with no vertices, such that the hyperbolic surface M obtained by gluing together the sides of P by Φ is incomplete. Let $\{S_n\}_{n=1}^\infty$ and $\{S'_n\}_{n=1}^\infty$ be sequences of disjoint lines of U^2 formed by Euclidean semicircles of unit radius whose centers lie on the real line $\mathbb{R}$ in the increasing order

$$S_1, S'_1, S_2, S'_2, \ldots$$

such that

$$\mathrm{dist}_U(S_n, S'_n) = 1/2^n = \mathrm{dist}_U(S'_n, S_{n+1})$$

for each n. Let P be the closed region of U^2 above and bounded by the family of lines $\{S_n, S'_n\}_{n=1}^\infty$. Then P is a convex polygon in U^2 whose sides are the lines $\{S_n, S'_n\}_{n=1}^\infty$.

Let x'_n be the point of S'_n nearest to S_{n+1} and let x_{n+1} be the point of S_{n+1} nearest to S'_n for each n. Then the geodesic segment $[x'_n, x_{n+1}]$ is orthogonal to both S'_n and S_{n+1} and has length $1/2^n$. Let g_1 be the composition of the reflection in the vertical line midway between S_1 and S'_1 followed by the reflection in S_1, and for each $n > 1$, let g_n be the composition of the reflection in the vertical line midway between S_n and S'_n followed by the reflection in S_n, and then followed by the translation along S_n so that

$$g_n(x'_n) = x_n.$$

Then $g_n(S'_n) = S_n$ and

$$\Phi = \{g_n, g_n^{-1}\}_{n=1}^\infty$$

is a proper $\mathrm{I}_0(U^2)$-side-pairing for P. Let $\pi : P \to M$ be the quotient map. Observe that the union of geodesic segments

$$[x'_1, x_2] \cup [x'_2, x_3] \cup \cdots$$

projects to a half-open geodesic section in M of length one. Hence, we have that $\{\pi(x_n)\}_{n=1}^\infty$ is a Cauchy sequence in M. Observe that this sequence does not converge in M, since each point of M has a neighborhood in M that contains at most one term of the sequence $\{\pi(x_n)\}$. Thus M is incomplete. Therefore P is not a fundamental polygon for the group Γ generated by Φ by Theorems 6.5.8 and 8.5.2.

Note that the same construction works in all dimensions. Just replace the semicircles with hemispheres all of whose centers are collinear.

Poincaré's Fundamental Polyhedron Theorem

Let $\mathcal{S}$ be the set of sides of an exact, convex, fundamental polyhedron P for a freely acting discrete group Γ of isometries of X. Then for each S in $\mathcal{S}$, we have the side-pairing relation

$$g_S g_{S'} = 1$$

of Γ. The expression SS' is called the word in $\mathcal{S}$ corresponding to the side-pairing relation $g_S g_{S'} = 1$ of Γ. Recall from §6.7 that each cycle of sides $\{S_i\}_{i=1}^{\ell}$ of P determines a cycle relation

$$(g_{S_1} g_{S_2} \cdots g_{S_\ell})^k = 1$$

of Γ, where k is the order of $g_{S_1} g_{S_2} \cdots g_{S_\ell}$.

If $X = S^n$, then $g_{S_1} g_{S_2} \cdots g_{S_\ell}$ leaves invariant a ridge R of P and so fixes a point of R by the Brouwer fixed point theorem; but Γ acts freely on S^n, therefore we must have $k = 1$. If $X = E^n$ or H^n, then Γ is torsion-free and so $k = 1$. Thus, we have the cycle relation

$$g_{S_1} g_{S_2} \cdots g_{S_\ell} = 1.$$

The expression $S_1 S_2 \cdots S_\ell$ is called the word in $\mathcal{S}$ corresponding to the above cycle relation of Γ. We are now ready to state Poincaré's fundamental polyhedron theorem for freely acting discrete groups of isometries of X.

Theorem 11.2.2. *Let Φ be a proper $\mathrm{I}(X)$-side-pairing for an n-dimensional convex polyhedron P in X such that the $(X, \mathrm{I}(X))$-manifold M obtained by gluing together the sides of P by Φ is complete. Then the group Γ generated by Φ is discrete and acts freely, P is an exact, convex, fundamental polyhedron for Γ, and if $\mathcal{S}$ is the set of sides of P and $\mathcal{R}$ is the set of words in $\mathcal{S}$ corresponding to all the side-pairing and cycle relations of Γ, then $(\mathcal{S}; \mathcal{R})$ is a group presentation for Γ under the mapping $S \mapsto g_S$.*

Proof: (1) By Theorem 11.2.1, the group Γ is discrete and acts freely, and P is an exact, convex, fundamental polyhedron for Γ.

(2) Let F be the group freely generated by the elements of $\mathcal{S}$. Then we have an epimorphism $\eta : F \to \Gamma$ defined by $\eta(S) = g_S$. By Theorem 6.7.7, the kernel of η contains the elements of $\mathcal{R}$. Let G be the quotient of F by the normal closure of the set $\mathcal{R}$ in F. Then η induces an epimorphism

$$\iota : G \to \Gamma.$$

We shall prove that ι is an isomorphism.

(3) Let $G \times P$ be the cartesian product of G and P. We topologize $G \times P$ by giving G the discrete topology and $G \times P$ the product topology. Then $G \times P$ is the topological sum of the subspaces

$$\{\{g\} \times P : g \in G\}.$$

Moreover, the mapping $(g, x) \mapsto \iota(g)x$ is a homeomorphism of $\{g\} \times P$ onto $\iota(g)P$ for each g in G.

(4) Two points (g, x) and (h, y) of $G \times P$ are said to be *paired* by Φ, written $(g, x) \simeq (h, y)$, if and only if $g^{-1}h$ is in $\mathcal{S}$ and $\iota(g)x = \iota(h)y$. Suppose $(g, x) \simeq (h, y)$. Then there is a side S of P such that $g^{-1}h = S$. As $S^{-1} = S'$ in G, we have that $(h, y) \simeq (g, x)$. Furthermore x is in $P \cap g_S(P) = S$ and $y = x'$ is in S'.

Two points (g, x) and (h, y) of $G \times P$ are said to be *related* by Φ, written $(g, x) \sim (h, y)$, if and only if there is a finite sequence, $(g_0, x_0), \ldots, (g_k, x_k)$, of points of $G \times P$ such that $(g, x) = (g_0, x_0)$, $(g_k, x_k) = (h, y)$, and

$$(g_{i-1}, x_{i-1}) \simeq (g_i, x_i) \quad \text{for } i = 1, \ldots, k.$$

Being related by Φ is obviously an equivalence relation on $G \times P$; moreover, if $(g, x) \sim (h, y)$, then $x \sim y$. Let $[g, x]$ be the equivalence class of (g, x) and let $\tilde{X}$ be the quotient space of $G \times P$ of equivalence classes.

(5) If $(g, x) \simeq (h, y)$, then obviously $(fg, x) \simeq (fh, y)$ for each f in G. Hence G acts on $\tilde{X}$ by $f[g, x] = [fg, x]$. For a subset A of P, set

$$[A] = \{[1, x] : x \in A\}.$$

Then, if g is in G, we have

$$g[A] = \{[g, x] : x \in A\}.$$

If (g, x) is in $G \times P^\circ$, then

$$[g, x] = \{(g, x)\}.$$

Consequently, the members of $\{g[P^\circ] : g \in G\}$ are mutually disjoint in $\tilde{X}$.

(6) We now show that $\tilde{X}$ is connected. Let $\pi : G \times P \to \tilde{X}$ be the quotient map. As π maps $\{g\} \times P$ onto $g[P]$, we have that $g[P]$ is connected. In view of the fact that

$$\tilde{X} = \cup\{g[P] : g \in G\},$$

it suffices to show that for any g in G, there is a finite sequence $g_0, \ldots, g_m$ in Γ such that $[P] = g_0[P]$, $g_m[P] = g[P]$, and $g_{i-1}[P]$ and $g_i[P]$ intersect for each $i > 0$. As G is generated by the elements of $\mathcal{S}$, there are sides S_i of P such that $g = S_1 \cdots S_m$. Let $g_0 = 1$ and $g_i = S_1 \cdots S_i$ for $i = 1, \ldots, m$. Now since

$$S_i = P \cap g_{S_i}(P),$$

we have that

$$[S_i] \subset [P] \cap S_i[P].$$

Therefore, we have

$$g_{i-1}[S_i] \subset g_{i-1}[P] \cap g_i[P].$$

Thus $\tilde{X}$ is connected.

(7) Let P_0 be P minus all its faces of dimension less than $n-3$. Set

$$\tilde{X}_0 = \cup\{g[P_0] : g \in G\}.$$

Then the same argument as in (6) shows that $\tilde{X}_0$ is connected.

(8) Let

$$\kappa : \tilde{X} \to X$$

be the function defined by $\kappa[g,x] = \iota(g)x$. Then κ is continuous, since $\kappa\pi : G \times P \to X$ is continuous. Moreover κ maps $g[P]$ homeomorphically onto $\iota(g)P$, since $\kappa\pi$ maps $\{g\} \times P$ homeomorphically onto $\iota(g)P$.

(9) Let

$$X_0 = \cup\{\gamma P_0 : \gamma \in \Gamma\}.$$

Then κ restricts to a surjection

$$\kappa_0 : \tilde{X}_0 \to X_0.$$

Hence X_0 is connected.

(10) We now show that $\kappa_0 : \tilde{X}_0 \to X_0$ is a covering projection. Let x be an arbitrary point of X_0; we need to find an open neighborhood U of x in X_0 that is evenly covered by κ_0. Let γ be an element of Γ such that x is in γP_0. Now since

$$\kappa_0 g = \iota(g)\kappa_0$$

for all g in G, we may assume that $\gamma = 1$.

Assume first that x is in P°. Then $U = P^\circ$ is an open neighborhood of x in X_0 that is evenly covered by κ_0 and the sheets over U are the members of

$$\{g[P^\circ] : g \in \mathrm{Ker}(\iota)\}.$$

Now assume that x is in the interior of a side S of P. Then we have

$$[1,x] = \{(1,x), (S,x')\}.$$

Hence, the set $[S^\circ]$ meets only $[P]$ and $S[P]$ among the members of

$$\{g[P] : g \in G\}.$$

Consequently

$$U = P^\circ \cup S^\circ \cup g_S P^\circ$$

is an open neighborhood of x in X_0 that is evenly covered by κ_0 and the sheets over U are the members of

$$\{g([P^\circ] \cup [S^\circ] \cup S[P^\circ]) : g \in \mathrm{Ker}(\iota)\}.$$

Now assume that x is in the interior of a ridge R of P. Let $\{S_i\}_{i=1}^{\ell}$ be the cycle of sides of P with $S_1 = S$ and $R = S'_\ell \cap S_1$. Let $x_1 = x$ and $x_{i+1} = g_{S_i}^{-1}(x_i)$ for $i = 1, \ldots, \ell-1$. Then $g_{S_\ell}(x_1) = x_\ell$ and

$$x = x_1 \simeq x_2 \simeq \cdots \simeq x_\ell \simeq x.$$

Therefore, we have

$$[x] = \{x_1, \ldots, x_\ell\}.$$

Now

$$(1, x) = (1, x_1) \simeq (S_1, x_2) \simeq \cdots \simeq (S_1 \cdots S_{\ell-1}, x_\ell).$$

As $S_1 \cdots S_\ell = 1$ in G, we have

$$(S_1 \cdots S_{\ell-1}, x_\ell) \simeq (1, x),$$

which closes the cycle of $(1, x)$. Therefore

$$[1, x] = \{(1, x_1), (S_1, x_2), \ldots, (S_1 \cdots S_{\ell-1}, x_\ell)\}.$$

Let $g_1 = 1$ and let $g_i = S_1 \cdots S_{i-1}$ for each $i = 2, \ldots, \ell$. The elements $\iota(g_1), \ldots, \iota(g_\ell)$ of Γ are distinct, since the polyhedra $\iota(g_1)P, \ldots, \iota(g_\ell)P$ form a cycle around their common ridge R of one revolution. See Figure 9.2.2. Therefore, the elements $g_1, \ldots, g_\ell$ of G are distinct. Now the set $[R^\circ]$ meets only $g_1[P], \ldots, g_\ell[P]$ among the members of $\{g[P] : g \in G\}$. Consequently

$$U = R^\circ \cup \bigcup_{i=1}^{\ell} \iota(g_i)S_i^\circ \cup \bigcup_{i=1}^{\ell} \iota(g_i)P^\circ$$

is an open neighborhood of x in X_0 that is evenly covered by κ_0 and the sheets over U are the members of

$$\{g([R^\circ] \cup \bigcup_{i=1}^{\ell} g_i[S_i^\circ] \cup \bigcup_{i=1}^{\ell} g_i[P^\circ]) : g \in \mathrm{Ker}(\iota)\}.$$

Thus κ_0 is a covering projection.

(11) Now X_0 is simply connected by a general position argument. Hence $\kappa_0 : \tilde{X}_0 \to X_0$ is a homeomorphism. Observe that κ maps $g[P^\circ]$ onto P° for all g in $\mathrm{Ker}(\iota)$ and the members of $\{g[P^\circ] : g \in \mathrm{Ker}(\iota)\}$ are mutually disjoint. Therefore $\mathrm{Ker}(\iota) = \{1\}$. Hence $\iota : G \to \Gamma$ is an isomorphism. Thus $(\mathcal{S}; \mathcal{R})$ is a group presentation for Γ under the mapping $S \mapsto g_S$. □

Theorem 11.2.2 gives a group presentation $(\mathcal{S}; \mathcal{R})$ for the group Γ generated by the side-pairing Φ of P. The presentation $(\mathcal{S}; \mathcal{R})$ can be simplified by eliminating each side-pairing relation $SS' = 1$ and exactly one of the generators S or S'. If S' is eliminated, then each occurrence of S' in a cycle relation is replaced by S^{-1}. Moreover, each cycle of sides $\{S_i\}_{i=1}^{\ell}$ determines 2ℓ cycles of sides by taking cyclic permutations of $\{S_i\}_{i=1}^{\ell}$ and their inverse orderings. The corresponding cycle transformations are all conjugate to each other or their inverses. Therefore, any one of the corresponding cycle relations is derivable from any one of the others. Hence, all but one of them can be eliminated from a presentation for Γ. Thus $(\mathcal{S}; \mathcal{R})$ can be simplified to a presentation with half the generators and one relation for each cycle of ridges of P.

Example 2. Consider the ideal square P in U^2 in Figure 9.8.6. Label the sides of P left to right by S, T, T', S'. Let M be the hyperbolic surface obtained by gluing the sides of P by the side-pairing Φ described in Example

2 of §9.8. Then M is a thrice-punctured sphere. Therefore M has three cusp points. It is clear that links for the cusp points of P can be chosen so that Φ pairs their endpoints. Hence M is complete. By Theorem 11.2.2, the group Γ generated by Φ has the presentation

$$(S, S', T, T'; SS', TT').$$

We eliminate the generators S' and T' and the side-pairing relations to obtain the presentation (S, T) for Γ. Thus Γ is a free group of rank two generated by g_S and g_T.

Example 3. Consider the regular octagon P in B^2 in Figure 9.2.3. Let M be the hyperbolic surface obtained by gluing the sides of P by the side-pairing Φ described in Example 4 of §9.2. Then M is a closed orientable surface of genus two. Observe that P has one cycle of vertices and therefore essentially one cycle of sides

$$\{S_1, T_1, S_1', T_2', S_2, T_2, S_2', T_1'\}.$$

Hence, the group Γ generated by Φ has the presentation

$$(S_1, T_1, S_2, T_2; S_1T_1S_1^{-1}T_2^{-1}S_2T_2S_2^{-1}T_1^{-1}).$$

Example 4. Consider a regular ideal octahedron P in B^3 with the gluing pattern for the Whitehead link complement in Figure 10.3.14. Then P has three cycles of edges and therefore essentially three cycles of sides

$$\{A, D', C, B'\}, \{B, C, D', C'\}, \{A, B, A', D'\}.$$

Therefore, the Whitehead link group has the presentation

$$(A, B, C, D; AD^{-1}CB^{-1}, BCD^{-1}C^{-1}, ABA^{-1}D^{-1}).$$

Exercise 11.2

1. Show that Theorem 11.2.2 does not hold for $X = S^1$ but does hold for $X = E^1$ or H^1.
2. Use the gluing pattern for the 3-torus M in Example 1 of §10.1 to find a presentation for $\pi_1(M)$.
3. Use the gluing pattern for the Poincaré dodecahedral space M in Figure 10.1.1 to find a presentation for $\pi_1(M)$.
4. Use the gluing pattern for the Seifert-Weber dodecahedral space M in Figure 10.1.2 to find a presentation for $\pi_1(M)$.
5. Use the gluing pattern for the figure-eight knot complement M in Figure 10.3.3 to find a presentation for $\pi_1(M)$.
6. Reduce the presentation from the previous exercise to a two-generator, one-relator presentation.
7. Reduce the presentation from Example 4 to a two-generator, one-relator presentation.

§11.3. Simplices of Maximum Volume

An n-simplex Δ^n in B^n is said to be *regular* if and only if every permutation of the vertices of Δ^n is induced by a Möbius transformation of B^n. In this section, we prove that an n-simplex Δ^n in B^n has maximum volume if and only if Δ^n is regular and ideal. As every simplex in B^n is contained in an ideal simplex, it suffices to consider only ideal simplices.

In dimension one, B^1 is the only ideal 1-simplex and B^1 is regular and of maximum length. Thus, we may assume that $n \geq 2$. In dimension two, all ideal triangles are congruent in B^2, and so all ideal triangles are regular and of maximum area. In dimension three, an ideal tetrahedron has maximum volume if and only if it is regular by Theorem 10.4.7. Thus, we are only concerned with dimensions $n \geq 4$.

Lemma 1. *The volume of an n-simplex Δ^n in D^n is given by*

$$\mathrm{Vol}(\Delta^n) = \int_{\Delta^n} \frac{dx_1 \cdots dx_n}{(1-|x|^2)^{(n+1)/2}}.$$

Proof: By Theorem 6.1.6, the element of hyperbolic volume of the projective disk model D^n is $dx_1 \cdots dx_n/(1-|x|^2)^{(n+1)/2}$. □

Let Δ^n be an ideal n-simplex in U^n with vertices $v_0, \ldots, v_n$. By replacing Δ^n with a congruent n-simplex, we may assume that $v_0 = \infty$. Since $v_1, \ldots, v_n$ all lie on an $(n-2)$-sphere in E^{n-1} and the group $\mathrm{S}(E^{n-1})$ acts transitively on the set of all $(n-2)$-spheres in E^{n-1}, we may assume, without loss of generality, that $v_1, \ldots, v_n$ are in S^{n-2}. Then the side of Δ^n, spanned by $v_1, \ldots, v_n$, lies in the northern hemisphere of S^{n-1}. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Since all the sides of Δ^n incident with ∞ are vertical, $\nu(\Delta^n)$ is a Euclidean $(n-1)$-simplex with deleted vertices. Therefore $\nu(\Delta^n)$ is an ideal $(n-1)$-simplex in D^{n-1}. We shall use this fact to set up an induction on the dimension n.

Lemma 2. *The volume of an ideal n-simplex Δ^n in U^n, with vertices $v_0, \ldots, v_n$ such that $v_0 = \infty$ and $v_1, \ldots, v_n$ are in S^{n-2}, is given by*

$$\mathrm{Vol}(\Delta^n) = \frac{1}{n-1} \int_{\nu(\Delta^n)} \frac{dx_1 \cdots dx_{n-1}}{(1-|x|^2)^{(n-1)/2}}.$$

Proof: By Theorem 4.6.7, the element of hyperbolic volume of the upper half-space model U^n is $dx_1 \cdots dx_n/(x_n)^n$. Therefore, we have

$$\begin{aligned}
\mathrm{Vol}(\Delta^n) &= \int_{\Delta^n} \frac{dx_1 \cdots dx_n}{(x_n)^n} \\
&= \int_{\nu(\Delta^n)} \left(\int_{(1-|\nu(x)|^2)^{\frac{1}{2}}}^{\infty} \frac{dx_n}{(x_n)^n} \right) dx_1 \cdots dx_{n-1} \\
&= \frac{1}{n-1} \int_{\nu(\Delta^n)} \frac{dx_1 \cdots dx_{n-1}}{(1-|\nu(x)|^2)^{(n-1)/2}}.
\end{aligned}$$

□

Lemma 3. *Let Δ^n be an ideal n-simplex in U^n, with vertices $v_0, \ldots, v_n$ such that $v_0 = \infty$ and $v_1, \ldots, v_n$ are in S^{n-2}, and let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then Δ^n is regular if and only if $\nu(\Delta^n)$ is Euclidean regular.*

Proof: Suppose that Δ^n is regular. To prove that $\nu(\Delta^n)$ is Euclidean regular, it suffices to show that the transposition of any two vertices v, w of $\nu(\Delta^n)$ is realized by a Euclidean isometry. Since Δ^n is regular, there is a Möbius transformation τ of U^n such that τ transposes v and w and fixes every other vertex of Δ^n. As τ fixes ∞, we have that τ is the Poincaré extension of a similarity of E^{n-1}. Moreover, since τ leaves invariant the Euclidean line segment $[v, w]$, we have that τ is a Euclidean isometry. Thus $\nu(\Delta^n)$ is Euclidean regular.

Conversely, suppose that $\nu(\Delta^n)$ is Euclidean regular. To prove that Δ^n is regular, it suffices to prove that the transposition of any vertex u of $\nu(\Delta^n)$ and ∞ is realized by a Möbius transformation of U^n. Since $\nu(\Delta^n)$ is Euclidean regular, every vertex $v \neq u$ of $\nu(\Delta^n)$ is the same Euclidean distance r from u. Let σ be the reflection of E^n in the sphere $S(u, r)$. Then $\sigma(u) = \infty$, $\sigma(\infty) = u$, and σ fixes all the other vertices of Δ^n. Thus Δ^n is regular. □

Lemma 4. *Let Δ^n_* be a regular Euclidean n-simplex inscribed in S^{n-1} and let $F : D^n \to E^n$ be the vector field defined by*

$$F(x) = \frac{x}{(1 - |x|^2)^{(n-1)/2}}.$$

Then the following divergence formula holds:

$$\int_{\Delta^n_*} (\mathrm{div} F) dV = \int_{\partial \Delta^n_*} (F \cdot \hat{n}) dS,$$

where $\hat{n}$ is the outward normal to the boundary of Δ^n_*.

Proof: We first calculate the divergence of F. Observe that

$$\begin{aligned}
\mathrm{div} F(x) &= \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\frac{x_i}{(1 - |x|^2)^{(n-1)/2}}\right) \\
&= \sum_{i=1}^{n} \left(\frac{1}{(1 - |x|^2)^{(n-1)/2}} + \frac{(n-1)x_i^2}{(1 - |x|^2)^{(n+1)/2}}\right) \\
&= \frac{n}{(1 - |x|^2)^{(n-1)/2}} + \frac{(n-1)|x|^2}{(1 - |x|^2)^{(n+1)/2}} \\
&= \frac{1}{(1 - |x|^2)^{(n-1)/2}} + \frac{(n-1)}{(1 - |x|^2)^{(n+1)/2}}.
\end{aligned}$$

By Theorem 6.3.26, the set Δ^n_* has finite volume in D^n. Therefore, by Lemma 1, the integral of $(1 - |x|^2)^{-(n+1)/2}$ over Δ^n_* is finite.

Next, observe that

$$0 \leq \frac{1}{(1-|x|^2)^{(n-1)/2}} \leq \frac{1}{(1-|x|^2)^{(n+1)/2}}.$$

Therefore, the integral of $(1-|x|^2)^{-(n-1)/2}$ over Δ_*^n is finite. Hence, the integral of $\operatorname{div} F$ over Δ_*^n is finite.

Now $\partial\Delta_*^n$ consists of $n+1$ regular Euclidean $(n-1)$-simplices $\partial_i\Delta_*^n$ for $i = 0, 1, \ldots, n$. Let v_i be the vertex of Δ_*^n opposite the side $\partial_i\Delta_*^n$. Since 0 is the centroid of Δ_*^n, we have that $\sum_{i=0}^n v_i = 0$. Hence, for each j, we have that

$$\sum_{i=0}^n v_i \cdot v_j = 0.$$

As $v_i \cdot v_j$, for $i \neq j$, is independent of i and j, we have

$$1 + n v_i \cdot v_j = 0$$

and so for all $i \neq j$, we have

$$v_i \cdot v_j = -1/n.$$

Let x be any point of $\partial_i\Delta_*^n$. Then there are coefficients $t_0, \ldots, t_n$ in the interval $[0, 1]$ such that

$$x = \sum_{j=0}^n t_j v_j, \quad \sum_{j=0}^n t_j = 1, \quad \text{and} \quad t_i = 0.$$

Hence

$$x \cdot \hat{n} = x \cdot (-v_i) = -\sum_{j=0}^n t_j v_j \cdot v_i = -v_j \cdot v_i = 1/n.$$

Let a_i and r_n be the center and radius of the circumscribed $(n-2)$-sphere for $\partial_i\Delta_*^n$. Then a_i is a scalar multiple of v_i. As $a_i \cdot -v_i = 1/n$, we have that $a_i = -v_i/n$. Now $0, a_i$, and any vertex $v_j \neq v_i$ form a right triangle with the right angle at a_i. Therefore

$$|a_i|^2 + r_n^2 = 1.$$

Hence, we have

$$r_n = (1 - 1/n^2)^{\frac{1}{2}}.$$

Let x be any point of $\partial_i\Delta_*^n$. Then $0, x$, and a_i form a right triangle with the right angle at a_i. Therefore

$$|a_i|^2 + |x - a_i|^2 = |x|^2.$$

Hence, we have

$$1 - |x|^2 = r_n^2 - |x - a_i|^2.$$

Therefore

$$\int_{\partial\Delta_*^n} (F \cdot \hat{n}) dS = \frac{n+1}{n} \int_{\partial_n\Delta_*^n} \frac{dS}{(r_n^2 - |x - a_n|^2)^{(n-1)/2}}.$$

Now $\partial_n \Delta_*^n$ is congruent to $r_n \Delta_*^{n-1}$. Hence, this integral transforms into

$$\frac{n+1}{n}\int_{\Delta_*^{n-1}} \frac{dx_1 \cdots dx_{n-1}}{(1-|x|^2)^{(n-1)/2}}.$$

Moreover, this integral is finite by Lemma 2. Thus, both integrals in the desired divergence formula are finite.

For each $i = 0, 1, \ldots, n$ and real number r such that $1/2 < r < 1$, let $\Delta_i^{n-1}(r)$ be the $(n-1)$-simplex obtained by intersecting Δ_*^n with the hyperplane normal to v_i and passing through the point rv_i. Then $\Delta_i^{n-1}(r)$ is a regular Euclidean $(n-1)$-simplex for each i. Let $\Delta_*^n(r)$ be the polyhedron obtained from Δ_*^n by truncating Δ_*^n along the $(n-1)$-simplices $\Delta_0^{n-1}(r), \ldots, \Delta_n^{n-1}(r)$. Then by the divergence theorem, we have

$$\int_{\Delta_*^n(r)} (\mathrm{div} F)dV = \int_{\partial\Delta_*^n(r)} (F\cdot\hat{n})dS.$$

Taking the limit as $r \to 1$ gives the formula

$$\int_{\Delta_*^n} (\mathrm{div} F)dV = \int_{\partial\Delta_*^n} (F\cdot\hat{n})dS + \lim_{r\to 1}\left(\sum_{i=0}^{n}\int_{\Delta_i^{n-1}(r)} (F\cdot\hat{n})dS\right).$$

Thus, it remains only to show that the last term is zero.

The hyperplane spanned by $\Delta_i^{n-1}(r)$ has the equation $x \cdot v_i = r$. Hence

$$\int_{\Delta_i^{n-1}(r)} (F\cdot\hat{n})dS = \int_{\Delta_i^{n-1}(r)} \frac{rdS}{(1-|x|^2)^{(n-1)/2}}.$$

Let $\Delta_i^n(r)$ be the n-simplex spanned by $\Delta_i^{n-1}(r)$ and v_i. Then $\Delta_i^n(r)$ is a regular Euclidean n-simplex. Let s be the Euclidean distance from the centroid c_i of $\Delta_i^n(r)$ to v_i. Since the Euclidean distance from c_i to the side $\Delta_i^{n-1}(r)$ of $\Delta_i^n(r)$ is s/n, we have

$$r + (s/n) + s = 1.$$

Hence, we have

$$s = (1-r)/(1+1/n).$$

Let s_n be the radius of the circumscribed $(n-2)$-sphere for $\Delta_i^{n-1}(r)$. Then

$$s_n = s(1-1/n^2)^{\frac{1}{2}} = (1-r)\left(\frac{1-1/n}{1+1/n}\right)^{\frac{1}{2}}.$$

Observe that for each x in $\Delta_i^{n-1}(r)$, we have

$$\begin{aligned}
|x|^2 &\leq r^2 + s_n^2 \\
&= r^2 + (1-r)^2\left(\frac{n-1}{n+1}\right) \\
&\leq r^2 + (1-r)^2 \\
&= 1 - 2r + 2r^2 \\
&= 1 + 2r(r-1) \\
&\leq 1 + (r-1) \ = \ r.
\end{aligned}$$

Then we have

$$\begin{aligned}\int_{\Delta_i^{n-1}(r)} \frac{rdS}{(1-|x|^2)^{(n-1)/2}} &\leq \int_{\Delta_i^{n-1}(v)} \frac{dS}{(1-r)^{(n-1)/2}} \\ &= \frac{\mathrm{Vol}(\Delta_i^{n-1}(r))}{(1-r)^{(n-1)/2}}.\end{aligned}$$

Now

$$\mathrm{Vol}(\Delta_i^{n-1}(r)) = s^{n-1}\mathrm{Vol}(\Delta_*^{n-1}) = k_n(1-r)^{n-1}$$

for some constant k_n depending only on n. Therefore

$$\int_{\Delta_i^{n-1}(v)} \frac{rdS}{(1-|x|^2)^{(n-1)/2}} \leq k_n(1-r)^{(n-1)/2}.$$

Taking the limit as $r \to 1$, we deduce that

$$\lim_{r\to 1}\int_{\Delta_i^{n-1}(r)} \frac{rdS}{(1-|x|^2)^{(n-1)/2}} = 0.$$

□

Lemma 5. *If Δ_*^n is a Euclidean regular ideal n-simplex in D^n, then*

$$\mathrm{Vol}(\Delta_*^n) = \frac{n}{n-1}\int_{\Delta_*^n} \frac{dx_1\cdots dx_n}{(1-|x|^2)^{(n-1)/2}}.$$

Proof: By Lemma 4, we have

$$\begin{aligned}\int_{\Delta_*^n} \frac{dx_1\cdots dx_n}{(1-|x|^2)^{(n-1)/2}} &+ (n-1)\int_{\Delta_*^n} \frac{dx_1\cdots dx_n}{(1-|x|^2)^{(n+1)/2}} \\ &= \frac{n+1}{n}\int_{\Delta_*^{n-1}} \frac{dx_1\cdots dx_{n-1}}{(1-|x|^2)^{(n-1)/2}}.\end{aligned}$$

By Lemmas 1-3, we have

$$\int_{\Delta_*^n} \frac{dx_1\cdots dx_n}{(1-|x|^2)^{(n-1)/2}} + (n-1)\mathrm{Vol}(\Delta_*^n) = \frac{(n+1)(n-1)}{n}\mathrm{Vol}(\Delta_*^n).$$

Hence

$$\frac{1}{(n-1)}\int_{\Delta_*^n} \frac{dx_1\cdots dx_n}{(1-|x|^2)^{(n-1)/2}} = \frac{1}{n}\mathrm{Vol}(\Delta_*^n).$$

□

Lemma 6. *If Δ_*^n is a Euclidean regular ideal n-simplex in D^n, then*

$$\frac{1}{n} - \frac{1}{n^2} \leq \frac{\mathrm{Vol}(\Delta_*^{n+1})}{\mathrm{Vol}(\Delta_*^n)} \leq \frac{1}{n}.$$

Proof: By Lemmas 1, 2, and 5, we have the formulas

$$\begin{aligned}\int_{\Delta_*^n} \frac{dx_1\cdots dx_n}{(1-|x|^2)^{(n+1)/2}} &= \mathrm{Vol}(\Delta_*^n), \\ \int_{\Delta_*^n} \frac{dx_1\cdots dx_n}{(1-|x|^2)^{n/2}} &= n\mathrm{Vol}(\Delta_*^{n+1}), \\ \int_{\Delta_*^n} \frac{dx_1\cdots dx_n}{(1-|x|^2)^{(n-1)/2}} &= \frac{n-1}{n}\mathrm{Vol}(\Delta_*^n).\end{aligned}$$

Hence
$$\frac{n-1}{n}\mathrm{Vol}(\Delta^n_*) \leq n\mathrm{Vol}(\Delta^{n+1}_*) \leq \mathrm{Vol}(\Delta^n_*). \qquad \square$$

Lemma 7. *Let $f : (0,1] \to \mathbb{R}$ be a continuous concave function, let c be the centroid of a Euclidean n-simplex Δ^n inscribed in S^{n-1}, and let Δ^n_* be a regular Euclidean n-simplex inscribed in S^{n-1}. Then*
$$\int_{\Delta^n} f(1-|x|^2)dV \leq \frac{\mathrm{Vol}_E(\Delta^n)}{\mathrm{Vol}_E(\Delta^n_*)}\int_{\Delta^n_*} f\big((1-|c|^2)(1-|x|^2)\big)dV$$
whenever both integrals are finite. Moreover, if f is strictly concave, then equality holds if and only if Δ^n is regular.

Proof: Let $v_0, \ldots, v_n$ be the vertices of Δ^n. Then
$$\Delta^n = \Big\{\sum_{i=0}^n t_i v_i : t_i \geq 0 \;\text{ and }\; \sum_{i=0}^n t_i = 1\Big\}.$$
Let Δ_n be the n-simplex in E^{n+1} given by
$$\Delta_n = \big\{(t_0, \ldots, t_n) : t_i \geq 0 \;\text{ and }\; \sum_{i=0}^n t_i = 1\big\}.$$
Let P be the hyperplane of E^{n+1} spanned by Δ_n and let $\alpha : P \to E^n$ be the affine bijection defined by
$$\alpha(t_0, \ldots, t_n) = \sum_{i=0}^n t_i v_i.$$
Upon changing variables by α, we have
$$\int_{\Delta^n} dV = \int_{\Delta_n} |\det \alpha'| dS.$$
Therefore, we have
$$|\det \alpha'| = \frac{\mathrm{Vol}_E(\Delta^n)}{\mathrm{Vol}_E(\Delta_n)}.$$
Upon changing variables by α, we have
$$\begin{aligned}\int_{\Delta^n} f(1-|x|^2)dV &= \int_{\Delta_n} f\Big(1-\big|\sum_{i=0}^n t_i v_i\big|^2\Big)|\det \alpha'|dS \\ &= \frac{\mathrm{Vol}_E(\Delta^n)}{\mathrm{Vol}_E(\Delta_n)}\int_{\Delta_n} f\Big(1-\big|\sum_{i=0}^n t_i v_i\big|^2\Big)dS.\end{aligned}$$

Let σ be a permutation of the set $\{0, \ldots, n\}$. As the Lebesgue measure S on P is invariant under the transformation $t_i \mapsto t_{\sigma(i)}$, we have
$$\int_{\Delta_n} f\Big(1-\big|\sum_{i=0}^n t_i v_i\big|^2\Big)dS = \int_{\Delta_n} f\Big(1-\big|\sum_{i=0}^n t_{\sigma(i)} v_i\big|^2\Big)dS.$$

Hence

$$\begin{aligned}
&\frac{\mathrm{Vol}_E(\Delta_n)}{\mathrm{Vol}_E(\Delta^n)}\int_{\Delta^n} f(1-|x|^2)dV \\
&= \frac{1}{(n+1)!}\sum_\sigma \int_{\Delta_n} f\Big(1-\big|\sum_{i=0}^n t_{\sigma(i)}v_i\big|^2\Big)dS \\
&= \int_{\Delta_n} \frac{1}{(n+1)!}\sum_\sigma f\Big(1-\big|\sum_{i=0}^n t_{\sigma(i)}v_i\big|^2\Big)dS \\
&\leq \int_{\Delta_n} f\bigg(\frac{1}{(n+1)!}\sum_\sigma\Big(1-\big|\sum_{i=0}^n t_{\sigma(i)}v_i\big|^2\Big)\bigg)dS.
\end{aligned}$$

Now

$$\Big|\sum_{i=0}^n t_{\sigma(i)}v_i\Big|^2 = \sum_{i\neq j} t_{\sigma(i)}t_{\sigma(j)}v_i\cdot v_j + \sum_{i=0}^n t_i^2.$$

Moreover

$$\begin{aligned}
&\frac{1}{(n+1)!}\sum_\sigma\sum_{i\neq j} t_{\sigma(i)}t_{\sigma(j)}v_i\cdot v_j \\
&= \sum_{i\neq j}\left(\frac{1}{(n+1)!}\sum_\sigma t_{\sigma(i)}t_{\sigma(j)}\right)v_i\cdot v_j \\
&= \sum_{i\neq j}\left(\frac{1}{n(n+1)}\sum_{k\neq\ell} t_k t_\ell\right)v_i\cdot v_j \\
&= \frac{1}{n(n+1)}\left(1-\sum_{i=0}^n t_i^2\right)\sum_{i\neq j} v_i\cdot v_j \\
&= \frac{1}{n(n+1)}\left(1-\sum_{i=0}^n t_i^2\right)\left[(n+1)^2|c|^2-(n+1)\right] \\
&= \frac{1}{n}\left(1-\sum_{i=0}^n t_i^2\right)\left((n+1)|c|^2-1\right).
\end{aligned}$$

Hence

$$\begin{aligned}
&\frac{1}{(n+1)!}\sum_\sigma\Big(1-\big|\sum_{i=0}^n t_{\sigma(i)}v_i\big|^2\Big) \\
&= 1-\frac{1}{(n+1)!}\sum_\sigma\Big|\sum_{i=0}^n t_{\sigma(i)}v_i\Big|^2 \\
&= \Big(1-\sum_{i=0}^n t_i^2\Big)+\frac{1}{n}\Big(1-\sum_{i=0}^n t_i^2\Big)\left(1+(n+1)|c|^2\right) \\
&= \Big(1-\sum_{i=0}^n t_i^2\Big)\Big(\frac{n+1}{n}\Big)\left(1-|c|^2\right).
\end{aligned}$$

Therefore

$$\frac{\mathrm{Vol}_E(\Delta_n)}{\mathrm{Vol}_E(\Delta^n)}\int_{\Delta^n} f(1-|x|^2)dV \leq \int_{\Delta_n} f\Big(\Big(\frac{n+1}{n}\Big)(1-|c|^2)\Big(1-\sum_{i=0}^n t_i^2\Big)\Big)dS$$

with equality if Δ^n is regular.

We now apply this last equality to $g(t) = f((1-|c|^2)t)$ and Δ^n_*. Then we have

$$\begin{aligned}&\frac{\mathrm{Vol}_E(\Delta_n)}{\mathrm{Vol}_E(\Delta^n_*)}\int_{\Delta^n_*} f((1-|c|^2)(1-|x|^2))dV \\ &= \int_{\Delta_n} f\Big((1-|c|^2)\Big(\frac{n+1}{n}\Big)\Big(1-\sum_{i=0}^n t_i^2\Big)\Big)dS.\end{aligned}$$

Therefore, we have

$$\int_{\Delta^n} f(1-|x|^2)dV \leq \frac{\mathrm{Vol}_E(\Delta^n)}{\mathrm{Vol}_E(\Delta^n_*)}\int_{\Delta^n_*} f((1-|c|^2)(1-|x|^2))dV.$$

Now assume that we have equality and f is strictly concave. Then

$$\frac{1}{(n+1)!}\sum_\sigma f\Big(1-\Big|\sum_{i=0}^n t_{\sigma(i)}v_i\Big|^2\Big) = f\Big(\frac{1}{(n+1)!}\sum_\sigma\Big(1-\Big|\sum_{i=0}^n t_{\sigma(i)}v_i\Big|^2\Big)\Big)$$

for all $(t_0,\ldots,t_n)$ in Δ_n. Therefore

$$\Big|\sum_{i=0}^n t_{\sigma(i)}v_i\Big| = \Big|\sum_{i=0}^n t_i v_i\Big|$$

for all $(t_0,\ldots,t_n)$ in Δ_n and all σ. Let $t_0 = t_1 = 1/2$ and $t_i = 0$ for $i > 1$. Then we find that

$$|v_0+v_1| = |v_i+v_j| \quad \text{for all } i \neq j.$$

Hence, we have

$$v_0\cdot v_1 = v_i\cdot v_j \quad \text{for all } i\neq j.$$

Therefore, we have

$$|v_0-v_1| = |v_i-v_j| \quad \text{for all } i\neq j.$$

Consequently Δ^n is regular. □

Theorem 11.3.1. *An n-simplex Δ^n in B^n has maximal volume if and only if it is regular and ideal.*

Proof: The proof is by induction on n. The theorem is true for $n = 1, 2, 3$, so assume that $n \geq 3$ and the theorem is true in dimension n. Now consider Δ^{n+1}. We may assume that Δ^{n+1} is ideal. We pass to the upper half-space model U^{n+1} and position Δ^{n+1} as in Lemma 2. Let $\Delta^n = \nu(\Delta^{n+1})$.

Let Δ^n_* be a regular ideal n-simplex in D^n and let

$$k_n = n\mathrm{Vol}(\Delta^{n+1}_*)/\mathrm{Vol}(\Delta^n_*).$$

Define $f : (0,1] \to \mathbb{R}$ by

$$f(t) = t^{-n/2} - k_n t^{-(n+1)/2}.$$

Then

$$f''(t) = \left(\frac{n}{2}\right)\left(\frac{n+2}{2}\right)t^{-(n+4)/2} - k_n\left(\frac{n+1}{2}\right)\left(\frac{n+3}{2}\right)t^{-(n+5)/2}.$$

Hence

$$f''(t) < 0 \quad \text{if and only if} \quad t < \frac{k_n(n+1)(n+3)}{n(n+2)}.$$

Therefore, if $1 < \frac{k_n(n+1)(n+3)}{n(n+2)}$ or equivalently $k_n > \frac{n(n+2)}{(n+1)(n+3)}$, then f is strictly concave. By Lemma 6, we have

$$k_n \geq (n-1)/n.$$

Now observe that

$$\frac{(n-1)}{n} \geq \frac{n(n+2)}{(n+1)(n+3)}$$

if and only if $n^2 - n > 3$, which is the case, since $n \geq 3$. Thus f is strictly concave.

For ease of notation, set

$$\ell_n = \frac{\mathrm{Vol}_E(\Delta^n)}{\mathrm{Vol}_E(\Delta^n_*)}.$$

We now apply Lemma 7 to f and Δ^n. By Lemmas 1 and 2, we have

$$\begin{aligned}
&n\mathrm{Vol}(\Delta^{n+1}) - k_n\mathrm{Vol}(\Delta^n)\\
&= \int_{\Delta^n} f(1-|x|^2)dV\\
&\leq \ell_n \int_{\Delta^n_*} f\big((1-|c|^2)(1-|x|^2)\big)dV\\
&= \ell_n(1-|c|^2)^{-n/2} n\mathrm{Vol}(\Delta^{n+1}_*) - \ell_n k_n(1-|c|^2)^{-(n+1)/2}\mathrm{Vol}(\Delta^n_*)\\
&= \ell_n(1-|c|^2)^{-n/2}\big[n\mathrm{Vol}(\Delta^{n+1}_*) - k_n(1-|c|^2)^{-1/2}\mathrm{Vol}(\Delta^n_*)\big]\\
&\leq \ell_n(1-|c|^2)^{-n/2}\big[n\mathrm{Vol}(\Delta^{n+1}_*) - k_n\mathrm{Vol}(\Delta^n_*)\big]\\
&= 0.
\end{aligned}$$

By the induction hypothesis, we have

$$\mathrm{Vol}(\Delta^n) \leq \mathrm{Vol}(\Delta^n_*)$$

and so

$$n\mathrm{Vol}(\Delta^{n+1}) \leq k_n\mathrm{Vol}(\Delta^n) \leq k_n\mathrm{Vol}(\Delta^n_*) = n\mathrm{Vol}(\Delta^{n+1}_*).$$

Thus $\mathrm{Vol}(\Delta^{n+1}_*)$ is maximal. If $\mathrm{Vol}(\Delta^{n+1}) = \mathrm{Vol}(\Delta^{n+1}_*)$, then we have by Lemma 7 that Δ^n is Euclidean regular and therefore Δ^{n+1} is regular by Lemma 3. □

Theorem 11.3.2. *The hyperbolic volume of a generalized n-simplex Δ^n in D^n is a continuous function of its vertices.*

Proof: For each positive integer j, let Δ_j^n be a generalized n-simplex in D^n with vertices $v_{0j}, \ldots, v_{nj}$ such that $(v_{0j}, \ldots, v_{nj}) \to (v_0, \ldots, v_n)$ where $v_0, \ldots, v_n$ are the vertices of a generalized n-simplex Δ^n in D^n. We need to prove that

$$\lim_{j\to\infty} \mathrm{Vol}(\Delta_j^n) = \mathrm{Vol}(\Delta^n).$$

Assume first that Δ_j^n is ideal for each j. Then Δ^n is ideal. This part of the proof is by induction on the dimension n. There is nothing to prove in dimension one, since D^1 is the only ideal 1-simplex in D^1. In dimension two, all ideal 2-simplices are congruent, and so the theorem is true in this case. Assume that $n > 2$ and this part of the theorem is true in dimension $n - 1$.

For each j, let A_j be the rotation of E^n that rotates v_{0j} to v_0 with no other nonzero angles of rotation. As $v_{0j} \to v_0$, we have that $A_j \to I$ in $\mathrm{O}(n)$. Hence $(A_j v_{0j}, \ldots, A_j v_{nj}) \to (v_0, \ldots, v_n)$. As

$$\mathrm{Vol}(A_j(\Delta_j^n)) = \mathrm{Vol}(\Delta_j^n),$$

we may replace Δ_j^n by $A_j(\Delta_j^n)$. Thus, we may assume, without loss of generality, that $v_{0j} = v_0$ for all j.

We now pass to the upper half-space model U^n of hyperbolic space and assume, without loss of generality, that $v_0 = \infty$ and $v_1, \ldots, v_n$ lie on S^{n-2} in E^{n-1}. For each j, the vertices $v_{1j}, \ldots, v_{nj}$ lie on an $(n-2)$-sphere $S(a_j, r_j)$ in E^{n-1}. Now as $(v_{1j}, \ldots, v_{nj}) \to (v_1, \ldots, v_n)$, we have that $a_j \to 0$ and $r_j \to 1$. Let

$$\phi_j = -r_j^{-1} a_j + r_j^{-1} I.$$

Then ϕ_j maps $S(a_j, r_j)$ onto S^{n-2}. Moreover $\phi_j \to I$ in $\mathrm{S}(E^{n-1})$. Hence $(\phi_j(v_{1j}), \ldots, \phi_j(v_{nj})) \to (v_1, \ldots, v_n)$. As

$$\mathrm{Vol}(\phi_j(\Delta_j^n)) = \mathrm{Vol}(\Delta_j^n),$$

we may replace Δ_j^n by $\phi_j(\Delta_j^n)$. Thus, we may assume, without loss of generality, that the vertices $v_1, \ldots, v_n$ lie on the sphere S^{n-2} for all j. By Lemma 2, we have

$$\mathrm{Vol}(\Delta^n) = \frac{1}{n-1} \int_{\nu(\Delta^n)} \frac{dx_1 \cdots dx_{n-1}}{(1 - |x|^2)^{(n-1)/2}},$$

where $\nu : U^n \to E^{n-1}$ is the vertical projection.

For each j, let χ_j be the characteristic function of the set $\nu(\Delta_j^n)$ and let χ be the characteristic function of $\nu(\Delta^n)$. Then $\{\chi_j\}$ converges to χ almost everywhere, and for each j, we have

$$\frac{\chi_j(x)}{(1 - |x|^2)^{(n-1)/2}} \leq \frac{\chi_j(x)}{(1 - |x|^2)^{n/2}}.$$

By the induction hypothesis, we have

$$\lim_{j\to\infty}\int_{D^{n-1}}\frac{\chi_j(x)dV}{(1-|x|^2)^{n/2}}=\int_{D^{n-1}}\frac{\chi(x)dV}{(1-|x|^2)^{n/2}}\ <\ \infty.$$

By the dominated convergence theorem, we deduce that

$$\lim_{j\to\infty}\int_{D^{n-1}}\frac{\chi_j(x)dV}{(1-|x|^2)^{(n-1)/2}}=\int_{D^{n-1}}\frac{\chi(x)dV}{(1-|x|^2)^{(n-1)/2}}.$$

Therefore

$$\lim_{j\to\infty}\mathrm{Vol}(\Delta_j^n)=\mathrm{Vol}(\Delta^n).$$

We now return to the general case. Without loss of generality, we may assume that 0 is the centroid of Δ^n. Then every vertex of Δ^n is nonzero. As the vertices of Δ_j^n converge to the vertices of Δ^n, we may assume that every vertex of Δ_j^n is nonzero for each j. For each j, let $\hat{\Delta}_j^n$ be the ideal n-simplex with vertices $\hat{v}_{0j},\ldots,\hat{v}_{nj}$, where $\hat{v}_{ij}=v_{ij}/|v_{ij}|$, and let $\hat{\Delta}^n$ be the ideal n-simplex with vertices $\hat{v}_0,\ldots,\hat{v}_n$, where $\hat{v}_i=v_i/|v_i|$. Then $(\hat{v}_{0j},\ldots,\hat{v}_{nj})\to(\hat{v}_0,\ldots,\hat{v}_n)$. Let $\chi_j,\hat{\chi}_j,\chi,\hat{\chi}$ be the characteristic functions for the sets $\Delta_j^n,\hat{\Delta}_j^n,\Delta^n,\hat{\Delta}^n$, respectively. Then $\chi_j\to\chi$ and $\hat{\chi}_j\to\hat{\chi}$ almost everywhere. Now as $\Delta_j^n\subset\hat{\Delta}_j^n$, we have that $\chi_j\leq\hat{\chi}_j$ for each j. Let

$$d\mu=\frac{dV}{(1-|x|^2)^{(n+1)/2}}$$

be the element of hyperbolic volume of D^n. By the first case, we have

$$\lim_{j\to\infty}\int_{D^n}\hat{\chi}_j d\mu=\int_{D^n}\hat{\chi}d\mu\ <\ \infty.$$

By the dominated convergence theorem, we deduce that

$$\lim_{j\to\infty}\int_{D^n}\chi_j d\mu=\int_{D^n}\chi d\mu.$$

Therefore, we have

$$\lim_{j\to\infty}\mathrm{Vol}(\Delta_j^n)=\mathrm{Vol}(\Delta^n).$$

□

Exercise 11.3

1. Prove that the edge length of a regular Euclidean n-simplex inscribed in S^{n-1} is $\sqrt{2(1+1/n)}$.
2. Prove that the volume of a regular Euclidean n-simplex inscribed in S^{n-1} is
$$\frac{(n+1)^{\frac{1}{2}}}{n!}\left(1+\frac{1}{n}\right)^{n/2}.$$
3. Let Δ^n be a Euclidean n-simplex inscribed in S^{n-1} and let Δ_*^n be a regular Euclidean n-simplex inscribed in S^{n-1}. Prove that
$$\mathrm{Vol}_E(\Delta^n)\leq\mathrm{Vol}_E(\Delta_*^n)$$
with equality if and only if Δ^n is regular.

4. Prove that a regular ideal 4-simplex in B^4 has volume

$$\frac{10\pi}{3}\arcsin\left(\frac{1}{3}\right) - \frac{\pi^2}{3} = .26889\ldots.$$

5. Fill in the details in the proof of Lemma 7 that

$$\sum_{\sigma} t_{\sigma(i)}t_{\sigma(j)} = (n-1)!\sum_{k\neq\ell} t_k t_\ell.$$

§11.4. The Gromov Invariant

In this section, we consider the Gromov invariant of a closed, orientable, hyperbolic manifold. As an application, we prove that two homotopy equivalent, closed, orientable, hyperbolic manifolds have the same volume.

Let X be a topological space and let $\mathrm{S}(X;\mathbb{R})$ be the *singular chain complex* of X with real coefficients. For each integer $k \geq 0$, the group of singular k-chains $\mathrm{S}_k(X;\mathbb{R})$ is a real vector space with a basis consisting of all continuous maps from the standard k-simplex Δ^k to X. Recall that a continuous map $\sigma : \Delta^k \to X$ is called a *singular k-simplex* in X.

Let c be a k-chain in $\mathrm{S}_k(X;\mathbb{R})$. Then for each singular k-simplex σ in X, there is a unique real number r_σ such that

$$c = \sum_{\sigma} r_\sigma \sigma.$$

Here $r_\sigma = 0$ for all but finitely many σ. The *simplicial norm* of c is defined to be the real number

$$\|c\| = \sum_{\sigma} |r_\sigma|.$$

If α is a homology class in $\mathrm{H}_k(X;\mathbb{R})$, the *simplicial norm* of α is defined to be the real number

$$\|\alpha\| = \inf\{\|c\| : c \text{ is a } k\text{-cycle representing } \alpha\}.$$

If α and β are in $\mathrm{H}_k(X;\mathbb{R})$ and t is in $\mathbb{R}$, then obviously

(1) $\|t\alpha\| = |t|\ \|\alpha\|$,

(2) $\|\alpha+\beta\| \leq \|\alpha\| + \|\beta\|$.

Lemma 1. *If $f : X \to Y$ is a continuous function and α is a homology class in $\mathrm{H}_k(X;\mathbb{R})$, then $\|f_*(\alpha)\| \leq \|\alpha\|$.*

Proof: Let c be a k-cycle representing α and write $c = \sum_{\sigma} r_\sigma\sigma$ as before. Then the homology class $f_*(\alpha)$ in $\mathrm{H}_k(Y;\mathbb{R})$ is represented by $f_*(c)$, where

$$f_*(c) = \sum_{\sigma} r_\sigma f\sigma.$$

As the maps $f\sigma : \Delta^k \to Y$ are not necessarily distinct, we have

$$\|f_*(c)\| \leq \sum_\sigma |r_\sigma| = \|c\|.$$

Therefore $\|f_*(\alpha)\| \leq \|\alpha\|$. □

Definition: The *Gromov invariant* of a closed, connected, orientable n-manifold M is the simplicial norm of a fundamental class of M in $\mathrm{H}_n(M;\mathbb{R})$. The Gromov invariant of M is denoted by $\|M\|$.

Theorem 11.4.1. *If M is a closed, connected, orientable, spherical or Euclidean n-manifold, with $n > 0$, then $\|M\| = 0$.*

Proof: Assume first that $M = S^n$ or T^n. Then M admits a map $f : M \to M$ of degree two. By Lemma 1, we have

$$(\deg f)\|M\| \leq \|M\|.$$

Consequently $\|M\| = 0$.

Now assume that M is arbitrary. Then M is finitely covered by $\tilde{M} = S^n$ or T^n. Let $\pi : \tilde{M} \to M$ be the covering projection. By Lemma 1, we have

$$(\deg \pi)\|M\| \leq \|\tilde{M}\| = 0.$$

As the degree of π is the order of the covering, we have that $\deg \pi \geq 1$ and so $\|M\| = 0$. □

Remark: Since the simplicial norm of a nonzero homology class may be zero, the simplicial norm on real singular homology is technically not a norm but only a pseudonorm.

Straight Singular k-Simplices

The *standard k-simplex* Δ^k is the k-simplex in E^n spanned by the vectors $0 = e_0, e_1, \ldots, e_k$. Let x be a point of Δ^k. Then we have

$$x = \sum_{i=1}^{k} x_i e_i$$

with $0 \leq x_i \leq 1$ for each i and $\sum_{i=1}^{k} x_i \leq 1$. Set

$$x_0 = 1 - \sum_{i=1}^{k} x_i.$$

Then $x_0, \ldots, x_k$ are the barycentric coordinates of x and we have

$$x = \sum_{i=0}^{k} x_i e_i.$$

A singular k-simplex σ in H^n is said to be *straight* if and only if for each x in Δ^k, we have

$$\sigma(x) = \sum_{i=0}^{k} x_i\sigma(e_i) \Big/ \left|\left|\left| \sum_{i=0}^{k} x_i\sigma(e_i) \right|\right|\right|.$$

The image of a straight singular k-simplex σ is the convex hull in H^n of the points $\sigma(e_0), \ldots, \sigma(e_k)$; moreover, σ is uniquely determined by the sequence of points $\sigma(e_0), \ldots, \sigma(e_k)$; furthermore, if g is an isometry of H^n, then $g\sigma$ is also a straight singular k-simplex.

Let $M = H^n/\Gamma$ be a hyperbolic space-form. A singular k-simplex σ in M is said to be *straight* if and only if σ lifts to a straight singular k-simplex $\tilde{\sigma}$ in H^n. By the previous remark, if some lift of σ is straight, then every lift of σ is straight, since any two lifts of σ differ by an element of Γ.

Given a singular k-simplex σ in M, we can associate to σ a straight singular k-simplex $\mathrm{Str}(\sigma)$ as follows: First lift σ to a singular k-simplex $\tilde{\sigma}$ in H^n. Let $\mathrm{Str}(\tilde{\sigma})$ be the unique straight singular k-simplex determined by the sequence of points $\tilde{\sigma}(e_0), \ldots, \tilde{\sigma}(e_k)$. Now let $\mathrm{Str}(\sigma) = \pi\ \mathrm{Str}(\tilde{\sigma})$ where $\pi : H^n \to M$ is the quotient map. Then $\mathrm{Str}(\sigma)$ is a straight singular k-simplex, and $\mathrm{Str}(\sigma)$ does not depend on the choice of the lift $\tilde{\sigma}$, since any two lifts of σ differ by an element of Γ.

The straightening operator Str on singular k-simplices in M extends linearly to a linear transformation

$$\mathrm{Str}_k : \mathrm{S}_k(M;\mathbb{R}) \to \mathrm{S}_k(M;\mathbb{R}).$$

Furthermore, since

$$\mathrm{Str}_{k-1}\partial_k = \partial_k \mathrm{Str}_k$$

for all k, we have that $\mathrm{Str} = \{\mathrm{Str}_k\}$ is a chain map.

Lemma 2. *The straightening chain map* $\mathrm{Str} : \mathrm{S}(M;\mathbb{R}) \to \mathrm{S}(M;\mathbb{R})$ *is chain homotopic to the identity.*

Proof: Let σ be a singular k-simplex in M. Lift σ to a singular k-simplex $\tilde{\sigma}$ in H^n. Since H^n is convex, there is a canonical homotopy

$$F_{\tilde{\sigma}} : \Delta^k \times [0,1] \to H^n$$

from $\tilde{\sigma}$ to $\mathrm{Str}(\tilde{\sigma})$ defined by

$$F_{\tilde{\sigma}}(x,t) = \frac{(1-t)\tilde{\sigma}(x) + t\mathrm{Str}(\tilde{\sigma}(x))}{|||(1-t)\tilde{\sigma}(x) + t\mathrm{Str}(\tilde{\sigma}(x))|||}.$$

If g is an isometry of H^n, then $F_{g\tilde{\sigma}} = gF_{\tilde{\sigma}}$. Therefore $F_{\tilde{\sigma}}$ projects to a homotopy

$$F_\sigma : \Delta^k \times [0,1] \to M$$

from σ to $\mathrm{Str}(\sigma)$ that does not depend on the choice of the lift $\tilde{\sigma}$.

Now $\Delta^k \times [0,1]$ has vertices

$$a_0 = (e_0, 0), \ldots, a_k = (e_k, 0), \ b_0 = (e_0, 1), \ldots, b_k = (e_k, 1).$$

For each $i = 0, \ldots, k$, let

$$\alpha_i : \Delta^{k+1} \to \Delta^k \times [0,1]$$

be the affine map that maps $e_0, \ldots, e_{k+1}$ to $a_0, \ldots, a_i, \ b_i, \ldots, b_k$, respectively. Define a linear transformation

$$F_k : \mathrm{S}_k(M; \mathbb{R}) \to \mathrm{S}_{k+1}(M; \mathbb{R})$$

by the formula

$$F_k(\sigma) = \sum_{i=0}^{k} (-1)^i F_\sigma \alpha_i.$$

A straightforward calculation shows that

$$\partial_{k+1} F_k(\sigma) + F_{k-1} \partial_k(\sigma) = \mathrm{Str}_k(\sigma) - \sigma.$$

Therefore, we have

$$\partial_{k+1} F_k + F_{k-1} \partial_k = \mathrm{Str}_k - id_k.$$

Thus $F = \{F_k\}$ is a chain homotopy from Str to the identity. □

Let $\mathrm{Str}_k(M; \mathbb{R})$ be the set of all straight singular k-chains in M. Then $\mathrm{Str}(M; \mathbb{R})$ is a chain subcomplex of $\mathrm{S}(M; \mathbb{R})$.

Theorem 11.4.2. *If M is a hyperbolic space-form, then the inclusion chain map*

$$i : \mathrm{Str}(M; \mathbb{R}) \to \mathrm{S}(M; \mathbb{R})$$

induces an isomorphism on homology.

Proof: The straightening chain map $\mathrm{Str} : \mathrm{S}(M; \mathbb{R}) \to \mathrm{Str}(M; \mathbb{R})$ is a chain homotopy inverse of i by Lemma 2. □

Remark: It follows from Theorem 11.4.2 that one can compute the real homology of a hyperbolic space-form M using only straight singular chains in M. Moreover, if c is any singular chain in M, then $\|\mathrm{Str}(c)\| \leq \|c\|$, and so one can also compute the simplicial norm of a real homology class of M using only straight singular cycles.

Lemma 3. *Let M be a compact, oriented, hyperbolic space-form H^n/Γ, with $n > 1$, and let V_n be the volume of a regular ideal n-simplex in H^n. Then*

$$\|M\| \geq \mathrm{Vol}(M)/V_n.$$

Proof: Let Ω_M be the hyperbolic volume form for M and let $c = \sum_\sigma r_\sigma \sigma$ be any straight singular n-cycle representing the fundamental class of M in $\mathrm{H}_n(M;\mathbb{R})$. We claim that

$$\int_c \Omega_M = \mathrm{Vol}(M).$$

First we show that the integral $\int_c \Omega_M$ depends only on the homology class of c. Let c' be any singular n-cycle homologous to c. Then there is singular $(n+1)$-chain b such that

$$c - c' = \partial b.$$

By Stokes's theorem, we have

$$\int_c \Omega_M - \int_{c'} \Omega_M = \int_{\partial b} \Omega_M = \int_b d\Omega_M = 0,$$

since $d\Omega_M = 0$. Thus $\int_c \Omega_M$ depends only on the homology class of c.

Let P be an exact, convex, fundamental polyhedron for Γ. Then P is compact by Theorem 6.5.10. Since P is exact, the barycentric subdivision of P projects to a subdivision of M into a finite number of hyperbolic n-simplices. Moreover, the second barycentric subdivision of P projects to a triangulation of M into a finite number of hyperbolic n-simplices $s_1, \ldots, s_m$ that barycentrically subdivides the first subdivision of M. For each $i = 1, \ldots, m$, let $\sigma_i : \Delta^n \to M$ be the straight singular n-simplex such that $\sigma_i(e_j)$ is the unique vertex of s_i contained in the jth skeleton of the first subdivision of M for each $j = 0, \ldots, n$, and let $r_i = 1$ or -1 according as σ_i is orientation preserving or not for each i. Then

$$c' = \sum_{i=1}^m r_i \sigma_i$$

is a straight singular n-cycle representing the fundamental class of M. Now

$$\int_{c'} \Omega_M = \sum_{i=1}^m r_i \int_{\sigma_i} \Omega_M = \sum_{i=1}^m r_i \int_{\Delta^n} \sigma_i^*(\Omega_M),$$

and for each i, we have

$$r_i \int_{\Delta^n} \sigma_i^*(\Omega_M) = \mathrm{Vol}(s_i).$$

Therefore, we have

$$\int_{c'} \Omega_M = \sum_{i=1}^m \mathrm{Vol}(s_i) = \mathrm{Vol}(M).$$

Thus, we have

$$\int_c \Omega_M = \int_{c'} \Omega_M = \mathrm{Vol}(M).$$

Next observe that

$$\begin{aligned}\int_c \Omega_M &= \sum_\sigma r_\sigma \int_{\Delta^n} \sigma^*(\Omega_M)\\ &= \sum_\sigma \pm r_\sigma \mathrm{Vol}(\tilde{\sigma}(\Delta^n))\\ &\leq \sum_\sigma |r_\sigma| \mathrm{Vol}(\tilde{\sigma}(\Delta^n)).\end{aligned}$$

Now by Theorem 11.3.1, we have

$$\mathrm{Vol}(\tilde{\sigma}(\Delta^n)) < V_n.$$

Therefore, we have

$$\mathrm{Vol}(M) = \int_c \Omega_M < \sum_\sigma |r_\sigma| V_n.$$

Dividing by V_n, we obtain the inequality

$$\mathrm{Vol}(M)/V_n < \|c\|.$$

Therefore, we deduce that

$$\mathrm{Vol}(M)/V_n \leq \|M\|. \qquad \square$$

Haar Measure

Let $G = \mathrm{I}(H^n)$ and let H be the subgroup of G of all elements that fix the point e_{n+1}. The left-invariant *Haar integral* of a function $\phi : G \to \mathbb{R}$ is given by the formula

$$\int_G \phi(g)dg = \int_{G/H} \left(\int_H \phi(gh)dh \right) d(gH),$$

where dh is the left-invariant *Haar measure* on the compact group H and $d(gH)$ is the left-invariant measure on G/H corresponding to hyperbolic volume in H^n under the homeomorphism from G/H to H^n given by Theorems 5.1.5 and 5.2.9. The Haar measure on a topological group is unique up to multiplication by a positive scalar. We shall normalize the Haar measure dg on G by normalizing the Haar measure dh on H so that

$$\int_H dh = 1.$$

Lemma 4. *Let x be a point of H^n, let R be an open (resp. closed) subset of H^n, and let*

$$S = \{g \in \mathrm{I}(H^n) : gx \in R\}.$$

Then S is open (resp. closed) and the Haar measure of S is the volume of the set R.

Proof: Assume first that $x = e_{n+1}$. As the evaluation map

$$\varepsilon : \mathrm{I}(H^n) \to H^n,$$

defined by $\varepsilon(g) = ge_{n+1}$, is continuous, $S = \varepsilon^{-1}(R)$ is open (resp. closed). Let χ_S be the characteristic function of the set S. Then

$$\begin{aligned}
\mathrm{Vol}(S) &= \int_G \chi_S(g)dg \\
&= \int_{G/H}\left(\int_H \chi_S(gh)dh\right)d(gH) \\
&= \int_{G/H} \chi_{S/H}(gH)d(gH) \quad = \quad \mathrm{Vol}(R).
\end{aligned}$$

Now let x be an arbitrary point of H^n. Set

$$S_0 = \{g \in \mathrm{I}(H^n) : ge_{n+1} \in R\}$$

and let f be an isometry of H^n such that $fx = e_{n+1}$. Then $S = S_0 f$. Hence S is open (resp. closed). It is a basic fact of the theory of Haar measure that the Haar measure on a group is both left- and right-invariant if the abelianization of the group is finite. Consequently, the Haar measure on $\mathrm{I}(H^n)$ is both left- and right-invariant because of Theorem 5.5.12. Therefore

$$\mathrm{Vol}(S) = \mathrm{Vol}(S_0 f) = \mathrm{Vol}(S_0) = \mathrm{Vol}(R). \qquad \square$$

Theorem 11.4.3. (Gromov's theorem) *Let M be a closed, connected, oriented, hyperbolic n-manifold, with $n > 1$, and let V_n be the volume of a regular ideal n-simplex in H^n. Then*

$$\|M\| = \mathrm{Vol}(M)/V_n.$$

Proof: Since M is complete, we may assume that M is a space-form H^n/Γ. Let P be a convex fundamental polyhedron for Γ. Then P is compact by Theorem 6.5.10. Choose a point x_0 in P° and let $u_0 = \pi(x_0)$ where $\pi : H^n \to H^n/\Gamma$ is the quotient map.

Let $\sigma : \Delta^n \to M$ be a straight singular n-simplex such that $\sigma(e_i) = u_0$ for each i. Then σ lifts to a unique straight singular n-simplex $\tilde{\sigma} : \Delta^n \to H^n$ such that $\tilde{\sigma}(e_0) = x_0$. As $\pi\tilde{\sigma}(e_i) = u_0$ for each i, we have that $\tilde{\sigma}(e_i)$ is in the Γ-orbit of x_0 for each i. Hence, there is a unique element f_i of Γ, with $f_0 = 1$, such that $\tilde{\sigma}(e_i) = f_i x_0$ for each i.

Given $\ell > 0$, choose points $x_1, \ldots, x_n$ of H^n such that $x_0, \ldots, x_n$ are the vertices of a regular n-simplex Δ^n_ℓ in H^n whose edge length is ℓ. For each $i = 0, \ldots, n$, let

$$S_i = \{g \in \mathrm{I}(H^n) : gx_i \in f_i(P^\circ)\}.$$

By Lemma 4, the set S_i is open and $\mathrm{Vol}(S_i) = \mathrm{Vol}(P)$. Let

$$S_\sigma = S_0 \cap \cdots \cap S_n.$$

Then S_σ is open and $\mathrm{Vol}(S_\sigma) \leq \mathrm{Vol}(P)$. As P is compact, $\mathrm{Vol}(P)$ is finite and therefore $\mathrm{Vol}(S_\sigma)$ is finite.

Suppose that g is in S_σ. Then $gx_\imath$ is in $f_\imath(P^\circ)$ for each $i = 0, \ldots, n$ and so

$$\begin{aligned} d(x_0, f_\imath x_0) &\leq d(x_0, gx_0) + d(gx_0, gx_\imath) + d(gx_\imath, f_\imath x_0) \\ &< \ell + 2\,\mathrm{diam}(P). \end{aligned}$$

Let $r = \ell + 2\,\mathrm{diam}(P)$. As $B(x_0, r)$ contains only finitely many elements of Γx_0, there are only finitely many σ such that the set S_σ is nonempty.

Suppose that S_σ is nonempty. Then if g is in S_σ, we have

$$d(\tilde{\sigma}(e_\imath), gx_\imath) < \mathrm{diam}(P)$$

for each $i = 0, \ldots, n$. Hence, the vertices of $\tilde{\sigma}(\Delta^n)$ are within a fixed distance from the corresponding vertices of the regular n-simplex $g\Delta^n_\ell$. By choosing ℓ sufficiently large, we may assume that $\tilde{\sigma}(\Delta^n)$ is a nondegenerate n-simplex in H^n.

For each σ, let $r_\sigma = \pm\mathrm{Vol}(S_\sigma)$ with the plus or minus sign according as σ preserves or reverses orientation. Set

$$c_\ell = \sum_\sigma r_\sigma \sigma.$$

Then c_ℓ is a straight singular n-chain in M.

For each $i = 0, \ldots, n$, let

$$T_\imath = \{g \in \mathrm{I}(H^n) : gx_\imath \in \Gamma\partial P\}.$$

By Lemma 4, the set $T_\imath$ is closed and

$$\mathrm{Vol}(T_\imath) = \mathrm{Vol}(\Gamma\partial P) = 0.$$

Now let

$$T = T_0 \cup \cdots \cup T_n.$$

Then T is closed and $\mathrm{Vol}(T) = 0$.

Suppose that g is in $S_0 - T$. Then there exists a unique element $f_\imath$ of Γ, with $f_0 = 1$, such that $gx_\imath$ is in $f_\imath P^\circ$ for each $i = 1, \ldots, n$. Let $\tilde{\sigma} : \Delta^n \to H^n$ be the straight singular n-simplex such that $\tilde{\sigma}(e_\imath) = f_\imath x_0$ for each i. Let $\sigma = \pi\tilde{\sigma}$. Then g is in S_σ. Consequently, we have

$$S_0 - T = \bigcup_\sigma S_\sigma.$$

Moreover, the sets $\{S_\sigma\}$ are pairwise disjoint. Therefore, we have

$$\mathrm{Vol}(S_0) = \mathrm{Vol}(S_0 - T) = \sum_\sigma \mathrm{Vol}(S_\sigma) = \sum_\sigma |r_\sigma|.$$

Hence, we have

$$\|c_\ell\| = \sum_\sigma |r_\sigma| = \mathrm{Vol}(S_0) = \mathrm{Vol}(P) = \mathrm{Vol}(M).$$

Now let $\tilde{\sigma} : \Delta^n \to H^n$ be an arbitrary, nondegenerate, straight, singular n-simplex such that $\tilde{\sigma}(e_i) = f_i x_0$ for some f_i in Γ for each $i = 0, \ldots, n$. Let

$$S_{\tilde{\sigma}} = \{g \in \mathrm{I}(H^n) : gx_i \in f_i(P^\circ) \text{ for } i = 0, \ldots, n\}$$

and let $r_{\tilde{\sigma}} = \pm\mathrm{Vol}(S_{\tilde{\sigma}})$ with the plus or minus sign according as $\pi\tilde{\sigma}$ preserves or reverses orientation. If f is in Γ, then $fS_{\tilde{\sigma}} = S_{f\tilde{\sigma}}$ and so $r_{f\tilde{\sigma}} = r_{\tilde{\sigma}}$. Thus, the infinite chain

$$\tilde{c}_\ell = \sum_{\tilde{\sigma}} r_{\tilde{\sigma}} \tilde{\sigma}$$

is Γ-equivariant. Now for each $\tilde{\sigma}$, there is an f in Γ such that $f\tilde{\sigma}(e_0) = x_0$. Therefore, we have

$$r_{\tilde{\sigma}} = r_{f\tilde{\sigma}} = r_{\pi(\tilde{\sigma})}.$$

Hence, the chain $\tilde{c}_\ell$ is the infinite chain in H^n that covers the chain c_ℓ in M. Therefore $\tilde{c}_\ell$ locally finite.

Now observe that

$$\partial\tilde{c}_\ell = \sum_{\tilde{\sigma}} r_{\tilde{\sigma}} \partial\tilde{\sigma}$$

is a locally finite chain. Hence, we have

$$\partial\tilde{c}_\ell = \sum_{\tau} s_\tau \tau,$$

where each τ is a straight singular $(n-1)$-simplex in H^n such that $\tau(e_i)$ is in Γx_0 for each i. For each such τ, let P_τ(resp. N_τ) be the set of all isometries of H^n that contribute positively (resp. negatively) to the coefficient s_τ of τ. Let ρ be the reflection of H^n in the hyperplane spanned by the image of τ. Then the symmetric difference of P_τ and ρN_τ is a subset of T. Hence P_τ and ρN_τ differ by a set of measure zero, and so $s_\tau = 0$. Therefore $\partial\tilde{c}_\ell = 0$. As $\partial\tilde{c}_\ell$ covers ∂c_ℓ, we deduce that $\partial c_\ell = 0$. Thus c_ℓ is a cycle.

Now since $\mathrm{H}_n(M;\mathbb{R})$ is generated by the fundamental class $[c]$ of M, there is a constant k_ℓ such that $[c_\ell] = k_\ell[c]$. Let Ω_M be the volume form of M. On the one hand,

$$\int_{c_\ell} \Omega_M = \int_{k_\ell c} \Omega_M = k_\ell \int_c \Omega_M = k_\ell \mathrm{Vol}(M)$$

and so

$$k_\ell = \frac{1}{\mathrm{Vol}(M)} \int_{c_\ell} \Omega_M.$$

On the other hand,

$$\int_{c_\ell} \Omega_M = \sum_\sigma r_\sigma \int_{\Delta^n} \sigma^*(\Omega_M) = \sum_\sigma |r_\sigma| \mathrm{Vol}(\tilde{\sigma}(\Delta^n)).$$

Let σ_ℓ be a simplex, with a nonzero coefficient in the sum $\sum r_\sigma \sigma$, such that $\tilde{\sigma}(\Delta^n)$ has least volume. Then

$$\begin{aligned} \int_{c_\ell} \Omega_M &\geq \Big(\sum_\sigma |r_\sigma|\Big) \mathrm{Vol}(\tilde{\sigma}_\ell(\Delta^n)) \\ &= \|c_\ell\| \mathrm{Vol}(\tilde{\sigma}_\ell(\Delta^n)) \\ &= \mathrm{Vol}(M)\mathrm{Vol}(\tilde{\sigma}_\ell(\Delta^n)). \end{aligned}$$

Hence, we have that

$$k_\ell \geq \mathrm{Vol}(\tilde{\sigma}_\ell(\Delta^n)).$$

Now as $[c_\ell/k_\ell]$ is the fundamental class of M, we deduce that

$$\|M\| \leq \|c_\ell\|/k_\ell \leq \mathrm{Vol}(M)/\mathrm{Vol}(\tilde{\sigma}_\ell(\Delta^n)).$$

Now there is an isometry g_ℓ of H^n such that $\tilde{\sigma}_\ell(e_i)$ is within a distance $\mathrm{diam}(P)$ from $g_\ell x_i$ for each $i = 0, \ldots, n$. Consequently

$$\lim_{\ell\to\infty} \mathrm{Vol}(\tilde{\sigma}_\ell(\Delta^n)) = V_n$$

by Theorem 11.3.2. Therefore

$$\|M\| \leq \mathrm{Vol}(M)/V_n.$$

As we have already established the reversed inequality in Lemma 3, the proof is complete. □

Theorem 11.4.4. *If M, N are homotopy equivalent, closed, connected, orientable, hyperbolic n-manifolds, with $n > 1$, then* $\mathrm{Vol}(M) = \mathrm{Vol}(N)$.

Proof: Let $f : M \to N$ be a homotopy equivalence and let $g : N \to M$ be a homotopy inverse of f. Let κ be a fundamental class of M. Then $f_*(\kappa)$ is a fundamental class of N and

$$g_*(f_*(\kappa)) = (gf)_*(\kappa) = \kappa.$$

Hence, by Lemma 1, we have

$$\|\kappa\| = \|g_*(f_*(\kappa))\| \leq \|f_*(\kappa)\| \leq \|\kappa\|.$$

Therefore, we have

$$\|M\| = \|\kappa\| = \|f_*(\kappa)\| = \|N\|.$$

Hence, by Theorem 11.4.3, we find that

$$\mathrm{Vol}(M) = \mathrm{Vol}(N).$$ □

Exercise 11.4

1. Explain why the proof of Lemma 3 breaks down in the spherical case where V_n is replaced by $\mathrm{Vol}(S^n)$.
2. Prove that the abelianization of $\mathrm{I}(H^n)$ has order two.
3. Let G be a topological group whose abelianization is finite. Prove that a left-invariant Haar measure on G is right-invariant.
4. Give a direct proof that the n-chain c_ℓ in the proof of Theorem 11.4.3 is a cycle.
5. Let M be a closed, connected, orientable, hyperbolic surface. Prove that $\|M\| = 2|\chi(M)|$.

§11.5. Measure Homology

In this section, we develop the theory of measure homology of a hyperbolic space-form. Let M be a differentiable manifold. For each integer $k \geq 0$, let $\mathrm{C}^\infty(\Delta^k, M)$ be the space of C^∞ singular k-simplices in M topologized with the C^1 *topology.* The C^1 topology is a larger topology than the compact-open topology that takes into account not only the proximity of functions but also of their first derivatives.

Let $\mathcal{C}_k(M)$ be the real vector space of all compactly supported, signed, *Borel measures* μ of bounded *total variation* $\|\mu\|$ on the space $\mathrm{C}^\infty(\Delta^k, M)$. Here

$$\|\mu\| = \int d\mu_+ + \int d\mu_-$$

where $\mu = \mu_+ - \mu_-$ is the *Jordan decomposition* of μ into its positive and negative variations.

For each $i = 0, \ldots, k$, let $\eta_i : \Delta^{k-1} \to \Delta^k$ be the ith face map. Then η_i induces a continuous function

$$\eta_i^* : \mathrm{C}^\infty(\Delta^k, M) \to \mathrm{C}^\infty(\Delta^{k-1}, M)$$

defined by $\eta_i^*(\sigma) = \sigma\eta_i$. Furthermore η_i^* induces a linear transformation

$$(\eta_i^*)_* : \mathcal{C}_k(M) \to \mathcal{C}_{k-1}(M)$$

defined by

$$\big((\eta_i^*)_*(\mu)\big)(B) = \mu\big((\eta_i^*)^{-1}(B)\big)$$

for each measure μ in $\mathcal{C}_k(M)$ and Borel subset B of $\mathrm{C}^\infty(\Delta^{k-1}, M)$. Define a linear transformation $\partial_k : \mathcal{C}_k(M) \to \mathcal{C}_{k-1}(M)$ by the formula

$$\partial_k = \sum_{i=0}^{k} (-1)^i (\eta_i^*)_*.$$

Lemma 1. *The system $\{\mathcal{C}_k(M), \partial_k\}$ is a chain complex.*

Proof: If $j < i$, then

$$\eta_i\eta_j = \eta_j\eta_{i-1}$$

and so we have

$$(\eta_j^*)_*(\eta_i^*)_* = (\eta_{i-1}^*)_*(\eta_j^*)_*.$$

With this identity, the usual calculation shows that $\partial_{k-1}\partial_k = 0$. □

The homology of the chain complex $\mathcal{C}(M) = \{\mathcal{C}_k(M), \partial_k\}$ is called the *measure homology* of M. Let $\mathrm{S}^\infty(M)$ be the subchain complex of $\mathrm{S}(M;\mathbb{R})$ of C^∞ singular chains in M. It is a basic fact of differential topology that the inclusion chain map from $\mathrm{S}^\infty(M)$ into $\mathrm{S}(M;\mathbb{R})$ induces an isomorphism on homology.

Given a C^∞ singular k-simplex $\sigma : \Delta^k \to M$, define an *atomic* Borel measure μ_σ on $C^\infty(\Delta^k, M)$ at σ by the formula

$$\mu_\sigma(B) = \begin{cases} 1 & \text{if } \sigma \text{ is in } B, \\ 0 & \text{otherwise.} \end{cases}$$

Define a linear transformation

$$m_k : S_k^\infty(M) \to \mathcal{C}_k(M)$$

by the formula

$$m_k\Big(\sum_\sigma r_\sigma \sigma\Big) = \sum_\sigma r_\sigma \mu_\sigma.$$

Lemma 2. *The family $\{m_k\}$ of linear transformations is a chain map from $S^\infty(M)$ to $\mathcal{C}(M)$.*

Proof: Let $\sigma : \Delta^k \to M$ be a C^∞ singular k-simplex. It suffices to show that

$$\partial m_k(\sigma) = m_{k-1}(\partial\sigma).$$

Observe that

$$\partial m_k(\sigma) = \partial\mu_\sigma = \sum_{i=0}^{k}(-1)^i(\eta_i^*)_*(\mu_\sigma),$$

whereas

$$m_{k-1}(\partial\sigma) = m_{k-1}\left(\sum_{i=0}^{k}(-1)^i\sigma\eta_i\right) = \sum_{i=0}^{k}(-1)^i\mu_{\sigma\eta_i}.$$

Moreover

$$\begin{aligned} (\eta_i^*)_*(\mu_\sigma)(B) &= \mu_\sigma\big((\eta_i^*)^{-1}(B)\big) \\ &= \begin{cases} 1 & \text{if } \sigma \text{ is in } (\eta_i^*)^{-1}(B), \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 1 & \text{if } \eta_i^*(\sigma) \text{ is in } B, \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 1 & \text{if } \sigma\eta_i \text{ is in } B, \\ 0 & \text{otherwise} \end{cases} \\ &= \mu_{\sigma\eta_i}(B). \end{aligned}$$

Thus, we have

$$\mu_{\sigma\eta_i} = (\eta_i^*)_*(\mu_\sigma).$$

Therefore, we have

$$\partial m_k(\sigma) = m_{k-1}(\partial\sigma).$$

□

Let $\Omega^k(M)$ be the real vector space of all C^∞ k-forms on M and let

$$d^k : \Omega^k(M) \to \Omega^{k+1}(M)$$

be the exterior differential. Then $\{\Omega^k(M), d^k\}$ is a cochain complex whose cohomology is the *de Rham cohomology* of M.

Let $\mathcal{D}_k(M)$ be the real vector space of all linear functionals on $\Omega^k(M)$. Define

$$\partial_k : \mathcal{D}_k(M) \to \mathcal{D}_{k-1}(M)$$

by the formula

$$(\partial_k f)(\omega) = f(d^{k-1}\omega).$$

Then $\{\mathcal{D}_k(M),\ \partial_k\}$ is a chain complex.

Lemma 3. *Let ω be a C^∞ k-form on M and let*

$$I_\omega : C^\infty(\Delta^k, M) \to \mathbb{R}$$

be the function defined by

$$I_\omega(\sigma) = \int_\sigma \omega.$$

Then I_ω is continuous.

Proof: For each point u of M, let $T(M,u)$ be the tangent space of M at u, and let $\Lambda^k(T(M,u))$ be the real vector space of all skew-symmetric k-linear functionals on $T(M,u)$. Set

$$\Lambda^k(T(M)) = \bigcup_{u\in M} \Lambda^k(T(M,u)).$$

Then $\Lambda^k(T(M))$ is a C^∞ vector bundle over M. A C^∞ k-form ω on M is a C^∞ section of this bundle.

Given σ in $C^\infty(\Delta^k, M)$, let

$$T(\sigma)^* : \Lambda^k(T(M)) \to \Lambda^k(T(\Delta^k))$$

be the induced map. Then

$$\int_\sigma \omega = \int_{\Delta^k} \sigma^*(\omega),$$

where $\sigma^*(\omega)$ is the C^∞ k-form on Δ^k defined by

$$\sigma^*(\omega) = T(\sigma)^* \circ \omega \circ \sigma.$$

Since the space $C^\infty(\Delta^k, M)$ is metrizable, we can prove the continuity of I_ω in terms of sequences. Suppose that $\sigma_i \to \sigma$ in $C^\infty(\Delta^k, M)$. Then we have that $T(\sigma_i) \to T(\sigma)$ in $C(T(\Delta^k), T(M))$, since $\sigma_i \to \sigma$ in the C^1 topology. Hence $T(\sigma_i)^* \to T(\sigma)^*$ in $C(\Lambda^k(T(M)), \Lambda^k(T(\Delta^k)))$. Since composition of maps is continuous with respect to the compact-open topology, we deduce that $\sigma_i^*(\omega) \to \sigma^*(\omega)$ in $C(\Delta^k, \Lambda^k(T(\Delta^k)))$.

Now we have

$$\sigma^*(\omega) = \sum_{i_1<\cdots<i_k} f_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

where $f_{i_1\cdots i_k}$ is in $C^\infty(\Delta^k, \mathbb{R})$. Likewise, we have

$$\sigma_i^*(\omega) = \sum_{i_1<\cdots<i_k} f^i_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

Then $f^i_{i_1\cdots i_k} \to f_{i_1\cdots i_k}$ in $C(\Delta^k, \mathbb{R})$ for each $i_1 \cdots i_k$. Hence $f^i_{i_1\cdots i_k} \to f_{i_1\cdots i_k}$ uniformly, since Δ^k is compact. Therefore

$$\lim_{i\to\infty} \int_{\Delta^k} f^i_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \int_{\Delta^k} f_{i_1\cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$$

by the Lebesgue dominated convergence theorem for each $i_1 \cdots i_k$. Hence

$$\lim_{i\to\infty} \int_{\Delta^k} \sigma_i^*(\omega) = \int_{\Delta^k} \sigma^*(\omega).$$

Thus I_ω is continuous. □

Let μ be a measure in $\mathcal{C}_k(M)$ and let K be the compact support of μ. Then the set $I_\omega(K)$ is bounded in $\mathbb{R}$ for each ω in $\Omega^k(M)$. As μ has bounded total variation, the integral $\int_K I_\omega d\mu$ is finite for each ω in $\Omega^k(M)$. Hence, we may define a linear functional

$$f_\mu : \Omega^k(M) \to \mathbb{R}$$

by the formula

$$f_\mu(\omega) = \int_{\sigma\in C^\infty(\Delta^k, M)} \left(\int_\sigma \omega\right) d\mu.$$

Define a linear transformation

$$\ell_k : \mathcal{C}_k(M) \to \mathcal{D}_k(M)$$

by $\ell_k(\mu) = f_\mu$.

Lemma 4. *The family $\{\ell_k\}$ of linear transformations is a chain map from $\mathcal{C}(M)$ to $\mathcal{D}(M)$.*

Proof: Let μ be a measure in $\mathcal{C}_k(M)$. Then

$$\begin{aligned}
\ell_{k-1}(\partial\mu) &= \ell_{k-1}\left(\sum_{i=0}^k (-1)^i (\eta_i^*)_*(\mu)\right) \\
&= \sum_{i=0}^k (-1)^i \ell_{k-1}\big((\eta_i^*)_*(\mu)\big) \\
&= \sum_{i=0}^k (-1)^i f_{(\eta_i^*)_*(\mu)}.
\end{aligned}$$

Now we have

$$\begin{aligned}
&\sum_{i=0}^{i}(-1)^i f_{(\eta_i^*)_*(\mu)}(\omega)\\
&= \sum_{i=0}^{k}(-1)^i \int_{\tau\in \mathrm{C}^\infty(\Delta^{k-1},M)} \left(\int_\tau \omega\right) d((\eta_i^*)_*(\mu))\\
&= \sum_{i=0}^{k}(-1)^i \int_{\sigma\in \mathrm{C}^\infty(\Delta^{k},M)} \left(\int_{\eta_i^*(\sigma)} \omega\right) d\mu\\
&= \int_{\sigma\in \mathrm{C}^\infty(\Delta^{k},M)} \left(\sum_{i=0}^{k}(-1)^i\int_{\sigma\eta_i} \omega\right) d\mu\\
&= \int_{\sigma\in \mathrm{C}^\infty(\Delta^{k},M)} \left(\int_{\partial\sigma} \omega\right) d\mu\\
&= \int_{\sigma\in \mathrm{C}^\infty(\Delta^{k},M)} \left(\int_{\sigma} d\omega\right) d\mu\\
&= f_\mu(d\omega)\\
&= \partial f_\mu(\omega).
\end{aligned}$$

Thus, we have

$$\ell_{k-1}(\partial\mu) = \partial\ell_k(\mu). \qquad \square$$

Theorem 11.5.1. *Let M be a differentiable manifold. Then the composition of the chain maps*

$$m_* : \mathrm{S}^\infty(M) \to \mathcal{C}(M) \quad\text{and}\quad \ell_* : \mathcal{C}(M) \to \mathcal{D}(M)$$

induces an isomorphism on homology.

Proof: Define a linear transformation

$$I^k : \Omega^k(M) \to \mathrm{Hom}(\mathrm{S}^\infty_k(M), \mathbb{R})$$

by the formula

$$(I^k(\omega))(c) = \int_c \omega.$$

Then $\{I^k\}$ is a cochain map that induces an isomorphism on cohomology by de Rham's theorem. By the universal coefficients theorem, the chain map

$$(I^*)^* : \mathrm{Hom}(\mathrm{Hom}(\mathrm{S}^\infty_*(M), \mathbb{R}), \mathbb{R}) \to \mathrm{Hom}(\Omega^*(M), \mathbb{R})$$

induces an isomorphism on homology. Consequently, the corresponding chain map

$$I_* : \mathrm{S}^\infty(M) \to \mathcal{D}(M)$$

induces an isomorphism on homology. Here

$$\begin{aligned}(I_k(c))(\omega) &= (c^*I^k)(\omega)\\ &= c^*(I^k(\omega))\\ &= (I^k(\omega))(c)\\ &= \int_c \omega.\end{aligned}$$

Given σ in $\mathrm{C}^\infty(\Delta^k, M)$, then

$$\ell_k m_k(\sigma) = \ell_k(\mu_\sigma) = f_{\mu_\sigma}.$$

Moreover

$$f_{\mu_\sigma}(\omega) = \int_{\tau\in \mathrm{C}^\infty(\Delta^k,M)} \left(\int_\tau \omega\right) d\mu_\sigma = \int_\sigma \omega,$$

since μ_σ is the atomic measure on $\mathrm{C}^\infty(\Delta^k, M)$ at σ. Therefore, we have that $\ell_* m_* = I_*$, and so $\ell_* m_*$ induces an isomorphism on homology. □

Straightening

Now assume that M is a hyperbolic space-form. Define a function

$$\mathrm{Str}^k : \mathrm{C}^\infty(\Delta^k, M) \to \mathrm{C}^\infty(\Delta^k, M)$$

by $\mathrm{Str}^k(\sigma) = \mathrm{Str}(\sigma)$.

Lemma 5. *The function* $\mathrm{Str}^k : \mathrm{C}^\infty(\Delta^k, M) \to \mathrm{C}^\infty(\Delta^k, M)$ *is continuous for each* k.

Proof: Let $\pi : H^n \to M$ be the quotient map. Then

$$\pi_* : \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{C}^\infty(\Delta^k, M)$$

is a continuous surjection, moreover π_* is an open map, since π is a local homeomorphism. Define a function

$$\widetilde{\mathrm{Str}}^k : \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{C}^\infty(\Delta^k, H^n)$$

by $\widetilde{\mathrm{Str}}^k(\sigma) = \mathrm{Str}(\sigma)$. As $\mathrm{Str}^k\pi_* = \pi_*\widetilde{\mathrm{Str}}^k$, it suffices to show that $\widetilde{\mathrm{Str}}^k$ is continuous.

The image of $\widetilde{\mathrm{Str}}^k$ is the set $\mathrm{Str}(\Delta^k, H^n)$ of straight singular k-simplices in H^n. The C^1 topology on $\mathrm{Str}(\Delta^k, H^n)$ is the same as the compact-open topology. Moreover, the function

$$\widetilde{\mathrm{Str}}^k : \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{Str}(\Delta^k, H^n)$$

is continuous with respect to the compact-open topology. Therefore $\widetilde{\mathrm{Str}}^k$ is continuous with respect to the C^1 topology, since the C^1 topology contains the compact-open topology. □

The continuous function

$$\mathrm{Str}^k : \mathrm{C}^\infty(\Delta^k, M) \to \mathrm{C}^\infty(\Delta^k, M)$$

induces a linear transformation

$$(\mathrm{Str}^k)_* : \mathcal{C}_k(M) \to \mathcal{C}_k(M)$$

defined by

$$\big((\mathrm{Str}^k)_*(\mu)\big)(B) = \mu\big((\mathrm{Str}^k)^{-1}(B)\big)$$

for each measure μ in $\mathcal{C}_k(M)$ and Borel subset B of $\mathrm{C}^\infty(\Delta^k, M)$.

Lemma 6. *The family $\{(\mathrm{Str}^k)_*\}$ of linear transformations is a chain map from $\mathcal{C}(M)$ to $\mathcal{C}(M)$.*

Proof: Observe that

$$\begin{aligned}
\partial_k(\mathrm{Str}^k)_* &= \sum_{i=0}^{k}(-1)^i(\eta_i^*)_*(\mathrm{Str}^k)_* \\
&= \sum_{i=0}^{k}(-1)^i(\eta_i^*\mathrm{Str}^k)_* \\
&= \sum_{i=0}^{k}(-1)^i(\mathrm{Str}^{k-1}\eta_i^*)_* \\
&= \sum_{i=0}^{k}(-1)^i(\mathrm{Str}^{k-1})_*(\eta_i^*)_* \quad = \quad (\mathrm{Str}^{k-1})_*\partial_k.
\end{aligned}$$

□

Theorem 11.5.2. *Let M be a hyperbolic space-form. Then the straightening chain map*

$$(\mathrm{Str}^*)_* : \mathcal{C}(M) \to \mathcal{C}(M)$$

is chain homotopic to the identity.

Proof: Given an element σ of $\mathrm{C}^\infty(\Delta^k, M)$, let $F_\sigma : \Delta^k \times [0,1] \to M$ be the homotopy from σ to $\mathrm{Str}(\sigma)$ constructed in Lemma 2 of §11.4. Define

$$F^k : \mathrm{C}^\infty(\Delta^k, M) \to \mathrm{C}^\infty(\Delta^k \times [0,1], M)$$

by $F^k(\sigma) = F_\sigma$. We claim that F^k is continuous. Define

$$\tilde{F}^k : \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{C}^\infty(\Delta^k \times [0,1], H^n)$$

by $\tilde{F}^k(\sigma) = F_{\tilde{\sigma}}$. Let $\pi : H^n \to M$ be the quotient map. Then

$$\begin{aligned}
\pi_* &: \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{C}^\infty(\Delta^k, M), \\
\pi_* &: \mathrm{C}^\infty(\Delta^k \times [0,1], H^n) \to \mathrm{C}^\infty(\Delta^k \times [0,1], M)
\end{aligned}$$

are continuous open surjections, since π is a local homeomorphism. As $\pi_* F^k = \tilde{F}^k \pi_*$, it suffices to show that $\tilde{F}^k$ is continuous.

The function
$$A : \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{C}^\infty(\Delta^k, H^n) \times \mathrm{C}^\infty(\Delta^k, H^n),$$
defined by the formula
$$A(\sigma) = (\sigma, \mathrm{Str}(\sigma)),$$
is continuous, since Str^k is continuous. The function
$$B : \mathrm{C}^\infty(\Delta^k, H^n) \times \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{C}^\infty(\Delta^k, H^n \times H^n),$$
defined by the formula
$$B(\sigma, \tau)(x) = (\sigma(x), \tau(x)),$$
is continuous. The function
$$C : \mathrm{C}^\infty(\Delta^k, H^n \times H^n) \to \mathrm{C}^\infty(\Delta^k \times [0,1], H^n \times H^n \times [0,1]),$$
defined by the formula
$$C(\sigma)(x,t) = (\sigma(x), t),$$
is continuous. The function
$$\phi : H^n \times H^n \times [0,1] \to H^n,$$
defined by the formula
$$\phi(x,y,t) = \frac{(1-t)x + ty}{|||(1-t)x + ty|||},$$
is C^∞. Therefore, the function
$$D : \mathrm{C}^\infty(\Delta^k \times [0,1], H^n \times H^n \times [0,1]) \to \mathrm{C}^\infty(\Delta^k \times [0,1], H^n),$$
defined by $D = \phi_*$, is continuous. Finally, the function
$$\tilde{F}^k : \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{C}^\infty(\Delta^k \times [0,1], H^n)$$
is continuous, since $\tilde{F}^k = DCBA$.

For each $i = 0, \ldots, k$, let
$$\alpha_i : \Delta^{k+1} \to \Delta^k \times [0,1]$$
be the affine map constructed in Lemma 2 of §11.4. Then
$$(\alpha_i)^* : \mathrm{C}^\infty(\Delta^k \times [0,1], M) \to \mathrm{C}^\infty(\Delta^{k+1}, M)$$
is continuous, since α_i is C^∞.

For each $i = 0, \ldots, k$, define a function
$$F_i^k : \mathrm{C}^\infty(\Delta^k, M) \to \mathrm{C}^\infty(\Delta^{k+1}, M)$$
by $F_i^k(\sigma) = F_\sigma \alpha_i$. Then F_i^k is continuous, since $F_i^k = \alpha_i^* F^k$.

Define a linear transformation $F_*^k : \mathcal{C}_k(M) \to \mathcal{C}_{k+1}(M)$ by the formula
$$F_*^k = \sum_{i=0}^k (-1)^i (F_i^k)_*.$$

Essentially the same calculation as in Lemma 2 of §11.4 shows that

$$\partial_{k+1}F_*^k + F_*^{k-1}\partial_k = (\mathrm{Str}^k)_* - id_k.$$

Thus $F_* = \{F_*^k\}$ is a chain homotopy from $(\mathrm{Str}^*)_*$ to the identity. □

Smearing

We now assume that the space-form $M = H^n/\Gamma$ is compact and orientable. Let $G = \mathrm{I}_0(H^n)$ be the group of orientation preserving isometries of H^n, and let H be the subgroup of G of all elements that fix the point e_{n+1}. The *Haar integral* of a function $\phi : G \to \mathbb{R}$ is given by the formula

$$\int_G \phi(g)dg = \int_{G/H}\left(\int_H \phi(gh)dh\right)d(gH),$$

where dh is the *Haar measure* on the compact group H and $d(gH)$ is the measure on G/H corresponding to hyperbolic volume in H^n under the homeomorphism from G/H to H^n given by Theorems 5.1.5 and 5.2.9. We shall normalize the Haar measure dg on G by normalizing the Haar measure dh on H so that

$$\int_H dh = 1.$$

The group G has a left-invariant metric. For example, the metric corresponding to the metric d on $\mathrm{M}_0(B^n)$, defined by

$$d(\phi, \psi) = D_B(\phi^{-1}, \psi^{-1}),$$

is left-invariant. Therefore Γ acts freely and discontinuously on G as a group of isometries by left multiplication by Theorem 5.3.4. Therefore, the quotient map

$$\kappa : G \to \Gamma\backslash G$$

is a covering projection by Theorem 8.1.3. Consequently, the Haar measure on G descends to a positive measure on $\Gamma\backslash G$ so that κ is locally measure preserving. The integral of a function $\phi : \Gamma\backslash G \to \mathbb{R}$, with respect to this measure, is given by the formula

$$\int_{\Gamma\backslash G}\phi(\Gamma g)d(\Gamma g) = \int_{(\Gamma\backslash G)/H}\left(\int_H \phi(\Gamma gh)dh\right)d(\Gamma gH),$$

where $d(\Gamma gH)$ is the measure on the double coset space

$$(\Gamma\backslash G)/H = \Gamma\backslash(G/H) = \Gamma\backslash H^n = M$$

corresponding to hyperbolic volume. The volume of $\Gamma\backslash G$ is given by

$$\begin{aligned}\mathrm{Vol}(\Gamma\backslash G) &= \int_{\Gamma\backslash G/H}\left(\int_H dh\right)d(\Gamma gH) \\ &= \int_{\Gamma\backslash G/H} d(\Gamma gH) = \mathrm{Vol}(M).\end{aligned}$$

The group G is homeomorphic to $H^n \times H$ by Theorems 5.1.5 and 5.2.9. Moreover, the corresponding action of Γ on $H^n \times H$ is given by

$$g(x,h) = (gx, *).$$

Let D be a Dirichlet polyhedron for Γ. Then $D^\circ \times H$ is a fundamental domain for the action of Γ on $H^n \times H$. As M is compact, D is compact. Therefore $D \times H$ is compact, and so $\Gamma \backslash G$ is compact.

Given σ in $\mathrm{Str}(\Delta^k, H^n)$, define a function

$$\sigma^* : \Gamma\backslash G \to \mathrm{Str}(\Delta^k, M)$$

by $\sigma^*(\Gamma g) = \pi g \sigma$, where $\pi : H^n \to M$ is the quotient map.

Lemma 7. *The function $\sigma^* : \Gamma\backslash G \to \mathrm{Str}(\Delta^k, M)$ is continuous.*

Proof: Let $\kappa : G \to \Gamma\backslash G$ be the quotient map. Then σ^* lifts to a function

$$\sigma^* : G \to \mathrm{Str}(\Delta^k, H^n)$$

defined by $\sigma^*(g) = g\sigma$. As $\pi_*\sigma^* = \sigma^*\kappa$, it suffices to show that

$$\sigma^* : G \to \mathrm{Str}(\Delta^k, H^n)$$

is continuous. Since the action of G on H^n,

$$\alpha : G \times H^n \to H^n,$$

given by $\alpha(g,x) = gx$, is continuous, the corresponding inclusion map $\hat{\alpha} : G \to \mathrm{C}(H^n, H^n)$ is continuous. As

$$\sigma^* : \mathrm{C}(H^n, H^n) \to \mathrm{C}(\Delta^k, H^n)$$

is continuous, its restriction

$$\sigma^* : G \to \mathrm{Str}(\Delta^k, H^n)$$

is continuous. □

Let σ be an element of $\mathrm{Str}(\Delta^k, H^n)$. The *smear* of σ is the positive Borel measure on $\mathrm{C}^\infty(\Delta^k, M)$ given by the formula

$$\mathrm{Smr}(\sigma) = (\sigma^*)_*(d(\Gamma g)).$$

In other words, if B is a Borel subset of $\mathrm{C}^\infty(\Delta^k, M)$, then $\mathrm{Smr}(\sigma)(B)$ is the volume of $(\sigma^*)^{-1}(B)$ in $\Gamma\backslash G$. As $\Gamma\backslash G$ is compact, the image of

$$\sigma^* : \Gamma\backslash G \to \mathrm{C}^\infty(\Delta^k, M)$$

is compact. Therefore $\mathrm{Smr}(\sigma)$ has compact support. Moreover

$$\|\mathrm{Smr}(\sigma)\| = \mathrm{Vol}(\Gamma\backslash G) = \mathrm{Vol}(M).$$

Thus, we have a function

$$\mathrm{Smr} : \mathrm{Str}(\Delta^k, H^n) \to \mathcal{C}_k(M).$$

Lemma 8. *If σ is in* $\mathrm{Str}(\Delta^k, H^n)$ *and f is in* $\mathrm{I}_0(H^n)$, *then*

$$\mathrm{Smr}(f\sigma) = \mathrm{Smr}(\sigma).$$

Proof: Define $f^* : \Gamma\backslash G \to \Gamma\backslash G$ by $f^*(\Gamma g) = \Gamma g f$. Then f^* is continuous, since right multiplication by f is continuous in G. Observe that

$$\begin{aligned}\mathrm{Smr}(f\sigma) &= ((f\sigma)^*)_*(d(\Gamma g)) \\ &= (\sigma^* f^*)_*(d(\Gamma g)) \quad = \quad (\sigma^*)_*(f^*)_*(d(\Gamma g)).\end{aligned}$$

Now since the Haar measure on G is right-invariant, the induced measure on $\Gamma\backslash G$ is invariant under right multiplication by G. Hence, if B is a Borel subset of $\Gamma\backslash G$, we have

$$(f^*)_*(d(\Gamma g))(B) = \mathrm{Vol}((f^*)^{-1}(B)) = \mathrm{Vol}(Bf^{-1}) = \mathrm{Vol}(B).$$

Therefore, we have

$$(f^*)_*(d(\Gamma g)) = d(\Gamma g).$$

Hence, we have

$$\mathrm{Smr}(f\sigma) = (\sigma^*)_*(d(\Gamma g)) = \mathrm{Smr}(\sigma).$$ □

The function

$$\mathrm{Smr} : \mathrm{Str}(\Delta^k, H^n) \to \mathcal{C}_k(M)$$

extends linearly to a linear transformation

$$\mathrm{Smr}_k : \mathrm{Str}_k(H^n) \to \mathcal{C}_k(M).$$

Lemma 9. *The family* $\{\mathrm{Smr}_k\}$ *of linear transformations is a chain map from* $\mathrm{Str}(H^n)$ *to* $\mathcal{C}(M)$.

Proof: Let σ be an element of $\mathrm{Str}(\Delta^k, H^n)$. It suffices to show that

$$\mathrm{Smr}_k(\partial\sigma) = \partial\mathrm{Smr}_k(\sigma).$$

Observe that

$$\mathrm{Smr}_k(\partial\sigma) = \mathrm{Smr}_k\left(\sum_{i=0}^{k}(-1)^i\sigma\eta_i\right) = \sum_{i=0}^{k}(-1)^i\mathrm{Smr}(\sigma\eta_i),$$

whereas

$$\partial\mathrm{Smr}_k(\sigma) = \sum_{i=0}^{k}(-1)^i(\eta_i^*)_*\mathrm{Smr}(\sigma).$$

Now observe that

$$\begin{aligned}\mathrm{Smr}(\sigma\eta_i) &= ((\sigma\eta_i)^*)_*(d(\Gamma g)) \\ &= (\eta_i^*\sigma^*)_*(d(\Gamma g)) \\ &= (\eta_i^*)_*(\sigma^*)_*(d(\Gamma g)) \quad = \quad (\eta_i^*)_*\mathrm{Smr}(\sigma).\end{aligned}$$

Therefore $\mathrm{Smr}_k(\partial\sigma) = \partial\mathrm{Smr}_k(\sigma)$. □

Let $\sigma : \Delta^n \to H^n$ be a straight singular n-simplex and let ρ be a reflection of H^n. The *average* of σ is the signed Borel measure on $C^\infty(\Delta^n, M)$ given by

$$\mathrm{Avg}(\sigma) = \frac{1}{2}\big(\mathrm{Smr}(\sigma) - \mathrm{Smr}(\rho\sigma)\big).$$

Theorem 11.5.3. *Let $M = H^n/\Gamma$ be a compact orientable space-form. If σ is in $\mathrm{Str}(\Delta^n, H^n)$, then $\mathrm{Avg}(\sigma)$ is a cycle in $\mathcal{C}_n(M)$.*

Proof: Observe that

$$\begin{aligned}
\partial\mathrm{Avg}(\sigma) &= \frac{1}{2}\big(\partial\mathrm{Smr}(\sigma) - \partial\mathrm{Smr}(\rho\sigma)\big)\\
&= \frac{1}{2}\big(\mathrm{Smr}(\partial\sigma) - \mathrm{Smr}(\partial\rho\sigma)\big)\\
&= \frac{1}{2}\left[\mathrm{Smr}\left(\sum_{i=0}^{k}(-1)^i\sigma\eta_i\right) - \mathrm{Smr}\left(\sum_{i=0}^{k}(-1)^i\rho\sigma\eta_i\right)\right]\\
&= \frac{1}{2}\sum_{i=0}^{k}(-1)^i\big(\mathrm{Smr}(\sigma\eta_i) - \mathrm{Smr}(\rho\sigma\eta_i)\big).
\end{aligned}$$

Moreover, we have

$$\mathrm{Smr}(\sigma\eta_i) = \mathrm{Smr}(\rho\sigma\eta_i),$$

since $\sigma\eta_i$ and $\rho\sigma\eta_i$ differ by an element of $\mathrm{I}_0(H^n)$. Hence $\partial\mathrm{Avg}(\sigma) = 0$. □

Representing the Fundamental Class

We now assume that the space-form $M = H^n/\Gamma$ is compact and oriented. Let c be a cycle in $S_n^\infty(M)$ that represents the fundamental class of M. Then the cycle $F_M = I_n(c)$ in $\mathcal{D}_n(M)$, defined by

$$F_M(\omega) = \int_c \omega,$$

represents the fundamental class of M in $\mathrm{H}_n(\mathcal{D}(M))$. The cycle F_M does not depend on the choice of c because $\mathcal{D}_{n+1}(M) = 0$. The cycle F_M is called the *fundamental cycle* of M in $\mathcal{D}_n(M)$.

A cycle μ in $\mathcal{C}_n(M)$ is said *to represent* a class κ in $\mathrm{H}_n(\mathcal{D}(M))$ if the cycle $\ell_n(\mu) = f_\mu$ in $\mathcal{D}_n(M)$ represents κ.

Lemma 10. *Let μ be a cycle in $\mathcal{C}_n(M)$, let Ω_M be the volume form of M, and let F_M be the fundamental cycle of M in $\mathcal{D}_n(M)$. Then μ represents the class $f_\mu(\Omega_M)\mathrm{Vol}(M)^{-1}[F_M]$ in $\mathrm{H}_n(\mathcal{D}(M))$.*

Proof: Since $[F_M]$ generates $\mathrm{H}_n(\mathcal{D}(M))$, there is a constant k such that $[f_\mu] = k[F_M]$. As $\mathcal{D}_{n+1}(M) = 0$, we have that $f_\mu = kF_M$. Hence

$$f_\mu(\Omega_M) = kF_M(\Omega_M) = k\mathrm{Vol}(M)$$

and so $k = f_\mu(\Omega_M)/\mathrm{Vol}(M)$. □

Theorem 11.5.4. *Let $M = H^n/\Gamma$ be a compact oriented space-form, let σ be in* $\mathrm{Str}(\Delta^n, H^n)$, *and let F_M be the fundamental cycle of M in $\mathcal{D}_n(M)$. Then* $\mathrm{Avg}(\sigma)$ *represents the class* $\pm\mathrm{Vol}(\sigma(\Delta^n))[F_M]$ *in* $\mathrm{H}_n(\mathcal{D}(M))$ *with the plus or minus sign according as $\pi\sigma$ preserves or reverses orientation.*

Proof: Observe that

$$\begin{aligned}
f_{\mathrm{Smr}(\sigma)}(\Omega_M) &= \int_{\tau\in \mathrm{C}^\infty(\Delta^n,M)}\left(\int_\tau \Omega_M\right) d(\mathrm{Smr}(\sigma))\\
&= \int_{\tau\in \mathrm{C}^\infty(\Delta^n,M)}\left(\int_\tau \Omega_M\right) d\big((\sigma^*)_*(d(\Gamma g))\big)\\
&= \int_{\Gamma g\in\Gamma\backslash G}\left(\int_{\sigma^*(\Gamma g)} \Omega_M\right) d(\Gamma g)\\
&= \int_{\Gamma g\in\Gamma\backslash G}\left(\int_{\pi g\sigma} \Omega_M\right) d(\Gamma g)\\
&= \int_{\Gamma g\in\Gamma\backslash G} \pm\mathrm{Vol}(g\sigma(\Delta^n))d(\Gamma g)\\
&= \int_{\Gamma\backslash G} \pm\mathrm{Vol}(\sigma(\Delta^n))d(\Gamma g)\\
&= \pm\mathrm{Vol}(\sigma(\Delta^n))\mathrm{Vol}(\Gamma\backslash G)\\
&= \pm\mathrm{Vol}(\sigma(\Delta^n))\mathrm{Vol}(M).
\end{aligned}$$

Hence

$$\begin{aligned}
f_{\mathrm{Avg}(\sigma)}(\Omega_M) &= \frac{1}{2}\left(f_{\mathrm{Smr}(\sigma)}(\Omega_M) - f_{\mathrm{Smr}(\rho\sigma)}(\Omega_M)\right)\\
&= \pm\mathrm{Vol}(\sigma(\Delta^n))\mathrm{Vol}(M).
\end{aligned}$$

Therefore $\mathrm{Avg}(\sigma)$ represents the class $\pm\mathrm{Vol}(\sigma(\Delta^n))[F_M]$ in $\mathrm{H}_n(\mathcal{D}(M))$ by Lemma 10. □

Exercise 11.5

1. Let M be a differentiable manifold. Prove that the total variation is a norm on $\mathcal{C}_k(M)$ for each k.

2. Let M be a hyperbolic space-form. Prove that $m_k : \mathrm{Str}_k(M) \to \mathcal{C}_k(M)$ is norm preserving for each k.

3. Let $i : \{e_0, \ldots, e_k\} \to \Delta^k$ be the inclusion map. Prove that

$$i^* : \mathrm{Str}(\Delta^k, H^n) \to \mathrm{C}(\{e_0, \ldots, e_k\}, H^n)$$

is a homeomorphism with the compact-open topology on $\mathrm{Str}(\Delta^k, H^n)$. Conclude that $\mathrm{Str}(\Delta^k, H^n)$ is homeomorphic to $(H^n)^{k+1}$.

4. Let $M = H^n/\Gamma$ be a hyperbolic space-form. Prove that $\mathrm{Str}(\Delta^k, M)$, with the compact-open topology, is homeomorphic to $(H^n)^{k+1}/\Gamma$, where Γ acts diagonally on the left of $(H^n)^{k+1}$ as a discontinuous group of isometries. Conclude that $\mathrm{Str}(\Delta^k, M)$ is a connected $(k+1)n$-dimensional manifold.
5. Prove that the C^1 topology on $\mathrm{Str}(\Delta^k, H^n)$ is the compact-open topology.
6. Prove that the straightening function
$$\mathrm{Str}^k : \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{Str}(\Delta^k, H^n)$$
is continuous with respect to the compact-open topology.
7. Let M be a hyperbolic space-form and let $\pi : H^n \to M$ be the quotient map. Prove that
$$\pi_* : \mathrm{C}^\infty(\Delta^k, H^n) \to \mathrm{C}^\infty(\Delta^k, M)$$
is an open map.
8. Let $\sigma : \Delta^n \to H^n$ be a straight singular n-simplex. Prove that the definition of $\mathrm{Avg}(\sigma)$ does not depend on the choice of the reflection ρ.
9. Prove that $\|\mathrm{Avg}(\sigma)\| = \mathrm{Vol}(M)$.
10. Let M be a compact, oriented, n-dimensional, hyperbolic space-form. Prove that $\|M\| = \inf\{\|\mu\| : \mu \text{ in } \mathcal{C}_n(M) \text{ represents the fundamental class of } M \text{ in } \mathrm{H}_n(\mathcal{D}(M))\}$.

§11.6. Mostow Rigidity

Let M and N be closed, connected, orientable, hyperbolic n-manifolds, with $n > 2$. In this section, we prove Mostow's rigidity theorem which states that a homotopy equivalence $\varphi : M \to N$ is homotopic to an isometry. Since M and N are complete, we may assume that M and N are hyperbolic space-forms, say $M = H^n/\Gamma$ and $N = H^n/\mathrm{H}$.

It is basic theorem of differential topology that any continuous function between differentiable manifolds is homotopic to a C^∞ map. Hence, we may assume that a homotopy equivalence $\varphi : M \to N$ is a C^∞ (*smooth*) map.

Lipschitz Conditions

Definition: A function $f : X \to Y$ between metric spaces satisfies a *Lipschitz condition* if and only if there is a constant $k > 0$ such that
$$d(f(x), f(y)) \leq k\,d(x,y) \quad \text{for all } x, y \text{ in } X.$$
The constant k is called a *Lipschitz constant* for f.

Lemma 1. *Let C be a compact convex subset of H^n and let $f : C \to H^n$ be a C^1 map. Then f satisfies a Lipschitz condition.*

Proof: Let x, y be distinct points of C and let $\alpha : [a,b] \to C$ be a geodesic arc from x to y. Then $f\alpha : [a,b] \to H^n$ is a C^1 curve from $f(x)$ to $f(y)$. We pass to the upper half-space model U^n of hyperbolic space. By Theorem 4.6.6, the element of hyperbolic arc length of U^n is $|dx|/x_n$. Observe that

$$\begin{aligned} d(f(x), f(y)) &\leq |f\alpha| \\ &= \int_a^b \frac{|(f\alpha)'(t)|}{(f\alpha(t))_n} dt \\ &= \int_a^b \frac{|f'(\alpha(t))\alpha'(t)|}{(f\alpha(t))_n} dt \\ &\leq \int_a^b \frac{|f'(\alpha(t))|\,|\alpha'(t)|}{(f\alpha(t))_n} dt \\ &= \int_a^b \frac{|f'(\alpha(t))|(\alpha(t))_n|\alpha'(t)|}{(f\alpha(t))_n(\alpha(t))_n} dt. \end{aligned}$$

Let k be the maximum value of the continuous function $|f'(x)|x_n/(f(x))_n$ on the compact set C. Then we have

$$\begin{aligned} d(f(x), f(y)) &\leq k\int_a^b \frac{|\alpha'(t)|dt}{(\alpha(t))_n} \\ &= k|\alpha| \;=\; k\,d(x,y). \end{aligned}$$

□

Lemma 2. *A C^1 map $\varphi : M \to N$ satisfies a Lipschitz condition.*

Proof: By Lemma 1 and Theorem 8.3.6, the map φ satisfies a Lipschitz condition locally, that is, for each point w of M, there is an $r(w) > 0$ and a $k(w) > 0$ such that

$$d(\varphi(u), \varphi(v)) \leq k(w)d(u,v) \quad \text{for all } u, v \text{ in } \overline{B}(w, r(w)).$$

As M is compact, there is a finite set of points $\{w_1, \ldots, w_\ell\}$ of M such that $\{B(w_i, r(w_i))\}$ covers M. Set

$$k = \max\{k(w_1), \ldots, k(w_\ell)\}.$$

Let u, v be distinct points of M. By Theorem 8.5.5, there is a geodesic arc $\alpha : [a,b] \to M$ joining u to v. Moreover, there is a partition

$$a = t_0 < \cdots < t_m = b$$

of the interval $[a,b]$ such that for each i, we have

$$\alpha([t_i, t_{i+1}]) \subset B(w_j, r(w_j)) \quad \text{for some } j.$$

Hence, we have

$$\begin{aligned} d(\varphi(u), \varphi(v)) &\leq \sum_{i=0}^{m-1} d(\varphi(\alpha(t_i)), \varphi(\alpha(t_{i+1}))) \\ &\leq \sum_{i=0}^{m-1} k\,d(\alpha(t_i), \alpha(t_{i+1})) \;=\; k\,d(u,v). \end{aligned}$$

□

By covering space theory, any map $\varphi : M \to N$ lifts to a map $\tilde{\varphi} : H^n \to H^n$ such that the following diagram commutes:

$$\begin{array}{ccc} H^n & \xrightarrow{\tilde{\varphi}} & H^n \\ \pi \downarrow & & \downarrow \eta \\ H^n/\Gamma & \xrightarrow{\varphi} & H^n/\mathrm{H}, \end{array}$$

where π and η are the quotient maps.

Lemma 3. *Let $\tilde{\varphi} : H^n \to H^n$ be a lift of a smooth homotopy equivalence $\varphi : M \to N$. Then $\tilde{\varphi}$ satisfies a Lipschitz condition and a Lipschitz constant for φ is also a Lipschitz constant for $\tilde{\varphi}$.*

Proof: Since $\eta : H^n \to N$ is a covering projection, we deduce from Theorem 8.3.6 that for each w in N and x in $\eta^{-1}(w)$ there is an $r(w) > 0$ such that η maps $B(x, r(w))$ isometrically onto $B(w, r(w))$. Let ϵ be a Lebesgue number for the covering $\{B(w, r(w))\}$ of the compact space N. Then η maps $B(x, \epsilon)$ isometrically onto $B(\eta(x), \epsilon)$ for each x in H^n.

Now as M is compact, $\varphi : M \to N$ is uniformly continuous. Hence, there is a $\delta > 0$ such that if $d(u, v) < \delta$, then $d(\varphi(u), \varphi(v)) < \epsilon$. Let x, y be points of H^n, with $d(x, y) < \delta$, and let $\alpha : [a, b] \to H^n$ be a geodesic arc from x to y. Then

$$\pi\alpha([a, b]) \subset B(\pi(x), \delta),$$

since π is a local isometry. Hence

$$\varphi\pi\alpha([a, b]) \subset B(\varphi\pi(x), \epsilon).$$

Next, observe that $\eta\tilde{\varphi}\alpha = \varphi\pi\alpha$ and η maps $B(\tilde{\varphi}(x), \epsilon)$ isometrically onto $B(\varphi\pi(x), \epsilon)$. Therefore, by unique path lifting, we have

$$\tilde{\varphi}\alpha([a, b]) \subset B(\tilde{\varphi}(x), \epsilon).$$

Let k be a Lipschitz constant for φ. Then we have

$$\begin{aligned} d(\tilde{\varphi}(x), \tilde{\varphi}(y)) &= d(\eta\tilde{\varphi}(x), \eta\tilde{\varphi}(y)) \\ &= d(\varphi\pi(x), \varphi\pi(y)) \\ &\leq k\,d(\pi(x), \pi(y)) \;=\; k\,d(x, y). \end{aligned}$$

Now assume that x and y are arbitrary points of H^n. Let

$$x = x_0, x_1, \ldots, x_m = y$$

be a partition of the geodesic segment $[x, y]$ such that $d(x_i, x_{i+1}) < \delta$ for each $i = 0, \ldots, m - 1$. Then

$$\begin{aligned} d(\varphi(x), \varphi(y)) &\leq \sum_{i=0}^{m-1} d(\varphi(x_i), \varphi(x_{i+1})) \\ &\leq \sum_{i=0}^{m-1} k\,d(x_i, x_{i+1}) \;=\; k\,d(x, y). \end{aligned}$$

□

Pseudo-isometries

Definition: Given a metric space X, a function $f : X \to X$ is a *pseudo-isometry* if and only if there are constants k and ℓ such that

$$k^{-1}d(x,y) - \ell \leq d(f(x), f(y)) \leq k\,d(x,y)$$

for all x, y in X; moreover, if $\ell = 0$, then f is called a *quasi-isometry.*

Theorem 11.6.1. *Let $M = H^n/\Gamma$ and $N = H^n/\mathrm{H}$ be compact orientable space-forms and let $\tilde{\varphi} : H^n \to H^n$ be a lift of a smooth homotopy equivalence $\varphi : M \to N$. Then $\tilde{\varphi}$ is a pseudo-isometry.*

Proof: Let $\psi : N \to M$ be a smooth homotopy inverse for φ and let $F : M \times [0,1] \to M$ be a homotopy from $\psi\varphi$ to id_M. Let $\tilde{\psi} : H^n \to H^n$ be a lift of ψ. By the covering homotopy theorem, F lifts to a map $\tilde{F} : H^n \times [0,1] \to H^n$ such that $\tilde{F}_0 = \tilde{\psi}\tilde{\varphi}$. As $\pi\tilde{F}_1 = F_1\pi = \pi$, we have that $\tilde{F}_1 = f$ for some element f of Γ. By replacing $\tilde{\psi}$ with $f^{-1}\tilde{\psi}$ and $\tilde{F}$ with $f^{-1}\tilde{F}$, if necessary, we may assume that $\tilde{F}_1 = id_{H^n}$. Then $\tilde{F}$ is a homotopy from $\tilde{\psi}\tilde{\varphi}$ to id_{H^n}. Now let g be an arbitrary element of Γ. Then we have

$$\begin{aligned} \pi\tilde{F}(g \times id) &= F(\pi \times id)(g \times id) \\ &= F(\pi g \times id) \\ &= F(\pi \times id) \;=\; \pi\tilde{F}. \end{aligned}$$

Hence, there is an element h of Γ such that $\tilde{F}(g \times id) = h\tilde{F}$. As $\tilde{F}_1 = id_{H^n}$, we find that $h = g$. Therefore $\tilde{F}$ is Γ-equivariant. In particular $\tilde{\psi}\tilde{\varphi} = \tilde{F}_0$ is Γ-equivariant.

Let D be a Dirichlet polyhedron for Γ. Then D is compact, since H^n/Γ is compact. Therefore $\tilde{F}(D \times [0,1])$ is compact. Let δ be the diameter of $\tilde{F}(D \times [0,1])$. If x is in D, then $\tilde{\psi}\tilde{\varphi}(x)$ and x are in $\tilde{F}(D \times [0,1])$ and so

$$d(\tilde{\psi}\tilde{\varphi}(x), x) \leq \delta.$$

As $\tilde{\psi}\tilde{\varphi}$ is Γ-equivariant, the above inequality holds for all x in H^n.

By Lemma 3, there is a constant $k > 0$ such that

$$d(\tilde{\varphi}(x), \tilde{\varphi}(y)) \leq k\,d(x,y) \quad \text{and} \quad d(\tilde{\psi}(x), \tilde{\psi}(y)) \leq k\,d(x,y)$$

for all x, y in H^n. Observe that

$$\begin{aligned} d(x,y) &\leq d(x, \tilde{\psi}\tilde{\varphi}(x)) + d(\tilde{\psi}\tilde{\varphi}(x), \tilde{\psi}\tilde{\varphi}(y)) + d(\tilde{\psi}\tilde{\varphi}(y), y) \\ &\leq 2\delta + k\,d(\tilde{\varphi}(x), \tilde{\varphi}(y)). \end{aligned}$$

Therefore, we have

$$d(\tilde{\varphi}(x), \tilde{\varphi}(y)) \geq k^{-1}d(x,y) - 2\delta/k.$$

Let $\ell = 2\delta/k$. Then for all x, y in H^n, we have

$$k^{-1}d(x,y) - \ell \leq d(\tilde{\varphi}(x), \tilde{\varphi}(y)) \leq k\,d(x,y).$$

□

Lemma 4. *Let $\gamma : [a,b] \to H^n$ be a C^1 curve, let s be the distance from the set $\gamma([a,b])$ to a hyperbolic line L of H^n, and let $\rho : H^n \to L$ be the nearest point retraction. Then*

$$|\rho\gamma| \leq (\cosh s)^{-1}|\gamma|.$$

Proof: We pass to the upper half-space model U^n of hyperbolic space. Now without loss of generality, we may assume that L is the positive nth axis. Then $\rho(x) = |x|e_n$ and

$$\cosh d(x, \rho(x)) = |x|/x_n.$$

Observe that

$$\begin{aligned}
|\rho\gamma| &= \int_a^b \frac{|(\rho\gamma)'(t)|}{(\rho\gamma(t))_n} dt \\
&= \int_a^b \frac{|\rho'(\gamma(t))\gamma'(t)|}{|\gamma(t)|} dt \\
&= \int_a^b \frac{|(\gamma(t)/|\gamma(t)|) \cdot \gamma'(t)|}{|\gamma(t)|} dt \\
&= \int_a^b \frac{|\gamma(t) \cdot \gamma'(t)|}{|\gamma(t)|^2} dt \\
&\leq \int_a^b \frac{|\gamma'(t)|}{|\gamma(t)|} dt \\
&\leq \int_a^b \frac{|\gamma'(t)|dt}{(\cosh s)(\gamma(t))_n} \quad = \quad (\cosh s)^{-1}|\gamma|.
\end{aligned}$$

□

Lemma 5. *Let $k > 0$ be a Lipschitz constant for a function $f : H^n \to H^n$ and let $\alpha : [a,b] \to H^n$ be a geodesic arc from x to y. Then $|f\alpha| \leq kd(x,y)$.*

Proof: Let $a = t_0 < t_1 < \cdots < t_m = b$ be a partition of $[a,b]$. Then

$$\sum_{i=1}^m d(f\alpha(t_{i-1}), f\alpha(t_i)) \leq \sum_{i=1}^m k\, d(\alpha(t_{i-1}), \alpha(t_i)) = k\, d(x,y).$$

By definition of $|f\alpha|$, we have that $|f\alpha| \leq k\, d(x,y)$. □

Lemma 6. *Let $f : H^n \to H^n$ be a pseudo-isometry. Then there exists a constant $r > 0$ such that if $\alpha : [a,b] \to H^n$ is a geodesic arc, then*

$$f\alpha([a,b]) \subset N([f\alpha(a), f\alpha(b)], r).$$

Proof: Let $\alpha : [a,b] \to H^n$ be a geodesic arc and let L be a hyperbolic line of H^n passing through $f\alpha(a)$ and $f\alpha(b)$. Let k and ℓ be constants such that

$$k^{-1}d(x,y) - \ell \leq d(f(x), f(y)) \leq k\, d(x,y)$$

for all x, y in H^n and set

$$s = \cosh^{-1}(k^2 + 1).$$

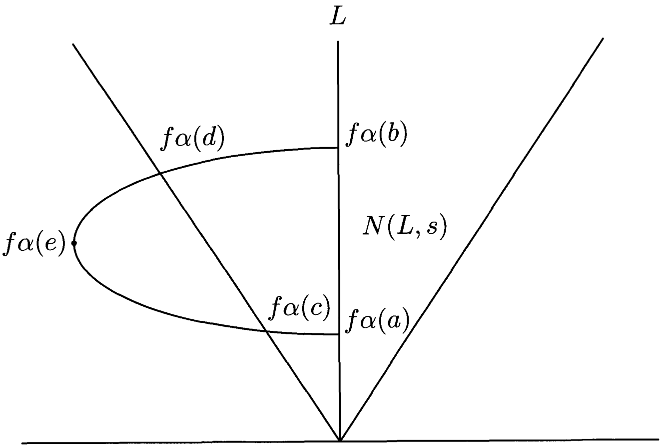


Figure 11.6.1. The pseudo-isometry f applied to the arc α

Suppose that $f\alpha(e)$ is not in $N(L,s)$. Then there is a largest subinterval $[c,d]$ of $[a,b]$ containing e such that $f\alpha([c,d])$ is disjoint from $N(L,s)$. See Figure 11.6.1. Let $p = \alpha(c)$ and $q = \alpha(d)$. Then

$$d(f(p),L) = s = d(f(q),L).$$

Let $\beta : [c,d] \to H^n$ be the restriction of α. We now establish an upper bound for the length of the curve $f\beta$. Let $\rho : H^n \to L$ be the nearest point retraction. By Lemmas 4 and 5, we have

$$\begin{aligned}
k^{-1}d(p,q) - \ell &\leq d(f(p),f(q)) \\
&\leq d(f(p),\rho f(p)) + d(\rho f(p),\rho f(q)) + d(\rho f(q),f(q)) \\
&\leq 2s + |\rho f\beta| \\
&\leq 2s + (k^2+1)^{-1}|f\beta| \\
&\leq 2s + (k^2+1)^{-1}k\,d(p,q).
\end{aligned}$$

Therefore, we have

$$d(p,q) \leq (2s+\ell)k(k^2+1) = m.$$

By Lemma 5, we have

$$|f\beta| \leq k\,d(p,q) \leq km.$$

Now set $t = s + km$. Then

$$f\beta([c,d]) \subset N(L,t).$$

Therefore $f\alpha(e)$ is in $N(L,t)$ and so

$$f\alpha([a,b]) \subset N(L,t).$$

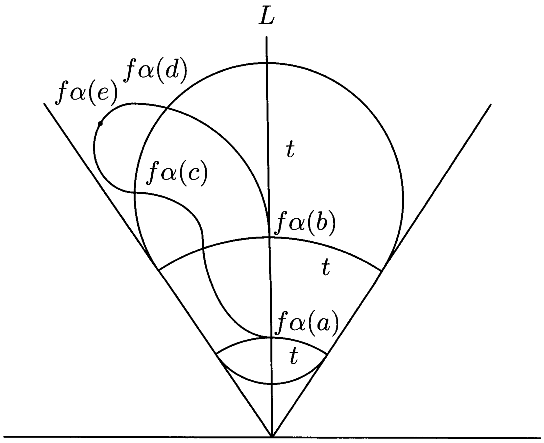


Figure 11.6.2. The pseudo-isometry f applied to the arc α

Suppose that $f\alpha(e)$ is not in $N([f\alpha(a), f\alpha(b)], t)$. Then there is a largest subinterval $[c, d]$ of $[a, b]$ containing e such that $f\alpha([c, d])$ is disjoint from $N([f\alpha(a), f\alpha(b)], t)$. See Figure 11.6.2. Let $p = \alpha(c)$ and $q = \alpha(d)$. Then

$$d(f(p), f\alpha(b)) \;=\; t \;=\; d(f(q), f\alpha(b))$$

or

$$d(f(p), f\alpha(a)) \;=\; t \;=\; d(f(q), f\alpha(a)).$$

Without loss of generality, we may assume that the former holds.

Let $\beta : [c, d] \to H^n$ be the restriction of α. We now establish an upper bound for the length of the curve $f\beta$. Observe that

$$d(f(p), f(q)) \leq d(f(p), f\alpha(b)) + d(f\alpha(b), f(q)) = 2t.$$

Therefore

$$k^{-1}d(p, q) - \ell \leq d(f(p), f(q)) \leq 2t.$$

Hence, we have

$$d(p, q) \leq k(2t + \ell) = j.$$

By Lemma 5, we have

$$|f\beta| \leq k\, d(p, q) \leq kj.$$

Now set $r = t + kj$. Then

$$f\beta([c, d]) \subset B(f\alpha(b), r).$$

Therefore $f\alpha(e)$ is in $B(f\alpha(b), r)$, and so

$$f\alpha([a, b]) \subset N([f\alpha(a), f\alpha(b)], r). \qquad \square$$

Lemma 7. *Let $f : B^n \to B^n$ be a pseudo-isometry. Then there exists a constant $r > 0$ such that for each hyperbolic ray R of B^n based at any point p, there is a unique hyperbolic ray R' of B^n based at $f(p)$ such that $f(R) \subset \overline{N}(R', r)$.*

Proof: Let R be a hyperbolic ray in B^n based at p and let $\lambda : \mathbb{R} \to B^n$ be a geodesic line such that $\lambda([0, \infty)) = R$. As f is a pseudo-isometry,

$$\lim_{i \to \infty} d(f\lambda(0), f\lambda(i)) = \infty.$$

Let r be the constant of Lemma 6. Then there is an $m > 0$ such that

$$d(f\lambda(0), f\lambda(i)) \geq r \quad \text{for all} \quad i \geq m.$$

Without loss of generality, we may assume that $f\lambda(0) = 0$. For each integer $i \geq m$, let R_i be the hyperbolic ray in B^n based at 0 and passing through $f\lambda(i)$. For each pair of integers i, j such that $j > i \geq m$, let x_{ij} be the point of R_j nearest to $f\lambda(i)$. As

$$f\lambda([0, i]) \subset f\lambda([0, j]) \subset N(R_j, r),$$

we find that

$$d(f\lambda(i), x_{ij}) < r.$$

Now the triangle $\triangle(0, f\lambda(i), x_{ij})$ has a right angle at x_{ij}. See Figure 11.6.3. Let α_{ij} be the angle of $\triangle$ at 0. Then by Formula 3.5.9, we have

$$\sinh d(f\lambda(i), x_{ij}) = \sinh d(0, f\lambda(i)) \sin \alpha_{ij}.$$

Therefore, we have

$$\sin \alpha_{ij} \leq \frac{\sinh r}{\sinh d(0, f\lambda(i))}.$$

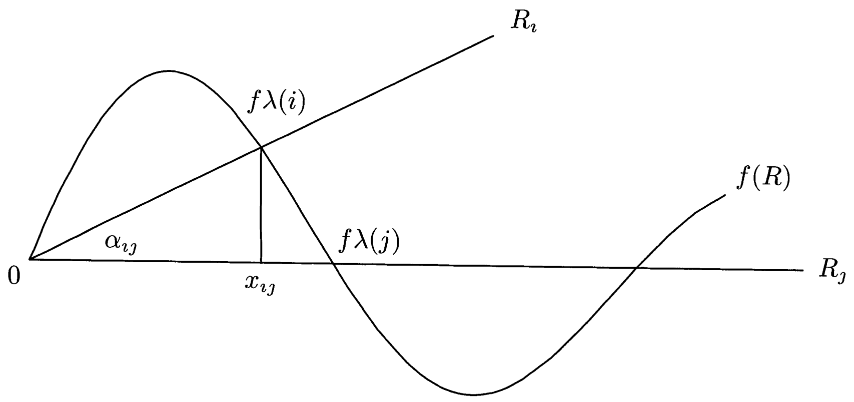


Figure 11.6.3. The pseudo-isometry f applied to the ray R

Hence, for each $\epsilon > 0$, there is an integer $k \geq m$ such that $\alpha_{ij} < \epsilon$ for all $j > i \geq k$. For each integer $i \geq m$, let

$$u_i = f\lambda(i)/|f\lambda(i)|.$$

Then $\{u_i\}$ is a Cauchy sequence in S^n, since if $i < j$, we have

$$d_S(u_i, u_j) = \alpha_{ij}.$$

Therefore $\{u_i\}$ converges to a point u in S^n.

Let R' be the ray based at 0 and ending at u. Then the sequence of rays $\{R_i\}$ converges to R' in E^n. Consequently, the sequence of neighborhoods $\{N(R_i, r)\}$ converges to $N(R', r)$ in E^n. If $i < j$, then

$$f\lambda([0,i)) \subset f\lambda([0,j]) \subset N(R_j, r).$$

Therefore, we have

$$f\lambda([0,i]) \subset \bigcap_{j>i} N(R_j, r) \subset \overline{N}(R', r).$$

Hence, we have

$$f(R) = f\lambda([0,\infty)) \subset \overline{N}(R', r). \qquad \square$$

Lemma 8. *Let $f : B^n \to B^n$ be a pseudo-isometry. Given a point u in S^{n-1}, let R be a ray in B^n ending at u, and let R' be a ray ending at u' such that $f(R) \subset \overline{N}(R', r)$ for some $r > 0$. Then u' is uniquely determined by u, and the function*

$$f_\infty : S^{n-1} \to S^{n-1},$$

defined by $f_\infty(u) = u'$, is injective.

Proof: Observe first that the point u' depends only on R and not on the choice of R', since if $\lambda : \mathbb{R} \to B^n$ is a geodesic line such that $\lambda([0,\infty)) = R$, then $f\lambda(i) \to u'$ as $i \to \infty$. Next, we show that u' depends only on u and not on the choice of R. Suppose that S is another ray ending at u and that S' is a ray ending at u'' such that $f(S) \subset \overline{N}(S', s)$ for some $s > 0$.

On the contrary, suppose that $u' \neq u''$. Let $\mu : R \to B^n$ be a geodesic line such that $\mu([0,\infty)) = S$. Then there exist $m > 0$ such that

$$d(f\lambda(i), f\mu(j)) \geq 1 \quad \text{for all} \quad i, j \geq m.$$

Let k and ℓ be constants such that

$$k^{-1}d(x,y) - \ell \leq d(f(x), f(y)) \leq k\,d(x,y)$$

for all x, y. As R and S are asymptotic, there exists $i, j \geq m$ such that

$$d(\lambda(i), \mu(j)) < 1/k.$$

Therefore, we have that

$$d(f\lambda(i), f\mu(j)) < 1,$$

which is a contradiction. Hence $u' = u''$. Thus u' depends only on u, and so we have a function $f_\infty : S^{n-1} \to S^{n-1}$ defined by $f_\infty(u) = u'$.

We now show that f_∞ is injective. On the contrary, suppose that u and v are distinct points of S^{n-1} such that $u' = v'$. Let R and S be rays in B^n ending at u and v, respectively, and let R' and S' be rays in B^n such that $f(R) \subset \overline{N}(R', r)$ for some $r > 0$ and $f(S) \subset \overline{N}(S', s)$ for some $s > 0$. Let λ, μ be geodesic lines as above. Then there exists $m > 0$ such that

$$d(\lambda(i), \mu(j)) \geq k(1 + r + s + \ell) \quad \text{for all} \quad i, j \geq m.$$

Since $u' = v'$, there exists $i, j \geq m$ such that

$$d(f\lambda(i), f\mu(j)) < 1 + r + s.$$

Hence, we have

$$\begin{aligned} d(\lambda(i), \mu(j)) &\leq k(d(f\lambda(i), f\mu(j)) + \ell) \\ &< k(1 + r + s + \ell), \end{aligned}$$

which is a contradiction. Thus f_∞ is injective. □

Lemma 9. *Let $f : B^n \to B^n$ be a pseudo-isometry. Then there exists a constant $r > 0$ such that for each hyperbolic line L of B^n, there is a unique hyperbolic line L' of B^n such that $f(L) \subset \overline{N}(L', r)$.*

Proof: Let L be a hyperbolic line of B^n with endpoints u and v, and let $\lambda : \mathbb{R} \to B^n$ be a geodesic line such that $\lambda(\mathbb{R}) = L$ and $\lambda(t) \to v$ as $t \to \infty$. Let $r > 0$ be the constant in Lemma 7. Then for each positive integer i, there is a ray R_i of B^n based at $f\lambda(i)$ such that

$$f\lambda((-\infty, i]) \subset \overline{N}(R_i, r).$$

Moreover, all the rays $\{R_i\}$ terminate at the same point u' of S^{n-1} that is the limit of the sequence $\{f\lambda(-i)\}$. Likewise, the sequence $\{f\lambda(i)\}$ converges to a point v' of S^{n-1}. By Lemma 8, we have that $u' \neq v'$. Hence, the sequence of rays $\{R_i\}$ converges to the hyperbolic line L' of B^n with endpoints u' and v'. Moreover, if $j > i > 0$, then

$$f\lambda((-\infty, i]) \subset f\lambda((-\infty, j]) \subset \overline{N}(R_j, r).$$

Therefore

$$f\lambda((-\infty, i]) \subset \bigcap_{j>i} \overline{N}(R_j, r) \subset \overline{N}(L', r).$$

Hence, we have

$$f(L) = f\lambda(\mathbb{R}) \subset \overline{N}(L', r).$$

□

Lemma 10. *Let $f : B^n \to B^n$ be a pseudo-isometry. Then there exists a constant $s > 0$ such that for each hyperplane P of B^n and hyperbolic line L orthogonal to P, the nearest point retraction $\rho : B^n \to L'$ maps $f(P)$ onto a geodesic segment of length at most s.*

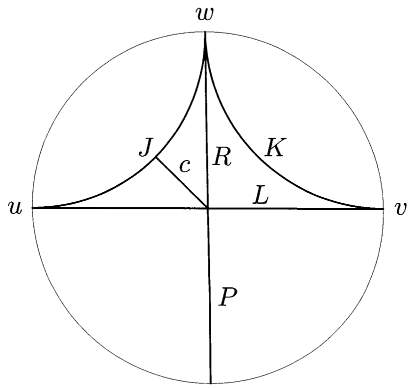


Figure 11.6.4. The ideal triangle with sides J, K, L

Proof: Let x be an arbitrary point of P. Without loss of generality, we may assume that P and L intersect at 0. Let R be a ray in P based at 0 and passing through x. Then there are two hyperbolic lines J and K of B^n that are asymptotic to both R and L. See Figure 11.6.4. The distance from 0 to either J or K is $c = \sinh^{-1}(1)$ by Formula 3.5.17.

Let R' be the ray based at $f(0)$ such that $f(R) \subset \overline{N}(R', r)$ as in Lemma 7, and let J', K', L' be the hyperbolic lines of B^n that remain within a distance r from $f(J), f(K), f(L)$, respectively, as in Lemma 9. By Lemma 8, the endpoint of R' is not an endpoint of L', and J' and L' are the two hyperbolic lines of B^n that are asymptotic to both R' and L'. See Figure 11.6.5. Let I be the hyperbolic line of B^n that is asymptotic to R' and perpendicular to L'. Let p be the nearest point of L' to $f(0)$ and let q be the intersection of I and L'.

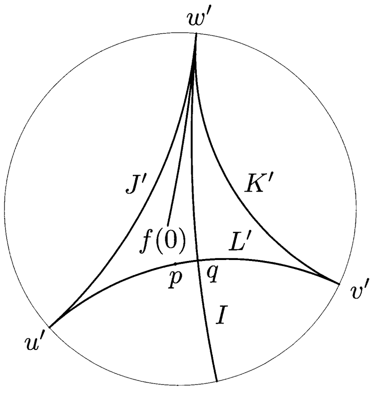


Figure 11.6.5. The ideal triangle with sides J', K', L'

Let k be a Lipschitz constant for f. Then the distance from $f(0)$ to J' and K' is at most $kc + r$, where r is the constant in Lemma 9. As $f(L) \subset N(L', r)$, the distance from p to J' and K' is at most $kc + 2r = b$. Since a geodesic segment from p to either J' or K' must cross I, we deduce from Formula 3.5.7 that $d(p, q) < b$.

Let y be a point of R' such that $d(f(x), y) \leq r$. Since ρ does not increase distances, $d(\rho f(x), \rho(y)) \leq r$. As $\rho(y)$ lies between p and q on L', we deduce that

$$d(\rho f(x), p) \leq d(\rho f(x), \rho(y)) + d(\rho(y), p) \leq r + b.$$

Therefore, the diameter of $\rho(f(P))$ is at most $s = 2(r + b)$. □

Given a pseudo-isometry $f : B^n \to B^n$, let $\overline{f} : \overline{B}^n \to \overline{B}^n$ be the function that extends both f and $f_\infty : S^{n-1} \to S^{n-1}$.

Theorem 11.6.2. *If $f : B^n \to B^n$ is a pseudo-isometry, then the function $\overline{f} : \overline{B}^n \to \overline{B}^n$ is continuous.*

Proof: This is clear if $n = 1$, so assume that $n > 1$. The function $\overline{f}$ is continuous in B^n, since f is continuous and B^n is open in $\overline{B}^n$. We now show that $\overline{f}$ is continuous at a point u of S^{n-1}. Let L be the hyperbolic line of B^n passing through 0 and ending at u. Let $r > 0$ be as large as the constants in Lemmas 9 and 10, and let L' be the hyperbolic line of B^n such that $f(L) \subset N(L', r)$. Let U' be the open neighborhood of $\overline{f}(u) = u'$ in $\overline{B}^n$ bounded by a hyperplane P' of B^n orthogonal to L'. Let H' be the half-space of B^n bounded by P' on the opposite side from U'. Let $\lambda : \mathbb{R} \to B^n$ be a geodesic line such that $\lambda(\mathbb{R}) = L$ and $\lambda(t) \to u$ as $t \to \infty$. Then $f\lambda(t) \to u'$ as $t \to \infty$. Let $\rho : B^n \to L'$ be the nearest point retraction. Then $\rho f\lambda(t) \to u'$ as $t \to \infty$. Hence, there is a constant $m > 0$ such that

$$d(\rho f\lambda(t), H') > 2r \quad \text{for all} \quad t \geq m.$$

Let P_t be the hyperplane of B^n orthogonal to L at $\lambda(t)$. Then by Lemma 10, we have

$$d(\rho f(P_t), H') > r \quad \text{for all} \quad t \geq m.$$

Let U be the open neighborhood of u in $\overline{B}^n$ bounded by P_m. In order to show that $\overline{f}$ is continuous at u, it suffices to show that $\overline{f}(U) \subset \overline{U}'$. Now since the nearest point retraction $\rho : B^n \to L'$ leaves H' invariant, the last inequality implies that $f(U \cap B^n) \subset U' \cap B^n$.

Let v be a point of $U \cap S^{n-1}$ and set $v' = \overline{f}(v)$. Let K be the hyperbolic line of B^n passing through 0 and ending at v, and let $\mu : \mathbb{R} \to B^n$ be a geodesic line such that $\mu(\mathbb{R}) = K$ and $\mu(t) \to v$ as $t \to \infty$. Then there is a constant c such that $\mu(t)$ is in $U \cap B^n$ for all $t \geq c$. Hence $f\mu(t)$ is in $U' \cap B^n$ for all $t \geq c$. Now since $f\mu(t) \to v'$ as $t \to \infty$, we deduce that v' is in $\overline{U}' \cap S^{n-1}$. Thus $\overline{f}(U) \subset \overline{U}'$ and so $\overline{f}$ is continuous at u. Thus $\overline{f}$ is continuous. □

Measure Homology

Let $\varphi : M \to N$ be a C^∞ map. Then φ induces a continuous function
$$\varphi_*^k : C^\infty(\Delta^k, M) \to C^\infty(\Delta^k, N)$$
defined by $\varphi_*^k(\sigma) = \varphi\sigma$. Furthermore φ_*^k induces a linear transformation
$$(\varphi_*^k)_* : \mathcal{C}_k(M) \to \mathcal{C}_k(N)$$
defined by
$$(\varphi_*^k)_*(\mu)(B) = \mu((\varphi_*^k)^{-1}(B))$$
for each μ in $\mathcal{C}_k(M)$ and Borel subset B of $C^\infty(\Delta^k, N)$.

Lemma 11. *The family $\{(\varphi_*^k)_*\}$ of linear transformations is a chain map from $\mathcal{C}(M)$ to $\mathcal{C}(N)$.*

Proof: Let μ be an element of $\mathcal{C}_k(M)$. Then we have
$$\begin{aligned}(\varphi_*^{k-1})_*(\partial\mu) &= (\varphi_*^{k-1})_*\Big(\sum_{i=0}^{k}(-1)^i(\eta_i^*)_*(\mu)\Big)\\ &= \sum_{i=0}^{k}(-1)^i(\varphi_*^{k-1})_*(\eta_i^*)_*(\mu)\\ &= \sum_{i=0}^{k}(-1)^i(\varphi_*^{k-1}\eta_i^*)_*(\mu),\end{aligned}$$
whereas
$$\begin{aligned}\partial(\varphi_*^k)_*(\mu) &= \sum_{i=0}^{k}(-1)^i(\eta_i^*)_*(\varphi_*^k)_*(\mu)\\ &= \sum_{i=0}^{k}(-1)^i(\eta_i^*\varphi_*^k)_*(\mu).\end{aligned}$$
Now observe that if σ is in $C^\infty(\Delta^k, M)$, then
$$(\varphi_*^{k-1}\eta_i^*)(\sigma) = \varphi(\sigma(\eta_i)) = (\varphi\sigma)\eta_i = \eta_i^*\varphi_*^k(\sigma).$$
Therefore, we have
$$\varphi_*^{k-1}\eta_i^* = \eta_i^*\varphi_*^k.$$
Thus, we have
$$(\varphi_*^{k-1})_*(\partial\mu) = \partial(\varphi_*^k)_*(\mu).$$
□

Let $\varphi : M \to N$ be a C^∞ map. Then φ induces a cochain map
$$\{\varphi_k^* : \Omega^k(M) \to \Omega^k(N)\},$$
which, in turn, induces a chain map
$$\{(\varphi_k^*)_* : \mathcal{D}_k(M) \to \mathcal{D}_k(N)\},$$
where
$$(\varphi_k^*)_*(f)(\omega) = f(\varphi_k^*(\omega)).$$

Lemma 12. *Let $\varphi : M \to N$ be a C^∞ map. Then the following diagram commutes for each k:*

$$\begin{array}{ccc}
\mathcal{C}_k(M) & \xrightarrow{\ell_k} & \mathcal{D}_k(M) \\
(\varphi_*^k)_* \downarrow & & \downarrow (\varphi_k^*)_* \\
\mathcal{C}_k(N) & \xrightarrow{\ell_k} & \mathcal{D}_k(N).
\end{array}$$

Proof: Let μ be an element of $\mathcal{C}_k(M)$ and let ω be in $\Omega^k(N)$. Then

$$\begin{aligned}
\ell_k(\varphi_*^k)_*(\mu)(\omega) &= f_{(\varphi_*^k)_*(\mu)}(\omega) \\
&= \int_{\tau\in \mathrm{C}^\infty(\Delta^k,N)} \left(\int_\tau \omega\right) d((\varphi_*^k)_*(\mu)) \\
&= \int_{\sigma\in \mathrm{C}^\infty(\Delta^k,M)} \left(\int_{\varphi_*^k(\sigma)} \omega\right) d\mu \\
&= \int_{\sigma\in \mathrm{C}^\infty(\Delta^k,M)} \left(\int_{\varphi\sigma} \omega\right) d\mu \\
&= \int_{\sigma\in \mathrm{C}^\infty(\Delta^k,M)} \left(\int_{\sigma} \varphi^*\omega\right) d\mu \\
&= f_\mu(\varphi_k^*(\omega)) \\
&= (\varphi_k^*)_*(f_\mu)(\omega) \quad = \quad (\varphi_k^*)_*\ell_k(\mu)(\omega).
\end{aligned}$$

Therefore, we have

$$\ell_k(\varphi_*^k)_* = (\varphi_k^*)_*\ell_k. \qquad \square$$

Theorem 11.6.3. *Let $M = B^n/\Gamma$ and $N = B^n/\mathrm{H}$ be compact orientable space-forms and let $\tilde{\varphi} : B^n \to B^n$ be a lift of a smooth homotopy equivalence $\varphi : M \to N$. If $u_0, \ldots, u_n$ are the vertices of a regular ideal n-simplex in B^n, then $\tilde{\varphi}_\infty(u_0), \ldots, \tilde{\varphi}_\infty(u_n)$ are the vertices of a regular ideal n-simplex in B^n.*

Proof: On the contrary, suppose that the ideal n-simplex spanned by $\tilde{\varphi}_\infty(u_0), \ldots, \tilde{\varphi}_\infty(u_n)$ is not regular. We pass to the upper half-space model U^n of hyperbolic space, and without loss of generality, we may assume that $u_i \neq \infty$ for each i. Let V_n be the volume of a regular ideal n-simplex. By Theorems 11.3.1, 11.3.2, and 11.6.2, there is an $\epsilon > 0$ and an $r > 0$ such that if $v_0, \ldots, v_n$ are the vertices of an n-simplex Δ in U^n, with $|u_i - v_i| < r$ for each i, then we have

$$\mathrm{Vol}(\mathrm{Str}(\tilde{\varphi}(\Delta))) < V_n - \epsilon.$$

Let

$$U_i = \{x \in U^n : |u_i - x| < r\}$$

and let

$$K_i = \{x \in U^n : |u_i - x| \leq r/2\}.$$

Let

$$U = \{g \in \mathrm{I}_0(U^n) : gK_\imath \subset U_\imath \quad \text{for each } i = 0, \ldots, n\}.$$

The set U is open in $\mathrm{I}_0(U^n)$, since the topology of $\mathrm{M}_0(U^n)$ corresponds under Poincaré extension to the compact-open topology on $\mathrm{M}_0(\hat{E}^{n-1})$.

Let $G = \mathrm{I}_0(U^n)$. Then the quotient map $\kappa : G \to \Gamma\backslash G$ is an open map, since it is a covering projection. Hence $\kappa(U)$ is an open subset of $\Gamma\backslash G$. Therefore $\mathrm{Vol}(\kappa(U)) > 0$.

Let ς be a straight singular n-simplex in M such that

$$|u_\imath - \tilde{\varsigma}(e_\imath)| \leq r/2 \quad \text{for each } i$$

and

$$\mathrm{Vol}(\tilde{\varsigma}(\Delta^n)) > V_n - \delta,$$

where

$$\delta = \epsilon \mathrm{Vol}(\kappa(U))/2\mathrm{Vol}(M).$$

Now if g is in U, then

$$\mathrm{Vol}(\mathrm{Str}(\tilde{\varphi} g \tilde{\varsigma})) < V_n - \epsilon < \mathrm{Vol}(\tilde{\varsigma}(\Delta^n)) + \delta - \epsilon,$$

whereas if g is not in U, then

$$\mathrm{Vol}(\mathrm{Str}(\tilde{\varphi} g \tilde{\varsigma})) < V_n < \mathrm{Vol}(\tilde{\varsigma}(\Delta^n)) + \delta.$$

We now assume that M and N are oriented so that $\varphi : M \to N$ is orientation preserving. By switching the indices of u_0 and u_1, if necessary, we may also assume that ς is orientation preserving.

Observe that

$$\begin{aligned}
& f_{(\mathrm{Str}^n)_*(\varphi^n_*)_*(\mathrm{Smr}(\tilde{\varsigma}))}(\Omega_N) \\
&= f_{(\mathrm{Str}^n \varphi^n_*)_*(\mathrm{Smr}(\tilde{\varsigma}))}(\Omega_N) \\
&= \int_{\tau \in \mathrm{C}^\infty(\Delta^n, N)} \left(\int_\tau \Omega_N \right) d\big((\mathrm{Str}^n \varphi^n_*)_*(\mathrm{Smr}(\tilde{\varsigma}))\big) \\
&= \int_{\sigma \in \mathrm{C}^\infty(\Delta^n, M)} \left(\int_{\mathrm{Str}^n \varphi^n_*(\sigma)} \Omega_N \right) d(\mathrm{Smr}(\tilde{\varsigma})) \\
&= \int_{\sigma \in \mathrm{C}^\infty(\Delta^n, M)} \left(\int_{\mathrm{Str}(\varphi\sigma)} \Omega_N \right) d\big((\tilde{\varsigma}^*)_*(d(\Gamma g))\big) \\
&= \int_{\Gamma g \in \Gamma\backslash G} \left(\int_{\mathrm{Str}(\varphi \tilde{\varsigma}^*(\Gamma g))} \Omega_N \right) d(\Gamma g) \\
&= \int_{\Gamma g \in \Gamma\backslash G} \left(\int_{\mathrm{Str}(\varphi \pi g \tilde{\varsigma})} \Omega_N \right) d(\Gamma g) \\
&= \int_{\Gamma g \in \Gamma\backslash G} \left(\int_{\mathrm{Str}(\eta \tilde{\varphi} g \tilde{\varsigma})} \Omega_N \right) d(\Gamma g)
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\Gamma g \in \Gamma\backslash G} \pm \mathrm{Vol}(\mathrm{Str}(\tilde{\varphi} g \tilde{\varsigma}))d(\Gamma g)\\
&< \big(\mathrm{Vol}(\tilde{\varsigma}(\Delta^n)) + \ \delta - \epsilon\big)\mathrm{Vol}(\kappa(U))\\
&\quad + \big(\mathrm{Vol}(\tilde{\varsigma}(\Delta^n)) + \delta\big)\big(\mathrm{Vol}(M) - \mathrm{Vol}(\kappa(U))\big)\\
&= (\mathrm{Vol}(\tilde{\varsigma}(\Delta^n)) + \delta)\mathrm{Vol}(M) - \epsilon\mathrm{Vol}(\kappa(U))\\
&= (\mathrm{Vol}(\tilde{\varsigma}(\Delta^n)) + \delta)\mathrm{Vol}(M) - 2\delta\mathrm{Vol}(M)\\
&= (\mathrm{Vol}(\tilde{\varsigma}(\Delta^n)) - \delta)\mathrm{Vol}(M).
\end{aligned}
$$

Let ρ be a reflection of U^n. Then we have

$$
\begin{aligned}
-f_{(\mathrm{Str}^n)_*(\varphi^n_*)_*(\mathrm{Smr}(\rho\tilde{\varsigma}))}(\Omega_N) &= -\int_{\Gamma g \in \Gamma\backslash G} \pm \mathrm{Vol}(\mathrm{Str}(\tilde{\varphi} g \rho \tilde{\varsigma}))d(\Gamma g)\\
&\leq V_n \mathrm{Vol}(M)\\
&< (\mathrm{Vol}(\tilde{\varsigma}(\Delta^n)) + \delta)\mathrm{Vol}(M).
\end{aligned}
$$

Therefore

$$
f_{\mathrm{Str}^n_*(\varphi^n_*)_*(\mathrm{Avg}(\tilde{\varsigma}))}(\Omega_N) < \mathrm{Vol}(\tilde{\varsigma}(\Delta^n))\mathrm{Vol}(M).
$$

Hence

$$
f_{\mathrm{Str}^n_*(\varphi^n_*)_*(\mathrm{Avg}(\tilde{\varsigma}))}(\Omega_N) = k\mathrm{Vol}(M) \quad \text{with } k < \mathrm{Vol}(\tilde{\varsigma}(\Delta^n)).
$$

Now by Lemma 10 of §11.5 and Theorem 11.4.4, we have

$$
\begin{aligned}
\ell_n \mathrm{Str}^n_*(\varphi^n_*)_*(\mathrm{Avg}(\tilde{\varsigma})) &= f_{\mathrm{Str}^n_*(\varphi^n_*)_*(\mathrm{Avg}(\tilde{\varsigma}))}(\Omega_N)\mathrm{Vol}(N)^{-1}F_N\\
&= k\mathrm{Vol}(M)\mathrm{Vol}(N)^{-1}F_N \ = \ kF_N;
\end{aligned}
$$

but by Theorems 11.5.2 and 11.5.4 and Lemma 12, we have

$$
\begin{aligned}
\ell_n \mathrm{Str}^n_*(\varphi^n_*)_*(\mathrm{Avg}(\tilde{\varsigma})) &= \ell_n(\varphi^n_*)_*(\mathrm{Avg}(\tilde{\varsigma}))\\
&= (\varphi^*_n)_*\ell_n(\mathrm{Avg}(\tilde{\varsigma}))\\
&= (\varphi^*_n)_*(\mathrm{Vol}(\tilde{\varsigma}(\Delta^n))F_M) \ = \ \mathrm{Vol}(\tilde{\varsigma}(\Delta^n))F_N,
\end{aligned}
$$

which is a contradiction. □

Rigidity

Lemma 13. *Let ρ be the reflection of B^n in the side S of a regular ideal n-simplex Δ in B^n. If $n > 2$, then Δ and $\rho\Delta$ are the only regular ideal n-simplices in B^n having S as a side.*

Proof: We pass to the upper half-space model U^n of hyperbolic space. Let $v_0, \ldots, v_n$ be the vertices of Δ with $v_0, \ldots, v_{n-1}$ the vertices of S. We may assume that $v_0 = \infty$ and $v_1, \ldots, v_n$ are in S^{n-2}. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then $\nu(\Delta)$ is a Euclidean regular $(n-1)$-simplex inscribed in S^{n-2} by Lemma 3 of §11.3; moreover $v_0, \ldots, v_{n-1}, v$ are the vertices of a regular ideal n-simplex if and only if $v_1, \ldots, v_{n-1}, v$ are the vertices of a Euclidean $(n-1)$-simplex in E^{n-1} inscribed in a unit $(n-2)$-sphere. Thus, either $v = v_n$ or v is the point obtained from v_n by reflecting E^{n-1} in the hyperplane spanned by $v_1, \ldots, v_{n-1}$. □

Lemma 14. *Let G be the group generated by the reflections in the sides of an ideal n-simplex Δ in B^n with vertices $u_0, \ldots, u_n$. Then the union of the orbits $Gu_0, \ldots, Gu_n$ is dense in S^{n-1}.*

Proof: On the contrary, assume that $U = \cup_{i=0}^n Gu_i$ is not dense in S^{n-1}. Then there is a point u of S^{n-1} and an open half-space H of B^n such that u is the center of the spherical disk

$$D = \overline{H} \cap S^{n-1} \quad \text{and} \quad D \subset S^{n-1} - U.$$

By Theorem 7.1.1, we have

$$\{g\Delta : g \in G\} = B^n.$$

Hence, there an element g of G such that $g\Delta$ meets H. Since $\overline{B}^n - H$ is hyperbolic convex, some vertex of $g\Delta$ meets D, which is a contradiction. Thus U is dense in S^{n-1}. □

Theorem 11.6.4. *Let $M = B^n/\Gamma$ and $N = B^n/\mathrm{H}$ be compact orientable space-forms, with $n > 2$, and let $\tilde{\varphi} : B^n \to B^n$ be a lift of a smooth homotopy equivalence $\varphi : M \to N$. Then $\tilde{\varphi}_\infty : S^{n-1} \to S^{n-1}$ is a Möbius transformation.*

Proof: Let Δ be a hyperbolic, regular, ideal n-simplex in B^n with vertices $u_0, \ldots, u_n$. By Theorem 11.6.3, we have that $\tilde{\varphi}_\infty(u_0), \ldots, \tilde{\varphi}_\infty(u_n)$ are the vertices of a regular ideal n-simplex Δ' in B^n. Let f be the unique Möbius transformation of S^{n-1} such that $fu_i = \tilde{\varphi}_\infty(u_i)$ for each i. Then $f^{-1}\tilde{\varphi}_\infty(u_i) = u_i$ for each i.

Let g_i be the reflection of B^n in the side of Δ opposite the vertex u_i. Then the points $u_0, \ldots, u_{i-1}, g_iu_i, u_{i+1}, \ldots, u_n$ are the vertices of the regular ideal n-simplex $g_i\Delta$ in B^n. Consequently, the points

$$\tilde{\varphi}_\infty(u_0), \ldots, \tilde{\varphi}_\infty(u_{i-1}), \tilde{\varphi}_\infty(g_iu_i), \tilde{\varphi}_\infty(u_{i+1}), \ldots, \tilde{\varphi}(u_n)$$

are the vertices of a regular ideal n-simplex $(g_i\Delta)'$ in B^n. Let h_i be the reflection of B^n in the side of Δ' opposite the vertex $\tilde{\varphi}_\infty(u_i)$. By Lemma 13, we have that

$$(g_i\Delta)' = h_i\Delta'.$$

Therefore, we have

$$\tilde{\varphi}_\infty(g_iu_i) = h_i\tilde{\varphi}_\infty(u_i).$$

Hence

$$f^{-1}\tilde{\varphi}_\infty(g_iu_i) = f^{-1}h_i\tilde{\varphi}_\infty(u_i) = f^{-1}h_iff^{-1}\tilde{\varphi}_\infty(u_i) = g_iu_i.$$

Thus $f^{-1}\tilde{\varphi}_\infty$ fixes g_iu_i for each i.

Let G be the group generated by $g_0, \ldots, g_n$. By induction, $f^{-1}\tilde{\varphi}_\infty$ fixes each point of $U = \cup_{i=0}^n Gu_i$. Moreover, the set U is dense in S^{n-1} by Lemma 14. Therefore $f^{-1}\tilde{\varphi}_\infty$ is the identity map of S^{n-1} by continuity. Hence $\tilde{\varphi}_\infty = f$. Thus $\tilde{\varphi}_\infty$ is a Möbius transformation of S^{n-1}. □

Theorem 11.6.5. (Mostow's rigidity theorem) *If $\varphi : M \to N$ is a homotopy equivalence between closed, connected, orientable, hyperbolic n-manifolds, with $n > 2$, then φ is a homotopic to an isometry.*

Proof: Without loss of generality, we may assume that M and N are hyperbolic space-forms, say $M = B^n/\Gamma$ and $N = B^n/\mathrm{H}$. Let $\pi : B^n \to M$ and $\eta : B^n \to N$ be the quotient maps. Let g be an element of Γ and let $\tilde{\varphi} : B^n \to B^n$ be a lift of φ. Then we have

$$\eta\tilde{\varphi}g = \varphi\pi g = \varphi\pi = \eta\tilde{\varphi}.$$

Hence, there is a unique element $\varphi_*(g)$ of H such that

$$\tilde{\varphi}g = \varphi_*(g)\tilde{\varphi}.$$

Moreover, if h is another element of Γ, then

$$\varphi_*(g)\varphi_*(h)\tilde{\varphi} = \varphi_*(g)\tilde{\varphi}h = \tilde{\varphi}gh.$$

Therefore, we have

$$\varphi_*(gh) = \varphi_*(g)\varphi_*(h).$$

Thus $\varphi_* : \Gamma \to \mathrm{H}$ is a homomorphism.

Let $\psi : N \to M$ be a homotopy inverse for φ. Then as in the proof of Theorem 11.6.1, we can choose a lift $\tilde{\psi} : B^n \to B^n$ such that $\tilde{\psi}\tilde{\varphi}$ is Γ-equivariant. Let g be an element of Γ. Then we have

$$g\tilde{\psi}\tilde{\varphi} = \tilde{\psi}\tilde{\varphi}g = \tilde{\psi}\varphi_*(g)\tilde{\varphi} = \psi_*(\varphi_*(g))\tilde{\psi}\tilde{\varphi}.$$

Therefore $g = \psi_*\varphi_*(g)$. Hence $\psi_*\varphi_* = id_\Gamma$. Therefore φ_* is injective and ψ_* is surjective. Moreover ψ_* is surjective regardless of the choice of $\tilde{\psi}$. By reversing the roles of φ and ψ, we obtain that φ_* is surjective. Therefore φ_* is an isomorphism.

Without loss of generality, we may assume that $\varphi : M \to N$ is smooth. By Theorems 11.6.1 and 11.6.2, we have

$$\tilde{\varphi}_\infty g = \varphi_*(g)\tilde{\varphi}_\infty$$

for each g in Γ by continuity. By Theorem 11.6.4, the map

$$\tilde{\varphi}_\infty : S^{n-1} \to S^{n-1}$$

is a Möbius transformation of S^{n-1}. Hence $\tilde{\varphi}_\infty$ extends to a Möbius transformation f of B^n such that $fg = \varphi_*(g)f$ for each g in Γ. Therefore

$$f\Gamma f^{-1} = \varphi_*(\Gamma) = \mathrm{H}.$$

By Theorem 8.1.5, we have that f induces an isometry $\overline{f} : M \to N$ defined by

$$\overline{f}(\Gamma x) = f\Gamma f^{-1}fx = \mathrm{H}fx.$$

We now pass to the hyperboloid model of hyperbolic space. Define a homotopy

$$F : H^n \times [0,1] \to H^n$$

by the formula

$$F(x,t) = \frac{(1-t)\tilde{\varphi}(x) + tf(x)}{|||(1-t)\tilde{\varphi}(x) + tf(x)|||}.$$

Then we have

$$F(g \times id) = \varphi_*(g)F \quad \text{for each } g \text{ in } \Gamma.$$

Hence F induces a homotopy

$$\overline{F} : M \times [0,1] \to N$$

from φ to $\overline{f}$. Thus φ is homotopic to an isometry. □

Corollary 1. *The hyperbolic structure on a closed, connected, orientable, hyperbolic n-manifold, with $n > 2$, is unique up to isometry homotopic to the identity.*

Exercise 11.6

1. Let k and ℓ be the constants in the definition of a pseudo-isometry of an unbounded metric space X. Prove that $k \geq 1$ and $\ell \geq 0$.
2. Let X be an unbounded metric space. Prove that a function $f : X \to X$ is a pseudo-isometry if and only if there are constants k and b such that
$$d(f(x), f(y)) \leq k\,d(x,y) \quad \text{for all } x, y \text{ in } X$$
and
$$k^{-1}d(x,y) \leq d(f(x), f(y)) \quad \text{if } d(x,y) \geq b.$$
3. Let L be a hyperbolic line of H^n and let $\rho : H^n \to L$ be the nearest point retraction. Prove that φ does not increase distances.
4. Let L be a hyperbolic line of D^n passing through 0, and let $\rho : D^n \to L$ be the nearest point retraction. Prove that ρ is the Euclidean orthogonal projection of D^n onto L.
5. Let ρ be as in Exercise 3 and let x, y, z be collinear points of H^n with y between x and z. Prove that $\rho(y)$ is between $\rho(x)$ and $\rho(z)$.
6. Let x be a point on a hyperbolic line L of H^n and suppose that $r > 0$. Prove that the sphere $S(x,r)$ is tangent to $\partial N(L,r)$.
7. Let $f : B^2 \to B^2$ be a pseudo-isometry. Prove that $f_\infty : S^1 \to S^1$ is a homeomorphism.
8. Let $u_0, \ldots, u_n$ be the vertices of a regular ideal n-simplex in B^n and let $v_0, \ldots, v_n$ be the vertices of a regular ideal n-simplex in B^n. Prove that there is a unique Möbius transformation of g of S^{n-1} such that $gu_i = v_i$ for each i.
9. Let G be the group generated by the reflections in the sides of a regular ideal n-simplex in B^n. Prove that G is discrete if and only if $n \leq 3$.
10. Let N^n/Γ and H^n/H be compact, orientable, hyperbolic space-forms, and let $\xi : \Gamma \to \mathrm{H}$ be an isomorphism. Prove that there is an element f of $\mathrm{I}(H^n)$ such that $\xi(g) = fgf^{-1}$ for each g in Γ.

§11.7. Historical Notes

§11.1. The Davis 120-cell space was constructed by Davis in his 1985 paper *A hyperbolic 4-manifold* [99]. Closed hyperbolic n-manifolds have been shown to exist in all dimensions n using the theory of quadratic forms. Examples can be found in Borel's 1963 paper *Compact Clifford-Klein forms of symmetric spaces* [53], Millson's 1976 paper *On the first Betti number of a constant negatively curved manifold* [289], and Gromov and Piatetski-Shapiro's 1988 paper *Non-arithmetic groups in Lobachevsky spaces* [171]. Moreover, it follows from the results of Millson's 1976 paper [289] that there are an infinite number of nonisometric, closed, hyperbolic n-manifolds for each dimension n.

Necessary and sufficient conditions for the complete gluing of a hyperbolic n-manifold from a finite family of polyhedra were essentially given by Seifert in his 1975 paper *Komplexe mit Seitenzuordnung* [371]. In particular, Theorem 11.1.2 implicitly appeared in this paper. Theorems 11.1.3 and 11.1.6 appeared in Thurston's 1979 lecture notes *The Geometry and Topology of 3-Manifolds* [389].

The hyperbolic 24-cell space was constructed by Tschantz and myself. Open, complete, hyperbolic n-manifolds of finite volume exist in all dimensions $n > 1$. Examples can be found in Millson's 1976 paper [289] and in Gromov and Piatetski-Shapiro's 1988 paper [171]. Moreover, it follows from the results of Millson's 1976 paper [289] that there are an infinite number of nonisometric, open, complete, hyperbolic n-manifolds of finite volume for each dimension $n > 1$. In contrast to dimension three, Wang has proved that for all $n > 3$, there are at most finitely many isometry classes of complete hyperbolic n-manifolds of volume less than any given bound in his 1972 paper *Topics on totally discontinuous groups* [403]. As references for n-dimensional hyperbolic manifolds, see Apanasov's 1991 treatise *Discrete Groups in Space and Uniformization Problems* [20] and Benedetti and Petronio's 1992 text *Lectures on Hyperbolic Geometry* [40].

§11.2. Theorem 11.2.1 for 3-dimensional compact polyhedra appeared in Weber and Seifert's 1933 paper *Die beiden Dodekaederräume* [405]. The 2- and 3-dimensional versions of Theorem 11.2.2 were essentially proved by Maskit in his 1971 paper *On Poincaré's theorem for fundamental polygons* [281].

§11.3. Theorem 11.3.1 was proved by Haagerup and Munkholm in their 1981 paper *Simplices of maximal volume in hyperbolic n-space* [174]. All the results in this section except for Theorem 11.3.2 appeared in this paper.

§11.4. The Gromov invariant was introduced by Gromov and Thurston in Thurston's 1979 lecture notes [389]. All the results of this section appeared in Thurston's lecture notes and in Gromov's 1982 paper *Volume and bounded cohomology* [169].

§11.5. Measure homology was introduced by Gromov and Thurston in Thurston's 1979 lecture notes [389]. All the results of this section appeared in Thurston's lecture notes and in Gromov's 1982 paper [169]. As a reference for the C^1 topology, see Hirsch's 1976 text *Differential Topology* [194].

§11.6. The concept of a *quasi-isometry* has it origins in Dehn's 1912 paper *Über unendliche diskontinuierliche Gruppen* [102] where he essentially proved that the fundamental group of a closed hyperbolic surface, with a word metric, quasi-isometrically embeds into the hyperbolic plane. For the n-dimensional version of this result, see Margulis' 1970 paper *Isometry of closed manifolds of constant negative curvature with the same fundamental group* [278] and Gromov and Pansu's 1991 survey *Rigidity of lattices: An introduction* [170]. See also Cannon's 1984 paper *The combinatorial structure of cocompact discrete hyperbolic groups* [67].

The concept of a *pseudo-isometry* was introduced by Mostow in his 1970 paper *The rigidity of locally symmetric spaces* [309]. Theorem 11.6.1 for homeomorphisms of a closed hyperbolic surface and the 2-dimensional version of Lemma 9 were essentially proved by Morse in his 1924 paper *A fundamental class of geodesics on any closed surface of genus greater than one* [307]. See also the Morse lemma in Gromov and Pansu's 1991 survey [170]. Theorem 11.6.2 for lifts of homeomorphisms of a closed hyperbolic surface appeared in Nielsen's 1924 paper *Über topologische Abbildungen geschlossener Fächen* [318]. See also Nielsen's 1927 paper *Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen* [319]. Lemmas 7-9 and Theorem 11.6.2 for quasi-isometries were proved by Efremovič and Tihomirova in their 1963 paper *Continuation of an equimorphism to infinity* [114]. Theorem 11.6.3 was proved by Gromov and appeared in Thurston's 1979 lecture notes [389] and in Munkholm's 1980 paper *Simplices of maximal volume in hyperbolic space, Gromov's norm, and Gromov's proof of Mostow's rigidity theorem (following Thurston)* [311]. Theorems 11.6.4 and 11.6.5 for diffeomorphisms were proved by Mostow in his 1968 paper *Quasi-conformal mappings in n-space and the rigidity of hyperbolic space forms* [308]. Theorems 11.6.4 and 11.6.5 were proved by Mostow in his 1973 study *Strong Rigidity of Locally Symmetric Spaces* [310].

All the essential material in this section appeared in Thurston's 1979 lecture notes [389] and in Munkholm's 1980 paper [311]. See also Gromov and Pansu's 1991 survey [170]. Mostow's rigidity theorem has been generalized to include complete hyperbolic n-manifolds, $n > 2$, of finite volume by Prasad in his 1973 paper *Strong rigidity of* $\mathbb{Q}$*-Rank 1 lattices* [344] and, more generally, to include complete hyperbolic n-manifolds, $n > 2$, whose volume grows radially slower than that of hyperbolic n-space by Sullivan in his 1980 paper *On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions* [384]. For a discussion on what happens to Mostow's rigidity theorem in dimension two, see Agard's 1985 article *Remarks on the boundary mapping for a Fuchsian group* [9].

CHAPTER 12

Geometrically Finite n-Manifolds

In this chapter, we study the geometry of geometrically finite hyperbolic n-manifolds. The chapter begins with a study of the limit set of a discrete group of Möbius transformations of B^n. In Section 12.3, we study geometrically finite groups of Möbius transformations of B^n. In Section 12.4, we study nilpotent groups of isometries of hyperbolic n-space. In Section 12.5, we prove the Margulis lemma. In Section 12.6, we apply the Margulis lemma to study the geometry of geometrically finite hyperbolic n-manifolds. In particular, we determine the global geometry of complete hyperbolic n-manifolds of finite volume.

§12.1. Limit Sets of Discrete Groups

In this section, we study some of the basic properties of the limit set of a discrete group of Möbius transformations of B^n. We shall denote the topological closure of a subset S of $\hat{E}^n$ by $\overline{S}$.

Definition: A point a of S^{n-1} is a *limit point* of a discrete subgroup Γ of $\mathrm{M}(B^n)$ if there is a point x of B^n and a sequence $\{g_i\}_{i=1}^{\infty}$ of elements of Γ such that $\{g_i x\}_{i=1}^{\infty}$ converges to a. The *limit set* of Γ is the set $L(\Gamma)$ of all limit points of Γ.

Theorem 12.1.1. *Let a be a point of S^{n-1} fixed by either a parabolic or hyperbolic element of a discrete subgroup Γ of $\mathrm{M}(B^n)$. Then a is a limit point of Γ.*

Proof: Let g be either a parabolic or hyperbolic element of Γ that fixes a. By replacing g with g^{-1}, if necessary, we may assume that a is the attractive fixed point of g. Then $g^i(0) \to a$ as $i \to \infty$. Hence a is a limit point of Γ. □

Theorem 12.1.2. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. Then for each point x of B^n, we have*

$$L(\Gamma) = \overline{\Gamma x} \cap S^{n-1}.$$

Proof: By definition, we have

$$\overline{\Gamma x} \cap S^{n-1} \subset L(\Gamma).$$

Suppose that a is a limit point of Γ. Then there is a sequence $\{g_i\}_{i=1}^{\infty}$ of elements of Γ and a point y of B^n such that $\{g_i y\}$ converges to a. Then we have

$$d(g_i x, g_i y) = d(x, y) \quad \text{for all } i.$$

Therefore $|g_i x - g_i y|$ goes to zero as $i \to \infty$ by Theorem 4.5.1. Hence

$$\lim_{i\to\infty} g_i x = \lim_{i\to\infty} g_i y = a.$$

Therefore a is in $\overline{\Gamma x} \cap S^{n-1}$. Thus $L(\Gamma) = \overline{\Gamma x} \cap S^{n-1}$. □

Theorem 12.1.3. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. Then the following are equivalent:*

(1) *The group Γ is elementary.*

(2) *The limit set $L(\Gamma)$ consists of 0, 1, or 2 points.*

(3) *The limit set $L(\Gamma)$ is finite.*

Proof: Suppose that Γ is elementary. Assume first that Γ is of elliptic type. Then Γ is finite by Theorem 5.5.2. Hence $\Gamma 0$ is finite and so $L(\Gamma)$ is empty. Assume next that Γ is of parabolic type. Let a be the fixed point of Γ. Then Γ leaves invariant the horosphere Σ based at a passing through 0 by Theorem 5.5.5. Hence

$$L(\Gamma) = \overline{\Gamma 0} \cap S^{n-1} \subset \overline{\Sigma} \cap S^{n-1} = \{a\}$$

and therefore $L(\Gamma) = \{a\}$ by Theorem 12.1.1.

Assume now that Γ is of hyperbolic type. Then Γ leaves invariant a hyperbolic line L of B^n by Theorem 5.5.6. Let a, b be the endpoints of L and let x be any point of L. Then

$$L(\Gamma) = \overline{\Gamma x} \cap S^{n-1} \subset \overline{L} \cap S^{n-1} = \{a, b\}.$$

By Theorem 5.5.8, the group Γ has a hyperbolic element h, and by Theorem 4.7.4, the axis of h is L. Hence $L(\Gamma) = \{a, b\}$ by Theorem 12.1.1. Thus (1) implies (2). Clearly (2) implies (3).

Suppose that $L(\Gamma)$ is finite. Assume first that $\Gamma 0$ is finite. As Γ is discontinuous, the stabilizer Γ_0 is finite. Therefore Γ is finite and so Γ is elementary. Now assume that $\Gamma 0$ is infinite. Then $\Gamma 0$ has a limit point a in the compact set $\overline{B}^n$. The point a is in S^{n-1}, since $\Gamma 0$ is a closed discrete subset of B^n by Theorem 5.3.4. As $\Gamma a \subset L(\Gamma)$, the orbit Γa is finite and so Γ is elementary. Thus (3) implies (1). □

Definition: A subset C of $\overline{B}^n$ is *hyperbolic convex* if and only if any two distinct points of C can be joined by either a hyperbolic line segment or a hyperbolic ray or a hyperbolic line contained in C.

Definition: The *hyperbolic convex hull* of a subset K of $\overline{B}^n$ is the intersection $C(K)$ of all the hyperbolic convex subsets of $\overline{B}^n$ that contain the set K.

Lemma 1. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$, let K be a closed Γ-invariant subset of S^{n-1}, and let $C(K)$ be its hyperbolic convex hull in $\overline{B}^n$. Then $C(K)$ is a closed Γ-invariant subset of $\overline{B}^n$.*

Proof: We pass to the projective disk model D^n. Then $C(K)$ is the Euclidean convex hull of K in E^n. It is a basic theorem in the theory of convex sets that the convex hull of a compact subset of E^n is compact. Hence $C(K)$ is compact and therefore $C(K)$ is closed.

Let g be in Γ. Then $C(K)$ is Γ-invariant, since

$$\begin{aligned} gC(K) &= g(\cap\{S : S \supset K \text{ and } S \text{ is a convex subset of } \overline{D}^n\}) \\ &= \cap\{gS : S \supset K \text{ and } S \text{ is a convex subset of } \overline{D}^n\} \\ &= \cap\{gS : gS \supset K \text{ and } gS \text{ is a convex subset of } \overline{D}^n\} \\ &= \cap\{S : S \supset K \text{ and } S \text{ is a convex subset of } \overline{D}^n\} \\ &= C(K). \end{aligned}$$

□

Theorem 12.1.4. *Let Γ be a nonelementary discrete subgroup of $\mathrm{M}(B^n)$. Then every nonempty, Γ-invariant, closed subset of S^{n-1} contains $L(\Gamma)$.*

Proof: Let K be a nonempty, Γ-invariant, closed subset of S^{n-1}. Then K is infinite, since Γ is nonelementary. Let $C(K)$ be the hyperbolic convex hull of K in $\overline{B}^n$. Then $C(K)$ is a Γ-invariant closed subset of $\overline{B}^n$ by Lemma 1. Moreover $C(K) \cap S^{n-1} = K$. Let x be any point of $C(K) \cap B^n$. Then $\Gamma x \subset C(K)$ and so

$$L(\Gamma) = \overline{\Gamma x} \cap S^{n-1} \subset C(K) \cap S^{n-1} = K.$$

Thus K contains $L(\Gamma)$. □

Lemma 2. *If Γ is a discrete subgroup of $\mathrm{M}(B^n)$ all of whose elements are elliptic, then Γ is finite.*

Proof: Every element of Γ is of finite order, since every element of Γ is elliptic and Γ is discontinuous. By Selberg's lemma, every finitely generated subgroup of Γ contains a torsion-free subgroup of finite index. Therefore, every finitely generated subgroup of Γ is finite. Given a finite subgroup H of Γ, let Fix(H) be the set of points fixed by every element of H. Then Fix(H) is an m-plane of B^n for some $m \geq 0$. Choose H such that dim Fix(H) is as small as possible. Now let g be any element of Γ and let K be the subgroup of Γ generated by g and the elements of H. Then K is finitely generated and therefore is finite. Now $\mathrm{Fix}(K) \subset \mathrm{Fix}(H)$. Hence, by the minimality of $\dim \mathrm{Fix}(H)$, we have that $\mathrm{Fix}(K) = \mathrm{Fix}(H)$. As g is arbitrary in Γ, we deduce that $\mathrm{Fix}(\Gamma) = \mathrm{Fix}(H)$. Thus Γ is elementary of elliptic type and so Γ is finite by Theorem 5.5.2. □

Theorem 12.1.5. *Let F be the set of all fixed points of nonelliptic elements of a discrete subgroup Γ of $\mathrm{M}(B^n)$. Then $\overline{F} = L(\Gamma)$.*

Proof: As $F \subset L(\Gamma)$, we have that $\overline{F} \subset L(\Gamma)$, since $L(\Gamma)$ is closed. If Γ is elementary, then $\overline{F} = L(\Gamma)$ by Theorem 12.1.3; and so we may assume that Γ is nonelementary. Then some element of Γ is nonelliptic by Lemma 2. Thus F is nonempty.

Let a be in F. Then a is fixed by some nonelliptic element h of Γ. If g is in Γ, then ghg^{-1} is nonelliptic and fixes ga. Hence F, and therefore $\overline{F}$, is Γ-invariant. Thus $\overline{F} = L(\Gamma)$, since $L(\Gamma)$ is the smallest nonempty, Γ-invariant, closed subset of S^{n-1} by Theorem 12.1.4. □

Lemma 3. *If g is either an elliptic or parabolic element of $\mathrm{M}(U^n)$ such that $g(\infty) \neq \infty$, then the isometric spheres of g and g^{-1} intersect.*

Proof: Let Σ_g and $\Sigma_{g^{-1}}$ be the isometric spheres of g and g^{-1}, respectively. By Theorem 4.4.4, the sphere Σ_g is orthogonal to E^{n-1} and $g = f\sigma$ where σ is the reflection in Σ and f is a Euclidean isometry that leaves U^n invariant. Now since

$$g^{-1} = \sigma f^{-1} = f^{-1}(f\sigma f^{-1}),$$

we find that $\Sigma_{g^{-1}} = f(\Sigma_g)$ by Theorem 4.3.3. Let H_g and $H_{g^{-1}}$ be the closed half-spaces of U^n bounded above by Σ_g and $\Sigma_{g^{-1}}$, respectively. Then

$$\begin{aligned} g(\overline{U}^n - \overline{H}_g \cup \overline{H}_{g^{-1}}) &\subset g(\overline{U}^n - \overline{H}_g) \\ &= f\sigma(\overline{U}^n - \overline{H}_g) \\ &\subset f(\overline{H}_g) \\ &= \overline{H}_{g^{-1}}. \end{aligned}$$

Hence g does not fix a point of the set $\overline{U}^n - (\overline{H}_g \cup \overline{H}_{g^{-1}})$. Therefore, the fixed points of g are in $\overline{H}_g \cup \overline{H}_{g^{-1}}$. By replacing g by g^{-1}, if necessary, we may assume that g fixes a point a of $\overline{H}_g$ and a is in H_g if g is elliptic. If a is in Σ_g, then Σ_g and $\Sigma_{g^{-1}}$ intersect at a. Assume next that a is inside Σ_g. Let Σ be the largest (horo)sphere (based) centered at a such that $\Sigma \subset H_g$. Then Σ meets Σ_g at a unique point b. As g leaves Σ invariant, we have that gb is in H_g, but gb is also in Σ_{g-1}. Therefore Σ_g and Σ_{g-1} intersect, since they have the same radius. □

Theorem 12.1.6. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$ all of whose elements are either elliptic or parabolic. Then Γ is elementary.*

Proof: If every element of Γ is elliptic, then Γ is elementary by Lemma 2. Now assume that Γ has a parabolic element f. We pass to the upper half-space model U^n and conjugate Γ in $\mathrm{M}(U^n)$ so that $f(\infty) = \infty$. Then f is a Euclidean isometry. We now prove that every element of Γ fixes ∞. On the contrary, suppose that g is an element of Γ such that $g(\infty) \neq \infty$. Let Σ_g be the isometric sphere of g. Then for each positive integer m, we have that $\Sigma_{f^m g} = \Sigma_g$ by Theorem 4.3.3. Moreover

$$\Sigma_{g^{-1}f^{-m}} = f^m g(\Sigma_{f^m g}) = f^m g(\Sigma_g) = f^m(\Sigma_{g-1}).$$

Since the cyclic group generated by f acts discontinuously on E^n, there is a positive integer m such that Σ_g and $f^m(\Sigma_{g-1})$ are disjoint. Hence $\Sigma_{f^m g}$ and $\Sigma_{g^{-1}f^{-m}}$ are disjoint. By Lemma 3, we have a contradiction. Thus, every element of Γ fixes ∞ and so Γ is elementary. □

Theorem 12.1.7. *Let F be the set of all fixed points of hyperbolic elements of a nonelementary discrete subgroup Γ of $\mathrm{M}(B^n)$. Then $\overline{F} = L(\Gamma)$.*

Proof: By Theorem 12.1.6, the set F is nonempty. Hence $\overline{F}$ is a nonempty, Γ-invariant, closed subset of $L(\Gamma)$, and therefore $\overline{F} = L(\Gamma)$ by Theorem 12.1.4. □

Theorem 12.1.8. *Let Γ be a nonelementary discrete subgroup of $\mathrm{M}(B^n)$. Then the limit set $L(\Gamma)$ is perfect and is therefore uncountable.*

Proof: Recall that a set is perfect if and only if it is closed and has no isolated points. On the contrary, suppose that $L(\Gamma)$ has an isolated point a. Then a is an isolated point of the set F of all fixed points of hyperbolic elements of Γ by Theorem 12.1.7. Hence a is fixed by some hyperbolic element h of Γ. As F is infinite, there is a b in F not fixed by h; but the set $\{h^k(b) : k \in \mathbb{Z}\}$ has a as a limit point, which is a contradiction. Thus $L(\Gamma)$ is perfect. It is well known that a nonempty perfect subset of E^n is uncountable. □

The Ordinary Set

We now study some of the basic properties of the complement of the limit set of a discrete group of Möbius transformations of B^n.

Definition: The *ordinary set* of a discrete subgroup Γ of $\mathrm{M}(B^n)$ is the set

$$O(\Gamma) = S^{n-1} - L(\Gamma).$$

A point of $O(\Gamma)$ is called an *ordinary point* of Γ.

Definition: A discrete subgroup Γ of $\mathrm{M}(B^n)$ is of the *first kind* if $O(\Gamma)$ is empty; otherwise Γ is of the *second kind.*

Example 1. Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$ such that B^n/Γ is compact. Then Γ is of the first kind. To see this, let x be a point of S^{n-1} and let P be a fundamental polyhedron for Γ containing 0. Then P is compact by Theorem 6.5.10. One can easily prove that there is a sequence $\{g_i\}_{i=1}^\infty$ of distinct elements of Γ such that $B(x, 1/i)$ contains $g_i P$ for all i. Therefore, the orbit $\Gamma 0$ accumulates at x. Thus $L(\Gamma) = S^{n-1}$.

Example 2. Every discrete elementary subgroup of $\mathrm{M}(B^n)$, with $n > 1$, is of the second kind by Theorem 12.1.3.

Example 3. Let Γ be a discrete subgroup of $\mathrm{M}(B^{n-1})$. Then Γ extends to a discrete subgroup $\tilde{\Gamma}$ of $\mathrm{M}(B^n)$ by Poincaré extension. Moreover

$$L(\tilde{\Gamma}) = L(\Gamma) \subset S^{n-2}$$

and so $\tilde{\Gamma}$ is of the second kind. In particular, if Γ is of the first kind, then $L(\tilde{\Gamma}) = S^{n-2}$.

Theorem 12.1.9. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$ of the second kind. Then*

(1) *the ordinary set $O(\Gamma)$ is an open dense subset of S^{n-1};*

(2) *the limit set $L(\Gamma)$ is a nowhere dense closed subset of S^{n-1}.*

Proof: (1) If Γ is elementary, then clearly $O(\Gamma)$ is a dense open subset of S^{n-1}. Now suppose that Γ is nonelementary. Then $\overline{O(\Gamma)}$ is a nonempty, Γ-invariant, closed subset of S^{n-1}. Therefore $\overline{O(\Gamma)}$ contains $L(\Gamma)$ by Theorem 12.1.4. Hence $\overline{O(\Gamma)} = S^{n-1}$.

(2) By (1), every neighborhood of a point in $L(\Gamma)$ contains a point of $O(\Gamma)$. Thus, the interior of $L(\Gamma)$ in S^{n-1} is empty and so $L(\Gamma)$ is nowhere dense in S^{n-1}. □

Nearest Point Retraction

Let K be a closed, nonempty, hyperbolic convex subset of $\overline{B}^n$. Let x be a point of $\overline{B}^n - K$. If K consists of a single point, then the *nearest point* of K to x is the single point of K, otherwise a *nearest point* of K to x is defined to be a point of K on the smallest (horo)sphere (based) centered at x that meets K. A nearest point y of K to x is unique, since if z were another nearest point, then the hyperbolic line segment $[y, z]$ would lie in K, and its interior (y, z) would lie inside the smallest (horo)sphere (based) centered at x that meets K, which would be a contradiction.

The *nearest point retraction* of $\overline{B}^n$ onto K is the function

$$\rho_K : \overline{B}^n \to K$$

defined such that $\rho_K(x)$ is the the nearest point of K to x. Note that if g is a Möbius transformation of B^n and x is a point of $\overline{B}^n$, then we have

$$\rho_{gK}(gx) = g\rho_K(x).$$

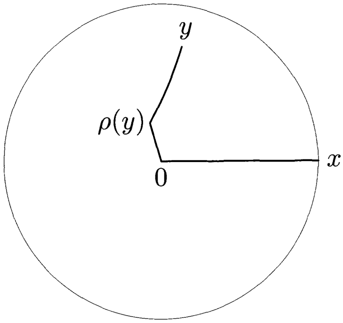


Figure 12.1.1. The nearest point retraction ρ applied to a point y

Lemma 4. *Let K be a closed, nonempty, hyperbolic convex subset of $\overline{B}^n$. Then the nearest point retraction $\rho_K : \overline{B}^n \to K$ is continuous on the set $B^n \cup (S^{n-1} - K)$.*

Proof: If $K \subset S^{n-1}$, then K is a single point, and so we may assume that K contains a point of B^n. Let x be a point of $B^n \cup (S^{n-1} - K)$. Then $\rho_K(x)$ is a point of B^n. By applying a Möbius transformation of B^n, we may assume, without loss of generality, that $\rho_K(x) = 0$. Let y be another point of $B^n \cup (S^{n-1} - K)$. In order to prove that $\rho = \rho_K$ is continuous at the point x, we will show that

$$|\rho(x) - \rho(y)| \leq |x - y|.$$

This is certainly true if $\rho(y) = 0$, so assume that $\rho(y) \neq 0$. As K is hyperbolic convex, the line segment $[0, \rho(y)]$ is contained in K. Moreover the angle between $[0, \rho(y)]$ and $[0, x]$ is at least $\pi/2$, since otherwise the smallest (horo)sphere (based) centered at x that meets K would meet the interior of $[0, \rho(y)]$ at a point of K nearer to x than 0. Likewise, by moving $\rho(y)$ to 0, we see that the angle between $[0, \rho(y)]$ and $[\rho(y), y]$ is at least $\pi/2$. Now let P and Q be the Euclidean hyperplanes passing through 0 and $\rho(y)$, respectively, perpendicular to $[0, \rho(y)]$. Then the points x and y are on opposite sides of the region between P and Q. Therefore

$$|\rho(x) - \rho(y)| \leq |x - y|.$$

See Figure 12.1.1. Hence ρ is continuous at the point x, and so ρ_K is continuous on the set $B^n \cup (S^{n-1} - K)$. □

Theorem 12.1.10. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. Then Γ acts discontinuously on $B^n \cup O(\Gamma)$.*

Proof: This is clear if $L(\Gamma)$ has 0 or 1 points, and so we assume that $L(\Gamma)$ has at least 2 points. Let $C(\Gamma)$ be the hyperbolic convex hull of $L(\Gamma)$ in

$\overline{B}^n$. Then $C(\Gamma)$ is a Γ-invariant closed subset of $\overline{B}^n$ by Lemma 1. Let

$$\rho : \overline{B}^n \to C(\Gamma)$$

be the nearest point retraction of $\overline{B}^n$ onto $C(\Gamma)$. Then $\rho(gx) = g\rho(x)$ for all g in Γ and x in $\overline{B}^n$, since $C(\Gamma)$ is Γ-invariant. Moreover ρ is continuous on $B^n \cup O(\Gamma)$ by Lemma 4.

Let K be a compact subset of $B^n \cup O(\Gamma)$. Then $\rho(K)$ is a compact subset of $C(\Gamma) - L(\Gamma)$. Let g be an element of Γ such that $K \cap gK \neq \emptyset$. Upon applying ρ to $K \cap gK$, we find that

$$\rho(K) \cap g\rho(K) \neq \emptyset.$$

By Theorem 5.3.5, the group Γ acts discontinuously on B^n. Therefore

$$\rho(K) \cap g\rho(K) \neq \emptyset$$

for only finitely many g in Γ, whence $K \cap gK \neq \emptyset$ for only finitely many g in Γ. Thus Γ acts discontinuously on $B^n \cup O(\Gamma)$. □

Remark: Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. The reason $O(\Gamma)$ is called the ordinary set of Γ is because $B^n \cup O(\Gamma)$ is the largest open subset of $\overline{B}^n$ on which Γ acts discontinuously. The proof is left as an exercise for the reader.

Theorem 12.1.11. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. Then for each x in $O(\Gamma)$, there is open neighborhood N of x in $B^n \cup O(\Gamma)$ such that for each g in Γ, either $N \cap gN = \emptyset$ or $gN = N$ and $gx = x$.*

Proof: Choose $r > 0$ so that

$$C(x, r) \cap \overline{B}^n \subset B^n \cup O(\Gamma).$$

Let $K = C(x,r) \cap \overline{B}^n$. Then K is a compact subset of $B^n \cup O(\Gamma)$. As Γ acts discontinuously on $B^n \cup O(\Gamma)$, there are only finitely many g in Γ such that $K \cap gK \neq \emptyset$. By shrinking r, if necessary, we may assume that $K \cap gK = \emptyset$ if $gx \neq x$. Now the stabilizer Γ_x is a finite group. By conjugating Γ in $\mathrm{M}(B^n)$, we may assume, without loss of generality, that Γ_x fixes 0. Then Γ_x is a subgroup of $\mathrm{O}(n)$ that fixes the line through 0 and x. Consequently, each element of Γ_x leaves $N = B(x, r) \cap \overline{B}^n$ invariant. □

Lemma 5. *Let P be a convex fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(B^n)$, and let $\{g_i\}_{i=1}^{\infty}$ be a sequence of distinct elements of Γ. Then the Euclidean diameter of $g_i\overline{P}$ goes to zero as i goes to infinity.*

Proof: Let $r > 0$. As $C(0, r)$ is compact, the ball $B(0, r)$ in B^n meets only finitely many members of $\{g\overline{P} : g \in \Gamma\}$, since P is locally finite. Therefore $\overline{B}^n - B(0, r)$ contains all but finitely many of the terms of $\{g_i\overline{P}\}_{i=1}^{\infty}$. As each $g_i\overline{P}$ is convex, the Euclidean diameters of all the $g_i\overline{P}$ in $B^n - B(0, r)$ are bounded above by a function of r that goes to zero as $r \to \infty$. Therefore $\mathrm{diam}_E(g_i\overline{P}) \to 0$ as $i \to \infty$. □

Theorem 12.1.12. *Let P be a convex fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(B^n)$. Then $\{gP : g \in \Gamma\}$ is a locally finite collection of subsets of $B^n \cup O(\Gamma)$.*

Proof: On the contrary, suppose that $\{gP : g \in \Gamma\}$ is not a locally finite collection of subsets of $B^n \cap O(\Gamma)$. Then there is a point a of $B^n \cap O(\Gamma)$ and a sequence $\{g_i\}_{i=1}^{\infty}$ of distinct elements of Γ such that $B(a, 1/i)$ contains a point x_i of g_iP. The point a is in $O(\Gamma)$, since $\{gP : g \in \Gamma\}$ is a locally finite collection of subsets of B^n. As the terms of $\{g_i\}$ are distinct, the Euclidean diameter of g_iP goes to zero as $i \to \infty$ by Lemma 5. As $x_i \to a$, we deduce that $g_ix \to a$ for any x in P. Therefore a is a limit point of Γ, which is a contradiction. □

Theorem 12.1.13. *Let P be a convex fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(B^n)$. Then*

$$O(\Gamma) = \bigcup_{g\in\Gamma} g(\overline{P} \cap O(\Gamma)).$$

Proof: Let x be a point of $O(\Gamma)$. Choose a sequence of points $\{x_i\}_{i=1}^{\infty}$ in B^n converging to x. Then for each i, there is a g_i in Γ such that x_i is in g_iP. Now only finitely many of the terms of $\{g_i\}_{i=1}^{\infty}$ are distinct by Theorem 12.1.12. Hence, there is a j such that x_i is in g_jP for infinitely many i. Therefore x is in $g_j\overline{P}$. Thus

$$O(\Gamma) = \bigcup_{g\in\Gamma} g(\overline{P} \cap O(\Gamma)).$$ □

We now give a characterization of the discrete subgroups of $\mathrm{M}(B^n)$ of the second kind in terms of the geometry of their convex fundamental polyhedra.

Theorem 12.1.14. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. Then the following are equivalent:*

(1) *The group Γ is of the second kind.*

(2) *Every convex fundamental polyhedron for Γ contains a closed half-space of B^n.*

(3) *The group Γ has a convex fundamental polyhedron that contains a closed half-space of B^n.*

Proof: Suppose that Γ is of the second kind. Let P be a convex fundamental polyhedron for Γ. By Theorem 12.1.13, we have

$$O(\Gamma) = \bigcup_{g\in\Gamma} g(\overline{P} \cap O(\Gamma)).$$

Now Γ is countable, since Γ is discrete. As $O(\Gamma)$ is locally compact, $O(\Gamma)$ is a Baire space. Therefore, one of the closed subsets $g(\overline{P} \cap O(\Gamma))$ of $O(\Gamma)$ has a nonempty interior in $O(\Gamma)$. Hence, the interior of $\overline{P} \cap O(\Gamma)$ in $O(\Gamma)$

is nonempty. Let x be a point of the interior of $\overline{P} \cap O(\Gamma)$. Then there is an $r > 0$ so that

$$B(x,r) \cap S^{n-1} \subset \overline{P} \cap O(\Gamma).$$

By convexity, the closed half-space of B^n bounded by $B(x,r) \cap S^{n-1}$ is contained in P. Thus (1) implies (2). Clearly (2) implies (3).

Suppose that Γ has a fundamental polyhedron that contains a closed half-space B^n. Then there is a point x of S^{n-1} and an $r > 0$ such that

$$B(x,r) \cap \overline{B}^n \subset \overline{P}.$$

As the sets $\{gP^\circ : g \in \Gamma\}$ are mutually disjoint, the sets

$$\{g(B(x,r) \cap \overline{B}^n) : g \in \Gamma\}$$

are mutually disjoint. Hence, no point of $B(x,r) \cap S^{n-1}$ is fixed by a nonidentity element of Γ. By Theorem 12.1.5, we have that

$$B(x,r) \cap S^{n-1} \subset O(\Gamma).$$

Therefore Γ is of the second kind. Thus (3) implies (1). □

Definition: Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. The *volume* of B^n/Γ is the volume of any proper fundamental domain for Γ in B^n.

Note that the volume of B^n/Γ is well defined, since all the proper fundamental domains for Γ have the same volume by Theorem 6.5.5. The next theorem follows immediately from Theorem 12.1.14.

Theorem 12.1.15. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$ such that the volume of B^n/Γ is finite. Then Γ is of the first kind.*

Theorem 12.1.16. *Let H be an infinite normal subgroup of a nonelementary discrete subgroup Γ of $\mathrm{M}(B^n)$. Then $L(\mathrm{H}) = L(\Gamma)$.*

Proof: Let F_H be the set of all fixed points of nonelliptic elements of H. Then F_H is nonempty by Lemma 2. Given an element h of H, let F_h be the fixed set of h. If g is in Γ, then $gF_h = F_{ghg^{-1}}$. Therefore F_H is a Γ-invariant subset of S^{n-1}. Hence $\overline{F}_\mathrm{H}$ is a nonempty, closed, Γ-invariant subset of S^{n-1}. Therefore

$$L(\Gamma) \subset \overline{F}_\mathrm{H} = L(\mathrm{H}) \subset L(\Gamma)$$

by Theorems 12.1.4 and 12.1.5. □

Example 4. In §10.3, we constructed a complete hyperbolic 3-manifold M of finite volume that is homeomorphic to the complement of the figure-eight knot K in $\hat{E}^3$. By Theorem 8.5.9, there is a discrete subgroup Γ of $\mathrm{M}(B^3)$ such that B^3/Γ is isometric to M. By Theorem 8.1.4, the group Γ is isomorphic to the fundamental group of B^3/Γ. It is a basic fact of knot theory that the commutator subgroup of $\pi_1(M)$ is a free group of rank 2 and the abelianization of $\pi_1(M)$ is infinite cyclic. Therefore, the commutator subgroup Γ' of Γ is a free group of rank 2 and Γ/Γ' is infinite cyclic. Now the group Γ/Γ' acts freely and discontinuously as a group of isometries on B^3/Γ' and the orbit space $(B^3/\Gamma')/(\Gamma/\Gamma')$ is B^3/Γ. By Theorem 8.1.3, the quotient map

$$\pi : B^3/\Gamma' \to B^3/\Gamma$$

is a local isometry and a covering projection. As π is an infinite covering, B^3/Γ' has infinite volume. Nevertheless Γ' is of the first kind because of Theorems 12.1.15 and 12.1.16.

Theorem 12.1.17. *Let Γ be a finitely generated, nonelementary, discrete subgroup of $\mathrm{M}(B^n)$ that leaves no m-plane of B^n invariant for $m < n-1$. Then the normalizer N of Γ in $\mathrm{M}(B^n)$ is discrete.*

Proof: Let $\{g_1, \ldots, g_m\}$ be a set of generators for Γ with $g_1 = 1$. Let x be a point of B^n that is fixed only by the identity element of Γ. Set

$$s = \mathrm{dist}(x, \Gamma x - \{x\}).$$

Let

$$U = \{\phi \in \mathrm{M}(B^n) : d(\phi(g_i x), g_i x) < s/2 \quad \text{for} \quad i = 1, \ldots, m\}.$$

Then U is an open neighborhood of the identity in $\mathrm{M}(B^n)$.

Suppose that h is an element of $\mathrm{N} \cap U$. Then we have

$$\begin{aligned} d(g_i^{-1}h^{-1}g_i hx, x) &= d(g_i hx, hg_i x) \\ &\leq d(g_i hx, g_i x) + d(g_i x, hg_i x) \\ &= d(hx, x) + d(g_i x, hg_i x) \quad < \quad s. \end{aligned}$$

Hence $g_i^{-1}h^{-1}g_i hx = x$ and so $g_i^{-1}h^{-1}g_i h = 1$. Therefore h and g_i commute for each $i = 1, \ldots, m$. As $g_1, \ldots, g_m$ generate Γ, we have that h commutes with every element of Γ.

Now let y be an arbitrary point of $L(\Gamma)$. Then there is a sequence $\{f_i\}$ of elements of Γ such that $f_i x \to y$. Observe that for each i, we have

$$d(f_i x, hf_i x) = d(f_i x, f_i hx) = d(x, hx).$$

Consequently, we have

$$\lim_{i\to\infty} |f_i x - hf_i x| = 0.$$

Therefore $hy = y$. Thus h is the identity on $L(\Gamma)$.

Let m be the least integer such that $L(\Gamma)$ is contained in an $(m-1)$-sphere of S^{n-1}. By conjugating Γ, we may assume that $L(\Gamma) \subset S^{m-1}$. As Γ leaves the convex hull $C(\Gamma)$ of $L(\Gamma)$ invariant, Γ also leaves $\overline{B}^m$ invariant, since $\overline{B}^m$ is the affine hull of $C(\Gamma)$. By our hypothesis, $m = n-1$ or n.

Assume first that $m = n$. Then we can choose points $y_0, \ldots, y_n$ of $L(\Gamma)$ that are the vertices of an ideal n-simplex. As $hy_i = y_i$ for each i, we deduce that $h = 1$. Hence $\mathrm{N} \cap U = \{1\}$.

Now assume that $m = n - 1$. Then we can conclude as above that h is the identity on B^{n-1}. Therefore h is either the identity or the reflection ρ of B^n in the hyperplane B^{n-1}. Therefore, we have

$$\mathrm{N} \cap (U - \{\rho\}) = \{1\}.$$

Hence, the identity is open in N, and therefore N is discrete. □

Classical Schottky Groups

Let Γ be a subgroup of $\mathrm{M}(B^n)$. An open subset D of B^n is called a Γ-*packing* if $D \cap gD = \emptyset$ for all $g \neq 1$ in Γ.

Theorem 12.1.18. *Let $\Gamma_1, \ldots, \Gamma_m$ be subgroups of $\mathrm{M}(B^n)$ whose union generates the group Γ, and let D_i be a Γ_i-packing for each $i = 1, \ldots, m$ such that $D = \bigcap_{i=1}^{m} D_i$ is nonempty and $D_i \cup D_j = B^n$ when $i \neq j$. Then*

(1) *the group Γ is the free product of the groups $\Gamma_1, \ldots, \Gamma_m$;*

(2) *the set D is a Γ-packing;*

(3) *the group Γ is discrete.*

Proof: (1) Let $g_k \neq 1$ be in Γ_{i_k} for each $k = 1, \ldots, \ell$ and suppose that $i_k \neq i_{k+1}$ for each $k = 1, \ldots, \ell - 1$. We now prove by induction that

$$g_\ell \cdots g_1(D) \subset B^n - D_{i_\ell}.$$

First of all,

$$g_1(D) \subset g_1(D_{i_1}) \subset B^n - D_{i_1}.$$

Assume that $k < \ell$ and

$$g_k \cdots g_1(D) \subset B^n - D_{i_k}.$$

Then we have

$$\begin{aligned} g_{k+1}g_k \cdots g_1(D) &\subset g_{k+1}(B^n - D_{i_k}) \\ &\subset g_{k+1}(D_{i_{k+1}}) \subset B^n - D_{i_{k+1}}. \end{aligned}$$

This completes the induction. Therefore

$$g_\ell \cdots g_1(D) \subset B^n - D_{i_\ell} \subset B^n - D.$$

This shows that $g_\ell \cdots g_1 \neq 1$. Therefore Γ is the free product of $\Gamma_1, \ldots, \Gamma_m$.

(2) Now suppose that $g \neq 1$ in Γ. Then there exist $g_1, \ldots, g_\ell$ as above so that $g = g_\ell \cdots g_1$. Hence $D \cap gD = \emptyset$ by (1). Thus D is a Γ-packing.

(3) Now let x be a point of D. Then x is open in Γx, since D is a Γ-packing by (2). Let $\varepsilon : \Gamma \to \Gamma x$ be the evaluation map at x. Then ε is continuous. Therefore $\varepsilon^{-1}(x) = 1$ is open in Γ, and so Γ is discrete. □

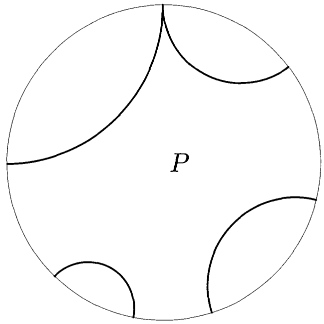


Figure 12.1.2. A Schottky polygon P in B^2

A *Schottky polyhedron* in B^n is a convex polyhedron P in B^n, with an even number of sides, each of which is a hyperplane of B^n. See Figure 12.1.2. Let Φ be a $\mathrm{M}(B^n)$-side-pairing for a Schottky polyhedron P in B^n, with $2m$ sides, such that no side of P is paired to itself. The group Γ generated by Φ is called a *classical Schottky subgroup* of $\mathrm{M}(B^n)$ of rank m.

Theorem 12.1.19. *Let Γ be a classical Schottky subgroup of $\mathrm{M}(B^n)$ of rank m. Then Γ is a free discrete subgroup of $\mathrm{M}(B^n)$ of rank m.*

Proof: Let Γ be generated by a $\mathrm{M}(B^n)$-side-pairing Φ for a Schottky polyhedron P in B^n, with $2m$ sides, such that no side of P is paired to itself. Then we can order the sides of P as follows:

$$S_1, \ldots, S_m, \ S'_1, \ldots, S'_m.$$

Moreover Γ is generated by the elements $g_{S_1}, \ldots, g_{S_m}$. Let $\Gamma_i = \langle g_{S_i} \rangle$ and let P_i be the convex polyhedron in B^n with S_i and S'_i as its only sides for each $i = 1, \ldots, m$. Then P_i° is a Γ_i-packing and Γ_i is infinite cyclic for each $i = 1, \ldots, m$. Moreover $P^\circ = \cap_{i=1}^m P_i^\circ$ is nonempty and $P_i^\circ \cup P_j^\circ = B^n$ when $i \neq j$. By Theorem 12.1.18, the group Γ is discrete and the free product of $\Gamma_1, \ldots, \Gamma_m$. Thus Γ is a free group of rank m. □

Example 5. Consider the Schottky polyhedron P in U^n whose sides are the vertical planes $x_1 = 1$ and $x_1 = 2$. Then the element h of $\mathrm{M}(U^n)$, defined by $hx = 2x$, pairs the sides of P. Observe that the set

$$\cup\{h^k(P) : k \in \mathbb{Z}\}$$

is the open half-space, $x_1 > 0$, in U^n. Therefore P is not a fundamental polyhedron for the Schottky group Γ generated by h.

Example 5 shows that a Schottky polyhedron P is not necessarily a fundamental polyhedron for a Schottky group generated by a side-pairing of P. On the other hand, we have the following theorem.

Theorem 12.1.20. *Let P be a Schottky polyhedron in B^n such that no two sides of P meet at infinity, and let Γ be a Schottky group generated by a $\mathrm{M}(B^n)$-side-pairing Φ for P such that no side is paired to itself. Then P is an exact, convex, fundamental polyhedron for Γ, and the inclusion of P into B^n induces an isometry from the hyperbolic n-manifold P/Γ, obtained by gluing together the sides of P by Φ, to the space-form B^n/Γ.*

Proof: The theorem follows immediately from Theorems 11.1.6 and 11.2.1, since P has no cusp points. □

We next show that the Schottky groups in Theorem 12.1.20 have interesting limit sets.

Theorem 12.1.21. *Let P be a Schottky polyhedron in B^n such that P has at least four sides and no two sides of P meet at infinity, and let Γ be a Schottky group generated by a $\mathrm{M}(B^n)$-side-pairing Φ for P such that no side is paired to itself. Then $L(\Gamma)$ is a Cantor set.*

Proof: Let S be a side of P. Since S and S' do not meet at infinity, the side-pairing transformation g_S is hyperbolic and its fixed points are on opposite sides of S and S'. Let T be a side of P distinct from S and S'. Then g_T is hyperbolic and its fixed points are on opposite sides of T and T'. Hence g_S and g_T do not have a common fixed point. Therefore Γ is nonelementary by Theorem 12.1.3. Hence $L(\Gamma)$ is perfect by Theorem 12.1.8. As every perfect, totally disconnected, compact, metric space is a Cantor set, it remains only to show that $L(\Gamma)$ is totally disconnected.

We begin by showing that

$$\overline{P} \cap S^{n-1} \subset O(\Gamma).$$

Assume first that $\overline{P}$ contains a point a fixed by some $g \neq 1$ in Γ. Then $\overline{P}$ and $g\overline{P}$ meet at a. Hence P and gP share a side S, and so $g = g_S$. As $g_S^{-1}(S) = S'$, the sides S and S' meet at infinity at a, which is a contradiction. Therefore $\overline{P}$ contains no fixed points of nonidentity elements of Γ.

Now assume that $\overline{P}$ contains a limit point b of Γ. As the interior of $\overline{P} \cap S^{n-1}$ is contained in $O(\Gamma)$, the point b is in the closure of a side S of P. Choose $r > 0$ so that

$$B(b,r) \cap S^{n-1} \subset (\overline{P} \cup g_S(\overline{P})) \cap S^{n-1}.$$

By Theorem 12.1.5, there is a point c of $B(b,r)$ that is fixed by a nonidentity element of Γ. As the interiors of $\overline{P} \cap S^{n-1}$ and $g_S(\overline{P}) \cap S^{n-1}$ are contained in $O(\Gamma)$, the point c must be in the closure of S. But $\overline{P}$ contains no fixed points of nonidentity elements of Γ, and so we have a contradiction. Thus, we have that

$$\overline{P} \cap S^{n-1} \subset O(\Gamma).$$

Let $\mathcal{P} = \{gP : g \in \Gamma\}$. Then $\mathcal{P}$ is an exact tessellation of B^n by Theorem 12.1.20. Therefore $\mathcal{P}$ is connected by Theorem 6.7.2. Define a sequence of convex polyhedra $P_1 \subset P_2 \subset \cdots$ inductively as follows. Let $P_1 = P$. Assume that P_i has been defined. Let P_{i+1} be the union of P_i and the polyhedra in $\mathcal{P}$ that share a side with P_i. Then for each i, the polyhedron P_i is a finite union of polyhedra in $\mathcal{P}$, and every side of P_i is a hyperplane of B^n. Moreover, since $\mathcal{P}$ is connected, we have

$$\bigcup_{i=1}^{\infty} P_i = \cup\mathcal{P}.$$

Now since

$$\overline{P} \cap S^{n-1} \subset O(\Gamma)$$

and $O(\Gamma)$ is Γ-invariant, we have that

$$\overline{P}_i \cap S^{n-1} \subset O(\Gamma)$$

for each i. Therefore

$$L(\Gamma) \subset \bigcap_{i=1}^{\infty} (\overline{B}^n - \overline{P}_i).$$

Let u and v be distinct limit points of Γ and let L be the hyperbolic line of B^n with end points u and v. Since

$$\bigcap_{i=1}^{\infty} (\overline{B}^n - \overline{P}_i) \subset S^{n-1},$$

there is an i such that $\overline{B}^n - \overline{P}_i$ does not contain L. Then by convexity, u and v lie in different components of $\overline{B}^n - \overline{P}_i$. Let U be the component of $S^{n-1} - \overline{P}_i$ containing u, and let V be the union of the remaining components of $S^{n-1} - \overline{P}_i$. Then U and V are disjoint open neighborhoods in S^{n-1} of u and v, respectively, such that

$$L(\Gamma) \subset U \cup V.$$

Therefore u and v lie in different components of $L(\Gamma)$. Thus $L(\Gamma)$ is totally disconnected. □

Exercise 12.1

1. Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$ with a parabolic element and let F be the set of all fixed points of parabolic elements of Γ. Prove that $L(\Gamma) = \overline{F}$.
2. Let Γ be a nonelementary discrete subgroup of $\mathrm{M}(B^n)$. Prove that Γ has an infinite number of hyperbolic elements, no two of which have a common fixed point.
3. Let g be an element of $\mathrm{M}(B^n)$ such that for some x in S^{n-1} and radius r with $0 < r < 1$, we have

$$g(C(x, r) \cap S^{n-1}) \subset B(x, r).$$

   Prove that g is hyperbolic and that g fixes a point of $B(x, r) \cap S^{n-1}$.

4. Let Γ be a nonelementary discrete subgroup of $M(B^n)$ and let x, y be distinct limit points of Γ. Prove that for each $r > 0$, there is a hyperbolic element h of Γ such that $B(x, r)$ contains one of the fixed points of h and $B(y, r)$ contains the other.
5. Prove that a perfect subset of E^n is uncountable.
6. Let Γ be a discrete subgroup of $M(B^n)$ such that B^n/Γ is compact. Prove that Γ is of the first kind by the argument sketched in Example 1.
7. Let Γ be a nonelementary discrete subgroup of $M(B^n)$, let P be an m-plane of B^n, with $m > 1$, and suppose that Γ leaves no ℓ-plane of B^n invariant for all $\ell < m$. Prove that Γ leaves P invariant if and only if $L(\Gamma) \subset \overline{P} \cap S^{n-1}$.
8. Let K be a closed hyperbolic convex subset of $\overline{B}^n$ that contains a point of B^n and let $\rho_K : \overline{B}^n \to K$ be the nearest point retraction. Prove that if x, y are in B^n, then $d(\rho_K(x), \rho_K(y)) \leq d(x, y)$.
9. Let K be a closed, nonempty, hyperbolic convex subset of $\overline{B}^n$. Prove that the nearest point retraction $\rho_K : \overline{B}^n \to K$ is continuous.
10. Let Γ be a discrete subgroup of $M(B^n)$ and let U be an open subset of S^{n-1} on which Γ acts discontinuously. Prove that $O(\Gamma)$ contains U. Conclude that $B^n \cup O(\Gamma)$ is the largest open subset of $\overline{B}^n$ on which Γ acts discontinuously.
11. Let Γ be a discrete subgroup of $M(B^n)$. Prove that a point x of S^{n-1} is in $O(\Gamma)$ if and only if there is an open neighborhood U of x in S^{n-1} such that $U \cap gU \neq \emptyset$ for only finitely many g in Γ.
12. Let Γ be a discrete subgroup of $M(B^n)$ and let H be a subgroup of Γ of finite index. Prove that $L(\mathrm{H}) = L(\Gamma)$.
13. Prove that the free group in Example 4 is not a classical Schottky subgroup of $M(B^3)$.
14. Let $g_1, \ldots, g_m$ be nonelliptic elements of $M(B^n)$ such that no two elements have a common fixed point. Prove that there are positive integers $k_1, \ldots, k_m$ such that $g_1^{k_1}, \ldots, g_m^{k_m}$ generate a classical Schottky group of rank m.
15. Let Γ be a nonelementary discrete subgroup of $M(B^n)$. Prove that Γ contains a classical Schottky group of rank m for each m.

§12.2. Limit Points of Discrete Groups

In this section, we study the basic properties of conical and cusped limit points of a discrete group of Möbius transformations of B^n.

Conical Limit Points

Definition: A point a of S^{n-1} is a *conical limit point* of a discrete subgroup Γ of $M(B^n)$ if there is a point x of B^n, a sequence $\{g_i\}_{i=1}^{\infty}$ of elements of Γ, a hyperbolic ray R in B^n ending at a, and an $r > 0$ such that $\{g_i x\}_{i=1}^{\infty}$ converges to a within the r-neighborhood $N(R, r)$ of R in B^n.

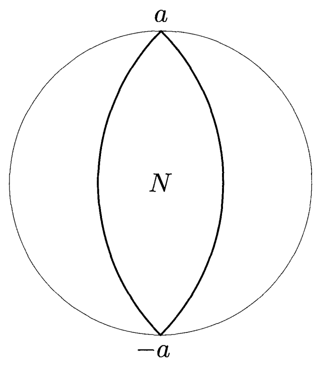


Figure 12.2.1. An r-neighborhood N of the line $(-a, a)$ of B^2

Figure 12.2.1 illustrates the r-neighborhood of a diameter of B^2. In the upper half-space model U^n, an r-neighborhood of a vertical line L of U^n is the interior of a Euclidean hypercone in U^n with L as its axis. Thus ∞ is a conical limit point of a discrete subgroup Γ of $\mathrm{M}(U^n)$ if and only if there is a point x of U^n and a sequence $\{g_i\}_{i=1}^\infty$ of elements of Γ such that $\{g_i x\}_{i=1}^\infty$ converges to ∞ within a Euclidean hypercone in U^n whose axis is a vertical line of U^n. See Figure 12.2.2.

Theorem 12.2.1. *Let a be a point of S^{n-1} fixed by a hyperbolic element h of a discrete subgroup Γ of $\mathrm{M}(B^n)$. Then a is a conical limit point of Γ.*

Proof: By replacing h with h^{-1}, if necessary, we may assume that a is the attractive fixed point of h. Let x be any point on the axis L of h. Then $\{h^i x\}_{i=1}^\infty$ converges to a within any r-neighborhood of L in B^n. Thus a is a conical limit point of Γ. $\square$

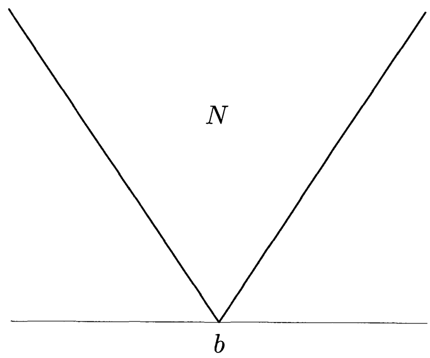


Figure 12.2.2. An r-neighborhood N of the line (b, ∞) of U^2

We next prove that the point x in the definition of a conical limit point plays no special role.

Theorem 12.2.2. *Let a be a conical limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$, let x be a point of B^n, let $\{g_i\}_{i=1}^\infty$ be a sequence of elements of Γ, let R be a hyperbolic ray in B^n ending in a, and let $r > 0$ be such that $\{g_i x\}_{i=1}^\infty$ converges to a within $N(R, r)$. Then for each point y of B^n, there is an $s > 0$ such that $\{g_i y\}_{i=1}^\infty$ converges to a within $N(R, s)$.*

Proof: Let $s = d(x, y) + r$. For each i, there is a point z_i on R such that $d(g_i x, z_i) < r$. Hence

$$d(g_i y, z_i) \leq d(g_i y, g_i x) + d(g_i x, z_i) < d(y, x) + r = s.$$

Hence $g_i y$ is in $N(R, s)$ for each i, and so $\{g_i y\}$ converges to a within $N(R, s)$. □

Theorem 12.2.3. *Let a be a limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$ and let $\{g_i\}_{i=1}^\infty$ be a sequence of distinct elements of Γ. Then the following are equivalent:*

(1) *For some (or each) hyperbolic ray R in B^n ending at a, there is an $r > 0$ such that $\{g_i(0)\}$ converges to a within $N(L, r)$.*

(2) *For some (or each) hyperbolic ray R in B^n ending at a, there is a compact subset K of B^n such that for all i, we have*

$$K \cap g_i^{-1} R \neq \emptyset.$$

Proof: Suppose there is a hyperbolic ray R ending at a such that $\{g_i(0)\}$ converges to a within $N(R, r)$ for some $r > 0$. Let S be another hyperbolic ray ending at a. Then there is an $s > 0$ such that $N(R, r) \subset N(S, s)$. Therefore $\{g_i(0)\}$ converges to a within $N(S, s)$. Thus, the quantifiers "for some" and "each" are equivalent in (1).

Let R be a hyperbolic ray ending at a. Then for any g in Γ and $r > 0$, one has $d(g(0), R) \leq r$ if and only if $g^{-1}R$ meets the compact set $C(0, r)$. Thus (1) and (2) are equivalent. □

Theorem 12.2.4. *A conical limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$ cannot lie on the Euclidean boundary of any convex fundamental polyhedron for Γ.*

Proof: On the contrary, suppose that a conical limit point a of Γ lies on the Euclidean boundary of a convex fundamental polyhedron P for Γ. By Theorem 6.3.21, there is a hyperbolic ray R in P ending at a. By Theorem 12.2.3, there is a sequence $\{g_i\}_{i=1}^\infty$ of distinct elements of Γ and a compact subset K of B^n such that for all i, we have

$$K \cap g_i P \neq \emptyset.$$

But this contradicts the fact that P is locally finite. □

Corollary 1. *A fixed point of a hyperbolic element of a discrete subgroup* Γ *of* $\mathrm{M}(B^n)$ *cannot lie on the Euclidean boundary of any convex fundamental polyhedron for* Γ.

Theorem 12.2.5. *Let* a *be a limit point of a discrete subgroup* Γ *of* $\mathrm{M}(B^n)$, *let* $\{g_i\}_{i=1}^{\infty}$ *be a sequence of distinct elements of* $\Gamma - \Gamma_0$, *and let* r_i *be the radius of the isometric sphere of* g_i *for each* i. *Then the following are equivalent:*

(1) $d(g_i(0), [0, a)) = O(1)$.

(2) $|a - g_i(0)| = O(1 - |g_i(0)|)$.

(3) $|a - g_i(0)| = O(r_i^2)$.

(4) *The sequence* $\{g_i(0)\}_{i=1}^{\infty}$ *converges to* a *and for some (or each)* $x \neq a$ *in* S^{n-1} *there is a* $\delta > 0$ *such that for all* i, *we have*
$$|g_i^{-1}x - g_i^{-1}a| \geq \delta.$$

Proof: Assume that (1) holds. Let $R = [0, a)$ and $L = (-a, a)$. Then $\{g_i(0)\}$ converges to a. Hence, for all sufficiently large i, we have
$$d(g_i(0), R) = d(g_i(0), L).$$
Let x be a point of B^n nearer to a than to $-a$, and let y (resp. z) be the foot of the hyperbolic (resp. Euclidean) perpendicular from x to L. See Figure 12.2.3.

Let α be the angle of the hyperbolic right triangle $\triangle(0, x, y)$ opposite the side $[x, y]$. Then we have
$$\sinh d(x, y) = \sinh d(0, x) \sin \alpha.$$

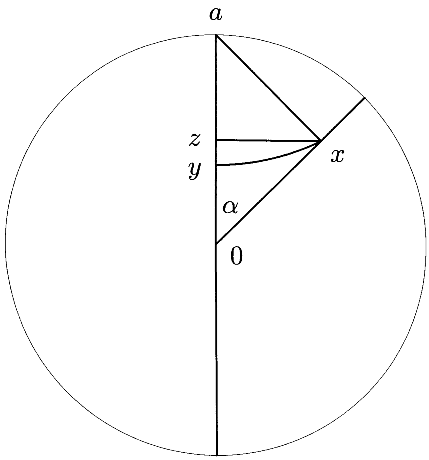


Figure 12.2.3. The points x, y, z in the proof of Theorem 12.2.5

Now as
$$d(0,x) = \log\left(\frac{1+|x|}{1-|x|}\right),$$
we have that
$$\sinh d(x,y) = \frac{2|x|}{1-|x|^2}\frac{|z-x|}{|x|} = \frac{2|z-x|}{1-|x|^2}.$$
Hence, we have
$$\sinh d(x,L) = \frac{2|x-z|}{1-|x|^2}.$$
Now observe that
$$\begin{aligned} |a-x| &\leq |a-z|+|z-x| \\ &\leq |(x/|x|)-z|+|z-x| \\ &\leq |(x/|x|)-x|+2|z-x| \\ &= 1-|x|+2|z-x|. \end{aligned}$$
Hence, we have
$$\begin{aligned} \frac{|a-x|}{1-|x|} &\leq 1+\frac{2|z-x|}{1-|x|} \\ &= 1+(1+|x|)\sinh d(x,L) \\ &\leq 1+2\sinh d(x,L) \\ &= 1+2\sinh d(x,R). \end{aligned}$$
Therefore (1) implies (2).

Assume that (2) holds. Then $\{g_\imath(0)\}$ converges to a. Let x be as above. Then
$$\sinh d(x,R) = 2\frac{|z-x|}{1-|x|^2} < 2\frac{|a-x|}{1-|x|^2} < 2\frac{|a-x|}{1-|x|}.$$
Therefore (2) implies (1). Thus (1) and (2) are equivalent.

By Theorem 4.4.7, the isometric sphere $S(a_\imath, r_\imath)$ of $g_\imath$ is orthogonal to S^{n-1} and $g_\imath = f_\imath\sigma_\imath$, where $\sigma_\imath$ is the reflection in $S(a_\imath, r_\imath)$ and $f_\imath$ is an orthogonal transformation. Observe that
$$a_\imath = \sigma_\imath(\infty) = \sigma_\imath f_\imath^{-1}(\infty) = g_\imath^{-1}(\infty).$$
By Theorem 4.4.2, we have
$$r_\imath^2 + 1 = |a_\imath|^2 = |f_\imath a_\imath|^2 = |f_\imath\sigma_\imath(\infty)|^2 = |g_\imath(\infty)|^2.$$
By Theorem 4.3.7, we find that $g_\imath(\infty)$ and $g_\imath(0)$ are inverse points with respect to S^{n-1}. Therefore
$$|g_\imath(0)|\,|g_\imath(\infty)| = 1.$$
Hence, we have that
$$r_\imath^2 + 1 = |g_\imath(0)|^{-2}.$$

Therefore, we have that

$$r_i^2 = \frac{1 - |g_i(0)|^2}{|g_i(0)|^2}.$$

Hence, we have

$$1 - |g_i(0)| < 1 - |g_i(0)|^2 < r_i^2.$$

Therefore, we have that

$$1 - |g_i(0)| = O(r_i^2).$$

Now let m be the minimum value of $|g_i(0)|$. Then we have

$$\begin{aligned} r_i^2 &\leq \frac{1 - |g_i(0)|^2}{m^2} \\ &= \frac{(1 - |g_i(0)|)(1 + |g_i(0)|)}{m^2} \\ &\leq \frac{2(1 - |g_i(0)|)}{m^2}. \end{aligned}$$

Hence, we have

$$r_i^2 = O(1 - |g_i(0)|).$$

Thus (2) and (3) are equivalent.

Now for each i, we have

$$g_i^{-1} = \sigma_i f_i^{-1} = f_i^{-1}(f_i \sigma_i f^{-1}).$$

Hence, the isometric sphere of g_i^{-1} is $S(f_i a_i, r_i)$ by Theorem 4.3.3. By Theorem 4.1.3, we deduce that for each $x \neq a$ in S^{n-1}, we have

$$|g_i^{-1}x - g_i^{-1}a| = \frac{r_i^2|x - a|}{|x - g_i(\infty)|\,|a - g_i(\infty)|}.$$

Let σ be the inversion in S^{n-1}. Then we have

$$\begin{aligned} |g_i^{-1}x - g_i^{-1}a| &= \frac{r_i^2|x - a|}{|x - g_i\sigma(0)|\,|a - g_i\sigma(0)|} \\ &= \frac{r_i^2|x - a|}{|x - \sigma g_i(0)|\,|a - \sigma g_i(0)|} \\ &= \frac{r_i^2|x - a|}{|\sigma(x) - \sigma g_i(0)|\,|\sigma(a) - \sigma g_i(0)|} \\ &= \frac{r_i^2|x - a|}{(|x - g_i(0)|/|g_i(0)|)(|a - g_i(0)|/|g_i(0)|)} \\ &= \frac{r_i^2|g_i(0)|^2\,|x - a|}{|a - g_i(0)|\,|x - g_i(0)|}. \end{aligned}$$

Now assume that (3) holds. Then there is a constant $k > 0$ such that

$$|a - g_i(0)| \leq k r_i^2.$$

Hence $\{g_i(0)\}$ converges to a, and for each $x \neq a$ in S^{n-1} and all sufficiently large i, we have

$$\begin{aligned} |g_i^{-1}x - g_i^{-1}a| &\geq \frac{|g_i(0)|^2\ |x-a|}{k|x-g_i(0)|} \\ &\geq \frac{3|x-a|}{4k|x-g_i(0)|} \\ &\geq \frac{3|x-a|}{4k(3|x-a|/2)} = 1/2k. \end{aligned}$$

Therefore (3) implies (4).

Now assume that (4) holds, that is, $\{g_i(0)\}$ converges to a and there is an $x \neq a$ in S^{n-1} and a $\delta > 0$ such that for all i, we have

$$|g_i^{-1}x - g_i^{-1}a| \geq \delta.$$

Then we have that

$$\frac{r_i^2|g_i(0)|^2\ |x-a|}{|a-g_i(0)|\ |x-g_i(0)|} \geq \delta.$$

Hence, for all sufficiently large i, we have

$$\begin{aligned} \frac{|a-g_i(0)|}{r_i^2} &\leq \frac{|g_i(0)|^2\ |x-a|}{\delta|x-g_i(0)|} \\ &\leq \frac{|x-a|}{\delta|x-g_i(0)|} \\ &\leq \frac{|x-a|}{\delta(|x-a|/2)} = 2/\delta. \end{aligned}$$

Therefore (4) implies (3). Thus (3) and (4) are equivalent. □

Cusped Limit Points

Let Γ be a discrete subgroup of $\mathrm{M}(U^n)$ such that ∞ is fixed by a parabolic element of Γ. Then the stabilizer Γ_∞ is an elementary group of parabolic type. Therefore Γ_∞ corresponds under Poincaré extension to a discrete subgroup of $\mathrm{I}(E^{n-1})$. By Theorems 5.4.6 and 7.4.2, there is a Γ_∞-invariant m-plane Q of E^{n-1} such that Q/Γ_∞ is compact. Let $r > 0$ and let $N(Q,r)$ be the r-neighborhood of Q in E^n. Then $N(Q,r)$ is invariant under Γ_∞. Now set

$$U(Q,r) = \overline{U}^n - \overline{N}(Q,r).$$

Then $U(Q,r)$ is an open Γ_∞-invariant subset of $\overline{U}^n$. Note that if $m = n-1$, then $U(Q,r)$ is a horoball based at ∞. The set $U(Q,r)$ is said to be a *cusped region* for Γ based at ∞ if and only if for all g in $\Gamma - \Gamma_\infty$, we have

$$U(Q,r) \cap gU(Q,r) = \emptyset.$$

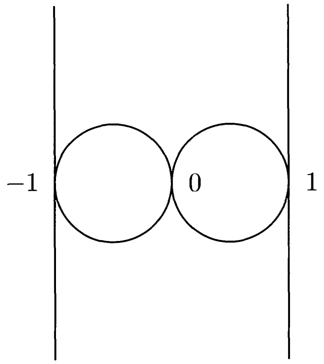


Figure 12.2.4. The four circles in Example 1

Example 1. Let P be the Schottky polyhedron in U^3 with four sides whose boundaries in $\hat{\mathbb{C}}$ are the four circles in Figure 12.2.4. We pair the two vertical sides of P and the two nonvertical sides of P by reflecting in the vertical plane midway between the two vertical sides, and then reflecting in the corresponding side of P. This side-pairing generates a Schottky subgroup Γ of $\mathrm{M}(U^3)$ of rank 2. The group Γ corresponds under Poincaré extension to the group in Example 2 at the end of §9.8.

Observe that the parabolic translation $f(z) = z + 2$ generates Γ_∞ and Γ_∞ leaves invariant the real axis $\mathbb{R}$ of $\mathbb{C}$. Let $r \geq 1/2$ and let $N(\mathbb{R}, r)$ be the r-neighborhood of $\mathbb{R}$ in E^3. Then Γ_∞ leaves $N(\mathbb{R}, r)$ invariant. Hence Γ_∞ leaves invariant the set

$$U(\mathbb{R}, r) = \overline{U}^3 - \overline{N}(\mathbb{R}, r).$$

Now since

$$U(\mathbb{R}, r) \subset \cup\{f^k(\overline{P}) : k \in \mathbb{Z}\},$$

we deduce that for all g in $\Gamma - \Gamma_\infty$, we have

$$U(\mathbb{R}, r) \cap gU(\mathbb{R}, r) = \emptyset.$$

Thus $U(\mathbb{R}, r)$ is a cusped region for Γ.

Let c be a point of $\hat{E}^{n-1}$ fixed by a parabolic element of a discrete subgroup Γ of $\mathrm{M}(U^n)$. A subset U of $\overline{U}^n$ is a *cusped region* for Γ based at c if and only if and only if upon conjugating Γ so that $c = \infty$, the set U transforms to a cusped region for Γ based at ∞.

Lemma 1. *If U is a cusped region based at c for a discrete subgroup Γ of $\mathrm{M}(U^n)$, then $U \subset U^n \cup O(\Gamma)$.*

Proof: On the contrary, suppose that there is a limit point a of Γ in U. Then there is a point x of U^n and a sequence $\{g_i\}_{i=1}^\infty$ of elements of Γ such that $\{g_i x\}$ converges to a. As U is an open neighborhood of a in $\overline{U}^n$, there is an integer j such that $g_i x$ is in U for all $i \geq j$. Since $U \cap gU = \emptyset$ for all g in $\Gamma - \Gamma_c$ and

$$g_i x = (g_i g_j^{-1}) g_j x,$$

we conclude that $g_i g_j^{-1}$ is in Γ_c for all $i \geq j$. Hence, there is an element f_i of Γ_c such that $g_i = f_i g_j$ for all $i \geq j$. Let $y = g_j x$. Then $\{f_i y\}_{i=j}^{\infty}$ converges to a. Hence a is a limit point of Γ_c. Therefore $a = c$. But c is not in U, and so we have a contradiction. □

Definition: A *cusped limit point* of a discrete subgroup Γ of $\mathrm{M}(B^n)$ is a fixed point c of a parabolic element of Γ such that after passing to the upper half-space model U^n, there is a cusped region U for Γ based at c.

Theorem 12.2.6. *Let c be a cusped limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$ and let P be a convex fundamental polyhedron for Γ. Then there is an element g of Γ such that c is in $g\overline{P}$.*

Proof: We pass to the upper half-space model U^n and conjugate Γ so that $c = \infty$. Then Γ has a cusped region $U(Q, r)$. By Lemma 1, we have that

$$U(Q, r) \subset U^n \cup O(\Gamma).$$

Hence, we have that

$$L(\Gamma) \subset \overline{N}(Q, r).$$

Therefore, by increasing r, if necessary, we may assume that

$$\overline{U}(Q, r) \subset U^n \cup O(\Gamma) \cup \{\infty\}.$$

We now prove that P meets only finitely many members of

$$\{gU(Q, r) : g \in \Gamma\}.$$

Set

$$C(Q, r) = \overline{U}(Q, r) - \big(U(Q, r+1) \cup \{\infty\}\big).$$

Then $C(Q, r)$ is a closed subset of E^n. Now as

$$U(Q, r) \cap gU(Q, r) = \emptyset \quad \text{for all } g \text{ in } \Gamma - \Gamma_\infty,$$

we have

$$\overline{U}(Q, r) \cap gU(Q, r+1) = \emptyset \quad \text{for all } g \text{ in } \Gamma - \Gamma_\infty.$$

Therefore, we have

$$C(Q, r) \subset \overline{U}^n - \bigcup_{g \in \Gamma} gU(Q, r+1).$$

Hence $U(Q, r+1)$ does not contain gP for any g in Γ. Therefore, if gP meets $U(Q, r)$, then gP meets $C(Q, r)$, since gP is connected. Thus, if P meets $g^{-1}U(Q, r)$, then gP meets $C(Q, r)$.

Let D be a Dirichlet polyhedron for Γ_∞ in Q. Then D is compact, since Q/Γ_∞ is compact. Let $\rho : E^n \to Q$ be the orthogonal projection. Then $\rho^{-1}(D)$ is closed in E^n, since ρ is continuous. Hence

$$K = C(Q,r) \cap \rho^{-1}(D)$$

is a closed subset of E^n; moreover K is bounded, since

$$K \subset \overline{N}(D, r+1).$$

Therefore K is compact. Furthermore

$$C(Q,r) = \cup\{fK : f \in \Gamma_\infty\}.$$

By Theorem 12.1.12, we have that $\{gP : g \in \Gamma\}$ is a locally finite family of subsets of $U^n \cup O(\Gamma)$. As K is a compact subset of $U^n \cup O(\Gamma)$, we deduce that K meets only finitely many Γ-images of P, say $g_1P, \ldots, g_kP$. Now suppose that P meets $g^{-1}U(Q,r)$. Then gP meets $C(Q,r)$. Hence, there is an f in Γ_∞ such that gP meets fK, and so $f^{-1}gP$ meets K. Therefore $f^{-1}g = g_i$ for some i, and so $g = fg_i$. Hence, we have

$$g^{-1}U(Q,r) = g_i^{-1}U(Q,r).$$

Thus P meets only

$$g_1^{-1}U(Q,r), \ldots, g_k^{-1}U(Q,r).$$

Now let $\{x_i\}_{i=1}^\infty$ be a sequence of points of $U(Q,r)$ such that the nth coordinate of x_i goes to infinity as $i \to \infty$. Then for each i, there is an h_i in Γ such that h_ix_i is in P. Then P meets $h_iU(Q,r)$. Hence, there is a j such that $h_iU(Q,r) = h_jU(Q,r)$ for infinitely many $i \geq j$. For all such i, we have that

$$h_j^{-1}h_iU(Q,r) = U(Q,r).$$

Hence, there is an f_i in Γ_∞ such that $h_j^{-1}h_i = f_i$. Therefore $h_i = h_jf_i$. Let $y_i = f_ix_i$. Then $h_jy_i = h_ix_i$, and so h_jy_i is in P. Therefore y_i is in $h_j^{-1}P$. As the nth coordinate of x_i goes to infinity as $i \to \infty$, we have that $f_ix_i \to \infty$. Therefore $y_i \to \infty$. Thus c is in $h_j^{-1}\overline{P}$. □

The next corollary follows immediately from Theorems 12.2.4 and 12.2.6.

Corollary 2. *A cusped limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$ is not a conical limit point of Γ.*

Lemma 2. *Let Γ be a discrete subgroups of $\mathrm{M}(U^n)$ such that ∞ is fixed by a parabolic element of Γ, let Q be a Γ_∞-invariant m-plane of E^{n-1} such that Q/Γ_∞ is compact, let P be a convex fundamental polyhedron for Γ, and let $\{x_i\}_{i=1}^\infty$ be a sequence of points of P converging to ∞. Then*

$$\lim_{i\to\infty} \mathrm{dist}_E(x_i, Q) = \infty.$$

Proof: By Theorem 5.4.5, the group Γ_∞ has a torsion-free subgroup of H of finite index. Then Q/H is compact by Lemma 1 of §7.4. Let D be a Dirichlet polyhedron for H. Then D is compact. Let $r > 0$ and let $M(D,r)$ be the r-neighborhood of D in E^{n-1}. Then $\overline{M}(D,r)$ is compact. Let $M(Q,r)$ be the r-neighborhood of Q in E^{n-1}. Then $M(Q,r)$ is convex. As $\overline{M}(D,r)$ projects onto $\overline{M}(Q,r)/\mathrm{H}$, we find that $\overline{M}(Q,r)/\mathrm{H}$ is compact. Hence $M(Q,r)/\mathrm{H}$ has finite volume in the space-form E^{n-1}/H.

Now since $x_i \to \infty$ in P, we have that ∞ is $\overline{P}$. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then $\nu(P^\circ)$ is an open convex subset of E^{n-1}. Hence $\nu(P^\circ) \cap M(Q,r)$ is an open convex subset of E^{n-1}. Now since $\nu(P^\circ) \cap M(Q,r)$ injects $M(Q,r)/\mathrm{H}$, we deduce that $\nu(P^\circ) \cap M(Q,r)$ has finite volume in E^{n-1}. Therefore $\nu(P^\circ) \cap M(Q,r)$ is bounded. Hence $\nu(P) \cap \overline{M}(Q,r)$ is compact.

We now show that

$$\lim_{i\to\infty} \mathrm{dist}_E(x_i, Q) = \infty.$$

Suppose that this is not the case. Then there is an $r > 0$ such that

$$\mathrm{dist}_E(x_i, Q) \leq r$$

for infinitely many i. Hence $\nu(x_i)$ is in the bounded subset $\nu(P) \cap \overline{M}(Q,r)$ of E^{n-1} for infinitely many i. As $x_i \to \infty$, we can conclude that there is an i such that the nth coordinate of x_i is greater than r, which is a contradiction. □

Definition: A *polyhedral wedge* in E^n is a convex polyhedron P in E^n such that the intersection of all its sides is nonempty.

Note that the intersection of all the sides of a polyhedral wedge in E^n is an m-plane of E^n. Also a polyhedral wedge in E^n has only finitely many sides, since the collection of its sides is locally finite. Figure 12.2.5 illustrates a polyhedral wedge in E^2.

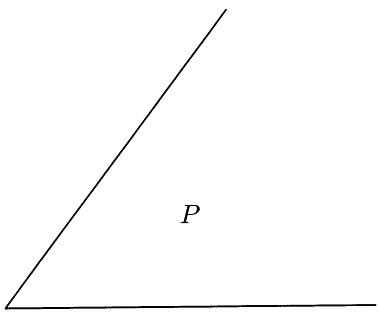


Figure 12.2.5. A polyhedral wedge P in E^2

Lemma 3. *Let P be an n-dimensional polyhedral wedge in E^n. Then there is an integer ℓ such that if $P_1, \ldots, P_k$ are polyhedra in E^n that are congruent to P, with mutually disjoint interiors, then $k \leq \ell$.*

Proof: Let z be the point in the intersection of all the sides of P nearest to the origin. The *normalized solid angle subtended by P* is defined to be

$$\omega(P) = \frac{\mathrm{Vol}(P \cap B(z,1))}{\mathrm{Vol}(B(z,1))}.$$

Given $r > 0$, let μ_r be the similarity of E^n defined by

$$\mu_r(x) = x/r$$

and let τ_r be the translation of E^n defined by

$$\tau_r(x) = x - z + z/r.$$

Then we have that

$$\mu_r(P) = \tau_r(P).$$

Observe that

$$\begin{aligned}
\lim_{r\to\infty} \frac{\mathrm{Vol}(P \cap B(0,r))}{\mathrm{Vol}(B(0,r))} &= \lim_{r\to\infty} \frac{\mathrm{Vol}(\mu_r(P) \cap B(0,1))}{\mathrm{Vol}(B(0,1))} \\
&= \lim_{r\to\infty} \frac{\mathrm{Vol}(P \cap B(\tau_r^{-1}(0),1))}{\mathrm{Vol}(B(\tau_r^{-1}(0),1))} \\
&= \lim_{r\to\infty} \frac{\mathrm{Vol}(P \cap B(z - z/r,1))}{\mathrm{Vol}(B(z - z/r,1))} \\
&= \frac{\mathrm{Vol}(P \cap B(z,1))}{\mathrm{Vol}(B(z,1))} \\
&= \omega(P).
\end{aligned}$$

Now let ℓ be the greatest integer less than or equal to $1/\omega(P)$. Suppose there are $\ell + 1$ polyhedra $P_0, \ldots, P_\ell$ in E^n that are congruent to P whose interiors are mutually disjoint. We shall derive a contradiction. First of all, $\omega(P_i) = \omega(P)$ for each i. Choose r sufficiently large so that for each i, we have

$$\left| \frac{\mathrm{Vol}(P_i \cap B(0,r))}{\mathrm{Vol}(B(0,r))} - \omega(P) \right| < \omega(P) - \frac{1}{\ell+1}.$$

Then for each i, we have

$$\mathrm{Vol}(P_i \cap B(0,r)) > \mathrm{Vol}(B(0,r))/(\ell+1).$$

Hence

$$\mathrm{Vol}\Big(\bigcup_{i=0}^{\ell} P_i \cap B(0,r) \Big) = \sum_{i=0}^{\ell} \mathrm{Vol}(P_i \cap B(0,r)) > \mathrm{Vol}(B(0,r)),$$

which is a contradiction. Thus ℓ is the desired upper bound. $\square$

Cusp Points

Let P be a convex polyhedron in B^n. A *cusp point* of P is a point c of $\overline{P} \cap S^{n-1}$ for which there is an open neighborhood N of c in E^n such that the intersection of the Euclidean closures of all the sides of P that meet N is c. If c is a cusp point of P, then the *cusp* of P incident with c is the union of all the sides of P incident with c. For example, the two vertical sides of the polyhedron P in Example 1 form a cusp of P with ∞ its cusp point. Likewise, the points $-1, 0$, and 1 are cusp points of P.

Suppose that c is a cusp point of P. Then there is a horosphere Σ based at c such that Σ meets only the sides of P incident with c. By Theorem 6.3.23, the set

$$L(c) = \Sigma \cap P$$

is a Euclidean convex polyhedron called the *link* of c in P. Note that the orientation preserving similarity class of $L(c)$ does not depend on the choice of Σ. If we pass to the upper half-space model U^n and conjugate Γ so that $c = \infty$, then there is a canonical way of representing $L(c)$. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then $L(c)$ is directly similar to νP. For example, the projection νP of the polyhedron P in Example 1 is the polygon in $\mathbb{C}$ whose two sides are the vertical straight lines in Figure 12.2.4.

An *ideal vertex* of a polyhedron P in B^n is a cusp point c of P such that $L(c)$ is compact. If P is 2-dimensional, then every cusp point of P is an ideal vertex. The cusp points of the 3-dimensional polyhedron P in Example 1 are not ideal vertices of P. If P is n-dimensional and has finite volume in B^n, then every cusp point of P is an ideal vertex of P.

Theorem 12.2.7. *Let c be a cusped limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$ and let P be a convex fundamental polyhedron for Γ such that c is in $\overline{P}$. Then c is a cusp point of P.*

Proof: First we show that there is an $r > 0$ such that $B(c, r)$ meets only the sides of P incident with c. Suppose that this is not the case. Then for each positive integer i, the ball $B(c, 1/i)$ meets a side S_i of P such that c is not in $\overline{S}_i$. Since $B(c, 1/i)$ is open, it contains a point x_i of S_i°. Then the sequence $\{x_i\}_{i=1}^\infty$ converges to c. By Lemma 1 of §6.6, there is an element $g_i \neq 1$ of Γ such that x_i is in $P \cap g_i P$ for each i. We now pass to the projective disk model D^n. By Theorem 6.3.20, we have

$$\begin{aligned} \overline{P} \cap g_i \overline{P} &\subset (P \cap g_i P) \cup (\overline{P} \cap S^{n-1}) \\ &\subset \partial P \cup (\overline{P} \cap S^{n-1}) = \partial \overline{P}. \end{aligned}$$

Hence $\overline{P} \cap g_i \overline{P}$ is a convex subset of $\partial \overline{P}$. By Theorem 6.2.6, the set $\overline{P} \cap g_i \overline{P}$ is contained in a side of the convex set $\overline{P}$. By Theorem 6.3.20, the sides of $\overline{P}$ are the Euclidean closures of the sides of P together with the points of

$\overline{P} \cap S^{n-1}$ that are not in the Euclidean closure of a side of P. Since x_i is in S_i°, we deduce that

$$\overline{P} \cap g_i\overline{P} \subset \overline{S}_i.$$

As c is not in S_i for all i, we have that $g_ic \neq c$ for all i.

We now pass to the upper half-space model U^n and conjugate Γ so that $c = \infty$. Let $U(Q, r)$ be a cusped region for Γ. Then g_ic is in $\overline{N}(Q, r)$ for each i by Lemma 1. Let D be a Dirichlet polyhedron for Γ_∞ in Q. Then D is compact, since Q/Γ_∞ is compact. Hence $\overline{N}(D, r)$ is compact. Now for each i, there is an element f_i of Γ_∞ such that f_ig_ic is in $\overline{N}(D, r)$. By passing to a subsequence, we may assume that $f_ig_ic \to b$ in E^{n-1}. By Lemma 2, we have that

$$\lim_{i\to\infty} \mathrm{dist}_E(x_i, Q) = \infty.$$

Hence, we have that

$$\lim_{i\to\infty} \mathrm{dist}_E(f_ix_i, Q) = \infty.$$

Therefore $f_ix_i \to c$.

We now show that infinitely many of the terms of $\{f_ig_i\}_{i=1}^\infty$ are distinct. Suppose that this is not the case. Then by passing to a subsequence, we may assume that there is an element h of Γ such that $f_ig_i = h$ for all i. As x_i is in g_iP, we have that f_ix_i is in hP for all i. As $f_ix_i \to c$, we find that c is in $h\overline{P}$. Hence $c = f_i^{-1}c$ is in $f_i^{-1}h\overline{P} = g_i\overline{P}$. Then c is in $\overline{P} \cap g_i\overline{P}$ and so c is in $\overline{S}_i$, which is a contradiction. Thus, infinitely many of the terms of $\{f_ig_i\}$ are distinct.

Let R_i be the ray in f_ig_iP joining f_ix_i to f_ig_ic. Then the sequence of rays $\{R_i\}$ converges to the line (b, c). Let x be any point of (b, c). Then $B(x, 1)$ meets all but finitely many of the rays $\{R_i\}$. Hence, the compact set $C(x, 1)$ meets all but finitely many terms of $\{f_ig_iP\}$ contrary to the local finiteness of P. We pass back to the conformal ball model B^n. Then we conclude that there is an $r > 0$ such that $B(c, r)$ meets only the sides of P incident with c.

Now since every point of the interior of $\overline{P} \cap S^{n-1}$ is an ordinary point of Γ, the limit point c is in $\partial\overline{P}$. Hence $B(c, r)$ meets at least one side of P incident with c. Let Σ be a horosphere based at c and contained in $B(c, r)$. We pass to the upper half-space model U^n with $c = \infty$. By Theorem 6.3.22, we have that Σ meets only the vertical sides of P. Hence $P \cap \Sigma$ is a Euclidean, $(n-1)$-dimensional, convex, polyhedron in Σ with at least one side by Theorem 6.3.23. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then νP is a Euclidean, $(n-1)$-dimensional, convex, polyhedron in E^{n-1} directly similar to $P \cap \Sigma$.

We now show that c is a cusp point of P. Suppose that this is not the case. Then the intersection of all the vertical sides of P is nonempty. Hence νP is a polyhedral wedge in E^{n-1}. Let f be a parabolic element of Γ_∞. Then f has infinite order. As the polyhedra $\{f^kP\}_{k=1}^\infty$ have mutually disjoint interiors in U^n, the polyhedra $\{\nu f^k P\}_{k=1}^{\infty}$ have mutually disjoint interiors in E^{n-1}. As $\nu f^k P = f^k \nu P$ for each k, the polyhedron $\nu f^k P$ is congruent to νP for each k. But this contradicts Lemma 3. Thus c is a cusp point of the polyhedron P. □

We next consider an example of a cusp point of a fundamental polygon that is not a limit point.

Example 2. Consider the Schottky polygon P in B^2 in Figure 12.2.6. The polygon P is invariant under the antipodal map of B^2. We pair the opposite sides of P by hyperbolic translations g, h along the diameters of B^2 joining the opposite sides of P. This side-pairing generates a Schottky group Γ of rank two. The polygon P obviously contains the Dirichlet polygon D for Γ centered at 0. Hence $P = D$, since P° is a Γ-packing.

The cusp point v of P is an ordinary point of Γ, since the open circular arcs (gu, v) and (v, hw) are subsets of $O(\Gamma)$ and limit points are not isolated. Thus v is not a limit point of Γ. The same argument also shows that $-v$ is not a limit point of Γ.

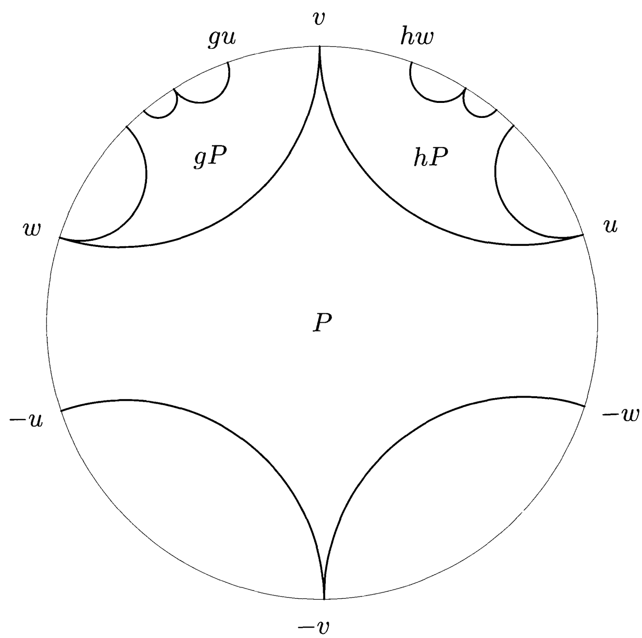


Figure 12.2.6. The polygon P and two of its translates

Exercise 12.2

1. Let a be a conical limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$. Prove that ga is a conical limit point of Γ for each g in Γ.
2. Let a be a limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$. Prove that a is a conical limit point of Γ if and only if there is a sequence $\{g_i\}_{i=1}^\infty$ of elements of Γ such that $\{g_i(0)\}_{i=1}^\infty$ converges to a within a Euclidean hypercone C whose vertex is a and whose axis passes through 0.
3. Let Γ be a nonelementary discrete subgroup of $\mathrm{M}(B^n)$. Prove that for each point a of S^{n-1}, there is a point $x \neq a$ of S^{n-1}, a positive real number δ, and a sequence $\{g_i\}_{i=1}^\infty$ of distinct elements of Γ such that $|g_i x - g_i a| \geq \delta$ for all i. Conclude that the hypothesis that $\{g_i(0)\}_{i=1}^\infty$ converges to a cannot be dropped from Theorem 12.2.5(4).
4. Let c be a cusped limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$. Prove that gc is a cusped limit point of Γ for each g in Γ.
5. Prove directly that a cusped limit point of Γ is not a conical limit point.
6. Let P be a polyhedral wedge in E^n. Prove that the intersection of all the sides of P is an m-plane of E^n.
7. Let P be a polyhedral wedge in E^n with at least two sides. Prove that every side of P is a polyhedral wedge.
8. Let P be a finite-sided, convex, fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(U^3)$ such that ∞ is in $\overline{P}$ and ∞ is fixed by a parabolic element of Γ. Let $\nu : U^3 \to E^2$ be the vertical projection. Prove that any two unbounded sides of νP are parallel.
9. Let P be an n-dimensional convex polyhedron in B^n of finite volume. Prove that every cusp point of P is an ideal vertex.

§12.3. Geometrically Finite Discrete Groups

In this section, we characterize the discrete subgroups of $\mathrm{M}(B^n)$ that have the property that every limit point is either conical or cusped in terms of the geometry of their convex fundamental polyhedra.

Geometrically Finite Convex Polyhedra

Definition: A convex polyhedron P in B^n is *geometrically finite* if and only if for each point x of $\overline{P} \cap S^{n-1}$ there is an open neighborhood N of x in E^n that meets only the sides of P incident with x.

Example 1. Every finite-sided convex polyhedron in B^n is geometrically finite.

Example 2. Let Q be a convex polyhedron in E^{n-1} with infinitely many sides and let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then the vertical prism $P = \nu^{-1}(Q)$ is a convex polyhedron in U^n with an infinite set of sides

$$\{\nu^{-1}(S) : S \text{ is a side of } Q\}.$$

The polyhedron P is geometrically finite in $\hat{E}^n$, since the set of sides of P is locally finite in E^n and every side of P is incident with ∞.

Theorem 12.3.1. *Let P be a geometrically finite convex polyhedron in B^n. Then*

(1) *if x is in $\overline{\partial P} \cap S^{n-1}$, then there is a side of P incident with x;*

(2) *if x is in $\overline{\partial P} \cap S^{n-1}$ and infinitely many sides of P are incident with x, then x is a cusp point of P;*

(3) *the polyhedron P has only finitely many cusp points;*

(4) *all but finitely many of the sides of P are incident with a cusp point of P.*

Proof: (1) Since P is geometrically finite, there is an $r > 0$ such that $B(x,r)$ meets only the sides of P incident with x. As x is in $\overline{\partial P}$, the ball $B(x,r)$ meets a side of P, which is therefore incident with c.

(2) Suppose that the set $\mathcal{S}(x)$ of all sides of P incident with x is infinite. Then the intersection of all the sides in $\mathcal{S}(x)$ is empty, since $\mathcal{S}(x)$ is locally finite. Therefore x is a cusp point of P.

(3) As $\overline{P} \cap S^{n-1}$ is compact, there are points $x_1, \ldots, x_m$ of $\overline{P} \cap S^{n-1}$ and radii $r_1, \ldots, r_m$ such that $B(x_i, r_i)$ meets only the sides of P incident with x_i for each i and

$$\overline{P} \cap S^{n-1} \subset \bigcup_{i=1}^{m} B(x_i, r_i).$$

Suppose that $B(x_i, r_i)$ contains a cusp point c of P. Then all the sides of P incident with c are incident with x_i. As the intersection of the Euclidean closures of all the sides of P incident with c is c, we have that $c = x_i$. Hence, all the cusp points of P are in the set $\{x_1, \ldots, x_m\}$.

(4) As $P - \bigcup_{i=1}^{m} B(x_i, r_i)$ is compact, all but finitely many sides of P meet $\bigcup_{i=1}^{m} B(x_i, r_i)$. Reindex so that $x_1, \ldots, x_k$ are all the cusp points of P. Then the ball $B(x_i, r_i)$ meets only finitely many sides of P for each $i = k+1, \ldots, m$ by (2). Hence, all but finitely many sides of P meet $\bigcup_{i=1}^{k} B(x_i, r_i)$. Thus, all but finitely many sides of P are incident with a cusp point of P. □

Remark: It follows from Theorem 12.3.1 that a geometrically finite convex polyhedron in B^n is finite-sided if and only if all its cusps are finite-sided. Thus, every geometrically finite convex polygon in B^2 is finite-sided.

Lemma 1. *Let P be a convex polyhedron in E^n. Then ∂P is disconnected if and only if ∂P is the union of two parallel hyperplanes of $\langle P \rangle$.*

Proof: Without loss of generality, we may assume that $\langle P \rangle = E^n$. Choose a point a of P° and $r > 0$ so that $C(a,r) \subset P$. Define a function

$$\rho : \partial P \to S(a,r)$$

by letting $\rho(x)$ be the intersection of the line segment $[a,x]$ with the sphere $S(a,r)$. Then we have

$$\rho(x) = a + \frac{r(x-a)}{|x-a|}.$$

Hence ρ is a continuous injection. Moreover ρ maps ∂P homeomorphically onto $\rho(\partial P)$, since ρ maps S homeomorphically onto $\rho(S)$ for each side S of P and the set of sides of P is locally finite. Therefore ∂P is disconnected if and only if $\rho(\partial P)$ is disconnected.

Let S be a side of P. Then for each point x of $\langle S \rangle$, the line segment $[a,x]$ intersects both ∂P and $S(a,r)$. Consequently $\rho(\partial P)$ contains the open hemisphere of $S(a,r)$ nearest to S whose boundary is parallel to S. As ∂P is the union of the sides of P, we deduce that $\rho(\partial P)$ is a union of open hemispheres of $S(a,r)$ whose boundaries are parallel to the sides of P. Consequently $\rho(\partial P)$ is disconnected if and only if $\rho(\partial P)$ is the union of two antipodal open hemispheres of $S(a,r)$. Therefore ∂P is disconnected if and only if P has exactly two parallel sides. Now P has exactly two parallel sides if and only if each side of P is a hyperplane of E^n by Theorem 6.3.6. Thus ∂P is disconnected if and only if ∂P is the union of two parallel hyperplanes of E^n. □

Lemma 2. *Let E and E' be two k-faces of a convex polyhedron P in E^n. Then there is a sequence $F_1, \ldots, F_\ell$ of $(k+1)$-faces of P such that E is a side of F_1, and E' is a side of F_ℓ, and F_i and F_{i+1} meet along a common side for each $i = 1, \ldots, \ell - 1$.*

Proof: Let $m = \dim P$. The proof is by induction on $m - k$. This is clear if $k = m - 1$, so assume that $k < m - 1$ and the theorem is true for $(k+1)$-faces of P. Let F and F' be $(k+1)$-faces of P such that E is a side of F and E' is a side of F'. If $F = F'$, then we are done, so assume that $F \neq F'$. Then by the induction hypothesis, there is a sequence $G_1, \ldots, G_\ell$ of $(k+2)$-faces of P such that F is a side of G_1, and F' is a side of G_ℓ, and G_i and G_{i+1} meet along a common side F_i for each $i < \ell$. Let $F_0 = F$ and $F_\ell = F'$. We may assume that ℓ is as small as possible. Then $F_i \neq F_{i+1}$ for each i. Since F has at least one side E, we have that ∂G_1 is connected by Lemma 1. Hence, there is a sequence $F_{11}, \ldots, F_{1\ell_1}$ of sides of G_1 such that $F_0 = F_{11}$, $F_{1\ell_1} = F_1$, and F_{1j} and F_{1j+1} meet along a common side for each $j < \ell_1$. By induction, there is a sequence $F_{i1}, \ldots, F_{i\ell_i}$ of sides of G_i such that $F_{i-1} = F_{i1}$, $F_{i\ell_i} = F_i$, and F_{ij} and F_{ij+1} meet along a common side for each i and $j < \ell_i$. □

Lemma 3. *Let P be a convex polyhedron in E^n. If some k-face of P is a k-plane of E^n, then every k-face of P is a k-plane of E^n.*

Proof: Let E and E' be two k-faces of P and suppose that E is a k-plane of E^n. By Lemma 2, there is a sequence $F_1, \ldots, F_\ell$ of $(k+1)$-faces of P such that E is a side of F_1, and E' is a side of F_ℓ, and F_i and F_{i+1} meet along a common side E_i for $i = 1, \ldots, \ell - 1$. We may assume that ℓ is as small as possible. Let $E_0 = E$ and $E_\ell = E'$. Then $E_i \neq E_{i+1}$ for each $i = 0, \ldots, \ell-1$. As E is both open and closed in ∂F, and $E \neq E_1$, we deduce that ∂F_1 is disconnected. Therefore E_1 is a k-plane of E^n by Lemma 1. By induction, we conclude that E_i is a k-plane for each $i = 1, \ldots, \ell$. Thus E' is a k-plane of E^n. □

Lemma 4. *If P is a convex polyhedron in E^n such that all but finitely many sides of P are polyhedral wedges, then P is finite-sided.*

Proof: Let $m = \dim P$. The proof is by induction on m. This is certainly true if $m = 0$, so assume that $m > 0$ and the theorem is true for all polyhedra in E^n of dimension $m - 1$. On the contrary, suppose that P has infinitely many sides. Then P has a side S that is a polyhedral wedge. Now the intersection of all the sides of S is a k-face of S that is a k-plane of E^n. Hence, every k-face of P is a k-plane by Lemma 3. Now every k-face of P is a face of only finitely many sides of P by Theorem 6.3.14, and by Lemma 3, every side of P has a k-face. Therefore, there are infinitely many k-faces of P.

Assume now that $k = m-2$. Then every side of P has either one side or two disjoint sides. Therefore P has at most two sides that are polyhedral wedges, which is a contradiction. Therefore, we may assume that $k < m-2$. Then every side of P has at least two sides by Lemma 3.

Let T be a side of P that is not a polyhedral wedge. Then all but finitely many of the sides of T are a side of a polyhedral wedge side of P. As every side of a polyhedral wedge, with at least two sides, is a polyhedral wedge, we have that all but finitely many of the sides of T are polyhedral wedges. By the induction hypothesis, T is finite-sided. Hence T has only finitely many k-faces by Theorem 6.3.14.

Now since all but finitely many of the sides of P are polyhedral wedges, and there are infinitely many k-faces of P, and each side of P has only finitely many k-faces, there is a k-face E of P such that all the sides of P containing E, say $S_1, \ldots, S_\ell$, are polyhedral wedges. As no other side of P meets S_i for each $i = 1, \ldots, \ell$, we find that $\bigcup_{i=1}^{\ell} S_i$ is both open and closed in ∂P. Hence ∂P is the union of the sides $S_1, \ldots, S_\ell$ by Lemma 1. But this contradicts the assumption that P has infinitely many sides. Thus P is finite-sided. □

Theorem 12.3.2. *Let c be the cusp point of an infinite-sided cusp of a geometrically finite, exact, convex, fundamental polyhedron P for a discrete subgroup Γ of $\mathrm{M}(B^n)$. Then c is fixed by a parabolic element of Γ.*

Proof: First, we prove that all but finitely many of the sides of P incident with c meet only the sides of P incident with c. On the contrary, suppose that $\{S_i\}_{i=1}^{\infty}$ is a sequence of distinct sides of P such that c is in $\overline{S}_i$ and S_i meets a side T_i of P such that c is not in $\overline{T}_i$ for all i. Let $g_i = g_{S_i}$ for each i. As $P \cap g_i P = S_i$, we find that c is in $g_i\overline{P}$ for each i. Now the terms of the sequence $\{g_i\}_{i=1}^{\infty}$ are distinct. Hence, the Euclidean diameter of $g_i\overline{P}$ goes to zero as $i \to \infty$. Now as c is a cusp point of P, there is an $r > 0$ such that $B(c,r)$ meets only the sides of P incident with c. Hence, there is a j such that

$$g_j P \subset B(c,r).$$

As $S_j \subset g_j P$, we find that $B(c,r)$ meets T_j, which is a contradiction. Thus, all but finitely many of the sides of P incident with c meet only the sides of P incident with c.

We say that a side S of P is *cusped* if ∂S is a cusp of S. We next prove that infinitely many of the sides of P incident with c are cusped and have c as their cusp point. We now pass to the upper half-space model U^n and conjugate Γ so that $c = \infty$. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then νP is an infinite-sided polyhedron in E^{n-1} whose sides are the vertical projections of the vertical sides of P. Now a vertical side S of P is cusped if and only if S meets only vertical sides of P and νS is not a polyhedral wedge. Moreover, all but finitely many of the vertical sides of P meet only vertical sides of P, and by Lemma 4, infinitely many of the sides of νP are not polyhedral wedges. Hence, infinitely many of the sides of P incident with c are cusped and have c as their cusp point.

Let S be a cusped side of P. Then S is paired to another cusped side S' of P by $g_{S'}$ and the unique cusp point of S is paired to the unique cusp point of S'. By Theorem 12.3.1, the polyhedron P has only finitely many cusp points and all but finitely many of the sides of P are incident with a cusp point of P. Consequently, there is a sequence $\{S_i\}_{i=1}^{\infty}$ of distinct cusped sides of P incident with c such that c is the cusp point of S_i for all i, and S_i' is incident with a cusp point c' of P for all i. Now since all but finitely many of the sides of P incident with c' meet only the sides of P incident with c', we may assume that S_i' meets only the sides of P incident with c' for each i. Then c' is the cusp point of S_i' for each i.

Let $h_i = g_{S_i'}$ for each i. Then the terms of the sequence $\{h_i\}_{i=1}^{\infty}$ are distinct. Moreover $h_i c = c'$ for each i. Hence $h_i c = h_1 c$ for all i. Therefore $h_i^{-1} h_1 c = c$ for all i. Hence, the stabilizer Γ_c is infinite. Therefore Γ_c is an infinite elementary group. By Theorem 12.2.4, the point c is not fixed by a hyperbolic element of Γ. Therefore Γ_c is of parabolic type. Hence c is fixed by a parabolic element of Γ. □

Let P be an exact, convex, fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(B^n)$ and let Φ be the Γ-side-pairing of P. Two points x, x' of $\overline{P}$ are said to be *paired* by Φ, written $x \simeq x'$, if and only if there is a side S of P such that x is in $\overline{S}$, and x' is in $\overline{S'}$, and $g_S(x') = x$. If $g_S(x') = x$, then $g_{S'}(x) = x'$. Therefore $x \simeq x'$ if and only if $x' \simeq x$. Two points x, y of $\overline{P}$ are said to be *related* by Φ, written $x \sim y$, if and only if either $x = y$ or there is a finite sequence $x_1, \ldots, x_m$ of points of P such that

$$x = x_1 \simeq x_2 \simeq \cdots \simeq x_m = y.$$

Being related by Φ is obviously an equivalence relation on the set $\overline{P}$. The equivalence classes of $\overline{P}$ are called *cycles.* If x is in $\overline{P}$, we denote the cycle containing x by $[x]$.

Theorem 12.3.3. *Let P be a geometrically finite, exact, convex, fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(B^n)$. Then for each point x of $\overline{P}$, we have that*

(1) *the cycle $[x]$ is finite;*

(2) $[x] = \overline{P} \cap \Gamma x$.

Proof: (1) By Theorem 6.7.5, we may assume that x is in $\overline{P} \cap S^{n-1}$. If x is in the interior of $\overline{P} \cap S^{n-1}$, then $[x] = \{x\}$. Hence, we may assume that x is in $\overline{\partial P} \cap S^{n-1}$.

Assume first that x is fixed by a parabolic element of Γ. Then by the same argument as at the end of the proof of Theorem 12.2.7, we deduce that x is a cusp point of P. As $[x] \subset \Gamma x$, every point of $[x]$ is fixed by a parabolic element of Γ. Hence, every point of $[x]$ is a cusp point of P. By Theorem 12.3.1, the polyhedron P has only finitely many cusp points. Thus $[x]$ is finite.

Assume now that x is not fixed by a parabolic element of Γ. By Theorem 12.3.1, there is a side S of P such that x is in $\overline{S}$. By Theorems 12.3.1 and 12.3.2, only finitely many sides of P are incident with x. Let k be the smallest dimension such that there is a k-face E of P such that x is in $\overline{E}$. Then for each side S of P incident with x, there is a k-face E of S such that that E is incident with x by Lemma 3 applied to the link of x in P. Now by Theorem 6.3.14, every k-face of P incident with x is an intersection of sides of P that are incident with x. Hence, there are only finitely many k-faces of P incident with x, say $E_1, \ldots, E_\ell$.

Assume first that $\ell = 1$. Then E_1 is the intersection of all the sides of P incident with x. Hence x is not a cusp point of P. Assume now that $\ell > 1$. Then the intersection of all the sides of P incident with x is empty, since $E_1 \cap E_2 = \emptyset$ by the minimality of k. Hence x is a cusp point of P. Thus $\ell > 1$ if and only if x is a cusp point of P. As x is not fixed by a parabolic element of Γ, no point of $[x]$ is fixed by a parabolic element of Γ. Therefore, each cusp point in $[x]$ is finite-sided by Theorem 12.3.2.

We say that x is *directly related* to a point y of $\overline{P}$ if there is an element g of Γ such that $y = gx$ and there are k-faces E and F of P such that x is in $\overline{E}$, y is in $\overline{F}$, and $F = gE$. As P is locally finite, there are only finitely many g in Γ such that

$$E_i \subset P \cap g^{-1}P$$

for each $i = 1, \ldots, \ell$. Hence x is directly related to only finitely many points of $\overline{P}$.

Now assume that $x \sim y$. Then there is a finite sequence $x_1, \ldots, x_m$ of points of $\overline{P}$ such that

$$x = x_1 \simeq x_2 \simeq \cdots \simeq x_m = y.$$

By induction on m, the integer k is the smallest dimension such that there is a k-face F of P such that y is in $\overline{F}$. Thus k depends only on $[x]$. If x is directly related to y, then y is one of only finitely many points, so assume that x is not directly related to y. Then $m > 2$ and one of the points $x_2, \ldots, x_{m-1}$ is a cusp point of P. Let j be the largest index such that x_j is a cusp point of P. Then x_j is directly related to y. As P has only finitely many cusp points and since each cusp point of P in $[x]$ is directly related to only finitely many points of $\overline{P}$, we conclude that y is one of only finitely many points. Thus $[x]$ is finite.

(2) By Theorem 6.7.5, we may assume that x is in $\overline{P} \cap S^{n-1}$. It is clear from the definition of $[x]$ that $[x] \subset \overline{P} \cap \Gamma x$. Let y be a point of $\overline{P} \cap \Gamma x$. Then there is an element f of Γ such that $y = fx$, whence x is in $f^{-1}\overline{P}$. We now pass to the upper half-space model U^n and conjugate Γ so that $x = \infty$. Let g be an element of Γ such that x is in $g\overline{P}$. Since gP is geometrically finite, a sufficiently high horizontal horosphere Σ will meet only the vertical sides of gP. Then $gP \cap \Sigma$ is a Euclidean, $(n-1)$-dimensional, convex polyhedron in Σ. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then νgP is a convex polyhedron in E^{n-1} directly similar to $gP \cap \Sigma$. Let

$$\mathcal{T} = \{\nu gP : g \in \Gamma \ \text{ and } \ x \in g\overline{P}\}$$

and let U be the union of all the polyhedra in $\mathcal{T}$. Then $\mathcal{T}$ is locally finite, since P is locally finite. Hence U is a closed subset of E^{n-1}. Now for any point z of U, there is a point w directly above z and an $r > 0$ such that

$$B(w, r) \subset \cup\{gP : g \in \Gamma \ \text{ and } \ x \in g\overline{P}\}.$$

Now $\nu B(w, r)$ is an open neighborhood of z in E^{n-1} contained in U. Hence U is an open subset of E^{n-1}. Thus U is both open and closed in E^{n-1} and therefore is all of E^{n-1}. As $\{gP : g \in \Gamma\}$ is an exact tessellation of U^n, we conclude that $\mathcal{T}$ is an exact tessellation of E^{n-1}.

Now by Theorem 6.7.2, the tessellation $\mathcal{T}$ is connected. Hence, there are elements $f_1, \ldots, f_m$ of Γ such that x is in the set $f_i^{-1}\overline{P}$ for each i and $\nu P = \nu f_1^{-1}P$, $\nu f_m^{-1}P = \nu f^{-1}P$, and $\nu f_{i-1}^{-1}P$ and $\nu f_i^{-1}P$ share a common side for each $i > 1$. Then $P = f_1^{-1}P$, $f_m^{-1}P = f^{-1}P$, and $f_{i-1}^{-1}P$ and $f_i^{-1}P$ share a common vertical side for each $i > 1$. Hence $f_1 = 1$, $f_m = f$, and

P and $f_{i-1}f_i^{-1}P$ share a common side S_i for each $i > 1$. We may assume that $f_{i-1} \neq f_i$ for each $i > 1$. Then $f_{i-1}f_i^{-1} = g_{S_i}$ for each $i > 1$. Let $x_1 = x$ and $x_i = f_i x$ for each $i > 1$. As x is in $f_i^{-1}\overline{P}$, we find that $f_i x$ is in $\overline{P}$. Hence x_i is in $\overline{P}$ for each i. Now observe that

$$g_{S_i}(x_i) = f_{i-1}f_i^{-1}(x_i) = f_{i-1}(x) = x_{i-1}.$$

Hence x_{i-1} is in $\overline{P} \cap g_{S_i}(\overline{P})$. Therefore x_{i-1} is in $\overline{S}_i$ and x_i is in $\overline{S}_i'$ for each $i > 1$. Hence

$$x = x_1 \simeq x_2 \simeq \cdots \simeq x_m = y.$$

Therefore $x \sim y$. Thus $[x] = P \cap \Gamma x$. □

Theorem 12.3.4. *Let P be a geometrically finite, exact, convex, fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(B^n)$. Then every point of $\overline{P} \cap S^{n-1}$ is either an ordinary point or a cusped limit point of Γ.*

Proof: Let x be a point of $\overline{P} \cap S^{n-1}$ and let g be an element of Γ such that x is in $g\overline{P}$. Then $g^{-1}x$ is in $\overline{P} \cap \Gamma x$. By Theorem 12.3.3, there are elements $g_1, \ldots, g_k$ of Γ such that

$$\overline{P} \cap \Gamma x = \{g_1^{-1}x, \ldots, g_k^{-1}x\}.$$

Hence $g^{-1}x = g_i^{-1}x$ for some i. Then $x = gg_i^{-1}x$ and so gg_i^{-1} is in Γ_x. Thus, we have that

$$g \in \Gamma_x g_1 \cup \cdots \cup \Gamma_x g_k.$$

Assume first that Γ_x is finite. Then g is one of only finitely many elements of Γ, say $g_1, \ldots, g_\ell$. We pass to the projective disk model D^n. Let $r > 0$ be less than the Euclidean distance from x to any side of $g_i P$ that does not contain x. Let y be a point of $D^n \cap B(x, r)$ and let $[x, y]$ be the line segment from x to y. From the proof of Theorem 12.3.3(2), we see that the line segment $[x, y]$ starts off at x and immediately enters $g_i P$ for some i. The ray can exit $g_i P$ only at one of its sides not containing x. As $[x, y] \subset B(x, r)$, we deduce that $[x, y] \subset g_i P$. Therefore y is in $g_i P$. Thus

$$D^n \cap B(x, r) \subset g_1 P \cup \cdots \cup g_\ell P.$$

Hence x is not a limit point of Γ, and so x is an ordinary point of Γ.

Now assume that Γ_x is infinite. Then Γ_x is an elementary group of either parabolic or hyperbolic type. By Theorem 12.2.4, the point x is not fixed by a hyperbolic element of Γ. Therefore Γ_x is of parabolic type. Hence x is the fixed point of a parabolic element of Γ.

We now pass to the upper half-space model U^n and conjugate Γ so that $x = \infty$. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then from the proof of Theorem 12.3.3(2), we have that

$$\mathcal{T} = \{\nu g P : g \in \Gamma \text{ and } x \in g\overline{P}\}$$

is an exact tessellation of E^{n-1}. As

$$\mathcal{T} = \{\nu g P : g \in \Gamma_x g_1 \cup \cdots \cup \Gamma_x g_k\},$$

we deduce that
$$E^{n-1} = \bigcup_{f\in\Gamma_\infty} f\Big(\bigcup_{i=1}^{k} \nu g_i P\Big).$$

By Theorems 5.4.6 and 7.4.2, there is a Γ_∞-invariant m-plane Q of E^{n-1} such that Q/Γ_∞ is compact. Since g_iP is geometrically finite for each $i = 1,\ldots,k$, there is an $r > 0$ such that $N(Q,r)$ contains every nonvertical side of $g_1P,\ldots,g_kP$. Let
$$U(Q,r) = \overline{U}^n - \overline{N}(Q,r).$$
Then we have that
$$U(Q,r) \subset \bigcup_{f\in\Gamma_\infty} f\Big(\bigcup_{i=1}^{k} g_i\overline{P}\Big).$$
Now since $g_jg_i^{-1}(\infty) \neq \infty$ for each i,j such that $i \neq j$, and since $g_jg_i^{-1}$ is continuous at ∞, we can increase r so that
$$g_jg_i^{-1}(U(Q,r)) \subset N(Q,r)$$
for each i,j such that $i \neq j$.

We claim that $U(Q,r)$ is a cusped region for Γ. On the contrary, suppose that there is an element g of $\Gamma - \Gamma_\infty$ such that
$$U(Q,r) \cap gU(Q,r) \neq \emptyset.$$
Since $U(Q,r)$ is an open subset of $\overline{U}^n$, there is a point y in the interior of $fg_i\overline{P}$ in $\overline{U}^n$ for some i and f in Γ_∞ such that gy is in $hg_j\overline{P}$ for some j and h in Γ_∞. Then we have
$$gfg_iP = hg_jP$$
and so $gfg_i = hg_j$. Then $i \neq j$ and
$$g = hg_jg_i^{-1}f^{-1}.$$
Therefore, we have
$$\begin{aligned} gU(Q,r) &= hg_jg_i^{-1}f^{-1}U(Q,r) \\ &= hg_jg_i^{-1}U(Q,r) \\ &\subset hN(Q,r) \\ &= N(Q,r), \end{aligned}$$
which is a contradiction. Hence $U(Q,r)$ is a cusped region for Γ. Thus x is a cusped limit point of Γ. □

Now since a geometrically finite convex polyhedron has only finitely many cusp points, the next corollary follows from Theorems 12.2.7 and 12.3.4.

Corollary 1. *If P is a geometrically finite, exact, convex, fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(B^n)$, then $\overline{P} \cap L(\Gamma)$ is finite.*

Geometrically Finite Groups

Definition: A discrete subgroup Γ of $\mathrm{M}(B^n)$ is *geometrically finite* if and only if Γ has a geometrically finite, exact, convex, fundamental polyhedron.

Remark: This is not the usual definition of a geometrically finite group. In the usual definition, polyhedra are finite-sided instead of geometrically finite. We shall prove that our new definition agrees with the usual definition when $n = 1, 2, 3$. The reason we have altered the usual definition is because the new definition seems to be the right definition when $n > 3$. This is justified by Theorem 12.3.5 and the examples below.

Theorem 12.3.5. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. Then the following are equivalent:*

(1) *The group Γ is geometrically finite.*

(2) *Every limit point of Γ is either conical or cusped.*

(3) *Every exact, convex, fundamental polyhedron for Γ is geometrically finite.*

Proof: Suppose that Γ is geometrically finite. Then Γ has a geometrically finite, convex, fundamental polyhedron P. By conjugating Γ, if necessary, we may assume that 0 is in P°. Let a be a limit point of Γ and let R be the ray in B^n from 0 to a. Assume first that R meets only finitely many members of the tessellation $\{gP : g \in \Gamma\}$. Then a is in $g\overline{P}$ for some g in Γ. Hence a is a cusped limit point of Γ by Theorem 12.3.4.

Now assume that R meets infinitely many members of the tessellation $\{gP : g \in \Gamma\}$. Then there is a sequence $\{g_i\}_{i=1}^\infty$ of distinct nonidentity elements of Γ and a sequence $\{x_i\}_{i=1}^\infty$ of points of P such that $g_i x_i$ is in $R \cap g_i P$ and $g_i x_i \to a$ as $i \to \infty$. Since the terms of $\{g_i\}$ are distinct, the Euclidean diameter of $g_i P$ goes to zero as $i \to \infty$. Hence $g_i(0) \to a$ as $i \to \infty$. By passing to a subsequence of $\{g_i\}$, we may assume that $g_i^{-1}(-a) \to b$ and $g_i^{-1}(a) \to c$.

Assume first that $b \neq c$. Then there exists a $\delta > 0$ such that for all i,

$$\left|g_i^{-1}(-a) - g_i^{-1}(a)\right| \geq \delta.$$

Hence a is a conical limit point of Γ by Theorem 12.2.5. Now assume that $b = c$. As x_i is on the line

$$g_i^{-1}(-a, a) = (g_i^{-1}(-a), g_i^{-1}(a)),$$

we deduce that $x_i \to c$. Hence c is in $\overline{P}$. Now since $L(\Gamma)$ is a Γ-invariant closed subset of S^{n-1} and $g_i^{-1}(a) \to c$, we have that c is a limit point of Γ. Therefore c is a cusped limit point of Γ by Theorem 12.3.4.

The point a is not in $g_i\overline{P}$, since the subray $[g_i x_i, a)$ of R is not contained in $g_i\overline{P}$, and the point $-a$ is not in $g_i\overline{P}$, since 0 is not in $g_i\overline{P}$. Hence, since

$\pm a$ is not in $g_i\overline{P}$ for all i, we have that $g_i^{-1}(\pm a)$ is not in $\overline{P}$ for all i. Therefore $g_i^{-1}(\pm a) \neq c$ for all i.

We now pass to the upper half-space model U^n and conjugate Γ so that $c = \infty$. Let $U(Q,r)$ be a cusped region for Γ based at ∞. As a is a limit point of Γ, we have that $g_i^{-1}(a)$ is a limit point of Γ for each i. Hence $g_i^{-1}(a)$ is in $\overline{N}(Q,r)$ for each i by Lemma 1 of §12.2. Now by Lemma 2 of §12.2, we have

$$\lim_{i\to\infty} \mathrm{dist}_E(x_i, Q) = \infty.$$

As x_i is on the line $(g_i^{-1}(-a), g_i^{-1}(a))$ and $g_i^{-1}(a)$ is in $\overline{N}(Q,r)$, we deduce that

$$\lim_{i\to\infty} \left|g_i^{-1}(-a) - g_i^{-1}(a)\right| = \infty.$$

Let D be a Dirichlet polyhedron for Γ_∞ in Q. Then D is compact. Hence $\overline{N}(D,r)$ is compact. Now for each i, there is an element f_i of Γ_∞ such that $f_i g_i^{-1}(a)$ is in $\overline{N}(D,r)$. As f_i is a Euclidean isometry,

$$\lim_{i\to\infty} \left|f_i g_i^{-1}(-a) - f_i g_i^{-1}(a)\right| = \infty.$$

Moreover $\{f_i g_i^{-1}(a)\}$ is bounded. We now pass back to the ball model B^n. Let $h_i = g_i f_i^{-1}$ for each i. Then there is a $\delta > 0$ such that for all i,

$$\left|h_i^{-1}(-a) - h_i^{-1}(a)\right| \geq \delta.$$

Now since the terms of $\{g_i\}$ are distinct, the Euclidean diameter of $g_i\overline{P}$ goes to zero as $i \to \infty$. Hence $g_i(c) \to a$, since $g_i(0) \to a$. Therefore

$$h_i(c) = g_i f_i^{-1}(c) = g_i(c) \to a.$$

As a is not in $g_i\overline{P}$ for all i, we have that $g_i(c) \neq a$ for all i. Hence $h_i(c) \neq a$ for all i. Therefore, infinitely many of the terms of $\{h_i\}$ are distinct. Hence, by passing to a subsequence, we may assume that all the terms of $\{h_i\}$ are distinct. Then the Euclidean diameter of $h_i\overline{P}$ goes to zero as $i \to \infty$. Hence $h_i(0) \to a$, since $h_i(c) \to a$. Therefore a is a conical limit point by Theorem 12.2.5. Thus (1) implies (2).

Now assume (2) that every limit point of Γ is either conical or cusped. Let P be an exact, convex, fundamental polyhedron for Γ. We now show that P is geometrically finite. Let x be a point of $\overline{P} \cap S^{n-1}$. Assume first that x is an ordinary point of Γ. Then by Theorem 12.1.12, there is an $r > 0$ such that $B(x,r)$ meets only finitely many members of $\{gP : g \in \Gamma\}$, say $g_1P, \ldots, g_kP$. By shrinking r, if necessary, we may assume that x is in $g_i\overline{P}$ for each $i = 1, \ldots, k$. Now suppose that $B(x,r)$ meets a side S of P. Then $B(x,r)$ meets g_SP. Hence $g_S = g_i$ for some i. Therefore x is $g_S\overline{P}$. By the argument in the proof of Theorem 12.2.7, we have that

$$\overline{P} \cap g_S(\overline{P}) = \overline{S}.$$

Therefore S is incident with x. Thus $B(x,r)$ meets only the sides of P incident with x.

Now assume that x is a limit point of Γ. Then x is not conical by Theorem 12.2.4. Therefore x is cusped by our hypothesis. Hence x is a cusp point of P by Theorem 12.2.7. Therefore, there is an $r > 0$ such that $B(x, r)$ meets only the sides of P incident with x by the definition of a cusp point. This completes the proof that P is geometrically finite. Thus (2) implies (3). Clearly (3) implies (1). □

Theorem 12.3.6. *If* Γ *is a geometrically finite discrete subgroup of* $\mathrm{M}(B^n)$ *with* $n = 1, 2, 3$, *then every exact, convex, fundamental polyhedron for* Γ *is finite-sided.*

Proof: Let P be an exact convex, fundamental polyhedron for Γ. Then P is geometrically finite by Theorem 12.3.5. By Theorem 12.3.1, it suffices to show that every cusp of P is finite-sided. On the contrary, suppose that c is the cusp point of an infinite-sided cusp of P. Then $n = 3$ and c is fixed by a parabolic element of Γ by Theorem 12.3.2. We now pass to the upper half-space model U^3 and conjugate Γ so that $c = \infty$. Let $\nu : U^3 \to E^2$ be the vertical projection. Then $\nu(P)$ is a convex polygon in E^2 whose sides are the vertical projections of the vertical sides of P. Hence $\nu(P)$ has infinitely many sides.

Assume first that E^2/Γ_∞ is compact. By Theorem 5.4.5, the group Γ_∞ has a torsion-free subgroup H of finite index. Then E^2/H is compact by Lemma 1 of §7.4. Now since $\nu(P^\circ)$ injects into the space-form E^2/H, we deduce that $\nu(P^\circ)$ has finite area. As $\nu(P^\circ)$ is convex, $\nu(P)$ is compact. Hence $\nu(P)$ has only finitely many sides, which is a contradiction.

Now assume that E^2/Γ_∞ is not compact. Then Γ_∞ has an infinite cyclic subgroup H of finite index by Theorem 5.4.6. Now H is generated by either a horizontal translation or a glide-reflection of E^3. Hence, by replacing H by a subgroup of index two, if necessary, we may assume that H is generated by a horizontal translation τ of E^3.

Let g be an element of Γ such that $P \cap gP$ is a vertical side S of P. We now show that there is at most one element $f \neq 1$ of H such that $P \cap fgP$ is a vertical side T of P. Let f be such an element. Then $f = \tau^m$ for some integer $m \neq 0$. By replacing τ by τ^{-1}, if necessary, we may assume that $m > 0$. Observe that f translates gP from the opposite side of $\langle S \rangle$ from P to the opposite side of $\langle T \rangle$ from P. Hence f translates a point of gP across $\langle S \rangle$ and a point of gP across $\langle T \rangle$.

Now let k be a nonzero integer. If $k < 0$, then $\tau^k gP$ lies on the opposite side of $\langle S \rangle$ from P, and so $P \cap \tau^k gP = \emptyset$. If $k > m$, then $\tau^k gP$ lies on the opposite side of $\langle T \rangle$ from P, and so $P \cap \tau^k gP = \emptyset$. Now suppose that $0 < k < m$. Choose points x in S° and y in T so that the Euclidean line segment $[x, y]$ is horizontal and sufficiently high enough so that $[x, y] \subset P$ and $[x, f^{-1}y] \subset gP$. Then $(x, y) \subset P^\circ$ and $(x, f^{-1}y) \subset gP^\circ$. Now observe that the line segments $\tau^k(x, f^{-1}y)$ and (x, y) intersect. See Figure 12.3.1. Hence P° and $\tau^k gP^\circ$ intersect. Therefore $\tau^k g = 1$ and so $P \cap \tau^k gP = P$.

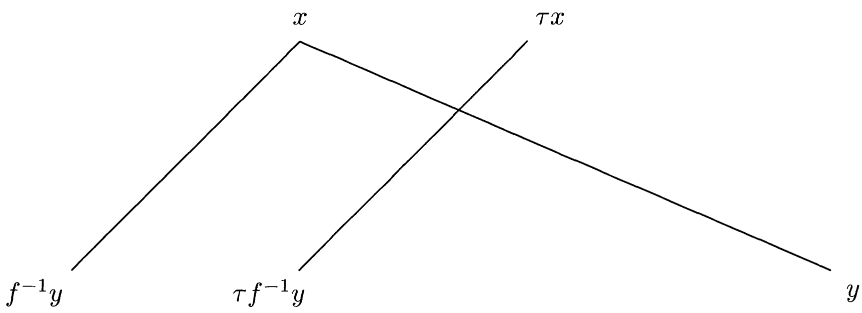


Figure 12.3.1. The line segment in the proof of Theorem 12.3.6

Thus f is the only nonidentity element of H such that $P \cap fgP$ is a vertical side of P.

Let $\{S_i\}_{i=1}^{\infty}$ be a sequence of distinct vertical sides of P. Then there is a sequence $\{g_i\}_{i=1}^{\infty}$ of distinct elements of Γ such that $P \cap g_i P = S_i$ for each i. Now each coset of H in Γ contains at most two terms of $\{g_i\}$. Hence, the terms of $\{g_i\}$ fall into infinitely many cosets of H in Γ. As H has finite index in Γ_∞, the terms of $\{g_i\}$ must fall into infinitely many cosets of Γ_∞ in Γ. But by the proof of Theorem 12.3.4, the terms of $\{g_i\}$ lie in only finitely many cosets of Γ_∞ in Γ, which is a contradiction. Thus P is finite-sided. □

Theorem 12.3.7. *If Γ is a geometrically finite discrete subgroup of $\mathrm{M}(B^n)$ with no parabolic elements, then every exact, convex, fundamental polyhedron for Γ is finite-sided.*

Proof: Let P be in exact, convex fundamental polyhedron for Γ. Then P is geometrically finite by Theorem 12.3.5. The polyhedron P has no infinite-sided cusps by Theorem 12.3.2. Therefore P is finite-sided by Theorem 12.3.1. □

Theorem 12.3.8. *If Γ is a geometrically finite discrete subgroup of $\mathrm{M}(B^n)$ of the first kind, then B^n/Γ has finite volume and every exact, convex, fundamental polyhedron for Γ is finite-sided.*

Proof: Let P be an exact, convex fundamental polyhedron for Γ. Then P is geometrically finite by Theorem 12.3.5. Let v be a point of $\overline{P} \cap S^{n-1}$. We claim that v is an ideal vertex of P. On the contrary, suppose that v is not an ideal vertex of P. We now pass to the upper half-space model U^n and conjugate Γ so that $v = \infty$. Since P is geometrically finite, there is an $r > 0$ so that $C(0, r)$ contains all the nonvertical sides of P. Let

$\nu : U^n \to E^{n-1}$ be the vertical projection. Then $\nu(P)$ is a noncompact convex polyhedron in E^{n-1}. Hence, the set $\nu(P^\circ)$ is unbounded. Therefore $\nu(P^\circ) - C(0,r)$ is a nonempty open subset of E^{n-1}. Hence, there is a point b of $\nu(P^\circ) - C(0,r)$ and an $s > 0$ so that

$$C(b,s) \cap E^{n-1} \subset \nu(P^\circ) - C(0,r).$$

Now since $C(0,r)$ contains all the nonvertical sides of P, we have that $C(b,s) \subset \overline{P}$. Therefore Γ is of the second kind by Theorem 12.1.14, which is a contradiction. Thus v is an ideal vertex of P. Hence P has finitely many sides and finite volume by Theorems 6.3.24 and 6.3.26. □

Theorem 12.3.9. *Every geometrically finite discrete subgroup of* $\mathrm{M}(B^n)$ *is finitely generated.*

Proof: Let Γ be a geometrically finite discrete subgroup of $\mathrm{M}(B^n)$. Then Γ has a geometrically finite, exact, convex, fundamental polyhedron P. By Theorem 6.7.3, the group Γ is generated by the Γ-side-pairing

$$\Phi = \{g_S : S \text{ is a side of } P\}.$$

If P is finite-sided, then Φ is a finite set, and we are done, so assume that P is infinite-sided. Then P has an infinite-sided cusp and its cusp point is fixed by a parabolic element of Γ by Theorems 12.3.1 and 12.3.2. Moreover P has only finitely many cusp points $c_1, \ldots, c_m$ that are fixed by a parabolic element of Γ, and all but finitely many sides of P, say $S_1, \ldots, S_\ell$, are incident with c_i for some i. Let Γ_i be the stabilizer of c_i for each i. Then Γ_i is an elementary group of parabolic type. Hence Γ_i is finitely generated for each i. Let $\{f_{ij}\}$ be a finite set of generators of Γ_i for each i. By Theorem 12.3.3, the cycle $[c_i]$ is finite for each i. Let $[c_i] = \{g_{ij}c_i\}$ for each i, and let Ψ be the union of the sets $\{f_{ij}\}$ and $\{g_{ij}\}$, for $i = 1, \ldots, m$, and $\{g_{S_k}\}$. Then Ψ is a finite subset of Γ.

We now show that Ψ generates Γ. Since Φ generates Γ, it suffices to show that $\Phi \subset \langle\Psi\rangle$. Let S be a side of P. If S is not incident with c_i for some i, then $S = S_k$ for some k, and so g_S is in Ψ. Assume now that S is incident with c_i for some i. Then $g_{S'}(c_i)$ is fixed by a parabolic element of Γ, and so $g_{S'}(c_i) = c_k$ for some k. As $c_i \simeq c_k$, we have that $c_k = g_{ij}c_i$ for some j. Then we have

$$g_{S'}(c_i) = g_{ij}(c_i),$$

whence $g_S g_{ij}$ is in Γ_i. Therefore g_S is $\langle\Psi\rangle$. This shows that $\Phi \subset \langle\Psi\rangle$. Thus Γ is finitely generated. □

It is well known that a discrete subgroup of $\mathrm{M}(B^2)$ is geometrically finite if and only if it is finitely generated. We next consider an example of a finitely generated discrete subgroup of $\mathrm{M}(B^3)$ that is not geometrically finite.

Example 3. Let Γ be the figure-eight knot group in Example 4 of §12.1. Then Γ is a discrete subgroup of $\mathrm{M}(B^3)$ such that B^3/Γ has finite volume. Let Γ' be the commutator subgroup of Γ. Then Γ' is a free group of rank two and a discrete subgroup of $\mathrm{M}(B^3)$ of the first kind such that B^3/Γ' has infinite volume. Therefore Γ' is not geometrically finite by Theorem 12.3.8.

We next consider an example that shows that Theorem 12.3.6 cannot be generalized to higher dimensions.

Example 4. Let θ be a real number such that θ/π is irrational and let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Then A is an irrational rotation with axis $\mathbb{R}$ in E^3. Let $f = e_1 + A$. Then f is an isometry of E^3 that leaves $\mathbb{R}$ invariant. The infinite cyclic group Γ generated by f is a discrete group of isometries of E^3. Let a be a point of E^3 and let $P(a)$ be the Dirichlet polyhedron for Γ with center a. If a is in $\mathbb{R}$, then $P(a)$ is the closed region between the two parallel planes orthogonal to $\mathbb{R}$ at a distance $1/2$ from a.

Assume now that a is not in $\mathbb{R}$. We claim that $P(a)$ has infinitely many sides. On the contrary, assume that $P = P(a)$ is finite-sided. Let S be a side of P. Then $\langle S\rangle$ is the perpendicular bisector of the line segment $[a, f^m a]$ for some integer $m \neq 0$. Consequently $\langle S\rangle$ intersects the line

$$L = \{(a_1, ta_2, ta_3) : t \in \mathbb{R}\}$$

passing through a and orthogonal to $\mathbb{R}$ below the ray

$$R = \{(a_1, ta_2, ta_3) : t \geq 1\}.$$

Hence P is contained in the closed half-space of E^3 bounded by $\langle S\rangle$ and containing R.

Now as a is an P°, there is an $r > 0$ so that $C(a, r) \subset P$. Define

$$\rho : \partial P \to S(a, r)$$

by letting $\rho(x)$ be the intersection of the line segment $[a, x]$ with the sphere $S(a, r)$. Then ρ is an injection. From the description of $\rho(\partial P)$ in Lemma 1, we deduce that $S(a, r) - \rho(\partial P)$ is a finite-sided convex polygon in $S(a, r)$ that contains the point $R \cap S(a, r)$ in its interior. Consequently, there is a solid cone C in E^3, with axis R, such that

$$C \cap S(a, r) \subset S(a, r) - \rho(\partial P).$$

Then $C \subset P^\circ$. Hence, the cones $\{f^m C\}_{m=1}^\infty$ are mutually disjoint; but the same argument as in the proof of Lemma 3 of §12.2 shows that this is impossible. Thus $P(a)$ is infinite-sided.

We now extend Γ to a discrete subgroup of $\mathrm{M}(U^4)$ by Poincaré extension. Let $\nu : U^4 \to E^3$ be the vertical projection. For each point u of U^4, let $P(u)$ be the Dirichlet polyhedron of Γ in U^4 with center u. Then $P(u)$ is a vertical prism over the polyhedron $\nu P(u)$ in E^3. Moreover, we have that $\nu P(u) = P(\nu(u))$. Therefore $P(u)$ is finite-sided if and only if $\nu(u)$ is in $\mathbb{R}$. Thus Γ is a geometrically finite discrete subgroup of $\mathrm{M}(U^4)$ such that some of its Dirichlet polyhedra are infinite-sided.

Example 5. We now consider an example of nonelementary, geometrically finite, discrete subgroup of $\mathrm{M}(U^4)$ such that some of its Dirichlet polyhedra are infinite-sided. Let P be the Schottky polyhedron in U^4 with two vertical sides

$$P(-e_1, 1/2) \cap U^4 \quad \text{and} \quad P(e_1, 1/2) \cap U^4,$$

and two nonvertical sides $S(-e_2, 1/2) \cap U^4$ and $S(e_2, 1/2) \cap U^4$. We pair the vertical sides of P by the element f of Example 4. Let L be the hyperbolic line of U^4 that is orthogonal to the nonvertical sides of P. We pair the nonvertical sides of P by the hyperbolic translation h of U^4, with axis L, that maps one side to the other. Let Γ be the subgroup of $\mathrm{M}(U^4)$ generated by f and h. Then Γ is a free discrete subgroup of $\mathrm{M}(U^4)$ of rank 2 by Theorem 12.1.19. Therefore Γ is a nonelementary subgroup of $\mathrm{M}(U^4)$.

For each point u of U^4, let $D(u)$ be the Dirichlet polyhedron for Γ with center u. Let v be the point of L midway between the nonvertical sides of P. Then $P = D(v)$, since P contains $D(v)$ and P° is a Γ-packing. Therefore P is an exact, convex, fundamental polyhedron for Γ with four sides. Hence Γ is geometrically finite.

Now as ∞ is a limit point of Γ in $\overline{P}$, we have that ∞ is cusped by Theorem 12.3.4. Hence, there is a cusped region U for Γ. Let B be a horoball based at ∞ and contained in U. We now show that $D(u)$ is infinite-sided for each u in B such that $\nu(u)$ is not in $\mathbb{R}$. Let u be such a point and let g an element of Γ such that the hyperplane

$$P_g(u) = \{x \in U^4 : d(x, u) = d(x, gu)\}$$

contains a side S of $D(u)$. If g is in Γ_∞, then S is a vertical side of $D(u)$. If g is not in Γ_∞, then gu is not in B, and so S is a nonvertical side of $D(u)$ with u above $\langle S \rangle$.

Now assume that $D(u)$ has only finitely many sides. Then $D(u)$ has only finitely many nonvertical sides, say $S_1, \ldots, S_m$. Let H_i be the closed half-space of U^4 bounded by $\langle S_i \rangle$ and containing u for each i. Let $P(u)$ be the Dirichlet polyhedron for Γ_∞ with center u. Then

$$D(u) = P(u) \cap \bigcap_{i=1}^{m} H_i.$$

But $P(u)$ has infinitely many sides, and so $D(u)$ has infinitely many vertical sides, which is a contradiction. Thus $D(u)$ has infinitely many sides.

Exercise 12.3

1. Let P be a finite-sided, convex, fundamental polygon for a discrete subgroup Γ of $\mathrm{M}(B^2)$. Prove that a cusp point c of P is a cusped limit point of Γ if and only if every element of $[c]$ is a cusp point of P.
2. Let P be a finite-sided, convex, fundamental polyhedron of finite volume for a discrete subgroup Γ of $\mathrm{M}(B^n)$. Prove that every ideal vertex of P is a cusped limit point of Γ.
3. Let P be a geometrically finite, convex, fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(B^n)$. A cusp of P is said to be *thin* if the link of its cusp point does not contain a Euclidean hypercone. Prove that a cusp point c of P is a cusped limit point of Γ if and only if every element of $[c]$ is a thin cusp point of P.
4. Let P be a convex fundamental polyhedron for a discrete subgroup Γ of $\mathrm{M}(B^n)$ and let b be a point of $\overline{P} \cap S^{n-1}$ for which there is no $r > 0$ such that $B(b, r)$ meets only the sides of P incident with b. Prove that b is a limit point of Γ that is neither conical nor cusped.
5. Prove that every elementary discrete subgroup of $\mathrm{M}(B^n)$ is geometrically finite.
6. Let H be a subgroup of finite index of a discrete subgroup Γ of $\mathrm{M}(B^n)$. Prove that H is geometrically finite if and only if Γ is geometrically finite.
7. Prove that every nonelementary, geometrically finite, discrete subgroup of $\mathrm{M}(B^n)$ contains a subgroup that is not geometrically finite.
8. Let Γ be a geometrically finite discrete subgroup of $\mathrm{M}(B^n)$. Prove that every convex, fundamental polyhedron for Γ is geometrically finite.
9. Let $D(u)$ be the Dirichlet polyhedron in Example 5 with $\nu(u)$ not in $\mathbb{R}$. Prove that $D(u)$ has infinitely many vertical sides.
10. Let x be an irrational number. Prove that there is a sequence $\{d_i/c_i\}_{i=1}^{\infty}$ of distinct rational numbers such that
$$|x - d_i/c_i| = O(|c_i|^{-2}).$$

§12.4. Nilpotent Groups

In this section, we study nilpotent subgroups of $\mathrm{I}(H^n)$. In particular, we prove that every discrete subgroup of $\mathrm{I}(H^n)$ generated by elements sufficiently close enough to the identity is abelian. As an application, we prove that a subgroup of $\mathrm{I}(H^n)$ is discrete if and only if all its abelian and two-generator subgroups are discrete.

Lemma 1. *Let A, B be in $\mathrm{O}(n)$ with $|B - I| < 2$. If A commutes with $[B, A]$, then A commutes with B.*

Proof: If A commutes with $[B, A] = BAB^{-1}A^{-1}$, then A commutes with BAB^{-1}, and so A commutes with B by Lemma 3 of §5.4. □

Lemma 2. *If G is a nilpotent subgroup of $\mathrm{O}(n)$ generated by elements A such that $|A - I| < 2$, then G is abelian.*

Proof: Let A and B be elements of G such that $|A-I| < 2$ and $|B-I| < 2$. On the contrary, assume that A and B do not commute. Consider a nested chain of commutators

$$D = [C_1, [C_2, \ldots, [C_m, C_{m+1}] \cdots]],$$

where $C_i = A$ or B for all i. Since G is nilpotent, there is a maximal length m such that $D \neq I$. Assume that m has this value. Then D commutes with A and B. Hence D commutes with $[C_2, \ldots, [C_m, C_{m+1}] \cdots]]$. Therefore $[C_2, \ldots, [C_m, C_{m+1}] \cdots]]$ commutes with C_1 by Lemma 1, which is a contradiction. Hence A and B commute. Therefore G is abelian. □

Lemma 3. *If G is a nilpotent subgroup of $\mathrm{S}(E^n)$ generated by elements $a + kA$ such that $|A - I| < 2$, then G is abelian.*

Proof: Define $\eta : G \to \mathrm{O}(n)$ by $\eta(a + kA) = A$. Then η is a homomorphism. Hence $\eta(G)$ is a nilpotent subgroup of $\mathrm{O}(n)$. By Lemma 2, we have that $\eta(G)$ is an abelian subgroup of $\mathrm{O}(n)$. Let $\phi = a + kA$ and $\psi = b + \ell B$ be in G with $|A - I| < 2$ and $|B - I| < 2$. Then A and B are in $\eta(G)$ and so A and B commute. Hence

$$\begin{aligned}
[\phi, \psi] &= \phi\psi\phi^{-1}\psi^{-1} \\
&= \phi\psi\phi^{-1}(-\ell^{-1}B^{-1}b + \ell^{-1}B^{-1}) \\
&= \phi\psi(-k^{-1}A^{-1}a - k^{-1}\ell^{-1}A^{-1}B^{-1}b + k^{-1}\ell^{-1}A^{-1}B^{-1}) \\
&= \phi(b - \ell k^{-1}BA^{-1}a - k^{-1}A^{-1}b + k^{-1}A^{-1}) \\
&= a + kAb - \ell Ba - b + I \\
&= (kA - I)b + (I - \ell B)a + I.
\end{aligned}$$

Now set

$$c = (kA - I)b + (I - \ell B)a.$$

Define a sequence $\{\phi_m\}$ in G by $\phi_1 = [\phi, [\phi, \psi]]$ and $\phi_m = [\phi, \phi_{m-1}]$. Then we have

$$\phi_1 = (kA - I)c + I,$$

and, in general, we have

$$\phi_m = (kA - I)^m c + I.$$

As G is nilpotent, $\phi_m = I$ for some m. Assume first that $k = 1 = \ell$. Then the same argument as at the end of the proof of Lemma 5 of §5.4 shows that ϕ and ψ commute. Now assume that one of k or ℓ is not 1. Without loss of generality, we may assume that $k \neq 1$. Then the null space of $kA - I$ is zero. Hence $(kA - I)^m c = 0$ implies that $c = 0$. Therefore ϕ and ψ commute. Thus G is abelian. □

Lemma 4. *If A is in* $\mathrm{PO}(n,1)$, *then*

$$|A|^2 = (n+1) + 4\sinh^2 d_H(e_{n+1}, Ae_{n+1}).$$

Proof: Let $A = (a_{ij})$. By Theorem 3.1.3, the columns (and rows) of A form a Lorentz orthonormal basis of $\mathbb{R}^{n,1}$. Therefore, we have

$$a_{i,1}^2 + \cdots + a_{i,n}^2 - a_{i,n+1}^2 = \begin{cases} 1 & \text{if } i < n+1 \\ -1 & \text{if } i = n+1 \end{cases}$$

and

$$a_{1,n+1}^2 + \cdots + a_{n,n+1}^2 - a_{n+1,n+1}^2 = -1.$$

Hence, we have

$$\begin{aligned}
|A|^2 &= \sum_{i=1}^{n+1}\sum_{j=1}^{n+1} a_{ij}^2 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n+1} a_{ij}^2 + \sum_{j=1}^{n+1} a_{n+1,j}^2 \\
&= \sum_{i=1}^{n}(1 + 2a_{i,n+1}^2) + (-1 + 2a_{n+1,n+1}^2) \\
&= (n-1) + 2\sum_{i=1}^{n+1} a_{i,n+1}^2 \\
&= (n-1) + 2(-1 + 2a_{n+1,n+1}^2) \\
&= (n-3) + 4a_{n+1,n+1}^2 \\
&= (n-3) + 4(-e_{n+1} \circ Ae_{n+1})^2 \\
&= (n-3) + 4\cosh^2 d_H(e_{n+1}, Ae_{n+1}) \\
&= (n+1) + 4\sinh^2 d_H(e_{n+1}, Ae_{n+1}).
\end{aligned}$$

□

Lemma 5. *Every nilpotent subgroup of* $\mathrm{M}(B^n)$ *fixes a point of* $\overline{B}^n$.

Proof: Let G be a nilpotent subgroup of $\mathrm{M}(B^n)$. Then G is elementary by Theorem 5.5.10. If G is of either elliptic or parabolic type, then G fixes a point of $\overline{B}^n$, so we may assume that G is of hyperbolic type. We pass to the upper half-space model U^n. By Theorem 5.5.6, we may conjugate G so that G leaves the set $\{0, \infty\}$ invariant.

We claim that G fixes both 0 and ∞. On the contrary, assume that G fixes neither 0 nor ∞. Let G_1 be the subgroup of G that fixes the nth axis L of U^n. We now show that G_1 is a normal subgroup of G. Let f be in G_1, let g be in G, and let y be in L. As g leaves L invariant, there is a point x of L such that $y = gx$. Then

$$gfg^{-1}y = gfx = gx = y.$$

Thus gfg^{-1} is in G_1 and so G_1 is a normal subgroup of G.

Let G_2 be the subgroup of G that fixes both 0 and ∞. Then G_2 is of index two in G. We now show that $G_1 \neq G_2$. On the contrary, suppose that $G_1 = G_2$. Let h be an element of $G - G_2$. Then h leaves L invariant and so fixes a point z of L. As G is generated by G_2 and h, we have that G fixes z, contrary to the assumption that G is of hyperbolic type. Therefore $G_1 \neq G_2$.

As G is nilpotent, G/G_1 is nilpotent. Therefore, the center of G/G_1 is nontrivial. Hence, there is an element g of $G_2 - G_1$ and an element h of $G - G_2$ such that

$$hgh^{-1} = g \mod G_1.$$

Now $g = kA$ for some $k > 0$, with $k \neq 1$, and A in $\mathrm{O}(n)$, with $A(e_n) = e_n$; and $h = \ell B\sigma$, for some $\ell > 0$, and B in $\mathrm{O}(n)$, with $B(e_n) = e_n$, and $\sigma(x) = x/|x|^2$. Then

$$\begin{aligned} hgh^{-1} &= \ell B\sigma kA\sigma\ell^{-1}B^{-1} \\ &= \ell Bk^{-1}A\ell^{-1}B^{-1} \quad = \quad k^{-1}BAB^{-1}. \end{aligned}$$

But we have that

$$k^{-1}BAB^{-1} \neq kA \mod G_1,$$

which is a contradiction. Hence $G = G_2$. □

Lemma 6. *If f is the parabolic translation of U^2 defined by $f(z) = z+1$, then f corresponds to the Möbius transformation g of B^2 defined by*

$$g(z) = \frac{(1+i/2)z + (1/2)}{(z/2) + (1-i/2)}$$

and g corresponds to the matrix A in $\mathrm{PO}(2,1)$ *defined by*

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1/2 & 1/2 \\ 1 & -1/2 & 3/2 \end{pmatrix}.$$

Proof: The standard transformation $\eta : U^2 \to B^2$ has the property that $\eta(0) = -i$, $\eta(i) = 0$, and $\eta(\infty) = i$. Therefore $\eta(z) = \frac{iz+1}{z+i}$. Hence $g = \eta f \eta^{-1}$ is given by the matrix product

$$\begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -i/2 & 1/2 \\ 1/2 & -i/2 \end{pmatrix} = \begin{pmatrix} 1+i/2 & 1/2 \\ 1/2 & 1-i/2 \end{pmatrix}.$$

Now let $\zeta : B^2 \to H^2$ be stereographic projection. Then g corresponds to the matrix A in $\mathrm{PO}(2,1)$ extending $\zeta g \zeta^{-1}$. From Formulas 4.5.2 and 4.5.3, we have that

$$Ae_3 = \zeta g \zeta^{-1}(e_3) = \zeta g(0) = \zeta(2/5, 1/5) = (1, 1/2, 3/2).$$

Therefore, we have

$$A = \begin{pmatrix} a_{11} & a_{12} & 1 \\ a_{21} & a_{22} & 1/2 \\ a_{31} & a_{32} & 3/2 \end{pmatrix}.$$

As f fixes ∞, we have that g fixes i. Consequently $(0,1,1)$ is an eigenvector of A. This, together with the fact that the second and third columns of A are Lorentz orthogonal, implies that

$$A = \begin{pmatrix} a_{11} & -1 & 1 \\ a_{21} & 1/2 & 1/2 \\ a_{31} & -1/2 & 3/2 \end{pmatrix}.$$

Finally, the first column of A can be derived from the information that the columns of A are Lorentz orthogonal and $\det A = 1$. □

Theorem 12.4.1. *Let G be a nilpotent subgroup of* $\mathrm{PO}(n,1)$ *generated by elements A such that $|A - I| < 2$. Then G is abelian.*

Proof: By Theorem 5.5.10, we have that G is an elementary subgroup of $\mathrm{PO}(n,1)$. Let A be an element of G such that $|A - I| < 2$. Then

$$|A - I|^2 = |A|^2 - 2\mathrm{tr}A + (n+1).$$

By Lemma 4, we have

$$|A|^2 = (n+1) + 4\sinh^2 d(e_{n+1}, Ae_{n+1}).$$

Therefore

$$|A - I|^2 = 2(n + 1 - \mathrm{tr}A + 2\sinh^2 d(e_{n+1}, Ae_{n+1})).$$

Assume first that G is of elliptic type. Then G is conjugate in $\mathrm{PO}(n,1)$ to a subgroup G' of $\mathrm{O}(n+1)$. Let A' be the element of G' corresponding to A. Then

$$|A' - I|^2 = 2(n + 1 - \mathrm{tr}A') = 2(n + 1 - \mathrm{tr}A).$$

Therefore, we have

$$|A' - I|^2 \leq |A - I|^2 < 4.$$

Therefore G' is abelian by Lemma 2. Hence G is abelian.

Now assume that G is not elliptic. Then G fixes a point on the sphere at infinity of H^n by Lemma 5. Hence, there is a subgroup G' of $\mathrm{S}(E^{n-1})$ whose Poincaré extension in $\mathrm{M}(U^n)$ corresponds to a conjugate of G in $\mathrm{PO}(n,1)$. Let $\phi = a + kA'$ be the element of G' corresponding to A. We shall prove that $|A' - I| < 2$. Now since

$$|A' - I| = |BA'B^{-1} - I|$$

for all B in $\mathrm{O}(n-1)$, we are free to conjugate ϕ in $\mathrm{S}(E^{n-1})$.

Assume first that ϕ is elliptic. Then by conjugating ϕ in $\mathrm{I}(E^{n-1})$, we may assume that $a = 0$ and $k = 1$. Let $\tilde{A}'$ be the Poincaré extension of A'. Then $\tilde{A}'$ is in $\mathrm{O}(n)$ and $\tilde{A}'e_n = e_n$. Let $\eta : U^n \to B^n$ be the standard transformation. Then $\eta = \sigma\rho$, where ρ is the reflection of E^n in E^{n-1} and

σ is the inversion in the sphere $S(e_n, \sqrt{2})$. Hence

$$\begin{aligned}
\eta\tilde{A}'\eta^{-1}(x) &= \sigma\rho\tilde{A}'\rho\sigma(x) \\
&= \sigma\tilde{A}'\sigma(x) \\
&= \sigma\tilde{A}'\left(e_n + \frac{2(x-e_n)}{|x-e_n|^2}\right) \\
&= \sigma\left(e_n + \frac{2(\tilde{A}'x-e_n)}{|\tilde{A}'x-e_n|^2}\right) \\
&= \sigma^2\tilde{A}'x \\
&= \tilde{A}'x.
\end{aligned}$$

Therefore $\eta\tilde{A}'\eta^{-1} = \tilde{A}'$. Hence A is conjugate in $\mathrm{PO}(n,1)$ to the block diagonal matrix

$$\begin{pmatrix} A' & 0 \\ 0 & I_2 \end{pmatrix},$$

where I_2 is the 2×2 identity matrix. Then we have

$$\begin{aligned}
|A'-I|^2 &= 2(n-1-\mathrm{tr}A') \\
&= 2(n+1-\mathrm{tr}A) \\
&\leq |A-I|^2 \\
&< 4.
\end{aligned}$$

Assume next that ϕ is parabolic. Then by conjugating ϕ in $\mathrm{S}(E^{n-1})$, we may assume that $a = e_{n-1}$, $k=1$, and $A'e_{n-1} = e_{n-1}$. By Lemma 6, we have that A is conjugate in $\mathrm{PO}(n,1)$ to the block diagonal matrix

$$\begin{pmatrix} A'' & 0 \\ 0 & B \end{pmatrix},$$

where A'' is the $(n-2)\times(n-2)$ matrix obtained from A' by deleting its last row and column, and B is the 3×3 matrix in Lemma 6. As $\mathrm{tr}B = 3$, we have that

$$\begin{aligned}
|A'-I|^2 &= 2(n-1-\mathrm{tr}A') \\
&= 2(n+1-\mathrm{tr}A) \\
&\leq |A-I|^2 \\
&< 4.
\end{aligned}$$

Assume now that ϕ is hyperbolic. Then by conjugating ϕ in $\mathrm{I}(E^{n-1})$, we may assume that $a=0$. Then A is conjugate in $\mathrm{PO}(n,1)$ to the block diagonal matrix

$$\begin{pmatrix} A' & 0 \\ 0 & C \end{pmatrix},$$

where

$$C = \begin{pmatrix} \cosh s & \sinh s \\ \sinh s & \cosh s \end{pmatrix}$$

and s is the hyperbolic distance translated by ϕ along its axis, that is, $s = |\log k|$. Let $\rho : H^n \to L$ be the nearest point retraction of H^n onto the axis L of A. It follows from Theorem 4.6.1 and Exercise 4.6.3 that for all x, y in H^n, we have

$$d(\rho(x), \rho(y)) \leq d(x, y).$$

Hence, we have

$$\begin{aligned} s &= d(\rho(e_{n+1}), A\rho(e_{n+1})) \\ &= d(\rho(e_{n+1}), \rho(Ae_{n+1})) \\ &\leq d(e_{n+1} Ae_{n+1}). \end{aligned}$$

Therefore

$$\begin{aligned} |A' - I|^2 &= 2(n - 1 - \mathrm{tr}A') \\ &= 2(n - 1 - \mathrm{tr}A + 2\cosh s) \\ &\leq 2(n - 1 - \mathrm{tr}A + 2\cosh^2 s) \\ &= 2(n + 1 - \mathrm{tr}A + 2\sinh^2 s) \\ &\leq 2(n + 1 - \mathrm{tr}A + 2\sinh^2 d(e_{n+1}, Ae_{n+1})) \\ &= |A - I|^2 \\ &< 4. \end{aligned}$$

Thus, in all three cases, we have that $|A' - I| < 2$. Therefore G' is abelian by Lemma 3. Hence G is abelian. □

Lemma 7. *Let A, B be matrices in* $\mathrm{GL}(n, \mathbb{C})$. *If $0 < |A - I| < 2 - \sqrt{3}$ and $0 < |B - I| < 2 - \sqrt{3}$, then*

$$\big|[A, B] - I\big| < \min\big\{|A - I|, |B - I|\big\}.$$

Proof: Suppose that $|A - I| < k < 1$ and $|B - I| < k < 1$. Observe that

$$A^{-1} - I = -(A - I) - (A - I)(A^{-1} - I).$$

Hence

$$|A^{-1} - I| \leq |A - I| + |A - I|\ |A^{-1} - I|.$$

Therefore

$$|A^{-1} - I| \leq \frac{|A - I|}{1 - |A - I|} < \frac{k}{1 - k}.$$

Let C be a complex $n \times n$ matrix. Then we have

$$CA^{-1} = C + C(A^{-1} - I).$$

Hence, we have

$$\begin{aligned} |CA^{-1}| &= |C + C(A^{-1} - I)| \\ &\leq |C| + |C(A^{-1} - I)| \\ &\leq |C|(1 + |A^{-1} - I|) \\ &< |C|\Big(1 + \frac{k}{1 - k}\Big) \ = \ \frac{|C|}{1 - k}. \end{aligned}$$

Let $A_1 = A - I$ and $B_1 = B - I$. Then we have

$$\begin{aligned}(ABA^{-1}B^{-1} - I) &= (AB - BA)A^{-1}B^{-1} \\ &= (A_1B_1 - B_1A_1)A^{-1}B^{-1}.\end{aligned}$$

Therefore, we have

$$\begin{aligned}|ABA^{-1}B^{-1} - I| &< \frac{|A_1B_1 - B_1A_1|}{(1-k)^2} \\ &\leq \frac{2|A-I|\,|B-I|}{(1-k)^2}.\end{aligned}$$

Now let $k = 2 - \sqrt{3}$. Then we have

$$|ABA^{-1}B^{-1} - I| < \frac{2|A-I|(2-\sqrt{3})}{(\sqrt{3}-1)^2} = |A - I|.$$

Likewise $|ABA^{-1}B^{-1} - I| < |B - I|$. □

Theorem 12.4.2. *Let Γ be a discrete subgroup of $\hat{\mathrm{SL}}(n,\mathbb{C})$ generated by elements A such that $|A - I| < 2 - \sqrt{3}$. Then Γ is nilpotent.*

Proof: Regard $\hat{\mathrm{SL}}(n,\mathbb{C})$ as a subset of $\mathbb{C}^{n^2}$. As

$$\hat{\mathrm{SL}}(n,\mathbb{C}) = \det{}^{-1}\{-1,1\},$$

we have that $\hat{\mathrm{SL}}(n,\mathbb{C})$ is closed in $\mathbb{C}^{n^2}$. As Γ is closed in $\hat{\mathrm{SL}}(n,\mathbb{C})$, the set Γ is closed in $\mathbb{C}^{n^2}$. Let

$$J = \Gamma \cap B(I, 2-\sqrt{3}) \quad \text{and} \quad K = \Gamma \cap C(I, 2-\sqrt{3}).$$

Then K is a compact discrete space. Therefore K and J are finite. Let m be the number of elements of J.

Suppose that $A_1, \ldots, A_k$ are elements of J. Define $[A_1] = A$ and

$$[A_1, \ldots, A_j] = [[A_1, \ldots, A_{j-1}], A_j]$$

and suppose that $[A_1, \ldots, A_j] \neq I$ for each $j = 1, \ldots, k$. Then by Lemma 7, we have that

$$\begin{aligned}|A_1 - I| &> |[A_1, A_2] - I| \\ &> |[A_1, A_2, A_3] - I| \\ &\;\;\vdots \\ &> |[A_1, \ldots, A_k] - I|.\end{aligned}$$

Therefore $A_1, [A_1, A_2], \ldots, [A_1, \ldots, A_k]$ are distinct nonidentity elements of J. Hence $k < m$. Consequently, any m-fold commutator of elements of J is trivial. By repeated application of the identities

$$\begin{aligned}[B, A] &= [A, B]^{-1}, \\ [A, B^{-1}] &= B^{-1}[B, A]B, \\ [A, BC] &= [A, B]B[A, C]B^{-1},\end{aligned}$$

we deduce that any m-fold commutator of elements of $\Gamma = \langle J \rangle$ is trivial. Thus Γ is nilpotent. □

Theorem 12.4.3. *Let Γ be a discrete subgroup of* $\mathrm{PO}(n, 1)$ *generated by elements A such that $|A - I| < 2 - \sqrt{3}$. Then Γ is abelian.*

Proof: By Theorem 12.4.2, we have that Γ is nilpotent, and by Theorem 12.4.1, we have that Γ is abelian. □

Theorem 12.4.4. *Let Γ be a subgroup of* $\mathrm{PO}(n, 1)$*. Then Γ is discrete if and only if*

(1) *every abelian subgroup of Γ is discrete; and*

(2) *every two-generator subgroup of Γ is discrete.*

Proof: Suppose that Γ is not discrete. Then there is a sequence $\{A_i\}_{i=1}^{\infty}$ of distinct elements of Γ such that $A_i \to I$. Without loss of generality, we may assume that $|A_i - I| < 2 - \sqrt{3}$ for all i. Let H be the subgroup of Γ generated by $\{A_i\}$. Then H is not discrete. If H is nonabelian, then A_i does not commute with A_j for some i, j, whence the subgroup generated by A_i and A_j is not discrete by Theorem 12.4.3. □

Exercise 12.4

1. A group Γ is said to be *locally discrete* if and only if every finitely generated subgroup of Γ is discrete. Prove that $\mathbb{Q}$ is an abelian, nondiscrete, locally discrete subgroup of $\mathbb{R}$.

2. Let
$$\Gamma = \left\{ \begin{pmatrix} \cos \pi x & -\sin \pi x \\ \sin \pi x & \cos \pi x \end{pmatrix} : x \in \mathbb{Q} \right\}.$$
Prove that Γ is an abelian, nondiscrete, locally discrete subgroup of $\mathrm{O}(2)$.

3. Let Γ be the group in Exercise 2, let H be a nonelementary discrete subgroup of $\mathrm{PO}(2, 1)$, and let
$$\mathrm{K} = \left\{ \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} : A \in \Gamma \ \text{ and } \ B \in \mathrm{H} \right\}.$$
Prove that K is a nonelementary, nondiscrete, locally discrete subgroup of $\mathrm{PO}(4, 1)$.

4. Let Γ be a nonelementary subgroup of $\mathrm{I}(H^n)$ such that Γ leaves no proper m-plane of H^n invariant. Prove that Γ is discrete if and only if every two-generator subgroup of Γ is discrete.

5. Find an example of a nondiscrete subgroup Γ of $\mathrm{PSL}(2, \mathbb{C})$ such that every abelian subgroup of Γ is discrete.

§12.5. The Margulis Lemma

In this section, we prove the Margulis lemma. We then use the Margulis lemma to prove the existence of Margulis regions for a discrete subgroup of $\mathrm{I}(H^n)$. As an application, we prove that for each dimension n there is a positive lower bound for the set of volumes of complete hyperbolic n-manifolds.

Definition: Given a discrete subgroup Γ of $\mathrm{I}(H^n)$, a point x of H^n, and $\epsilon > 0$, let $\Gamma_\epsilon(x)$ be the subgroup of Γ generated by the set

$$\{g \in \Gamma : d(gx, x) \leq \epsilon\}.$$

Theorem 12.5.1. (The Margulis lemma). *For each dimension n, there is an $\epsilon > 0$ such that for every discrete subgroup Γ of $\mathrm{I}(H^n)$ and for every point x of H^n, the group $\Gamma_\epsilon(x)$ is elementary.*

Proof: We pass to the conformal ball model B^n. Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. Let x be a point of B^n and let τ be the hyperbolic translation of B^n by x. Then for each $\epsilon > 0$, we have

$$\begin{aligned} \tau^{-1}\Gamma_\epsilon(x)\tau &= \tau^{-1}\langle g \in \Gamma : d(gx, x) \leq \epsilon\rangle\tau \\ &= \langle \tau^{-1}g\tau \in \tau^{-1}\Gamma\tau : d(\tau^{-1}g\tau(0), 0) \leq \epsilon\rangle \\ &= (\tau^{-1}\Gamma\tau)_\epsilon(0). \end{aligned}$$

Thus we may assume, without loss of generality, that $x = 0$. Let $\Gamma_\epsilon = \Gamma_\epsilon(0)$. For each positive integer ℓ, set

$$K_\ell = \{g \in \mathrm{M}(B^n) : d(g(0), 0) \leq 1/\ell\}.$$

Observe that K_ℓ corresponds to the subset $C(0, 1/\ell) \times \mathrm{O}(n)$ of $B^n \times \mathrm{O}(n)$ under the homeomorphism $\Phi : B^n \times \mathrm{O}(n) \to \mathrm{M}(B^n)$ of Theorem 5.2.8. Therefore K_ℓ is compact for each ℓ. The set K_ℓ obviously contains the identity I for each ℓ. Moreover K_ℓ is invariant under the inversion map of $\mathrm{M}(B^n)$ for each ℓ, since

$$d(g(0), 0) = d(0, g^{-1}(0)).$$

Let K_ℓ^ℓ be the set of all elements of $\mathrm{M}(B^n)$ of the form $g_1 \cdots g_\ell$ with g_i in K_ℓ for each $i = 1, \ldots, \ell$. Observe that if g_i is in K_ℓ for each $i = 1, \ldots, \ell$, then

$$\begin{aligned} d(g_1 \cdots g_\ell(0), 0) &\leq \sum_{i=1}^{\ell} d(g_1 \cdots g_{\ell+1-i}(0), g_1 \cdots g_{\ell-i}(0)) \\ &= \sum_{i=1}^{\ell} d(g_{\ell+1-i}(0), 0) \ \leq \ 1. \end{aligned}$$

Therefore $K_\ell^\ell \subset K_1$ for each ℓ.

Let U be the open neighborhood of I in $\mathrm{M}(B^n)$ corresponding to the open set

$$\{A \in \mathrm{PO}(n,1) : |A - I| < 2 - \sqrt{3}\}.$$

As $\mathrm{M}(B^n)$ is a topological group, with respect to the metric D_B, defined by Formula 5.2.1, there is an $r > 0$ such that if $B = B(I, r)$, then $B^{-1}B \subset U$. As the metric D_B is right-invariant, $Bg = B(g, r)$ for each g in $\mathrm{M}(B^n)$. By Lemma 6 of §5.4, there is a maximum number m of elements of the compact metric space K_1 with mutual distances at least r. Hence, we can have at most m mutually disjoint open balls in K_1 of radius r. Therefore, we can have at most m mutually disjoint right translates of B in $\mathrm{M}(B^n)$ by elements of K_1.

Let $\epsilon = 1/(m+1)$ and let $\mathrm{H} = \langle \Gamma_\epsilon \cap U \rangle$. Then H is an abelian subgroup of Γ_ϵ by Theorem 12.4.3. Let $Bf_1, \ldots, Bf_k$ be mutually disjoint right translates of B by elements of $\Gamma_\epsilon \cap K_1$ with k as large as possible. Then $k \leq m$. We now show that $\{\mathrm{H}f_i\}_{i=1}^k$ contains a full set of cosets for H in Γ_ϵ. Let g be in Γ_ϵ. As Γ_ϵ is generated by $\Gamma \cap K_{m+1}$, we can write $g = g_1 \cdots g_\ell$ with g_i in $\Gamma \cap K_{m+1}$ for each i. We assume that ℓ is as small as possible. We call ℓ the length of g.

Assume first that $\ell \leq m+1$. Then g is in $K_{m+1}^{m+1} \subset K_1$, and so g is in $\Gamma_\epsilon \cap K_1$. Therefore Bg meets Bf_i for some i. Hence gf_i^{-1} is in $B^{-1}B \subset U$. Therefore gf_i^{-1} is in H and so $\mathrm{H}g = \mathrm{H}f_i$. Now assume that $\ell > m+1$. Let $h_i = g_1 \cdots g_i$ for each $i = 1, \ldots, m+1$. Then h_i is in $K_{m+1}^{m+1} \subset K_1$ for each i. Consequently, the sets $\{Bh_i\}_{i=1}^{m+1}$ cannot all be disjoint; say Bh_i meets Bh_j with $i < j$. Let $\alpha = h_i$, $\beta = g_{i+1} \cdots g_j$, and $\gamma = g_{j+1} \cdots g_\ell$. Then $g = \alpha\beta\gamma$ with $B\alpha \cap B\alpha\beta \neq \emptyset$. Hence $\alpha(\alpha\beta)^{-1}$ is in $B^{-1}B \subset U$. Therefore $\alpha\beta^{-1}\alpha^{-1}$ is H and

$$\mathrm{H}g = \mathrm{H}(\alpha\beta^{-1}\alpha^{-1})(\alpha\beta\gamma) = \mathrm{H}\alpha\gamma.$$

Let $g' = \alpha\gamma$. Then $\mathrm{H}g = \mathrm{H}g'$ and the length of g' is less than the length of g. By induction, it follows that $\mathrm{H}g = \mathrm{H}g''$ with the length of g'' at most $m+1$. Hence $\mathrm{H}g = \mathrm{H}f_i$ for some i by the previous argument. Thus $\{\mathrm{H}f_i\}_{i=1}^k$ contains a full set of cosets for H in Γ_ϵ. Hence

$$[\Gamma_\epsilon : \mathrm{H}] \leq k \leq m.$$

Therefore Γ_ϵ is elementary by Theorem 5.5.9. □

Definition: The n-dimensional *Margulis constant* is the supremum c_n of all $\epsilon > 0$ that satisfy the n-dimensional Margulis lemma.

Note that the Margulis constant c_n is finite for each $n > 1$, since there are nonelementary, discrete subgroups Γ of $\mathrm{I}(H^n)$ such that H^n/Γ is compact for each $n > 1$.

Margulis Regions

Let Γ be a discrete subgroups of $\mathrm{M}(B^n)$. For each $r > 0$, set

$$V(\Gamma, r) = \{x \in B^n : d(x, gx) < r \quad \text{for some nonelliptic } g \text{ in } \Gamma\}.$$

Lemma 1. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$. Then $V(\Gamma, r)$ is a Γ-invariant open subset of B^n for each $r > 0$.*

Proof: Let x be a point of $V(\Gamma, r)$. Then there is a nonelliptic element g of Γ such that $d(x, gx) < r$. Let f be any element of Γ. Then

$$d(fx, fgf^{-1}fx) = d(x, gx) < r.$$

As fgf^{-1} is nonelliptic, fx is in $V(\Gamma, r)$. Thus $V(\Gamma, r)$ is Γ-invariant.
Now let $s = (r - d(x, gx))/2$. Then for each y in $B(x, s)$, we have

$$\begin{aligned} d(y, gy) &\leq d(y, x) + d(x, gx) + d(gx, gy) \\ &= 2d(x, y) + d(x, gx) \quad < \quad r. \end{aligned}$$

Therefore $V(\Gamma, r)$ contains $B(x, s)$. Thus $V(\Gamma, r)$ is open. □

Lemma 2. *Let Γ be an infinite, elementary, discrete subgroup of $\mathrm{M}(B^n)$. Then $V(\Gamma, r)$ is connected for each $r > 0$.*

Proof: Assume first that Γ is of parabolic type. We pass to the upper half-space model U^n and assume without loss of generality that Γ fixes ∞. Then Γ is the Poincaré extension of a discrete subgroup of $\mathrm{I}(E^{n-1})$ by Theorem 5.5.5. Hence, by Theorem 4.6.1, we deduce that for each x in U^n and each nonelliptic g in Γ, we have

$$\cosh d(x, gx) = 1 + \frac{|x - gx|^2}{2x_n^2}.$$

It is clear from the above formula that for all y sufficiently high enough on the vertical ray $[x, \infty)$ we have that $d(y, gy) < r$. Hence $V(\Gamma, r)$ contains $[y, \infty)$ for some y in $[x, \infty)$. Moreover, if x is in $V(\Gamma, r)$, then $V(\Gamma, r)$ contains $[x, \infty)$.

On the contrary, suppose that $V(\Gamma, r)$ is disconnected. Then there exist disjoint, nonempty, open subsets M and N of $V(\Gamma, r)$ such that

$$V(\Gamma, r) = M \cup N.$$

By Lemma 1, the sets M and N are open in U^n. Now no point of M is vertically above a point of N and vice versa. Let $\nu : U^n \to E^{n-1}$ be the vertical projection. Then $\nu(M)$ and $\nu(N)$ are disjoint, nonempty, open subsets of E^{n-1} such that

$$E^{n-1} = \nu(M) \cup \nu(N),$$

which is a contradiction. Thus $V(\Gamma, r)$ is connected.

Assume now that Γ is of hyperbolic type. Then without loss of generality, we may assume that Γ leaves invariant the positive nth axis in U^n. Let x be a point of $V(\Gamma, r)$. Then there is a nonelliptic g in Γ such that $d(x, gx) < r$. As g fixes both 0 and ∞, we have that g is hyperbolic. Hence, there is a positive constant k, with $k \neq 1$, and an A in $\mathrm{O}(n-1)$ such that $g = k\tilde{A}$. By Theorem 4.6.1, we have

$$\cosh d(x, gx) = 1 + \frac{|x - k\tilde{A}x|^2}{2kx_n^2}.$$

Now $|x|e_n$ is the nearest point to x on the positive nth axis. Let y be any point on the geodesic segment $[|x|e_n, x]$. Then $|y| = |x|$ and $y_n \geq x_n$. Observe that

$$\begin{aligned}
y \cdot \tilde{A}y &= (\nu(y) + y_n e_n) \cdot (A\nu(y) + y_n e_n) \\
&= \nu(y) \cdot A\nu(y) + y_n^2 \\
&= |\nu(y)|^2 \cos\theta(\nu(y), A\nu(y)) + y_n^2 \\
&= |\nu(y)|^2 \cos\theta(\nu(x), A\nu(x)) + y_n^2 \\
&= |\nu(x)|^2 \cos\theta(\nu(x), A\nu(x)) + x_n^2 \\
&\quad + (y_n^2 - x_n^2)(1 - \cos\theta(\nu(x), A\nu(x))) \\
&\geq |\nu(x)|^2 \cos\theta(\nu(x), A\nu(x)) + x_n^2 \quad = \quad x \cdot \tilde{A}x.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\frac{|y - k\tilde{A}y|^2}{y_n^2} &= \frac{|y|^2 - 2ky \cdot \tilde{A}y + k^2|y|^2}{y_n^2} \\
&\leq \frac{|x|^2 - 2kx \cdot \tilde{A}y + k^2|x|^2}{x_n^2} \\
&= \frac{|x - k\tilde{A}x|^2}{x_n^2}.
\end{aligned}$$

Therefore, we have

$$d(y, gy) \leq d(x, gx) < r.$$

Hence $V(\Gamma, r)$ contains the geodesic segment $[|x|e_n, x]$. Since $V(\Gamma, r)$ also contains the positive nth axis, $V(\Gamma, r)$ is connected. □

Lemma 3. *If Γ is a discrete subgroup of $\mathrm{M}(B^n)$ and $r > 0$, then*

$$V(\Gamma, r) = \cup\{V(\Gamma_a, r) : \ a \text{ is a fixed point of a nonelliptic element of } \Gamma\}.$$

Proof: Clearly, we have

$$V(\Gamma_a, r) \subset V(\Gamma, r)$$

for each point a fixed by a nonelliptic element of Γ. Now let x be an arbitrary point of $V(\Gamma, r)$. Then there is a nonelliptic element g of Γ such that $d(x, gx) < r$. Let a be a fixed point of g. Then g is in Γ_a, and so x is in $V(\Gamma_a, r)$. Thus $V(\Gamma, r)$ is the union of the sets $\{V(\Gamma_a, r)\}$. □

Theorem 12.5.2. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$ and suppose that $0 < r \leq c_n$, where c_n is the Margulis constant. Then the set of connected components of $V(\Gamma, r)$ is*

$$\{V(\Gamma_a, r) : a \text{ is a fixed point of a nonelliptic element of } \Gamma\}.$$

Proof: By Lemmas 1-3, it suffices to show that any two members of $\{V(\Gamma_a, r)\}$ are either disjoint or coincide. Suppose that a and b are two points fixed by nonelliptic elements of Γ, and suppose that x is in both $V(\Gamma_a, r)$ and $V(\Gamma_b, r)$. Then there are nonelliptic elements g and h of Γ, fixing a and b, respectively, such that $d(x, gx) < r$ and $d(x, hx) < r$. Hence g and h are in $\Gamma_s(x)$ with $s < c_n$. As $\Gamma_s(x)$ is elementary, g and h have the same fixed points by Theorems 5.5.3 and 5.5.6. Therefore $\Gamma_a = \Gamma_b$ by Theorem 5.5.4. □

Definition: Suppose that $0 < r \leq c_n$, where c_n is the Margulis constant. A component $V(\Gamma_a, r)$ of $V(\Gamma, r)$ is called a *Margulis region* for Γ based at the point a.

Parabolic Fixed Points

Lemma 4. *Let Γ be an elementary discrete subgroup of $\mathrm{M}(U^n)$ of parabolic type that fixes ∞, let Q be a Γ-invariant m-plane of E^{n-1} such that Q/Γ is compact, let P be the vertical $(m+1)$-plane of U^n above Q, and let $P_t = \{x \in P : x_n \geq t\}$. Then for each $r > 0$, there is a $t > 0$ such that*

$$N(P_t, r/3) \subset V(\Gamma, r).$$

Proof: Let $r > 0$. Since Q/Γ is compact, there is a compact Dirichlet polyhedron D for Γ in Q. From the proof of Lemma 2, we know that $V(\Gamma, r/3)$ intersects each vertical line of U^n in an open vertical ray ending at ∞, and if x is in $V(\Gamma, r/3)$, then $V(\Gamma, r/3)$ contains the ray $[x, \infty)$. Hence, since $V(\Gamma, r/3)$ is open and D is compact, there is a $t > 0$ such that

$$D \times \{t\} \subset V(\Gamma, r/3).$$

Since Γ leaves $\partial P_t = Q \times \{t\}$ invariant and $V(\Gamma, r/3)$ is Γ-invariant, we have that $\partial P_t \subset V(\Gamma, r/3)$. Therefore $P_t \subset V(\Gamma, r/3)$.

Now let x be an arbitrary point of $N(P_t, r/3)$. Then there is a point y of P_t such that $d(x, y) < r/3$. As y is in $V(\Gamma, r/3)$, there is a nonelliptic g in Γ such that $d(y, gy) < r/3$. Observe that

$$\begin{aligned} d(x, gx) &= d(x, y) + d(y, gy) + d(gy, gx) \\ &< r/3 + r/3 + r/3 \;\; = \;\; r. \end{aligned}$$

Therefore x is in $V(\Gamma, r)$. Thus

$$N(P_t, r/3) \subset V(\Gamma, r).$$

□

Lemma 5. *Let a be a conical limit point of a discrete subgroup Γ of $\mathrm{M}(B^n)$ and let R be a hyperbolic ray in B^n ending at a. Then for each $r > 0$, there is a point x of B^n and a sequence $\{g_i\}_{i=1}^\infty$ of distinct elements of Γ such that $\{g_i x\}_{i=1}^\infty$ converges to a within $N(R, r)$.*

Proof: By Theorem 12.2.3, there is a sequence $\{g_i\}_{i=1}^\infty$ of distinct elements of Γ and a compact subset K of B^n such that $\{g_i(0)\}_{i=1}^\infty$ converges to a and for all i, we have

$$K \cap g_i^{-1}R \neq \emptyset.$$

For each i, choose a point x_i on R such that $g_i^{-1}x_i$ is in K. As K is compact, the set $\{g_i^{-1}x_i\}$ has a limit point x in K. By passing to a subsequence, we may assume that $\{g_i^{-1}x_i\}_{i=1}^\infty$ converges to x within $B(x, r)$. As $g_i^{-1}x_i$ is in $B(x, r)$ for each i, we have that x_i is in $B(g_i x, r)$ for each i. Hence $g_i x$ is in $N(R, r)$ for each i. As $g_i(0) \to a$, we have that $g_i x \to a$ within $N(R, r)$. □

Theorem 12.5.3. *Let b be a point of S^{n-1} fixed by a parabolic element of a discrete subgroup Γ of $\mathrm{M}(B^n)$. Then b is not a conical limit point of Γ.*

Proof: We pass to the upper half-space model U^n and assume, without loss of generality, that $b = \infty$. Let $V(\Gamma_\infty, r)$ be a Margulis region for Γ based at ∞. Then by Lemma 4, there is a $t > 0$ such that

$$N(P_t, r/3) \subset V(\Gamma_\infty, r).$$

Let R be a vertical ray in P_t that ends at ∞. On the contrary, assume that ∞ is a conical limit point of Γ. Then by Lemma 5, there is a sequence $\{g_i\}_{i=1}^\infty$ of distinct elements of Γ and a point x of U^n such that $\{g_i x\}_{i=1}^\infty$ converges to ∞ within $N(R, r/3)$. Hence $g_i x \to \infty$ within $V(\Gamma_\infty, r)$. Now by Theorem 12.5.2, the set $V(\Gamma_\infty, r)$ is Γ_∞-invariant and is moved disjointly away from itself by elements of $\Gamma - \Gamma_\infty$. Therefore, the elements of $\{g_i x\}$ are translates of each other by elements of Γ_∞, and so all have the same nth coordinate and therefore lie in a compact subset of $N(R, r/3)$. Hence $\{g_i x\}$ cannot converge to ∞, which is a contradiction. Thus b is not a conical limit point of Γ. □

The next corollary follows immediately from Theorems 12.3.5 and 12.5.3.

Corollary 1. *Every point fixed by a parabolic element of a geometrically finite discrete subgroup Γ of $\mathrm{M}(B^n)$ is a cusped limit point of Γ.*

The Thick and Thin Parts of a Hyperbolic Space-Form

Let $M = B^n/\Gamma$ be a hyperbolic space-form and let $r > 0$. The *r-thin part* of M is the set

$$V(M, r) = V(\Gamma, r)/\Gamma.$$

The r-thin part of M is an open subset of M. Let

$$T(\Gamma, r) = B^n - V(\Gamma, r).$$

The r-*thick part* of M is the set

$$T(M, r) = T(\Gamma, r)/\Gamma.$$

The r-thick part of M is a closed subset of M whose complement is the r-thin part of M.

Theorem 12.5.4. *For each dimension n, there is a $\delta > 0$ such that for each hyperbolic space-form B^n/Γ, there is a point x of B^n such that the quotient map $\pi : B^n \to B^n/\Gamma$ maps $B(x, \delta)$ isometrically onto $B(\pi(x), \delta)$.*

Proof: Let c_n be the Margulis constant. By Theorem 12.5.2, the set $T(\Gamma, c_n)$ is nonempty. Let x be any point of $T(\Gamma, c_n)$. Then $d(x, gx) \geq c_n$ for every $g \neq 1$ in Γ. Then for every $g \neq 1$ in Γ, we have

$$B(x, c_n/2) \cap gB(x, c_n/2) = \emptyset.$$

Hence π maps $B(x, c_n/2)$ bijectively onto $B(\pi(x), c_n/2)$. Therefore, by the triangle inequality, π maps $B(x, c_n/4)$ isometrically onto $B(\pi(x), c_n/4)$. $\square$

Corollary 2. *For each dimension n, there is a positive lower bound for the set of volumes of complete hyperbolic n-manifolds.*

Exercise 12.5

1. Let Γ be an elementary discrete subgroup of $\mathrm{M}(B^n)$ all of whose nonidentity elements are parabolic translations. Prove that $V(\Gamma, r)$ is a horoball in B^n for each $r > 0$.

2. Let Γ be an elementary discrete subgroup of $\mathrm{M}(B^n)$ generated by a hyperbolic translation h of B^n with axis L and translation length ℓ on L. Prove that for each $r > \ell$, there is an $s > 0$ such that $V(\Gamma, r) = N(L, s)$.

3. Let Γ be an elementary discrete subgroup of $\mathrm{M}(B^3)$ generated by a hyperbolic element h with axis L and translation length ℓ on L. Prove that for each $r > \ell$, there is an $s > 0$ such that $V(\Gamma, r) = N(L, s)$.

4. Let Γ be the Poincaré extension of the Klein bottle group generated by the translation $\tau(z) = z + 1$ and the glide reflection $\rho(z) = -\overline{z} + 1 + i$. Describe the subset $V(\Gamma, r)$ of U^3 for each $r > 0$.

5. Let Γ be an infinite, elementary, discrete subgroup of $\mathrm{M}(B^n)$. Prove that $\overline{V}(\Gamma, r) \cap S^{n-1} = L(\Gamma)$ for each $r > 0$.

6. Prove that a geometrically finite discrete subgroup Γ of $\mathrm{M}(B^n)$ has only finitely many conjugacy classes of parabolic elements.

§12.6. Geometrically Finite Manifolds

In this section, we study the geometry of geometrically finite hyperbolic manifolds.

Lemma 1. *If $\triangle(x,y,z)$ is a generalized hyperbolic triangle whose angles at x and y are greater than $\pi/4$, then $d(x,y) < \cosh^{-1}(3)$.*

Proof: Let α, β, γ be the angles of $\triangle(x,y,z)$ at x, y, z, respectively, and let $c = d(x,y)$. Then by Theorems 3.5.4 and 3.5.6, we have

$$\cosh c = \frac{\cos\alpha\cos\beta + \cos\gamma}{\sin\alpha\sin\beta}.$$

As $\alpha, \beta > \pi/4$ and $\alpha + \beta + \gamma < \pi$, we have that $\alpha, \beta < 3\pi/4$. Hence $\sin\alpha$, $\sin\beta > 1/\sqrt{2}$ and $|\cos\alpha|$, $|\cos\beta| < 1/\sqrt{2}$. Therefore

$$\cosh c < \frac{(1/\sqrt{2})^2 + 1}{(1/\sqrt{2})^2} = 3. \qquad \square$$

Lemma 2. *If $\triangle(x,y,z)$ is a hyperbolic triangle, with $d(y,z) \geq 2d(x,y)$, then the angle of $\triangle(x,y,z)$ at z is less than $\pi/4$.*

Proof: Let α, β, γ be the angles of $\triangle(x,y,z)$ at x, y, z, respectively, and let a, b, c be the lengths of the opposite sides. Then we have that $a \geq 2c$. Now as

$$d(y,z) \leq d(y,x) + d(x,z),$$

we find that

$$d(x,z) \geq d(y,z) - d(x,y) \geq d(x,y).$$

Therefore $b \geq c$.

On the contrary, assume that $\gamma \geq \pi/4$. By the law of sines, we have

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

As $b \geq c$, we have that $\sin\beta \geq \sin\gamma$. Assume first that $\gamma \geq \pi/2$. Then $\beta \geq \pi - \gamma$. As $\alpha + \beta + \gamma < \pi$, we have a contradiction. Therefore $\gamma < \pi/2$, and so $\beta \geq \gamma$.

Now as $a \geq 2c$, we have

$$\sinh a \geq \sinh 2c = 2\sinh c\cosh c \geq 2\sinh c.$$

Therefore

$$\sin\alpha \geq 2\sin\gamma \geq 2\sin\gamma\cos\gamma = \sin 2\gamma.$$

As $\gamma \geq \pi/4$, we have that $2\gamma \geq \pi/2$. Hence $\alpha \geq \pi - 2\gamma$. Therefore

$$\alpha + \beta + \gamma \geq \pi - \gamma + \beta \geq \pi,$$

which is a contradiction. It follows that $\gamma < \pi/4$. $\square$

Definition: Two subsets A and B of B^n are said to be *r-near* for some $r > 0$ if and only if $A \subset N(B, r)$ and $B \subset N(A, r)$.

Let K be a closed, nonempty, hyperbolic convex subset of $\overline{B}^n$ and let
$$\rho_K : \overline{B}^n \to K$$
be the nearest point retraction.

Lemma 3. *For each $r > 0$, there is an $s > 0$ such that if K and L are closed, nonempty, convex, r-near subsets of B^n, then for all x in $\overline{B}^n$,*
$$d(\rho_K(x),\ \rho_L(x)) < s.$$

Proof: Set
$$s = \max\{2r,\ \cosh^{-1}(3)\}.$$
Let x be a point of $\overline{B}^n$, let $y = \rho_K(x)$, and let $z = \rho_L(x)$. If $d(y, z) < 2r$, then $d(y, z) < s$, so assume that $d(y, z) \geq 2r$. Then $x \neq y$, since if $x = y$, then x is in K and $d(x, z) < r$. Likewise $x \neq z$. Hence, the points x, y, z are distinct.

Now since z is in L and $L \subset N(K, r)$, there is a point w in $K \cap B(z, r)$. As $d(y, z) > r$, we have that $w \neq y$. As K is convex, the geodesic segment $[y, w]$ lies in K. Since y is the nearest point of K to x, the angle between $[x, y]$ and $[y, w]$ is at least $\pi/2$. As $d(z, w) < r$ and $d(y, z) \geq 2r$, the angle between $[y, z]$ and $[y, w]$ is less than $\pi/4$ by Lemma 2. Without loss of generality, we may assume that $y = 0$. Then by Theorem 2.1.2, we have
$$\theta(x, w) \leq \theta(x, z) + \theta(z, w).$$
Hence
$$\theta(x, z) \geq \theta(x, w) - \theta(z, w) > \pi/2 - \pi/4 = \pi/4.$$
Therefore, the angle between $[x, y]$ and $[y, z]$ is greater than $\pi/4$. Likewise, the angle between $[y, z]$ and $[z, x]$ is greater than $\pi/4$. Therefore
$$d(y, z) < \cosh^{-1}(3) \leq s$$
by Lemma 1. □

Lemma 4. *Let K and L be closed, nonempty, hyperbolic convex subsets of $\overline{B}^n$ and let C be a closed convex subset of B^n such that $K \cap C$ and $L \cap C$ are r-near. Let s be as in Lemma 3 and let B be a subset of C such that $N(B, s) \subset C$. Then*
$$\rho_K^{-1}(B) \subset \rho_L^{-1}(C).$$

Proof: Let x be a point of $\rho_K^{-1}(B)$. Then $\rho_K(x)$ is in $K \cap B$. Therefore
$$\rho_{K\cap C}(x) = \rho_K(x).$$
By Lemma 3, we have
$$d(\rho_{K\cap C}(x), \rho_{L\cap C}(x)) < s.$$

As $N(B,s) \subset C$, we deduce that $\rho_{L\cap C}(x)$ is in C°. We next show that

$$\rho_{L\cap C}(x) = \rho_L(x).$$

On the contrary, suppose that $\rho_{L\cap C}(x) = y$ and $\rho_L(x) = z$ with $y \neq z$. Then z is nearer to x than y. As L is convex, the geodesic segment $[y,z]$ lies in L. After positioning y at the origin, we see that every point on the open segment (y,z) is nearer to x than y. But (y,z) meets C° contrary to the fact that y is the nearest point of $L\cap C$ to x. Therefore, we have $\rho_{L\cap C}(x) = \rho_L(x)$. Hence $\rho_L(x)$ is in C. Thus $\rho_K^{-1}(B) \subset \rho_L^{-1}(C)$. □

Theorem 12.6.1. *Let Γ be a discrete subgroup of $\mathrm{M}(U^n)$ such that ∞ is fixed by a parabolic element of Γ. Let Q be a Γ_∞-invariant m-plane of E^{n-1} such that Q/Γ_∞ is compact. Then ∞ is a cusped limit point of Γ if and only if there is a $t > 0$ such that*

$$L(\Gamma) \subset \overline{N}(Q,t).$$

Proof: This is clear if Γ is elementary, so assume that Γ is nonelementary. If ∞ is a cusped limit point of Γ, then by Lemma 1 of §12.2, there is a $t > 0$ such that

$$L(\Gamma) \subset \overline{N}(Q,t).$$

Conversely, suppose that there is a $t > 0$ such that $L(\Gamma) \subset \overline{N}(Q,t)$. Let $V = V(\Gamma_\infty, r)$ be a Margulis region for Γ based at ∞. Then for each g in Γ, either $V \cap gV = \emptyset$ or $gV = V$ and $g(\infty) = \infty$. Let $\nu : U^n \to E^{n-1}$ be the vertical projection and let K be the closure of $\nu^{-1}(Q)$ in $\overline{U}^n$. Let L be the hyperbolic convex hull of $L(\Gamma)$ and let R be the closure of $\nu^{-1}(\overline{N}(Q,t))$ in $\overline{U}^n$. Then K, L, R are closed hyperbolic convex subsets of $\overline{U}^n$. Since Q/Γ_∞ is compact, there is a closed horoball C based at ∞ such that $R \cap C \subset V$. As R contains $L(\Gamma)$, we have that $L \subset R$. Consequently, there is a $r > 0$ such that $K\cap C$ and $L \cap C$ are r-near.

Let s be as in Lemma 4 and let B be the horoball contained in C such that ∂B is at a distance s from ∂C. Then $N(B,s) \subset C$. By Lemma 4, we have that $\rho_K^{-1}(B) \subset \rho_L^{-1}(C)$. Now as

$$L \cap C \subset R \cap C \subset V,$$

we find that $\rho_L^{-1}(C) \subset \rho_L^{-1}(V)$. Therefore, we have that $\rho_K^{-1}(B) \subset \rho_L^{-1}(V)$.

Now observe that the set $\rho_K^{-1}(B)$ has the shape of a cusped region for Γ and for each g in Γ, we have

$$g\rho_L^{-1}(V) = \rho_L^{-1}(gV).$$

Consequently, for each g in $\Gamma - \Gamma_\infty$, we have

$$\rho_L^{-1}(V) \cap g\rho_L^{-1}(V) = \rho_L^{-1}(V \cap gV) = \emptyset.$$

Hence $\rho_K^{-1}(B)$ is a cusped region for Γ. Thus ∞ is a cusped limit point of the group Γ. □

Corollary 1. *If Γ is a discrete subgroup of* $\mathrm{M}(U^n)$ *such that ∞ is fixed by a parabolic element of Γ and E^{n-1}/Γ_∞ is compact, then ∞ is a cusped limit point of Γ.*

Lemma 5. *If Γ is an elementary discrete subgroup of* $\mathrm{M}(U^n)$ *of parabolic type that fixes ∞, then for each point w of E^{n-1} and $r > 0$ there is a unique point x of U^n directly above w such that x is in $\partial V(\Gamma, r)$. Moreover,*

$$\partial V(\Gamma, r) \subset \{x \in U^n : d(x, gx) = r \text{ for some parabolic } g \text{ in } \Gamma\}.$$

Proof: Let w be a point of E^{n-1}, let x a point directly above w, and let f be a parabolic element of Γ. By Theorem 4.6.1, we have

$$\cosh d(x, fx) = 1 + \frac{|x - fx|^2}{2x_n^2}.$$

By increasing the value of x_n, if necessary, we may assume that we have $d(x, fx) \leq r$. Now there are only finitely many parabolic elements g of Γ such that

$$C(x, r/2) \cap gC(x, r/2) \neq \emptyset,$$

since Γ is discontinuous. Hence, there are only finitely many parabolic elements g of Γ such that $d(x, gx) \leq r$. By replacing f with another parabolic element of Γ, if necessary, we may assume that

$$d(x, fx) \leq d(x, gx)$$

for all parabolic elements g of Γ. By increasing the value of x_n, if necessary, we may assume that $d(x, fx) = r$ and $d(x, gx) \geq r$ for all parabolic elements of Γ. If y is a point of U^n directly above x, then $d(y, fy) < r$. Therefore x is in $\partial V(\Gamma, r)$.

Now suppose that x is a point on $\partial V(\Gamma, r)$. Then there is a sequence of points $\{y_i\}_{i=1}^\infty$ of $V(\Gamma, r)$ converging to x within $B(x, r/2)$. Now for each i, there is a parabolic element g_i of Γ such that $d(y_i, g_i y_i) < r$. Observe that

$$\begin{aligned} d(x, g_i x) &\leq d(x, y_i) + d(y_i, g_i y_i) + d(g_i y_i, g_i x) \\ &< (r/2) + r + (r/2) \\ &= 2r. \end{aligned}$$

Now as Γ is discontinuous, there are only finitely many elements g of Γ such that

$$C(x, r) \cap gC(x, r) \neq \emptyset.$$

Consequently, the sequence $\{g_i\}_{i=1}^\infty$ can take on only finitely many values. By passing to a subsequence, we may assume that $g_i = g$ for all i. As $d(y_i, gy_i) < r$, we have by continuity that $d(x, gx) \leq r$. But x is not in $V(\Gamma, r)$, and so $d(x, gx) = r$. If y is a point of U^n directly above x, then x is in $V(\Gamma, r)$. Therefore x is the only point in $\partial V(\Gamma, r)$ directly above w. □

Geometrically Finite Hyperbolic Manifolds

Definition: A hyperbolic n-manifold M is *geometrically finite* if and only if M has a finite number of connected components and each component of M is isometric to a space-form B^n/Γ such that Γ is geometrically finite.

Remark: It follows from Theorem 8.1.5 that a hyperbolic space-form B^n/Γ is geometrically finite if and only if Γ is geometrically finite.

Let $M = B^n/\Gamma$ be a hyperbolic space-form and let $C(\Gamma)$ be the hyperbolic convex hull of the limit set of Γ. Then $C(\Gamma) \cap B^n$ is a closed, convex, Γ-invariant subset of B^n. The *convex core* of M is the set

$$C(M) = (C(\Gamma) \cap B^n)/\Gamma.$$

The convex core $C(M)$ is a geodesically connected closed subset of M. It is an exercise to prove that $C(M)$ is a deformation retract of M when M in nonelementary.

Theorem 12.6.2. *Let $M = B^n/\Gamma$ be a hyperbolic space-form. Then the following are equivalent:*

(1) *The hyperbolic manifold M is geometrically finite.*

(2) *The open set $N(C(M), r)$ has finite volume for each $r > 0$.*

(3) *The closed set $C(M) \cap T(M, r)$ is compact for each $r > 0$.*

Proof: Suppose that M is geometrically finite. Then Γ is geometrically finite. Hence Γ has a geometrically finite, exact, convex, fundamental polyhedron P. Set

$$B(\Gamma) = C(\Gamma) \cap B^n.$$

We shall prove that $N(C(M), r)$ has finite volume by proving that the set $N(B(\Gamma), r) \cap P$ has finite volume. Now we have

$$\overline{N}(B(\Gamma), r) \cap S^{n-1} = L(\Gamma).$$

Hence

$$\overline{N}(B(\Gamma), r) \cap \overline{P} \cap S^{n-1} = L(\Gamma) \cap \overline{P}.$$

By Theorems 12.2.7, 12.3.1, and 12.3.4, we have that $L(\Gamma) \cap \overline{P}$ is a finite set of cusped limit points of Γ, say $c_1, \ldots, c_k$. For each i, let U_i be a cusped region for Γ based at c_i. By Lemma 2 of §12.2, we have that

$$\big(N(B(\Gamma), r) \cap P\big) - \bigcup_{i=1}^{k} U_i$$

is a bounded subset of B^n, and so it has finite volume. Hence, to prove that the set $N(B(\Gamma), r) \cap P$ has finite volume, it suffices to show that the set $N(B(\Gamma), r) \cap P \cap U_i$ has finite volume for each i.

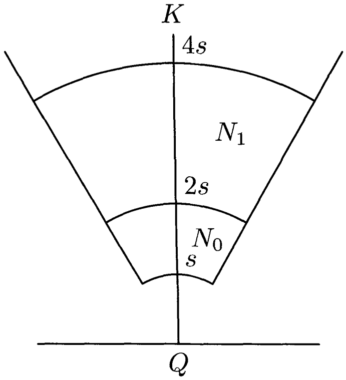


Figure 12.6.1. The subdivision of $N(K,s) \cap \rho_K^{-1}(\nu^{-1}(C) \cap B(s))$

We now pass to the upper half-space model U^n and conjugate Γ so that $c_1 = \infty$. Then there is a Γ_∞-invariant m-plane Q of E^{n-1} and an $s > 0$ such that $U_1 = U(Q,s)$. By Lemma 1 of §12.2, we have that

$$L(\Gamma) \subset \overline{N}(Q,s).$$

Let $\nu : U^n \to E^{n-1}$ be the vertical projection and let R be the closure in $\overline{U}^n$ of $\nu^{-1}(\overline{N}(Q,s))$. Then R is a closed hyperbolic convex subset of $\overline{U}^n$ containing $L(\Gamma)$. Therefore $C(\Gamma) \subset R$. Let $K = \nu^{-1}(Q)$. Then by increasing s, if necessary, we may assume that

$$N(B(\Gamma), r) \cap U_1 \subset N(K,s) \cap U_1.$$

Let D be a Dirichlet polyhedron for Γ_∞ in Q and let

$$B(t) = \{x \in U^n : x_n > t\}.$$

Observe that the set

$$N(K,s) \cap \rho_K^{-1}(\nu^{-1}(D^\circ) \cap B(s))$$

is a fundamental domain for Γ_∞ in $N(K,s) \cap U_1$. We now show that

$$\mathrm{Vol}(N(K,s) \cap \rho_K^{-1}(\nu^{-1}(D) \cap B(s))) < \infty.$$

As Q/Γ_∞ is compact, D is compact. Hence, there is an m-cube C in Q containing D. By conjugating Γ, we may assume that $Q = E^m$. Then $K = E^{m+1}$. Let $\mu : E^n \to E^n$ be defined by $\mu(x) = 2x$. Then μ is an isometry of U^n that leaves K invariant. For each $i = 0, 1, 2, \ldots$, let

$$N_i = N(K,s) \cap \rho_K^{-1}(\nu^{-1}(C) \cap (B(2^i s) - B(2^{i+1}s))).$$

Observe that N_0 is bounded, and so it has finite volume. See Figure 12.6.1. Since $\mu(C)$ can be subdivided into 2^m cubes congruent to C, we deduce that $\mu(N_i)$ can be subdivided into 2^m regions each congruent to N_{i+1}. Therefore

$$\mathrm{Vol}(N_{i+1}) = \frac{1}{2^m}\mathrm{Vol}(N_i).$$

Hence, by induction, we have

$$\mathrm{Vol}(N_i) = \left(\frac{1}{2^m}\right)^i \mathrm{Vol}(N_0).$$

Therefore

$$\begin{aligned}\mathrm{Vol}\Big(\bigcup_{i=0}^{\infty} N_i\Big) &= \sum_{i=0}^{\infty} \mathrm{Vol}(N_i) \\ &= \mathrm{Vol}(N_0)\sum_{i=0}^{\infty}\left(\frac{1}{2^m}\right)^i \\ &= \mathrm{Vol}(N_0)\left(\frac{2^m}{2^m-1}\right) < \infty.\end{aligned}$$

Hence

$$\mathrm{Vol}(N(K,s)\cap \rho_K^{-1}(\nu^{-1}(C)\cap B(s))) < \infty.$$

As $D \subset C$, we have that

$$\mathrm{Vol}(N(K,s)\cap \rho_K^{-1}(\nu^{-1}(D)\cap B(s))) < \infty.$$

Therefore, we have that

$$\mathrm{Vol}((N(K,s)\cap U_1)/\Gamma_\infty) < \infty.$$

As $N(B(\Gamma),r)\cap U_1 \subset N(K,s)\cap U_1$, we have that

$$\mathrm{Vol}((N(B(\Gamma),r)\cap U_1)/\Gamma_\infty) < \infty.$$

Since U_1 is a cusped region for Γ based at ∞, we deduce that

$$\mathrm{Vol}(N(B(\Gamma),r)\cap P \cap U_1) < \infty.$$

Likewise, we have that

$$\mathrm{Vol}(N(B(\Gamma),r)\cap P \cap U_i) < \infty$$

for each $i > 1$. Hence

$$\mathrm{Vol}(N(B(\Gamma),r)\cap P) < \infty.$$

Therefore, we have that

$$\mathrm{Vol}(N(C(M),r)) < \infty.$$

Thus (1) implies (2).

Now assume that $N(C(M),r)$ has finite volume for each $r > 0$. On the contrary, suppose that $C(M)\cap T(M,r)$ is not compact for some $r > 0$. Choose a sequence of points $\{u_i\}_{i=1}^{\infty}$ of $C(M)\cap T(M,r)$ inductively as follows: Let u_1 be any point of $C(M)\cap T(M,r)$. Assume that $u_1,\ldots,u_m$ have been chosen so that the balls $\{B(u_i,r/2)\}_{i=1}^{m}$ are mutually disjoint. Since the set $\bigcup_{i=1}^{m} C(u_i,r)$ is compact, it cannot contain $C(M)\cap T(M,r)$. Hence, there is a point u_{m+1} of $C(M)\cap T(M,r)$ such that the balls $\{B(u_i,r/2)\}_{i=1}^{m+1}$

are mutually disjoint. It follows by induction that there is a sequence $\{u_i\}_{i=1}^{\infty}$ of points of $C(M) \cap T(M,r)$ such that the balls $\{B(u_i, r/2)\}_{i=1}^{\infty}$ are mutually disjoint.

Now for each i, choose x_i in $C(\Gamma) \cap T(\Gamma, r)$ such that $\pi(x_i) = u_i$ where $\pi : B^n \to M$ is the quotient map. Now as x_i is in $T(\Gamma, r)$, we have that $d(x_i, gx_i) \geq r$ for all $g \neq 1$ in Γ. Hence, we have

$$B(x_i, r/2) \cap gB(x_i, r/2) = \emptyset$$

for all $g \neq 1$ in Γ. Consequently π maps $B(x_i, r/2)$ bijectively onto $B(u_i, r/2)$. Therefore

$$\mathrm{Vol}(B(u_i, r/2)) = \mathrm{Vol}(B(x_i, r/2)) = \mathrm{Vol}(B(0, r/2)).$$

Now as

$$B(u_i, r/2) \subset N(C(M), r/2)$$

for each i, we deduce that $N(C(M), r/2)$ has infinite volume, which is a contradiction. Therefore $C(M) \cap T(M,r)$ must be compact for all r. Thus (2) implies (3).

Now assume that $C(M) \cap T(M,r)$ is compact for each $r > 0$. We shall prove that Γ is geometrically finite by showing that every limit point of Γ is either conical or cusped. Let a be a limit point of Γ. If a is fixed by a hyperbolic element of Γ, then a is a conical limit point of Γ by Theorem 12.2.1.

Assume next that a is fixed by a parabolic element of Γ. We pass to the upper half-space model U^n and conjugate Γ so that $a = \infty$. Let Q be a Γ_∞-invariant m-plane of E^{n-1} such that Q/Γ_∞ is compact. We shall prove that a is a cusped limit point of Γ by showing that there is an $s > 0$ such that

$$L(\Gamma) \subset \overline{N}(Q, s).$$

On the contrary, suppose that there is no such s. Then there is a sequence $\{x_i\}_{i=1}^{\infty}$ of points of $L(\Gamma) - \{\infty\}$ such that $\mathrm{dist}_E(x_i, Q)$ goes to ∞ with i. Let $V(\Gamma_\infty, r)$ be a Margulis region for Γ based at ∞. Then for each i, there is a point y_i of U^n directly above x_i such that y_i is in $\partial V(\Gamma_\infty, r/2)$ by Lemma 5. Moreover y_i is in $V(\Gamma_\infty, r)$ for each i by Lemma 5. Furthermore y_i is in $C(\Gamma)$ for each i, since $C(\Gamma)$ is convex. Clearly $\mathrm{dist}_E(y_i, Q)$ goes to ∞ with i.

Let $\pi : U^n \to M$ be the quotient map. Then the sequence $\{\pi(y_i)\}_{i=1}^{\infty}$ has a limit point in the compact set $C(M) \cap \partial T(M, r/2)$. By passing to a subsequence, we may assume that $\{\pi(y_i)\}$ converges to a point w. Let z be a point of $C(\Gamma) \cap \partial T(\Gamma, r/2)$ such that $\pi(z) = w$. As $\pi(y_i) \to w$, there is a g_i in Γ such that $\{g_i y_i\}_{i=1}^{\infty}$ converges to z. Now z is in $V(\Gamma, r)$. Hence z is in $V(\Gamma_b, r)$ for some fixed point b of a nonidentity element of Γ by Lemma 3 of §12.5. As $g_i y_i \to z$, there is a j such that $g_j y_j$ is in $V(\Gamma_b, r)$. Now since y_j is in $V(\Gamma_\infty, r)$, we have that $g_j y_j$ is in $V(\Gamma_{g_j(\infty)}, r)$. By Theorems 5.5.4 and 12.5.2, we deduce that $g_j(\infty) = b$. Now by replacing z by $g_j^{-1} z$, we may assume that z is in $V(\Gamma_\infty, r)$; by passing to a subsequence, we may assume that $g_i y_i$ is in $V(\Gamma_\infty, r)$ for all i. Then g_i is in Γ_∞ for all i. Hence $\mathrm{dist}_E(g_i y_i, Q)$ goes to ∞ with i, and so $\{g_i y_i\}$ diverges to ∞, which is a contradiction. Thus, there is a $s > 0$ such that

$$L(\Gamma) \subset \overline{N}(Q, s).$$

Hence, by Theorem 12.6.1, we have that a is a cusped limit point of Γ.

Assume now that the point a is not fixed by a nonidentity element of Γ. Let R be a hyperbolic ray in B^n starting in $C(\Gamma)$ and ending at a. Then $R \subset C(\Gamma)$, since $C(\Gamma)$ is convex. Let r be the Margulis constant. Then no subray of R is contained in a component of $V(\Gamma, r)$, since otherwise its endpoint a would be fixed by a nonidentity element of Γ. Therefore, the set $R \cap T(\Gamma, r)$ is unbounded. Let $\{x_i\}_{i=1}^\infty$ be a sequence of points of $R \cap T(\Gamma, r)$ converging to a. Then the sequence $\{\pi(x_i)\}_{i=1}^\infty$ has a limit point in the compact set $C(M) \cap T(M, r)$. By passing to a subsequence, we may assume that $\{\pi(x_i)\}$ converges to a point u of M. Choose a point y of B^n such that $\pi(y) = u$. As $\pi(x_i) \to u$, there is an element g_i of Γ such that the sequence $\{g_i x_i\}_{i=1}^\infty$ converges to y. Since $x_i \to a$, infinitely many of the terms of $\{g_i\}$ are distinct. Hence, by passing to a subsequence, we may assume that the terms of $\{g_i\}$ are distinct. As $g_i x_i \to y$, there is an $s > 0$ such that

$$g_i R \cap C(y, s) \neq \emptyset$$

for all i. Therefore a is a conical limit point of Γ by Theorem 12.2.3. Thus, every limit point of Γ is either conical or cusped. Hence Γ is geometrically finite by Theorem 12.3.5. Thus (3) implies (1). □

Theorem 12.6.3. *Every complete hyperbolic n-manifold of finite volume is geometrically finite.*

Proof: Let M be a complete hyperbolic n-manifold of finite volume. By Theorem 12.5.4, there is a positive lower bound for the set of volumes of complete hyperbolic n-manifolds. Therefore M has a finite number of connected components. Thus, we may assume that M is connected. By Theorem 8.5.9, we may assume that M is a space-form B^n/Γ of finite volume. Then $C(M) = M$ by Theorem 12.1.15. Hence M is geometrically finite by Theorem 12.6.2. □

Theorem 12.6.4. *Let $M = B^n/\Gamma$ be a nonelementary geometrically finite space-form such that Γ leaves no m-plane of B^n invariant for $m < n-1$. Then the group $\mathrm{I}(M)$ of isometries of M is finite.*

Proof: An isometry ϕ of $M = B^n/\Gamma$ lifts to an isometry $\tilde{\phi}$ of B^n such that $\tilde{\phi}\Gamma\tilde{\phi}^{-1} = \Gamma$. Moreover $\tilde{\phi}$ is unique up to composition with an element of Γ. Conversely, if ψ is an isometry of B^n such that $\psi\Gamma\psi^{-1} = \Gamma$, then ψ induces an isometry of M. Let N be the normalizer of Γ in $\mathrm{M}(B^n)$. We conclude that $\mathrm{I}(M)$ is isomorphic to N/Γ.

The group Γ is finitely generated by Theorem 12.3.9. Therefore N is discrete by Theorem 12.1.17. Now by Theorem 12.1.16, we have that $L(\Gamma) = L(\mathrm{N})$. Therefore N leaves $L(\Gamma)$ invariant. Hence N also leaves invariant the set

$$B(\Gamma) = C(\Gamma) \cap B^n.$$

Therefore N leaves invariant the set $N(B(\Gamma), 1)$.

Since the set $N(B(\Gamma), 1)$ is open, there is a point x of $N(B(\Gamma), 1)$ that is not fixed by any $g \neq 1$ in N. Let D be the Dirichlet domain for N centered at x. Set

$$E = D \cap N(B(\Gamma), 1).$$

Then E is a fundamental domain for the action of N on $N(B(\Gamma), 1)$. Let $\{h_i\}$ be a set of Γ-coset representatives in N. Then

$$F = \cup h_i E$$

is a fundamental region for the action of Γ on $N(B(\Gamma), 1)$. Let $\partial_N F$ be the boundary of F in $N(B(\Gamma), 1)$. As D is a locally finite fundamental domain for N, we have

$$\partial_N F \subset \cup h_i \partial D.$$

Therefore, we have

$$\mathrm{Vol}(\partial_N F) = 0.$$

Hence, we have

$$\mathrm{Vol}(F) = \mathrm{Vol}(N(B(\Gamma), 1)/\Gamma).$$

By Theorem 12.6.2, we have that

$$\mathrm{Vol}(F) = \mathrm{Vol}(N(C(M), 1)) < \infty.$$

Now since

$$[\mathrm{N} : \Gamma] = \mathrm{Vol}(F)/\mathrm{Vol}(E),$$

we deduce that N/Γ is finite. Therefore $\mathrm{I}(M)$ is finite. □

Corollary 2. *Every complete hyperbolic n-manifold of finite volume, with $n > 1$, has a finite group of isometries.*

Proof: Let M be a complete hyperbolic n-manifold of finite volume. Then M has a finite number of connected components. Therefore, we may assume that M is connected. By Theorem 8.5.9, we may assume that M is a space-form B^n/Γ of finite volume. The group Γ is nonelementary, since every elementary hyperbolic space-form has infinite volume. By Theorem 12.6.3, the group Γ is geometrically finite. By Theorem 12.1.15, the group Γ is of the first kind. Therefore Γ leaves no proper m-plane of B^n invariant. Hence $\mathrm{I}(M)$ is finite by Theorem 12.6.4. □

The Ideal Boundary of a Hyperbolic Manifold

Let M be a complete, connected, hyperbolic n-manifold. Then there is a torsion-free discrete subgroup Γ of $\mathrm{M}(B^n)$ and an isometry $\xi : M \to B^n/\Gamma$. The orbit space $O(\Gamma)/\Gamma$ is called the *ideal boundary* of M. Let $\overline{M}$ be the union of M and its ideal boundary, and let

$$\overline{\xi} : \overline{M} \to (B^n \cup O(\Gamma))/\Gamma$$

be the extension of ξ that is the identity on $O(\Gamma)/\Gamma$. We topologize $\overline{M}$ so that $\overline{\xi}$ is a homeomorphism. We shall prove that $\overline{M}$ is an n-manifold-with-boundary. The first step is to prove that $\overline{M}$ is Hausdorff.

Lemma 6. *Let Γ be a group acting discontinuously on a locally compact Hausdorff space X. Then the orbit space X/Γ is Hausdorff.*

Proof: Let x and y be points of X such that Γx and Γy are disjoint. As X is locally compact, there are open neighborhoods U and V of x and y, respectively, such that $\overline{U}$ and $\overline{V}$ are compact and disjoint. Since $\{x\} \cup \overline{V}$ is compact, only finitely many elements of Γx meet $\overline{V}$. Hence $W = V - \Gamma x$ is an open neighborhood of y.

Let N be an open neighborhood of y such that $\overline{N} \subset W$. Then Γx and $\Gamma \overline{N}$ are disjoint, since Γx and $\overline{N}$ are disjoint. Now since $\overline{U} \cup \overline{N}$ is compact, at most finitely many Γ-images of $\overline{N}$ meet $\overline{U}$. Hence $M = \overline{U} - \Gamma\overline{N}$ is an open neighborhood of x. Moreover ΓM and ΓN are disjoint, since M and ΓN are disjoint. Therefore X/Γ is Hausdorff. □

Theorem 12.6.5. *Let Γ be a torsion-free discrete subgroup of $\mathrm{M}(B^n)$ of the second kind. Then the quotient map*

$$\pi : B^n \cup O(\Gamma) \to (B^n \cup O(\Gamma))/\Gamma$$

is a covering projection and the orbit space $(B^n \cup O(\Gamma))/\Gamma$ is an n-manifold-with-boundary $O(\Gamma)/\Gamma$.

Proof: Since Γ is torsion-free, it acts freely on $B^n \cup O(\Gamma)$ by Theorems 8.2.1 and 12.1.11. Therefore π is a covering projection by Theorems 8.1.3 and 12.1.11. Now by Lemma 6, the orbit space $(B^n \cup O(\Gamma))/\Gamma$ is Hausdorff. Hence $(B^n \cup O(\Gamma))/\Gamma$ is an n-manifold-with-boundary, since $B^n \cup O(\Gamma)$ is an n-manifold-with-boundary. The boundary of $(B^n \cup O(\Gamma))/\Gamma$ is $O(\Gamma)/\Gamma$.
□

Corollary 3. *Let M be a complete, connected, hyperbolic n-manifold, and let $\overline{M}$ be the union of M and its ideal boundary. Then $\overline{M}$ is an n-manifold-with-boundary.*

Cusps

Let Γ be a torsion-free, elementary, discrete subgroup of $\mathrm{M}(B^n)$ of parabolic type and let U be a cusped region in B^n for Γ. A hyperbolic n-manifold isometric to U/Γ is called an n-dimensional *cusp*.

Lemma 7. *Let Γ be a discrete subgroup of $\mathrm{M}(B^n)$ and let C be the set of cusped limit points of Γ. Then for each point c in C, there is a cusped region $U(c)$ based at c for Γ such that the regions $\{\overline{U}(c) : c \in C\}$ are mutually disjoint and $gU(c) = U(gc)$ for each g in Γ and c in C.*

Proof: This is clear if Γ is elementary, so assume that Γ is nonelementary. For each c in C, let $V(c) = V(\Gamma_c, r)$ be a Margulis region for Γ based at c. Then the regions $\{V(c) : c \in C\}$ are mutually disjoint and $gV(c) = V(gc)$ for each g in Γ and c in C. Let $\rho : \overline{B}^n \to C(\Gamma)$ be the nearest point retraction. Then the regions $\{\rho^{-1}(V(c)) : c \in C\}$ are mutually disjoint. As in the proof of Theorem 12.6.1, there is a cusped region $U(c)$ for Γ based at c such that $\overline{U}(c) \subset \rho^{-1}(V(c))$ for each c in C. Then the regions $\{\overline{U}(c) : c \in C\}$ are mutually disjoint. Now as ρ is Γ-equivariant, we have

$$g\rho^{-1}(V(c)) = \rho^{-1}(V(gc))$$

for each g in Γ and c in C. Consequently, we can choose $U(c)$ so that $gU(c) = U(gc)$ for each g in Γ and c in C. □

Theorem 12.6.6. *Let M be a connected, geometrically finite, hyperbolic n-manifold and let $\overline{M}$ be the union of M and its ideal boundary. Then there is a compact connected n-manifold-with-boundary $\overline{M}_0$ in $\overline{M}$ such that $\overline{M} - \overline{M}_0$ is the disjoint union of a finite number of cusps.*

Proof: Since M is geometrically finite, we may assume that M is a space-form B^n/Γ and

$$\overline{M} = (B^n \cup O(\Gamma))/\Gamma$$

with Γ geometrically finite. Let C be the set of cusped limit points of Γ. By Lemma 7, there is a cusped region $U(c)$ based at c for Γ for each c in C such that the regions $\{\overline{U}(c) : c \in C\}$ are mutually disjoint and $gU(c) = U(gc)$ for each g in Γ and c in C. Let

$$(B^n \cup O(\Gamma))_0 = (B^n \cup O(\Gamma)) - \bigcup_{c \in C} U(c).$$

It is clear from the geometry of a cusped region that there is a continuous retraction

$$\rho : B^n \cup O(\Gamma) \to (B^n \cup O(\Gamma))_0.$$

Hence $(B^n \cup O(\Gamma))_0$ is a closed, connected, Γ-invariant subset of $B^n \cup O(\Gamma)$. Moreover $(B^n \cup O(\Gamma))_0$ is an n-manifold-with-boundary. Let

$$\pi : B^n \cup O(\Gamma) \to \overline{M}$$

be the quotient map and set

$$\overline{M}_0 = \pi((B^n \cup O(\Gamma))_0).$$

Then $\overline{M}_0$ is a closed connected subset of $\overline{M}$. Moreover $\overline{M}_0$ is an n-manifold-with-boundary.

Let P be an exact, convex, fundamental polyhedron for Γ. Then P is geometrically finite. Hence P has only finitely many cusp points that are cusped limit points of Γ, say $c_1, \ldots, c_m$. It follows from Theorems 12.2.6 and 12.2.7 that for each c in C, there is a g in Γ such that $gc = c_i$ for some i. Therefore C is partitioned into only finitely many Γ-orbits. Consequently $\overline{M} - \overline{M}_0$ has only finitely many components and for each component K of $\overline{M} - \overline{M}_0$ there is an injective local isometry

$$\iota : U(c)/\Gamma_c \to K$$

for some c in $\{c_1, \ldots, c_m\}$. By shrinking the regions $\{U(c_i)\}$, if necessary, we may assume that ι is an isometry, whence K is a cusp.

Now let

$$\overline{P}_0 = \overline{P} - \bigcup_{i=1}^{m} \big(\{c_i\} \cup U(c_i)\big).$$

Then $\overline{P}_0$ is a closed subset of $\overline{B}^n$ by Lemma 2 of §12.2. Therefore $\overline{P}_0$ is compact. By Theorem 12.3.4, we have that $\overline{P}_0$ is a subset of $B^n \cup O(\Gamma)$. Hence $\pi(\overline{P}_0)$ is compact. Now as $\overline{M}_0 \subset \pi(\overline{P}_0)$, we deduce that $\overline{M}_0$ is compact. □

The next corollary gives the global geometry of a complete hyperbolic n-manifold of finite volume.

Corollary 4. *Let M be a complete hyperbolic n-manifold of finite volume. Then there is a compact n-manifold-with-boundary M_0 in M such that $M - M_0$ is the disjoint union of a finite number of cusps.*

Proof: The manifold M has a finite number of connected components by Theorem 12.6.3. Thus, we may assume that M is connected. By Theorem 12.6.6, there is a compact connected n-manifold-with-boundary M_0 in M such that $M - M_0$ is the disjoint union of a finite number of cusps, since $\overline{M} = M$. □

Exercise 12.6

1. Fill in the details of the proof of Theorem 12.6.1 of the existence of the closed horoball C.
2. Fill in the details of the proof of Theorem 12.6.1 that the sets $K \cap C$ and $L \cap C$ are r-near.
3. Let M be a nonelementary hyperbolic space-form. Prove that $C(M)$ is a deformation retract of M.

4. Let M be a geometrically finite hyperbolic space-form. Prove that the infimum of the set of lengths of simple closed geodesics of M is positive.

5. Let H^n/Γ be a geometrically finite space-form. Prove that for all sufficiently small values of r, we have

$$V(\Gamma, r) = \cup\{V(\Gamma_a, r) : a \text{ is a fixed point of a parabolic element of } \Gamma\}$$

6. Let M be a geometrically finite hyperbolic space-form. Prove that $T(M, r)$ is a deformation retract of M for all sufficiently small values of r.

7. Let M be a geometrically finite hyperbolic space-form. Prove that the set $C(M) \cap T(M, r)$ is a deformation retract of M for all sufficiently small values of r.

8. Fill in the details in the proof of Theorem 12.6.6 of the existence and continuity of the retraction ρ.

9. Prove that the set of similarity types of the links of the cusp points of a complete hyperbolic n-manifold M of finite volume is an isometry invariant of M.

§12.7. Historical Notes

§12.1. Poincaré introduced the limit set of a discrete group of linear fractional transformations of the unit disk in his 1882 paper *Sur les fonctions fuchsiennes* [331]. Theorems 12.1.1, 12.1.3, 12.1.9, 12.1.13, and 12.1.18 appeared in Vol. I of Fricke and Klein's 1897 *Vorlesungen über die Theorie der automorphen Functionen* [139]. Theorems 12.1.2 and 12.1.12 appeared in Fubini's 1908 treatise *Introduzione alla teoria dei gruppi discontinui e delle funzioni automorfe* [144]. Theorem 12.1.4 appeared in Ford's 1927 paper *On the foundations of the theory of discontinuous groups of linear transformations* [135]. Theorem 12.1.5 appeared in Vol. II of Appell, Goursat, and Fatou's 1930 treatise *Théorie des Fonctions Algébriques* [22]. The 3-dimensional version of Theorem 12.1.6 was proved by van Vleck in his 1919 paper *On the combination of non-loxodromic substitutions* [394]. Theorem 12.1.6 appeared in Apanasov's 1975 paper *Kleinian groups in space* [16]. Theorem 12.1.7 appeared in Lehner's 1964 survey *Discontinuous Groups and Automorphic Functions* [255]. Theorems 12.1.8, 12.1.14, and 12.1.15 appeared in Poincaré's 1882 paper [331]. Theorem 12.1.10 appeared in Poincaré's 1883 *Mémoir sur les groupes kleinéens* [332]. The convex hull of the limit set of a torsion-free discrete group of Möbius transformations of the unit disk was introduced by Koebe in his 1928 paper *Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen III* [243]. See also Nielsen's 1940 paper *Über Gruppen linearer Transformationen* [320]. As a reference for nearest point retractions onto convex sets, see Bishop and O'Neill's 1969 paper *Manifolds of negative curvature* [49].

The 2-dimensional version of Theorem 12.1.16 appeared in Greenberg's 1960 paper *Discrete groups of motions* [163], and the n-dimensional version appeared in Chen and Greenberg's 1974 paper *Hyperbolic spaces* [80]. Theorem 12.1.17 for closed surface groups was proved by Poincaré in his 1885 paper *Sur un théorème de M. Fuchs* [334]. Theorem 12.1.17 is a consequence of a general result in Wang's 1967 paper *On a maximality property of discrete subgroups with fundamental domain of finite measure* [401]. Two-dimensional Schottky groups were introduced by Schottky in his 1877 paper *Ueber die conforme Abbildung mehrfach zusammenhängender ebener Flächen* [367]. The 2-dimensional versions of Theorems 12.1.19 and 12.1.20 are consequences of general results in Poincaré's 1882 paper [331]. Limit sets of 3-dimensional Schottky groups were considered by Poincaré in his 1883 memoir [332]. Theorem 12.1.21 essentially appeared in Fricke's 1894 paper *Die Kreisbogenvierseite und das Princip der Symmetrie* [138]. See also Vol. I of Fricke and Klein's 1897 treatise [139]. For a discussion of the fractal nature of limit sets of Schottky groups, see Mandelbrot's 1983 paper *Self-inverse fractals osculated by sigma-discs and the limit sets of inversion groups* [273].

§12.2. Conical limit points of Fuchsian groups were introduced by Hedlund in his 1936 paper *Fuchsian groups and transitive horocycles* [187]. In particular, Theorem 12.2.1 for Fuchsian groups of the first kind appeared in this paper. A conical limit point is also called a *point of approximation*. The 3-dimensional versions of Theorems 12.2.1-12.2.5 were proved by Beardon and Maskit in their 1974 paper *Limit points of Kleinian groups and finite-sided fundamental polyhedra* [35]. Corollary 1 appeared in Vol. I of Fricke and Klein's 1897 treatise [139]. Cusped limit points in dimension three were introduced by Beardon and Maskit in their 1974 paper [35]. Theorems 12.2.6 and 12.2.7 for Fuchsian groups were proved by Klein in his 1883 paper *Neue Beiträge zur Riemannschen Functionentheorie* [233]. The 3-dimensional versions of Theorems 12.2.6 and 12.2.7 for rank two parabolic fixed points appeared in Vol. I of Fricke and Klein 1897 treatise [139]. Theorems 12.2.6 and 12.2.7 for dimension $n > 3$ seem to be new. Corollary 2 was proved by Beardon and Maskit in their 1974 paper [35]. As references for the theory of limit sets, see Nicholls' 1988 survey article *The limit set of a discrete group of hyperbolic motions* [316] and his 1989 treatise *The Ergodic Theory of Discrete Groups* [317].

§12.3. The concept of a geometrically finite convex polyhedron is new. Theorems 12.3.1 and 12.3.2 are new. The 3-dimensional versions of Theorems 12.3.3-12.3.9 were proved by Beardon and Maskit in their 1974 paper [35]. Theorem 12.3.4 for finite-sided polyhedra appeared in Apanasov's 1982 paper *Geometrically finite groups of transformations of space* [18]. Theorems 12.3.3-12.3.9 for dimension $n > 3$ seem to be new.

§12.4. Lemma 1 was proved by Frobenius in his 1911 paper *Über den von L. Bieberbach gefundenen Beweis eines Satzes von C. Jordan* [142]. Lemmas 2, 3, 5, and Theorem 12.4.1 were proved by Bowditch in his 1993 paper *Geometrical finiteness for hyperbolic groups* [55]. Lemma 4 appeared in Beardon and Wilker's 1984 paper *The norm of a Möbius transformation* [36]. Lemma 6 appeared in Greenberg's 1962 paper *Discrete subgroups of the Lorentz group* [164]. Lemma 7 was proved by Zassenhaus in his 1938 paper *Beweis eines Satzes über diskrete Gruppen* [421]. Moreover, Theorem 12.4.2 has its origins in this paper. Theorem 12.4.2, without an explicit bound, appeared in Każdan and Margulis' 1968 paper *A proof of Selberg's conjecture* [213]. See also Wang's 1969 paper *Discrete nilpotent subgroups of Lie groups* [402]. Theorem 12.4.2 for real matrices appeared in Martin's 1989 paper *On discrete Möbius groups in all dimensions* [280]. Theorem 12.4.3, without an explicit bound, appeared in Bowditch's 1993 paper [55]. The 2- and 3-dimensional versions of Theorem 12.4.4 were proved by Jørgensen in his 1977 paper *A note on subgroups of* $\mathrm{SL}(2, \mathbb{C})$ [211]. For improvements of Theorem 12.4.4, under additional hypothesis, see Martin's 1989 paper [280] and Abikoff and Haas' 1990 paper *Nondiscrete groups of hyperbolic motions* [3].

§12.5. The 3-dimensional versions of Theorems 12.5.1 and 12.5.2 appeared in Thurston's 1979 notes *The Geometry and Topology of 3-Manifolds* [389] and Gromov's 1981 paper *Hyperbolic manifolds according to Thurston and Jørgensen* [168]. See also Gromov's 1978 paper *Manifolds of negative curvature* [167]. The Margulis lemma has its origins in Każdan and Margulis' 1968 paper [213] and appeared in Gromov's 1978 paper [167] and in Thurston's 1979 notes [389]. The existence of parabolic Margulis regions in dimension two was established by Shimizu in his 1963 paper *On discontinuous groups operating on the product of the upper half planes* [374]. See also Leutbecher's 1967 paper *Über Spitzen diskontinuierlicher Gruppen von lineargebrochenen Transformationen* [256]. The existence of hyperbolic Margulis regions in dimension two was essentially established by Keen in her 1974 paper *Collars on Riemann surfaces* [216]. See also Halpern's 1981 paper *A proof of the collar lemma* [176] and Basmajian's 1992 paper *Generalizing the hyperbolic collar lemma* [32]. Hyperbolic Margulis regions in dimension three were studied by Brooks and Matelski in their 1982 paper *Collars in Kleinian groups* [60] and by Gallo in his 1983 paper *A 3-dimensional hyperbolic collar lemma* [145]. Lemmas 4 and 5 and Theorem 12.5.3 were proved by Susskind and Swarup in their 1992 paper *Limit sets of geometrically finite hyperbolic groups* [385]. The thick and thin parts of a hyperbolic space-form were introduced by Thurston in his 1979 lecture notes [389]. Theorem 12.5.4 was essentially proved by Każdan and Margulis in their 1968 paper [213]. See also Wang's 1969 paper [402]. Theorem 12.5.4 for Fuchsian groups appeared in Marden's 1974 paper *Universal properties of Fuchsian groups in the Poincaré metric* [275] and in Sturm and Shinnar's 1974 paper *The maximal inscribed ball of a Fuchsian group* [383]. Theorem 12.5.4 appeared in Apanasov's 1975 paper *A universal property of Kleinian groups in the hyperbolic metric* [17] and in Wielenberg's 1977 paper *Discrete Moebius groups: fundamental polyhedra* *and convergence* [411]. For a lower bound on the radius in Theorem 12.5.4, see Martin's 1989 paper *Balls in hyperbolic manifolds* [279].

§12.6. Lemmas 1-4, and Theorem 12.6.1 were proved by Bowditch in his 1993 paper [55]. Corollary 1 for Fuchsian groups was implicitly proved by Klein in his 1883 paper [233]. The 2- and 3-dimensional versions of Corollary 1 appeared implicitly in Vol. I of Fricke and Klein's 1897 treatise [139]. Corollary 1 appeared in Wielenberg's 1977 paper [411] and in Apanasov's 1982 paper [18]. The convex core of a hyperbolic surface was introduced by Löbell in his 1927 thesis *Die überall regulären unbegrenzten Flächen fester Krümmung* [265]. See also Koebe's 1928 paper [243] and Löbell's 1929 paper *Über die geodätischen Linien der Clifford-Kleinschen Flächen* [266]. Theorem 12.6.2 has its origins in Nielsen's 1940 paper [320]. The 2-dimensional version of Theorem 12.6.2 was proved by Nielsen and Fenchel in their 1959 manuscript *Discontinuous Groups of Non-Euclidean Motions* [321]. The convex core of a hyperbolic 3-manifold was introduced by Löbell in his 1931 paper *Beispiele geschlossener dreidimensionaler Clifford-Kleinscher Räume negativer Krümmung* [268]. The 3-dimensional version of Theorem 12.6.2 was proved by Thurston in his 1979 lecture notes [389]. Theorem 12.6.2 was essentially proved by Bowditch in his 1993 paper [55]. The 2-dimensional version of Theorem 12.6.3 was proved by Siegel in his 1945 paper *Some remarks on discontinuous groups* [376]. Theorem 12.6.3 was proved by Garland and Raghunathan in their 1970 paper *Fundamental domains for lattices in* ($\mathbb{R}$-)*rank 1 semisimple Lie groups* [147]. See also Margulis' 1969 paper *On the arithmeticity of discrete groups* [277] and Selberg's 1970 paper *Recent developments in the theory of discontinuous groups of motions of symmetric spaces* [373]. The 2-dimensional version of Theorem 12.6.4 was proved by Löbell in his 1930 paper *Ein Satz über die eindeutigen Bewegungen Clifford-Kleinscher Flächen in sich* [267]. Theorem 12.6.4 for dimension $n > 2$ seems to be new. Corollary 2 for closed surfaces was proved by Poincaré in his 1885 paper [334], and for closed n-manifolds by Lawson and Yau in their 1972 paper *Compact manifolds of nonpositive curvature* [254]. Corollary 2 was proved by Avérous and Kobayashi in their 1976 paper *On automorphisms of spaces of nonpositive curvature with finite volume* [29]. Lemma 6 appeared in Thurston's 1979 lecture notes [389]. Theorem 12.6.5 and Corollary 3 appeared in Marden's 1974 paper *The geometry of finitely generated Kleinian groups* [276]. Lemma 7 was proved by Bowditch in his 1993 paper [55]. The 3-dimensional version of Theorem 12.6.6 was proved by Marden in his 1974 paper [276]. Theorem 12.6.6 for manifolds with a finite-sided fundamental polyhedron appeared in Apanasov's 1983 paper *Geometrically finite hyperbolic structures on manifolds* [19]. See also Tukia's 1985 paper *On isomorphisms of geometrically finite Möbius groups* [392]. Theorem 12.6.6 was proved by Bowditch in his 1993 paper [55]. The 2-dimensional version of Corollary 4 was proved by Nielsen and Fenchel in their 1959 manuscript [321]. Corollary 4 was proved by Garland and Raghunathan in their 1970 paper [147].

CHAPTER 13

Geometric Orbifolds

In this chapter, we study the geometry of geometric orbifolds. We begin by studying the geometry of an orbit space of a discrete group of isometries of a geometric space. In Section 13.2, we study orbifolds modeled on a geometric space X via a group G of similarities of X. Such an orbifold is called an (X, G)-orbifold. In particular, if Γ is a discrete group of isometries of X, then the orbit space X/Γ is an (X, G)-orbifold for any group G of similarities of X containing Γ. In Section 13.3, we study the role of metric completeness in the theory of (X, G)-orbifolds. In particular, we prove that if M is a complete (X, G)-orbifold, with X simply connected, then there is a discrete subgroup Γ of G of isometries of X such that M is isometric to X/Γ. In Section 13.4, we prove the gluing theorem for geometric orbifolds. The chapter ends with a proof of Poincaré's fundamental polyhedron theorem.

§13.1. Orbit Spaces

In this section, we study the geometry of an orbit space X/Γ of a discrete group Γ of isometries of a geometric space X.

Theorem 13.1.1. *Let Γ be a discontinuous group of isometries of a metric space X and let $\pi : X \to X/\Gamma$ be the quotient map. Then for each point x of X, the map π induces a homeomorphism from $B(x,r)/\Gamma_x$ onto $B(\pi(x), r)$ for all r such that*

$$0 < r \leq \frac{1}{2}\mathrm{dist}(x, \Gamma x - \{x\}).$$

Moreover π induces an isometry from $B(x,r)/\Gamma_x$ onto $B(\pi(x), r)$ for all r such that

$$0 < r \leq \frac{1}{4}\mathrm{dist}(x, \Gamma x - \{x\}).$$

Proof: Let x be an arbitrary point of X. Then we have

$$\pi(B(x,r)) = B(\pi(x), r)$$

for each $r > 0$ by Theorem 6.5.2. Hence π is an open map. Set

$$s = \frac{1}{2}\text{dist}(x, \Gamma x - \{x\})$$

and suppose that $0 < r \leq s$. Then by the triangle inequality, we have

$$B(x,r) \cap gB(x,r) = \emptyset \quad \text{for all } g \text{ in } \Gamma - \Gamma_x.$$

Therefore π induces a homeomorphism from $B(x,r)/\Gamma_x$ onto $B(\pi(x), r)$.

Now suppose that $0 < r \leq s/2$. Let y and z be points of $B(x,r)$ with

$$d(y,z) = \text{dist}(y, \Gamma_x z)$$

and suppose that g is in $\Gamma - \Gamma_x$. Then we have

$$\begin{aligned} 2s &\leq d(x, gx) \\ &\leq d(x,y) + d(y, gz) + d(gz, gx) \\ &\leq (s/2) + d(y, gz) + (s/2). \end{aligned}$$

Hence, we have that

$$d(y, gz) \geq s > d(y,z).$$

Therefore, we have that

$$d(y,z) = \text{dist}(y, \Gamma z).$$

Hence, we have that

$$\text{dist}(\Gamma_x y, \Gamma_x z) = \text{dist}(\Gamma y, \Gamma z).$$

Thus π maps $B(x,r)/\Gamma_x$ isometrically onto $B(\pi(x), r)$. □

Theorem 13.1.2. *Let Γ be a discontinuous group of isometries of a geodesically connected and geodesically complete metric space X and let $\pi : X \to X/\Gamma$ be the quotient map. Then Γ is the set of all isometries ϕ of X such that $\pi\phi$ agrees with π on a nonempty open set; in particular, Γ is the group of all isometries ϕ of X such that $\pi\phi = \pi$.*

Proof: Let ϕ be an isometry of X such that $\pi\phi$ agrees with π on a nonempty open set U. Let x be a point of U such that the order of Γ_x is as small as possible. Set

$$s = \frac{1}{2}\text{dist}(x, \Gamma x - \{x\}).$$

Then by the triangle inequality, we have

$$B(x,r) \cap gB(x,r) = \emptyset \quad \text{for all } g \text{ in } \Gamma - \Gamma_x.$$

Let y be a point of $B(x,s) \cap U$. Then $\Gamma_y \subset \Gamma_x$, and so $\Gamma_y = \Gamma_x$. Therefore, every element of Γ_x fixes each point of the nonempty open set $B(x,s) \cap U$. Hence $\Gamma_x = \{1\}$ by Theorem 8.3.2.

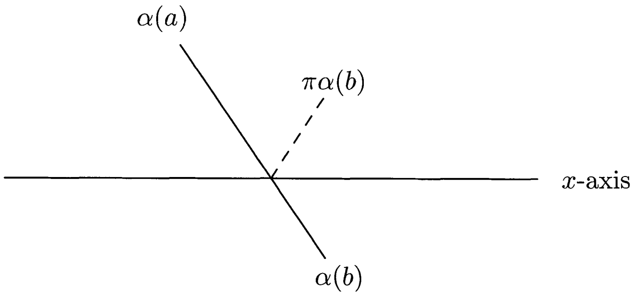


Figure 13.1.1. The image of a geodesic arc

Now as $\pi\phi(x) = \pi(x)$, there is an element g of Γ such that $\phi(x) = gx$. Hence $g^{-1}\phi$ is an isometry of X that fixes the point x. Now for each point y of $B(x,s) \cap U$, we have

$$\pi g^{-1}\phi(y) = \pi\phi(y) = \pi(y)$$

and so $g^{-1}\phi(y)$ is in

$$\Gamma y \cap B(x,s) = \{y\}.$$

Therefore $g^{-1}\phi$ is the identity on the open set $B(x,s) \cap U$. Hence $\phi = g$ by Theorem 8.3.2. □

Let Γ be a discrete group of isometries of a geometric space X and let $\pi : X \to X/\Gamma$ be the quotient map. If $\alpha : [a,b] \to X$ is a geodesic arc, then $\pi\alpha : [a,b] \to X/\Gamma$ is not necessarily a geodesic curve. For example, let $X = E^2$ and let Γ be the group generated by the reflection of E^2 in the x-axis. Then X/Γ is isometric to the closed half-plane $\overline{U}^2$. Observe that if $\alpha(a)$ and $\alpha(b)$ lie on opposite sides of the x-axis, then $\pi\alpha$ fails to be a geodesic curve at the point where α crosses the x-axis. See Figure 13.1.1. However, if $\alpha(a)$ or $\alpha(b)$ lies on the x-axis, then $\pi\alpha$ is a geodesic arc.

Lemma 1. *Let Γ be a finite group of isometries of a metric space X and let $\pi : X \to X/\Gamma$ be the quotient map. Let $\alpha : [a,b] \to X$ be a geodesic arc such that $\alpha(a)$ is fixed by every element of Γ. Then $\pi\alpha : [a,b] \to X/\Gamma$ is a geodesic arc.*

Proof: Observe that for each t in the interval $[a,b]$, we have

$$\begin{aligned} d_\Gamma(\pi\alpha(a), \pi\alpha(t)) &= \operatorname{dist}(\Gamma\alpha(a), \Gamma\alpha(t)) \\ &= \operatorname{dist}(\Gamma\alpha(a), \alpha(t)) \\ &= d(\alpha(a), \alpha(t)) \;=\; t-a. \end{aligned}$$

Now if $a \leq s < t \leq b$, then we have

$$d_\Gamma(\pi\alpha(a), \pi\alpha(s)) = s - a.$$

Hence, we have

$$\begin{aligned} d_\Gamma(\pi\alpha(s), \pi\alpha(t)) &\geq d_\Gamma(\pi\alpha(a), \pi\alpha(t)) - d_\Gamma(\pi\alpha(a), \pi\alpha(s)) \\ &= (t-a) - (s-a) \\ &= t - s. \end{aligned}$$

Moreover, we have

$$d_\Gamma(\pi\alpha(s), \pi\alpha(t)) = \text{dist}(\Gamma\alpha(s), \Gamma\alpha(t)) \leq d(\alpha(s), \alpha(t)) = t - s.$$

Therefore, we have

$$d_\Gamma(\pi\alpha(s), \pi\alpha(t)) = t - s.$$

Thus $\pi\alpha$ is a geodesic arc. □

Theorem 13.1.3. *Let Γ be a discontinuous group of isometries of a metric space X and let $\pi : X \to X/\Gamma$ be the quotient map. If $\alpha : [a, b] \to X$ is a geodesic arc, then $\pi\alpha : [a, b] \to X/\Gamma$ is a piecewise geodesic curve.*

Proof: For each point x of X, set

$$r(x) = \frac{1}{4}\text{dist}(x, \Gamma x - \{x\}).$$

Then the collection of open intervals

$$\{B(t, r(\alpha(t))) : a \leq t \leq b\}$$

covers $[a, b]$. Now as $[a, b]$ is compact, there is a partition $\{t_0, \ldots, t_m\}$ of $[a, b]$ such that for each $i = 1, \ldots, m$, we have

$$[t_{i-1}, t_i] \subset B(t, r(\alpha(t)))$$

for some t in $[a, b]$. Hence, by Theorem 13.1.1 and Lemma 1, we deduce that $\pi\alpha$ restricted to $[t_{i-1}, t_i]$ is either a geodesic arc if t is not in (t_{i-1}, t_i) or the product of two geodesic arcs joined at t if t is in (t_{i-1}, t_i). Thus $\pi\alpha$ is a piecewise geodesic curve. □

Note that Theorem 13.1.3 implies that $\pi : X \to X/\Gamma$ preserves the length of a geodesic arc $\alpha : [a, b] \to X$. The next theorem says that π preserves the length of any curve $\gamma : [a, b] \to X$.

Theorem 13.1.4. *Let Γ be a discontinuous group of isometries of a metric space X and let $\pi : X \to X/\Gamma$ be the quotient map. If $\gamma : [a, b] \to X$ is a curve, then $|\pi\gamma| = |\gamma|$.*

Proof: For each point x of X, set

$$r(x) = \frac{1}{4}\text{dist}(x, \Gamma x - \{x\}).$$

Then the collection of open balls

$$\mathcal{B} = \{B(\gamma(t), r(\gamma(t))) : a \leq t \leq b\}$$

covers $\gamma([a,b])$. Now as $\gamma([a,b])$ is compact, there is a partition $\{t_0,\ldots,t_m\}$ of $[a,b]$ such that for each i, there is a ball $B(x_i,r_i)$ in $\mathcal{B}$ such that

$$[t_{i-1},t_i] \subset B(x_i,r_i).$$

Moreover, by Theorem 13.1.1, there is a finite subgroup Γ_i of Γ such that π induces an isometry from $B(x_i,r_i)/\Gamma_i$ onto $B(\pi(x_i),r_i)$. Let γ_i be the restriction of γ to the interval $[t_{i-1},t_i]$, and let $\pi_i : X \to X/\Gamma_i$ be the quotient map. If the theorem is true for finite groups, then we would have

$$|\pi\gamma_i| = |\pi_i\gamma_i| = |\gamma_i|,$$

and it would then follow from the additivity of arc length that $|\pi\gamma| = |\gamma|$. Thus, we may assume that Γ is finite.

The proof now proceeds by induction on the order $|\Gamma|$ of Γ. The theorem is certainly true if $|\Gamma| = 1$. Assume that $|\Gamma| > 1$ and the theorem is true for all groups of order less than $|\Gamma|$. Let F be the set of points of X that are fixed by all the elements of Γ. If the image of γ is disjoint from F, then by the previous argument and the induction hypothesis, we can conclude that $|\pi\gamma| = |\gamma|$. Thus, we may assume that there is a number c in the interval $[a,b]$ such that $\gamma(c)$ is in F.

Now let $P = \{t_0,\ldots,t_m\}$ be an arbitrary partition of $[a,b]$. Then

$$\begin{aligned}
\ell(\pi\gamma,P) &= \sum_{i=1}^{m} d_\Gamma(\pi\gamma(t_{i-1}),\pi\gamma(t_i)) \\
&\leq \sum_{i=1}^{m} d(\gamma(t_{i-1}),\gamma(t_i)) \\
&= \ell(\gamma,P) \\
&\leq |\gamma|.
\end{aligned}$$

Hence $|\pi\gamma| \leq |\gamma|$.

On the contrary, suppose that $|\pi\gamma| < |\gamma|$. Then there is a partition $\{t_0,\ldots,t_m\}$ of $[a,b]$ such that

$$|\pi\gamma| < \sum_{i=1}^{m} d(\gamma(t_{i-1}),\gamma(t_i)).$$

Let γ_i be the restriction of γ to the interval $[t_{i-1},t_i]$. Then we have that

$$|\pi\gamma_i| < d(\gamma(t_{i-1}),\gamma(t_i))$$

for at least one index i. Thus, by replacing γ with γ_i, we may assume, without loss of generality, that

$$|\pi\gamma| < d(\gamma(a),\gamma(b)).$$

Now as the point $\gamma(c)$ is in F, we have

$$\begin{aligned}
d_\Gamma(\pi\gamma(a),\pi\gamma(c)) &= \mathrm{dist}(\Gamma\gamma(a),\Gamma\gamma(c)) \\
&= \mathrm{dist}(\gamma(a),\Gamma\gamma(c)) \\
&= d(\gamma(a),\gamma(c)).
\end{aligned}$$

Likewise, we have

$$d_\Gamma(\pi\gamma(c), \pi\gamma(b)) = d(\gamma(c), \gamma(b)).$$

Hence, we have

$$\begin{aligned} |\pi\gamma| &\geq d_\Gamma(\pi\gamma(a), \pi\gamma(c)) + d_\Gamma(\pi\gamma(c), \pi\gamma(b)) \\ &= d(\gamma(a), \gamma(c)) + d(\gamma(a), \gamma(c)) \\ &\geq d(\gamma(a), \gamma(b)), \end{aligned}$$

which is a contradiction. Thus $|\pi\gamma| = |\gamma|$. □

Theorem 13.1.5. *Let Γ be a discontinuous group of isometries of a finitely compact metric space X. If X is geodesically connected, then X/Γ is geodesically connected.*

Proof: Let Γx and Γy be distinct Γ-orbits and let $\ell = d_\Gamma(\Gamma x, \Gamma y)$. Now $\ell = \text{dist}(x, \Gamma y)$ and $B(x, \ell + 1)$ contains only finitely many points of Γy, since $\overline{B}(x, \ell + 1)$ is compact. Hence, there is an element g of Γ such that $\ell = d(x, gy)$.

Let $\alpha : [0, \ell] \to X$ be a geodesic arc from x to gy and let $\pi : X \to X/\Gamma$ be the quotient map. We now show that $\pi\alpha : [0, \ell] \to X/\Gamma$ is a geodesic arc from Γx to Γy. Suppose that $0 \leq s < t \leq \ell$. Then

$$d_\Gamma(\pi\alpha(s), \pi\alpha(t)) \leq d(\alpha(s), \alpha(t)) = t - s,$$

since π does not increase distances. Now observe that

$$\begin{aligned} \ell &= d_\Gamma(\pi\alpha(0), \pi\alpha(\ell)) \\ &\leq d_\Gamma(\pi\alpha(0), \pi\alpha(s)) + d_\Gamma(\pi\alpha(s), \pi\alpha(t)) + d_\Gamma(\pi\alpha(t), \pi\alpha(\ell)) \\ &\leq s + (t - s) + (\ell - t) \quad = \quad \ell. \end{aligned}$$

Hence, we have that

$$d_\Gamma(\pi\alpha(s), \pi\alpha(t)) = t - s.$$

Thus $\pi\alpha$ is a geodesic arc from Γx to Γy. □

Theorem 13.1.6. *Let Γ be a discrete group of isometries of a geometric space X and let $\pi : X \to X/\Gamma$ be the quotient map. If $\alpha : [a, b] \to X/\Gamma$ is a geodesic arc and x is a point of X such that $\pi(x) = \alpha(a)$, then there is a geodesic arc $\tilde{\alpha} : [a, b] \to X$ such that $\tilde{\alpha}(a) = x$ and $\pi\tilde{\alpha} = \alpha$; moreover, $\tilde{\alpha}$ is unique up to multiplication by an element of the stabilizer Γ_x.*

Proof: Since X is a geometric space, there is a $k > 0$ such that any point in the ball $B(x, k)$ distinct from x is joined to x by a unique geodesic segment. Set

$$s = \frac{1}{2}\text{dist}(x, \Gamma x - \{x\})$$

and let

$$r = \min\{k, s/2\}.$$

Suppose that c is a number such that $a < c \leq b$ and $c - a < r$. Then

$$d_\Gamma(\alpha(a), \alpha(c)) = c - a < r$$

and so there is a point z in $B(x, r)$ such that $\Gamma z = \alpha(c)$ and

$$d(x, z) = d_\Gamma(\alpha(a), \alpha(c)) = c - a.$$

Let t be a number such that $a < t < c$. Then we have that

$$d_\Gamma(\alpha(a), \alpha(t)) = t - a < r.$$

Hence, there is a point y in $B(x, r)$ such that $\Gamma y = \alpha(t)$ and $d(x, y) = t - a$. As $r \leq s/2$, we have that $d(y, z) < s$. Now, if g is in $\Gamma - \Gamma_x$, then

$$B(x, s) \cap gB(x, s) = \emptyset$$

and so $d(gy, z) \geq s$, since $r \leq s/2$. Therefore, by replacing y with gy for some g in Γ_x, we may assume that $d(y, z) = c - t$. As $r \leq k$, there is a unique geodesic segment $[x, z]$ in X joining x to z. Let $[x, y]$ be a geodesic segment in X joining x to y, and let $[y, z]$ be a geodesic segment in X joining y to z. Then we have

$$d(x, y) + d(y, z) = (t - a) + (c - t) = c - a = d(x, z).$$

Therefore, by Theorem 1.4.3, we have

$$[x, y] \cup [y, z] = [x, z].$$

Hence y lies on $[x, z]$ at a distance $t - a$ from x. Consequently

$$\tilde{\alpha}_{a,c} : [a, c] \to X,$$

defined by $\tilde{\alpha}_{a,c}(a) = x$ and $\tilde{\alpha}_{a,c}(t) = y$ and $\tilde{\alpha}_{a,c}(c) = z$, is a geodesic arc such that $\tilde{\alpha}_{a,c}(a) = x$ and $\pi\tilde{\alpha}_{a,c}(c) = \alpha(t)$ for all t in $[a, c]$.

Next suppose that

$$\hat{\alpha}_{a,c} : [a, c] \to X$$

is another geodesic arc such that $\hat{\alpha}_{a,c}(a) = x$ and $\pi\hat{\alpha}_{a,c}(c) = \alpha(t)$ for all t in $[a, c]$. Then we have

$$\pi\hat{\alpha}_{a,c}(c) = \alpha(c) = \Gamma z.$$

Now as

$$d(x, \hat{\alpha}_{a,c}(c)) = c - a < r < s,$$

there is a g in Γ_x such that $\hat{\alpha}_{a,c}(c) = gz$. Moreover, as

$$d(x, gz) = d(x, z) < r \leq k,$$

there is a unique geodesic segment $[x, gz]$ in X joining x to gz. Therefore $\hat{\alpha}_{a,c} = g\tilde{\alpha}_{a,c}$.

Next let ℓ be the supremum of all real numbers c such that $a < c \leq b$ and there is a geodesic arc $\tilde{\alpha}_{a,c} : [a, c] \to X$ such that $\tilde{\alpha}_{a,c}(a) = x$ and $\pi\tilde{\alpha}_{a,c}(t) = \alpha(t)$ for all t in $[a, c]$ and $\tilde{\alpha}_{a,c}$ is unique up to multiplication by an element of Γ_x.

Now since Γ_x is finite, there is an increasing sequence

$$a < c_1 < c_2 < \cdots$$

converging to ℓ such that $\tilde{\alpha}_{a,c_j}$ extends $\tilde{\alpha}_{a,c_i}$ for all $i < j$. Define

$$\tilde{\alpha}_{a,\ell} : [a, \ell] \to X$$

by $\tilde{\alpha}_{a,\ell}(t) = \tilde{\alpha}_{a,c_i}(t)$ if $a \leq t \leq c_i$ and

$$\tilde{\alpha}_{a,\ell}(\ell) = \lim_{i \to \infty} \tilde{\alpha}_{a,c_i}(c_i),$$

which exists, since $\{\tilde{\alpha}_{a,c_i}(c_i)\}$ is a Cauchy sequence. Clearly $\tilde{\alpha}_{a,\ell}$ preserves distances on $[a, \ell)$. Observe that if $a \leq t < \ell$, then we have

$$\begin{aligned} d(\tilde{\alpha}_{a,\ell}(t), \tilde{\alpha}_{a,\ell}(\ell)) &= d\big(\tilde{\alpha}_{a,\ell}(t), \lim_{i \to \infty} \tilde{\alpha}_{a,c_i}(c_i)\big) \\ &= \lim_{i \to \infty} d(\tilde{\alpha}_{a,\ell}(t), \tilde{\alpha}_{a,c_i}(c_i)) \\ &= \lim_{i \to \infty} |c_i - t| \quad = \quad \ell - t. \end{aligned}$$

Thus $\tilde{\alpha}_{a,\ell}$ preserves distances and therefore $\tilde{\alpha}_{a,\ell}$ is a geodesic arc. Clearly $\pi\tilde{\alpha}_{a,\ell}(t) = \alpha(t)$ for all t in $[a, \ell)$. As the quotient map $\pi : X \to X/\Gamma$ is continuous, $\pi\tilde{\alpha}_{a,\ell}(\ell) = \alpha(\ell)$.

Now suppose that $\hat{\alpha}_{a,\ell} : [a, \ell] \to X$ is another geodesic arc such that $\hat{\alpha}_{a,\ell}(a) = x$ and $\pi\hat{\alpha}_{a,\ell}(t) = \alpha(t)$ for all t. Then for each i, there is a g_i in Γ_x such that $\hat{\alpha}_{a,\ell}$ extends $g_i\tilde{\alpha}_{a,c_i}$. As Γ_x is finite, there is a g in Γ_x such that $g = g_i$ for infinitely many i. Thus, by passing to a subsequence, we may assume that $\hat{\alpha}_{a,\ell}$ extends $g\tilde{\alpha}_{a,c_i}$ for all i. Therefore $\hat{\alpha}_{a,\ell} = g\tilde{\alpha}_{a,\ell}$ by continuity.

We claim that $\ell = b$. On the contrary, suppose $\ell < b$. Let $z = \tilde{\alpha}_{a,\ell}(\ell)$. By the first part of the proof, there is a geodesic arc $\tilde{\alpha}_{\ell,d} : [\ell, d] \to X$ such that $\tilde{\alpha}_{\ell,d}(\ell) = z$ and $\pi\tilde{\alpha}_{\ell,d}(t) = \alpha(t)$ for all t in $[\ell, d]$. Define

$$\tilde{\alpha}_{a,d} : [a, d] \to X$$

by $\tilde{\alpha}_{a,d} = \tilde{\alpha}_{a,\ell}\tilde{\alpha}_{\ell,d}$. Then $\tilde{\alpha}_{a,d}(a) = x$ and $\pi\tilde{\alpha}_{a,d}(t) = \alpha(t)$ for all t in $[a, d]$. Let $w = \tilde{\alpha}_{a,d}(d)$. Then we have

$$\begin{aligned} d(x, w) &\geq \operatorname{dist}(x, \Gamma w) \\ &= d_\Gamma(\alpha(a), \alpha(d)) \\ &= d - a \\ &= (\ell - a) + (d - \ell) \\ &= d(x, z) + d(z, w) \quad \geq \quad d(x, w). \end{aligned}$$

Therefore, we have

$$d(x, w) = d(x, z) + d(z, w)$$

and so $\tilde{\alpha}_{a,d}$ is a geodesic arc by Theorem 1.4.3.

Now suppose that $\hat{\alpha}_{a,d} : [a, d] \to X$ is another geodesic arc such that $\hat{\alpha}_{a,d}(a) = x$ and $\pi\hat{\alpha}_{a,d}(t) = \alpha(t)$ for all t. Then there is an element g of Γ_x

such that $g\hat{\alpha}_{a,d}$ extends $\tilde{\alpha}_{a,\ell}$. Let $v = \hat{\alpha}_{a,d}(d)$ and let $[x,z], [x,w]$, and $[x,v]$ be the images of $\tilde{\alpha}_{a,\ell}, \tilde{\alpha}_{a,d}$, and $\hat{\alpha}_{a,d}$, respectively. As the geodesic segments $g[x,v]$ and $[x,w]$ both extend the geodesic segment $[x,z]$, we deduce that $g[x,v] = [x,w]$, since X is geodesically complete. Therefore $g\hat{\alpha}_{a,d} = \tilde{\alpha}_{a,d}$. Thus $\tilde{\alpha}_{a,d}$ is unique up to multiplication by an element of Γ_x. But $d > \ell$, which contradicts the supremacy of ℓ. Therefore, we must have $\ell = b$. Thus, there is a geodesic arc $\tilde{\alpha} : [a,b] \to X$ such that $\tilde{\alpha}(a) = x$, $\pi\tilde{\alpha} = \alpha$, and $\tilde{\alpha}$ is unique up to multiplication by an element of Γ_x. □

Theorem 13.1.7. *Let Γ be a discrete group of isometries of a geometric space X and let $\pi : X \to X/\Gamma$ be the quotient map. If $\gamma : [a,b] \to X/\Gamma$ is a rectifiable curve and x is a point of X such that $\pi(x) = \gamma(a)$, then there is a curve $\tilde{\gamma} : [a,b] \to X$ such that $\tilde{\gamma}(a) = x$ and $\pi\tilde{\gamma} = \gamma$.*

Proof: Since $\gamma : [a,b] \to X/\Gamma$ is uniformly continuous, for each positive integer j, there is a $\delta_j > 0$ such that if s,t are in $[a,b]$, with $|s-t| < \delta_j$, then we have that

$$d_\Gamma(\gamma(s), \gamma(t)) < 1/j.$$

Construct a sequence of partitions $P_j = \{t_{ij}\}$ of $[a,b]$ such that $|P_j| < \delta_j$ for each j and

$$\lim_{j\to\infty} \ell(\gamma, P_j) = |\gamma|.$$

Set

$$\ell_{ij} = d_\Gamma(\gamma(t_{ij}), \gamma(t_{i+1,j})) \quad \text{for each } i,j.$$

By Theorem 13.1.5, there is a geodesic arc $\alpha_{ij} : [0, \ell_{ij}] \to X/\Gamma$ starting at $\gamma(t_{ij})$ and ending at $\gamma(t_{i+1,j})$. Define $\gamma_j : [a,b] \to X/\Gamma$ by

$$\gamma_j(t) = \alpha_{ij}\left(\frac{t - t_{ij}}{t_{i+1,j} - t_{ij}}\ell_{ij}\right) \quad \text{if } t \text{ is in } [t_{ij}, t_{i+1,j}].$$

Let $\mathrm{C}([a,b],X)$ be the set of all continuous functions from $[a,b]$ to X. Define a metric D on $\mathrm{C}([a,b],X)$ by the formula

$$D(\alpha,\beta) = \sup\{d(\alpha(t),\beta(t)) : t \in [a,b]\}.$$

Then the metric topology determined by D is the compact-open topology. Likewise, define a metric D_Γ on $\mathrm{C}([a,b],X/\Gamma)$.

We now show that the sequence $\{\gamma_j\}$ converges to γ in $\mathrm{C}([a,b],X/\Gamma)$. Observe that if t is in $[t_{ij}, t_{i+1,j}]$, then

$$\begin{aligned} d_\Gamma(\gamma(t),\gamma_j(t)) &\leq d_\Gamma(\gamma(t),\gamma(t_{ij})) + d_\Gamma(\gamma(t_{ij}),\gamma_j(t)) \\ &< 1/j + d_\Gamma(\gamma(t_{ij}),\gamma(t_{i+1,j})) \\ &< 1/j + 1/j \;=\; 2/j. \end{aligned}$$

Hence, we have that

$$D_\Gamma(\gamma,\gamma_j) \leq 2/j.$$

Therefore $\gamma_j \to \gamma$.

Now by Theorem 13.1.6, the curve $\gamma_j : [a,b] \to X/\Gamma$ lifts to a curve $\tilde{\gamma}_j : [a,b] \to X$, with respect to π, such that $\tilde{\gamma}_j(a) = x$ for each j. We next show that the sequence $\{\tilde{\gamma}_j\}$ is equicontinuous. Let $\epsilon > 0$. Then there is a positive integer m such that

$$D_\Gamma(\gamma, \gamma_j) < \epsilon/3 \quad \text{for all } j > m.$$

For each t in $[a,b]$, let $\gamma_{a,t}$ be the restriction of γ to $[a,t]$ and let $\lambda(t) = |\gamma_{a,t}|$. Then $\lambda : [a,b] \to \mathbb{R}$ is continuous. Now since $\tilde{\gamma}_1, \ldots, \tilde{\gamma}_m$ and λ are uniformly continuous, there is a $\delta > 0$ such that if s, t are in $[a,b]$, with $s < t$ and $t - s < \delta$, then

$$d(\tilde{\gamma}_j(s), \tilde{\gamma}_j(t)) < \epsilon$$

for $j = 1, \ldots, m$ and

$$\lambda(t) - \lambda(s) < \epsilon/3.$$

Now suppose that $j > m$ and s, t are in $[a,b]$, with $s < t$ and $t - s < \delta$. Let $\gamma_{s,t}$ be the restriction of γ to $[s,t]$. Suppose that s is in $[t_{k-1,j}, t_{kj}]$ and t is in $[t_{\ell j}, t_{\ell+1,j}]$. Then we have

$$\begin{aligned}
&d(\tilde{\gamma}_j(s), \tilde{\gamma}_j(t)) \\
&\le d(\tilde{\gamma}_j(s), \tilde{\gamma}_j(t_{kj})) + \sum_{i=k}^{\ell-1} d(\tilde{\gamma}_j(t_{ij}), \tilde{\gamma}_j(t_{i+1,j})) + d(\tilde{\gamma}_j(t_{\ell j}), \tilde{\gamma}_j(t)) \\
&= d_\Gamma(\gamma_j(s), \gamma_j(t_{kj})) + \sum_{i=k}^{\ell-1} d_\Gamma(\gamma_j(t_{ij}), \gamma_j(t_{i+1,j})) + d_\Gamma(\gamma_j(t_{\ell j}), \gamma_j(t)) \\
&\le d_\Gamma(\gamma_j(s), \gamma(s)) + d_\Gamma(\gamma(s), \gamma(t_{kj})) + \sum_{i=k}^{\ell-1} d_\Gamma(\gamma(t_{ij}), \gamma(t_{i+1,j})) \\
&\quad + d_\Gamma(\gamma(t_{\ell j}), \gamma(t)) + d_\Gamma(\gamma(t), \gamma_j(t)) \\
&< \epsilon/3 + |\gamma_{s,t}| + \epsilon/3 \\
&= \epsilon/3 + \lambda(t) - \lambda(s) + \epsilon/3 \\
&< \epsilon/3 + \epsilon/3 + \epsilon/3 \;=\; \epsilon.
\end{aligned}$$

Thus $\{\tilde{\gamma}_j\}$ is equicontinuous.

Now observe that if t is in $[a,b]$, then we have

$$d(\tilde{\gamma}_j(a), \tilde{\gamma}_j(t)) \le |\tilde{\gamma}_j| = |\gamma_j| \le |\gamma|.$$

Thus, the image of $\tilde{\gamma}_j$ is contained in $\overline{B}(x, |\gamma|)$ for each j. It follows by the Arzela-Ascoli theorem that the sequence $\{\tilde{\gamma}_j\}$ has a limit point $\tilde{\gamma}$ in $\mathrm{C}([a,b], X)$. By passing to a subsequence, we may assume that $\tilde{\gamma}_j \to \tilde{\gamma}$. Then $\tilde{\gamma}_j(a) \to \tilde{\gamma}(a)$ and so $\tilde{\gamma}(a) = x$. Now the induced map

$$\pi_* : \mathrm{C}([a,b], X) \to \mathrm{C}([a,b], X/\Gamma)$$

is continuous. Therefore $\pi_*(\tilde{\gamma}_j) \to \pi_*(\tilde{\gamma})$. Hence $\gamma_j \to \pi\tilde{\gamma}$. Therefore $\pi\tilde{\gamma} = \gamma$. □

Exercise 13.1

1. Let Γ be a discrete group of isometries of a geometric space X and let x be a point of X. The point Γx of X/Γ is called a *ordinary point* of X/Γ if $\Gamma_x = \{1\}$, otherwise Γx is called a *singular point* of X/Γ. Prove that the set of all ordinary points of X/Γ is a connected, open, dense subset of X/Γ. Conclude that the set of all singular points of X/Γ is a closed nowhere dense subset of X/Γ.
2. Let Γ_1 and Γ_2 be discrete groups of isometries of $X = S^n, E^n$, or H^n. Prove that X/Γ_1 is isometric to X/Γ_2 if and only if Γ_1 is conjugate to Γ_2 in $\mathrm{I}(X)$.
3. Let Γ be a discrete group of isometries of $X = S^n, E^n$, or H^n. The *volume* of X/Γ is the volume of any proper fundamental domain for Γ in X. Prove that $\mathrm{Vol}(X/\Gamma)$ is an isometry invariant of X/Γ.
4. Let Γ be a discrete group of isometries of E^n. Prove that E^n/Γ has finite volume if and only if Γ is a crystallographic group.
5. Let Γ be an elementary discrete group of isometries of H^n. Prove that H^n/Γ has infinite volume.

§13.2. (X, G)-Orbifolds

Let G a group of similarities of a geometric space X and let M be a Hausdorff space. An (X, G)-*orbifold atlas* for M is defined to be a family of functions
$$\Phi = \{\phi_i : U_i \to X/\Gamma_i\}_{i\in\mathcal{I}},$$
called *charts*, satisfying the following conditions:

(1) The set U_i, called a *coordinate neighborhood*, is an open connected subset of M, and Γ_i is a discrete group of isometries of X for each i.

(2) The chart ϕ_i maps the coordinate neighborhood U_i homeomorphically onto an open subset of X/Γ_i for each i.

(3) The coordinate neighborhoods $\{U_i\}_{i\in\mathcal{I}}$ cover M.

(4) If U_i and U_j overlap, then the function
$$\phi_j\phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j),$$
called a *coordinate change*, has the property that if x and y are points of X such that
$$\phi_j\phi_i^{-1}(\Gamma_i x) = \Gamma_j y,$$
then there is an element g of G such that $gx = y$ and g lifts $\phi_j\phi_i^{-1}$ in a neighborhood of x, that is,
$$\phi_j\phi_i^{-1}(\Gamma_i w) = \Gamma_j g w$$
for all w in a neighborhood of x.

Theorem 13.2.1. *Let Φ be an (X,G)-orbifold atlas for M. Then there is a unique maximal (X,G)-orbifold atlas for M containing Φ.*

Proof: Let $\Phi = \{\phi_i : U_i \to X/\Gamma_i\}$ and let $\overline{\Phi}$ be the set of all functions $\phi : U \to X/\Gamma$ such that

(1) the set U is an open connected subset of M, and Γ is a discrete group of isometries of X;

(2) the function ϕ maps U homeomorphically onto an open subset of X/Γ;

(3) the function
$$\phi\phi_i^{-1} : \phi_i(U_i \cap U) \to \phi(U_i \cap U)$$
has the property that if w and x are points of X such that
$$\phi\phi_i^{-1}(\Gamma_i w) = \Gamma x,$$
then there is an element g of G such that $gw = x$ and g lifts $\phi\phi_i^{-1}$ in a neighborhood of w.

Clearly $\overline{\Phi}$ contains Φ. Suppose that $\phi : U \to X/\Gamma$ and $\psi : V \to X/\mathrm{H}$ are in $\overline{\Phi}$. Consider the function
$$\psi\phi^{-1} : \phi(U \cap V) \to \psi(U \cap V).$$
Suppose that x and y are points of X such that $\psi\phi^{-1}(\Gamma x) = \mathrm{H}y$. Let
$$\phi_i : U_i \to X/\Gamma_i$$
be in Φ such that $\phi^{-1}(\Gamma x)$ is in U_i. Then there is a point w of X such that
$$\phi_i^{-1}(\Gamma_i w) = \phi^{-1}(\Gamma x) = \psi^{-1}(\mathrm{H}y).$$
Hence, there are elements g and h of G such that $gw = x$ and $hw = y$, and g and h lift $\phi\phi_i^{-1}$ and $\psi\phi_i^{-1}$, respectively, in a neighborhood of w. Observe that $hg^{-1}x = y$ and hg^{-1} lifts $\psi\phi_i^{-1}\phi_i\phi^{-1} = \psi\phi^{-1}$ in a neighborhood of x. Thus $\overline{\Phi}$ is an (X,G)-orbifold atlas for M. Clearly $\overline{\Phi}$ contains every (X,G)-orbifold atlas for M containing Φ, and so $\overline{\Phi}$ is the unique maximal (X,G)-atlas for M containing Φ. □

Definition: An (X,G)-*orbifold structure* for a Hausdorff space M is a maximal (X,G)-orbifold atlas for M.

Definition: An (X,G)-*orbifold* M is a Hausdorff space M together with an (X,G)-orbifold structure for M.

Definition: A *geometric orbifold* is an (X,G)-orbifold such that X is an n-dimensional geometry.

Example 1. Let Γ be a discrete group of isometries of a geometric space X and let G be any group of similarities of X containing Γ. Then the identity map $\iota : X/\Gamma \to X/\Gamma$ constitutes an (X, G)-orbifold atlas for X/Γ. By Theorem 13.2.1, this atlas determines an (X, G)-orbifold structure for X/Γ, called the *induced* (X, G)-orbifold structure. Thus X/Γ together with the induced (X, G)-orbifold structure is an (X, G)-orbifold.

Example 2. An $(S^n, \mathrm{I}(S^n))$-orbifold is called a *spherical n-orbifold.*

Example 3. A $(E^n, \mathrm{I}(E^n))$-orbifold is called a *Euclidean n-orbifold.*

Example 4. An $(H^n, \mathrm{I}(H^n))$-orbifold is called a *hyperbolic n-orbifold.*

Example 5. A $(E^n, \mathrm{S}(E^n))$-orbifold is called a *Euclidean similarity n-orbifold.*

Definition: A *chart* for an (X, G)-orbifold M is an element $\phi : U \to X/\Gamma$ of the (X, G)-structure of M.

Theorem 13.2.2. *Let $\phi : U \to X/\Gamma$ be a chart for an (X, G)-orbifold M. Then Γ is a subgroup of G.*

Proof: By Theorem 6.5.15, the group Γ has a fundamental domain D in X. Let $\pi : X \to X/\Gamma$ be the quotient map. Then D contains a point x of the open set $\pi^{-1}(\phi(U))$, since ΓD is dense in X. Let f be an arbitrary element of Γ and set $y = fx$. Then $\Gamma x = \Gamma y$. Hence, there is an element g of G such that $gx = y$ and g lifts the identity map $\phi\phi^{-1}$ of $\phi(U)$ in a neighborhood of x. Therefore πg agrees with π in a nonempty open set. Hence g is in Γ by Theorem 13.1.2. As x is in D, the stabilizer Γ_x is trivial. Therefore $fx = gx$ implies that $f = g$. Hence f is in G. Thus Γ is a subgroup of G. □

Order of a Point

Let u be a point of an (X, G)-orbifold M. A *chart* for (M, u) is a chart $\phi : U \to X/\Gamma$ for M such that u is in U. Suppose that $\phi_i : U_i \to X/\Gamma_i$ and $\phi_j : U_j \to X/\Gamma_j$ are charts for (M, u). Then there are points x and y of X such that $\phi_i(u) = \Gamma_i x$ and $\phi_j(u) = \Gamma_j y$. Hence $\phi_j\phi_i^{-1}(\Gamma_i x) = \Gamma_j y$. Therefore, there is an element g of G such that $gx = y$ and g lifts $\phi_j\phi_i^{-1}$ in a neighborhood of x. Let Γ_x be the stabilizer of x in Γ_i and let Γ_y be the stabilizer of y in Γ_j. Let f be an element of Γ_x. Then we have that $gfg^{-1}y = y$ and gfg^{-1} lifts the identity map $(\phi_j\phi_i^{-1})(\phi_i\phi_j^{-1})$ of $\phi_j(U_i \cap U_j)$ in a neighborhood of y. Therefore gfg^{-1} is in Γ_y by Theorem 13.1.2. Thus $g\Gamma_x g^{-1} \subset \Gamma_y$. By reversing the roles of x and y, we deduce that $g^{-1}\Gamma_y g \subset \Gamma_x$. Therefore $g\Gamma_x g^{-1} = \Gamma_y$. Hence, the conjugacy class of Γ_x in G depends only on the point u.

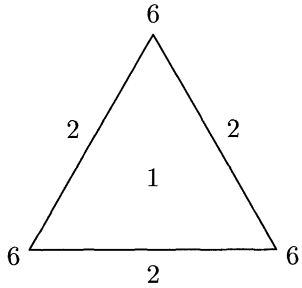


Figure 13.2.1. A Euclidean orbifold

The *order* of the point u of the orbifold M is the order of the stabilizer Γ_x. As Γ_x is determined up to conjugacy by u, the order of u does not depend on the choices of ϕ_i and x.

Example 6. Let Γ be the discrete group of isometries of E^2 generated by the reflections in the sides of an equilateral triangle $\triangle$. By Theorem 6.5.8, the inclusion map $\iota : \triangle \to E^2$ induces a homeomorphism $\kappa : \triangle \to E^2/\Gamma$. Consequently, we can pull back the Euclidean orbifold structure of E^2/Γ onto $\triangle$ by κ. Then the vertices of $\triangle$ have order six. The interior points of the sides of $\triangle$ have order two, and the interior points of $\triangle$ have order one. See Figure 13.2.1.

Theorem 13.2.3. *Let $\phi : U \to X/\Gamma$ be a chart for (M, u), let x be a point of X such that $\phi(u) = \Gamma x$, and let Γ_x be the stabilizer of x in Γ. Then there is an open neighborhood V of u in U such that ϕ restricted to V lifts to a chart $\psi : V \to X/\Gamma_x$ for (M, u).*

Proof: If $\Gamma_x = \Gamma$, then we may take $V = U$. Thus, we may assume that Γ_x is a proper subgroup of Γ. Set

$$s = \frac{1}{2}\mathrm{dist}(x, \Gamma x - \{x\}).$$

By Theorem 13.1.1, the quotient map $\pi : X \to X/\Gamma$ induces a homeomorphism

$$\eta : B(x, s)/\Gamma_x \to B(\pi(x), s).$$

Let $V = \phi^{-1}(B(\pi(x), s))$. Then V is an open neighborhood of u in U. Define $\psi : V \to X/\Gamma_x$ by $\psi(v) = \eta^{-1}\phi(v)$. Then ψ lifts the restriction of ϕ to V.

As the ball $B(x, s)$ is connected, $B(\pi(x), s))$ is also connected. Therefore V is connected. The function ϕ maps V homeomorphically onto $B(\pi(x), s)$ and η^{-1} maps $B(\pi(x), s)$ homeomorphically onto the open subset $B(x, s)/\Gamma_x$ of X/Γ_x. Therefore ψ maps V homeomorphically onto an open subset of X/Γ_x.

Now suppose that $\phi_i : U_i \to X/\Gamma_i$ is a chart for M. Consider the function

$$\psi\phi_i^{-1} : \phi_i(U_i \cap V) \to \psi(U_i \cap V).$$

Suppose that y and z are points of X such that

$$\psi\phi_i^{-1}(\Gamma_i y) = \Gamma_x z.$$

Then we have that

$$\eta^{-1}\phi\phi_i^{-1}(\Gamma_i y) = \Gamma_x z.$$

Hence, we have that

$$\phi\phi_i^{-1}(\Gamma_i y) = \Gamma z.$$

As ϕ and ϕ_i are charts for M, there is an element g of G such that $gy = z$ and g lifts $\phi\phi_i^{-1}$ in a neighborhood W of y. This means that

$$\phi\phi_i^{-1}(\Gamma_i w) = \Gamma g w$$

for all w in W. Let $\pi_i : X \to X/\Gamma_i$ be the quotient map and let

$$W' = W \cap \pi_i^{-1}(\phi_i(U_i \cap V)).$$

Then W' is a neighborhood of y in X, and for all w in W', we have

$$\psi\phi_i^{-1}(\Gamma_i w) = \eta^{-1}\phi\phi_i^{-1}(\Gamma_i w) = \eta^{-1}(\Gamma g w) = \Gamma_x g w.$$

Thus g lifts $\psi\phi_i^{-1}$ in a neighborhood of y. Therefore $\psi : V \to X/\Gamma_x$ is a chart for (M, u). □

An *ordinary point* of an (X, G)-orbifold M is a point of M of order one, and a *singular point* of M is a point of M of order greater than one. The *ordinary set* of M is the set $\Omega(M)$ of all ordinary points of M, and the *singular set* of M is the set $\Sigma(M)$ of all singular points of M.

Example 7. Consider the Euclidean orbifold structure on the equilateral triangle $\triangle$ in Example 6. Then $\Omega(\triangle) = \triangle^\circ$ and $\Sigma(\triangle) = \partial\triangle$.

Theorem 13.2.4. *Let M be an (X, G)-orbifold. Then the ordinary set $\Omega(M)$ is an open dense subset of M and the singular set $\Sigma(M)$ is a closed nowhere dense subset of M.*

Proof: Let u be an ordinary point of M. By Theorem 13.2.3, there is a chart $\phi : U \to X$ for (M, u). Then the order of each point of U is one. Hence $U \subset \Omega(M)$. Thus $\Omega(M)$ is open in M.

Now let v be an arbitrary point of M and let $\psi : V \to X/\Gamma$ be a chart for (M, v). Let D be a fundamental domain for Γ in X. Then ΓD is an open dense subset of X. Let W be an open neighborhood of v in V. Then there is an element g of Γ such that $\psi(W) \cap gD$ is nonempty. Now each point of $\psi^{-1}(gD)$ has order one. Therefore W contains an ordinary point of M. Thus $\Omega(M)$ is dense in M. As $\Sigma(M)$ is the complement of $\Omega(M)$ in M, we conclude that $\Sigma(M)$ is a closed nowhere dense subset of M. □

Theorem 13.2.5. *Let $\phi_j\phi_i^{-1} : \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$ be a coordinate change of an (X, G)-orbifold M. Then $\phi_j\phi_i^{-1}$ lifts to an element of G on each connected component over its domain.*

Proof: Let $\pi_i : X \to X/\Gamma_i$ be the quotient map and let C be a connected component of $\pi_i^{-1}(\phi_i(U_i \cap U_j))$. Let w be a point of C. Then there is an open neighborhood W of w in C and an element g of G such that g lifts $\phi_j\phi_i^{-1}$ on W. Let x be an arbitrary point of C. Then there are open subsets $W_1, \ldots, W_m$ of C such that $W = W_1$, the sets W_k and W_{k+1} overlap for $k = 1, \ldots, m-1$, the point x is in W_m, and $\phi_j\phi_i^{-1}$ lifts to an element g_k of G on W_k for each k.

It suffices to prove that we can replace g_m by g. The proof is by induction on m. This is certainly true if $m = 1$, so assume that $m > 1$, and we can replace g_{m-1} by g. By Theorem 13.2.4, the open set

$$\phi_j\phi_i^{-1}(\pi_i(W_{m-1} \cap W_m))$$

contains an ordinary point $\Gamma_j z$ of X/Γ_j. Then the stabilizer of z in Γ_j is trivial. Hence, there is an $r > 0$ such that

$$B(z, r) \cap B(fz, r) = \emptyset$$

for all $f \neq 1$ in Γ_j.

Let y be a point of $W_{m-1} \cap W_m$ such that

$$\phi_j\phi_i^{-1}(\Gamma_i y) = \Gamma_j z.$$

Then there is an $s > 0$ such that

$$B(y, s) \subset W_{m-1} \cap W_m,$$

$$gB(y, s) \subset B(gy, r),$$

$$g_m B(y, s) \subset B(g_m y, r).$$

Now observe that

$$\Gamma_j g y = \phi_j\phi_i^{-1}(\Gamma_i y) = \Gamma_j g_m y.$$

Hence, there is an element h of Γ_j such that $gy = hg_m y$. Moreover, if y' is in $B(y, s)$, then

$$\Gamma_j g y' = \phi_j\phi_i^{-1}(\Gamma_i y') = \Gamma_j g_m y'.$$

Hence, there is an element h' of Γ_j such that $gy' = h' g_m y'$. Observe that gy' is in $B(gy, r)$ and $g_m y'$ is in $B(g_m y, r)$. Hence $h' g_m y'$ is in the set

$$B(hg_m y, r) \cap B(h' g_m y, r).$$

Now since $\Gamma_j g_m y = \Gamma_j z$, the stabilizer of $g_m y$ in Γ_j is trivial, and so

$$hB(g_m y, r) \cap h'B(g_m y, r) = \emptyset \quad \text{unless } h = h'.$$

Hence $h = h'$. Therefore $gy' = hg_m y'$ for all y' in $B(y, s)$. Hence $g = hg_m$ by Theorem 8.3.2. Thus, we may replace g_m by g. □

Metric (X, G)-Orbifolds

Definition: A *metric (X, G)-orbifold* is a connected (X, G)-orbifold M such that G is a group of isometries of X.

Let $\gamma : [a, b] \to M$ be a curve in a metric (X, G)-orbifold M. We now defined the *X-length* of γ. Assume first that $\gamma([a, b])$ is contained in a coordinate neighborhood U. Let $\phi : U \to X/\Gamma$ be a chart for M. The *X-length* of γ is defined to be

$$\|\gamma\| = |\phi\gamma|.$$

We now show that the X-length of γ does not depend on the choice of the chart ϕ. Suppose that $\psi : V \to X/\mathrm{H}$ is another chart for M such that V contains $\gamma([a, b])$.

Assume first that $\phi\gamma$ is rectifiable. Then the curve $\phi\gamma : [a, b] \to X/\Gamma$ lifts to a curve $\widetilde{\phi\gamma} : [a, b] \to X$ by Theorem 13.1.7. Now by Theorem 13.2.5, there is an isometry g in G that lifts $\psi\phi^{-1}$ on $\widetilde{\phi\gamma}([a, b])$. Hence

$$|\phi\gamma| = |\widetilde{\phi\gamma}| = |g\widetilde{\phi\gamma}| = |\psi\phi^{-1}\phi\gamma| = |\psi\gamma|.$$

Now assume that $\phi\gamma$ is nonrectifiable. Then $\psi\gamma$ is nonrectifiable; otherwise, we could lift $\psi\gamma : [a, b] \to X/\mathrm{H}$ to a curve $\widetilde{\psi\gamma} : [a, b] \to X$ and $g^{-1}\widetilde{\psi\gamma}$ would be a rectifiable curve that lifts $\phi\gamma$, contrary to Theorem 13.1.4. Therefore, we have that

$$|\phi\gamma| = \infty = |\psi\gamma|.$$

Thus, the X-length of γ is well defined when the image of γ lies in a coordinate neighborhood of M.

Now assume that $\gamma : [a, b] \to M$ is an arbitrary curve. As $\gamma([a, b])$ is compact, there is a partition

$$a = t_0 < t_1 < \cdots < t_m = b$$

of $[a, b]$ such that $\gamma([t_{i-1}, t_i])$ is contained in a coordinate neighborhood U_i for each $i = 1, \ldots, m$. Let γ_{t_{i-1},t_i} be the restriction of γ to $[t_{i-1}, t_i]$. The *X-length* of γ is defined to be

$$\|\gamma\| = \sum_{i=1}^{m} \|\gamma_{t_{i-1},t_i}\|.$$

The X-length of γ does not depend on the choice of the partition $\{t_i\}$, since if

$$a = s_0 < s_1 < \cdots < s_\ell = b$$

is another partition such that $\gamma([s_{i-1}, s_i])$ is contained in a coordinate neighborhood V_i, then there is a third partition

$$a = r_0 < r_1 < \cdots < r_k = b$$

such that $\{r_i\} = \{s_i\} \cup \{t_i\}$, and therefore

$$\sum_{i=1}^{m} \|\gamma_{t_{i-1},t_i}\| = \sum_{i=1}^{k} \|\gamma_{r_{i-1},r_i}\| = \sum_{i=1}^{\ell} \|\gamma_{s_{i-1},s_i}\|.$$

Definition: A curve γ in a metric (X, G)-orbifold M is *X-rectifiable* if and only if $\|\gamma\| < \infty$.

Lemma 1. *Any two points of a metric (X, G)-orbifold M can be joined by an X-rectifiable curve in M.*

Proof: Define a relation on M by $u \sim v$ if and only if u and v are joined by an X-rectifiable curve in M. Clearly, this is an equivalence relation on M. Let $[u]$ be an equivalence class and suppose that v is in $[u]$. Let $\psi : V \to X/\mathrm{H}$ be a chart for (M, v). Then there is an $r > 0$ such that $\psi(V)$ contains $B(\psi(v), r)$. Let $\mathrm{H}x$ be an arbitrary point of $B(\psi(v), r)$. As X/H is geodesically connected, there is a geodesic arc $\alpha : [a, b] \to X/\mathrm{H}$ from $\psi(v)$ to $\mathrm{H}x$. Clearly $B(\psi(v), r)$ contains $\alpha([a, b])$. Hence $\psi^{-1}\alpha : [a, b] \to M$ is an X-rectifiable curve from v to $\psi^{-1}(\mathrm{H}x)$. This shows that $[u]$ contains the open set $\psi^{-1}(B(\psi(v), r))$. Thus $[u]$ is open in M. As M is connected, $[u]$ must be all of M. Thus, any two points of M can be joined by an X-rectifiable curve. □

Theorem 13.2.6. *Let M be a metric (X, G)-orbifold. Then the function $d : M \times M \to \mathbb{R}$, defined by*

$$d(u, v) = \inf_{\gamma} \|\gamma\|,$$

where γ varies over all X-rectifiable curves from u to v, is a metric on M.

Proof: By Lemma 1, the function d is well defined. Clearly d is nonnegative and $d(u, u) = 0$ for all u in M. To see that d is nondegenerate, let u, v be distinct points of M. Since M is Hausdorff, there is a chart $\phi : U \to X/\Gamma$ for (M, u) such that v is not in U. Choose $r > 0$ such that $\phi(U)$ contains $C(\phi(u), r)$. By Theorems 6.5.2 and 8.1.2, the set

$$S(\phi(u), r) = \{\Gamma x \in X/\Gamma : d_\Gamma(\phi(u), \Gamma x) = r\}$$

is compact. Hence, the set $T = \phi^{-1}(S(\phi(u), r))$ is closed in M, since M is Hausdorff.

Let $\gamma : [a, b] \to M$ be an arbitrary X-rectifiable curve from u to v. Since $\gamma([a, b])$ is connected and contains both u and v, it must meet T. Hence, there is a first point c of the open interval (a, b) such that $\gamma(c)$ is in T. Let $\gamma_{a,c}$ be the restriction of γ to $[a, c]$. Then the image of $\gamma_{a,c}$ is contained in $\phi^{-1}(C(\phi(u), r))$. Consequently, we have

$$\|\gamma\| \geq \|\gamma_{a,c}\| = |\phi\gamma_{a,c}| \geq d_\Gamma(\phi(u), \phi\gamma(c)) = r.$$

Therefore $d(u,v) \geq r > 0$. Thus d is nondegenerate. The rest of the proof follows the proof of Theorem 8.3.4. $\square$

Let M be a metric (X,G)-orbifold. Then the metric d, in Theorem 13.2.6, is called the *induced metric* on M. Henceforth, we shall assume that a metric (X,G)-orbifold is a metric space with the induced metric.

Theorem 13.2.7. *Let $\phi : U \to X/\Gamma$ be a chart for a metric (X,G)-orbifold M, let Γx be a point of $\phi(U)$, and let $r > 0$ be such that $\phi(U)$ contains the ball $B(\Gamma x, r)$. Then ϕ^{-1} maps $B(\Gamma x, r)$ homeomorphically onto $B(\phi^{-1}(\Gamma x), r)$.*

Proof: The proof is the same as the proof of Theorem 8.3.5 with x replaced by Γx. $\square$

Corollary 1. *If M is a metric (X,G)-orbifold, then the topology of M is the metric topology determined by the induced metric.*

Theorem 13.2.8. *Let $\phi : U \to X/\Gamma$ be a chart for a metric (X,G)-orbifold M, let Γx be a point of $\phi(U)$, and let $r > 0$ be such that $\phi(U)$ contains the ball $B(\Gamma x, r)$. Then ϕ^{-1} maps $B(\Gamma x, r/2)$ isometrically onto $B(\phi^{-1}(\Gamma x), r/2)$; therefore ϕ is a local isometry.*

Proof: The proof is the same as the proof of Theorem 8.3.6 with x replaced by Γx. $\square$

Exercise 13.2

1. Let $\phi : U \to X/\Gamma$ be a chart for an (X,G)-orbifold M and let g be an element of G. Show that the function $\overline{g} : X/\Gamma \to X/g\Gamma g^{-1}$, defined by
$$\overline{g}(\Gamma x) = g\Gamma g^{-1} gx,$$
is a similarity and that $\overline{g}\phi : U \to X/g\Gamma g^{-1}$ is a chart for M.

2. Let M be an (X,G)-orbifold. Prove that the (X,G)-orbifold structure of M contains a unique (X,G)-manifold structure for $\Omega(M)$.

3. Let $\gamma : [a,b] \to M$ be a curve in a metric (X,G)-orbifold. Prove that the X-length of γ is the same as the length of γ with respect to the induced metric.

4. Let Γ be a discrete group of isometries of a geometric space X. Prove that $\Omega(X/\Gamma)$ is a geodesically convex subset of X/Γ.

5. Let Γ be a discrete group of isometries of geometric space X. Show that the induced metric on X/Γ and $\Omega(X/\Gamma)$ is the orbit space metric d_Γ. Conclude that X/Γ is the metric completion of the metric (X,Γ)-manifold $\Omega(X/\Gamma)$.

§13.3. Developing Orbifolds

In this section, we study the role of metric completeness in the theory of (X, G)-orbifolds. In particular, we prove that if M is a complete (X, G)-orbifold, with X simply connected, then there is a discrete subgroup Γ of G of isometries of X such that M is (X, G)-equivalent to X/Γ.

(X, G)-Paths

Let M be an (X, G)-orbifold. Informally, an (X, G)-path over M is a list of data that describes a piecewise lifting of a curve in M to X. The formal definition goes as follows: Let x and y be points of X and let $\phi : U \to X/\Gamma$ and $\psi : V \to X/\mathrm{H}$ be charts for M such that Γx is in $\phi(U)$ and $\mathrm{H}y$ is in $\psi(V)$. An (X, G)-*path* over M from (x, ϕ) to (y, ψ) is a sequence

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

such that there is a partition $\{s_0, \ldots, s_m\}$ of the unit interval $[0, 1]$ so that $\alpha_i : [s_{i-1}, s_i] \to X$ is a curve and $\phi_i : U_i \to X/\Gamma_i$ is a chart for M such that if $\pi_i : X \to X/\Gamma_i$ is the quotient map, then

$$\pi_i\alpha_i([s_{i-1}, s_i]) \subset \phi_i(U_i)$$

for each i, and $g_0, \ldots, g_m$ are elements of G such that

(1) $x = g_0\alpha_1(0)$ and g_0 lifts $\phi\phi_1^{-1}$ in a neighborhood of $\alpha_1(0)$,

(2) $\alpha_i(s_i) = g_i\alpha_{i+1}(s_i)$ and g_i lifts $\phi_i\phi_{i+1}^{-1}$ in a neighborhood of $\alpha_{i+1}(s_i)$ for each $i = 1, \ldots, m-1$, and

(3) $\alpha_m(1) = g_m y$ and g_m lifts $\phi_m\psi^{-1}$ in a neighborhood of y.

Observe that

(1) $\phi^{-1}(\Gamma x) = \phi_1^{-1}\pi_1\alpha_1(0)$,

(2) $\phi_i^{-1}\pi_i\alpha_i(s_i) = \phi_{i+1}^{-1}\pi_{i+1}\alpha_{i+1}(s_i)$ for each $i = 1, \ldots, m-1$,

(3) $\phi_m^{-1}\pi_m\alpha_m(1) = \psi^{-1}(\mathrm{H}y)$,

and A describes the piecewise lifting of the curve

$$\overline{A} = (\phi_1^{-1}\pi_1\alpha_1)\cdots(\phi_m^{-1}\pi_m\alpha_m)$$

in M from the point $\phi^{-1}(\Gamma x)$ to the point $\psi^{-1}(\mathrm{H}y)$.

Example: Let $\alpha : [0, 1] \to X$ be the constant curve at the point x. Then

$$I = \{1, \alpha, \phi, 1\}$$

is an (X, G)-path over M from (x, ϕ) to (x, ϕ) called the *constant* (X, G)-*path* over M at (x, ϕ).

We now consider five operations on an (X,G)-path

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}.$$

1. Subdivision

For some index j, add a point s of the open interval (s_{j-1}, s_j) to the partition $\{s_0, \ldots, s_m\}$ and replace α_j in A by

$$\alpha_j|_{[s_{j-1},s]},\ \phi_j, 1,\ \alpha_j|_{[s,s_j]}.$$

2. Junction

Junction is the opposite operation of subdivision.

3. Translation

For some index j, if $\psi : V_j \to X/\mathrm{H}_j$ is a chart for M such that

$$\phi_j^{-1}\pi_j\alpha_j([s_{j-1}, s_j]) \subset V_j$$

and if f_j is an element of G that lifts

$$\psi_j\phi_j^{-1} : \phi_j(U_j \cap V_j) \to \psi_j(U_j \cap V_j)$$

in the component containing $\alpha_j([s_{j-1}, s_j])$, replace g_{j-1}, α_j, ϕ_j, g_j in A by

$$g_{j-1}f_j^{-1},\ f_j\alpha_j,\ \psi_j,\ f_jg_j.$$

Example: Let g be an element of G. Then g induces a similarity

$$\overline{g} : X/\Gamma_j \to X/g\Gamma_jg^{-1},$$

defined by

$$\overline{g}(\Gamma_jx) = g\Gamma_jg^{-1}gx,$$

such that the following diagram commutes:

$$\begin{array}{ccc} X & \xrightarrow{g} & X \\ \downarrow & & \downarrow \\ X/\Gamma_j & \xrightarrow{\overline{g}} & X/g\Gamma_jg^{-1} \end{array}$$

where the vertical maps are the quotient maps. Observe that the function

$$\overline{g}\phi_j : U_j \to X/g\Gamma_jg^{-1}$$

is a chart for M, since g lifts $(\overline{g}\phi_j)\phi_j^{-1}$. Hence, by translation, we may replace g_{j-1}, α_j, ϕ_j, g_j in A by

$$g_{j-1}g^{-1},\ g\alpha_j,\ \overline{g}\phi_j,\ gg_j.$$

Thus, we are free to translate by any element of G.

4. Reparameterization

For some increasing homeomorphism $h : [0,1] \to [0,1]$ such that $h(s_i) = t_i$ for $i = 0, \dots, m$, replace α_i by β_i, defined by

$$\beta_i(t) = \alpha_i(h^{-1}(t)) \quad \text{for } t_{i-1} \le t \le t_i \text{ and } i = 1, \dots, m.$$

5. Small Homotopy

Replace α_i by β_i for each $i = 1, \dots, m$ when there is a homotopy

$$H_i : [s_{i-1}, s_i] \times [0,1] \to X$$

from α_i to β_i such that

$$\pi_i H_i([s_{i-1}, s_i] \times [0,1]) \subset \phi_i(U_i)$$

and for all t, we have

(1) $\alpha_1(0) = H_1(0,t) = \beta_1(0)$,

(2) $H_i(s_i, t) = g_i H_{i+1}(s_i, t)$ for $i = 1, \dots, m-1$,

(3) $\alpha_m(1) = H_m(1,t) = \beta_m(1)$.

Homotopic (X, G)-Paths

Two (X, G)-paths A and B over M from (x, ϕ) to (y, ψ) are said to be *homotopic*, written $A \simeq B$, if and only if there is a finite sequence of the above five operations taking A to B. Being homotopic is obviously an equivalence relation among the set of (X, G)-paths over M from (x, ϕ) to (y, ψ). We shall denote the homotopy class of A by $[A]$.

Now let

$$\begin{aligned} A &= \{g_0, \alpha_1, \phi_1, g_1, \dots, g_{m-1}, \alpha_m, \phi_m, g_m\}, \\ B &= \{h_0, \beta_1, \psi_1, h_1, \dots, h_{n-1}, \beta_n, \psi_n, h_n\} \end{aligned}$$

be (X, G)-paths over M from (x, ϕ) to (y, ψ) and (y, ψ) to (z, χ), respectively. The *product* AB of A and B is the (X, G)-path over M from (x, ϕ) to (z, χ),

$$\{g_0, \alpha'_1, \phi_1, g_1, \dots, g_{m-1}, \alpha'_m, \phi_m, g_m h_0, \beta'_1, \psi_1, h_1, \dots, h_{n-1}, \beta'_n, \psi_n, h_n\},$$

where

$$\alpha'_i(s) = \alpha_i(2s) \quad \text{for } s_{i-1}/2 \le s \le s_i/2 \text{ and } i = 1, \dots, m$$

and

$$\beta'_j(s) = \beta_j(2s - 1) \quad \text{for } (1 + s_{j-1})/2 \le s \le (1 + s_j)/2 \text{ and } j = 1, \dots, n.$$

In order to simplify notation, we shall drop the primes in AB and ignore reparameterization. Observe that if $A \simeq A'$ and $B \simeq B'$, then $AB \simeq A'B'$. Hence, we may define the product

$$[A][B] = [AB].$$

Fundamental Orbifold Group

Let M be an (X, G)-orbifold. The *fundamental orbifold group* of M, based at (x, ϕ), is the set $\pi_1^o(M, x, \phi)$ of homotopy classes of (X, G)-paths over M from (x, ϕ) to (x, ϕ) together with the multiplication of homotopy classes.

Theorem 13.3.1. *Let M be an (X, G)-orbifold. Then $\pi_1^o(M, x, \phi)$ is a group.*

Proof: The multiplication of $\pi_1^o(M, x, \phi)$ satisfies the associative law, since homotopy includes reparameterization. Let $I = \{1, \alpha, \phi, 1\}$ be the constant (X, G)-path over M at (x, ϕ), and let

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

be an (X, G)-path over M from (x, ϕ) to (x, ϕ). Then we have

$$\begin{aligned} IA &= \{1, \alpha, \phi, 1\}\{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\} \\ &= \{1, \alpha, \phi, g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}. \end{aligned}$$

By translation, we have

$$IA \simeq \{g_0, g_0^{-1}\alpha, \phi_1, 1, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}.$$

Hence, by junction, we have

$$IA \simeq \{g_0, (g_0^{-1}\alpha)\alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}.$$

Now by small homotopy, we have

$$IA \simeq \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\} = A.$$

Likewise, we have that $AI \simeq A$. Hence, we have

$$[I][A] = [A] = [A][I].$$

Thus $[I]$ is the identity element of $\pi_1^o(M, x, \phi)$.

Given A as above, let

$$A^{-1} = \{g_m^{-1}, \alpha_m^{-1}, \phi_m, g_{m-1}^{-1}, \alpha_{m-1}^{-1}, \phi_{m-1}, g_{m-2}^{-1}, \ldots, g_1^{-1}, \alpha_1^{-1}, \phi_1, g_0^{-1}\}.$$

Then we have that

$$[A][A^{-1}] = [I] = [A^{-1}][A].$$

Hence $[A^{-1}]$ is the inverse of $[A]$. Thus $\pi_1^o(M, x, \phi)$ is a group. □

Holonomy

Let M be an (X, G)-orbifold, let x be a point of X, and let $\phi : U \to X/\Gamma$ be a chart for M such that Γx is in $\phi(U)$. Let

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

be an (X, G)-path over M from (x, ϕ) to (x, ϕ). Then the element $g_0 \cdots g_m$ of G depends only on $[A]$. Hence, we may define a homomorphism

$$\eta : \pi_1^o(M, x, \phi) \to G$$

by the formula

$$\eta([A]) = g_0 \cdots g_m.$$

The homomorphism η is called the *holonomy* of M determined by (x, ϕ).

Let Γ be a discrete group of isometries of X. Then the orbit space X/Γ is an (X, Γ)-orbifold such that the identity map

$$\iota : X/\Gamma \to X/\Gamma$$

is a chart for X/Γ.

Theorem 13.3.2. *Let Γ be a discrete group of isometries of a simply connected geometric space X. Then for any point x of X, the holonomy*

$$\eta : \pi_1^o(X/\Gamma, x, \iota) \to \Gamma$$

is an isomorphism.

Proof: We first show that η is surjective. Let g be an element of Γ. Then there is a curve $\alpha : [0, 1] \to X$ from x to gx. Observe that $A = \{1, \alpha, \iota, g\}$ is an (X, Γ)-path over X/Γ from (x, ι) to (x, ι) and $\eta([A]) = g$. Thus η is surjective.

We now show that η is injective. Let

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

be an (X, Γ)-path over X/Γ from (x, ι) to (x, ι) such that $g_0 \cdots g_m = 1$. Observe that by translation, we have

$$A \simeq \{1, g_0\alpha_1, \iota, g_0g_1, \alpha_2, \phi_2, g_3, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}.$$

Continuing in this way, we deduce that

$$A \simeq \{1, g_0\alpha_1, \iota, 1, g_0g_1\alpha_2, \iota, 1, \ldots, 1, g_0 \cdots g_{m-1}\alpha_m, \iota, 1\}.$$

Hence, by junction, we have

$$A \simeq \{1, (g_0\alpha_1)(g_0g_1\alpha_2) \cdots (g_0 \cdots g_{m-1}\alpha_m), \iota, 1\}.$$

Now since X is simply connected, the closed curve

$$(g_0\alpha_1)(g_0g_1\alpha_2) \cdots (g_0 \cdots g_{m-1}\alpha_m)$$

is null homotopic. Therefore $A \simeq I$. Thus η is injective. □

Universal Orbifold Covering Space

Let M be an (X,G)-orbifold. Let x,y,z be a points of X and suppose that $\phi: U \to X/\Gamma$, $\psi: V \to X/\mathrm{H}$, and $\chi: W \to X/\mathrm{K}$ are charts for M such that Γx is in $\phi(U)$, $\mathrm{H}y$ is in $\psi(V)$, and $\mathrm{K}z$ is in $\chi(W)$. An (X,G)-path J over M from (y,ψ) to (z,χ) is said to be *constant* if and only if $J = \{1,\beta,\psi,f\}$, where $\beta: [0,1] \to X$ is the constant curve at y.

Let A be an (X,G)-path over M from (x,ϕ) to (y,ψ) and let B be an (X,G)-path over M from (x,ϕ) to (z,χ). We say that A is *related* to B, written $A \sim B$, if and only if there is a constant (X,G)-path J over M from (y,ψ) to (z,χ) such that $AJ \simeq B$.

Lemma 1. *Being related is an equivalence relation among the set of all (X,G)-paths over M that start at (x,ϕ).*

Proof: As $AI \simeq A$, we have that $A \sim A$. Suppose that $A \sim B$ as above. Then there is a constant (X,G)-path $J = \{1,\beta,\psi,f\}$ over M from (y,ψ) to (z,χ). Let $J' = \{1,\gamma,\chi,f^{-1}\}$, where $\gamma: [0,1] \to X$ is the constant curve at z. Then J' is a constant (X,G)-path over M from (z,χ) to (y,ψ). Observe that

$$J' = \{1,\gamma,\chi,f^{-1}\} \simeq \{f^{-1}, f\gamma, \psi, 1\} = J^{-1}.$$

Therefore, we have that

$$BJ' \simeq AJJ^{-1} \simeq A.$$

Hence $B \sim A$.

Now suppose that $A \sim B$ and $B \sim C$. Then there is a constant (X,G)-path $K = \{1,\gamma,\chi,g\}$ over M such that $BK \simeq C$. Observe that

$$\begin{aligned} JK &= \{1,\beta,\psi,f\}\{1,\gamma,\chi,g\} \\ &= \{1,\beta,\psi,f,\gamma,\chi,g\} \\ &\simeq \{1,\beta,\psi,1,f\gamma,\psi,fg\} \\ &\simeq \{1,\beta f\gamma,\psi,fg\} \quad = \quad \{1,\beta,\psi,fg\} \end{aligned}$$

and the last (X,G)-path is constant. Moreover, we have that

$$AJK \simeq BK \simeq C.$$

Therefore $A \sim C$. Thus, being related is an equivalence relation. □

The *universal orbifold covering space* of M, based at (x,ϕ), is the set $\tilde{M}$ of all equivalence classes of (X,G)-paths over M starting at (x,ϕ). Let A be an (X,G)-path over M starting at (x,ϕ). The equivalence class of A will be denoted by $\langle A\rangle$. Define a function $\kappa: \tilde{M} \to M$ by

$$\kappa(\langle A\rangle) = \overline{A}(1).$$

The function κ is called the *universal orbifold covering projection* of $\tilde{M}$.

We now define a topology on $\tilde{M}$. Let A be an (X,G)-path over M from (x,ϕ) to (y,ψ), and let N be an open neighborhood of $\overline{A}(1)$ in M. Let $\langle A,N\rangle$ be the set of all equivalence classes of the form $\langle AB\rangle$, where B is an (X,G)-path over M starting at (y,ψ) such that $\overline{B}([0,1])\subset N$. Observe that if J is a constant (X,G)-path over M starting at (y,ψ), then $\langle A\rangle = \langle AJ\rangle$. Therefore $\langle A\rangle$ is in $\langle A,N\rangle$. Moreover, if $\langle A''\rangle$ is in $\langle A,N\rangle\cap\langle A',N'\rangle$, then $\overline{A}''(1)$ is in $N\cap N'$ and

$$\langle A'', N\cap N'\rangle \subset \langle A,N\rangle\cap\langle A',N'\rangle.$$

Consequently, the set of all subsets of $\tilde{M}$ of the form $\langle A,N\rangle$ form a basis for a topology on $\tilde{M}$. Henceforth, we shall regard $\tilde{M}$ to be topologized with this topology.

Lemma 2. *If A' is an (X,G)-path over M such that $\langle A'\rangle$ is in $\langle A,N\rangle$, then*

$$\langle A',N\rangle = \langle A,N\rangle.$$

Proof: Since $\langle A'\rangle$ is in $\langle A,N\rangle$, there is an (X,G)-path B over M such that $A'\sim AB$ and $\overline{B}([0,1])\subset N$. Hence, there is a constant (X,G)-path J such that $A'J\simeq AB$. Now if B' is an (X,G)-path over M starting where A' ends such that $\overline{B}'([0,1])\subset N$, then

$$A'B' \simeq A'JJ^{-1}B' \simeq ABJ^{-1}B'.$$

Therefore, we have

$$\langle A',N\rangle \subset \langle A,N\rangle.$$

Now as

$$A\simeq ABB^{-1}\simeq A'JB^{-1},$$

we have that $\langle A\rangle$ is in $\langle A',N\rangle$. Therefore, we have

$$\langle A,N\rangle \subset \langle A',N\rangle$$

by the previous argument. Thus $\langle A',N\rangle = \langle A,N\rangle$. □

Lemma 3. *Let M be an (X,G)-orbifold. Then a universal orbifold covering projection $\kappa:\tilde{M}\to M$ is a continuous open map. Moreover, if M is connected, then κ is surjective.*

Proof: Suppose that $\tilde{M}$ is based at (x,ϕ), let A be an (X,G)-path over M from (x,ϕ) to (y,ψ), and let N be an open neighborhood of $\overline{A}(1)$ in M. Then κ is continuous at $\langle A\rangle$, since

$$\kappa(\langle A,N\rangle)\subset N.$$

To show that κ is open, it suffices to show that $\kappa(\langle A,N\rangle)$ is open in M. Now since $\overline{A}(1)=\psi^{-1}(\mathrm{H}y)$, we find that $\psi^{-1}(\mathrm{H}y)$ is in $V\cap N$, and so $\mathrm{H}y$ is in $\psi(V\cap N)$. Let $s>0$ be such that

$$B(\mathrm{H}y,s)\subset\psi(V\cap N).$$

Then $\psi^{-1}(B(\mathrm{H}y, s))$ is an open neighborhood of $\overline{A}(1)$ in N and

$$\psi^{-1}(B(\mathrm{H}y, s)) \subset \kappa(\langle A, N \rangle),$$

since any geodesic arc in X/H from $\mathrm{H}y$ to a point of $B(\mathrm{H}y, s)$ lifts to a geodesic arc in X from y to a point of $B(y, s)$. Now $\langle A, N \rangle = \langle A', N \rangle$ for all A' in $\langle A, N \rangle$ by Lemma 2. Therefore, by the same argument, $\overline{A}'(1)$ has an open neighborhood contained in $\kappa(\langle A, N \rangle)$ for each $\langle A' \rangle$ in $\langle A, N \rangle$. Thus $\kappa(\langle A, N \rangle)$ is open in M.

By a similar argument, $M - \kappa(\tilde{M})$ is open in M. Hence $\kappa(\tilde{M})$ is both open and closed in M. Therefore, if M is connected, κ is surjective. □

Lemma 4. *Let M be an (X, G)-orbifold. Then every universal orbifold covering space $\tilde{M}$ of M is connected.*

Proof: Let $\tilde{M}$ be the universal orbifold covering space of M based at (x, ϕ). Let

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

be an (X, G)-path over M from (x, ϕ) to (y, ψ) and let I be the constant (X, G)-path over M at (x, ϕ). We claim that there is a curve in $\tilde{M}$ from $\langle I \rangle$ to $\langle A \rangle$. The proof is by induction on m.

Assume first that $m = 1$. Then

$$A = \{g_0, \alpha_1, \phi_1, g_1\}.$$

Let $J = \{1, \beta, \psi, g_1^{-1}\}$ be the constant (X, G)-path over M from (y, ψ) to $(\alpha_1(1), \phi_1)$. Then we have

$$J \simeq \{g_1^{-1}, g_1\beta, \phi_1, 1\}.$$

Hence, we have

$$\begin{aligned} AJ &\simeq \{g_0, \alpha_1, \phi_1, g_1\}\{g_1^{-1}, g_1\beta, \phi_1, 1\} \\ &= \{g_0, \alpha_1, \phi_1, 1, g_1\beta, \phi_1, 1\} \\ &= \{g_0, \alpha_1 g_1\beta, \phi_1, 1\} \\ &= \{g_0, \alpha_1, \phi_1, 1\}. \end{aligned}$$

Consequently, we may assume that $g_1 = 1$ and $(y, \psi) = (\alpha_1(1), \phi_1)$.

Now for each t in $[0, 1]$, define $\alpha_t : [0, 1] \to X$ by

$$\alpha_t(s) = \alpha_1(ts)$$

and define an (X, G)-path A_t over M from (x, ϕ) to $(\alpha_1(t), \phi_1)$ by

$$A_t = \{g_0, \alpha_t, \phi_1, 1\}.$$

Observe that α_0 is the constant curve at $\alpha_1(0)$ and

$$A_0 = \{g_0, \alpha_0, \phi_1, 1\} \simeq \{1, g_0\alpha_0, \phi, g_0\} \sim I.$$

Hence $\langle A_0 \rangle = \langle I \rangle$. Define $\gamma : [0, 1] \to \tilde{M}$ by $\gamma(t) = \langle A_t \rangle$.

We now show that γ is continuous at a point t. Let N be an open neighborhood of $\overline{A}(t)$ in M. Now since $\overline{A} = \phi_1^{-1}\pi_1\alpha_1$ is continuous at t, there is an $\epsilon > 0$ such that

$$\overline{A}(B(t,\epsilon) \cap [0,1]) \subset N.$$

We claim that

$$\gamma(B(t,\epsilon) \cap [0,1]) \subset \langle A_t, N\rangle.$$

Let r be in $B(t,\epsilon) \cap [0,1]$. Define a curve $\beta_r : [0,1] \to X$ by

$$\beta_r(s) = \alpha_1((1-s)t + sr)$$

and define an (X,G)-path B_r over M from $(\alpha_1(t),\phi_1)$ to $(\alpha_1(r),\phi_1)$ by

$$B_r = \{1, \beta_r, \phi_1, 1\}.$$

Then we have

$$\begin{aligned} A_tB_r &= \{g_0, \alpha_t, \phi_1, 1\}\{1, \beta_r, \phi_1, 1\} \\ &\simeq \{g_0, \alpha_t\beta_r, \phi_1, 1\} \\ &\simeq \{g_0, \alpha_r, \phi_1, 1\} \quad = \quad A_r \end{aligned}$$

and

$$\overline{B}_r([0,1]) = \phi_1^{-1}\pi_1\beta_r([0,1]) \subset N.$$

Hence $\gamma(r) = \langle A_r\rangle$ is in $\langle A_t, N\rangle$. Therefore

$$\gamma(B(t,\epsilon) \cap [0,1]) \subset \langle A_t, N\rangle$$

and so γ is continuous at t. Thus γ is a curve in $\tilde{M}$ from $\langle I\rangle$ to $\langle A\rangle$.

Now assume that $m > 1$ and let

$$A_{m-1} = \{g_0, \alpha_1', \phi_1, g_1, \ldots, g_{m-2}, \alpha_{m-1}', \phi_{m-1}, 1\}$$

be the (X,G)-path over M from (x,ϕ) to $(\alpha_{m-1}(s_{m-1}), \phi_{m-1})$ determined by A by reparameterization. Then by the induction hypothesis, $\langle I\rangle$ can be joined to $\langle A_{m-1}\rangle$ by a curve in $\tilde{M}$. Let $\tilde{M}'$ be the universal orbifold covering space of M based at $(\alpha_{m-1}(s_{m-1}), \phi_{m-1})$. Define a function

$$(A_{m-1})_* : \tilde{M}' \to \tilde{M}$$

by the formula

$$(A_{m-1})_*(\langle A'\rangle) = \langle A_{m-1}A'\rangle.$$

Then we have

$$(A_{m-1})_*(\langle A', N\rangle) = \langle A_{m-1}A', N\rangle.$$

Hence $(A_{m-1})_*$ is a homeomorphism with inverse $(A_{m-1}^{-1})_*$.

Let I_{m-1} be the constant (X,G)-path over M at $(\alpha_{m-1}(s_{m-1}), \phi_{m-1})$ and let

$$A_{m-1}' = \{g_{m-1}, \alpha_m', \phi_m, g_m\}$$

be the (X,G)-path over M from $(\alpha_{m-1}(s_{m-1}), \phi_{m-1})$ to (y,ψ) determined by A by reparameterization. Then by the case $m = 1$, we have that $\langle I_{m-1}\rangle$ can be joined to $\langle A_{m-1}'\rangle$ by a curve $\gamma : [0,1] \to \tilde{M}'$. Now

$$(A_{m-1})_*\gamma : [0,1] \to \tilde{M}$$

is a curve from $\langle A_{m-1}\rangle$ to $\langle A\rangle$. Hence $\langle I\rangle$ can be joined to $\langle A\rangle$ by a curve in $\tilde{M}$. Thus $\tilde{M}$ is connected. □

Lemma 5. *Let M be an (X,G)-orbifold. Then every universal orbifold covering space $\tilde{M}$ of M is Hausdorff.*

Proof: Let $\tilde{M}$ be the universal orbifold covering space of M based at (x,ϕ) and let $\kappa : \tilde{M} \to M$ be the universal covering projection. Let $\langle A \rangle$ and $\langle A' \rangle$ be distinct points of $\tilde{M}$. Assume first that $\kappa(\langle A \rangle)$ and $\kappa(\langle A' \rangle)$ are distinct. As M is Hausdorff, there are disjoint open neighborhoods N and N' of $\kappa(\langle A \rangle)$ and $\kappa(\langle A' \rangle)$, respectively. The projection κ is continuous by Lemma 3. Hence $\kappa^{-1}(N)$ and $\kappa^{-1}(N')$ are disjoint open neighborhoods of $\langle A \rangle$ and $\langle A' \rangle$, respectively. Thus, we may assume that $\kappa(\langle A \rangle) = \kappa(\langle A' \rangle)$.

Suppose that A is an (X,G)-path over M from (x,ϕ) to (y,ψ), where $\psi : V \to X/\mathrm{H}$. Let $r > 0$ be such that

(1) $B(\mathrm{H}y, r) \subset \psi(V)$,

(2) $r \leq \frac{1}{2}\mathrm{dist}(y, \mathrm{H}y - \{y\})$,

(3) $B(y,r)$ is simply connected.

Now set
$$N = \psi^{-1}(B(\mathrm{H}y, r)).$$
Then N is an open neighborhood of $\psi^{-1}(\mathrm{H}y) = \kappa(\langle A \rangle)$ in M.

We claim that $\langle A, N \rangle$ and $\langle A', N \rangle$ are disjoint open neighborhoods of $\langle A \rangle$ and $\langle A' \rangle$, respectively. On the contrary, suppose that $\langle A, N \rangle$ meets $\langle A', N \rangle$. Then $\langle A, N \rangle = \langle A', N \rangle$ by Lemma 2. Hence $\langle A' \rangle = \langle AB \rangle$ for some (X,G)-path B over M from (y,ψ) to (z,χ) such that $\overline{B}([0,1]) \subset N$. Suppose that
$$B = \{h_0, \beta_1, \psi_1, h_1, \ldots, h_{n-1}, \beta_n, \psi_n, h_n\}.$$
Then by Theorem 13.2.5, there is an element f_i of G such that f_i lifts the coordinate change
$$\psi\psi_i^{-1} : \psi_i(V_i \cap V) \to \psi(V_i \cap V)$$
in the component containing $\beta_i([s_{i-1}, s_i])$. Then by translation, we have
$$B \simeq \{h_0 f_1^{-1}, f_1\beta_1, \psi, f_1 h_1 f_2^{-1}, \ldots, f_{n-1}h_{n-1}f_n^{-1}, f_n\beta_n, \psi, f_n h_n\}.$$
Now since we are free to replace B by any element of $\langle B \rangle$, we may assume, without loss of generality, that $\psi_i = \psi$ for all i to begin with. Then each h_i lifts $\psi\psi^{-1}$, and so h_i is in H for each i. Hence, by translation, we have
$$\begin{aligned} B &\simeq \{1, h_0\beta_1, \psi, h_0h_1, \beta_2, \psi, h_2, \ldots, h_{n-1}, \beta_n, \psi, h_n\} \\ &\simeq \{1, h_0\beta_1, \psi, 1, h_0h_1\beta_2, \psi, h_0h_1h_2, \beta_3, \psi, h_3, \ldots, h_{n-1}, \beta_n, \psi, h_n\} \\ &\vdots \\ &\simeq \{1, h_0\beta_1, \psi, 1, h_0h_1\beta_2, \psi, 1, \ldots, 1, h_0\cdots h_{n-1}\beta_n, \psi, h_0\cdots h_n\}. \end{aligned}$$
Hence, we may assume that $h_i = 1$ for $i = 1, \ldots, n-1$. Then by junction, we have that
$$B \simeq \{1, \beta_1\cdots\beta_n, \psi, h_n\}.$$

Hence, we may assume that

$$B = \{1, \beta, \psi, h\}.$$

Let $J = \{1, \gamma, \chi, h^{-1}\}$ be the constant (X, G)-path over M from (z, χ) to $(\beta(1), \psi)$. Then we have

$$J \simeq \{h^{-1}, h\gamma, \psi, 1\}.$$

Hence, we have

$$\begin{aligned} BJ &\simeq \{1, \beta, \psi, h\}\{h^{-1}, h\gamma, \psi, 1\} \\ &= \{1, \beta, \psi, 1, h\gamma, \psi, 1\} \\ &\simeq \{1, \beta h\gamma, \psi, 1\} \\ &\simeq \{1, \beta, \psi, 1\}. \end{aligned}$$

Hence, we may assume that $h = 1$ and $(z, \chi) = (\beta(1), \psi)$.

Now as

$$\kappa(\langle AB \rangle) = \kappa(\langle A \rangle),$$

we have that

$$\psi^{-1}(\mathrm{H}\beta(1)) = \psi^{-1}(\mathrm{H}y).$$

Hence $\mathrm{H}\beta(1) = \mathrm{H}y$ and so there is an element f of H such that $f\beta(1) = y$. Let $\eta : X \to X/\mathrm{H}$ be the quotient map. Then we have

$$\eta(\beta([0, 1])) \subset B(\mathrm{H}y, r).$$

Hence, we have

$$\beta([0, 1]) \subset \eta^{-1}(B(\mathrm{H}y, r)) = \bigcup_{h \in \mathrm{H}} B(hy, r).$$

Now since

$$r \leq \frac{1}{2}\mathrm{dist}(y, \mathrm{H}y - \{y\}),$$

any two balls in $\{B(hy, r) : h \in \mathrm{H}\}$ are disjoint or coincide. Moreover

$$B(hy, r) = B(y, r)$$

if and only if h is in the stabilizer H_y of y. As $\beta(0) = y$ and $\beta([0, 1])$ is connected, we deduce that

$$\beta([0, 1]) \subset B(y, r).$$

As $f\beta(1) = y$, we must have that f is in H_y. Therefore $\beta(1) = y$. Thus β is a closed curve. Now since $B(y, r)$ is simply connected, β is null homotopic in $B(y, r)$. Therefore $AB \simeq A$. Thus, we have

$$\langle A' \rangle = \langle AB \rangle = \langle A \rangle,$$

which is a contradiction. Therefore $\langle A, N \rangle$ and $\langle A', N \rangle$ are disjoint open neighborhoods of $\langle A \rangle$ and $\langle A' \rangle$ in $\tilde{M}$, respectively. Thus $\tilde{M}$ is Hausdorff. □

The Developing Map

Let

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

be an (X, G)-path over M from (x, ϕ) to (y, ψ). Then the point $g_0 \cdots g_m y$ of X depends only on $[A]$. Moreover, if $J = \{1, \beta, \psi, f\}$ is a constant (X, G)-path over M from (y, ψ) to (z, χ), then we have

$$g_0 \cdots g_m f z = g_0 \cdots g_m y,$$

since $fz = y$, and so $g_0 \cdots g_m y$ depends only on $\langle A \rangle$.

Let $\tilde{M}$ be the universal orbifold covering space of M based at (x, ϕ). The *developing map* determined by (x, ϕ) is the function $\delta : \tilde{M} \to X$ defined by

$$\delta(\langle A \rangle) = g_0 \cdots g_m y.$$

Lemma 6. *Let $\tilde{M}$ be a universal orbifold covering space of an (X, G)-orbifold M. Then the developing map $\delta : \tilde{M} \to X$ is a local homeomorphism.*

Proof: Let δ be determined by (x, ϕ) and let

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

be an (X, G)-path over M from (x, ϕ) to (y, ψ), where $\psi : V \to X/\mathrm{H}$. Let $r > 0$ be such that

(1) $B(\mathrm{H}y, r) \subset \psi(V)$,

(2) $r \leq \frac{1}{2}\mathrm{dist}(y, \mathrm{H}y - \{y\})$,

(3) $B(y, r)$ is simply connected.

Now set

$$N = \psi^{-1}(B(\mathrm{H}y, r)).$$

Then N is an open neighborhood of $\psi^{-1}(\mathrm{H}y) = \overline{A}(1)$ in M. Let

$$g = g_0 \cdots g_m.$$

We claim that δ maps the set $\langle A, N \rangle$ bijectively onto the ball $gB(y, r)$. Let $\langle A' \rangle$ be an element of $\langle A, N \rangle$. By the argument in Lemma 5, we have that $\langle A' \rangle = \langle AB \rangle$, where

$$B = \{1, \beta, \psi, 1\}$$

is an (X, G)-path over M from (y, ψ) to $(\beta(1), \psi)$ such that

$$\beta([0, 1]) \subset B(y, r).$$

Hence $\delta(\langle A' \rangle) = g\beta(1)$ is in $gB(y, r)$. Moreover, since we may take β to be any rescaled geodesic arc in $B(y, r)$, we have that

$$\delta(\langle A, N \rangle) = gB(y, r).$$

Now suppose that
$$B' = \{1, \beta', \psi, 1\}$$
is another (X, G)-path over M from (y, ψ) to $(\beta'(1), \psi)$ such that
$$\beta'([0,1]) \subset B(y,r) \quad \text{and} \quad \delta(\langle AB\rangle) = \delta(\langle AB'\rangle).$$
Then $g\beta(1) = g\beta'(1)$. Hence $\beta(1) = \beta'(1)$. Now since $B(y,r)$ is simply connected, β is homotopic to β' in $B(y,r)$ by a homotopy keeping the endpoints fixed. Hence $B \simeq B'$ and so $\langle AB\rangle = \langle AB'\rangle$. Thus δ maps $\langle A, N\rangle$ injectively onto $gB(y,r)$.

Now since the sets of the form $\langle A, N\rangle$ form a basis for the topology of $\tilde{M}$, we deduce that $\delta : \tilde{M} \to X$ is a local homeomorphism. □

It follows from Lemmas 5 and 6 that a developing map $\delta : \tilde{M} \to X$ induces an $(X, \{1\})$-manifold structure on $\tilde{M}$. We shall regard the universal orbifold covering space $\tilde{M}$ to be an $(X, \{1\})$-manifold whose charts are the restrictions of δ. Then $\tilde{M}$ has a metric such that $\delta : \tilde{M} \to X$ is an $(X, \{1\})$-map and therefore a local isometry. Thus, we have the following theorem.

Theorem 13.3.3. *If $\tilde{M}$ is a universal orbifold covering space of an (X, G)-orbifold M, then $\tilde{M}$ is an $(X, \{1\})$-manifold such that the developing map $\delta : \tilde{M} \to X$ is an $(X, \{1\})$-map.*

Observe that the fundamental orbifold group $\pi_1^o(M, x, \phi)$ of an (X, G)-orbifold M acts on the universal orbifold covering space $\tilde{M}$ of M based at (x, ϕ) by the formula
$$[C]\langle A\rangle = \langle CA\rangle.$$

Theorem 13.3.4. *Let $\tilde{M}$ be the universal orbifold covering space based at (x, ϕ) of a connected (X, G)-orbifold M. Then $\pi_1^o(M, x, \phi)$ acts effectively and discontinuously on $\tilde{M}$ via similarities, and the universal orbifold covering projection $\kappa : \tilde{M} \to M$ induces a homeomorphism*
$$\overline{\kappa} : \tilde{M}/\pi_1^o(M, x, \phi) \to M.$$

Proof: We first show that $\pi_1^o(M, x, \phi)$ acts effectively on $\tilde{M}$. Suppose that A is an (X, G)-path over M from (x, ϕ) to (y, ψ), and $[C]$ is an element of $\pi_1^o(M, x, \phi)$, and $[C]\langle A\rangle = \langle A\rangle$. Then $\langle CA\rangle = \langle A\rangle$. Hence, there is a constant (X, G)-path $J = \{1, \beta, \psi, f\}$ over M from (y, ψ) to (y, ψ) such that $CAJ \simeq A$. Now $fy = y$ and f lifts $\psi\psi^{-1}$ in a neighborhood of y. Hence f is in the stabilizer H_y.

Observe that the homotopy classes of the form $[J]$, with J as above, form a subgroup of $\pi_1^o(M, y, \psi)$ isomorphic to H_y via the holonomy
$$\eta : \pi_1^o(M, y, \psi) \to G,$$
and since $[C] = [AJ^{-1}A^{-1}]$, this subgroup of $\pi_1^o(M, y, \psi)$ is isomorphic to the stabilizer of $\langle A\rangle$ via the change of base point isomorphism
$$[A]_* : \pi_1^o(M, y, \psi) \to \pi_1^o(M, x, \phi).$$

Thus, the stabilizer of $\langle A\rangle$ is isomorphic to the finite group H_y. In particular, if $\overline{A}(1)=\psi^{-1}(\mathrm{H}y)$ is an ordinary point of M, then the stabilizer of $\langle A\rangle$ is trivial. Hence $\pi_1^o(M,x,\phi)$ acts effectively on $\tilde{M}$.

We next show that $\pi_1^o(M,x,\phi)$ acts on $\tilde{M}$ via similarities. Let $[C]$ be an element of $\pi_1^o(M,x,\phi)$. Then we have

$$\delta([C]\langle A\rangle)=\delta(\langle CA\rangle)=\eta([C])\delta(\langle A\rangle).$$

Hence, the following diagram commutes:

$$\begin{array}{ccc} \tilde{M} & \xrightarrow{\delta} & X \\ {[C]_*}\downarrow & & \downarrow \eta([C])_* \\ \tilde{M} & \xrightarrow{\delta} & X \end{array}$$

Now as δ is a local isometry and $\eta([C])_*$ is a similarity, we deduce that $[C]_*$ is a local similarity, all of whose local scale factors are the same. As $[C]_*$ is a bijection, we conclude that $[C]_*$ is a similarity by the same argument as in the proof of Theorem 8.5.8. Thus $\pi_1^o(M,x,\phi)$ acts on $\tilde{M}$ via similarities.

We next show that the $\pi_1^o(M,x,\phi)$-orbits are the fibers of $\kappa:\tilde{M}\to M$. If $[C]$ is in $\pi_1^o(M,x,\phi)$, then

$$\kappa([C]\langle A\rangle)=\kappa(\langle A\rangle).$$

Hence, we have

$$\pi_1^o(M,x,\phi)\langle A\rangle\subset\kappa^{-1}(\kappa(\langle A\rangle)).$$

Now let B be an (X,G)-path over M from (x,ϕ) to (z,χ) such that

$$\kappa(\langle A\rangle)=\kappa(\langle B\rangle)$$

Suppose that $\chi:W\to X/\mathrm{K}$. Then

$$\psi^{-1}(\mathrm{H}y)=\chi^{-1}(\mathrm{K}z).$$

Let f be an element of G such that $fz=y$ and f lifts $\psi\chi^{-1}$ in a neighborhood of z and let

$$J=\{1,\beta,\psi,f\}$$

be the constant (X,G)-path over M from (y,ψ) to (z,χ). Then $B(AJ)^{-1}$ is an (X,G)-path over M from (x,ϕ) to (x,ϕ) and we have

$$\begin{aligned}[B(AJ)^{-1}]\langle A\rangle &= \langle B(AJ)^{-1}A\rangle \\ &= \langle BJ^{-1}A^{-1}A\rangle \\ &= \langle BJ^{-1}\rangle \ = \ \langle B\rangle.\end{aligned}$$

Hence $\langle B\rangle$ is in $\pi_1^o(M,x,\phi)\langle A\rangle$. Therefore

$$\pi_1^o(M,x,\phi)\langle A\rangle=\kappa^{-1}(\kappa(\langle A\rangle)).$$

Thus, the $\pi_1^o(M,x,\phi)$-orbits are the fibers of κ.

We next show that $\pi_1^o(M,x,\phi)$ acts discontinuously on $\tilde{M}$. First of all, the $\pi_1^o(M,x,\phi)$-orbits are closed, since they are the fibers of $\kappa:\tilde{M}\to M$. Let A be an (X,G)-path over M from (x,ϕ) to (y,ψ), where $\psi:V\to X/\mathrm{H}$. Let $r>0$ be such that

(1) $B(\mathrm{H}y, r) \subset \psi(V)$,

(2) $r \leq \frac{1}{2}\mathrm{dist}(y, \mathrm{H}y - \{y\})$,

(3) $B(y, r)$ is simply connected.

Now set
$$N = \psi^{-1}(B(\mathrm{H}y, r)).$$
Then N is an open neighborhood of $\psi^{-1}(\mathrm{H}y) = \kappa(\langle A\rangle)$ in M. By the argument in Lemma 6, we have
$$\langle A, N\rangle \cap \kappa^{-1}(\kappa(\langle A\rangle)) = \langle A\rangle.$$
Hence $\langle A\rangle$ is open in $\kappa^{-1}(\kappa(\langle A\rangle))$. Thus, the $\pi_1^o(M, x, \phi)$-orbits are discrete. Therefore $\pi_1^o(M, x, \phi)$ acts discontinuously on $\tilde{M}$ by Theorem 5.3.4.

Now $\kappa : \tilde{M} \to M$ is a continuous open surjection by Lemma 3, and the fibers of κ are the $\pi_1^o(M, x, \phi)$-orbits. Therefore κ induces a homeomorphism
$$\overline{\kappa} : \tilde{M}/\pi_1^o(M, x, \phi) \to M. \qquad \square$$

Theorem 13.3.5. *Let $\tilde{M}$ be the universal orbifold covering space based at (x, ϕ) of a connected (X, G)-orbifold M and let G_1 be the group of isometries in G. Then the following are equivalent:*

(1) *The group $\pi_1^o(M, x, \phi)$ acts on $\tilde{M}$ via isometries.*

(2) *The image of the holonomy $\eta : \pi_1^o(M, x, \phi) \to G$ is contained in G_1.*

(3) *The (X, G)-orbifold structure Φ of M contains an (X, G_1)-orbifold structure Φ_1 for M containing ϕ.*

Proof: Let $[C]$ be an element of $\pi_1^o(M, x, \phi)$. Then we have the commutative diagram
$$\begin{array}{ccc} \tilde{M} & \xrightarrow{\delta} & X \\ [C]_* \downarrow & & \downarrow \eta([C])_* \\ \tilde{M} & \xrightarrow{\delta} & X. \end{array}$$
Now by Theorem 13.3.4, the map $[C]_*$ is a similarity. As δ is a local isometry, $[C]_*$ is an isometry if and only if $\eta([C])_*$ is an isometry. Thus (1) and (2) are equivalent.

Suppose that the image of the holonomy $\eta : \pi_1^o(M, x, \phi) \to G$ is contained in G_1. Let
$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$
be an (X, G)-path over M from (x, ϕ) to (y, ψ), where $\psi : V \to X/\mathrm{H}$. Let
$$g = g_0 \cdots g_m$$
and let
$$\overline{g} : X/\mathrm{H} \to X/g\mathrm{H}g^{-1}$$

be the induced similarity. Define a function

$$\psi_A : V \to X/g\mathrm{H}g^{-1}$$

by $\psi_A = \overline{g}\psi$. We claim that the totality of such maps $\{\psi_A\}$ is an (X, G_1)-orbifold atlas for M.

Suppose that

$$B = \{h_0, \beta_1, \psi_1, h_1, \ldots, h_{n-1}, \beta_n, \psi_n, h_n\}$$

is an (X, G)-path over M from (x, ϕ) to (z, χ), where $\chi : W \to X/\mathrm{K}$, and let

$$h = h_0 \cdots h_n.$$

Suppose that gy' and hz' are points of X such that

$$\psi_B\psi_A^{-1}(g\mathrm{H}g^{-1}gy') = h\mathrm{K}h^{-1}hz'.$$

Then we have that

$$\chi\psi^{-1}(\mathrm{H}y') = \mathrm{K}z'.$$

Now as V is connected, there is a rectifiable curve $\overline{\gamma} : [0,1] \to X/\mathrm{H}$ from $\mathrm{H}y$ to $\mathrm{H}y'$ such that

$$\overline{\gamma}([0,1]) \subset \psi(V).$$

The curve $\overline{\gamma}$ lifts to a curve $\gamma : [0,1] \to X$ starting at y by Theorem 13.1.7. Let $C = \{1, \gamma, \psi, 1\}$ be the corresponding (X, G)-path over M from (y, ψ) to $(\gamma(1), \psi)$. By replacing A by AC and y by $\gamma(1)$, we may assume that $\mathrm{H}y = \mathrm{H}y'$. Likewise, we may assume that $\mathrm{K}z = \mathrm{K}z'$. Let e be an element of H such that $ey = y'$, and let k be an element of K such that $kz = z'$. Then e and k are in G_1 by Theorem 13.2.2.

Now since $\chi\psi^{-1}(\mathrm{H}y) = \mathrm{K}z$, there is an element f of G such that $fy = z$ and f lifts $\chi\psi^{-1}$ in a neighborhood of y. Let $J = \{1, \beta, \psi, f^{-1}\}$ be the constant (X, G)-path over M from (y, ψ) to (z, χ). Now (2) implies that $\eta([AJB^{-1}])$ is an element of G_1. Hence $gf^{-1}h^{-1}$ is an element of G_1. Observe that

$$hkfe^{-1}g^{-1} = (hkh^{-1})(hfg^{-1})(ge^{-1}g^{-1})$$

is in G_1,

$$(hkfe^{-1}g^{-1})(gy') = hkfe^{-1}y' = hkfy = hkz = hz',$$

and $hkfe^{-1}g^{-1}$ lifts $\chi_B\psi_A^{-1}$ in a neighborhood of gy'. Thus $\{\psi_A\}$ is an (X, G_1)-orbifold atlas for M. Moreover $\{\psi_A\}$ is obviously contained in the (X, G)-orbifold structure Φ of M. Now as $\phi_I = \phi$, we find that ϕ is in $\{\psi_A\}$. Thus (2) implies (3).

Now suppose that the (X, G)-orbifold structure Φ of M contains an (X, G)-orbifold structure Φ_1 for M containing ϕ. Let

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

be an (X,G)-path over M from (x,ϕ) to (x,ϕ) with partition $\{s_0,\ldots,s_m\}$ of $[0,1]$. We claim that $g_0\cdots g_m$ is in G_1. By subdivision, we may assume that there is a chart $\psi_i : V_i \to X/\mathrm{H}_i$ in Φ_1 such that

$$\alpha_i([s_{i-1},s_i]) \subset V_i \quad \text{for each } i=1,\ldots,m.$$

Hence, by translation, we may assume that $\phi_i=\psi_i$ for each i. Now since $\phi : U \to X/\Gamma$ is in Φ_1, there is an element h_0 of G_1 such that $h_0\alpha_1(0)=x$ and h_0 lifts $\phi\phi_1^{-1}$ in a neighborhood of $\alpha_1(0)$. Hence $g_0h_0^{-1}x = x$ and $g_0h_0^{-1}$ lifts $\phi\phi_1^{-1}(\phi_1\phi^{-1})$ in a neighborhood of x. Therefore $g_0h_0^{-1}$ is in the stabilizer Γ_x. Now Γ is a subgroup of G_1 by Theorem 13.2.2. Therefore g_0 is in G_1. Likewise $g_1,\ldots,g_m$ are in G_1. Hence $\eta([A]) = g_0\cdots g_m$ is in G_1. Thus, the image of η is contained in G_1 and so (3) implies (2). □

Theorem 13.3.6. *Let $\tilde{M}$ be the universal orbifold covering space based at (x,ϕ) of a connected (X,G)-orbifold M and let G_1 be the group of isometries in G. Suppose that $\pi_1^o(M,x,\phi)$ acts on $\tilde{M}$ via isometries. Then the (X,G)-orbifold structure Φ of M contains an (X,G_1)-orbifold structure Φ_1 for M containing ϕ, and if M together with Φ_1 is considered to be a metric (X,G_1)-orbifold, then the universal orbifold covering projection $\kappa : \tilde{M} \to M$ induces an isometry*

$$\overline{\kappa} : \tilde{M}/\pi_1^o(M,x,\phi) \to M.$$

Proof: The (X,G)-orbifold structure Φ of M contains an (X,G_1)-orbifold structure Φ_1 for M containing ϕ by Theorem 13.3.5. Consider M together with Φ_1 to be an (X,G_1)-orbifold. Let $\langle A\rangle$ be an arbitrary point of $\tilde{M}$ and suppose that

$$A = \{g_0,\alpha_1,\phi_1,g_1,\ldots,g_{m-1},\alpha_m,\phi_m,g_m\}$$

is an (X,G)-path over M from (x,ϕ) to (y,ψ), where $\psi : V \to X/\mathrm{H}$. Now let $\chi : W \to X/\mathrm{K}$ be in Φ_1 such that $\psi^{-1}(\mathrm{H}y)$ is in W. Let z be a point of X such that

$$\psi^{-1}(\mathrm{H}y) = \chi^{-1}(\mathrm{K}z).$$

Then there is a constant (X,G)-path $J = \{1,\beta,\psi,f\}$ over M from (y,ψ) to (z,χ). Now by replacing A by AJ and ψ by χ, we may assume that ψ is in Φ_1. Then the same argument as at the end of the proof of Theorem 13.3.5 shows that $g = g_0\cdots g_m$ is in G_1.

Let $r>0$ be such that

(1) $B(\mathrm{H}y,2r) \subset \psi(V)$,

(2) $r \le \frac{1}{4}\mathrm{dist}(y,\mathrm{H}y-\{y\})$,

(3) $r \le \frac{1}{4}\mathrm{dist}(\langle A\rangle,\pi_1^o(M)\langle A\rangle - \{\langle A\rangle\})$,

(4) $B(y,2r)$ is simply connected.

Now set
$$N = \psi^{-1}(B(\mathrm{H}y, r)).$$

Then N is an open neighborhood of $\psi^{-1}(\mathrm{H}y) = \kappa(\langle A\rangle)$ in M. By the argument in Lemma 6, the developing map $\delta : \tilde{M} \to X$ maps the set $\langle A, \psi^{-1}(B(\mathrm{H}y, 2r))\rangle$ homeomorphically onto the ball $B(gy, 2r)$. Hence, by Theorem 8.3.6, we have that
$$\langle A, N\rangle = B(\langle A\rangle, r)$$
and δ maps $B(\langle A\rangle, r)$ isometrically onto $B(gy, r)$.

Suppose that $[C]$ is in the stabilizer of $\langle A\rangle$. Then there is a constant (X, G)-path J over M from (y, ψ) to (y, ψ) such that
$$\begin{aligned}\delta([C]\langle A\rangle) &= \delta(\langle CA\rangle)\\ &= \delta([AJ])\\ &= \delta([AJA^{-1}A])\\ &= g\eta([J])g^{-1}\delta([A])\end{aligned}$$
with $\eta([J])$ in the stabilizer H_y. Hence δ induces an isometry $\overline{\delta}$ such that the following diagram commutes:
$$\begin{array}{ccc}B(\langle A\rangle, r) & \xrightarrow{\delta} & B(gy, r)\\ \downarrow & & \downarrow\\ B(\langle A\rangle, r)/\pi_1^o(M)_{\langle A\rangle} & \xrightarrow{\overline{\delta}} & B(gy, r)/g\mathrm{H}_y g^{-1}\\ \downarrow & & \downarrow\\ B(\pi_1^o(M)\langle A\rangle, r) & \xrightarrow{\overline{g}\psi\overline{\kappa}} & B(g\mathrm{H}g^{-1}gy, r),\end{array}$$
where the vertical maps are induced by quotient maps. Now by Theorem 13.1.1, the bottom vertical maps are isometries. Therefore $\overline{g}\psi\overline{\kappa}$ is an isometry. Observe that ψ maps $B(\kappa(\langle A\rangle), r)$ isometrically onto $B(\mathrm{H}y, r)$ by Theorem 13.2.8. Now as g is an isometry, the map
$$\overline{g} : X/\mathrm{H} \to X/g\mathrm{H}g^{-1}$$
is an isometry. Hence $\overline{g}$ maps $B(\mathrm{H}y, r)$ isometrically onto $B(g\mathrm{H}g^{-1}gy, r)$. Therefore $\overline{\kappa}$ maps $B(\pi_1^o(M)\langle A\rangle, r)$ isometrically onto $B(\kappa(\langle A\rangle), r)$. Thus $\overline{\kappa}$ is a local isometry.

Now as $\overline{\kappa} : \tilde{M}/\pi_1^o(M) \to M$ is a homeomorphism, $\overline{\kappa}$ induces an (X, G_1)-orbifold structure on $\tilde{M}/\pi_1^o(M)$. We claim that the orbit space metric d_π on $\tilde{M}/\pi_1^o(M)$ agrees with the induced (X, G_1)-orbifold metric d on $\tilde{M}/\pi_1^o(M)$. First of all, d_π and d agree locally, since $\overline{\kappa}$ is a local isometry; moreover, $d_\pi \leq d$, since arc length with respect to d_π is the same as X-length. On the contrary, suppose that $\langle A\rangle$ and $\langle B\rangle$ are points of $\tilde{M}$ such that
$$d_\pi(\pi_1^o(M)\langle A\rangle, \pi_1^o(M)\langle B\rangle) < d(\pi_1^o(M)\langle A\rangle, \pi_1^o(M)\langle B\rangle).$$

Then we have
$$\mathrm{dist}(\langle A\rangle, \pi_1^o(M)\langle B\rangle) < d(\pi_1^o(M)\langle A\rangle, \pi_1^o(M)\langle B\rangle).$$

Hence, there is an X-rectifiable curve $\gamma : [0,1] \to \tilde{M}$ from $\langle A \rangle$ to a point in $\pi_1^o(M)\langle B \rangle$ such that

$$\|\gamma\| < d(\pi_1^o(M)\langle A \rangle, \pi_1^o(M)\langle B \rangle).$$

Let $\varpi : \tilde{M} \to \tilde{M}/\pi_1^o(M)$ be the quotient map. Then $\|\varpi\gamma\| = \|\gamma\|$ by Theorem 13.1.4. Therefore, we have

$$\|\varpi\gamma\| < d(\pi_1^o(M)\langle A \rangle, \pi_1^o(M)\langle B \rangle),$$

which is a contradiction. Hence $d_\pi = d$. Thus $\overline{\kappa}$ is an isometry. □

Complete (X, G)-Orbifolds

We now define a notion of completeness for (X, G)-orbifolds.

Definition: An (X, G)-orbifold M is *complete* if and only if every universal orbifold covering space $\tilde{M}$ of M is a complete metric space.

Theorem 13.3.7. *Let M be a metric (X, G)-orbifold. Then M is complete if and only if M is a complete metric space.*

Proof: Suppose that M is complete. Let $\tilde{M}$ be the universal orbifold covering space of M based at (x, ϕ). Then $\tilde{M}$ is a complete metric space. Hence $\tilde{M}$ is geodesically complete by Theorem 8.5.7. Therefore, the developing map $\delta : \tilde{M} \to X$ is a covering projection by Theorem 8.5.6. Furthermore, the proof of Theorem 8.5.6 shows that there is an $r > 0$ such that $B(w, 2r)$ is evenly covered by δ for all w in X. Now δ maps $\overline{B}(\langle A \rangle, r)$ homeomorphically onto $\overline{B}(\delta(\langle A \rangle), r)$ for all $\langle A \rangle$ in $\tilde{M}$. Hence $\overline{B}(\langle A \rangle, r)$ is compact for all $\langle A \rangle$ in $\tilde{M}$. Now the quotient map

$$\varpi : \tilde{M} \to \tilde{M}/\pi_1^o(M, x, \phi)$$

maps $B(\langle A \rangle, r)$ onto $B(\varpi(\langle A \rangle), r)$ by Theorem 6.5.2. As $\overline{B}(\langle A \rangle, r)$ is compact, we deduce that

$$\varpi(\overline{B}(\langle A \rangle, r)) = \overline{B}(\varpi(\langle A \rangle), r).$$

Hence $\overline{B}(\varpi(\langle A \rangle), r)$ is compact for all $\langle A \rangle$ in $\tilde{M}$. Therefore $\tilde{M}/\pi_1^o(M, x, \phi)$ is a complete metric space by Theorem 8.5.1. Hence M is a complete metric space by Theorem 13.3.6.

Conversely, suppose that M is a complete metric space. Then we have that $\tilde{M}/\pi_1^o(M, x, \phi)$ is a complete metric space by Theorem 13.3.6. Hence $\tilde{M}$ is a complete metric space by Theorem 8.5.3. Thus M is complete. □

Definition: An (X, G)-orbifold structure Φ for a Hausdorff space M is *complete* if and only if M, with the (X, G)-orbifold structure Φ, is a complete (X, G)-orbifold.

Theorem 13.3.8. *Let M be an (X, G)-orbifold and let G_1 be the group of isometries in G. Then M is complete if and only if the (X, G)-orbifold structure of M contains a complete (X, G_1)-orbifold structure for M.*

Proof: Without loss of generality, we may assume that M is connected. Suppose that M is complete. Then the universal orbifold covering space $\tilde{M}$ of M based at (x, ϕ) is a complete metric space. Let $[C]$ be an element of $\pi_1^o(M, x, \phi)$. Then the map $[C]_* : \tilde{M} \to \tilde{M}$ is a similarity by Theorem 13.3.4. We claim that $[C]_*$ is an isometry. On the contrary, suppose that $[C]_*$ is not an isometry. Then $[C]_*$ has a fixed point $\langle A \rangle$ in $\tilde{M}$ by Theorem 8.5.4. Now by Theorem 13.3.4, the stabilizer of $\langle A \rangle$ is a finite group of isometries, which is a contradiction. Hence $[C]_*$ is an isometry. Thus $\pi_1^o(M, x, \phi)$ acts on $\tilde{M}$ via isometries. Therefore, by Theorem 13.3.5, the (X, G)-orbifold structure of M contains an (X, G_1)-orbifold structure for M containing ϕ. Consider M to be an (X, G_1)-orbifold with this structure. Then by Theorem 13.3.6, the universal orbifold covering projection $\kappa : \tilde{M} \to M$ induces an isometry

$$\overline{\kappa} : \tilde{M}/\pi_1^o(M, x, \phi) \to M.$$

The developing map $\delta : \tilde{M} \to X$ is a covering projection by Theorems 8.5.6 and 8.5.7. Hence, there is an $r > 0$ such that $\overline{B}(\langle A \rangle, r)$ is compact for all $\langle A \rangle$ in $\tilde{M}$. Therefore $\overline{B}(\pi_1^o(M)\langle A \rangle, r)$ is compact for all $\langle A \rangle$ in $\tilde{M}$. Hence $\tilde{M}/\pi_1^o(M)$ is a complete metric space by Theorem 8.5.1. Therefore M is a complete metric space. Hence M is a complete (X, G_1)-orbifold by Theorem 13.3.7. Thus, the (X, G)-orbifold structure of M contains a complete (X, G_1)-orbifold structure for M.

Conversely, suppose that the (X, G)-orbifold structure Φ of M contains a complete (X, G_1)-orbifold structure Φ_1 for M. Consider M together with Φ_1 to be an (X, G_1)-orbifold. Let ϕ be a chart in Φ_1 and let $\tilde{M}$ be the universal (X, G)-orbifold covering space of M based at (x, ϕ). Then by Theorems 13.3.5 and 13.3.6, the group $\pi_1^o(M, x, \phi)$ acts on $\tilde{M}$ via isometries, and the universal orbifold covering projection $\kappa : \tilde{M} \to M$ induces an isometry $\overline{\kappa} : \tilde{M}/\pi_1^o(M) \to M$. Now M is a complete metric space by Theorem 13.3.7. Hence $\tilde{M}/\pi_1^o(M)$ is a complete metric space. Therefore $\tilde{M}$ is a complete metric space by Theorem 8.5.3.

Now suppose that $\tilde{M}'$ is the (X, G)-orbifold covering space of M based at (y, ψ). Then there is an (X, G)-path A over M from (x, ϕ) to (y, ψ), since M is connected. Let $A_* : \tilde{M}' \to \tilde{M}$ be the change of base point homeomorphism defined by

$$A_*(\langle A' \rangle) = \langle AA' \rangle.$$

Suppose that

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

and let $g = g_0 \cdots g_m$. Then we have a commutative diagram

$$\begin{array}{ccc} \tilde{M}' & \xrightarrow{A_*} & \tilde{M} \\ \delta' \downarrow & & \downarrow \delta \\ X & \xrightarrow{g_*} & X, \end{array}$$

where the vertical maps are the developing maps. As g_* is a similarity, we deduce that A_* is a similarity. Hence $\tilde{M}'$ is a complete metric space. Thus M is complete. □

Definition: A function $\xi : M \to N$ between (X, G)-orbifolds is an (X, G)-*map* if and only if ξ is continuous and for each chart $\phi : U \to X/\Gamma$ for M and chart $\psi : V \to X/\mathrm{H}$ for N such that U and $\xi^{-1}(V)$ overlap, the function

$$\psi\xi\phi^{-1} : \phi(U \cap \xi^{-1}(V)) \to \psi(\xi(U) \cap V)$$

has the property that if x and y are points of X such that

$$\psi\xi\phi^{-1}(\Gamma x) = \mathrm{H}y,$$

then there is an element g of G such that $gx = y$ and g lifts $\psi\xi\phi^{-1}$ in a neighborhood of x.

Theorem 13.3.9. *An injection $\xi : M \to N$ between (X, G)-orbifolds is an (X, G)-map if and only if for each point u of M, there is a chart $\phi : U \to X/\Gamma$ for (M, u) such that ξ maps U homeomorphically onto an open subset of N and $\phi\xi^{-1} : \xi(U) \to X/\Gamma$ is a chart for N.*

Proof: Suppose that $\xi : M \to N$ is an (X, G)-map and u is an arbitrary point of M. Let $\psi : V \to X/\mathrm{H}$ be a chart for $(N, \xi(u))$. Since ξ is continuous at u, there is a chart $\phi : U \to X/\Gamma$ for (M, u) such that $\xi(U) \subset V$. Then

$$\psi\xi\phi^{-1} : \phi(U) \to \psi\xi(U)$$

lifts to an element of G on each component over $\phi(U)$. Hence $\psi\xi(U)$ is open in X/H, and so $\xi(U)$ is open in N. Therefore ξ is an open map. Hence ξ maps U homeomorphically onto $\xi(U)$.

Now consider the map

$$(\phi\xi^{-1})\psi^{-1} : \psi(\xi(U)) \to \phi\xi^{-1}(\xi(U))$$

and suppose that

$$\phi\xi^{-1}\psi^{-1}(\mathrm{H}y) = \Gamma x.$$

Then we have

$$\psi\xi\phi^{-1}(\Gamma x) = \mathrm{H}y.$$

Hence, there is an element g of G such that $gx = y$ and g lifts $\psi\xi\phi^{-1}$ in a neighborhood of x. Therefore $g^{-1}y = x$ and g^{-1} lifts $\phi\xi^{-1}\psi^{-1}$ in a neighborhood of y. As $\xi(U) \subset V$ and $\psi : V \to X/\mathrm{H}$ is a chart for N, we deduce that $\phi\xi^{-1} : \xi(U) \to X/\Gamma$ is a chart for N.

Conversely, suppose that for each point u of M, there is a chart $\phi : U \to X/\Gamma$ for (M, u) such that ξ maps U homeomorphically onto an open subset of N and $\phi\xi^{-1} : \xi(U) \to X/\Gamma$ is a chart for N. Then ξ is continuous. Let $\chi : W \to X/\mathrm{K}$ and $\psi : V \to X/\mathrm{H}$ be charts for M and N, respectively, such that W and $\xi^{-1}(V)$ overlap. Now let u be an arbitrary point of the set $W \cap \xi^{-1}(V)$. Then there is a chart $\phi : U \to X/\Gamma$ for (M, u) such that ξ maps U homeomorphically onto an open subset of N and $\phi\xi^{-1} : \xi(U) \to X/\Gamma$ is a chart for N. Consider the function

$$\psi\xi\chi^{-1} : \chi(W \cap \xi^{-1}(V)) \to \psi(\xi(W) \cap V).$$

Suppose that y and z are points of X such that

$$\psi\xi\chi^{-1}(\mathrm{K}z) = \psi\xi(u) = \mathrm{H}y.$$

Now since

$$\psi\xi\chi^{-1} = (\psi\xi\phi^{-1})(\phi\chi^{-1})$$

and $\phi\chi^{-1}$ and $\psi(\phi\xi^{-1})^{-1}$ are coordinate changes for M and N, respectively, there is an element h of G such that $hz = y$ and h lifts $\psi\xi\chi^{-1}$ in a neighborhood of z. Thus ξ is an (X, G)-map. □

Definition: A function $\xi : M \to N$ between (X, G)-orbifolds is an (X, G)-*equivalence* if and only if ξ is a bijective (X, G)-map.

Note that the inverse of an (X, G)-equivalence is also an (X, G)-equivalence. Two (X, G)-orbifolds M and N are said to be (X, G)-*equivalent* if and only if there is an (X, G)-equivalence $\xi : M \to N$. Note that an (X, G)-equivalence $\xi : M \to N$ between metric (X, G)-orbifolds is an isometry.

Theorem 13.3.10. *Let G be a group of similarities of a simply connected geometric space X and let M be a complete connected (X, G)-orbifold. Let $\eta : \pi_1^o(M) \to G$ be a holonomy of M and let $\delta : \tilde{M} \to X$ be the corresponding developing map. Then δ is an $(X, \{1\})$-equivalence, η maps $\pi_1^o(M)$ isomorphically onto a discrete group Γ of isometries of X, and δ induces an (X, G)-equivalence from M to X/Γ.*

Proof: Now $\delta : \tilde{M} \to X$ is a covering projection by Theorems 8.5.6 and 8.5.7. Therefore δ is a homeomorphism, since X is simply connected. Hence δ is an $(X, \{1\})$-equivalence and so is an isometry. Now $\pi_1^o(M)$ corresponds to the group of covering transformations of the universal orbifold covering projection $\kappa : \tilde{M} \to M$ which corresponds via δ to the image of η. By Theorems 13.3.4, 13.3.5, and 13.3.8, the group $\pi_1^o(M)$ acts discontinuously on $\tilde{M}$ via isometries. Therefore η maps $\pi_1^o(M)$ isomorphically onto a discrete group Γ of isometries of X. By Theorem 13.3.4, we deduce that δ induces a homeomorphism $\overline{\delta}$ such that the following diagram commutes:

$$\begin{array}{ccc} \tilde{M} & \xrightarrow{\delta} & X \\ \kappa \downarrow & & \downarrow \pi \\ M & \xrightarrow{\overline{\delta}} & X/\Gamma, \end{array}$$

where π is the quotient map.

We claim that $\overline{\delta} : M \to X/\Gamma$ is a chart for M. Let $\psi : V \to X/\mathrm{H}$ be a chart for M and let y and z be points of X such that

$$\overline{\kappa}\psi^{-1}(\mathrm{H}y) = \Gamma z.$$

Now since $\kappa : \tilde{M} \to M$ is onto, there is an (X, G)-path

$$A = \{g_0, \alpha_1, \phi_1, g_1, \ldots, g_{m-1}, \alpha_m, \phi_m, g_m\}$$

over M from (x, ϕ) to (y, ψ). Let $g = g_0 \cdots g_m$. Then g is in G and

$$\begin{aligned} \overline{\delta}\psi^{-1}(\mathrm{H}y) &= \overline{\delta}\kappa(\langle A \rangle) \\ &= \pi\delta(\langle A \rangle) \\ &= \pi(gy) \quad = \quad \Gamma g y. \end{aligned}$$

Hence, there is an element f of Γ such that $fgy = z$. Let $r > 0$ such that

$$B(\mathrm{H}y, r) \subset \psi(V).$$

Suppose that $y' \neq y$ is in $B(y, r)$. Then there is a rescaled geodesic arc $\beta : [0, 1] \to X$ from y to y', and $\{1, \beta, \psi, 1\}$ is an (X, G)-path over M from (y, ψ) to (y', ψ). Observe that

$$\begin{aligned} \overline{\delta}\psi^{-1}(\mathrm{H}y') &= \overline{\delta}\kappa(\langle AB \rangle) \\ &= \pi\delta(\langle AB \rangle) \\ &= \pi(gy') \quad = \quad \Gamma g y'. \end{aligned}$$

Hence fg lifts $\overline{\delta}\psi^{-1}$ on $B(y, r)$. Thus $\overline{\delta} : M \to X/\Gamma$ is a chart for M. It now follows from Theorem 13.3.9, with $U = M$, that $\overline{\delta} : M \to X/\Gamma$ is an (X, G)-equivalence. □

Exercise 13.3

1. Let M be a connected (X, G)-orbifold. Prove that there is an (X, G)-path over M from any (x, ϕ) to any (y, ψ).

2. Let Γ be a discrete group of isometries of a geometric space X and let $\iota : X/\Gamma \to X/\Gamma$ be the identity map. Define a function $\zeta : \pi_1(X, x) \to \pi_1^o(X/\Gamma, x, \iota)$ by $\zeta([\alpha]) = [\{1, \alpha, \iota, 1\}]$. Prove that ζ is a homomorphism and that the following sequence is exact:

$$1 \longrightarrow \pi_1(X, x) \xrightarrow{\zeta} \pi_1^o(X/\Gamma, x, \iota) \xrightarrow{\eta} \Gamma \longrightarrow 1.$$

3. Let $\tilde{M}$ be the universal orbifold covering space based at (x, ϕ) of an (X, G)-orbifold M and let $\kappa : \tilde{M} \to M$ be the universal orbifold covering projection. Let A be an (X, G)-path over M from (x, ϕ) to (y, ψ) and let N be an open neighborhood of $\kappa(\langle A \rangle)$ in M. Prove that $\kappa(\langle A, N \rangle)$ is the component of N containing $\kappa(\langle A \rangle)$.

4. Let $\kappa : \tilde{M} \to M$ be as in Exercise 3 with M connected. Prove that κ restricts to a covering projection $\kappa_1 : \kappa^{-1}(\Omega(M)) \to \Omega(M)$.

5. Prove that a connected (X, G)-orbifold M is complete if and only if every (or some) developing map $\delta : \tilde{M} \to X$ for M is a covering projection.

§13.4. Gluing Orbifolds

In this section, we shall construct n-dimensional spherical, Euclidean, and hyperbolic orbifolds by gluing together n-dimensional convex polyhedra. Let $X = S^n, E^n$, or H^n with $n > 0$.

Definition: A *disjoint set of n-dimensional convex polyhedra* of X is a set of functions

$$\Xi = \{\xi_P : P \in \mathcal{P}\}$$

indexed by a set $\mathcal{P}$ such that

(1) the function $\xi_P : X \to X_P$ is a similarity for each P in $\mathcal{P}$;

(2) the index P is an n-dimensional convex polyhedron in X_P for each P in $\mathcal{P}$;

(3) the polyhedra in $\mathcal{P}$ are mutually disjoint.

Let Ξ be a disjoint set of n-dimensional convex polyhedra of X and let G be a group of similarities of X.

Definition: A *G-side-pairing* for Ξ is a set of functions

$$\Phi = \{\phi_S : S \in \mathcal{S}\}$$

indexed by the collection $\mathcal{S}$ of all the sides of the polyhedra in $\mathcal{P}$ such that for each side S of a polyhedron P in $\mathcal{P}$

(1) there is a polyhedron P' in $\mathcal{P}$ such that the function $\phi_S : X_{P'} \to X_P$ is a similarity;

(2) the similarity $g_S = \xi_P^{-1}\phi_S\xi_{P'}$ is in G;

(3) there is a side S' of P' such that $\phi_S(S') = S$;

(4) the similarities ϕ_S and $\phi_{S'}$ satisfy the relation $\phi_{S'} = \phi_S^{-1}$;

(5) the polyhedrons P and $\phi_S(P')$ are situated so that $P \cap \phi_S(P') = S$.

Let Φ be a G-side-pairing for Ξ. The pairing of side points by elements of Φ generates an equivalence relation on the set $\Pi = \cup_{P\in\mathcal{P}} P$ whose equivalence classes are called the *cycles* of Φ. Topologize Π with the direct sum topology and let M be the quotient space of Π of cycles. The space M is said to be obtained by gluing together the polyhedra of Ξ by Φ.

The cycle of a point x of Π is denoted by $[x]$. Recall that a *ridge* of a polyhedron P is a side of a side of P. If x is in the interior of a ridge of a polyhedron in $\mathcal{P}$, then every point of $[x]$ is in the interior of a ridge of a polyhedron in $\mathcal{P}$, in which case $[x]$ is called a *ridge cycle* of Φ.

Let $[x] = \{x_1, \ldots, x_m\}$ be a finite ridge cycle of Φ and let P_i be the polyhedron in $\mathcal{P}$ containing x_i for each i. The point x_i is in exactly two sides of P_i. Hence x_i is paired to at most two other points of $[x]$ for each i. Therefore, we can reindex $\{x_1, \ldots, x_m\}$ so that

$$x_1 \simeq x_2 \simeq \cdots \simeq x_m.$$

The ridge cycle $[x]$ is said to be *dihedral* if there is a side S of P_1 containing x_1 such that ϕ_S is the reflection of X_{P_1} in $\langle S \rangle$ and there is a side T of P_m containing x_m such that ϕ_T is the reflection of X_{P_m} in $\langle T \rangle$, otherwise $[x]$ is said to be *cyclic*. Let θ_i be the dihedral angle of P_i along the ridge containing x_i for each i. The *dihedral angle sum* of the ridge cycle $[x]$ is defined to be

$$\theta[x] = \theta_1 + \cdots + \theta_m.$$

Definition: A G-side-pairing Φ for Ξ is *subproper* if and only if each cycle of Φ is finite, each dihedral ridge cycle of Φ has dihedral angle sum a submultiple of π, and each cyclic ridge cycle has dihedral angle sum a submultiple of 2π.

Theorem 13.4.1. *Let G be a group of similarities of X and let M be a space obtained by gluing together a disjoint set Ξ of n-dimensional convex polyhedra of X by a subproper G-side-pairing Φ. Then M is an (X, G)-orbifold such that the natural injection of P° into M is an (X, G)-map for each polyhedron P of Ξ.*

Proof: The proof is by induction on the dimension n. In order to simplify notation, we shall assume that G is a group of isometries of X and leave the proof of the general case to the reader. This restriction only affects the Euclidean case of the theorem. By changing the scale of X_P for each P in $\mathcal{P}$, we may assume that each $\xi_P : X \to X_P$ in Ξ is an isometry. In order to simplify the notation, we shall further assume that $X_P = X$ and $\xi_P = 1$ for each P in $\mathcal{P}$ and leave the proof of the general case to the reader.

Let x a point of Π and let $[x] = \{x_1, \ldots, x_m\}$. Let P_i be the polyhedron in $\mathcal{P}$ containing x_i for each i and let $\delta(x)$ be the minimum of π, the distance from x_i to x_j for each $i \neq j$, and the distance from x_i to any side of P_i not containing x_i for each i.

Let r be a real number such that $0 < r < \delta(x)/2$. Then for each i, the set $P_i \cap S(x_i, r)$ is a spherical $(n-1)$-dimensional polyhedron in the sphere $S(x_i, r)$, and the polyhedra $\{P_i \cap S(x_i, r)\}$ are disjoint. Observe that the side-pairing Φ restricts to a subproper $\mathrm{I}(S^{n-1})$-side-pairing of the polyhedra $\{P_i \cap S(x_i, r)\}$. Let $\Sigma(x, r)$ be the space obtained by gluing together the polyhedra $\{P_i \cap S(x_i, r)\}$. Then $\Sigma(x, r)$ has a spherical $(n-1)$-orbifold structure by inspection if $n = 1, 2$, or by induction if $n > 2$. Moreover $\Sigma(x, r)$ is compact, since $[x]$ is a finite cycle. Therefore $\Sigma(x, r)$ is a complete spherical orbifold by Theorem 13.3.7. Furthermore $\Sigma(x, r)$ is connected if $n > 1$.

Now by inspection if $n = 1$, or since Φ is subproper if $n = 2$, or by Theorem 13.3.10 if $n > 2$, there is, for each i, a finite subgroup Γ_i of G that fixes the point x_i such that the restriction of the quotient map $\pi : \Pi \to M$ to the polyhedron $P_i \cap S(x_i, r)$ extends to a continuous function

$$\kappa_i : S(x_i, r) \to \Sigma(x, r)$$

such that κ_i induces an isometry

$$\overline{\kappa}_i : S(x_i, r)/\Gamma_i \to \Sigma(x, r).$$

Moreover Γ_i does not depend on the choice of r. Let

$$\pi_i : S(x_i, r) \to S(x_i, r)/\Gamma_i$$

be the quotient map. Then we have $\kappa_i = \overline{\kappa}_i\pi_i$. For each i, j, the isometry

$$\overline{\kappa}_j^{-1}\overline{\kappa}_i : S(x_i, r)/\Gamma_i \to S(x_j, r)/\Gamma_j$$

lifts to an isometry

$$\xi_{ij} : S(x_i, r) \to S(x_j, r)$$

such that

$$\begin{aligned} \kappa_j\xi_{ij} &= \overline{\kappa}_j\pi_j\xi_{ij} \\ &= \overline{\kappa}_j\overline{\kappa}_j^{-1}\overline{\kappa}_i\pi_i \\ &= \overline{\kappa}_i\pi_i \quad = \quad \kappa_i. \end{aligned}$$

Moreover ξ_{ij} is unique up to left multiplication by the restriction of an element of Γ_j by Theorem 13.1.2. The isometry ξ_{ij} extends to an isometry g_{ij} of X that is unique up to left multiplication by an element of Γ_j. We may assume that $g_{ii} = 1$ for each i.

Suppose that the element g_S of Φ pairs the side $S' \cap S(x_i, r)$ of the polyhedron $P_i \cap S(x_i, r)$ to the side $S \cap S(x_j, r)$ of $P_j \cap S(x_j, r)$. Then g_S restricts to an isometry

$$\overline{g}_S : S(x_i, r) \to S(x_j, r).$$

Observe that κ_i agrees with $\kappa_j\overline{g}_S$ on the open set

$$U_S = \left(P_i^\circ \cup (S')^\circ \cup g_S^{-1}(P_j^\circ)\right) \cap S(x_i, r).$$

Hence, on the open set $\xi_{ij}(U_S)$, the map $\kappa_j\overline{g}_S\xi_{ij}^{-1}$ agrees with $\kappa_i\xi_{ij}^{-1} = \kappa_j$. Therefore $\overline{g}_S\xi_{ij}^{-1}$ is the restriction of an element of Γ_j by Theorem 13.1.2. Hence $g_S g_{ij}^{-1}$ is in Γ_j, and so we may assume that $g_{ij} = g_S$. If $i = j$, then the assumption that $g_{ij} = g_S$ will conflict with the previous assumption that $g_{ii} = 1$, but this will not matter, since we only need to specify g_{ij} up to left multiplication by an element of Γ_j, and in this case g_S is in Γ_j.

Now suppose that

$$x_i = x_{i_1} \simeq x_{i_2} \simeq \cdots \simeq x_{i_p} = x_j.$$

Then we have

$$\begin{aligned}\kappa_j\xi_{i_{p-1}i_p}\xi_{i_{p-2}i_{p-1}}\cdots\xi_{i_1i_2} &= \kappa_{i_{p-1}}\xi_{i_{p-2}i_{p-1}}\cdots\xi_{i_1i_2}\\ &\ \ \vdots\\ &= \kappa_{i_2}\xi_{i_1i_2} = \kappa_i.\end{aligned}$$

Hence, we deduce that

$$\xi_{ij}(\xi_{i_{p-1}i_p}\xi_{i_{p-2}i_{p-1}}\cdots\xi_{i_1i_2})^{-1}$$

is the restriction of an element of Γ_j. Therefore, we have that

$$g_{ij}(g_{i_{p-1}i_p}g_{i_{p-2}i_{p-1}}\cdots g_{i_1i_2})^{-1}$$

is an element of Γ_j. Hence, we may assume that

$$g_{ij} = g_{i_{p-1}i_p}g_{i_{p-2}i_{p-1}}\cdots g_{i_1i_2}.$$

Define

$$U(x,r) = \bigcup_{i=1}^{m}\pi(P_i\cap B(x_i,r)).$$

Since the set

$$\pi^{-1}(U(x,r)) = \bigcup_{i=1}^{m}P_i\cap B(x_i,r)$$

is open in Π, we have that $U(x,r)$ is an open subset of M.

Suppose that $x = x_k$ and let $\Gamma_x = \Gamma_k$. Define a function

$$\psi_x : \bigcup_{i=1}^{m}P_i\cap B(x_i,r)\to B(x,r)/\Gamma_x$$

by the rule

$$\psi_x(z) = \Gamma_x g_{ik}(z) \ \text{ if } z \text{ is in } P_i\cap B(x_i,r).$$

Suppose that $g_S(x_i) = x_j$. Then we may assume that $g_{ij} = g_S$. Let y be a point of $S\cap B(x_j,r)$ and let $y' = g_S^{-1}(y)$. Then y' is a point of $S'\cap B(x_i,r)$. Observe that

$$\kappa_k\xi_{jk}\xi_{ij} = \kappa_j\xi_{ij} = \kappa_i = \kappa_k\xi_{ik}.$$

Therefore, we have that

$$\xi_{ik}(\xi_{jk}\xi_{ij})^{-1}$$

is the restriction of an element of Γ_x. Hence, we have that

$$g_{ik}(g_{jk}g_{ij})^{-1}$$

is an element of Γ_x. Therefore, we have

$$\begin{aligned}\psi_x(y) &= \Gamma_x g_{jk}(y)\\ &= \Gamma_x g_{jk}g_S(y')\\ &= \Gamma_x g_{jk}g_{ij}(y')\\ &= \Gamma_x g_{ik}(y') = \psi_x(y').\end{aligned}$$

Consequently ψ_x induces a continuous function

$$\phi_x : U(x,r)\to B(x,r)/\Gamma_x.$$

For each t such that $0 < t < r$, the function ϕ_x restricts to a map

$$\overline{\phi}_x : \Sigma(x,t) \to S(x,t)/\Gamma_x.$$

Let z be a point of $P_i \cap S(x_i,t)$. Then we have

$$\begin{aligned} \overline{\phi}_x \pi(z) &= \psi_x(z) \\ &= \pi_k \xi_{ik}(z) \\ &= \overline{\kappa}_k^{-1} \overline{\kappa}_i \pi_i(z) \\ &= \overline{\kappa}_k^{-1} \kappa_i(z) \;=\; \overline{\kappa}_k^{-1} \pi(z). \end{aligned}$$

Therefore $\overline{\phi}_x = \overline{\kappa}_k^{-1}$. Hence $\overline{\phi}_x$ is an isometry. Consequently ϕ_x is a bijection with a continuous inverse defined by the rule

$$\phi_x^{-1}(\Gamma_x z) = \pi g_{ik}^{-1}(z) \ \text{ if } z \text{ is in } g_{ik}(P_i \cap B(x_i,r)).$$

Hence ϕ_x is a homeomorphism. The same argument as in the proof of Theorem 9.2.2 shows that M is Hausdorff.

Next, we show that

$$\{\phi_x : U(x,r) \to B(x,r)/\Gamma_x \mid x \text{ is in } \Pi \text{ and } r < \delta(x)/4\}$$

is an (X,G)-atlas for M. By construction, $U(x,r)$ is an open connected subset of M and ϕ_x is a homeomorphism. Moreover $U(x,r)$ is defined for each point $\pi(x)$ of M and sufficiently small radius r. Consequently $\{U(x,r)\}$ is an open cover of M.

Suppose that the sets $U(x,r)$ and $U(y,s)$ overlap and $r < \delta(x)/4$ and $s < \delta(y)/4$. Let w be in $B(x,r)$ and z be in $B(y,s)$ such that

$$\phi_y \phi_x^{-1}(\Gamma_x w) = \Gamma_y z.$$

We need to find an element g of G such that $gw = z$ and g lifts $\phi_y\phi_x^{-1}$ in a neighborhood of w.

Let $F(x)$ be the face of the polyhedron in $\mathcal{P}$ that contains x in its interior. By reversing the roles of x and y, if necessary, we may assume that

$$\dim F(x) \geq \dim F(y)$$

with equality only if $r \leq s$. As before, we have

$$\begin{aligned} \pi^{-1}(U(x,r)) &= \bigcup_{i=1}^{m} P_i \cap B(x_i,r), \\ \pi^{-1}(U(y,s)) &= \bigcup_{j=1}^{n} Q_j \cap B(y_j,s). \end{aligned}$$

Now for some i and j, the set $P_i \cap B(x_i,r)$ meets $Q_j \cap B(y_j,s)$. Then $P_i = Q_j$ and $d(x_i,y_j) < r + s$. We claim that y_j is in every side of P_i that contains x_i. On the contrary, suppose that y_j is not in a side of P_i that contains x_i. Then $s < d(x_i,y_j)/4$. Therefore x_i is in every side of P_i that contains y_j, otherwise we would have the contradiction that $r < d(x_i,y_j)/4$. Hence $F(x_i)$ is a proper face of $F(y_j)$, which is a contradiction. Therefore y_j is in every side of P_i that contains x_i. This implies that for each i, the set $P_i \cap B(x_i,r)$ meets $Q_j \cap B(y_j,s)$ for some j.

We claim that the set $P_i \cap B(x_i, r)$ meets $Q_j \cap B(y_j, s)$ for only one index j. On the contrary, suppose that $P_i \cap B(x_i, r)$ meets $Q_j \cap B(y_j, s)$ and $Q_k \cap B(y_k, s)$. Then $P_i = Q_j = Q_k$. Now since y_j and y_k are in every side of P_i that contains x_i, we have that $F(y_j)$ and $F(y_k)$ are faces of $F(x_i)$.

Assume first that

$$\dim F(x) > \dim F(y).$$

Then $F(y_j)$ and $F(y_k)$ are proper faces of $F(x_i)$. Consequently, we have

$$r < d(x_i, y_j)/4, \quad r < d(x_i, y_k)/4, \quad \text{and} \quad s < d(y_j, y_k)/4,$$

which leads to the contradiction

$$\begin{aligned} d(x_i, y_j) + d(x_i, y_k) &< (r+s) + (r+s) \\ &< d(x_i, y_j)/4 + d(x_i, y_k)/4 + 2d(y_j, y_k)/4 \\ &< d(x_i, y_j) + d(x_i, y_k). \end{aligned}$$

Now assume that

$$\dim F(x) = \dim F(y).$$

Then $r \leq s$. Observe that

$$s < d(y_j, y_k)/4 \leq (d(x_i, y_j) + d(x_i, y_k))/4 < 2(r+s)/4$$

and so $s < r$, which is a contradiction. Therefore $P_i \cap B(x_i, r)$ meets $Q_j \cap B(y_j, s)$ for only one index $j = i'$.

Let g_{ij} and h_{ij} be the elements of G constructed as before for x and y. Suppose that g_S pairs the side $S' \cap S(x_i, r)$ of $P_i \cap S(x_i, r)$ to the side $S \cap S(x_j, r)$ of $P_j \cap S(x_j, r)$. Then we may assume that $g_{ij} = g_S$. Now $g_S(x_i) = x_j$, and so x_i is in S'. As $P_i \cap B(x_i, r)$ meets $P_i \cap B(y_{i'}, s)$, we have that $y_{i'}$ is also in S'. Now observe that $g_S(P_i \cap B(x_i, r))$ meets $g_S(P_i \cap B(y_{i'}, s))$. Hence $P_j \cap B(x_j, r)$ meets $P_j \cap B(g_S y_{i'}, s)$. Therefore $g_S y_{i'} = y_{j'}$. Hence, we may assume that $g_{ij} = h_{i'j'}$.

Now suppose that

$$x_i = x_{i_1} \simeq x_{i_2} \simeq \cdots \simeq x_{i_p} = x_j.$$

Then we deduce from the previous argument that

$$y_{i'} = y_{i'_1} \simeq y_{i'_2} \simeq \cdots \simeq y_{i'_p} = y_{j'}$$

and so we may assume that

$$\begin{aligned} g_{ij} &= g_{i_{p-1}i_p} g_{i_{p-2}i_{p-1}} \cdots g_{i_1 i_2} \\ &= h_{i'_{p-1}i'_p} h_{i'_{p-2}i'_{p-1}} \cdots h_{i'_1 i'_2} \quad = \quad h_{i'j'}. \end{aligned}$$

Next, observe that

$$\begin{aligned} &U(x,r) \cap U(y,s) \\ &= \pi\Big(\bigcup_{i=1}^{m} P_i \cap B(x_i, r)\Big) \cap \pi\Big(\bigcup_{j=1}^{n} Q_j \cap B(y_j, s)\Big) \\ &= \pi\Big(\Big[\bigcup_{i=1}^{m} P_i \cap B(x_i, r)\Big] \cap \Big[\bigcup_{j=1}^{n} Q_j \cap B(y_j, s)\Big]\Big) \\ &= \pi\Big(\bigcup_{i=1}^{m} \bigcup_{j=1}^{n} \big[P_i \cap B(x_i, r) \cap Q_j \cap B(y_j, s)\big]\Big) \\ &= \pi\Big(\bigcup_{i=1}^{m} P_i \cap B(x_i, r) \cap B(y_{i'}, s)\Big). \end{aligned}$$

Let $x = x_k$ and $y = y_\ell$. Then

$$\phi_x\big(U(x,r)\cap U(y,s)\big) = \bigcup_{\imath=1}^{m} \Gamma_x g_{\imath k}\big(P_\imath \cap B(x_\imath, r)\cap B(y_{\imath'}, s)\big)$$

and

$$\phi_y\big(U(x,r)\cap U(y,s)\big) = \bigcup_{\imath=1}^{m} \Gamma_y h_{\imath'\ell}\big(P_\imath \cap B(x_\imath, r)\cap B(y_{\imath'}, s)\big).$$

Now if v is a point of the set

$$g_{\imath k}\big(P_\imath \cap B(x_\imath, r)\cap B(y_{\imath'}, s)\big),$$

then we have

$$\begin{aligned}\phi_y\phi_x^{-1}(\Gamma_x v) &= \phi_y(\pi(g_{\imath k}^{-1}v))\\ &= \Gamma_y h_{\imath'\ell}g_{\imath k}^{-1}v\\ &= \Gamma_y h_{\imath'\ell}h_{\imath'k'}^{-1}v\\ &= \Gamma_y h_{\imath'\ell}h_{k'\imath'}v \quad = \quad \Gamma_y h_{k'\ell}v.\end{aligned}$$

Therefore, the element $h_{k'\ell}$ lifts $\phi_y\phi_x^{-1}$. Hence, there is an element f of Γ_y such that $fh_{k'\ell}w = z$. Let $g = fh_{k'\ell}$. Then g is an element of G such that $gw = z$ and g lifts $\phi_y\phi_x^{-1}$ in a neighborhood of w. This completes the proof that $\{\phi_x\}$ is an (X,G)-atlas for M.

The same argument as in the proof of Theorem 9.2.2 shows that the (X,G)-structure of M has the property that the natural injection map of P° into M is an (X,G)-map for each P in $\mathcal{P}$. □

Example 1. Let $\triangle$ be a triangle in S^2, E^2, or H^2 with angles $\alpha, \alpha, 2\pi/3$ at its vertices x, y, z, respectively. See Figure 13.4.1 Let $L = [x,z]$, $R = [y,z]$, $S = [x,y]$ be the sides of $\triangle$. Pair side L to side R by the rotation g_R about z of $2\pi/3$, pair side R to side L by $g_L = g_R^{-1}$, and pair side S to itself by the reflection g_S in the line $\langle S\rangle$. Consider the side-pairing $\Phi = \{g_L, g_R, g_S\}$. The point z forms a cyclic ridge cycle whose angle sum is $2\pi/3$. The points x and y form a dihedral ridge cycle whose angle sum is 2α.

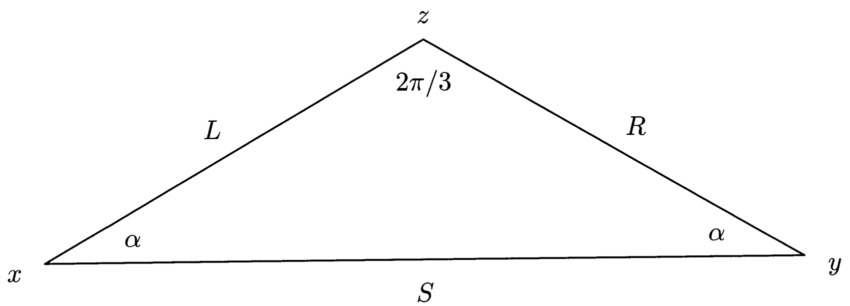


Figure 13.4.1. A triangle in S^2, E^2, or H^2

Assume that Φ is subproper. Then there is a positive integer k such that $2\alpha = \pi/k$. Observe that the angle sum of $\triangle$ is

$$\frac{2\pi}{3} + 2\alpha = \frac{2\pi}{3} + \frac{\pi}{k},$$

which is greater than, equal to, or less than π, according as k is less than, equal to, or greater than three. Thus $\triangle$ is spherical if $\alpha = \pi/2, \pi/4$, Euclidean if $\alpha = \pi/6$, and hyperbolic if $\alpha = \pi/2k$ with $k > 3$.

Let M be the space obtained from $\triangle$ by gluing together its sides according to Φ. Then by Theorem 13.4.1, we have that M is a 2-dimensional orbifold that is spherical if $\alpha = \pi/2, \pi/4$, Euclidean if $\alpha = \pi/6$, and hyperbolic if $\alpha = \pi/2k$ with $k > 3$. Topologically, M is a disk. The singular set of M consists of a point of order 3 in the interior of M, corresponding to z, and the boundary of M, which consists of a point of order $2k$, corresponding to $\{x, y\}$, and an open edge of points of order 2, corresponding to S°.

Example 2. Let Q be a quadrilateral in E^2 whose vertices are in cyclic order w, x, y, z, and whose angles are $\alpha, \alpha, \beta, \beta$, respectively. See Figure 13.4.2. As $2\alpha + 2\beta = 2\pi$, we have that $\alpha + \beta = \pi$. Let $S = [w, x]$, $R = [x, y]$, $T = [y, z]$, $L = [z, w]$. Then the sides S and T are parallel. Pair side T to side S by the composition g_S of the vertical translation from T to S followed by a change of scale, pair side S to side T by $g_T = g_S^{-1}$, pair side L to itself by the reflection g_L in the line $\langle L \rangle$, and pair side R to itself by the reflection g_R in the line $\langle R \rangle$. Consider the side-pairing $\Phi = \{g_L, g_R, g_S, g_T\}$. Then $\{w, z\}$ and $\{x, y\}$ are dihedral ridge cycles whose angle sum is π. Therefore Φ is subproper.

Let M be the space obtained from Q by gluing together its sides according to Φ. Then M is a Euclidean similarity 2-orbifold by Theorem 13.4.1. Topologically, M is a cylinder. The singular set of M is its boundary and all the singular points of M have order two.

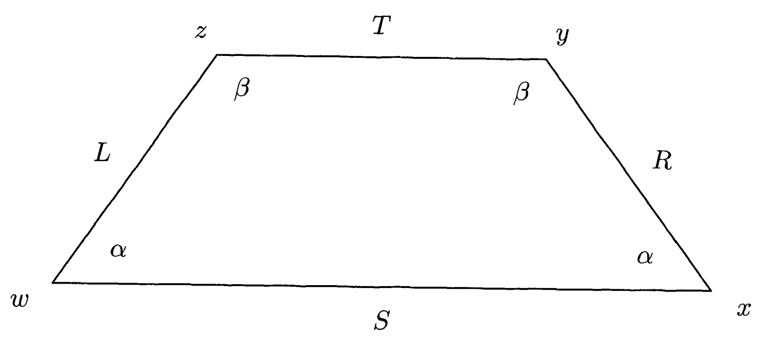


Figure 13.4.2. A quadrilateral in E^2

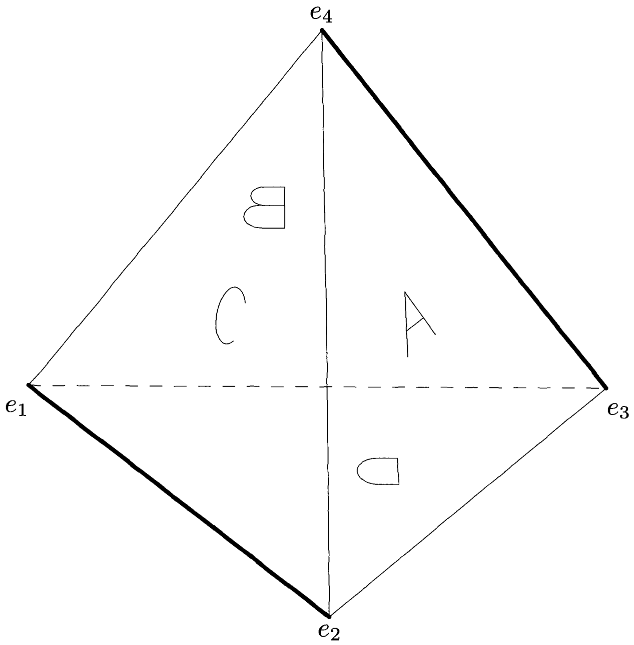


Figure 13.4.3. A right-angled regular tetrahedron in S^3

Example 3. Let P be the regular spherical tetrahedron in S^3 whose vertices are the vectors e_1, e_2, e_3, e_4. All the proper dihedral angles of P are $\pi/2$. Let A, B, C, D be the side of P opposite the vertex e_1, e_2, e_3, e_4, respectively. See Figure 13.4.3. Pair the side B to the side A by a rotation g_A of $\pi/2$ about their common edge $[e_3, e_4]$. Pair the side A to the side B by $g_B = g_A^{-1}$. Pair the side D to the side C by a rotation C of $\pi/2$ about their common edge $[e_1, e_2]$. Pair the side C to the side D by $g_D = g_C^{-1}$. Consider the side-pairing $\Phi = \{g_A, g_B, g_C, g_D\}$. Observe that each point on the open edges (e_1, e_2) and (e_3, e_4) forms a ridge cycle whose dihedral angle sum is $\pi/2$. All the remaining interior edge points of P fall into ridge cycles whose dihedral angle sum is 2π. Therefore Φ is subproper.

Let M be the space obtained from P by gluing together its sides according to Φ. Then M is a spherical 3-orbifold by Theorem 13.4.1. Topologically, M is a 3-sphere. This can be seen by first gluing side A to side B. This yields a 3-ball with the edge $[e_1, e_2]$ glued together at its ends to form the equator of the ball. The edge $[e_3, e_4]$ becomes the north-south diameter of the ball. The sides C and D become the northern and southern hemispheres of the ball. Now gluing the northern and southern hemispheres by a rotation about the equator yields a 3-sphere. The north-south diameter of the ball glues together at its ends to form a circle that simply links the equator. The singular set of M is therefore two simply linked circles, and all the singular points of M have order four.

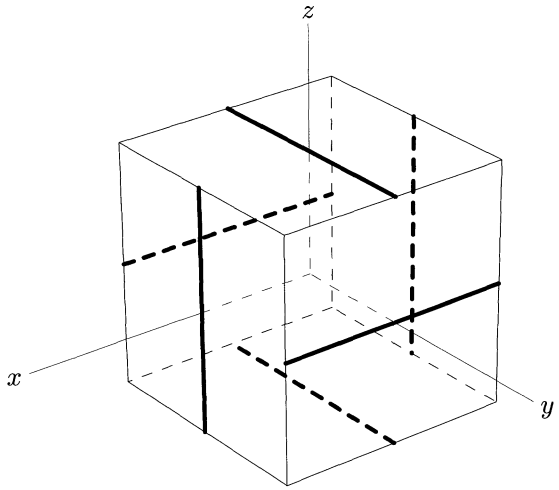


Figure 13.4.4. A cube in E^3

Example 4. Let P be the cube in E^3 with vertices $(\pm 1, \pm 1, \pm 1)$. Pair the $x = \pm 1$ side of P to itself by the rotation of π about the line $y = 0$, $x = \pm 1$, respectively. Pair the $y = \pm 1$ side of P to itself by the rotation of π about the line $z = 0$, $y = \pm 1$, respectively. Pair the $z = \pm 1$ side of P to itself by the rotation of π about the line $x = 0$, $z = \pm 1$, respectively. The axes of these six rotations intersect P in six line segments that bisect the sides of P as indicated in Figure 13.4.4. Consider the side-pairing Φ consisting of these six rotations. The endpoints of the six axis line segments fall into ridge cycles whose dihedral angle sum is π, and all the other interior edge points of P fall into ridge cycles whose dihedral angle sum is 2π. Therefore Φ is subproper.

Let M be the space obtained from P by gluing together its sides according to Φ. Then M is a Euclidean 3-orbifold by Theorem 13.4.1. Topologically, M is a 3-sphere. This can be seen by gluing together the sides of P one at a time. The six axis line segments are glued together to form the Borromean rings. See Figure 10.3.20. This is beautifully illustrated in the video *Not Knot*. The singular set of M is therefore the Borromean rings, and all the singular points of M have order two.

Example 5. Let P be a regular hyperbolic dodecahedron P in H^3 all of whose proper dihedral angles are $\pi/2$ as in Example 4 of §7.1. We pass to the projective disk model D^3 and center P at the origin. Then P is also a Euclidean regular dodecahedron. Choose three pairs of opposite edges of P that are perpendicular to each other. For example, the six horizontal and vertical edges in Figure 13.4.5. Each side of P shares exactly one of these edges with another side of P. For each of these six edges, pair the two sides

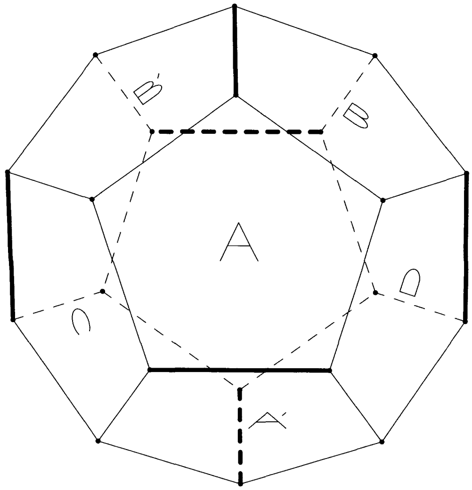


Figure 13.4.5. A right-angled regular dodecahedron in D^3

of P that share this edge by a rotation of $\pi/2$ about the edge. Consider the side-pairing Φ consisting of these 12 rotations. Observe that each point in the interior of these six edges forms a ridge cycle whose dihedral angle sum is $\pi/2$, and all the remaining interior edge points of P fall into ridge cycles whose dihedral angle sum is 2π. Therefore Φ is subproper.

Let M be the space obtained from P by gluing together its sides according to Φ. Then M is a hyperbolic 3-orbifold by Theorem 13.4.1. Topologically, M is a 3-sphere. This can be seen by gluing together the sides of P one at a time. The six edges are glued together to form the Borromean rings. The singular set of M is therefore the Borromean rings, and all the singular points of M have order four.

Complete Gluing of Orbifolds

We now consider gluing together polyhedra to form a complete orbifold. We begin with the complete gluing theorem for Euclidean orbifolds.

Theorem 13.4.2. *Let M be a Euclidean n-orbifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, finite-sided, n-dimensional, convex polyhedra in E^n by a subproper $\mathrm{I}(E^n)$-side-pairing Φ. Then M is complete.*

Proof: The proof is the same as the proof of Theorem 11.1.2 with the exception that the constant $1/3$ must be replaced by $1/4$ as in the proof of Theorem 13.4.1. □

Let M be a hyperbolic n-orbifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, finite-sided, n-dimensional, convex polyhedra in B^n by a subproper $\mathrm{M}(B^n)$-side-pairing Φ. We shall determine necessary and sufficient conditions such that M is complete. We may assume, without loss of generality, that no two polyhedrons in $\mathcal{P}$ meet at infinity. Then Φ extends to a side-pairing of the $(n-1)$-dimensional sides of the Euclidean closures of the polyhedra in $\mathcal{P}$ which, in turn, generates an equivalence relation on the union of the Euclidean closures of the polyhedra in $\mathcal{P}$. The equivalence classes are called *cycles.* We denote the cycle containing a point x by $[x]$. The cycle of a cusp point of a polyhedron in $\mathcal{P}$ is called a *cusp point* of M. As each polyhedron in $\mathcal{P}$ has only finitely many cusp points, M has only finitely many cusp points.

Let c be a cusp point of a polyhedron in $\mathcal{P}$. Let b be a point in $[c]$ and let P_b be the polyhedron in $\mathcal{P}$ containing b in its Euclidean closure. The *link* of b is the $(n-1)$-dimensional, Euclidean, convex polyhedron $L(b)$ obtained by intersecting P_b with a horosphere Σ_b based at b that meets only the sides of P_b incident with b. We shall assume that the horospheres $\{\Sigma_b : b \in [c]\}$ have been chosen small enough so that the links of the points in $[c]$ are mutually disjoint. Then Φ determines a subproper $\mathrm{S}(E^{n-1})$-side-pairing for $\{L(b) : b \in [c]\}$ as in §10.2. Let $L[c]$ be the space obtained by gluing together the polyhedra $\{L(b)\}$ by this side-pairing. The space $L[c]$ is called the *link of the cusp point* $[c]$ of M.

Theorem 13.4.3. *The link $L[c]$ of a cusp point $[c]$ of M is a connected, Euclidean, similarity $(n-1)$-orbifold.*

Proof: The space $L[c]$ is a $(E^n, \mathrm{S}(E^{n-1}))$-orbifold by Theorem 13.4.1. It follows directly from the definition of a cycle that $L[c]$ is connected. □

Theorem 13.4.4. *The link $L[c]$ of a cusp point $[c]$ of M is complete if and only if the links $\{L(b)\}$ for the points in $[c]$ can be chosen so that Φ restricts to a side-pairing for $\{L(b)\}$.*

Proof: If links for the points in $[c]$ can be chosen so that Φ restricts to a side-pairing for $\{L(b)\}$, then this side-pairing for $\{L(b)\}$ is a $\mathrm{I}(E^{n-1})$-side-pairing, and so $L[c]$ is complete by Theorem 13.4.2. The converse is proved by the same argument as in the proof of Theorem 10.2.2. □

Theorem 13.4.5. *If the link $L[c]$ of a cusp point $[c]$ of M is complete, then there is a horoball $B(c)$ based at the point c, a discrete subgroup Γ_c of $\mathrm{M}(B^n)$ leaving $B(c)$ invariant, and an injective local isometry*

$$\iota : B(c)/\Gamma_c \to M$$

compatible with the projection of P_c to M.

Proof: The proof is the same as the proof of Theorem 10.2.3. □

Theorem 13.4.6. *Let M be a hyperbolic n-orbifold obtained by gluing together a finite family $\mathcal{P}$ of disjoint, finite-sided, n-dimensional, convex polyhedra in B^n by a subproper $\mathrm{M}(B^n)$-side-pairing Φ. Then M is complete if and only if $L[c]$ is complete for each cusp point $[c]$ of M.*

Proof: The proof is the same as the proof of Theorem 11.1.6. □

Example 6. Let $\triangle$ be a generalized triangle in H^2 with angles $0, 0, 2\pi/3$ at its vertices x, y, z, respectively. See Figure 13.4.6. Let $L = (x, z]$, $R = (y, z]$, $S = (x, y)$ be the sides of $\triangle$. Pair side L to side R by the rotation g_R about z of $2\pi/3$, pair side R to side L by $g_L = g_R^{-1}$, and pair side S to itself by the reflection g_S in the line $\langle S \rangle$. Consider the side-pairing $\Phi = \{g_L, g_R, g_S\}$. The point z forms a cyclic ridge cycle whose angle sum is $2\pi/3$. Therefore Φ is subproper.

Let M be the space obtained from $\triangle$ by gluing together its sides according to Φ. Then M is a hyperbolic 2-orbifold by Theorem 13.4.1. The cusp points x and y of $\triangle$ form a cusp point of M. Let $L(x)$ and $L(y)$ be disjoint links for x and y that are equidistant from z. Then Φ restricts to a side-pairing for $L(x)$ and $L(y)$. Therefore $L[x]$ is complete by Theorem 13.4.4. Hence M is complete by Theorem 13.4.6.

Topologically, M is a disk with a point removed from its boundary that corresponds to the cusp point $\{x, y\}$. The singular set of M consists of a point of order 3 in the interior of M, corresponding to z, and the boundary of M, all of whose points have order 2.

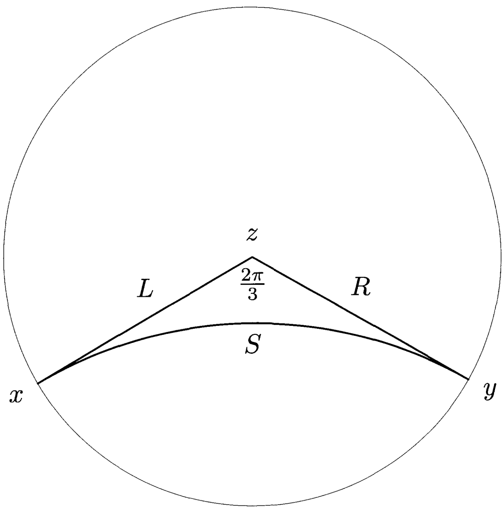


Figure 13.4.6. A generalized triangle in B^2

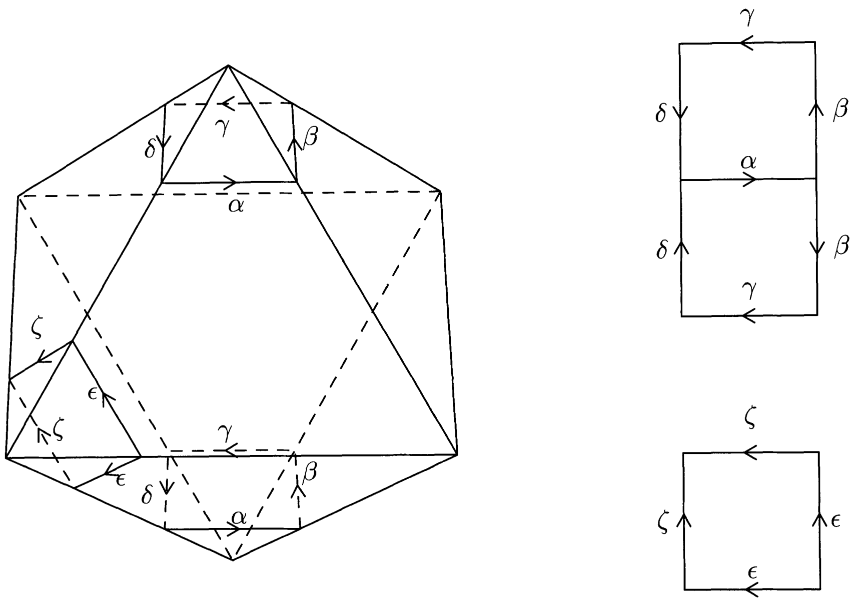


Figure 13.4.7. The links of the cusp points of M

Example 7. Let P be the regular, ideal, hyperbolic octahedron in B^3 with vertices at $\pm e_1, \pm e_2, \pm e_3$. See Figure 10.3.12. All the proper dihedral angles of P are $\pi/2$. For each horizontal edge of P, pair the two sides of P that share this edge by a rotation of $\pi/2$ about the edge. Consider the side-pairing Φ consisting of these eight rotations. Observe that each point on a horizontal edge of P forms a ridge cycle whose dihedral angle sum is $\pi/2$, and all the remaining edge points of P fall into ridge cycles whose dihedral angle sum is π. Therefore Φ is subproper.

Let M be the space obtained from P by gluing together its sides according to Φ. Then M is a hyperbolic 3-orbifold by Theorem 13.4.1. Observe that M has five cusps. Each of the four equatorial cusps of P yields a cusp of M, and the northern and southern cusps of P form the fifth cusp of M. Choose disjoint links for the cusps of P that are equidistant from the origin. Then Φ restricts to a side-pairing for these links. Therefore, each link of M is complete by Theorem 13.4.4. Hence M is complete by Theorem 13.4.6.

Each link of M is topologically a 2-sphere. This can be seen from Figure 13.4.7. Consequently, M is topologically a 3-sphere minus five points. The singular set of M consists of eight lines whose points have order either two or four as indicated in Figure 13.4.8.

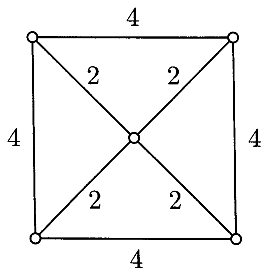


Figure 13.4.8. The singular set of M

Exercise 13.4

1. Let Φ be a G-side-pairing for a finite set of disjoint, n-dimensional, compact, convex polyhedra of X. Prove that Φ has finite cycles.
2. Let P be an exact fundamental polyhedron for a discrete group Γ of isometries of X. Prove that the side-pairing of P determined by Γ is subproper.
3. Let Φ be a G-side-pairing for an n-dimensional convex polyhedron P in X. Prove that Φ is subproper if and only if Φ has finite cycles and every ridge R of P satisfies the conclusions of Theorem 6.7.7.
4. Explain in detail how the hypothesis that Φ is subproper is used in the proof of Theorem 13.4.1.
5. Prove directly that the space obtained by gluing together the sides of the quadrilateral in Example 2 is a Euclidean similarity 2-orbifold.
6. Prove that the Euclidean similarity orbifold in Example 2 is complete if and only if $\alpha = \beta$.
7. Let $M(\alpha)$ be the Euclidean similarity orbifold in Example 2. Prove that $M(\alpha)$ and $M(\alpha')$ are similar if and only if $\alpha = \alpha'$ or $\alpha = \pi - \alpha'$.
8. Describe all the subproper side-pairings for the ideal octahedron in Example 7 that yield a complete hyperbolic orbifold.
9. Position the quadrilateral Q in Example 2 in $\mathbb{C}$ so that the similarity g_S is multiplication by a positive real number. Find all the values of the angle α of Q so that the side-pairing Φ generates a discrete group Γ of isometries of $\mathbb{C}^*$ with fundamental polygon Q.
10. Generalize Theorem 10.5.6 so that the conclusion is as follows: The metric completion $\overline{M}$ is a hyperbolic 3-orbifold if and only if the image of the holonomy for the link L of the cusp point of M contains $2\pi i$.
11. Generalize Theorem 10.5.8 so that the conclusion is as follows: The metric completion $\overline{M}$ is a hyperbolic 3-orbifold if and only if the Dehn surgery invariant of M is a pair (p, q) of integers.
12. Generalize Theorem 10.5.9 so that the greatest common divisor d of p and q may be greater than one and the conclusion is as follows: The metric completion $\overline{M}$ is a hyperbolic 3-orbifold homeomorphic to the 3-manifold $M_{(p/d,q/d)}$ obtained from $\hat{E}^3$ by $(p/d, q/d)$-Dehn surgery on K.

13. Generalize Theorem 10.5.10 so that the greatest common divisor d of p and q may be greater than one and the conclusion is as follows: $M_{(p/d,q/d)}$ has a hyperbolic 3-orbifold structure whose singular set is a simple closed curve all of whose points have order d when $d > 1$.
14. Prove that if $d > 4$, then S^3 has a hyperbolic orbifold structure whose singular set is the figure-eight knot all of whose points have order d.
15. Prove that if $d > 4$, then the d-fold cyclic branched covering of the figure-eight knot has a hyperbolic 3-manifold structure.

§13.5. Poincaré's Theorem

In this section, we prove Poincaré's fundamental polyhedron theorem for discrete groups of isometries of $X = S^n, E^n$, or H^n with $n > 1$. We begin by proving a weak version of Poincaré's theorem.

Theorem 13.5.1. *Let Φ be a subproper $\mathrm{I}(X)$-side-pairing for an n-dimensional, convex polyhedron P in X such that the $(X, \mathrm{I}(X))$-orbifold M obtained from P by gluing together the sides of P by Φ is complete. Then the group Γ generated by Φ is discrete, P is an exact, convex, fundamental polyhedron for Γ, and the inclusion of P into X induces an isometry from M to X/Γ.*

Proof: The quotient map $\pi : P \to M$ maps P° homeomorphically onto an open subset U of M. Let $\phi : U \to X$ be the inverse of π. From the construction of M, we have that ϕ is locally a chart for M. Therefore ϕ is a chart for M.

Let x be a point of P°, let $\tilde{M}$ be the universal orbifold covering space of M based at (x, ϕ), let $\kappa : \tilde{M} \to M$ be the universal orbifold covering projection, and let $\delta : \tilde{M} \to X$ be the corresponding developing map. By Theorem 13.3.10, the map δ is an isometry. Let $\zeta = \kappa\delta^{-1}$. Then $\zeta : X \to M$ extends π on P°, and so ζ extends π by continuity.

Let $\eta : \pi_1^o(M, x, \phi) \to \mathrm{I}(X)$ be the holonomy of M. Then by Theorem 13.3.10, the image of η is a discrete group Γ of isometries of X and the map $\delta : \tilde{M} \to X$ induces an isometry $\overline{\delta} : M \to X/\Gamma$ such that $\overline{\delta}\zeta : X \to X/\Gamma$ is the quotient map.

Now as U is a simply connected subset of $\Omega(M)$, it is evenly covered by κ and ζ. Hence, the members of $\{gP^\circ : g \in \Gamma\}$ are mutually disjoint. As $\pi(P) = M$, we have

$$X = \cup\{gP : g \in \Gamma\}.$$

Therefore P° is a fundamental domain for Γ.

Let g_S be an element of Φ. Choose a point y in the interior of the side S of P. Then there is a point y' in the interior of the side S' of P such that $g_S(y') = y$. Since $\pi(y') = y$, there is an element g of Γ such that $g(y') = y$.

If $y' \neq y$, then $g \neq 1$. If $y' = y$, then $\pi(y)$ is a singular point of M of order two, and so we may assume that $g \neq 1$. Now since gS' does not extend into P°, we must have that gS' lies on the hyperplane $\langle S \rangle$.

Assume first that $S' \neq S$. Then $\pi : P \to M$ maps S° injectively into M. Therefore, we must have that $g = g_S$ in a neighborhood of y' in S'. Hence $g = g_S$ on $\langle S' \rangle$. Furthermore, since gP lies on the opposite side of S from P, we deduce that $g = g_S$ by Theorem 4.3.6.

Assume now that $S' = S$. Then g_S has order two. We may assume that y is an ordinary point of the orbifold $\langle S \rangle / \langle g_S \rangle$. Then π maps a neighborhood of y in S injectively into M. Therefore, the same argument as before shows that $g = g_S$. Thus Γ contains Φ. Therefore P/Γ is a quotient of M.

Now by Theorem 6.5.8, the inclusion map of P into X induces a continuous bijection from P/Γ to X/Γ. The composition of the induced maps

$$X/\Gamma \to M \to P/\Gamma \to X/\Gamma$$

restricts to the identity map of P° and so is the identity map by continuity. Therefore $M = P/\Gamma$.

Now since $\zeta : X \to M$ induces an isometry from X/Γ to $M = P/\Gamma$, the inclusion map of P into X induces an isometry from P/Γ to X/Γ. Therefore P is locally finite by Theorem 6.5.8. Hence P is an exact, convex, fundamental polyhedron for Γ. Finally Φ generates Γ by Theorem 6.7.3. □

In order to apply Theorem 13.5.1, we need to know that the orbifold M is complete. If $X = S^n$, then M is always complete, since M is compact. If $X = E^n$ and the polyhedron P is finite-sided, then M is complete by Theorem 13.4.2. If $X = H^n$ and P is finite-sided, then easily verifiable necessary and sufficient conditions for M to be complete are given by Theorems 13.4.4 and 13.4.6. If $X = H^n$ and P has infinitely many sides, then M may fail to be complete even though the conditions of Theorem 13.4.6 are satisfied. This phenomenon is exhibited by the Example 1 of §11.2. In contrast, we have the following general reflection theorem, where M is always complete.

Theorem 13.5.2. *Let P be an n-dimensional convex polyhedron in X all of whose dihedral angles are submultiples of π. Then the group Γ generated by the reflections of X in the sides of P is a discrete reflection group with respect to the polyhedron P.*

Proof: The orbifold M obtained by gluing together the sides of P by the reflections in the sides of P is just P. Moreover M is isometric to P, since P is a convex subset of X. Now as P is a closed subset of X, we have that P and M are complete. Therefore, the group Γ generated by the reflections of X in the sides of P is a discrete reflection group with respect to the polyhedron P by Theorem 13.5.1. □

Poincaré's Fundamental Polyhedron Theorem

Let $\mathcal{S}$ be the set of sides of an exact, convex, fundamental polyhedron P for a discrete group Γ of isometries of X. Then for each S in $\mathcal{S}$, we have the side-pairing relation

$$g_S g_{S'} = 1$$

of Γ. The expression SS' is called the word in $\mathcal{S}$ corresponding to the side-pairing relation $g_S g_{S'} = 1$ of Γ. Recall from §6.7 that each cycle of sides $\{S_i\}_{i=1}^{\ell}$ of P determines a cycle relation

$$(g_{S_1} g_{S_2} \cdots g_{S_\ell})^k = 1$$

of Γ, where k is the order of $g_{S_1} g_{S_2} \cdots g_{S_\ell}$. The expression $(S_1 S_2 \cdots S_\ell)^k$ is called the word in $\mathcal{S}$ corresponding to the above cycle relation of Γ. We are now ready to state Poincaré's fundamental polyhedron theorem.

Theorem 13.5.3. *Let Φ be a subproper $\mathrm{I}(X)$-side-pairing for an n-dimensional, convex polyhedron P in X such that the $(X, \mathrm{I}(X))$-orbifold M obtained from P by gluing together the sides of P by Φ is complete. Then the group Γ generated by Φ is discrete, P is an exact, convex, fundamental polyhedron for Γ, and if $\mathcal{S}$ is the set of sides of P and $\mathcal{R}$ is the set of words in $\mathcal{S}$ corresponding to all the side-pairing and cycle relations of Γ, then $(\mathcal{S}; \mathcal{R})$ is a group presentation for Γ under the mapping $S \mapsto g_S$.*

Proof: The proof is essentially the same as the proof of Theorem 11.2.2. The only difference is in the construction of the neighborhood U of an interior ridge point x of P in step (10), where ℓ is replaced by $k\ell$. □

Theorem 13.5.3 gives a group presentation $(\mathcal{S}; \mathcal{R})$ for the group Γ generated by the side-pairing Φ. The presentation $(\mathcal{S}; \mathcal{R})$ can be simplified by eliminating each side-pairing relation $SS' = 1$ such that $S \neq S'$ and exactly one of the generators S or S'. If S' is eliminated, then each occurrence of S' in a cycle relation is replaced by S^{-1}. Moreover, each cycle of sides $\{S_i\}_{i=1}^{\ell}$ determines 2ℓ cycles of sides by taking cyclic permutations of $\{S_i\}_{i=1}^{\ell}$ and their inverse orderings. The corresponding cycle transformations are all conjugate to each other or their inverses. Therefore, any of the corresponding cycle relations is derivable from any of the others. Hence, all but one of them can be eliminated from a presentation for Γ. Thus $(\mathcal{S}; \mathcal{R})$ can be simplified to a presentation with the generators of the form $S = S'$ and half the generators of the form $S \neq S'$, and the side-pairing relations of the form $S^2 = 1$, and one cycle relation for each cycle of ridges of P.

Example 1. Consider the triangle $\triangle$ in S^2, E^2 or H^2 in Figure 13.4.1. Let Γ be the group generated by the side-pairing for $\triangle$ described in Example 1 of §13.4. The triangle has two cycles of vertices. By Theorem 13.5.3, the group Γ has the presentation

$$(L, R, S; LR, S^2, R^3, (RS)^{2k}).$$

We eliminate the generator L and the side-pairing relation $RL = 1$ to obtain the presentation

$$(R, S; S^2, R^3, (RS)^{2k})$$

for the group Γ.

Example 2. Consider the regular tetrahedron P in S^3 in Figure 13.4.3. Let Γ be the group generated by the side-pairing for P described in Example 3 of §13.4. The tetrahedron has three cycles of edges. By Theorem 13.5.3, the group Γ has the presentation

$$(A, B, C, D; AB, CD, B^4, C^4, ADBC).$$

We eliminate the generators A and D and the side-pairing relations $AB = 1$ and $CD = 1$ to obtain the presentation

$$(B, C; B^4, C^4, B^{-1}C^{-1}BC)$$

for Γ. Therefore Γ is the direct product of two cyclic groups of order 4.

Theorem 13.5.4. *Let P be an exact, convex, fundamental polyhedron for a discrete group Γ of isometries of X, let $\mathcal{S}$ be the set of sides of P, and let $\mathcal{R}$ be the set of all the side-pairing and cycle relations of Γ with respect to the Γ-side-pairing of P. Then $(\mathcal{S}; \mathcal{R})$ is a group presentation for Γ under the mapping $S \mapsto g_S$.*

Proof: Let M be the orbifold obtained by gluing the sides of P by the Γ-side-pairing of P. Then the inclusion of P into X induces an isometry from M to X/Γ by Theorem 13.5.1. Therefore M is complete. Hence $(\mathcal{S}; \mathcal{R})$ is a group presentation for Γ under the mapping $S \mapsto g_S$ by Theorem 13.5.3. $\square$

Exercise 13.5

1. Show that Theorem 13.5.3 does not hold for $X = S^1$ but does hold for $X = E^1$ or H^1.
2. Find a presentation for the discrete group of isometries of E^3 corresponding to the Euclidean orbifold in Example 4 of §13.4.
3. Find a presentation for the discrete group of isometries of H^3 corresponding to the hyperbolic orbifold in Example 5 of §13.4.
4. Find a presentation for the discrete group of isometries of H^3 corresponding to the hyperbolic orbifold in Example 7 of §13.4.
5. Let Γ be a discrete group of isometries of X. Prove that the dimension of $\Sigma(X/\Gamma)$ is $n - 1$ if and only if Γ contains a reflection of X.

§13.6. Historical Notes

§13.1. Theorem 13.1.7 was essentially proved by Floyd in his 1950 paper *Some characterizations of interior maps* [134]. See also Armstrong's 1968 paper *The fundamental group of the orbit space of a discontinuous group* [24].

§13.2. Spherical, Euclidean, and hyperbolic 2-orbifolds were studied by Koebe in his 1930 paper *Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen V* [245]. Two-dimensional spherical, Euclidean, and hyperbolic orbit spaces were studied by Nielsen and Fenchel in their 1959 manuscript *Discontinuous Groups of Non-Euclidean Motions* [321]. Differentiable n-orbifolds were introduced by Satake in his 1956 paper *On a generalization of the notion of manifold* [356]. These orbifolds were called *V-manifolds* by Satake. The term *orbifold* was introduced by Thurston in his 1979 lecture notes *The Geometry and Topology of 3-Manifolds* [389].

§13.3. The homotopy theory of (X, G)-paths was developed by Haefliger in his 1990 paper *Orbi-espaces* [175]. In particular, Theorem 13.3.2 appeared in this paper. The concept of the developing map of an orbifold was introduced by Koebe in his 1930 paper [245]. In particular, Theorem 13.3.10 for groups of isometries of S^2, E^2, or H^2, without reflections, appeared in this paper. Theorem 13.3.10 for groups of isometries appeared in Thurston's 1979 lecture notes [389].

§13.4. The hyperbolic 2-orbifold obtained by gluing together the sides of a fundamental polygon of a Fuchsian group was introduced by Poincaré in his 1882 paper *Théorie des groupes fuchsiens* [330]. Theorems 13.4.2-13.4.6 were essentially proved by Seifert in his 1975 paper *Komplexe mit Seitenzuordnung* [371]. For some interesting examples of hyperbolic 3-orbifolds, see Weber and Seifert's 1933 paper *Die beiden Dodekaederräume* [405], Meyerhoff's 1985 paper *The cusped hyperbolic 3-orbifold of minimum volume* [287], Adams' 1992 paper *Noncompact hyperbolic 3-orbifolds of small volume* [7], and Hilden, Lozano, and Montesinos' 1992 papers *The arithmeticity of figure eight knot orbifolds* [192] and *On the Borromean orbifolds: Geometry and arithmetic* [191]. For a beautiful illustration of a sequence of geometric 3-orbifolds converging to the complement of the Borromean rings, see Epstein and Gunn's 1991 video *Not Knot* [116].

It is an interesting fact due to Thurston that every closed orientable 3-manifold has a hyperbolic orbifold structure. In fact, every closed orientable 3-manifold is an orbifold covering space of the hyperbolic orbifold in Example 5. For a discussion, see Hilden, Lozano, Montesinos, and Whitten's 1987 paper *On universal groups and 3-manifolds* [193].

§13.5. The 2-dimensional version of Poincaré's theorem for finite-sided polygons appeared in Poincaré's 1882 paper [330]. See also de Rham's 1971 paper *Sur les polygones générateurs de groupes fuchsiens* [103]. The 3-dimensional version of Poincaré's theorem for finite-sided polyhedra of infinite volume appeared in Poincaré's 1883 *Mémoire sur les groupes des* *kleinéens* [332]. The 2- and 3-dimensional versions of Poincaré's theorem, for side-pairings such that the stabilizer of a face fixes the face pointwise, were proved by Maskit in his 1971 paper *On Poincaré's theorem for fundamental polygons* [281]. Theorem 13.5.1, for finite-sided polyhedra and side-pairings such that the stabilizer of a face fixes the face pointwise, was proved by Seifert in his 1975 paper [371]. The n-dimensional version of Poincaré's theorem, for finite-sided polyhedra of finite volume and side-pairings such that the stabilizer of a face fixes the face pointwise, was proved by Morokuma in his 1978 paper *A characterization of fundamental domains of discontinuous groups acting on real hyperbolic spaces* [306]. The n-dimensional version of Poincaré's theorem appeared in Maskit's 1988 treatise *Kleinian Groups* [282]. For a computer implementation of the 3-dimensional version of Poincaré's theorem, see Riley's 1983 paper *Applications of a computer implementation of Poincaré's theorem on fundamental polyhedra* [352].

Bibliography

1. Abikoff, W., *The Real Analytic Theory of Teichmüller Space*, Lecture Notes in Math., 820, Springer-Verlag, Berlin (1980).
2. Abikoff, W., The uniformization theorem, *Amer. Math. Monthly*, 88 (1981), 574-592.
3. Abikoff, W. and Haas, A., Nondiscrete groups of hyperbolic motions, *Bull. London Math. Soc.*, 22 (1990), 233-238.
4. Adams, C., The noncompact hyperbolic 3-manifold of minimal volume, *Proc. Amer. Math. Soc.*, 100 (1987), 601-606.
5. Adams, C., Volumes of N-cusped hyperbolic 3-manifolds, *J. London Math. Soc.*, 38 (1988), 555-565.
6. Adams, C., SnapPea, The Weeks hyperbolic 3-manifold program, *Notices Amer. Math. Soc.*, 37 (1990), 273-275.
7. Adams, C., Noncompact hyperbolic 3-orbifolds of small volume, In: *Topology' 90*, edited by B. Apanasov, W. D. Neumann, A. W. Reid, and L. Siebenmann, de Gruyter, Berlin (1992), 1-15.
8. Adams, C., Hildebrand, M., and Weeks, J., Hyperbolic invariants of knots and links, *Trans. Amer. Math. Soc.*, 326 (1991), 1-56.
9. Agard, S., Remarks on the boundary mapping for a Fuchsian group, *Ann. Acad. Sci. Fenn. Ser. A. I. Math.*, 10 (1985), 1-13.
10. Aitchison, I. R. and Rubinstein, J. H., An introduction to polyhedral metrics of non-positive curvature on 3-manifolds, In: *Geometry of Low-Dimensional Manifolds, Vol. 2*, edited by S. K. Donaldson and C. B. Thomas, London Math. Soc. Lecture Note Ser., 151, Cambridge Univ. Press, Cambridge (1990), 127-161.
11. Aitchison, I. R. and Rubinstein, J. H., Combinatorial cubings, cusps, and the dodecahedral knots, In: *Topology' 90*, edited by B. Apanasov, W. D. Neumann, A. W. Reid, and L. Siebenmann, de Gruyter, Berlin (1992), 17-26.
12. Aleksandrov, A. D., On the filling of space by polyhedra (Russian), *Vestnik Leningrad Univ. Ser. Mat. Fiz. Khim.*, 9 (1954), 33-43.
13. Alexander, J. W., Note on two three-dimensional manifolds with the same group, *Trans. Amer. Math. Soc.*, 20 (1919), 339-342.
14. Alperin, R. C., An elementary account of Selberg's lemma, *Enseign. Math.*, 33 (1987), 269-273.
15. Andreev, E. M., Intersection of plane boundaries of a polytope with acute angles, *Math. Notes*, 8 (1970), 761-764.
16. Apanasov, B. N., Kleinian groups in space, *Siberian Math. J.*, 16 (1975), 679-684.

17. Apanasov, B. N., A universal property of Kleinian groups in the hyperbolic metric, *Soviet Math. Dokl.*, 16 (1975), 1418-1421.
18. Apanasov, B. N., Geometrically finite groups of transformations of space, *Siberian Math. J.*, 23 (1982), 771-780.
19. Apanasov, B. N., Geometrically finite hyperbolic structures on manifolds, *Ann. Glob. Anal. Geom.*, 1 (1983), 1-22.
20. Apanasov, B. N., *Discrete Groups in Space and Uniformization Problems*, Kluwer, Dordrecht (1991).
21. Apanasov, B. N. and Gutsul, I. S., Greatly symmetric totally geodesic surfaces and closed hyperbolic 3-manifolds which share a fundamental polyhedron, In: *Topology' 90*, edited by B. Apanasov, W. D. Neumann, A. W. Reid, and L. Siebenmann, de Gruyter, Berlin (1992), 37-53.
22. Appell, P., Goursat, É., and Fatou, P., *Théorie des Fonctions Algébriques II, 2nd Ed.*, Gauthier-Villars, Paris (1930).
23. Archimedes, *The Works of Archimedes*, edited by T. L. Heath, Cambridge Univ. Press, Cambridge (1897) (Dover, New York).
24. Armstrong, M. A., The fundamental group of the orbit space of a discontinuous group, *Proc. Camb. Phil. Soc.*, 64 (1968), 299-301.
25. Atiyah, M. F. and Macdonald, I. G., *Introduction to Commutative Algebra*, Addison-Wesley, Reading, MA (1969).
26. Auslander, L., An account of the theory of crystallographic groups, *Proc. Amer. Math. Soc.*, 16 (1965), 1230-1236.
27. Auslander, L. and Kuranishi, M., On the holonomy group of locally Euclidean spaces, *Ann. of Math.*, 65 (1957), 411-415.
28. Autonne, L., Sur l'Hermitien, *Rend. Circ. Mat. Palermo*, 16 (1902), 104-128.
29. Avérous, G. and Kobayashi, S., On automorphisms of spaces of nonpositive curvature with finite volume, In: *Differential Geometry and Relativity*, edited by M. Cahen and M. Flato, D. Reidel, Dordrecht (1976), 19-26.
30. Barbarin, P., Études de géométrie analytique non Euclidienne, *Acad. Roy. Belg. Mem.*, 60 (1901), 1-168.
31. Basmajian, A., Constructing pairs of pants, *Ann. Acad. Sci. Fenn. Ser. A. I. Math.*, 15 (1990), 65-74.
32. Basmajian, A., Generalizing the hyperbolic collar lemma, *Bull. Amer. Math. Soc.*, 27 (1992), 154-158.
33. Beardon, A. F., Fundamental domains for Kleinian groups, In: *Discontinuous Groups and Riemann Surfaces*, edited by L. Greenberg, Ann. of Math. Studies, 79, Princeton Univ. Press, Princeton, NJ (1974), 31-41.
34. Beardon, A. F., *The Geometry of Discrete Groups*, Graduate Texts in Math., 91, Springer-Verlag, New York (1983).
35. Beardon, A. F. and Maskit, B., Limit points of Kleinian groups and finite-sided fundamental polyhedra, *Acta Math.*, 132 (1974), 1-12.
36. Beardon, A. F. and Wilker, J. B., The norm of a Möbius transformation, *Math. Proc. Camb. Phil. Soc.*, 96 (1984), 301-308.
37. Bellavitis, G., Teoria della figure inverse, e loro uso nella geometria elementare, *Ann. Sci. Regno Lombardo-Veneto*, 6 (1836), 126-141.
38. Beltrami, E., Saggio di interpetrazione della geometria non-euclidea, *Giorn. Mat.*, 6 (1868), 248-312 (Essai d'interprétation de la géométrie non euclidienne, *Ann. Sci. École Norm. Sup.*, 6 (1869), 251-288; *Translation of*

Beltrami's Essay on the Interpretation of Non-Euclidean Geometry, J. Stillwell, Monash Univ., Clayton, Australia, 1982).

39. Beltrami, E., Teoria fondamentale degli spazii di curvatura costante, *Ann. Mat. Pura Appl.*, 2 (1868), 232-255 (Théorie fondamentale des espaces de courbure constante, *Ann. Sci. École Norm. Sup.* , 6 (1869), 347-375; *Translation of Beltrami's Theory of Spaces of Constant Curvature*, J. Stillwell, Monash Univ., Clayton, Australia, 1982).
40. Benedetti, R. and Petronio, C., *Lectures on Hyperbolic Geometry*, Universitext, Springer-Verlag, Berlin (1992).
41. Berestovskii, V. N., Homogeneous Busemann G-spaces, *Siberian Math. J.*, 23 (1982), 141-150.
42. Berger, M., Convexity, *Amer. Math. Monthly*, 97 (1990), 650-678.
43. Bers, L. and Gardiner, F. P., Fricke Spaces, *Adv. Math.*, 62 (1986), 249-284.
44. Best, L. A., On torsion-free discrete subgroups of $\mathrm{PSL}(2,\mathbb{C})$ with compact orbit space., *Canad. J. Math.*, 23 (1971), 451-460.
45. Bianchi, L., Sulle forme differenziali quadratiche indefinite, *Atti. Accad. Naz. Lincei. Mem.*, 5 (1888), 539-603.
46. Bieberbach, L., Über die Bewegungsgruppen der Euklidischen Räume I, *Math. Ann.*, 70 (1911), 297-336.
47. Bieberbach, L., Über die Bewegungsgruppen der Euklidischen Räume II, *Math. Ann.*, 72 (1912), 400-412.
48. Birkhoff, G. and Bennett, M. K., Felix Klein and his "Erlanger Programm", In: *History and Philosophy of Modern Mathematics*, edited by W. Aspray and P. Kitcher, Minnesota Studies Philos. Sci., 11, Univ. Minnesota Press, Minneapolis (1988), 145-176.
49. Bishop, R. L. and O'Neill, B., Manifolds of negative curvature, *Trans. Amer. Math. Soc.*, 145 (1969), 1-49.
50. Böhm, J. and Hertel, E., *Polyedergeometrie in n-dimensionalen Räumen konstanter Krümmung*, Birkhäuser, Basel (1981).
51. Bolyai, J., Appendix, scientiam spatii absolute veram exhibens, In: *Tentamen Juventutem studiosam in elementa Matheseos purae*, W. Bolyai, Maros Vásárhelyini (1832) (*Appendix*, edited by F. Kárteszi, North-Holland Math. Studies, 138, North-Holland, Amsterdam, 1987).
52. Bonola, R., *Non-Euclidean Geometry*, translated by H. S. Carslaw, Open Court, Chicago (1912) (Dover, New York, 1955).
53. Borel, A., Compact Clifford-Klein forms of symmetric spaces, *Topology*, 2 (1963), 111-122.
54. Bourbaki, N., *Groupes et Algèbres de Lie*, Éléments de mathématique, 34, Hermann, Paris (1968).
55. Bowditch, B. H., Geometrical finiteness for hyperbolic groups, *J. Funct. Anal.*, 113 (1993), 245-317.
56. Boy, W., Über die Curvatura integra und die Topologie geschlossener Flächen, *Math. Ann.*, 57 (1903), 151-184.
57. Boyer, C. B., Early rectifications of curves, In: *Mélanges Alexandre Koyré I*, Histoire de la pensée, 12, Hermann, Paris (1964), 30-39.
58. Brody, E. J., The topological classification of the lens spaces, *Ann. of Math.*, 71 (1960), 163-184.
59. Brøndsted, A., *An Introduction to Convex Polytopes*, Graduate Texts in Math., 90, Springer-Verlag, New York (1983).

60. Brooks, R. and Matelski, J. P., Collars in Kleinian groups, *Duke Math. J.*, 49 (1982), 163-182.
61. Brown, H., Bülow, R., Neubüser, J., Wondratschek, H., and Zassenhaus, H., *Crystallographic Groups of Four-Dimensional Space*, Wiley, New York (1978).
62. Busemann, H., Spaces with non-positive curvature, *Acta Math.*, 80 (1948), 259-310.
63. Busemann, H., *The Geometry of Geodesics*, Pure Appl. Math., 6, Academic Press, New York (1955).
64. Busemann, H. and Kelly, P. J., *Projective Geometry and Projective Metrics*, Pure Appl. Math., 3, Academic Press, New York (1953).
65. Buser, P., A geometric proof of Bieberbach's theorems on crystallographic groups, *Enseign. Math.*, 31 (1985), 137-145.
66. Cajori, F., Generalizations in geometry as seen in the history of developable surfaces, *Amer. Math. Monthly*, 36 (1929), 431-437.
67. Cannon, J. W., The combinatorial structure of cocompact discrete hyperbolic groups, *Geom. Dedicata*, 16 (1984), 123-148.
68. Cartan, É., La théorie des groupes continus et la géométrie, *Encyclop. Sci. Math.* (1915) (*Oeuvres de É. Cartan*, Vol. III, Gauthier-Villars, Paris (1952), 1727-1861).
69. Cartan, É., L'application des espaces de Riemann et l'analysis situs, *Assoc. Avanc. Sci. Lyon* (1926), 53 (*Oeuvres de É. Cartan*, Vol. III, Gauthier-Villars, Paris (1952), 993-995).
70. Cartan, É., Le rôle de la théorie des groupes de Lie dans l'évolution de la géométrie moderne, *C. R. Congrès Intern. Oslo I* (1936), 92-103 (*Oeuvres de É. Cartan*, Vol. III, Gauthier-Villars, Paris (1952), 1373-1384).
71. Cauchy, A., *Cours d'Analysis*, Paris (1821).
72. Cauchy, A., Sur l'équation à l'aide de laquelle on détermine les inégalités séculaires des mouvements des planètes, *Exercices de Math.*, 4 (1829) (*Oeuvres de Cauchy* (2) 9, 174-195).
73. Cauchy, A., Mémoire sur les lieux analytiques, *C. R. Acad. Sci. Paris*, 24 (1847), 885-887.
74. Cayley, A., Chapters in the analytical geometry of (n) dimensions, *Cambridge Math. J.*, 4 (1843), 119-127.
75. Cayley, A., A memoir on the theory of matrices, *Philos. Trans. Roy. Soc. London*, 148 (1858), 17-37.
76. Cayley, A., A sixth memoir upon quantics, *Philos. Trans. Roy. Soc. London*, 149 (1859), 61-91.
77. Charlap, L. S., *Bieberbach Groups and Flat Manifolds*, Universitext, Springer-Verlag, New York (1986).
78. Chasles, M., Note sur les propriétés générales du système de deux corps semblables entr'eux, *Bull. Sci. Math.*, 14 (1830), 321-326.
79. Chein, M., Recherche des graphes des matrices de Coxeter hyperboliques d'ordre ≤ 10, *Rev. Francaise Informat. Rech. Opération*, 3 (1969), 3-16.
80. Chen, S. S. and Greenberg, L., Hyperbolic Spaces, In: *Contributions to Analysis*, edited by L. V. Ahlfors, I. Kra, B. Maskit, and L. Nirenberg, Academic Press, New York (1974), 49-87.
81. Clausen, T., Ueber die Function $\sin\phi + \frac{1}{2^2}\sin 2\phi + \frac{1}{3^2}\sin 3\phi + etc.$, *J. Reine Angew. Math.*, 8 (1832), 298-300.

82. Clifford, W., Preliminary sketch of biquaternions, *Proc. London Math. Soc.*, 4 (1873), 381-395.
83. Cohn-Vossen, S., Existenz kürzester Wege, *Dokl. Akad. Nauk. SSSR*, 3 (1935), 339-342.
84. Cox, H., Homogeneous coordinates in imaginary geometry and their application to systems of forces, *Quart. J. Pure Appl. Math.*, 18 (1882), 178-215.
85. Coxeter, H. S. M., Groups whose fundamental regions are simplexes, *J. London Math. Soc.*, 6 (1931), 133-136.
86. Coxeter, H. S. M., The polytopes with regular-prismatic vertex figures II, *Proc. London Math. Soc.*, 34 (1932), 126-189.
87. Coxeter, H. S. M., Discrete groups generated by reflections, *Ann. of Math.*, 35 (1934), 588-621.
88. Coxeter, H. S. M., The complete enumeration of finite groups of the form $R_i^2 = (R_iR_j)^{k_{ij}} = 1$, *J. London Math. Soc.*, 10 (1935), 21-25.
89. Coxeter, H. S. M., The functions of Schläfli and Lobatschefsky, *Quart. J. Math. Oxford*, 6 (1935), 13-29.
90. Coxeter, H. S. M., Regular honeycombs in hyperbolic space, *Proc. Intern. Congr. Math., 1954, Amsterdam*, 3 (1956), 155-169.
91. Coxeter, H. S. M., *Non-Euclidean Geometry, 5th Ed.*, Univ. Toronto Press, Toronto (1968).
92. Coxeter, H. S. M., *Regular Polytopes, 3rd Ed.*, Dover, New York (1973).
93. Coxeter, H. S. M., Gauss as a geometer, *Historia Math.*, 4 (1977), 379-396.
94. Coxeter, H. S. M., Angels and devils, In: *The Mathematical Gardner*, edited by D. A. Klarner, Wadsworth, Belmont, CA (1981), 197-209.
95. Coxeter, H. S. M. and Moser, W. O. J., *Generators and Relations for Discrete Groups, 4th Ed.*, Ergeb. Math. Grenzgeb., 14, Springer-Verlag, Berlin (1980).
96. Coxeter, H. S. M. and Whitrow, G. J., World-structure and non-Euclidean honeycombs, *Proc. Royal Soc. London*, A201 (1950), 417-437.
97. Crowe, M. J., *A History of Vector Analysis*, Univ. of Notre Dame Press, Notre Dame, IN (1967).
98. Culler, M. and Shalen, P. B., Paradoxical decompositions, 2-generator Kleinian groups, and volumes of hyperbolic 3-manifolds, *J. Amer. Math. Soc.*, 5 (1992), 231-288.
99. Davis, M. W., A hyperbolic 4-manifold, *Proc. Amer. Math. Soc.*, 93 (1985), 325-328.
100. Dedekind, R., Schreiben an Herrn Borchardt über die Theorie der elliptischen Modulfunktionen, *J. Reine Angew. Math.*, 83 (1877), 265-292.
101. Dehn, M., Die Eulersche Formel im Zusammenhang mit dem Inhalt in der Nicht-Euklidischen Geometrie, *Math. Ann.*, 61 (1905), 561-586.
102. Dehn, M., Über unendliche diskontinuierliche Gruppen, *Math. Ann.*, 71 (1912), 116-144 (On infinite discontinuous groups, In: *Papers on Group Theory and Topology*, translated by J. Stillwell, Springer-Verlag, New York (1987), 133-178).
103. de Rham, G., Sur les polygones générateurs de groupes fuchsiens, *Enseign. Math.*, 17 (1971), 49-61.
104. Desargues, G., *Brouillon project d'une atteinte aux événements des recontres du cone avec un plan*, Paris (1639) (Rough draft of an essay on the results of taking plane sections of a cone, In: *The Geometrical Work of*

Girard Desargues, J. V. Field and J. J. Gray, Springer-Verlag, New York (1987), 69-143).

105. Descartes, R., De solidorum elementis, In: *Oeuvres inédites de Descartes*, Vol. 2, Foucher de Careil, Paris (1860), 214-234 (The elements of solids, In: *Descartes on Polyhedra*, P. J. Federico, Sources Hist. Math. Phys. Sci., 4, Springer-Verlag, New York, 1982).
106. Dieudonné, J., *History of Algebraic Geometry*, translated by J. D. Sally, Wadsworth, Monterey, CA (1985).
107. Dirichlet, G. L., Über die Reduction der positiven quadratischen Formen mit drei unbestimmten ganzen Zahlen, *J. Reine Angew. Math.*, 40 (1850), 209-227.
108. Douady, A., L'espace de Teichmüller, In: *Travaux de Thurston sur les surfaces*, edited by A. Fathi, F. Laudenbach, and V. Poénaru, Astérisque, 66-67, Soc. Math. France, Paris (1979), 127-137.
109. Dugac, P., Histoire des espaces complets, *Rev. Hist. Sci.*, 37 (1984), 3-28.
110. Dugundji, J., *Topology*, Allyn and Bacon, Boston (1966).
111. Dyck, W., Gruppentheoretische Studien, *Math. Ann.*, 20 (1882), 1-44.
112. Dyck, W., Vorläufige Mittheilungen über die durch Gruppen linearer Transformationen gegebenen regulären Gebietseintheilungen des Raumes, *Ber. Verh. Säch. Akad. Wiss. Leipzig. Math.-Phys. Kl.*, (1883), 61-75.
113. Dyck, W., Beiträge zur Analysis situs, *Math. Ann.*, 32 (1888), 457-512.
114. Efremovič, V. A. and Tihomirova, E. S., Continuation of an equimorphism to infinity, *Soviet Math. Dokl.*, 4 (1963), 1494-1496.
115. Ehresmann, C., Sur les espaces localement homogènes, *Enseign. Math.*, 35 (1936), 317-333.
116. Epstein, D. and Gunn, C., *Not Knot*, video, Jones and Bartlett, Boston (1991).
117. Escher, M. C., *Art and Science*, edited by H. S. M. Coxeter, M. Emmer, R. Penrose, and M. L. Teuber, North-Holland, Amsterdam (1986).
118. Euclid, *The Thirteen Books of Euclid's Elements, Vol. 1-3, 2nd Ed.*, edited by T. L. Heath, Cambridge Univ. Press, Cambridge (1926) (Dover, New York, 1956).
119. Euler, L., De linea brevissima in superficie quacunque duo quaelibet puncta iungente, *Comment. Acad. Sci. Petrop.*, 3 (1732), 110-124 (*Opera omnia*, (1) 25, 1-12).
120. Euler, L., Principes de la trigonométrie sphérique tirés de la méthode des plus grands et plus petits, *Mem. Acad. Sci. Berlin*, 9 (1755), 223-257 (*Opera omnia*, (1) 27, 277-308).
121. Euler, L., Elementa doctrinae solidorum, *Novi Comment. Acad. Sci. Petrop.*, 4 (1758), 109-140 (*Opera omnia* (1) 26, 71-93).
122. Euler, L., Demonstratio nonnullarum insignium proprietatum quibus solida hedris planis inclusa sunt praedita, *Novi Comment. Acad. Sci. Petrop.*, 4 (1758), 140-160 (*Opera omnia* (1) 26, 94-108).
123. Euler, L., Considerationes de traiectoriis orthogonalibus, *Novi Comment. Acad. Sci. Petrop.*, 14 (1770), 46-71 (*Opera omnia* (1) 28, 99-119).
124. Euler, L., Problema algebraicum ob affectiones prorsus singulares memorabile, *Novi Comment. Acad. Sci. Petrop.*, 15 (1771), 75-106 (*Opera omnia*, (1) 6, 286-315).

125. Euler, L., De solidis quorum superficiem in planum explicare licet, *Novi Comment. Acad. Sci. Petrop.*, 16 (1772), 3-34 (*Opera Omnia* (1) 28, 161-186).
126. Euler, L., Formulae generales pro translatione quacunque corporum rigidorum, *Novi Comment. Acad. Sci. Petrop.*, 20 (1776), 189-207 (*Opera omnia*, (2) 9, 84-98).
127. Euler, L., De mensura angulorum solidorum, *Acta Acad. Sci. Petrop.*, 2 (1781), 31-54 (*Opera omnia*, (1) 26, 204-223).
128. Euler, L., Trigonometria sphaerica universa ex primis principiis breviter et dilucide derivata, *Acta Acad. Sci. Petrop.*, 3 (1782), 72-86 (*Opera omnia*, (1) 26, 224-236).
129. Euler, L., De centro similitudinis, *Nova Acta Acad. Sci. Petrop.*, 9 (1795), 154-165 (*Opera Omnia* (1) 26, 276-285).
130. Farkas, D. R., Crystallographic groups and their mathematics, *Rocky Mountain J. Math.*, 11 (1981), 511-551.
131. Fenchel, W., Bemerkungen zur allgemeinen Theorie der diskontinuierlichen Transformationsgruppen, *Proc. XIII Congr. Math. Scand.*, (1957), 77-85.
132. Fenchel, W., *Elementary Geometry in Hyperbolic Space*, de Gruyter Studies Math., 11, de Gruyter, Berlin (1989).
133. Fleming, W., *Functions of Several Variables*, Undergraduate Texts in Math., Springer-Verlag, New York (1977).
134. Floyd, E. E., Some characterizations of interior maps, *Ann. of Math.*, 51 (1950), 571-575.
135. Ford, L. R., On the foundations of the theory of discontinuous groups of linear transformations, *Proc. Nat. Acad. U.S.A.*, 13 (1927), 286-289.
136. Ford, L. R., *Automorphic Functions, 2nd Ed.*, Chelsea, New York (1951).
137. Fréchet, M., Sur quelques points du calcul fonctionnel, *Rend. Circ. Mat. Palermo*, 22 (1906), 1-74.
138. Fricke, R., Die Kreisbogenvierseite und das Princip der Symmetrie, *Math. Ann.*, 44 (1894), 565-599.
139. Fricke, R. and Klein, F., *Vorlesungen über die Theorie der automorphen Functionen*, Vols. I, II, B. G. Teubner, Leipzig (1897, 1912).
140. Fried, D., Closed similarity manifolds, *Comment. Math. Helv.*, 55 (1980), 576-582.
141. Frobenius, G., Über die principale Transformation der Thetafunctionen mehrerer Variabeln, *J. Reine Angew. Math.*, 95 (1883), 264-296.
142. Frobenius, G., Über den von L. Bieberbach gefundenen Beweis eines Satzes von C. Jordan, *Sitzungsber. Preuss. Akad. Wiss.*, (1911), 241-248.
143. Fubini, G., Sulla teoria dei gruppi discontinui, *Ann. Mat. Pura Appl.*, 11 (1905), 159-186.
144. Fubini, G., *Introduzione alla teoria dei gruppi discontinui e delle funzioni automorfe*, Pisa (1908).
145. Gallo, D., A 3-dimensional hyperbolic collar lemma, In: *Kleinian Groups and Related Topics*, edited by D. M. Gallo and R. M. Porter, Lecture Notes in Math., 971, Springer-Verlag, Berlin (1983), 31-35.
146. Galois, É., Mémoire sur les conditions de résolubilité des équations par radicaux, *J. Math. Pures Appl.*, 11 (1846), 417-433.
147. Garland, H. and Raghunathan, M. S., Fundamental domains for lattices in ($\mathbb{R}$)-rank 1 semisimple Lie groups, *Ann. of Math.*, 92 (1970), 279-326.

148. Gauss, C. F., Disquisitiones generales circa superficies curvas, *Comment. Soc. Reg. Sci. Göttingen. Rec.*, 6 (1828), 99-146 (General investigations of curved surfaces, In: *150 Years after Gauss' "Disquisitiones generales circa superficies curvas", 2nd Ed.*, P. Dombrowski, Astérisque, 62 (1979), Soc. Math. France, Paris).
149. Gauss, C. F., Anzeige. Untersuchungen über die Eigenschaften der positiven ternären quadratischen Formen von Ludwig August Seeber, *Göttingische gelehrete Anzeigen*, (1831) (*Carl Friedrich Gauss Werke*, Vol. II, Göttingen (1876), 188-196).
150. Gauss, C. F., Grundlagen der Geometrie, In: *Carl Friedrich Gauss Werke*, Vol. VIII, B. G. Teubner, Leipzig (1900), 157-268.
151. Gehring, F. W. and Martin, G. J., Inequalities of Möbius transformations and discrete groups, *J. Reine Angew. Math.*, 418 (1991), 31-76.
152. Gibbs, J., *Elements of Vector Analysis*, New Haven, CT (1881) (*The Collected Works of J. Willard Gibbs*, Vol. II, Pt. 2, Longmans, Green and Co., New York (1928), 17-50).
153. Gieseking, H., *Analytische Untersuchungen über topologische Gruppen*, Dissertation, Univ. Münster (1912).
154. Goldman, W. M., Geometric structures on manifolds and varieties of representations, In: *Geometry of Group Representations*, edited by W. M. Goldman and A. R. Magid, Contemporary Math., 74, Amer. Math. Soc., Providence, RI (1988), 169-198.
155. Goursat, É., Sur les substitutions orthogonales et les divisions régulières de l'espace, *Ann. Sci. École Norm. Sup.*, 6 (1889), 9-102.
156. Grassmann, H., *Die lineale Ausdehnungslehre*, Leipzig (1844) (*Gesammelte mathematische und physikalische Werke*, Vol. 1, Pt. 1, B. G. Teubner, Leipzig, 1894).
157. Grassmann, H., *Die lineale Ausdehnungslehre. Vollständig und in strender Form bearbeitet*, Berlin (1862) (*Gesammelte mathematische und physikalische Werke*, Vol. 1, Pt. 2, B. G. Teubner, Leipzig, 1896).
158. Grassmann, H., On the Ausdehnungslehre, In: *A Source Book In Mathematics*, D. E. Smith, Dover, New York (1959), 684-696.
159. Gray, J., Non-Euclidean geometry – a re-interpretation, *Historia Math.*, 6 (1979), 236-258.
160. Gray, J., *Linear Differential Equations and Group Theory from Riemann to Poincaré*, Birkhäuser, Boston (1986).
161. Gray, J., The discovery of non-Euclidean geometry, In: *Studies in the History of Mathematics*, edited by E. R. Phillips, MAA Studies Math., 26, Math. Assoc. America, Washington DC (1987), 37-60.
162. Green, G., On the determination of the exterior and interior attractions of ellipsoids of variable densities, *Trans. Cambridge Philos. Soc.*, 5 (1835), 395-430.
163. Greenberg, L., Discrete groups of motions, *Canad. J. Math.*, 12 (1960), 414-426.
164. Greenberg, L., Discrete subgroups of the Lorentz group, *Math. Scand.*, 10 (1962), 85-107.
165. Greenberg, L., Commensurable groups of Moebius transformations, In: *Discontinuous Groups and Riemann Surfaces*, edited by L. Greenberg, Ann. of Math. Studies, 79, Princeton Univ. Press, Princeton, NJ (1974), 227-237.

166. Greenberg, M. J., *Euclidean and Non-Euclidean Geometries, 3rd Ed.*, W. H. Freeman, New York (1993).
167. Gromov, M., Manifolds of negative curvature, *J. Diff. Geometry*, 13 (1978), 223-230.
168. Gromov, M., Hyperbolic manifolds according to Thurston and Jørgensen, In: *Séminaire Bourbaki vol. 1979/80*, Lecture Notes in Math., 842, Springer-Verlag, Berlin (1981), 40-53.
169. Gromov, M., Volume and bounded cohomology, *Inst. Hautes Études Sci. Publ. Math.*, 56 (1982), 5-99.
170. Gromov, M. and Pansu, P., Rigidity of lattices: An introduction, In: *Geometric Topology: Recent Developments*, edited by P. de Bartolomeis and F. Tricerri, Lecture Notes in Math., 1504, Springer-Verlag, Berlin (1991), 39-137.
171. Gromov, M. and Piatetski-Shapiro, I., Non-arithmetic groups in Lobachevsky spaces, *Inst. Hautes Études Sci. Publ. Math.*, 66 (1988), 93-103.
172. Grünbaum, B., *Convex Polytopes*, Pure Appl. Math., 16, Interscience, London (1967).
173. Gucul, I. S., On a series of compact three-dimensional manifolds of constant negative curvature, *Soviet Math. Dokl.*, 20 (1979), 996-999.
174. Haagerup, U. and Munkholm, H. J., Simplices of maximal volume in hyperbolic n-space, *Acta. Math.*, 147 (1981), 1-11.
175. Haefliger, A., Orbi-espaces, In: *Sur les groupes hyperboliques d'après Mikhael Gromov*, edited by E. Ghys and P. de la Harpe, Progress Math., 83, Birkhäuser, Boston (1990), 203-213.
176. Halpern, N., A proof of the collar lemma, *Bull. London Math. Soc.*, 13 (1981), 141-144.
177. Hamilton, W., On quaternions; or on a new system of imaginaries in algebra, *Philos. Mag.*, 25-36 (1844-1850), (*The Mathematical Papers of Sir William Rowen Hamilton*, Vol. III, Cambridge Univ. Press, Cambridge (1967), 227-297).
178. Hantzsche, W. and Wendt, H., Dreidimensionale euklidische Raumformen, *Math. Ann.*, 110 (1935), 593-611.
179. Harvey, W. J., Spaces of discrete groups, In: *Discrete Groups and Automorphic Functions*, edited by W. J. Harvey, Academic Press, London (1977), 295-348.
180. Hausdorff, F., Analytische Beiträge zur nichteuklidischen Geometrie, *Ber. Verh. König. Sächs. Ges. Wiss. Leipzig. Math.-Phys. Kl.*, 51 (1899), 161-214.
181. Hausdorff, F., *Grundzüge der Mengenlehre*, de Gruyter, Leipzig (1914).
182. Hawkins, T., Another look at Cayley and the theory of matrices, *Arch. Internat. Hist. Sci.*, 27 (1977), 82-112.
183. Hawkins, T., Weierstrass and the theory of matrices, *Arch. Hist. Exact. Sci.*, 17 (1977), 119-163.
184. Hawkins, T., Non-Euclidean geometry and Weierstrassian mathematics: The background to Killing's work on Lie algebras, *Historia Math.*, 7 (1980), 289-342.
185. Hawkins, T., The Erlanger Programm of Felix Klein: Reflections on its place in the history of mathematics, *Historia Math.*, 11 (1984), 442-470.

186. Heath, T., *A History of Greek Mathematics*, Vols. I, II, Clarendon Press, Oxford (1921) (Dover, New York, 1981).
187. Hedlund, G. A., Fuchsian groups and transitive horocycles, *Duke Math. J.*, 2 (1936), 530-542.
188. Helgason, S., *Differential Geometry, Lie Groups, and Symmetric Spaces*, Pure Appl. Math., 80, Academic Press, New York (1978).
189. Hermite, C., Sur la théorie des formes quadratiques, *J. Reine Angew. Math.*, 47 (1854), 343-368.
190. Hilbert, D., Ueber Flächen von constanter Gaussscher Krümmung, *Trans. Amer. Math. Soc.*, 2 (1901), 87-99.
191. Hilden, H. M., Lozano, M. T., and Montesinos, J. M., On the Borromean orbifolds: Geometry and arithmetic, In: *Topology' 90*, edited by B. Apanasov, W. D. Neumann, A. W. Reid, and L. Siebenmann, de Gruyter, Berlin (1992), 133-167.
192. Hilden, H. M., Lozano, M. T., and Montesinos, J. M., The arithmeticity of the figure eight knot orbifolds, In: *Topology' 90*, edited by B. Apanasov, W. D. Neumann, A. W. Reid, and L. Siebenmann, de Gruyter, Berlin (1992), 169-183.
193. Hilden, H. M., Lozano, M. T., Montesinos, J. M., and Whitten, W. C., On universal groups and three-manifolds, *Invent. Math.*, 87 (1987), 441-456.
194. Hirsch, M. W., *Differential Topology*, Graduate Texts in Math., 33, Springer-Verlag, New York (1976).
195. Hodgson, C. D., Meyerhoff, G. R., and Weeks, J. R., Surgeries on the Whitehead link yield geometrically similar manifolds, In: *Topology' 90*, edited by B. Apanasov, W. D. Neumann, A. W. Reid, and L. Siebenmann, de Gruyter, Berlin (1992), 195-206.
196. Hodgson, C. D. and Rivin, I., A characterization of compact convex polyhedra in hyperbolic 3-space, *Invent. Math.*, 111 (1993), 77-111.
197. Hodgson, C. D., Rivin, I., and Smith, W. D., A characterization of convex hyperbolic polyhedra and of convex polyhedra inscribed in the sphere, *Bull. Amer. Math. Soc.*, 27 (1992), 246-251.
198. Hopf, H., Zum Clifford-Kleinschen Raumproblem, *Math. Ann.*, 95 (1926), 313-339.
199. Hopf, H. and Rinow, W., Ueber den Begriff der vollständigen differentialgeometrischen Fläche, *Comment. Math. Helv.*, 3 (1931), 209-225.
200. Houzel, C., The birth of non-Euclidean geometry, In: *1830-1930: A Century of Geometry*, edited by L. Boi, D. Flament, and J.-M. Salanskis, Lecture Notes in Physics, 402, Springer-Verlag, Berlin (1992), 3-21.
201. Humphreys, J. E., *Reflection Groups and Coxeter Groups*, Cambridge Studies Adv. Math., 29, Cambridge Univ. Press, Cambridge (1990).
202. Jacobi, C. G. J., De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares in alias binas transformandis, quae solis quadratis variabilium constant, *J. Reine Angew. Math.*, 12 (1834), 1-69.
203. Jordan, C., La déformation des surfaces, *J. Math. Pures Appl.*, 11 (1866), 105-109.
204. Jordan, C., Recherches sur les polyèdres, *C. R. Acad. Sci. Paris*, 62 (1866), 1339-1341.
205. Jordan, C., Sur les groupes de mouvements, *C. R. Acad. Sci. Paris*, 65 (1867), 229-232.

206. Jordan, C., Mémoire sur les groupes de mouvements, *Ann. Mat. Pura Appl.*, 2 (1869), 167-215, 322-345.
207. Jordan, C., Essai sur la géométrie à n dimensions, *Bull. Soc. Math. France*, 3 (1875), 103-174.
208. Jordan, C., Mémoire sur les équations différentielles linéaires à intégrale algébrique, *J. Reine Angew. Math.*, 84 (1878), 89-215.
209. Jordan, C., Sur la détermination des groupes d'ordre fini contenus dans le groupe linéaire, *Atti Accad. Sci. Napoli*, 8 (1879), 1-41.
210. Jordan, C., *Cours d'Analyse, Vol. I, 2nd Ed.*, Gauthier-Villars, Paris (1893).
211. Jørgensen, T., A note on subgroups of $\mathrm{SL}(2,\mathbb{C})$, *Quart. J. Math. Oxford*, 28 (1977), 209-212.
212. Jørgensen, T., Compact 3-manifolds of constant negative curvature fibering over the circle, *Ann. of Math.*, 106 (1977), 61-72.
213. Každan, D. A. and Margulis, G. A., A proof of Selberg's conjecture, *Math. USSR-Sbornik*, 4 (1968), 147-152.
214. Keen, L., On Fricke moduli, In: *Advances in the Theory of Riemann Surfaces*, edited by L. V. Ahlfors, L. Bers, H. M. Farkas, R. C. Gunning, I. Kra, and H. E. Rauch, Ann. of Math. Studies, 66, Princeton Univ. Press, Princeton, NJ (1971), 205-224.
215. Keen, L., A correction to "On Fricke moduli", *Proc. Amer. Math. Soc.*, 40 (1973), 60-62.
216. Keen, L., Collars on Riemann surfaces, In: *Discontinuous groups and Riemann Surfaces*, edited by L. Greenberg, Ann. of Math. Studies, 79, Princeton Univ. Press, Princeton, NJ (1974), 263-268.
217. Kellerhals, R., On the volume of hyperbolic polyhedra, *Math. Ann.*, 285 (1989), 541-569.
218. Kellerhals, R., The dilogarithm and volumes of hyperbolic polytopes, In: *Structural Properties of Polylogarithms*, edited by L. Lewin, Math. Surveys Monographs, 37, Amer. Math. Soc., Providence, RI (1991), 301-336.
219. Killing, W., Ueber zwei Raumformen mit constanter positiver Krümmung, *J. Reine Angew. Math.*, 86 (1878), 72-83.
220. Killing, W., Die Rechnung in den Nicht-Euklidischen Raumformen, *J. Reine Angew. Math.*, 89 (1880), 265-287.
221. Killing, W., *Die Nicht-Euklidischen Raumformen in analytischer Behandlung*, B. G. Teubner, Leipzig (1885).
222. Killing, W., Ueber die Clifford-Klein'schen Raumformen, *Math. Ann.*, 39 (1891), 257-278.
223. Klee, V., Some characterizations of convex polyhedra, *Acta Math.*, 102 (1959), 79-107.
224. Klein, F., Ueber die sogenannte Nicht-Euklidische Geometrie, *Math. Ann.*, 4 (1871), 573-625 (Sur la géométrie dite non Euclidienne, *Ann. Fac. Sci. Toulouse*, 11 (1897), 1-62).
225. Klein, F., Ueber Liniengeometrie und metrische Geometrie, *Math. Ann.*, 5 (1872), 257-277.
226. Klein, F., *Vergleichende Betrachtungen über neuer geometrische Forschungen*, Erlangen (1872) (*Math. Ann.*, 43 (1893), 63-100; A comparative review of recent researches in geometry, *Bull. New York Math. Soc.*, 2 (1893), 215-249).

227. Klein, F., Ueber die sogenannte Nicht-Euklidische Geometrie (Zweiter Aufsatz.), *Math. Ann.*, 6 (1873), 112-145.
228. Klein, F., Bemerkungen über den Zusammenhang der Flächen, *Math. Ann.*, 7 (1874), 549-557.
229. Klein, F., Ueber binäre Formen mit linearen Transformationen in sich selbst, *Math. Ann.*, 9 (1876), 183-208.
230. Klein, F., Ueber den Zusammenhang der Flächen, *Math. Ann.*, 9 (1876), 476-483.
231. Klein, F., Ueber die Transformation der elliptischen Functionen und die Auflösung der Gleichungen fünften Grades, *Math. Ann.*, 14 (1879), 111-172.
232. Klein, F., *Ueber Riemanns Theorie der algebraischen Functionen und ihrer Integrale*, B. G. Teubner, Leipzig (1882) (*On Riemann's Theory of Algebraic Functions and their Integrals*, translated by F. Hardcastle, Macmillan and Bowes, Cambridge, 1893).
233. Klein, F., Neue Beiträge zur Riemannschen Functionentheorie, *Math. Ann.*, 21 (1883), 141-218.
234. Klein, F., Zur Nicht-Euklidischen Geometrie, *Math. Ann.*, 37 (1890), 544-572.
235. Klein, F., Ueber den Begriff des functionentheoretischen Fundamentalbereichs, *Math. Ann.*, 40 (1891), 130-139.
236. Klein, F., Über die geometrischen Grundlagen der Lorentzgruppe, *Jber. Deutsch. Math. Verein.*, 19 (1910), 281-300.
237. Klein, F., *Vorlesungen über nicht-euklidische Geometrie*, Grundlehren Math. Wiss., 26, Springer-Verlag, Berlin (1928).
238. Klein, F., *Development of Mathematics in the 19th Century*, translated by M. Ackerman, Lie Groups: History, Frontiers, and Applications, Vol. 9, Math Sci Press, Brookline, MA (1979).
239. Koebe, P., Ueber die Uniformisierung beliebiger analytischer Kurven, *Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl.*, (1907), 191-210.
240. Koebe, P., Allgemeine Theorie der Riemannschen Mannigfaltigkeiten, *Acta Math.*, 50 (1927), 27-157.
241. Koebe, P., Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen. I, *Sitzungsber. Preuss. Akad. Wiss. Berlin. Math.-Phys. Kl.*, (1927), 164-196.
242. Koebe, P., Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen. II, *Sitzungsber. Preuss. Akad. Wiss. Berlin. Math.-Phys. Kl.*, (1928), 345-384.
243. Koebe, P., Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen. III, *Sitzungsber. Preuss. Akad. Wiss. Berlin. Math.-Phys. Kl.*, (1928), 385-442.
244. Koebe, P., Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen. IV, *Sitzungsber. Preuss. Akad. Wiss. Berlin. Math.-Phys. Kl.*, (1929), 414-457.
245. Koebe, P., Riemannsche Mannigfaltigkeiten und nichteuklidische Raumformen. V, *Sitzungsber. Preuss. Akad. Wiss. Berlin. Math.-Phys. Kl.*, (1930), 304-364.
246. Koecher, M. and Roelcke, W., Diskontinuierliche und diskrete Gruppen von Isometrien metrischer Räume, *Math. Z.*, 71 (1959), 258-267.

247. Kuiper, N. H., Compact spaces with a local structure determined by the group of similarity transformations in E^n, *Indag. Math.*, 12 (1950), 411-418.
248. Kulkarni, R. S., Conjugacy classes in M(n), In: *Conformal Geometry*, edited by R. S. Kulkarni and U. Pinkall, Aspects of Math., E12, Vieweg, Braunschweig (1988), 41-63.
249. Lagrange, J., Nouvelle solution du problème du mouvement de rotation d'un corps de figure quelconque qui n'est animé par aucune force accélératrice, *Nouv. Mem. Acad. Roy. Sci. Bel.-Let. Berlin* (1773) (*Oeuvres de Lagrange*, 3, 579-616).
250. Lagrange, J., Sur l'attraction des sphéroides elliptiques, *Nouv. Mem. Acad. Roy. Sci. Bel.-Let. Berlin* (1773) (*Oeuvres de Lagrange*, 3, 619-658).
251. Lambert, J. H., Observations trigonométriques, *Mem. Acad. Sci. Berlin*, 24 (1770), 327-354 (*Opera Mathematica*, Vol. 2, Orell Füssli Verlag, Zürich (1948), 245-269).
252. Lambert, J. H., Theorie der Parallellinien, *Mag. Reine Angew. Math.*, (1786), 137-164, 325-358 (In: *Die Theorie der Parallellinien von Euklid bis auf Gauss*, F. Engel and P. Stäckel, B. G. Teubner, Leipzig (1895), 152-208).
253. Lannér, F., On complexes with transitive groups of automorphisms, *Med. Lunds Univ. Math. Sem.*, 11 (1950), 1-71.
254. Lawson, H. B. and Yau, S. T., Compact manifolds of nonpositive curvature, *J. Diff. Geometry*, 7 (1972), 211-228.
255. Lehner, J., *Discontinuous Groups and Automorphic Functions*, Math. Surveys, 8, Amer. Math. Soc., Providence, RI (1964).
256. Leutbecher, A., Über Spitzen diskontinuierlicher Gruppen von lineargebrochenen Transformationen, *Math. Z.*, 100 (1967), 183-200.
257. Lhuilier, M., Mémoire sur la polyédrométrie; contenant une démonstration directe du Théorème d'Euler sur les polyèdres, et un examen des diverses exceptions auxquelles ce théorème est assujetti, *Ann. Math. Pures Appl.*, 3 (1813), 169-189.
258. Lie, S., Ueber diejenige Theorie eines Raumes mit beliebig vielen Dimensionen, die der Krümmungs-Theorie des gewöhnlichen Raumes entspricht, *Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl.*, (1871), 191-209.
259. Liouville, J., Note au suject de l'article précédent, *J. Math. Pures Appl.*, 12 (1847), 265-290.
260. Liouville, J., Sur le théorème de M. Gauss, concernant le produit des deux rayons de courbure principaux en chaque point d'une surface, In: *Application de l'Analyse à la Géométrie*, G. Monge, Paris (1850), 583-600.
261. Liouville, J., Extension au cas des trois dimensions de la question du tracé géographique, In: *Application de l'Analyse à la Géométrie*, G. Monge, Paris (1850), 609-616.
262. Lobachevsky, N. I., On the principles of geometry (Russian), *Kazanskij Vestnik*, 25-28 (1829-1830) (Ueber die Anfangsgründe der Geometrie, In: *Nikolaj Iwanowitsch Lobatschefskij. Zwei geometrische Abhandlungen*, edited by F. Engel, B. G. Teubner, Leipzig (1898), 1-66).
263. Lobachevsky, N. I., *Application of Imaginary Geometry to Certain Integrals* (Russian), Kazan (1836) (*N. J. Lobatschefskijs Imaginäre Geometrie und Anwendung der imaginären Geometrie auf einige Integrale*, edited by H. Liebmann, B. G. Teubner, Leipzig, 1904).

264. Lobatschewsky, N., Géométrie imaginaire, *J. Reine Angew. Math.*, 17 (1837), 295-320.
265. Löbell, F., *Die überall regulären unbegrenzten Flächen fester Krümmung*, Dissertation, Univ. Tübingen (1927).
266. Löbell, F., Über die geodätischen Linien der Clifford-Kleinschen Flächen, *Math. Z.*, 30 (1929), 572-607.
267. Löbell, F., Ein Satz über die eindeutigen Bewegungen Clifford-Kleinscher Flächen in sich, *J. Reine Angew. Math.*, 162 (1930), 114-124.
268. Löbell, F., Beispiele geschlossener dreidimensionaler Clifford-Kleinscher Räume negativer Krümmung, *Ber. Verh. Sächs. Akad. Wiss. Leipzig. Math.-Phys. Kl.*, 83 (1931), 167-174.
269. Loewy, A., Ueber bilineare Formen mit conjugirt imaginären Variabeln, *Math. Ann.*, 50 (1898), 557-576.
270. Lohne, J. A., Essays on Thomas Harriot, *Arch. Hist. Exact Sci.*, 20 (1979), 189-312.
271. Lorentz, H. A., Electromagnetic phenomena in a system moving with any velocity less than that of light, *Proc. Acad. Sci. Amsterdam*, 6 (1904), 809 (In: *The Principle of Relativity*, H. A. Lorentz, A. Einstein, H. Minkowski, and H. Weyl, Dover, New York (1952), 9-34).
272. Magnus, W., *Noneuclidean Tesselations and their Groups*, Pure Appl. Math., 61, Academic Press, New York (1974).
273. Mandelbrot, B. B., Self-inverse fractals osculated by sigma-discs and the limit sets of inversion groups, *Math. Intelligencer*, 5 (1983), 9-17.
274. Mangler, W., Die Klassen von topologischen Abbildungen einer geschlossenen Fläche auf sich, *Math. Z.*, 44 (1938), 541-554.
275. Marden, A., Universal properties of Fuchsian groups in the Poincaré metric, In: *Discontinuous Groups and Riemann Surfaces*, edited by L. Greenberg, Ann. of Math. Studies, 79, Princeton Univ. Press, Princeton, NJ (1974), 315-339.
276. Marden, A., The geometry of finitely generated Kleinian groups, *Ann. of Math.*, 99 (1974), 383-462.
277. Margulis, G. A., On the arithmeticity of discrete groups, *Soviet Math. Dokl.*, 10 (1969), 900-902.
278. Margulis, G. A., Isometry of closed manifolds of constant negative curvature with the same fundamental group, *Soviet Math. Dokl.*, 11 (1970), 722-723.
279. Martin, G. J., Balls in hyperbolic manifolds, *J. London Math. Soc.*, 40 (1989), 257-264.
280. Martin, G. J., On discrete Möbius groups in all dimensions: A generalization of Jørgensen's inequality, *Acta Math.*, 163 (1989), 253-289.
281. Maskit, B., On Poincaré's theorem for fundamental polygons, *Adv. Math.*, 7 (1971), 219-230.
282. Maskit, B., *Kleinian Groups*, Grundlehren Math. Wiss., 287, Springer-Verlag, Berlin (1988).
283. Massey, W. S., *Algebraic Topology: An Introduction*, Graduate Texts in Math., 56, Springer-Verlag, New York (1967).
284. Matveev, S. V. and Fomenko, A. T., Constant energy surfaces of Hamiltonian systems, enumeration of three-dimensional manifolds in increasing

order of complexity, and computation of volumes of closed hyperbolic manifolds, *Russian Math. Surveys*, 43 (1988), 3-24.
285. McMullen, C., Riemann surfaces and the geometrization of 3-manifolds, *Bull. Amer. Math. Soc.*, 27 (1992), 207-216.
286. Menger, K., Untersuchungen über allgemeine Metrik. Vierte Untersuchung. Zur Metrik der Kurven, *Math. Ann.*, 103 (1930), 466-501.
287. Meyerhoff, R., The cusped hyperbolic 3-orbifold of minimum volume, *Bull. Amer. Math. Soc.*, 13 (1985), 154-156.
288. Miller, A. I., A study of Henri Poincaré's "Sur la dynamique de l'électron", *Arch. Hist. Exact Sci.*, 10 (1973), 208-328.
289. Millson, J. J., On the first Betti number of a constant negatively curved manifold, *Ann. of Math.*, 104 (1976), 235-247.
290. Milnor, J., Hyperbolic geometry: The first 150 years, *Bull. Amer. Math. Soc.*, 6 (1982), 9-24.
291. Milnor, T. K., Efimov's theorem about complete immersed surfaces of negative curvature, *Adv. Math.*, 8 (1972), 474-543.
292. Minding, F., Wie sich entscheiden läfst, ob zwei gegebene krumme Flächen auf einander abwickelbar sind oder nicht, *J. Reine Angew. Math.*, 19 (1839), 370-387.
293. Minkowski, H., Die Grundgleichungen für die elektromagnetischen Vorgänge in bewegten Körpern, *Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl.*, (1908), 53-111.
294. Minkowski, H., Raum and Zeit, *Jber. Deutsch. Math.-Verein.*, 18 (1909), 75-88 (Space and Time, In: *The Principle of Relativity*, H. A. Lorentz, A. Einstein, H. Minkowski, and H. Weyl, Dover, New York (1952), 73-96).
295. Minkowski, H., Theorie der konvexen Körper, insbesondere Begründung ihres Oberflächenbegriffs, *Gesammelte Abhandlungen von Hermann Minkowski*, Vol. 2, B. G. Teubner, Leipzig (1911), 131-229.
296. Minkowski, H., Das Relativitätsprinzip, *Ann. Physik*, 47 (1915), 927-938.
297. Möbius, A. F., Ueber eine neue Verwandtschaft zwischen ebenen Figuren, *Ber. Verh. Königl. Säch. Ges. Wiss. Math.-Phys. Kl.*, 5 (1852), 14-24 (*J. Reine Angew. Math.* 52 (1856), 218-228).
298. Möbius, A. F., Die Theorie der Kreisverwandtschaft in rein geometrischer Darstellung, *Abh. Königl. Sächs. Ges. Wiss. Math.-Phys. Kl.*, 2 (1855), 529-595.
299. Möbius, A. F., Theorie der elementaren Verwandtschaft, *Ber. Verh. Sächs. Akad. Wiss. Leipzig. Math.-Phys. Kl.*, 15 (1863), 18-57.
300. Möbius, A. F., Ueber die Bestimmung des Inhaltes eines Polyëders, *Ber. Verh. Sächs. Akad. Wiss. Leipzig. Math.-Phys. Kl.*, 17 (1865), 31-68.
301. Möbius, A. F., Zur Theorie der Polyëder und der Elementarverwandtschaft, *Gesammelte Werke*, Vol. 2, S. Hirzel, Leipzig (1886), 513-560.
302. Moise, E. E., *Geometric Topology in Dimensions 2 and 3*, Graduate Texts in Math., 47, Springer-Verlag, New York (1977).
303. Molnár, E., Two hyperbolic football manifolds, In: *Differential Geometry and Applications*, Proceedings of the Conference, 1988, Dubrovik, edited by N. Bokan, I. Čomić, J. Nikić, and M. Prvanović, Univ. Novi Sad, Novi Sad (1989), 217-241.
304. Moore, E. H., An universal invariant for finite groups of linear substitutions, *Math. Ann.*, 50 (1898), 213-219.

305. Morgan, J. W., On Thurston's uniformization theorem for three-dimensional manifolds, In: *The Smith Conjecture*, edited by J. W. Morgan and H. Bass, Pure Appl. Math., 112, Academic Press, New York (1984), 37-125.
306. Morokuma, T., A characterization of fundamental domains of discontinuous groups acting on real hyperbolic spaces, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 25 (1978), 157-183.
307. Morse, H. M., A fundamental class of geodesics on any closed surface of genus greater than one, *Trans. Amer. Math. Soc.*, 26 (1924), 25-60.
308. Mostow, G. D., Quasi-conformal mappings in n-space and the rigidity of hyperbolic space forms, *Inst. Hautes Études Sci. Publ. Math.*, 34 (1968), 53-104.
309. Mostow, G. D., The rigidity of locally symmetric spaces, *Actes, Congrès Intern. Math.*, 2 (1970), 187-197.
310. Mostow, G. D., *Strong Rigidity of Locally Symmetric Spaces*, Ann. of Math. Studies, 78, Princeton Univ. Press, Princeton, NJ (1973).
311. Munkholm, H. J., Simplices of maximal volume in hyperbolic space, Gromov's norm, and Gromov's proof of Mostow's rigidity theorem (following Thurston), In: *Topology Symposium, Siegen 1979*, edited by U. Koschorke and W. D. Neumann, Lecture Notes in Math., 788, Springer-Verlag, Berlin (1980), 109-124.
312. Myrberg, P. J., Die Kapazität der singulären Menge der linearen Gruppen, *Ann. Acad. Sci. Fenn. Ser. A. I. Math.-Phys.*, 10 (1941), 1-19.
313. Naber, G. L., *The Geometry of Minkowski Spacetime*, Appl. Math. Sci., 92, Springer-Verlag, New York (1992).
314. Neumann, W. D. and Zagier, D., Volumes of hyperbolic three-manifolds, *Topology*, 24 (1985), 307-332.
315. Newcomb, S., Elementary theorems relating to the geometry of a space of three dimensions and of uniform positive curvature in the fourth dimension, *J. Reine Angew. Math.*, 83 (1877), 293-299.
316. Nicholls, P. J., The limit set of a discrete group of hyperbolic motions, In: *Holomorphic Functions and Moduli II*, edited by D. Drasin, C. J. Earle, F. W. Gehring, I. Kra, and A. Marden, Springer-Verlag, New York (1988), 141-164.
317. Nicholls, P. J., *The Ergodic Theory of Discrete Groups*, London Math. Soc. Lecture Notes Ser., 143, Cambridge Univ. Press, Cambridge (1989).
318. Nielsen, J., Über topologische Abbildungen geschlossener Flächen, *Abh. Math. Sem. Univ. Hamburg*, 3 (1924), 246-260.
319. Nielsen, J., Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen, *Acta Math.*, 50 (1927), 189-358 (Investigations in the topology of closed orientable surfaces I, In: *Jakob Nielsen: Collected Mathematical Papers*, Vol. 1, Birkhäuser, Boston (1986), 223-341).
320. Nielsen, J., Über Gruppen linearer Transformationen, *Mitt. Math. Ges. Hamburg*, 8 (1940), 82-104.
321. Nielsen, J. and Fenchel, W., *Discontinuous Groups of Non-Euclidean Motions*, manuscript, Univ. California, Berkeley, CA (1959).
322. Nowacki, W., Die euklidischen, dreidimensionalen, geschlossenen und offenen Raumformen, *Comment. Math. Helv.*, 7 (1934), 81-92.
323. Oliver, R. K., On Bieberbach's analysis of discrete Euclidean groups, *Proc. Amer. Math. Soc.*, 80 (1980), 15-21.

324. Patterson, B. C., The origins of the geometric principle of inversion, *Isis*, 19 (1933), 154-180.
325. Penrose, R., The geometry of the universe, In: *Mathematics Today, Twelve Informal Essays*, edited by L. A. Steen, Springer-Verlag, New York (1978), 83-125.
326. Plücker, J., Analytisch-geometrische Aphorismen, *J. Reine Angew. Math.*, 11 (1834), 219-225.
327. Poincaré, H., Sur les fonctions fuchsiennes, *C. R. Acad. Sci. Paris*, 92 (1881), 333-335 (On Fuchsian functions, In: *Papers on Fuchsian Functions*, translated by J. Stillwell, Springer-Verlag, New York (1985), 47-50).
328. Poincaré, H., Sur les fonctions fuchsiennes, *C. R. Acad. Sci. Paris*, 92 (1881), 1484-1487.
329. Poincaré, H., Sur les groupes kleinéens, *C. R. Acad. Sci. Paris*, 93 (1881), 44-46.
330. Poincaré, H., Théorie des groupes fuchsiens, *Acta Math.*, 1 (1882), 1-62 (Theory of Fuchsian groups, In: *Papers on Fuchsian Functions*, translated by J. Stillwell, Springer-Verlag, New York (1985), 55-127).
331. Poincaré, H., Sur les fonctions fuchsiennes, *Acta Math.*, 1 (1882), 193-294 (On Fuchsian functions, In: *Papers on Fuchsian Functions*, translated by J. Stillwell, Springer-Verlag, New York (1985), 128-254).
332. Poincaré, H., Mémoire sur les groupes kleinéens, *Acta Math.*, 3 (1883), 49-92 (Memoir on Kleinian groups, In: *Papers on Fuchsian Functions*, translated by J. Stillwell, Springer-Verlag, New York (1985), 255-304).
333. Poincaré, H., Sur les groupes des équations linéaires, *Acta Math.*, 4 (1884), 201-311 (On the groups of linear equations, In: *Papers on Fuchsian Functions*, translated by J. Stillwell, Springer-Verlag, New York (1985), 357-483).
334. Poincaré, H., Sur un théorème de M. Fuchs, *Acta Math.*, 7 (1885), 1-32.
335. Poincaré, H., Sur les hypothèses fondamentales de la géométrie, *Bull. Soc. Math. France*, 15 (1887), 203-216.
336. Poincaré, H., Analysis situs, *J. École Polytech.*, 1 (1895), 1-121.
337. Poincaré, H., Cinquième complément à l'analysis situs, *Rend. Circ. Mat. Palermo*, 18 (1904), 45-110.
338. Poincaré, H., Sur la dynamique de l'électron, *Rend. Circ. Mat. Palermo*, 21 (1906), 129-176.
339. Poincaré, H., Sur l'uniformisation des fonctions analytiques, *Acta Math.*, 31 (1907), 1-63.
340. Poincaré, H., *Papers on Fuchsian Functions*, translated by J. Stillwell, Springer-Verlag, New York (1985).
341. Poncelet, J. V., *Traité des propriétés projectives des figures*, Paris (1822).
342. Pont, J.-C., *La Topologie Algébrique des origines à Poincaré*, Presses Universitaires de France, Paris (1974).
343. Pontrjagin, L., *Topological Groups*, translated by E. Lehmer, Princeton Univ. Press, Princeton, NJ (1939).
344. Prasad, G., Strong rigidity of $\mathbb{Q}$-rank 1 lattices, *Invent. Math.*, 21 (1973), 255-286.
345. Pyenson, L., Hermann Minkowski and Einstein's Special Theory of Relativity, *Arch. Hist. Exact Sci.*, 17 (1977), 71-95.

346. Ranum, A., Lobachefskian polygons trigonometrically equivalent to the triangle, *Jber. Deutsch. Math. Verein.*, 21 (1912), 228-240.

347. Ratcliffe, J. G., Euler characteristics of 3-manifold groups and discrete subgroups of SL(2, $\mathbb{C}$), *J. Pure Appl. Algebra*, 44 (1987), 303-314.

348. Riemann, B., Theorie der Abel'schen Functionen, *J. Reine Angew. Math.*, 54 (1857), 115-155.

349. Riemann, B., Über die Hypothesen, welche der Geometrie zu Grunde liegen, *Abh. Ges. Wiss. Göttingen*, 13 (1867), 133-152.

350. Riley, R., A quadratic parabolic group, *Math. Proc. Camb. Phil. Soc.*, 77 (1975), 281-288.

351. Riley, R., An elliptical path from parabolic representations to hyperbolic structures, In: *Topology of Low-Dimensional Manifolds*, edited by R. Fenn, Lecture Notes in Math., 722, Springer-Verlag, Berlin (1979), 99-133.

352. Riley, R., Applications of a computer implementation of Poincaré's theorem on fundamental polyhedra, *Math. Comp.*, 40 (1983), 607-632.

353. Rosenfeld, B. A., *A History of Non-Euclidean Geometry*, translated by A. Shenitzer, Studies Hist. Math. Phys. Sci., 12, Springer-Verlag, New York (1988).

354. Rowe, D. E., Klein, Lie, and the "Erlanger Programm", In: *1830-1930: A Century of Geometry*, edited by L. Boi, D. Flament, and J.-M. Salanskis, Lecture Notes in Physics, 402, Springer-Verlag, Berlin (1992), 45-54.

355. Saccheri, G., *Euclides ab omni naevo vindicatus: sive conatus geometricus quo stabiliuntur prima ipsa universae geometriae principia*, Milan (1733) (*Euclides vindicatus*, translated by G. B. Halsted, Open Court, Chicago, 1920).

356. Satake, I., On a generalization of the notion of manifold, *Proc. Nat. Acad. Sci. U.S.A.*, 42 (1956), 359-363.

357. Schilling, F., Ueber die geometrische Bedeutung der Formeln der sphärischen Trigonometrie im Falle complexer Argumente, *Math. Ann.*, 39 (1891), 598-600.

358. Schilling, F., Beiträge zur geometrischen Theorie der Schwarz'schen s-Function, *Math. Ann.*, 44 (1894), 161-260.

359. Schläfli, L., Réduction d'une intégrale multiple, qui comprend l'arc de cercle et l'aire du triangle sphérique comme cas particuliers, *J. Math. Pures Appl.*, 20 (1855), 359-394.

360. Schläfli, L., On the multiple integral $\int^n dxdy \cdots dz$, whose limits are $p_1 = a_1x + b_1y + \cdots + h_1z > 0$, $p_2 > 0, \ldots, p_n > 0$, and $x^2 + y^2 + \cdots + z^2 < 1$, *Quart. J. Pure Appl. Math.*, 2 (1858), 269-301.

361. Schläfli, L., On the multiple integral $\int^n dxdy \cdots dz$, whose limits are $p_1 = a_1x + b_1y + \cdots + h_1z > 0$, $p_2 > 0, \ldots, p_n > 0$, and $x^2 + y^2 + \cdots + z^2 < 1$, *Quart. J. Pure Appl. Math.*, 3 (1860), 54-68, 97-108.

362. Schläfli, L., *Theorie der vielfachen Kontinuität*, Zürich (1901).

363. Scholz, E., *Geschichte des Mannigfaltigkeitsbegriffs von Riemann bis Poincaré*, Birkhäuser, Boston (1980).

364. Scholz, E., The rise of symmetry concepts in the atomistic and dynamistic schools of crystallography, 1815-1830, *Rev. Hist. Sci.*, 42 (1989), 109-122.

365. Scholz, E., Crystallographic symmetry concepts and group theory (1850-1880), In: *The History of Modern Mathematics II*, edited by D. E. Rowe and J. McCleary, Academic Press, Boston (1989), 3-27.

366. Scholz, E., Riemann's vision of a new approach to geometry, In: *1830-1930: A Century of Geometry*, edited by L. Boi, D. Flament, and J.-M. Salanskis, Lecture Notes in Physics, 402, Springer-Verlag, Berlin (1992), 22-34.
367. Schottky, F., Ueber die conforme Abbildung mehrfach zusammenhängender ebener Flächen, *J. Reine Angew. Math.*, 83 (1877), 300-351.
368. Schreier, O., Abstrakte kontinuierliche Gruppen, *Abh. Math. Sem. Univ. Hamburg*, 4 (1925), 15-32.
369. Schwarz, H. A., Ueber diejenigen Fälle, in welchen die Gaussische hypergeometrische Reihe eine algebraische Function ihres vierten Elementes darstellt, *J. Reine Angew. Math.*, 75 (1873), 292-335.
370. Scott, P., The geometries of 3-manifolds, *Bull. London Math. Soc.*, 15 (1984), 401-487.
371. Seifert, H., Komplexe mit Seitenzuordnung, *Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. II*, (1975), 49-80.
372. Selberg, A., On discontinuous groups in higher-dimensional symmetric spaces, In: *Contributions to Function Theory*, edited by K. Chandrasekharan, Tata Inst. of Fund. Research, Bombay (1960), 147-164.
373. Selberg, A., Recent developments in the theory of discontinuous groups of motions of symmetric spaces, In: *Proceedings of the 15th Scandinavian Congress, Oslo 1968*, edited by K. E. Aubert and W. Ljunggren, Lecture Notes in Math., 118, Springer-Verlag, Berlin (1970), 99-120.
374. Shimizu, H., On discontinuous groups operating on the product of the upper half planes, *Ann. of Math.*, 77 (1963), 33-71.
375. Siegel, C. L., Discontinuous groups, *Ann. of Math.*, 44 (1943), 674-689.
376. Siegel, C. L., Some remarks on discontinuous groups, *Ann. of Math.*, 46 (1945), 708-718.
377. Sommerville, D. M. Y., *Bibliography of Non-Euclidean Geometry, 2nd Ed.*, Chelsea, New York (1970).
378. Spivak, M., *A Comprehensive Introduction to Differential Geometry, 2nd Ed.*, Vol. II, Publish or Perish, Berkeley, CA (1979).
379. Stäckel, P., Bermerkungen zur Geschichte der geodätischen Linien, *Abh. Sächs. Akad. Wiss. Leipzig. Math.-Natur. Kl.*, 45 (1893), 444-467.
380. Steinitz, E., Bedingt konvergente Reihen und konvexe Systeme, *J. Reine Angew. Math.*, 143 (1913), 128-175.
381. Steinitz, E., Bedingt konvergente Reihen und konvexe Systeme (Fortsetzung), *J. Reine Angew. Math.*, 144 (1914), 1-40.
382. Steinitz, E., Bedingt konvergente Reihen und konvexe Systeme (Schluss), *J. Reine Angew. Math.*, 146 (1916), 1-52.
383. Sturm, J. and Shinnar, M., The maximal inscribed ball of a Fuchsian group, In: *Discontinuous Groups and Riemann Surfaces*, edited by L. Greenberg, Ann. of Math. Studies, 79, Princeton Univ. Press, Princeton, NJ (1974), 439-443.
384. Sullivan, D., On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions, In: *Riemann Surfaces and Related Topics*, edited by I. Kra and B. Maskit, Ann. of Math. Studies, 97, Princeton Univ. Press, Princeton, NJ (1980), 465-496.
385. Susskind, P. and Swarup, G. A., Limit sets of geometrically finite hyperbolic groups, *Amer. J. Math.*, 114 (1992), 233-250.

386. Taurinus, F. A., *Geometriae prima elementa*, Cologne (1826) (Stücke aus den Geometriae prima elementa, In: *Die Theorie der Parallellinien von Euklid bis auf Gauss*, F. Engel and P. Stäckel, B. G. Teubner, Leipzig (1895), 267-286).
387. Teichmüller, O., Extremale quasikonforme Abbildungen und quadratische Differentiale, *Abh. Preuss. Akad. Wiss. Math-Nat. Kl.*, 22 (1939), 1-197.
388. Thomson, W., Extrait d'une lettre de M. William Thomson à M. Liouville, *J. Math. Pures Appl.*, 10 (1845), 364-367.
389. Thurston, W. P., *The Geometry and Topology of Three-Manifolds*, lecture notes, Princeton Univ., Princeton, NJ (1979).
390. Thurston, W. P., Three dimensional manifolds, Kleinian groups, and hyperbolic geometry, *Bull. Amer. Math. Soc.*, 6 (1982), 357-381.
391. Traub, G., *The Development of the Mathematical Analysis of Curve Length from Archimedes to Lebesgue*, Dissertation, New York Univ. (1984).
392. Tukia, P., On isomorphisms of geometrically finite Möbius groups, *Inst. Hautes Études. Sci. Publ. Math.*, 61 (1985), 171-214.
393. van Dantzig, D. and van der Waerden, B. L., Über metrisch homogene Räume, *Abh. Math. Sem. Univ. Hamburg*, 6 (1928), 367-376.
394. van Vleck, E. B., On the combination of non-loxodromic substitutions, *Trans. Amer. Math. Soc.*, 20 (1919), 299-312.
395. Veblen, O. and Whitehead, J. H. C., A set of axioms for differential geometry, *Proc. Nat. Acad. Sci. U.S.A.*, 17 (1931), 551-561.
396. Vesnin, A. Y., Three-dimensional hyperbolic manifolds of Löbell type, *Siberian Math. J.*, 28 (1987), 731-734.
397. Vinberg, E. B., Discrete groups generated by reflections in Lobacevskii spaces, *Math. USSR-Sbornik*, 1 (1967), 429-444.
398. Vinberg, E. B., Hyperbolic reflection groups, *Russian Math. Surveys*, 40 (1985), 31-75.
399. von Helmholtz, H., On the origin and significance of geometrical axioms, *Mind*, 1 (1876), 302-321 (In: *The World of Mathematics I*, J. R. Newman, Simon and Schuster, New York (1956), 647-668).
400. Wall, C. T. C., Geometries and geometric structures in real dimension 4 and complex dimension 2, In: *Geometry and Topology*, edited by J. Alexander and J. Harer, Lecture Notes in Math., 1167, Springer-Verlag, Berlin (1985), 268-292.
401. Wang, H.-C., On a maximality property of discrete subgroups with fundamental domain of finite measure, *Amer. J. Math.*, 89 (1967), 124-132.
402. Wang, H.-C., Discrete nilpotent subgroups of Lie groups, *J. Diff. Geometry*, 3 (1969), 481-492.
403. Wang, H.-C., Topics on totally discontinuous groups, In: *Symmetric Spaces*, edited by W. M. Boothby and G. L. Weiss, Pure Appl. Math., 8, Marcel Dekker, New York (1972), 459-487.
404. Waterman, P. L., Purely elliptic Möbius groups, In: *Holomorphic Functions and Moduli II*, edited by D. Drasin, C. J. Earle, F. W. Gehring, I. Kra, and A. Marden, Springer-Verlag, New York (1988), 173-178.
405. Weber, C. and Seifert, H., Die beiden Dodekaederräume, *Math. Z.*, 37 (1933), 237-253.
406. Weeks, J. R., *The Shape of Space*, Pure Appl. Math., 96, Marcel Dekker, New York (1985).

407. Weeks, J. R., Convex hulls and isometries of cusped hyperbolic 3-manifolds, *Topology Appl.*, 52 (1993), 127-149.
408. Weil, A., On discrete subgroups of Lie groups, *Ann. of Math.*, 72 (1960), 369-384.
409. Whitehead, J. H. C., Locally homogeneous spaces in differential geometry, *Ann. of Math.*, 33 (1932), 681-687.
410. Whitehead, J. H. C., On the covering of a complete space by the geodesics through a point, *Ann. of Math.*, 36 (1935), 679-704.
411. Wielenberg, N., Discrete Moebius groups: fundamental polyhedra and convergence, *Amer. J. Math.*, 99 (1977), 861-877.
412. Wielenberg, N., The structure of certain subgroups of the Picard group, *Math. Proc. Camb. Phil. Soc.*, 84 (1978), 427-436.
413. Wielenberg, N., Hyperbolic 3-manifolds which share a fundamental polyhedron, In: *Riemann Surfaces and Related Topics*, edited by I. Kra and B. Maskit, Ann. of Math. Studies, 97, Princeton Univ. Press, Princeton, NJ (1980), 505-513.
414. Witt, E., Spiegelungsgruppen und Aufzählung halbeinfacher Liescher Ringe, *Abh. Math. Sem. Univ. Hamburg*, 14 (1941), 289-322.
415. Woepcke, F., Recherches sur l'histoire des sciences mathématiques ches les orientaux, d'après des traités inédits arabes et persans, *J. Asiatique*, 5 (1855), 309-359.
416. Wolf, J. A., *Spaces of Constant Curvature, 5th Ed.*, Publish or Perish, Wilmington, DE (1984).
417. Wolpert, S., The Fenchel-Nielsen deformation, *Ann. of Math.*, 115 (1982), 501-528.
418. Wussing, H., *The Genesis of the Abstract Group Concept*, translated by A. Shenitzer, MIT Press, Cambridge, MA (1984).
419. Yaglom, I. M., *A Simple Non-Euclidean Geometry and Its Physical Basis*, translated by A. Shenitzer, Springer-Verlag, New York (1979).
420. Yaglom, I. M., *Felix Klein and Sophus Lie*, translated by S. Sossinsky, Birkhäuser, Boston (1988).
421. Zassenhaus, H., Beweis eines Satzes über diskrete Gruppen, *Abh. Math. Sem. Univ. Hamburg*, 12 (1938), 289-312.
422. Zassenhaus, H., Über einen Algorithmus zur Bestimmung der Raumgruppen, *Comment. Math. Helv.*, 21 (1948), 117-141.

Index

Graduate Texts in Mathematics

continued from page ii

97 KOBLITZ. Introduction to Elliptic Curves and Modular Forms. 2nd ed.
98 BRÖCKER/TOM DIECK. Representations of Compact Lie Groups.
99 GROVE/BENSON Finite Reflection Groups. 2nd ed.
100 BERG/CHRISTENSEN/RESSEL. Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions.
101 EDWARDS. Galois Theory.
102 VARADARAJAN. Lie Groups, Lie Algebras and Their Representations.
103 LANG. Complex Analysis. 3rd ed.
104 DUBROVIN/FOMENKO/NOVIKOV. Modern Geometry--Methods and Applications. Part II.
105 LANG. $SL_2(\mathbf{R})$.
106 SILVERMAN. The Arithmetic of Elliptic Curves.
107 OLVER. Applications of Lie Groups to Differential Equations. 2nd ed.
108 RANGE. Holomorphic Functions and Integral Representations in Several Complex Variables.
109 LEHTO. Univalent Functions and Teichmüller Spaces.
110 LANG. Algebraic Number Theory.
111 HUSEMÖLLER. Elliptic Curves.
112 LANG. Elliptic Functions.
113 KARATZAS/SHREVE. Brownian Motion and Stochastic Calculus. 2nd ed.
114 KOBLITZ. A Course in Number Theory and Cryptography. 2nd ed.
115 BERGER/GOSTIAUX. Differential Geometry: Manifolds, Curves, and Surfaces.
116 KELLEY/SRINIVASAN. Measure and Integral. Vol. I.
117 SERRE. Algebraic Groups and Class Fields.
118 PEDERSEN. Analysis Now.
119 ROTMAN. An Introduction to Algebraic Topology.
120 ZIEMER. Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation.
121 LANG. Cyclotomic Fields I and II. Combined 2nd ed.
122 REMMERT. Theory of Complex Functions.
Readings in Mathematics
123 EBBINGHAUS/HERMES et al. Numbers.
Readings in Mathematics
124 DUBROVIN/FOMENKO/NOVIKOV. Modern Geometry--Methods and Applications. Part III.
125 BERENSTEIN/GAY. Complex Variables: An Introduction.
126 BOREL. Linear Algebraic Groups.
127 MASSEY. A Basic Course in Algebraic Topology.
128 RAUCH. Partial Differential Equations.
129 FULTON/HARRIS. Representation Theory: A First Course.
Readings in Mathematics
130 DODSON/POSTON. Tensor Geometry.
131 LAM. A First Course in Noncommutative Rings.
132 BEARDON. Iteration of Rational Functions.
133 HARRIS. Algebraic Geometry: A First Course.
134 ROMAN. Coding and Information Theory.
135 ROMAN. Advanced Linear Algebra.
136 ADKINS/WEINTRAUB. Algebra: An Approach via Module Theory.
137 AXLER/BOURDON/RAMEY. Harmonic Function Theory.
138 COHEN. A Course in Computational Algebraic Number Theory.
139 BREDON. Topology and Geometry.
140 AUBIN. Optima and Equilibria. An Introduction to Nonlinear Analysis.
141 BECKER/WEISPFENNING/KREDEL. Grobner Bases. A Computational Approach to Commutative Algebra.
142 LANG. Real and Functional Analysis. 3rd ed.
143 DOOB. Measure Theory.
144 DENNIS/FARB. Noncommutative Algebra.

145 VICK. Homology Theory. An Introduction to Algebraic Topology. 2nd ed.
146 BRIDGES. Computability: A Mathematical Sketchbook.
147 ROSENBERG. Algebraic K-Theory and Its Applications.
148 ROTMAN. An Introduction to the Theory of Groups. 4th ed.
149 RATCLIFFE. Foundations of Hyperbolic Manifolds.