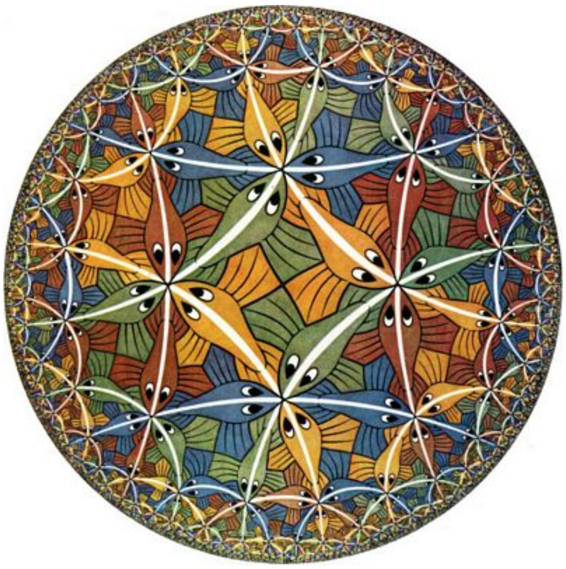
University of Cape Town

MAM3000W - Project

Hyperbolic Geometry & Kleinian Groups



Author:
Greg Jackson

Supervisor:
Dr. Jesse Ratzkin

Updated September 30, 2013

Plagiarism Declaration

(i) I know that plagiarism is wrong. Plagiarism is to use anothers work and to pretend that it is ones own.

(ii) Each significant contribution to, and quotation in, this project from the work of other people has been been cited and referenced.

(iii) This report is my own work.

(iv) I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Name: ______________________________

Student number: ______________________________

Signature: ______________________________

Date: ______________________________

Figure Credits

The artwork used on the cover page is an MC Escher painting titled *Circle Limit III*. All remaining images were hand drawn by the author.

Contents

I. INTRODUCTION

A. Motivation

It is to the ancient Greeks that we owe the first steps in establishing a mathematically grounded theorem of geometry. *Euclid's Elements* [8] remains one of the most important pieces of mathematical literature, despite the modern picture of geometry. As is always the aim with mathematics, an accumulation of effort has resulted in considerable generalisation and extension. Today the pillar of geometry is the study of topology. When a child entertains himself with a ball of modelling clay, he is performing topological manipulations. It is imperative that the clay ball is deformed without tearing or breaking it. A relevant investigation, is how the local characteristics of the geometry are related to the global nature.

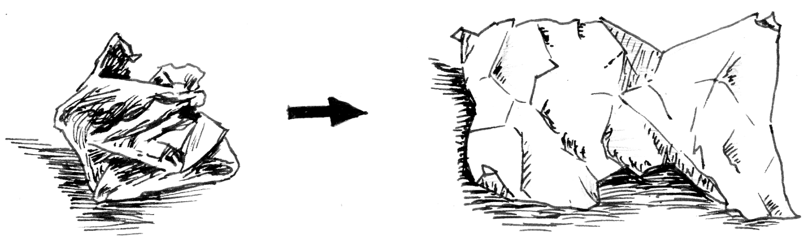
Run-of-the-mill surfaces, both closed (such as the sphere) and those with boundary (such as a disk), have two sides. By this we mean that we can paint the two sides a different colour to distinguish them. If the surface is closed, then these two colours will never touch, and if the surface has a boundary then the colours will meet only at the boundary curve. For example, a torus may be painted red on the outside and blue on the inside. A disk, may be red on the top and blue on the bottom. Möbius is famous for discovering the existence of a surface with only *one* side, the so called Möbius strip. This may be constructed from a rectangle by identifying one of the edges of a rectangle with a half-twist. The Möbius strip demonstrates a wealth of bizarre facts. For instance, if the strip is cut down its centre line, it will produce another strip, now with two sides!

This project is mainly directed at the study of the group structure associated with geometry, in particular the hyperbolic variety. After one is bored with the flat world of Euclidean geometry, it is natural to consider the sphere - especially since we live on one! We shall be concerned with making use of the ideas of transformations [6] in the realms of non-standard geometries. Classification of shapes & surfaces is a daunting task, especially as the number of dimensions increases. We intend to review some of the progress that has been made in this direction, following the work of Euler, Gauss and Poincaré. We adopt the route laid out in recent lectures in circulation by Thurston [27], which introduce such topics. Thurston's geometrization conjecture is at the centre of the modern study of 3-manifolds [25, 26]. However, hyperbolic geometry is not restricted to the cerebral ideas of sophisticated mathematics [23].

It has become clear that the classification of surfaces is intimately connected with the study of groups. Moreover, by observing how certain group operations interact with various geometric objects one can get a handle on certain crucial invariants. The main difficulties arise when points of these geometric objects exhibit singularities. And this is the main reason that recent efforts to classify the 3-manifolds has been such a challenge and has produced a variety of new approaches and ideas [18].

B. An example from origami

The following demonstration was presented to me by Dr. Tadashi Tokieda [28]. Imagine that you take a piece of paper and crumple it up. For the purpose of this example, suppose that we flatten the paper by squashing it down on a plane surface.. If we then open up the crumpled page, we will see a network consisting of edges and vertices. The edges may be either valleys or ridges depending on which way the crease was folded.



It would seem that such a careless process would result in no emergent patterns. Consider first a vertex with a degree of 4, (the general case is similar). Let the angles about this point be α, β, γ and δ. It is a triviality to note that these angles must sum to 2π, what is less obvious, is that $\alpha + \gamma = \beta + \delta$. This relation can be confirmed by observing the folding process involved in Fig. 1. In general, the vertex must have the sums of alternating angles being the same.

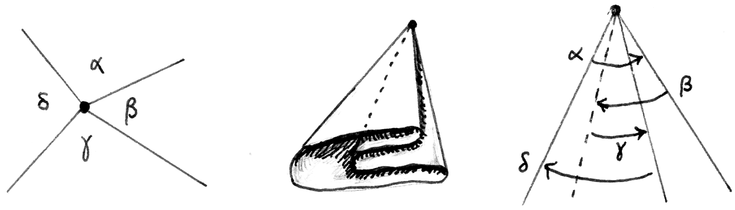


FIG. 1: Explanation for the alternating sum of angles at each vertex.

Paper folding (origami) provides a rich avenue of mathematical study. It differs quite drastically from the confined space of $\mathbb{R}^2$. Indeed, the classical problem of trisecting an angle becomes almost a triviality. It is with this optimistic attitude that we shall proceed in an analysis of various geometric objects.

C. History & Context

Euclid's five postulates read as follows [8]:

(i) *A straight line segment can be drawn between any two points.*

(ii) *Any straight line segment can be indefinitely extended in a straight line.*

(iii) *Given any straight line segment, a circle can be drawn having the segment as a radius and one of the endpoints as its centre.*

(iv) *All right angles are congruent.*

(v) *If two lines are drawn which intersect a third in such a way that the sum of the inner angles on one side is less than two right angles, then the two lines must inevitably intersect each other if extended far enough.*

The last one is known as the *parallel postulate*, and it cannot be proven as a theorem.

For centuries mathematicians tried to find a proof of the parallel postulate in terms of other Euclidean axioms [3]. This was because of the widespread belief that the parallel postulate was artificial compared to the other axioms. Earlier attempts were made at a proof by contradiction by the Jesuit Saccheri (1667-1733) and later Lambert (1728-1777) [5]. Of course, as we now know, their conclusions actually amounted to theorems pertaining to non-Euclidean geometry. Unfortunately such a framework for geometry had not yet been developed.

While Euclid's axioms comfort the intuition, this brings nil to proof. This is because the parallel postulate *is* really independent from the others. A fact that was demonstrated by Bolyai (1802-1860) and Lobachevsky (1793-1856) by doing the unthinkable and constructing a perfectly valid geometry in which the parallel postulate did not hold. A simple but relevant such model was demonstrated by Klein (1849-1925) and we shall discuss it shortly.

In the middle of the nineteenth century, geometric study took a turn which resulted in a new subject we know as *topology*. In this field, one is primarily concerned with the properties of shapes & figures which hold true when such objects are deformed to the extend that metric and projective properties are ruined in the process. We have already mentioned Möbius (1790-1868) and his work on the "one-sided" geometry. It was from the investigation into analytic functions of one complex variable that Bernhard Riemann (1826-1866) realised that these topological results are of critical importance.

II. PRELIMINARIES

A. Toolbox for Topology

Manifolds give the baseline for smooth surfaces with which we shall be working. Before introducing them in full rigour, there are a few basic building blocks which we need:

$$n\text{-ball } B^n := \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \sum x_i^2 \leq 1\} \quad \sim$$ 

$$n\text{-sphere } S^n := \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} : \sum x_i^2 = 1\} \quad \sim$$ 

$$\text{unitinterval } I := [0, 1] \quad \sim$$

Note that we distinguish balls as being *solid* (include their interior), while spheres are *shells* (only boundary). We shall follow previous customs [2, 20], and use the symbol "∂X" to denote the boundary of X. We shall use the symbol "$\simeq$" for "homeomorphic to".

Example The $n-1$-sphere is the boundary of the n-ball, $\partial B^n = S^{n-1}$. For the Möbius strip,

$\partial\Big($  $\Big) = S^1.$

Example The unit interval is homeomorphic to the 1-ball, $I \simeq B^1$.

We now discuss various ways in which we can combine spaces into complicated forms. The first of these is the standard Cartesian product, which has a convenient geometric interpretation.

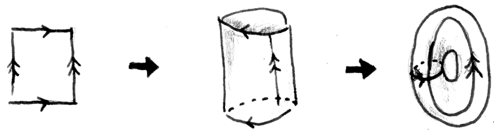
Example The unit cube in n dimensions is simply three copies of the unit interval, $I^n = I \times I \times \ldots \times I \simeq B^n$.

Example The n dimensional torus is similarly a product of 1-spheres (circles), $\mathbb{T}^n = (S^1)^n = S^1 \times S^1 \times \ldots \times S^1$.

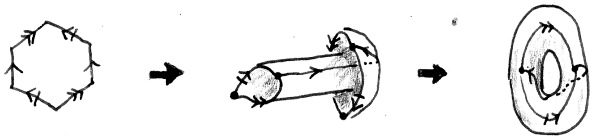
$\mathbb{T}^2 =$  (doughnut) $\simeq$  (trefoil) $\simeq$  (mug).

The next way of joining space is the so called *quotient* of two spaces. We shall make this precise later. Loosely speaking, one seeks to identify certain points in a space through equivalence classes. The following examples will make this clear.

Example Instead of representing the torus as a direct product, it is also possible to arrive at its shape by identifying opposite sides of a rectangle in a parallel fashion as shown:

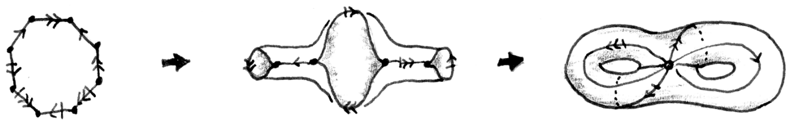


Note that the rectangle tiles the entire plane, and one frequently writes $\mathbb{T}^2 \simeq \mathbb{R}^2/\mathbb{Z}^2$, to indicate that we take real numbers "modulo" one. As a matter of fact, it is possible to produce the torus with a regular hexagon:



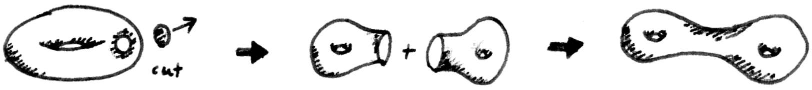
However, this cannot be done for a regular octagon, since it is impossible to tile the plane with such a shape. Nevertheless, identifying sides can produce a perfectly legitimate manifold:

In fact, we shall see later that the hyperbolic plane *can* be tiled with octagons! In some sense this tiling shows that a genus two (and beyond!) object can be given *hyperbolic structure*. Such a tiling may be found on the Escher picture reproduced on the cover page, by joining the noses of the fishes to their fins.



The last tool we shall describe, is the *connected sum* of two manifolds. This is defined for two manifolds M_1 and M_2. The connected sum is denoted $M_1 \# M_2$ and it is obtains be deleting a ball on the interior of each manifold, and "attaching" them together at the resulting boundary spheres. Both M_1 and M_2 need to be of the same dimension for this to work.

Example The connected sum of two tori is $\mathbb{T}^2 \# \mathbb{T}^2$,



An *(oriented) Riemannian* surface of genus g, (denoted Σ_g) is simply the connected sum of g tori:

B. Group Actions

Definition A group G *acts on* a set X if there is a map $\Phi : G \times X \to X$ where

(i) $\Phi(e, x) = x$ for all $x \in X$, where e is the identity in G.

(ii) $\Phi(g, \Phi(h, x)) = \Phi(gh, x)$ for all $g, h \in G$ and all $x \in X$.

This motivates us to denote $\Phi(g, x)$ by gx.

Definition Suppose X is a topological space, and for each $g \in G$, the mapping $\lambda_g : X \to X$ given by the action of G on X, i. e. $\lambda_g(x) = gx$. If λ_g is a homeomorphism, then we say X is a G-space. If we further have that $gx = x$ iff $g = e$, then G is said to *act freely* on X.

Definition If G acts on X, we may form the quotient X/G as a set of equivalence classes on X where $[x_1] = [x_2]$ iff $x_2 = gx_1$ for some $g \in G$. Then X/G is called the *orbit space* of X over G, and the disjoin *orbits* are simple the cosets Gx.

Lemma 2.1. *Suppose that X is a G-space then the map $\pi : X \to X/G$ defined by $\pi(x) = [x]$ is an open map.*

Proof. Let U be an open subset of X. We know that λ_g defined above is a homeomorphism for each $g \in G$, implying $\lambda_g(U)$ is open. Then observe that

$$\pi^{-1}(\pi(U)) = \bigcup_{g \in G} g(U), = \bigcup_{g \in G} \lambda_g(U), \tag{2.1}$$

which is plainly open in X. Since π is continuous, the preimage of every open set is open and the preimage of every closed set is closed. It follows that $\pi(U)$ is open in X/G. □

Proposition 2.2. *If X is a compact Hausdorff G-space with G a finite group, then X/G is also a compact Hausdorff space.*

Proof. Define $\pi : X \to X/G$ by the equivalence class $\pi(x) = [x]$. An open set in X/G are plainly reached as images of open sets in X meaning that π is continuous. The continuous image of a compact space is compact, therefore X/G is compact. It remains to show that X/G is Hausdorff. Take $[x], [y] \in X/G$ such that $[x] \neq [y]$. This means that the preimages $\pi^{-1}([x])$ and $\pi^{-1}([y])$ are disjoint subsets of X. Furthermore, they are finite since G is finite. Using the Hausdorff property of X, we can construct A, B open in X such that $\pi^{-1}([x]) \subset A$,

$\pi^{-1}([y]) \subset B$ and $A \cap B = \emptyset$. Notice now that $\pi(X \setminus A)$ is closed in X/G, since its preimage (in the same was as Eq. (2.1))

$$\pi^{-1}(\pi(X \setminus A)) = \bigcup_{g \in G} \lambda_g(X \setminus A),$$

is a finite union of closed sets and thus closed (in X). In exactly the same way $\pi(X \setminus B)$ is closed (in X/G). Then $V = (X/G) \setminus \pi(X \setminus A)$ and $W = (X/G) \setminus \pi(X \setminus B)$ are both open in X/G. These two sets are furthermore disjoint since (through De Morgan's laws),

$$V \cap W = (X/G) \setminus (\pi(X \setminus A) \cup \pi(X \setminus B)) = (X/G) \setminus \pi(X \setminus (A \cap B)) = (X/G) \setminus \pi(X) = \emptyset.$$

On the other hand, $\pi^{-1}([x]) \subset A$ implies that $[x] \notin \pi(X \setminus A)$. Thus by definition $[x] \in V$. Similarly $[y] \in W$, fulfilling the Hausdorff requirement in X/G. □

It may seem reasonable that we would only require the base space X to be Hausdorff for the quotient to be Hausdorff. The archetypical example is the line with two origins. This is the quotient space of two copies of the real line $\mathbb{R} \times \{a\}$ and $\mathbb{R} \times \{b\}$. The equivalence is $(x, a) \sim (x, b)$ if $x \neq 0$. Therefore, where the image of $(0, a)$ is 0_a and the image of $(0, b)$ is 0_b, we find that all neighbourhoods of 0_a intersect all neighbourhoods of 0_b. The action is plainly not by homeomorphism. So to emphasise the point, consider the following example.

Example Let $\mathbb{Z}$ act on $\mathbb{R}^2 \setminus \{0\}$ via multiplication by the matrix

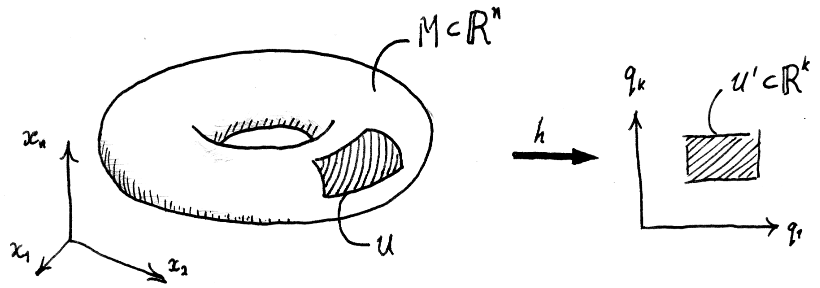
$$A = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix}$$

More precisely, the action $n \cdot \boldsymbol{v} = A^n \boldsymbol{v}$. Then the images of $(0, 1)^T$ and $(1, 0)^T$ in the quotient don't have disjoint neighbourhoods.

C. Manifolds

Definition An n-dimensional (topological) manifold M is a Hausdorff space with countable basis for its topology that is locally Euclidean. In this sense we demand that for each point $x \in M$, there is an open subset $U \ni x$ and a homeomorphism $h : U \to U' = h(U) \subset \mathbb{R}^n$.

We call the pair (h, U) an *(inner) chart* of M or a local coordinate system of M near x. And in keeping with the theme, we call a collection of charts and *atlas*.



Example Perhaps the simplest nontrivial n dimensional manifold is the n-sphere, embedded in $\mathbb{R}^{n+1}$, denoted $S^n = \{\boldsymbol{x} \in \mathbb{R}^{n+1} : \boldsymbol{x}^2 = 1\}$. A possible atlas is

$$A = \{(S^n \setminus \{(0, \ldots, 0, 1)\}, h_1), (S^n \setminus \{0, \ldots, 0, -1\}, h_2)\},$$

where the homeomorphisms are

$$h_1(x_1, \ldots, x_{n+1}) = \left(\frac{2x_1}{1 - x_{n+1}}, \ldots, \frac{2x_n}{1 - x_{n+1}} \right),$$

$$h_2(x_1, \ldots, x_{n+1}) = \left(\frac{2x_1}{1 + x_{n+1}}, \ldots, \frac{2x_n}{1 + x_{n+1}} \right).$$

It is simple to verify that these are the stereographic projections from the north and south poles respectively. Observe that range(h_1) = range$(h_2) = \mathbb{R}^n$, suggesting the opinion that S^n is nothing by $\mathbb{R}^n$ plus a "point at infinity".

Example It is easy to make a Möbius strip: take a ribbon, give it a half-twist and glue the ends together. As on might expect, the Möbius strip is a manifold. More precisely, we are considering the set $[0,1)\times(-1,1)$ with the ends identified in the reverse order, $(0,t)\sim(1,-t)$ for $-1<t<1$. Thus our point set is $M=[0,1)\times(-1,1)$. We can take a two map atlas $A=\{(U_1,h_1),(U_2,h_2)\}$ where our open sets are

$$U_1=\left(\frac{1}{8},\frac{7}{8}\right)\times(-1,1), \qquad h_1(x,y)=(x,y),$$
$$U_2=(\left[0,\frac{1}{2}\right)\cup\left(\frac{3}{4},1\right))\times(-1,1) \qquad h_2(x,y)=\begin{cases}(x,y) & 0\le x<\frac{1}{4},\\ (x-1,-y) & \frac{3}{4}<x<1.\end{cases}$$

Theorem 2.3. *If G is a finite group acting freely on a G-space X, and X a compact n-manifold, then X/G is a compact n-manifold.*

Proof. The conditions of lemma (2.2) are satisfied and we only need to show that each $x\in X/G$ has a neighbourhood homeomorphic to $\mathbb{R}^n$. First we shall label the elements of the group $G=\{g_0,g_1,\dots,g_m\}$ using the fact that it is finite. Let W_i be the neighbourhood containing $g_ix=\lambda_{g_i}(x)$, such that $W_0\cap W_i=\emptyset$ for $i\neq 0$, by the Hausdorff property (the reason for this choice will become clear later). It follows that the set

$$W=\bigcap_{i=0}^{m} g_i^{-1}W_i,$$

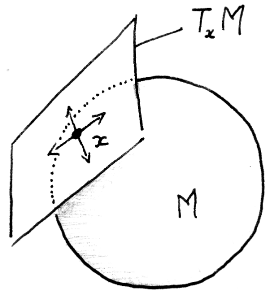
is a neighbourhood of x since the λ_{g_i} are homeomorphisms. Now since X is a manifold, there is a neighbourhood $U\ni x$ such that there is a homeomorphism $h:U\to h(U)\subset\mathbb{R}^n$. In this way, the continuous image of the intersections, $h(W\cap U)$ is open in $\mathbb{R}^n$. Therefore we can find $r>0$ such that $B(h(x),r)\subset h(W\cap U)$. Furthermore, $A:=h^{-1}(B(h(x),r))$ is an open neighbourhood of x in U.

We now claim that the restriction $\pi|_A : A\to\pi(A)$ is a homeomorphism. The function will be bijective if we can confirm that it is injective (surjectivity is trivial). Suppose $\pi|_A(x)=\pi|_A(y)$. Then $x=g_ky$ for some $k\in\{0,1,\dots,m\}$. This restriction enables $x,y\in U$, so that $x\in U_0$ and $y\in g_k^{-1}U_k$. Therefore we conclude that $x=g_ky\in U_0\cap U_k$. This is only possible if $k=0$, by our original construction of U_k. But this means exactly that $x=y$. By lemma 2.1 we know that π is open and continuous (and therefore so is its restriction). Therefore we have $\pi(A)\cong A\cong B(h(x),r)\subset\mathbb{R}^n$. Therefore $\pi(A)$ is the required neighbourhood, proving that X/G is a manifold. □

Definition The tangent bundle over a manifold M is the manifold $TM=\bigcup_{x\in M}T_xM$, where T_xM is the tangent space to M at x, endowed with the local product topology.

A point of TM is a pair $(x,\boldsymbol{v})$ where $x\in M$ and $\boldsymbol{v}\in T_xM$.

Example A vector field $\boldsymbol{v}$ on M is a map $\boldsymbol{v}:M\to TM$ taking $x\mapsto(x,\boldsymbol{v}(x))$.



D. Curvature

Before proceeding, it is important to cement certain basic concepts from differential geometry applied to surfaces [20]. The *Gaussian curvature* at a point on a surface is an indicator of the intrinsic geometry at that point. The following intuitive definition of curvature follows [10].

Definition Let p be a point on the surface Σ and consider a small neighbourhood $U\ni p$. We shall describe this neighbourhood as a ball $U=B(p,r)$. Then consider the set of unit normal vectors for each point in U (the surface at the U will look like a porcupine). Now translate these vectors to the origin, so that their endpoints produce a "patch" on the surface of a unit sphere. If the area of the patch is $A(r)$, then the curvature is

$$\kappa := \lim_{r\to 0} \frac{A(r)}{\text{Area}(B(p,r))}. \tag{2.2}$$

The quantity $A(r)$ is signed by convention. If the boundary on the unit sphere is traced out anti-clockwise, then the area is positive. Clockwise oriented boundaries produce negative areas.

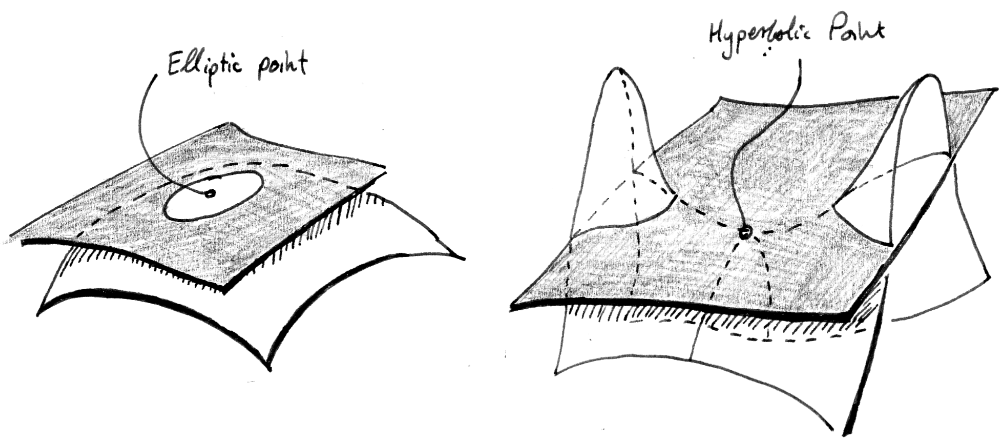
This definition makes it clear that for the flat Euclidean plane, $A(r) = 0$ always, so that $\kappa = 0$. The surface of a sphere of radius R produces a constant ratio in Eq.(2.2) and results in $\kappa = 1/R^2$. Another, more mechanical way of obtaining the curvature is to arrange the coordinate system so that the point under scrutiny is at the origin, and the tangent plane being horizontal. This means that Σ is *locally* the graph of a functions $f(x,y)$. The arrangement means that $f(0,0) = f_x(0,0) = f_y(0,0) = 0$. Then the curvature is the determinant of the Hessian matrix:

$$\kappa = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx}f_{yy} - f_{xy}f_{yx}. \tag{2.3}$$

Example Consider the surface described by the functions $f(x,y) = Ax^2 + 2Bxy + Cy^2$. This will be the leading order contribution for any point arranged such that it is at the origin and flat. Then the curvature at the origin is

$$\kappa = 4(AC - B^2).$$

Let's now set $B = 0$ for simplicity. Then if A, B have the same sign, the curvature of positive. This is the shape of a paraboloid, and we call the point *elliptic*. On the other hand, if A, B have opposite sign, the graph describes a hyperboloid and we call the point *hyperbolic*.



The definitions we have described are extrinsically based, in other words we must hold the surface still while we make the measurements. In fact, the Gaussian curvature is actually an invariant of the surface - it remains the same after deformations (without stretching). This fundamental theorem was proved by Gauss and he called it the "Theorema Egregium" (extraordinary theorem). We shall state it here without proof [11].

Theorem 2.4. (Gauss' Theorema Egregium) *Taking x, y to be coordinates of a surface in $\mathbb{E}^3$ corresponding to $F(x,y)$, centred at the origin with zero first partial derivatives, then the metric admits a power expansion in the normal coordinates (u,v) as*

$$ds^2 = du^2 + dv^2 - \frac{\kappa}{3}(u\,dv - v\,du)^2 + \ldots. \tag{2.4}$$

This shows that the curvature is obtainable through the metric, and is therefore a intrinsic invariant.

From this deep result, we can show how the curvature is calculable at a point p in terms of both the length of the circumference $C_p(r)$, the circumference of a circle of radius r centred at p with area $A_p(r)$.

Theorem 2.5. (Bertrand-Diquet-Puiseux theorem) *For a smooth surface in $\mathbb{E}^3$, with C_p and A_p defined above, then the following equality holds:*

$$\lim_{r\to 0^+} 3\frac{2\pi r - C_p(r)}{\pi r^3} = \kappa(p) = \lim_{r\to 0^+} 12\frac{\pi r^2 - A_p(r)}{\pi r^4}. \tag{2.5}$$

Proof. First we need to recast Eq. (2.4) in normal polar coordinated, $u = r\cos\theta$ and $v = r\sin\theta$. Then one easily finds

$$\mathrm{d}r = \frac{u\mathrm{d}u + v\mathrm{d}v}{\sqrt{u^2+v^2}} \qquad \text{and} \qquad \mathrm{d}\theta\frac{u\mathrm{d}v - v\mathrm{d}u}{u^2+v^2}.$$

Therefore

$$\mathrm{d}r^2 \;=\; \mathrm{d}u^2 + \mathrm{d}v^2 - \frac{(u\mathrm{d}v - v\mathrm{d}u)^2}{u^2+v^2}.$$

Substituting these into (2.4) to find that

$$\mathrm{d}s^2 \;=\; \mathrm{d}r^2 + \mathrm{d}\theta^2\left[r^2 - \frac{\kappa}{3}r^4 + \mathcal{O}(r^5)\right] \tag{2.6}$$

Now we are in a position to calculate lengths and areas on the surface.

$$\begin{aligned}
C_p(r) &= \int_0^{2\pi}\mathrm{d}\theta\sqrt{r^2 - \kappa r^4/3 + \mathcal{O}(r^5)} \\
&= \int_0^{2\pi}\mathrm{d}\theta\; r\left(1 - \frac{\kappa}{6}r^2 + \mathcal{O}(r^3)\right) \\
&= 2\pi r\left(1 - \frac{\kappa}{6}r^2\right) + \mathcal{O}(r^4), \\
A_p(r) &= \int_0^{r}\mathrm{d}r'\int_0^{2\pi}\mathrm{d}\theta\sqrt{(r')^2 - \kappa(r')^4/3 + \mathcal{O}((r')^5)} \\
&= \int_0^{r}\mathrm{d}\;r'\int_0^{2\pi}\mathrm{d}\theta\; r'\left(1 - \frac{\kappa}{6}(r')^2 + \mathcal{O}((r')^3)\right) \\
&= \pi r^2\left(1 - \frac{\kappa}{12}r^2\right) + \mathcal{O}(r^5).
\end{aligned}$$

Taking the relevant limits one finds that the only surviving term is the curvature κ. $\square$

III. MODEL SPACES

A. Projective Plane

The question naturally arises whether we can make any progress in geometry before defining lengths and angles and other magnitude related qualities. One can easily conceive of drastic transformations which destroy these properties. What (if anything) survives? For example, consider a stationary cube on the ground. As we look at it we can move around constantly changing our perspective. We observe that angles change, lengths are stretched or compressed, bisectors not longer bisect and areas dilate. Such transformations are termed *perspective* transformations. Clearly they are obedient enough so that concurrency and collinearity are preserved.

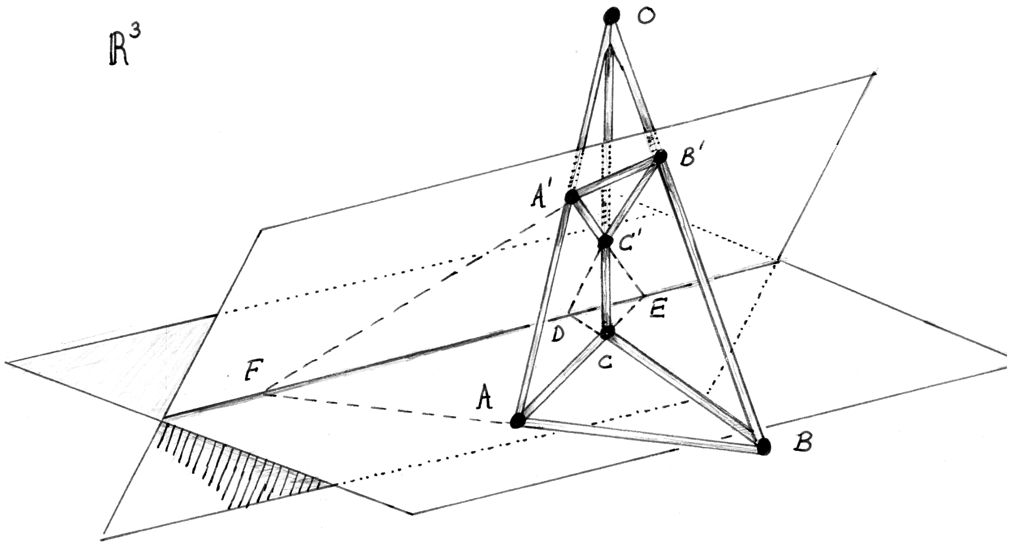
Note that perspectivity is not a transitive operation. We call consecutive perpectivities a *projective transformation.* We now lay out the two most important axioms of projective geometry:

1. *Given any two distinct points in a plane, there is a unique line that passes through them.*
2. *Given any two distinct lines in a plane, there is a unique point that lies on both of them.*

It is then apparent that the words "point" and "line" are interchangeable in any theorem. This is called the principle of *duality*, and for every theorem proved, we get another for free (sometimes both collapse to the same statement).

Theorem 3.1. (Desargues) *In a plane two triangles ABC and $A'B'C'$ are situated to that the straight lines joining corresponding vertices are concurrent in a point O, if and only if the corresponding sides, if extended, will intersect in three collinear points.*

Proof. The argument is visually based. It is remarkable that the theorem is true if the triangles lie in two different (non-parallel) planes. Suppose that AA', BB' and CC' intersect at the point O. This forms a tripod in three dimensions. Now AB and $A'B'$ lie in the same plane and hence intersect at some point F. Similarly AC and $A'C'$ intersect at E, and BC and $B'C'$ intersect at D. Since E, F and D are extensions of the sides of both of the triangles, they lie in the same plane as each of these triangles. Hence they lie on the intersection of these two planes.



Now simply regard the above picture as a perspective drawing of the $\mathbb{R}^3$ configuration and the forwards direction of the theorem is proved. The reverse is immediate by the duality, since three points being collinear is dual to three lines meeting at a point. □

We now introduce a very important quantity called the *cross ratio.* It is defined classically for four points on the line as a "double ratio",

$$(A,B;C,D) \;:=\; \frac{AC/BC}{AD/BD} = \frac{AC}{BC} : \frac{AD}{BD} = \frac{AC\cdot BD}{AD\cdot BC}. \tag{3.1}$$

Note that the line segments are directed (signed). We note that this may also be written in terms of complex numbers, in terms of the function of four complex variables,

$$(z_1,z_2;z_3,z_4) \;:=\; \frac{z_1-z_3}{z_2-z_3} : \frac{z_1-z_4}{z_2-z_4} = \frac{(z_1-z_3)(z_2-z_4)}{(z_2-z_3)(z_1-z_4)}. \tag{3.2}$$

Before we elaborate on the purpose of this quantity, we note certain symmetries. The permutations of the subscripts under which the value does not change are (1 2)(3 4), (1 3)(2 4), (1 4)(2 3) and of course the identity. These constitute the klein four-group $\boldsymbol{V} \trianglelefteq S_4$. This reduces the number of distinct value from 24 to 6. Each of these different values may be expressed in terms of the others:

$$\begin{aligned} &(z_1, z_1; z_3, z_4) = \lambda, && (z_1, z_2; z_4, z_3) = \lambda^{-1}, \\ &(z_1, z_3; z_4, z_2) = (1-\lambda)^{-1}, && (z_1, z_3; z_2, z_4) = 1-\lambda, \\ &(z_1, z_4; z_3, z_2) = \lambda(\lambda-1)^{-1}, && (z_1, z_4; z_2, z_3) = (\lambda-1)\lambda^{-1}. \end{aligned}$$

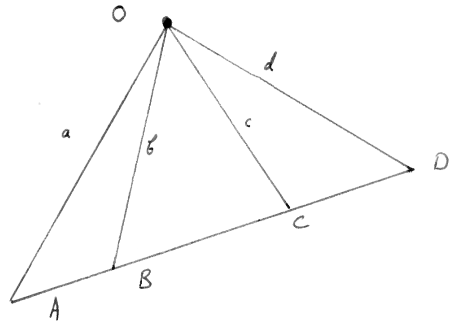
The above formulae in terms of λ form the dihedral group of order six with respect to composition D_6. It is worth mentioning that one of the points is allowed to be the point at infinity in $\hat{\mathbb{Q}}$.

In the spirit of duality, we may also define the cross ratio of four line as follows. Let a, b, c, d be lines all intersecting at the point O. Then define

$$(a, b; c, d) \ := \ \frac{\sin(a\hat{O}c)/\sin(b\hat{O}c)}{\sin(a\hat{O}d)/\sin(b\hat{O}d)}. \tag{3.3}$$

As a motivation for this, let a line l pass through a at A, b at B, c at C and d at D and note that

$$\begin{aligned} |COA| &= h\frac{AC}{2} = \frac{1}{2}OC.OA\sin(a\hat{O}c), \\ |BOC| &= h\frac{BC}{2} = \frac{1}{2}OB.OC\sin(b\hat{O}c), \\ |AOD| &= h\frac{AD}{2} = \frac{1}{2}OA.OD\sin(a\hat{O}d), \\ |BOD| &= h\frac{BD}{2} = \frac{1}{2}OB.OD\sin(b\hat{O}d). \end{aligned}$$



Taking the relevant ratios of the above verifies that

$$(A, B; C, D) \ = \ (a, b; c, d).$$

This fact leads to the theorem below.

Theorem 3.2. *The cross ratio is invariant under projections.*

B. Spherical Geometry

Proposition 3.3. *Consider a sphere $S \subset \mathbb{R}^3$ with radius r. A triangle, with vertices A, B, C on the surface of S, has interior angles α, β and γ given by*

$$\alpha + \beta + \gamma \ = \ \pi + \frac{|ABC|}{r^2}, \tag{3.4}$$

where $|ABC|$ is the area of this triangle on the surface of the sphere.

Proof. The edges of the triangle are the same as the segments of the great circles passing through pairs of points. Construct A', B' and C' as shown above.

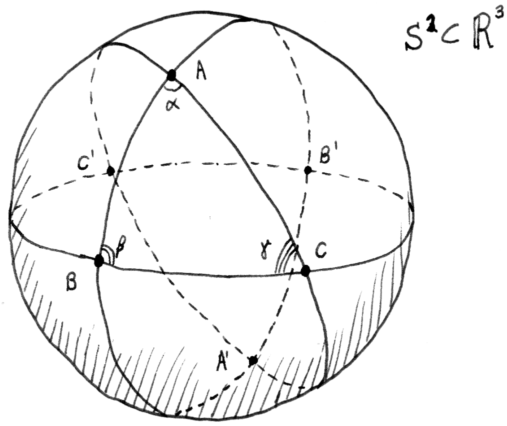
Then it is clear that $|ABC| = |A'B'C'|$. Note that the area of a spherical lune (at angle α) between great circles of the sphere is determined by the angular fraction of the whole sphere, $\alpha(4\pi r^2)/2\pi = 2\alpha r^2$. Then the surface can be tiled by ABC, $A'B'C'$ and three lunes from the exterior angles:

$$4\pi r^2 \ = \ 2|ABC| + 2(\pi-\alpha)r^2 + 2(\pi-\beta)r^2 + 2(\pi-\gamma)r^2 \quad \Leftrightarrow \quad \text{Eq.(3.4)}.$$

□

Corollary 3.4. *For a convex n-polygon on the surface of S, the sum of the angles is given by*

$$\sum_{i=1}^{n} \theta_i \ = \ (n-2)\pi + \frac{|\text{Area}|}{r^2}. \tag{3.5}$$



An n-sphere can be attained via the quadratic form on $\mathbb{R}^{n+1}$ given by

$$Q^{+}(\boldsymbol{x}) \;=\; x_0^2 + x_1^2 + x_2^2 + \ldots + x_n^2, \qquad \text{where } \boldsymbol{x} = (x_0, \ldots, x_n) \in \mathbb{R}^{n+1}. \tag{3.6}$$

The space inherits a Euclidean metric $\mathrm{d}x^2 = \Sigma_{i=0}^{n} \mathrm{d}x_i^2$. Restricting the value of this quadratic form to unity gives the unit sphere $S^n = \{\boldsymbol{x} \in \mathbb{R}^{n+1} : Q^{+}(\boldsymbol{x}) = 1\}$. This surface has a Riemannian metric of constant positive curvature $\kappa = +1$. The isometries of S^n will be the linear transformations of $n+1$ dimensional Euclidean space which preserve Q^{+}, which will be the group of orthogonal transformations $O(n+1)$.

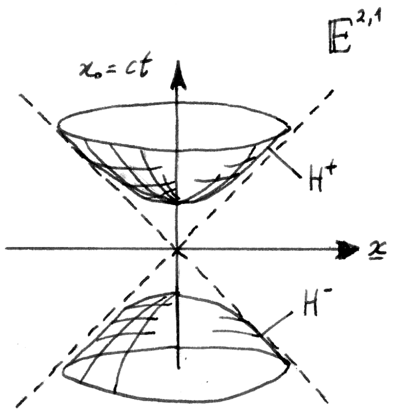
C. Hyperboloid Model

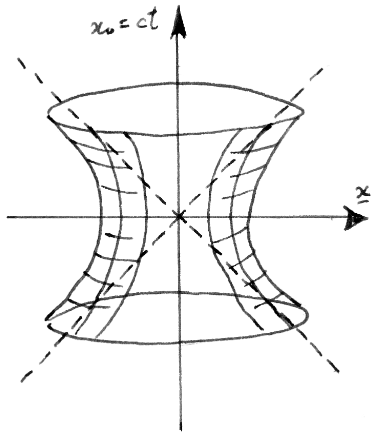
We now arrive at the first topic for which this project takes its title. Hyperbolic geometry can be constructed in several different ways.

As we have seen, a sphere in euclidean space with radius r has constant curvature $1/r^2$. Therefore it is not unreasonable to imagine that hyperbolic space should be a sphere of imaginary radius $r = i$. This would have constant curvature $1/(i)^2 = -1$. Now we use a quadratic form defined by

$$Q^{-}(\boldsymbol{x}) \;=\; -x_0^2 + x_1^2 + x_2^2 + \ldots + x_n^2, \tag{3.7}$$

in $\mathbb{R}^{n+1}$. This is results in the metric $\mathrm{d}x^2 = -\mathrm{d}x_0^2 + \mathrm{d}x_1^2 + \ldots \mathrm{d}x_n^2$. Special relativity makes use of this metric space (with $n = 3$), which is referred to as *Lorentz space* and denoted $\mathbb{E}^{n,1}$. Physically, the x_0 axis represents time with x_i while $i = 1, 2, 3$ represents the three spatial dimensions. Associating the "length" of the vector by $\sqrt{Q^{-}(\boldsymbol{x})}$, as in the Euclidean case. Therefore a vector may have both real and imaginary length, when $Q^{-} = 1$ we have the *one sheeted hyperboloid* and if $Q^{-} = -1$ we get the *two sheeted hyperboloid* H.





Notice that H is separated into two disjoint components, and upper sheet H^{+} ($x_0 > 0$) and a lower sheet H^{-} ($x_0 < 0$). We identify n dimensional hyperbolic space $\mathbb{H}^n$ with the upper sheet H^{+}, and this is called the *hyperboloid model*.

D. Beltrami-Klein Disk & Poincaré Disk

There are two models in which hyperbolic space is contained inside a disk. We shall discuss both here.

The Beltrami-Klein model of hyperbolic geometry is outlined as follows. The "space" consists of only points interior to a circle, (ignore all points outside). Straight lines are nothing but chords of this circle. The tricky matter here is to define a distance which is consistent. We imagine that the plane is subjected to projective transformations, such that the Klein circle and its interior remain inside. We should now require that the distance we use is invariant under projective transformations. Take two points P and Q, for which we wish to measure the distance between. It is important to involve the cross section, and this emerges naturally by extending the usual straight line PQ to intersect the circle at O and S as shown in Fig. 2.

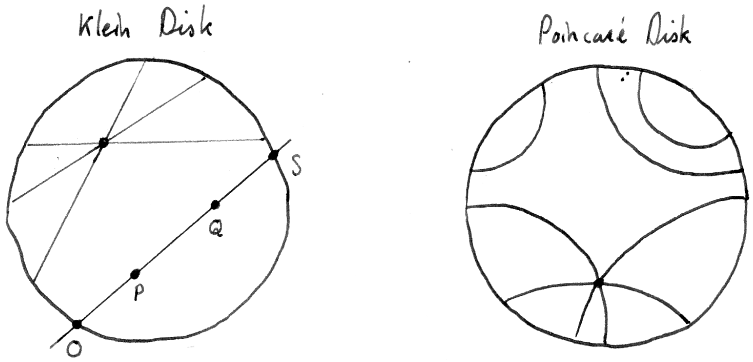


FIG. 2: Left: The Beltrami-Klein Disk model. Various straight lines (geodesics) are shown. Right: The Poincaré disk model. Note that the geodesics appear as circles orthogonal to the boundary from the point of view of $\mathbb{E}^2$.

Perhaps the positive number $(O, S; Q, P)$ will suffice as a distance? Then from standard metric space requirements, for three points P, Q, R on a line the distances should add $PQ + QR = PR$. However this is certainly not true in general for cross ratios. Instead, we have

$$(O, S; Q, P)(O, S; R, Q) \;=\; \frac{QO/QS}{PO/PS} \cdot \frac{RO/RS}{QO/QS} = \frac{RO/RS}{PO/PS} = (O, S; R, P). \tag{3.8}$$

Therefore, to recover additivity from multiplicativity we define the distance by

$$d_K(P, Q) \;:=\; \frac{1}{2} \ln(O, S; Q, P), \tag{3.9}$$

where the factor of $1/2$ ensures that the curvature is -1.

The details of the Poincaré model will be easily attainable shortly. For now we will remark that "straight lines" (geodesics) in this model, consist of circular arcs which are orthogonal to S^1.

E. Upper-Half Plane

Next we discuss the framework for the *upper half-plane model* of hyperbolic non-Euclidean geometry. We shall identify $\mathbb{R}^2$ with $\mathbb{C}$, through the bijection:

$$\mathbb{R}^2 \ni (x, y) \mapsto x + iy =: z \in \mathbb{C}.$$

Definition The *upper-half plane*, denoted $\mathbb{H}^2$, is the set of complex numbers whose imaginary part is positive:

$$\mathbb{H}^2 := \{z \in \mathbb{C} : \Im(z) > 0\}. \tag{3.10}$$

Furthermore, the *circle at infinity* (in the extended complex plane) is the boundary of $\mathbb{H}^2$, and consists of the reals axis together with the point ∞,

$$\partial\mathbb{H}^2 := \{z \in \mathbb{C} : \Im(z) = 0\} \cup \{\infty\}. \tag{3.11}$$

In order to make measurements in hyperbolic space, we require a notion of "length". This is traditionally done through a line integral of arc lengths. Bearing similarity with the Klein disk, we ensure that the length grows arbitrarily as one approaches $\partial\mathbb{H}^2$.

Definition Let $\gamma : [0,1] \to \mathbb{H}^2$ be a partially smooth curve in the upper half plane. Then the *hyperbolic length* of γ is obtained by integrating as follows.

$$\ell_{\mathbb{H}^2}(\gamma) := \int_\gamma \mathrm{d}z \frac{1}{\Im(z)} = \int_0^1 \mathrm{d}t \frac{|\gamma'(t)|}{\Im(\gamma(t))}. \tag{3.12}$$

Observe that if $\Im(\gamma(t))$ is constant for $t \in [0,1]$, then $\ell_{\mathbb{H}^2}$ recovers the usual Euclidean distance, scaled by this constant. On the other hand, we can examine the distance between i and the point yi where $0 < y < 1$. Parametrise the curve $\gamma(t) = (1-t)i + tyi$, yielding

$$\ell_{\mathbb{H}^2}(\gamma) = \int_0^1 \mathrm{d}t \frac{|y-1|}{1+t(y-1)} = \int_0^{y-1} \mathrm{d}u \frac{-1}{1+u} = \ln\frac{1}{y}.$$

This explains why we call $\partial\mathbb{H}^2$ the boundary at infinity.

Definition Let $z, w \in \mathbb{H}^2$. We can now define the *hyperbolic distance* between z and w to be

$$d_{\mathbb{H}^2}(z,w) := \inf_{\gamma \in Q} \ell_{\mathbb{H}^2}(\gamma), \tag{3.13}$$

where Q is the set of all partially smooth paths between z and w, i. e. $\gamma(0) = z$ and $\gamma(1) = w$.

To show that $(\mathbb{H}^2, d_{\mathbb{H}^2})$ is a metric space the triangle inequality is easily verified.

The relationship between the four different models presented here is shown in Fig. 3. The hemisphere model has not been discussed, but it is simply related to the others. Its presence makes it clear how straight lines in the Poincaré model are mapped to circular arcs in the Klein disk. One will notice that the connection between the Poincaré disk and the hyperboloid model is identical to the familiar stereographic projection between the projective plane, $\mathbb{C}P^2$, and the unit sphere S^2. This lends itself to a natural generalisation in $\mathbb{H}^n$, a detailed analysis may be found in [15]. This also enables the metrics defined on the different spaces to be derived from one another.

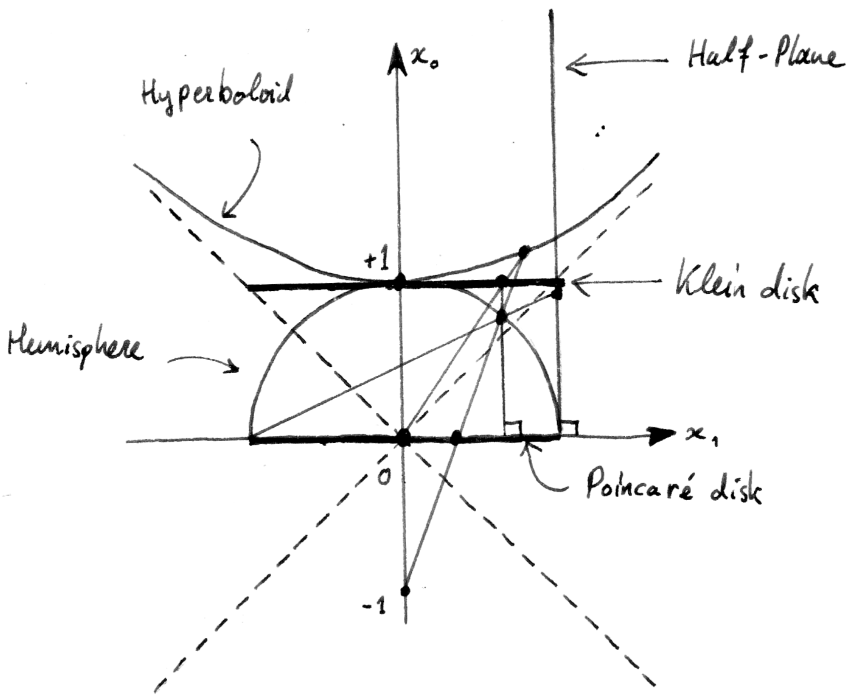


FIG. 3: The rules for transforming between the different models for 2 dimensional hyperbolic space. The figure should be rotated about the x_0 axis to view the correct 3 dimensional embedding.

We now move to study triangles in $\mathbb{H}^2$. Spherical triangles demonstrate a strong dependency between angles and area, and a similar result holds in hyperbolic geometry. The following elementary argument is due to Gauss [25]. An *ideal* triangle is one which has all of its vertices at the boundary of hyperbolic space $\partial\mathbb{H}^2$. Then,

Proposition 3.5. *All ideal triangles are congruent, and have an area of π.*

Proof. We work in the upper half plane model of $\mathbb{H}^2$. Here it is possible to use inversions to transform an ideal triangle for that its vertices are at $(-1,0)$, $(1,0)$ and ∞. This is shown in Fig. 4.

Therefore, written in Cartesian coordinates, the area we seek is of the region $R = \{(x,y) \in \mathbb{R}^2 : -1 \leq x \leq 1, y \geq \sqrt{1-x^2}\}$. Then we know how to calculate the hyperbolic area,

$$\mathrm{Area}(R) \;=\; \int_{-1}^{1} \mathrm{d}x \int_{\sqrt{1-x^2}}^{\infty} \mathrm{d}y \frac{1}{y^2} = \int_{-1}^{1} \mathrm{d}x \frac{1}{\sqrt{1-x^2}} = \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta \;\cos\theta \frac{1}{\cos\theta} = \pi. \tag{3.14}$$

And the theorem is proved. □

Proposition 3.6. *A triangle with vertices A, B, C in the hyperbolic plane $\mathbb{H}^2$ has interior angles α, β and γ given by*

$$\alpha + \beta + \gamma \;=\; \pi - |ABC|, \tag{3.15}$$

where $|ABC|$ is the hyperbolic area of the triangle.

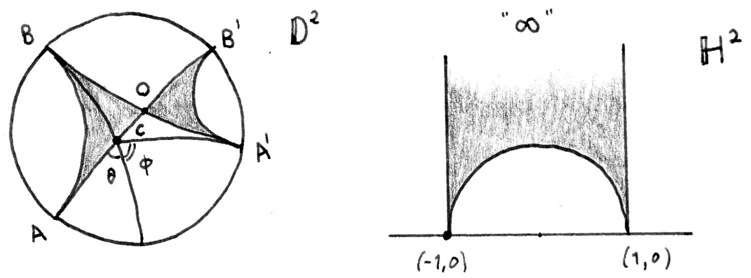
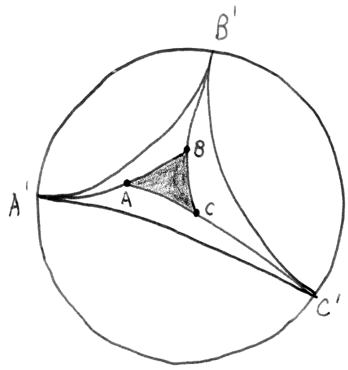


FIG. 4: Right: ideal triangle in the upper-half plane model $\mathbb{H}^2$. Left: Triangles with two vertices on $\partial\mathbb{H}^2$ and their respective angles.

Proof. First we prove the theorem for triangles with 2 vertices on $\partial\mathbb{H}^2$. Since the angle at this boundary is zero, we may label a triangle by it's exterior non-zero angle. Let $A(\theta)$ be the area of such a triangle. The key observation is that A is additive $A(\theta+\phi) = A(\theta) + A(\phi)$. Consider the left diagram of Fig. 4.

The two shaded regions in the Beltrami-Klein model are of equal area (since they are reflections in the point O). Furthermore the region COA' is common and it follows that A is additive. But A is also continuous, and hence by Cauchy's theorem it is linear. Setting $\phi = \pi$ we immediately get $A(\theta) = \theta$.

The proof is completed by noting that an arbitrary hyperbolic triangle is a sum or difference of ideal triangles and triangles with two vertices on $\partial\mathbb{H}^2$.



In the figure, we find

$$|ABC| \;=\; |A'B'C'| - |A'AC'| - |B'BA'| - |C'CB'| = \pi - \alpha - \beta - \gamma,$$

as required. □

Corollary 3.7. *For a planar hyperbolic n-polygon in $\mathbb{H}^2$, the area of the polygon is related to the sum of the interior angles θ_i by*

$$|\text{Area}| \;=\; (n-2)\pi - \sum_{i=1}^{n} \theta_i. \tag{3.16}$$

IV. THE EULER CHARACTERISTIC

A. Regular polyhedra

Theorem 4.1. (Euler) *The number of vertices v, edges e and faces f of a convex polyhedron are related by the formula*

$$v - e + f \;=\; 2. \tag{4.1}$$

Proof. Denote the convex polyhedron by P and project its surface from an interior point to the surface of a sphere centred at that point. This procedure produces a network on the surface of a sphere, with faces of the polyhedron identified with convex spherical polygons, P_j where $j \in \{1, 2, \ldots, f\}$.

Suppose that the spherical polygon P_j has n_j sides. Then by Eq.(3.5), we may relate the area of P_j to the sum of the interior angles θ_{ij} where $i \in \{1, 2, \ldots, n_j\}$. Without loss of generality, the radius of the sphere may be set to unity. The total number of vertices may be counted in terms of the number of edges, since each edge contributes to two vertices, hence

$$2\pi e \;=\; \sum_{j=1}^{f} n_j. \tag{4.2}$$

In terms of angles, each vertex contributes 2π to the total sum over all angles:

$$2\pi v = \sum_{j=1}^{f}\sum_{i=1}^{n_j}\theta_{ij} \overset{\text{Eq.(3.5)}}{=} \sum_{j=1}^{f}(n_j-2)\pi + \sum_{i=1}^{f}|P_j| = 2\pi e - 2\pi f + 4\pi. \tag{4.3}$$

This reduces to to the Euler formula, Eq.(4.1). □

In efforts to generalise this statement, we shall consider expressions similar to $v - e + f$ in a more generic setting. Of course the value will not always be equal to 2 (consider for example a triangulation of the torus), instead we shall call its values χ the *Euler characteristic* which is intimately connected with the topological structure. As an example we shall consider the five colour theorem [5]:

Theorem 4.2. (The five colour theorem) *Every subdivision of the sphere can be* properly coloured *by using at most five different colours. By properly coloured we mean that no two regions having a whole segment of their boundaries in common receive the same colour.*

Proof. It is sufficient to consider subdivision whose regions are bounded by simple closed polygons of circular arcs. If we replace every vertex at which more than three arcs meet by a small circle, and join the interior of each such circle to one of the regions which meet at the vertex, a new map is obtained with the same number of regions but only vertices of degree three. Such a division will be called a *regular map* on the sphere. If this can be properly coloured with five colours, then we simply need to shrink the circles down to a point to recover the original map.

The first step is to prove that each regular map will contain a least one region which is bounded by fewer than six sides. By construction of a regular map, each arc has two ends, and each end is connected to exactly three arcs. Hence we know $2e = 3v$. Denote by f_n the number of regions with n sides, and therefore n vertices. Summing over all faces we obtain:

$$2e = 3v \;=\; 2f_2 + 3f_3 + 4f_4 + \ldots . \tag{4.4}$$

Now combining this with Eq.(4.1), we conclude that $3f - e = 6$. And therefore re-expressing (4.4) yields

$$12 \;=\; 6(f_2 + f_3 + f_4 + \ldots) - (2f_2 + 3f_3 + 4f_4 + \ldots) = \sum_{n=1}^{\infty}(6-n)f_n\,.$$

Clearly one of the terms in the sum must be positive, and therefore at lease one of f_2, f_3, f_4 or f_5 is positive.

Now let M be a regular map one the sphere with k regions. We complete the proof in two parts.

Case 1. The map M contains a region R with five sides. Call these five neighbouring regions Q_1, Q_2, Q_3, Q_4 and Q_5. The Jordan curve theorem applied to the sphere guarantees us that we can find a pair among these which do not touch. Without loss of generality, take these to be Q_1 and Q_2. Now remove the sides of R adjoining Q_1 and Q_2. This produces a new map with $k-2$ regions. If this can be properly coloured with five colours, then so can M, since when the boundaries are restored, R will be touching regions of a most four different colours (as Q_1 and Q_2 are the same colour). Hence we simply paint R the fifth colour.

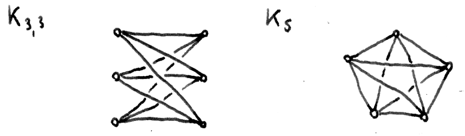
Case 2. The map M contains a region R with 2,3 or 4 sides. If it has four sides, then again we can find a pair of adjacent regions which do not touch. Remove the boundary between R and one of these regions. Otherwise if R has 2 or 3 sides then it doesn't matter which boundary we remove. Either way we get a new map with $k-1$ regions. If we can find a proper colouring for this map, then we can find a colouring for M, since R is neighbour to at most four different colours.

In this fashion, if M is a regular map with k regions, we may always reduce this to a regular map of $k-1$ or $k-2$ regions. This process may be sequentially applied until we arrive at a map with five or fewer regions, which can of course be properly coloured with five colours. Returning step by step to M, we find that indeed M itself can be properly coloured with five colours. □

Another example from the field of graph theory will hint at the topological relevance of the Euler characteristic. A *planar graph* is a graph whose vertices can be represented by points in the Euclidean plane, and whose edges are simple (non-intersecting) curves in the same plane.

Theorem 4.3. (Kuratowksi's theorem) *A finite graph is planar if and only if it does not contain a subgraph homeomorphic to K_5 or $K_{3,3}$.*

Here K_5 denotes the complete graph with five vertices, and $K_{3,3}$ denotes the bipartite graph of six vertices.



We shall not prove this theorem in full, however Euler's formula can determine one direction of the statement:

Proposition 4.4. *The two graphs K_5 and $K_{3,3}$ are nonplanar.*

Proof. Observe that if a graph is planar, then it may be embedded on the sphere without edge intersection. Thus we will obtain a subdivision of the sphere, and Euler's formula is applicable!

First we do away with K_5. Notice that there are $v = 5$ vertices and $e = \binom{5}{2} = 10$ edges. From Eq.(4.1), there must be $f = 7$ faces. Each face is bounded by at least three edges, while each edge bounds exactly two faces. Therefore $3f \leq 2e$, meaning the number of faces is at most,

$$f \;\leq\; \frac{2e}{3} = 6 + \frac{2}{3} < 7 = f,$$

giving a contradiction.

For the bipartite graph $K_{3,3}$, there are $v = 6$ vertices and $e = 3^2 = 9$ edges, and hence by (4.1) there are $f = 5$ faces. Note that bipartite graphs cannot have odd length cycles, and therefore each face is contained in at least 4 edges, i. e. $4f \leq 2e$. Hence we find

$$f \;\leq\; \frac{2e}{4} = 4 + \frac{1}{2} < 5 = f,$$

giving the second contradiction. □

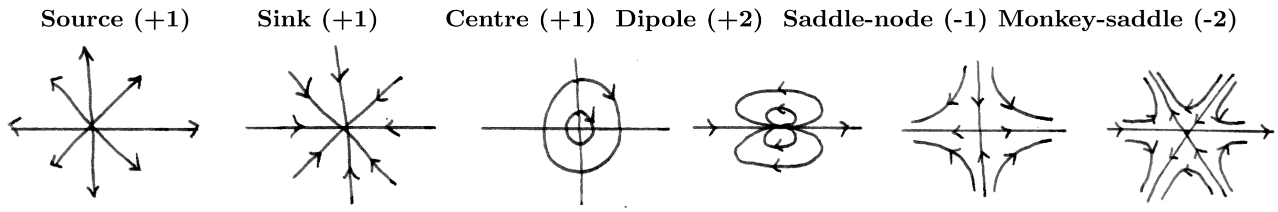
B. Relation to the Theory of Vector Fields

This section elucidates the relationship between singularities and the Euler characteristic [13]. Note that in what follows, we make no statement about the genus of the surface. This connection shall be discussed in the next section.

Definition Let Σ be a closed surface with continuous first derivatives at every point. This enables the existence of a tangent plane at every point. The normal of this tangent plane will vary continuously over the surface. Consider a field of normalised tangent vectors defined and continuous at all but finitely many points on the surface. These points of discontinuity are the so called *singular* points. We call such a construction a *regular vector field* on Σ.

Enclose an isolated singularity in a small region such that this region contains no other singularities of the vector field. We shall let the boundary of this region be a simply closed curve. The surface has continuous first derivatives and it is therefore possible to define a continuous tangent vector field on this boundary. The index is then the change in angle between the tangent vector field and the original, as one traverses the simple closed curve divided by 2π. Since the curve is closed this will be an integers.

Example Consider the following singular points described by vector fields in the plane. The index of the singular point is the value in brackets.

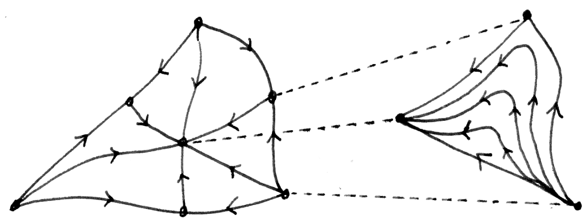


Proposition 4.5. (Poincaré index theorem) *If Σ is a smooth surface and $\boldsymbol{v}$ is a regular vector field on Σ with isolated zeros, the Euler characteristic of Σ is*

$$\chi(\Sigma) \;=\; \sum_{z\in\text{zeros}} \text{index}_{\boldsymbol{v}}(z). \tag{4.5}$$

Consequently $\chi(\Sigma)$ is independent of the triangulation and the sum of the indicies at the zeros is independent of $\boldsymbol{v}$.

Proof. We shall first find the value of $\chi(\Sigma)$ for a specific vector field. Then we shall complete the proof by showing that this is invariant the sum on the right hand side of (4.5) does not depend on the choice of $\boldsymbol{v}$. Construct $\boldsymbol{v}$ as follows. Divide Σ into a network in such a manner that all regions are bounded by three edges (this is a *triangulation* of the surface). In each region, introduce four singular points, as shown in the diagram below.



There is a *sink* at the centre and three *saddle-nodes* on the edges. The three vertices are *sources*. Breaking up the triangle into six sub-triangles to ensure that the vector field is continuous on its interior. Notice that the boundaries of the larger triangle allow for the vector field $\boldsymbol{v}$ to be defined continuously over Σ except at the isolated singularities by simply tiling.

Recalling that the index of sinks and sources is $+1$ and the index of saddle-nodes is -1. Then each face contains a sink and therefore contributes $+1$ to the sum of the indexes. Each edge contains a saddle-node and contributes -1, while each vertex is a source and hence contributes $+1$. Hence we have found,

$$\sum_{z\in\text{zeros}} \text{index}_{\boldsymbol{v}}(z) \;=\; v+f-e = \chi(\Sigma).$$

It remains to show that this sum is independent of $\boldsymbol{v}$. Consider two vector fields $\boldsymbol{v}$ and $\boldsymbol{w}$ on Σ. Proceed by subdividing Σ by a network with the condition that the singularities of both vector fields to not fall on the edges or vertices of this division. Furthermore, by making the regions small enough, we are able to ensure that each region has at most one singularity of $\boldsymbol{v}$ or $\boldsymbol{w}$ in its interior. Now that the regions separate the singular points, the difference in indexes at a point z for two vector fields $\boldsymbol{v}$ and $\boldsymbol{w}$ can be arrived at by the change in relative directions of the two vector fields in traversing the boundary of a region in the positive sense (divided by 2π). However, if we sum over all regions, every edge in the network will be traversed twice: once for each of the two regions bounded by the edge. Therefore, the total sum is reduced to zero by cancellations each term. Thus:

$$\sum_{z\in\text{zeros}} \text{index}_{\boldsymbol{v}}(z) - \text{index}_{\boldsymbol{w}}(z) \;=\; 0,$$

where the set of zeros is over $\boldsymbol{v}$ and $\boldsymbol{w}$. This completes the proof. ∎

C. Generic surfaces

Two-dimensional surfaces display many important but still simple facts. In our earlier discussion of topological transformations, it was mentioned that objects should not be torn or punctured in their manipulation. This amounts to preserving a property called the *genus* of a surface. Roughly speaking, this is equal to the number of "holes" a surface has. More precisely,

Definition The genus g of a surface Σ is the largest number of non-intersecting closed cuts that can be made on Σ without separating it into two parts.

From a topological point of view, a closed surface is completely characterised by its genus $g = 0, 1, 2, \ldots$.

Proposition 4.6. *Suppose that a closed surface Σ of genus g is divided into a number of regions by marking a number of vertices on Σ and joining them by curved arcs. Then*

$$2(1 - g) \; = \; v - e + f, \tag{4.6}$$

where v is the number of vertices, e is the number of arcs and f is the number of regions.

We have as a matter of fact already proved this for a sphere (cf. Eq.(4.1)), as this is an object with $g = 0$.

Proof. To prove the formula (4.6), we may treat Σ as a sphere with g handles. It is possible to continously deform a g genus object into this form and this operation will keep $v - e + f$ unchanged. Note that it is implicit in the statement of the theorem, that a region obtained by such a subdivision may not entirely include a handle. Therefore it is possible to arrange that the deformation of the sufrace includes closed curves C_i and K_i (contained in Σ) where the i-th handle joins the sphere. We now cut the surface along each C_i, thereby reducing the surface to a zero genus object.

Each handle now has a free edge bounded by C'_i which was previously joined to C_i. Because of this, C_i has the same number of vertices and arcs as C'_i. Clearly the boundary of any region has the same number of vertices as edges, and therefore the quantity $v - e + f$ remains unchanged through this cutting process.

We now continuously deform the resultant object (its handles severed), into a sphere. What remains is a sphere with $2g$ regions omitted, precisely those bounded by C_i and C'_i. We know that Eq.(4.1) holds for a sphere, so that

$$v - e + (f + 2g) \; = \; 2 \qquad \Leftrightarrow \qquad \text{Eq.(4.6)}.$$

□

Example Consider the doughnut with two holes ($g = 2$) shown in Fig. 5. This can be subdivided & continuously deformed to produce the "concrete brick".

Evidently in this case $f = 2 \times 7 + 3 \times 4 = 26$, $e = 4 \times 2 \times 5 + 2 \times 6 + 4 = 56$ and $v = 2 \times (6 + 8) = 28$. Therefore

$$v - e + f \; = \; 28 - 56 + 26 = -2 = 2(1 - g).$$

We are now in a position address a topic raised in [22], namely that you "can't comb a hairy ball". Indeed were this possible, there would exist a vector field on the surface without singularities which we now know to be impossible by theorem 4.5. Moreover, there is only one surface for which there are no singularities. For $\chi(\Sigma) = 0$ to hold we must have $g = 1$. Therefore the only possible surface with this property is the torus. It is simple to produce a construction to prove the existance.

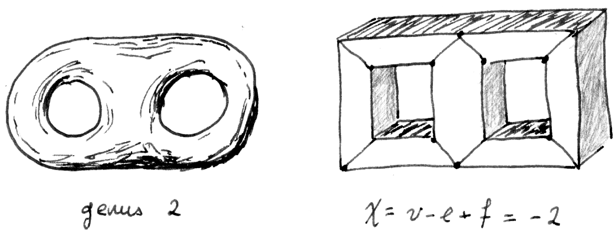


FIG. 5: A "figure-eight" surface transformed into a rigid polyhedron on the right.

D. Higher dimensions

The Euler characteristic we have thus far described applies to 2-dimensional surfaces embedded in $\mathbb{R}^3$. Formula (4.1) suggests a natural generalisation for the Euler characteristic of an n-dimensional manifold M which may be *triangulated*. To clarify the meaning of this statement, we need to generalise the triangle to arbitrary dimension.

The "3-dimensional triangle" is more commonly known as a *tetrahedron*, i. e. a triangular based pyramid. Call the n-dimensional tetrahedron $\mathcal{T}_n$. From the structural point of view, these are the simplest figures requiring the given dimension n. Notice that the tetrahedron has 4 vectors, 6 segments and 4 triangular faces. It is then clear how to proceed: each $\mathcal{T}_n$ contains $n+1$ vertices. Moreover, each each subset of $i+1$ vertices of a $\mathcal{T}_n$, $(i \in \{0, 1, \ldots, n\})$ determines another copy of $\mathcal{T}_i$. Therefore a combinatorial approach is needed, and we see that there are

$$\binom{n+1}{i+1} = \frac{(n+1)!}{(n-i)!(i+1)!}, \tag{4.7}$$

ways of choosing the vertices to form a $\mathcal{T}_i$ in a $\mathcal{T}_n$. For example, $\mathcal{T}_4$ is a 4-dimensional object with 5 vertices, 10 line segments, 10 triangular faces and 5 solid tetrahedra.

Consider a triangulation of M involving β_0 vertices, β_1 edges, β_2 faces,…, β_n n-dimensional tetrahedra. The Euler characteristic of M is then

$$\chi(M) = \sum_{k=0}^{n} (-1)^k \beta_k. \tag{4.8}$$

Example The n-dimensional tetrahedron $\mathcal{T}_n$ can be used to triangulate the $(n-1)$-sphere, S^{n-1}. The circle and the sphere are simple:

$\chi(S^1) = \chi($  $) = 3 - 3 = 0,$ and $\chi(S^2) = \chi($  $) = 4 - 6 + 4 = 2.$

We are in fact equipped to determine $\chi(S^n)$ using what we have shown. In this case, the number of k-tetrahedra $\mathcal{T}_k$ is β_k given by choosing $k+1$ vertices from $n+1$ via Eq. (4.7). Therefore

$$\chi(S^n) = \sum_{k=0}^{n} (-1)^k \binom{n+1}{k+1} = \begin{cases} 0 & \text{if } n \text{ odd}, \\ 2 & \text{if } n \text{ even}, \end{cases}$$

using the binomial theorem to make the parity distinction.

V. GEOMETRIC GROUP STRUCTURES

A. Flows on the billiard table

What follows is a motivational example involving translation surfaces and their properties. A standard billiards table is a flat rectangular region in $\mathbb{R}^2$. Cue balls are directed along the surface travelling at a unit speed until colliding elastically with either the wall or another ball. We shall simply examine the trajectory of an isolated ball moving in the planar domain.

From now on, the pool table will be the rectangle $U := (0, a) \times (0, b)$. Denote by θ_i the incident angle of a trajectory just before it strikes the boundary ∂U and θ_f the angle it rebounds at. If the reflection occurs on a horizontal side of U, then

$$\theta_f = -\theta_i \ =: \ S_h(\theta_i).$$

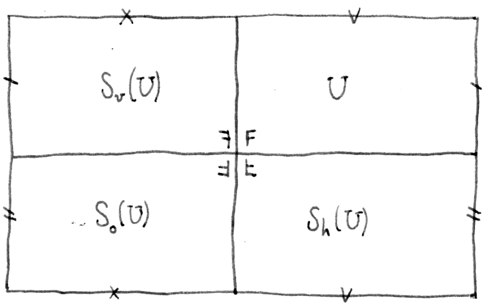
And for reflections on a vertical component of the boundary,

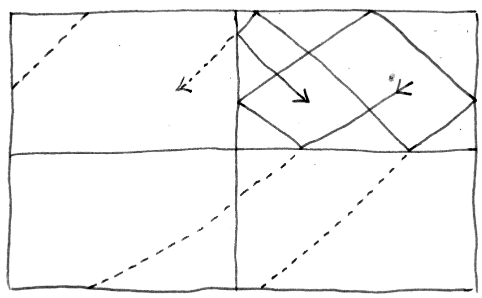
$$\theta_f = \pi - \theta_i =: S_v(\theta_i).$$

It is plain that S_h and S_v are commuting involutions (modulo 2π) and hence generate the Klein four-group $\boldsymbol{V} = \mathbb{Z}_2 \times \mathbb{Z}_2$.These functions are naturally associated with linear symmetries in $\mathbb{R}^2$, defined by

$$S_h(x, y) = (x, -y) \qquad \text{and} \qquad S_v(x, y) = (-x, y),$$

which are reflections in the x-axis and the y-axis respectively. The composition of these two operations is $S_o := S_h \circ S_v = \circ S_v \circ S_h$. The diagram shows the surface constructed through symmetric images of U, namely $S_h(U)$, $S_v(U)$ and $S_o(U)$. Sides are identified as shown in the diagram: for each $g \in \boldsymbol{V} = \{e, S_h, S_v, S_o\}$, the upper/lower side of $g(U)$ is identified with the lower/upper side of $S_h \circ g(U)$ respectively. Similarly the left/right side of $g(U)$ is identified with the right/left side of $S_v \circ g(U)$.





This is nothing but the 2-dimensional flat torus $\mathbb{T}^2 = S^1 \times S^1$. We have therefore found that trajectories of a billiard ball on a pool table are exactly the *linear flow* on a torus! The trajectory is simply parametrised by $t \in \mathbb{R}$,

$$F^t_{u,v}(x, y) := (x + ut, y + vt). \tag{5.1}$$

These may be further classified by the following proposition (the proof is omitted):

Proposition 5.1. *If the ratio $u/v \in \mathbb{Q}$, then every trajectory described by the flow $F^t_{u,v}$ is periodic. Alternatively, if the ratio u/v is irrational then any given trajectory is dense and equidistributed in $\mathbb{T}^2$.*

B. Crystallographic Group

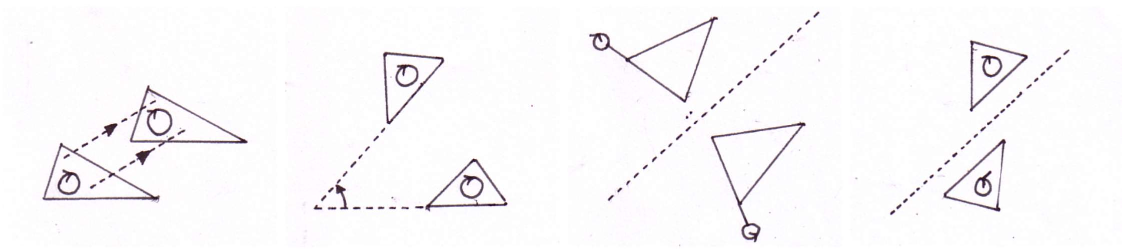
Before proceeding to more complicated environments, we shall discuss the symmetry groups in $\mathbb{R}^2$, known as the *wallpaper groups*. It has been known for centuries [16] that there are only 17 distinct groups of this variety, no two of which are isomorphic.

Recall that an *isometry* is a distance preserving map between metric spaces. In particular a *Euclidean plane isometry* is a function $\tau : \mathbb{R}^2 \to \mathbb{R}^2$ such that for any two points x, y in the plane,

$$d(x, y) \ = \ d(\tau(x), \tau(y)),$$

where d denotes the standard Euclidean distance. When considering isometries from the plane to itself these may be entirely classified (the orientation is shown by arrowheads on circles):

Translations - Rotations - Reflections - Glide Reflections



Proposition 5.2. *Every orientation preserving (proper) isometry of $\mathbb{R}^2$ is either a translation or a rotation. Contrariwise, every orientation reversing (improper) isometry of $\mathbb{R}^2$ is a reflection or a glide reflection.*

Moreover, every isometry of the plane may be reduced to a composition of reflections. For example, a translation by a distance $2d$ may be arrived at by first reflecting in a line perpendicular to the translation direction moving a distance d, and then reflecting again in the same fashion to complete the translation. In a similar way, a rotation by 2θ may be obtained by reflecting in a line at an angle $\theta/2$ and another in a line at an angle $3\theta/2$. More is true:

Proposition 5.3. *Any Euclidean plane isometry can be reduced to a composition of reflections in three different mirrors.*

Proof. The isometry τ is entirely determined by its effect on three non-collinear points. Suppose x, y and z are such points. If we are not dealing with the identity, we may assume without loss of generality that one of the points is not fixed: $x \neq \tau(x)$. Hence we set our first reflection to be in the perpendicular bisector of the line segment joining x and $\tau(x)$. Call this reflection f. All further reflections will keep $\tau(x)$ fixed. If $\tau(y) = f(y)$ and $\tau(z) = f(z)$ then we are done. Otherwise we may say $\tau(y) \neq f(y)$. Take h to be a reflection in the angle bisector of the angle $f(y)$ and $\tau(y)$ makes at $\tau(x)$. In the projective plane, lines are regarded as generalised circles. Then $\tau(y) = h \circ f(y)$, so that it remains to map $h \circ f(z)$ to $\tau(z)$. If it is not already so, we simply need to reflect in the line through $\tau(x)$ and $\tau(y)$. □

The set of all isometries in n dimensional Euclidean space forms a group, named $E(n)$, with respect to composition. We are at the moment studying those with $n = 2$. Note that the composition of two proper, or of two improper, isometries is a proper isometry. And unsurprisingly, the composition of a proper or improper isometry (in either order), is an improper isometry.

In order to move to the wallpaper patterns, we need to formally introduce the *discrete* isometries. This shall be unpacked in Sec. VI, but for now it will suffice to motivate this as follows.

Definition A subgroup G of $E(n)$ is called discrete iff for each point $x \in \mathbb{R}^n$ there exists a ball $B(x, r) \subset \mathbb{R}^n$ such for all $g \in G$, we have either $g(x) = x$ or $g(x) \notin B(x, r)$.

Definition A plane symmetry group is a discrete subgroup of $E(2)$ which contains two independent translations. For ease, we shall now work in the complex plane $\mathbb{C} \cong \mathbb{R}^2$. We shall define the *lattice group* in a natural way:

$$G(\omega_1, \omega_2) \;=\; \{z \mapsto z + \omega_1 m + \omega_2 n : m, n \in \mathbb{Z}\} = \langle z \mapsto z + \omega_1, z \mapsto z + \omega_2 \rangle, \tag{5.2}$$

with associated lattice $\Lambda(\omega_1, \omega_2) = \{\omega_1 m + \omega_2 n : m, n \in \mathbb{Z}\}$, where $\omega_1, \omega_2 \in \mathbb{C}$. There is a further condition that $\omega_1/\omega_2 \notin \mathbb{R}$, i. e.the complex numbers are not parallel.

The motivation for this is that if t and s represent translations by ω_1 and ω_2 respectively, then all elements in the group are of the form $t^i s^j$. Working backwards, any discrete subgroup of $E(2)$ consisting of two independent translations must have two independent "smallest" translations.

Next we wish to include rotations in our subgroup of $E(2)$. This is characterised by the following theorem.

Theorem 5.4. *(The Crystallographic Restriction) Let G be a plane symmetry group. Then each rotation of G has order* $1, 2, 3, 4$ *or* 6.

Proof. Let t be the smallest translation in G and let r be the smallest (positive) rotation. Let r have centre X and rotate anticlockwise by $2\pi/m$, $m \in \mathbb{N}$. Take another point $Y = t(X)$ and define

$$Z \ = \ rtr^{-1}t^{-1}(Y) = rt(X) = r(Y).$$

Then $rtr^{-1}r^{-1}$ is a translation sending Y to Z. Since t is the smallest translation, YZ is at least equal to $XY = XZ$, meaning that $2\pi/m \geq \pi/3$. In other words, $m \leq 6$.

It remains to eliminate $m = 5$. To see this, note that if the centre of rotation lies on a lattice point, through t^2 we construct a rotation by π (called the diad rotation). Therefore $2\pi/m$ must divide π, i. e. m is even. Hence for both $m = 3$ and $m = 5$ (being the smallest such rotations), the centre of rotation cannot lie on a lattice point. For such a lattice, we will thusly obtain a tiling on the plane in m-gons. This rules $m = 5$ out of the picture. □

We arrive at the conclusion that there are only five basic lattice types which constitute a plane symmetry group. From this point, it is a simple matter to identify which symmetries are associated with which lattices. For example, the lattice comprising of equilateral triangles enjoys a 6-fold symmetry.

Definition An n-dimensional crystallographic point group K is a group of isometries of $\mathbb{R}^n$ which fixes a point x and leaves some n-dimensional lattice unchanged.

We now return to the $n = 2$ dimensional case. Observe that a crystallographic point group cannot contain translation or glide reflections. Infact, the elements of the group are either reflections in a line through x or rotations centred at x. Thus, K can be nothing more than the cyclic groups C_n or the dihedral groups D_n, where theorem 5.4 implies $n = 1, 2, 3, 4$ or 6. This definition is motivated by the next result.

Proposition 5.5. *Each plane group G gives rise to a crystallographic point group as a homomorphic image.*

We require the following lemma to ease the proof.

Lemma 5.6. *$E(n)$ is the semidirect product of $O(n)$ extended by T, the subgroup of translations in $\mathbb{R}^2$.*

Proof. Any element s of $E(n)$ may be written $s(x) = Ax + v$ where $A \in O(n)$ and $v \in \mathbb{R}^n$. A translation $t \in T$ may be represented by $t(x) = x + w$, for some $w \in \mathbb{R}^2$. Then

$$t^s(x) = A^{-1}(Ax + v + w - v) = x + A^{-1}w,$$

which is plainly a translation. Thus if t is a translation and u is any isometry, $u^{-1}tu$ is also a translation. □

Returning to the full proof of 5.5, set $n = 2$.

Proof. Now we can combine this with the second isomorphism theorem to conclude that $H := T \cap G$ is a normal subgroup of G, since plane groups are in turn subgroups of $E(2)$. It remains to find a map ϕ such that $\ker \phi = H$.

We now fix a point x in the plane. Consider $g \in G$ and write t as the translation carrying x to $g(x)$. Thus $s := t^{-1}g$ leaves x invariant, and so must be a rotation about x or a reflection with an axis passing through x. Thus $g = ts$ where t is a translation and s leaves x unchanged. It is elementary to show that the set of all s which leaves x invariant forms a group, we denote K. Similarly that the map $\phi : G \to K$, defined by $\phi(g) = \phi(ts) = s$, is a group homomorphism. Furthermore, this construction gives $\ker \phi = H$. Therefore the first isomorphism theorem gives $K \cong G/H$.

The lattice formed which is left unchanged is $\Lambda = \{t(x) : t \in T\}$. Take some $x_0 \in \Lambda$, so that $x_0 = t_0(x)$ for some $t_0 \in T$. Thus for $ts = g \in G$ gives

$$s(x_0) \ = \ st_0(x) = t^{-1}gt_0(x) = t^{-1}gt_0g^{-1}g(x) = t^{-1}(gt_0g^{-1})ts(x) \in \Lambda,$$

since $s(x) = x$ and the conjugate of t_0 is a translation. □

Pursuing this analysis will produce the 17 unique plane groups, and images showing the symmetries may be found in [1].

C. Möbius Transformations

Definition A Möbius transformation of the Riemann sphere $\hat{\mathbb{C}}$ is a map of the form

$$f(z) \ := \ \frac{az+b}{cz+d}, \tag{5.3}$$

where a, b, c and d are complex numbers with the proviso that $ad-bc \neq 0$. We let $f(\infty) = a/c$ and $f(-d/r) = \infty$.

The Möbius transformations form a collection of bijections from the extended complex plane to itself. Furthermore, these transformations are homeomorphisms and form a group under compositions as we shall see [14]. In the following example, we illustrate the analogy between the Möbius transformations and $GL(2,\mathbb{C})$, the set of 2×2 invertible matrices with elements in $\mathbb{C}$. The question was posed to me by a fellow student.

Example The function $f(z) = 1/z$ is special for a number of reasons, one of which is the fact that if you apply it to a number twice, you get back the original number, i.e. $f(f(z)) = 1/(1/z) = z$. Do other functions behave in similar ways? Consider $f(z) = 1/(1-z)$. After applying and reapplying this function we find the sequence

$$x \to \frac{1}{1-z} \to \frac{1-z}{z} \to z, \tag{5.4}$$

which will then repeat. For brevity we shall write this nested composition of functions as $f^n(z)$ when the function is applied n times. The set of consecutive compositions of a function is called the *orbit*. For the previous function, we found that $f(f(f(z))) = f^3(z) = z$.

Now the question; For any natural number n, does there exist a rational function f, from the Riemann sphere[1] to itself, such that the orbit is periodic with length n? By rational function, we simply mean a function which is the ratio of polynomials. We've already found such examples where $n = 2$ and $n = 3$. The function $f(z) = z$ is the trivial case providing $n = 1$.

A little searching yields $f(z) = (z+1)/(1-z)$ which has a period of 4;

$$z \to \frac{1+z}{1-z} \to \frac{-1}{z} \to \frac{1+z}{z-1} \to \frac{z-1}{z+1} \to z. \tag{5.5}$$

It seems that linear fractional functions might be good enough. Indeed, the problem is made easier by noting the link between such functions (Möbius functions) and their relation to 2×2 matrices. Square matrices are particularly useful and the 2×2 variety multiply in the following way,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{pmatrix}.$$

On the other hand, if we compose two different Möbius functions we get

$$\frac{a\left(\frac{\alpha x+\beta}{\gamma x+\delta}\right)+b}{c\left(\frac{\alpha x+\beta}{\gamma x+\delta}\right)+d} = \frac{(a\alpha + b\gamma)x + (a\beta + b\delta)}{(c\alpha + d\gamma)x + (\beta + d\delta)}.$$

Notice that the parameters are changed in exactly the same way. Thus composition of Möbius functions can be represented by multiplication of 2×2 matrices. This allows the question about functions to be changed into one about matrices. We can view a 2×2 matrix as an operation on a cartesian plane. If we want an operation of period n, the simplest to consider would be a rotation of $2\pi/n$ around the origin. After applying this n times, an object will rotate through one entire revolution, returning it to where it began.

Rotation anticlockwise about the origin by an angle θ is given by the formula

$$(x, y) \mapsto (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta),$$

which can be rewritten in terms of matrices as

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}.$$

To answer the question posed, the function

$$f(x) \ = \ \frac{x\cos(2\pi/n) - \sin(2\pi/n)}{x\sin(2\pi/n) + \cos(2\pi/n)}$$

must have a period of length n with respect to function composition.

---

[1] We need this proviso to do away with division by zero.

The requirement that $ad - bc \neq 0$ translates to a non-zero determinant in the 2×2 matrix picture. We have seen that the Möbius transformations are closed under composition, suggesting there may be a group structure to them.

Proposition 5.7. *The Möbius transformations form a group of conformal maps of the Riemann sphere $\hat{\mathbb{C}}$.*

Proof. To find the inverse of f according to Eq.(5.3), define

$$g(z) := \frac{dz - b}{-cz + a}.$$

Observe that $ad - (-c)(-b) = ad - cb \neq)$, so that g is a Möbius transformation. Simple substitution shows that $f(g(z)) = g(f(z)) = z$ for all $z \in \hat{\mathbb{C}}$. Note that $f_{\bar{z}}(z) = 0$, and therefore to verify that f is conformal, we need to prove that $f_z \neq 0$. This is readily checked: Firstly, if $r = 0$ we have $f_z = p/s \neq 0$. Otherwise, if $r \neq 0$, then $f_z(z) = (ad - bc)/(cz + d)^2 \neq 0$ for any $z \in \mathbb{C}$. For the point at infinity, write $w = 1/z$ and then note that $f_z(z) = -(ps - qr)w_z/(r + sw)^2$ by the chain rule. Evaluating at $w = 0$ we clearly see that $f_z(z) \neq 0$, completing the proof. □

An intuitive reason for what has been shown above can be reached by decomposing a Möbius function. A Möbius transformation is equivalent to a sequence of simpler transformations. Define

$$\begin{aligned} f_1(z) &:= z + \frac{d}{c}, \\ f_2(z) &:= \frac{1}{z}, \\ f_3(z) &:= \frac{bc - ad}{c^2} z, \\ f_4(z) &:= z + \frac{a}{c}. \end{aligned}$$

The first and last functions (f_1 and f_4) represent simple translations by d/c and a/c respectively. The third transformation $f_3$is a spiral transformations (a rotation & dilation). The second function f_2 represents an inversion in the unit circle S^1 combined with a reflection in the real axis. It is an algebraic exercise to check that

$$f_4 \circ f_3 \circ f_2 \circ f_1(z) = \frac{az + b}{cz + d} \overset{\text{Eq.(5.3)}}{=} f(z).$$

The Möbius group is usually denoted $\text{Aut}(\hat{\mathbb{C}})$ as it is simply the automorphism group of the Riemann sphere. With every invertible complex 2×2 matrix $\mathfrak{H}$ we can associate the Möbius transformation f,

$$\mathfrak{H} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \mapsto \quad f(z) = \frac{az + b}{cz + d}. \tag{5.6}$$

This identification means that the map $\pi : GL(2, \mathbb{C}) \to \text{Aut}(\hat{\mathbb{C}})$ which sends $\mathfrak{H}$ to f is a group homomorphism. However this is not a bijection since any matrix obtained from a scalar multiple of $\mathfrak{H}$ determines the same transformation. This tells use that $\ker(\pi) = \{\lambda I : \lambda \in \mathbb{C}\}$. Thus the first isomorphism theory provides a description of the Möbius group,

$$\text{Aut}(\hat{\mathbb{C}}) \cong GL(2, \mathbb{C})/((\mathbb{C} \setminus \{0\})I) =: PGL(2, \mathbb{C}),$$

which is called the *projective general linear* group. This name comes from the action of $PGL(2, \mathbb{C})$ on the complex projective $\mathbb{C}P^2$, as discussed in Sec. III, is identical to that of the Möbius group on the Riemann sphere. The bijection is arrived at from the homogeneous coordinates $[z_1, z_2]$ and the ratio z_1/z_2.

Remark Given a set of three distinct points z_1, z_2 and z_3 on the Riemann sphere, and a second set of distinct points w_1, w_2 and w_3, then there exists precisely one Möbius transformation f satisfying $f(z_i) = w_i$ for $i = 1, 2, 3$. Consider the definition

$$h_1(z) := \frac{(z - z_1)(z_2 - z_3)}{(z - z_3)(z_2 - z_1)}, \tag{5.7}$$

which is easily checked to be a Möbius transformation. Then $h_1(z_1) = 0$, $h_1(z_2) = 1$ and $h_1(z_3) = \infty$. If h_2 is defined in a similar way to map w_1, w_2 and w_3 to $0, 1$ and ∞ respectively, then setting $f = h_2^{-1} \circ h_1$ gives the requirement. Moreover, the expression in Eq.(5.7) bears similarity to the cross ratio (3.2). It is an important characteristic of the Möbius function that it leaves the cross ratio invariant. This can be seen by simply checking that it is conserved through h_1 above.

Example Consider the Möbius transformation given by $q(z) = i(1+z)/(1-z)$. The image of the unit circle is found as

$$q(e^{i\theta}) \;=\; i\frac{1+e^{i\theta}}{1-e^{i\theta}} = \frac{1}{\tan(\theta/2)} \in \mathbb{R}.$$

The point $z = 1$ is mapped to ∞. In other words, the unit circle is mapped to the extended real axis. It is a simple check to find that q defined above maps B^2 to $\mathbb{H}^2$.

Corollary 5.8. *Four distinct points z_1, z_2, z_3 and z_4 belong to the same line or circle iff their cross ratio is real. In the projective plane, lines are regarded as generalised circles.*

Using the notation in (5.6), we see that the vector $(z_1, z_2)^T \in \mathbb{C}^2$ is an eigenvector of $\mathfrak{H}$ precisely when $f(z_1/z_2) = z_1/z_2$. Then z_1/z_2 is a fixed point of f. Note that from the Jordan Canonical form of $\mathfrak{H}$, we know that $\mathfrak{H}$ is conjugate to

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \text{ for some } \lambda \neq 0, \quad \text{ or to } \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

for when the eigenvalues of $\mathfrak{H}$ are distinct and both equal to 1 respectively. This has analogous results for Möbius transformations.

Theorem 5.9. *A non-identiy Möbius transformation has either 1 or 2 fixed points in $\hat{\mathbb{C}}$. If it has 1, then it is conjugate to $z \mapsto z+1$. If it has 2, then it is conjugate to $z \mapsto kz$ for some $k/neq0, 1$.*

Proof. First suppose that the Möbius transformation f has two fixed points u and v. We have the leisure of picking another g such that $g(u) = 0$ and $g(v) = \infty$. Then by construction $g \circ f \circ g^{-1}$ fixes 0 and ∞. We therefore know that $g \circ f \circ g^{-1}(z) = \lambda^2 z$ for some $\lambda \neq 1, 0, -1$.

Otherwise, f has precisely one fixed point, u. This time pick g such that $g(u) = \infty$. Then $g \circ f \circ g^{-1}$ fixes only the point ∞. Thus $g \circ f \circ g^{-1}(z) = z + q$. However, we can then form the expression $q^{-1}g \circ f \circ (q^{-1}g)^{-1}(z) = z + 1$. □

This characterisation affords its own nomenclature. A non-identity is named

(i) *parabolic* if it is conjugate to $z \mapsto z+1$, ($\Leftrightarrow \mathrm{tr}\mathfrak{H} = \pm 2$)

(ii) *elliptic* if it is conjugate to $z \mapsto kz$ for $|k| = 1$ ($k \neq 1$), ($\Leftrightarrow \mathrm{tr}\mathfrak{H} \in (-2, 2)$)

(iii) *hyperbolic* if it is conjugate $z \mapsto kz$ for $k \in \mathbb{R}^+$ ($k \neq 0, 1$), ($\Leftrightarrow \mathrm{tr}H \in (-\infty, -2) \cup (2, \infty)$)

(iv) *loxodromic* if it is conjugate $z \mapsto kz$ for $|k| \neq 1$ and $k \notin \mathbb{R}^+$. ($\Leftrightarrow \mathrm{tr}\mathfrak{H} \notin \mathbb{R}$)

The connection to the trace of $\mathfrak{H}$ is easily seen by observing that for different eigenvectors λ and λ^{-1},

$$\mathrm{tr}\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} = \lambda + \lambda^{-1} \geq 2,$$

from the arithmetic-geometric mean inequality when $\lambda > 0$. The negative case is a trivial modification and the boundaries are readily checked.

We now disclose our purpose for the discussion of Möbius transformations, and the reason why they are of core importance to Hyperbolic geometry [4]. Indeed the following proposition determines which of the Möbius transformations of $\mathbb{H}^2$ are isometries of $\mathbb{H}^2$. First note that with such a restriction $a, b, c \in \mathbb{R}$, an associated Möbius transformation will leave the real line unchanged. Moreover, we find that $\mathbb{H}^2$ is mapped to itself:

$$\Im\left(\frac{az+b}{cz+d}\right) = \frac{1}{2i}\left(\frac{az+b}{cz+d} - \frac{a\bar{z}+b}{c\bar{z}+d}\right) = \Im(z)\left(\frac{ad-bc}{|cz+d|^2}\right) = \frac{\Im(z)}{|cz+d|^2}, \tag{5.8}$$

meaning that is $\Im(z) > 0$ implies $\Im(f(z)) > 0$.

Definition The special linear group $SL(2, \mathbb{R})$ is the group of all real 2×2 matrices with unit determinant,

$$SL(2, \mathbb{R}) := \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{R} \quad \text{and} \quad ad - bc = 1\}. \tag{5.9}$$

Observe that a Möbius transformation remains unchanged if the variables a, b, c and d have their sign flipped. Therefore we define $PSL(2, \mathbb{R})$ as the set of matrices in $SL(2, \mathbb{R})$ with two matrices A, B identified iff $A = -B$. This is called the *projective special linear group*.

Remark The projective special linear group emerges naturally as the quotient

$$PSL(2,\mathbb{R}) \;\cong\; SL(2,\mathbb{R})/\{I,-I\}.$$

Proposition 5.10. *Let f be a Möbius transformation of $\mathbb{H}$ and let $z, w \in \mathbb{H}$. Then the distance is unchanged:*

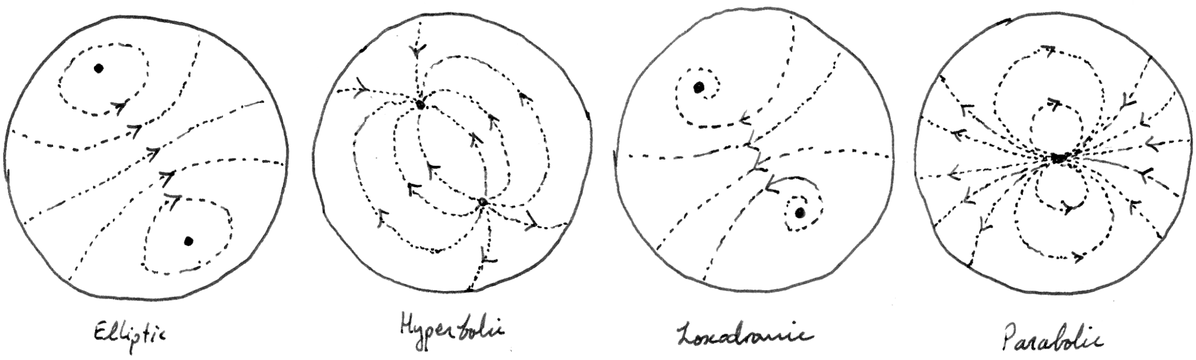
$$d_{\mathbb{H}}(f(z), f(w)) \;=\; d_{\mathbb{H}}(z,w). \tag{5.10}$$

Proof. It will suffice to show that the length of a path γ from z to w keeps its value. Note that $f\circ\gamma$ is a path from z to w, suggesting we apply the chain rule:

$$\ell_{\mathbb{H}}(f\circ\gamma) \;=\; \int_0^1 \mathrm{d}t \frac{|f'(\gamma(t))||\gamma'(t)|}{\Im(f\circ\gamma(t))} = \int_0^1 \mathrm{d}t \frac{|\gamma'(t)|}{\Im(\gamma(t))|} = \ell_{\mathbb{H}}(\gamma),$$

where we have used the previous results (5.8) and the chain rule, to simplify the expression. □

The classification of Möbius transformations carries over to the isometries of $\mathbb{H}^2$. Their actions on the Klein disk are shown below:



We can do better. The next theorem identifies all isometries of $\mathbb{H}^2$ in terms of Möbius transformations plus reflection in the imaginary axis.

Theorem 5.11. *The group* $\mathrm{Isom}(\mathbb{H}^2)$ *is generated by Möbius transformations from $PSL(2,\mathbb{R})$ together with $z \mapsto -\bar{z}$.*

Proof. Let ϕ be an isometry of $\mathbb{H}^2$. Then geodesics are mapped to geodesics, so in particular the imaginary axis I is mapped to a geodesic. Now take another isometry $\psi \in PSL(2,\mathbb{R}$ such that $\psi\circ\phi(I) = I$. Then we can ensure that ψ fixes the entire axis by composing it with dilations ($z \mapsto kz$) and inversions ($z \mapsto -1/z$) to fix say i, and hence all of I. We know that it takes three independent points to fix an isometry (by triangulation). Since we have entirely determined a geodesic in $\mathbb{H}^2$, it remains to determine the orientation. Hence there are two options: $\psi\circ\phi(z) = z$ or $\psi\circ\phi(z) = -\bar{z}$, that is, $\psi\circ\phi$ is either the identity, or it is a reflection in the imaginary axis.

If $\psi\circ\phi = \mathrm{Id}$, then $\phi = \psi^{-1} \in PSL(2,\mathbb{R})$. Otherwise, if $\psi\circ\phi(z) = -\bar{z}$ for all $z \in \mathbb{H}^2$, then

$$\phi(z) \;=\; \frac{a\bar{z}+b}{c\bar{z}+d}, \tag{5.11}$$

where $ad - bc = -1$. This verfies the theorem statement. □

And so we have classified all the isometries of $\mathbb{H}^2$. It is the sign of the determinant in Eq. (5.11) which characterises the *orientation* of an isometry. The isometries which preserve the orientation are $PSL(2,\mathbb{R})$, which hence form a subgroup of $\mathrm{Isom}(\mathbb{H}^2)$ of index two.

VI. KLEINIAN & FUCHSIAN GROUPS

The connection between the theory of Kleinian groups and Hyperbolic geometry was made by Thurston [24]. The result is that when a Kleinian group G is isomorphic to the fundamental group of π_1 of a hyperbolic 3-manifold, then the quotient space $\mathbb{H}^3/G$ is a model for that manifold. We shall not prove this result, but we shall make a start at the description of Kleinian groups and their properties.

A. Hyperbolic 3-space

Let Q be a sphere orthogonal to the unit sphere S^2. Then $\gamma := Q \cap S^2$ a circle. It is well known that orthogonal spheres are fixed by inversion. Hence an inversion in Q fixes γ and maps S^2 onto itself. In fact, it is not hard to see that B^3 will also be mapped to itself. Notice that for any plane p which passes through the centre of Q intersects it in a circle γ', then restricted to p, the map is nothing but inversion in γ'. We have found an isometry:

Proposition 6.1. *Let ϕ be an inversion in $Q \perp S^2$; Then ϕ is an orientation reversing isometry for the hyperbolic metric.*

It quickly becomes clear that many of the propositions for 2-dimensional inversions hold in the 3-dimensional case as well. For example: for each point $x \in B^3$ there is an inversion in a sphere orthogonal to S^2 which interchanges the origin and x. Moreover, we attain a Poincaré model using B^3 for $\mathbb{H}^3$. Between any two distinct points in B^3, there is a unique path with the shortest hyperbolic length. This is an arc of a circle which is orthogonal to S^2.

Let's now bring Möbius transformations into the game. Let $f \in \text{Aut}(\hat{\mathbb{C}})$. Then it is well-known that f can be written as the composition of inversions in an even number of circles, $\gamma_1, \gamma_2, \gamma_3, \ldots, \gamma_{2N}$. For each of these circles there is a unique sphere Q_n orthogonal to S^3 such that $Q_n \cap S^3 = \gamma_n$. Let ϕ_n be an inversion in Q_n. Then we know that $\phi_{2N} \circ \ldots \circ \phi_1$ acts on S^2 as f. This scheme extends f, to some function $\widetilde{f}$ from $\mathbb{R}^3 \cup \{\infty\}$ to itself. The map agress on S^2 with f and maps B^3 to itself. The function $\widetilde{f}$ is called the *Poincaré extension* of the Möbius transformation f.

Proposition 6.2. *For every Möbius function $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$, the extension $\widetilde{f}$ maps the unit ball to itself and is an orientation preserving isometry for the hyperbolic metric.*

Proof. In the notation above, we know that each ϕ_n is an orientation *reversing* isometry. Thus $\widetilde{f}$ is an even number of orientation reversing isomtries and therefore and orientation preserving isometry. □

It follows from this that every Möbius function has a unique extension. The next theorem will come as not surprise, based on our discussion of $\mathbb{H}^2$. The proof follows [4].

Theorem 6.3. *Every orientation preserving isometry of hyperbolic 3-space in B^3 is $\widetilde{f}$ for some Möbius transformation f.*

Proof. Suppose that $\phi : B^3 \to B^3$ is an orientation preserving isometry for B^3 equipped with the hyperbolic metric. Then we know that there is an inversion ψ in a sphere orthogonal to S^2 such that $\psi \circ \phi(\mathbf{0}) = \mathbf{0}$. We determined the form of $d_{\mathbb{H}^2}$, and to achieve this in B^2 we simply need to apply the Möbius transformation

$$z \mapsto i\frac{1+z}{1-z},$$

which sends B^2 to $\mathbb{H}^2$, as in Sec. V C. In Sec. III E we found that in $\mathbb{H}^2$, the distance between i and iy where $0 < y < 1$ is $\ln(1/y)$. Notice that the origin in B^2 is mapped to i in $\mathbb{H}^2$. This means, that when we go up to B^3, if we take $\boldsymbol{u} \in S^2$, then (where $0 < t < 1$),

$$d_{B^3}(\mathbf{0}, t\boldsymbol{u}) \;=\; \ln\left(\frac{1+t}{1-t}\right).$$

Thus the curve $\sigma(t) := t\boldsymbol{u}$ is a geodesic ($0 \leq t < 1$). Therefore $\psi \circ \phi \circ \sigma$ is also a geodesic which starts at the origin. This enables us to write

$$\psi \circ \phi(t\boldsymbol{u}) \;=\; t\boldsymbol{v},$$

where $\boldsymbol{v} \in S^2$, where $\boldsymbol{v} = g(\boldsymbol{u})$.

Consider two unit vectors $\boldsymbol{x}, \boldsymbol{y} \in S^2$, then as $t \to 0$ the vectors $t\boldsymbol{x}$ and $t\boldsymbol{y}$ become separated by a euclidean distance $t2|\boldsymbol{x}-\boldsymbol{y}| + \mathcal{O}(t^2)$. Therefore:

$$\lim_{t\to 0^+} \frac{d_{B^3}(t\boldsymbol{x}, t\boldsymbol{y})}{t} = 2|\boldsymbol{x}-\boldsymbol{y}|.$$

Substituting the isometry $\psi \circ \phi$ into the above gives $|g(\boldsymbol{x}) - g(\boldsymbol{y})| = |\boldsymbol{x}-\boldsymbol{y}|$. This forces the function g to preserve the Euclidean inner product, and hence $g \in O(3)$. Generically this means that g is a combination of reflections through the origin (recall that in Sec. V B we mention that any rotation can be achieved through reflections). These reflections are included in the space of inversions in spheres orthogonal to S^2, giving $\psi \circ \phi$, and hence ϕ as such a composition of inversions. □

It is worth remarking at this stage that the upper half-space model,

$$\mathbb{H}^3 := \{(x,y,z) \in \mathbb{R}^3 : z > 0\}. \tag{6.1}$$

The boundary $\partial\mathbb{H}^3 = \hat{\mathbb{C}}$. The hyperbolic metric is attained in and identical way to $d_{\mathbb{H}^2}$. The results from B^3 carry through by an inversion in the sphere

$$Q := \{\boldsymbol{x} \in \mathbb{R}^3 : |\boldsymbol{x} + (0,0,1)| = \sqrt{2}\}.$$

This map travels bijectively between B^3 and $\mathbb{H}^3$. Hence we have found

$$PGL(2,\mathbb{C}) \cong \mathrm{Isom}^+(\mathbb{H}^3).$$

B. Discrete Subgroups

We have already remarked that the Möbius transformations are naturally identified with the space of all 2×2 matrices, denoted by $GL(2,\mathbb{C})$. As a finite dimensional vector space over $\mathbb{C}$, we may define a norm which will induce a metric. This norm will induce a metric, and we are able to study properties of $GL(2,\mathbb{C})$, as a metric space.

Definition A subgroup G of $SL(2,\mathbb{C})$ is called discrete if there is a neighbourhood $N \in SL(2,\mathbb{C})$ such that $G \cap N = \{Id\}$.

This is the same definition we made in Sec. V B, in the discussion on crystallographic groups. The method is not different, we are now describing isometries of $\mathbb{H}^3$.

Remark An equivalent formulation of the above defintion reads as follow. A subgroup G of $\mathcal{M}$ is discrete iff there is no sequence of distinct elements $g_n \in G$ with $g_n \to Id$.

Example Consider $G < SL(2,\mathbb{C})$ and let $f \in SL(2,\mathbb{C})$. Then G and $fGf^{-1} := \{fgf^{-1} : g \in G\}$ are either both discrete or both non-discrete. This is based on the observation that if $g_n \to Id$, then $fg_nf^{-1} \to Id$ by continuity. Hence if G is not discrete then neither is fGf^{-1}. Similarly, if fGf^{-1} is not discrete and $fg_nf^{-1} \to Id$, then $g_n \to Id$.

The above results hold equally for subgroups of $PSL(2,\mathbb{C})$, and we shall be primarily concerned with them as our consideration of isometries limits us to the Möbius groups.

Definition A discrete subgroup of $PSL(2,\mathbb{C})$ is called a Kleinian group. A discrete subgroup of $PSL(2,\mathbb{R})$ (or Γ) is called a Fuchsian group. Evidently a Fuchsian group is a Kleinian group.

Note that a finite subgroup of $PSL(2,\mathbb{C})$ is automatically a Kleinian group.

Proposition 6.4. *Every finite subgroup of $PSL(2,\mathbb{C})$ is conjugate in the Möbius group to a subgroup of $SO(3)$.*

Proof. Consider a finite Kleinian group G acting on the unit ball B^3. The orbit $W = G\boldsymbol{p}$ is finite for any $\boldsymbol{p} \in B^3$, and therefore there is a unique closed ball $\bar{B}(\boldsymbol{p}_0, r)$ of smallest hyperbolic radius r containing W. (This fact is slightly subtle, but the proof mirrors that for Euclidean space).

We know that $g \in G$ acts isometrically on B^3 and hence can only permute elements of W. We therefore find

$$W = g(W) \subset g(\bar{B}(\boldsymbol{p}_0, r)) = \bar{B}(g\boldsymbol{p}_0, r),$$

for each $g \in G$. This ball has the same smallest radius, and uniqueness thus gives that $g\boldsymbol{p}_o = \boldsymbol{p}_0$ for each $g \in G$. In other words, $\boldsymbol{p}_0$ is a fixed point for all elements of G.

Now conjugate G by a Möbius transformation f taking $\boldsymbol{p}_0 \mapsto 0$. Then fGf^{-1} is a group of isometries which fix the origin. It follows that $fGf^{-1} < SO(3)$, which means exactly that G is conjugate to a subgroup of $SO(3)$. □

Therefore, finite Kleinian groups are in some sense "ordinary" and we must pursue infinite groups to find interesting examples.

C. Discontinuous Subgroups

Definition A subgroup G of $PSL(2,\mathbb{C})$ acts discontinuously at a point $\boldsymbol{x} \in \mathbb{H}^3$ if there is some neighbourhood $U \ni \boldsymbol{x}$ such that there are only finitely many $g \in G$ for which $U \cap g(U)$ is non-empty.

This splits the Riemann sphere into two categories for a given G. The *regular set* $\Omega(G) \subset \hat{\mathbb{C}}$ is the set of all points at which G acts discontinuously. The complement $\Lambda(G) := \hat{\mathbb{C}} \setminus \Omega(G)$ is called the *limit set* of G. The regular set is open in $\mathbb{C}$ (by definition), and if it is also non-empty, we call G a *discontinuous group*. Note from the definition that $\Omega(G)$ is invariant under action by G and therefore so is $\Lambda(G)$. Since $\Omega(G)$ is always open, $\Lambda(G)$ is closed.

Example If a Kleinian group acts discontinuously at $\boldsymbol{x} \in \mathbb{H}^3$, then the *stabilizer*, defined by

$$\mathrm{Stab}(\boldsymbol{x}) \ := \ \{g \in G : g(\boldsymbol{x}) = \boldsymbol{x}\},$$

is a finite subgroup of G and hence conjugate to a subgroup of $SO(3)$.

Proposition 6.5. *If G is discontinuous then it is discrete.*

Proof. Assume the contrary and let $\boldsymbol{x} \in \mathbb{H}^3$. Let g_n be a discontinuous sequence of elements of G converging to the identity, $g_n \to Id$. This means that for all $\epsilon > 0$, there is some integer N_ϵ such that $d_{\mathbb{H}^3}(g_n(\boldsymbol{x}) - \boldsymbol{x}) < \epsilon$ for all $n > N_\epsilon$. Therefore, if U is a neighbourhood of $\boldsymbol{x}$ containing $B(\boldsymbol{x}, \epsilon)$, then $g_n(U) \cap U$ is non-empty for $n > N_\epsilon$. This means that G does not act discontinuously at $\boldsymbol{x}$. But $\boldsymbol{x}$ was arbitrary. Therefore we have shown that if G is not discrete then it is not continuous. □

The converse of the above is not true, and an example can be found in [14].

Consider $g \in G$ being a loxodromic or hyperbolic transformation. In this case there are two fixed points: one "source" and one "sink". Therefore $g^n(\boldsymbol{x})$ tends to one fixed points whilte $g^{-n}(\boldsymbol{x})$ tends to the other one as $n \to n$. Therefore both fixed points are in $\Lambda(G)$. The fixed points of an elliptic transformation need not lie in the limit set.

As we did for the crystallographic groups, we may imagine that $\mathcal{R}$ is a fundamental set for G acting on $\mathbb{H}^3$. Then hyperbolic 3-space is tiled by the copies $g(\mathcal{R})$ where $g \in G$. This gives another way of determining whether G acts discontinuously at a point $\boldsymbol{x} \in \mathbb{H}^3$: precisely when there is a neighbourhood of $\boldsymbol{x}$ which meets only finitely many copies $\mathcal{R}$. For the Euclidean plane groups, the limit set is trivial. It is simply the point at infinity. Hyperbolic geometry on the other hand offers a rich structure to the limit set $\Lambda(G)$. Beautiful fractal patterns have been observed from the action of Kleinian groups [19].

VII. CONCLUSION

A. Back to curvature

Returning back to our notion of curvature, one may note the similarity between propositions (3.3) and (3.6), we see that the angles are closely related to the area. We now follow the discussion in [10]. Consider a surface Σ embedded in $\mathbb{E}^3$ and triangulate it. Recover a polyhedra by "flattening" the triangles. Make no assumption about the genus, so that the Euler characteristic is $\chi(\Sigma) = v - e + f$. We are in a similar position as for Eq.(4.4), and therefore $2e = 3f \Leftrightarrow f = 2(e - f)$.

Each vertex is surrounded by planes which make some angle α with this vertex. We define the *angle defect* at vertex vi as the amount by which the sum of these angles differs from 2π,

$$\delta(v) \;=\; 2\pi - \sum_i \alpha_i(v),$$

where the sum runs over all adjoining planes to v. The total angular defect can be calculated by recalling that the sum of the angles in each face f is π, since they are nothing but flat Euclidean triangles. Hence the total angular defect is

$$\sum_k \delta(v_k) \;=\; 2\pi v - \pi f = 2\pi(v - e + f) = 2\pi\chi(\Sigma). \tag{7.1}$$

The above formula amounts to the Gauss-Bonnet theorem for polyhedra. This theorem holds in a smooth setting as well.

Theorem 7.1. (Gauss-Bonnet theorem for closed surfaces) *If Σ is a compact two dimensional Riemannian manifold (without boundary), let κ be the Gaussian curvature of Σ. Then*

$$\int_\Sigma dA\ \kappa = 2\pi\chi(\Sigma), \tag{7.2}$$

where dA is the area element of the surface.

Proof. (outline) The compactness of the Riemannian manifold allows us to take and infinitesimal triangulation. The claim is that the angular defect becomes the curvature. Observe that one of the limits in the Bertrand-Diquet-Puiseux theorem appears to take this form 2.5. We shall try a different tack. Define the angular defect of a polygon to be the amount by which the sum of its angles fails to make the sum of angles of a Euclidean polygon with the same number of edges,

$$\delta(v_1, v_2, \ldots, v_n) \;=\; (n-2)\pi - \sum_{k=1}^{n} \alpha_k$$

This measures the size of the rotation suffered by a vector during a counterclockwise parallel transport about the polygon. Then, provided the following limit exists, the curvature *should* be equal to

$$\kappa(x) \;=\; \lim_{\mathcal{P}\to x} \frac{\delta(\mathcal{P})}{\mathrm{Area}(\mathcal{P})}, \tag{7.3}$$

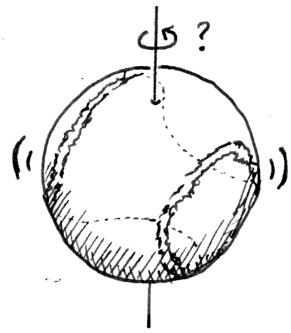
where the above limit assumes that the polygon $\mathcal{P}$ can be shrunk to a point. Then the total curvature over a surface Σ is given by

$$\lim_{\mathcal{P}\to x} \sum \delta(\mathcal{P}) \;=\; \int_\Sigma \mathrm{d}A\kappa.$$

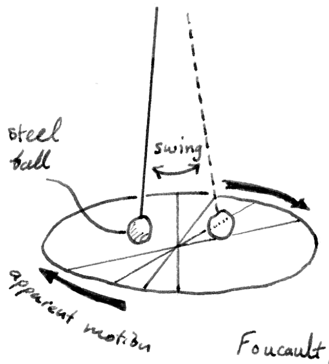
The above statements are made precise by taking into consideration the additivity of δ over polygons. The compactness of Σ permits the limits to be passed, and enables the curvature to equate to the Euler characteristic. It should be noted that the definition in (7.3) is consistent with our first definition of curvature. □

A geodesic curve γ on a surface Σ is such that the tangent vector remains tangent under parallel transport along γ. Geodesics are the "straight lines" on Σ, and we have already used this fact frequently. The *geodesic curvature* is defined as follows for an oriented curve on Σ. Approximate this curve by geodesic segments γ_i to obtain a polygonal shape $\mathcal{P}$ on Σ. Now tracing out the tangent vector to the γ_i as one moves along the boundary of $\mathcal{P}$, we see that the vector rotates by and angle α_i at each of the corners (signed so that $\alpha_i > 0$ if γ_i turns left and $\alpha_i < 0$ if γ_i turns right). Then the geodesic curvature is the total rotation $\sum_i \alpha_i$. The formulation leads to an inclusion of surfaces with boundary to the Gauss-Bonnet theorem [10]: *the total geodesic curvature of a simple Riemannian surface (no holes!) Σ with boundary $\partial\Sigma$, plus the total curvature of Σ is equal to Σ.* We now present two examples to demonstrate this form of the Gauss-Bonnet theorem.

Example A tennis ball has a closed curve etched on its surface which symmetrically divides the surface into two equal pieces. Mark a point on this curve and place the ball on a flat surface touching at this point. Roll the ball in such a way that the point of contact is always on the curve (moving in one direction along the curve) until it again returns to the marked point. The vertical axis perpendicular to the plane is the same for initial and final positions and therefore the tennis ball has made some revolution about this vertical axis. The Gauss-Bonnet formula tells us what this angle is. The angle is nothing but the total curvature along this curve. The shape of this curve is complicated, but we know that the total curvature on its interior is 2π (half the total of 4π for a sphere). Since $\chi(B^2) = 1$, we find that the angle is $2\pi\chi(S^2) - 2\pi = 0$. So the sphere makes a zero angle of revolution.



Wobbling tennis ball



Foucault pendulum

Example Another related example is the Foucault pendulum which demonstrates the rotation of the earth. If the pendulum is situated at the north pole, then the plane of motion remains fixed. Therefore with respect to the earth it rotates by $\pi/12 \approx 15°$ per hour. Now suppose that the pendulum is suspended from a point at a latitude of ψ. Then the total curvature of the polar cap is simply the area for a unit sphere, namely $2\pi(1 - \sin\psi)$. As in the case for the tennis ball, we conclude that total rotation of the plane of motion of the pendulum is $2\pi - 2\pi(1 - \sin\psi) = 2\pi\sin\psi$. For Paris, $\psi \approx 48°$ which gives a rotation of $\pi\sin\psi/12 \sim 11°$ per hour.

B. Final comments

We have seen that surfaces of constant curvature obey some degree of order. The following theorem is a precise description of all surfaces with constant curvature which are further (a) compact and (b) orientable.

Theorem 7.2. (Poincaré-Koebe Uniformisation Theorem) *Let Σ be a compact orientable surface of constant curvature without boundary. Then there is a covering space M and a discrete group of isometries Γ such that the quotient M/Γ is homeomorphic to the surface Σ. Furthermore, the covering space M is completely determined by the curvature as follows.*

(i) if the curvature of Σ is zero, then M is the Euclidean plane $\mathbb{E}^2$,

(ii) if the curvature of Σ is positive, then M is the sphere S^2,

(iii) if the curvature of Σ is negative, then M is the hyperbolic plane $\mathbb{H}^2$.

It was Thurston [26] who first made the bold conjecture of a uniformisation of 3-manifolds. In it, he laid out eight model geometries in a similar manner to the three above. Unsurprisingly, the situation with 3-manifolds is considerably more complicated than those of only 2 dimensions which we have largely been discussing. Extensive work was done in an attempt to prove the geometrization conjecture and much progress was made by Hamilton [12] who used the *Ricci flow* to make certain statements about the geometries. Only recently, and somewhat out of the blue, did Perelman [21] prove the conjecture using the same methods introduced by Hamilton.

One might wonder how things are going on the ($n \geq 4$)-manifold side of things. Here we encounter the so called *exotic* spaces which make such a classification impossible! Milnor [17] developed the structure of differentiable manifolds of dimension ≥ 5. It was Freedman and Donaldson who produced results on 4-manifolds [7, 9] for which they received the Fields Medal at the 1986 International Congress of Mathematicians. In some sense therefore, the 3 and 4 dimensional manifolds have proven to be the most subtle.

The eight geometries laid out by Thurston include, as one might expect, the more familiar $\mathbb{E}^3$, S^3 and $\mathbb{H}^3$. In 1982 Thurston proved the theorem for a large class of manifolds, (the *Haken* manifolds) and this is what led him to make the conjecture.

[1] R.B.J.T. Allenby. *Rings, Fields and Groups: An Introduction to Abstract Algebra.* Modular Mathematics Series. Edward Arnold, 1991.

[2] A. Armstrong. *Basic Topology.* Undergraduate Texts in Mathematics. Springer, 2010.

[3] E.T. Bell. *The Development of Mathematics.* Dover books explaining science. McGraw-Hill book Company, Incorporated, 1945.

[4] T. K. Carne. *Geometry and groups.* Lecture notes, Cambridge University. 2006.

[5] R. Courant and H. Robbins. *What is Mathematics?: An Elementary Approach to Ideas and Methods.* Oxford Paperbacks. Oxford University Press, 1996.

[6] H.S.M. Coxeter and S.L. Greitzer. *Geometry Revisited.* Number v. 19 in Anneli Lax New Mathematical Library. Mathematical Association of America, 1967.

[7] Simon K Donaldson. An application of gauge theory to four dimensional topology. *Journal of Differential Geometry*, 18(2):279–315, 1983.

[8] Euclid, T.L. Heath, and D. Densmore. *Euclid's Elements: All Thirteen Books Complete in One Volume.* Green Lion Press, 2002.

[9] Michael Hartley Freedman. The topology of four-dimensional manifolds. *Journal of Differential Geometry*, 17(3):357–453, 1982.

[10] D.B. Fuchs and S. Tabachnikov. *Mathematical Omnibus: Thirty Lectures on Classic Mathematics.* American Mathematical Society, 2007.

[11] Karl Friedrich Gauss. *General investigations of curved surfaces of 1827 and 1825.* 1902.

[12] Richard S Hamilton. Three-manifolds with positive ricci curvature. *J. Differential Geom*, 17(2):255–306, 1982.

[13] H. Hopf and S.S. Chern. *Differential Geometry in the Large: Seminar Lectures New York University 1946 and Stanford University 1956.* Lecture Notes in Mathematics. Springer, 1989.

[14] L. Keen and N. Lakic. *Hyperbolic Geometry from a Local Viewpoint.* London Mathematical Society Student Texts. Cambridge University Press, 2007.

[15] J.M. Lee. *Riemannian Manifolds: An Introduction to Curvature.* Graduate Texts in Mathematics. Springer, 1997.

[16] E.H. Lockwood and R.H. Macmillan. *Geometric Symmetry.* Cambridge University Press, 1978.

[17] John Milnor. On manifolds homeomorphic to the 7-sphere. *The Annals of Mathematics*, 64(2):399–405, 1956.

[18] John W Morgan and Gang Tian. *Ricci flow and the Poincaré conjecture.* 2006.

[19] D. Mumford, C. Series, and D. Wright. *Indra's Pearls: The Vision of Felix Klein.* Indra's Pearls: The Vision of Felix Klein. Cambridge University Press, 2002.

[20] B. O'Neill. *Elementary Differential Geometry, Revised 2nd Edition.* Elementary Differential Geometry Series. Elsevier Science, 2006.

[21] Grisha Perelman. The entropy formula for the ricci flow and its geometric applications, 2002. *arXiv preprint math.DG/0211159*, 1, 2006.

[22] I. Stewart. *Professor Stewart's Hoard of Mathematical Treasures.* Profile Books, 2010.

[23] D. Taimina. *Crocheting Adventures with Hyperbolic Planes.* Ak Peters Series. A. K. Peters, 2009.

[24] W.P. Thurston. *The geometry and topology of three-manifolds.* Princeton lecture notes. 1980.

[25] W.P. Thurston. Three dimensional manifolds, kleinian groups and hyperbolic geometry. *Bulletin (New Series) of the American Mathematical Society*, 6(3):357–379, 1982.

[26] W.P. Thurston. On the geometry and dynamics of diffeomorphisms of surfaces. *Bulletin (New Series) of the American Mathematical Society*, 19(2):417–431, 1988.

[27] W.P. Thurston and S. Levy. *Three-dimensional Geometry and Topology.* Number v. 1 in Luis A.Caffarelli. Princeton University Press, 1997.

[28] T. Tokieda. *Science from a sheet of paper?* Lectures at the 2012 International Summer School in Mathematics for Young Students. 2012.