Introduction to hyperbolic geometry

D. Massart

September 13, 2023

Abstract

Course to be given at the Nesin Village, september, 2023.
Bu ders Eylül, 2023, Nesin köyüde verilecek.

Contents

Part I
Two-dimensional hyperbolic geometry

1 Introduction

I guess most of you are already familiar with, and much better than me at, Euclidean geometry, which rests on the following five axioms :

1. through any two points goes a line (beware : uniqueness is not formally required, but is used by Euclid)

2. any straight line segment may be extended to infinity

3. given any two points A and B, there is a circle centered at A through B (again, uniqueness is implicitly understood)

4. all right angles are equal (in modern language, given any two of them, there is an isometry that sends one to the other)

5. if two straight lines d_1 and d_2 intersect a third one d, so that the sum of the angles on one side of d at the intersection points is less than the sum of two right angles, then d_1 and d_2 intersect on the same side of d.

As you see the fifth one is a bit of an ugly duckling and its status (axiom or theorem) has been questioned almost since it was stated, but the question was only settled in the XIXth century. The fifth axiom is often replaced by its equivalent formulation (due to Playfair) which says "through a given point there is at most one parallel to a given line". Note that the better known phrasing "through any given point goes exactly one parallel to a given line" is only equivalent to the fifth postulate if you assume uniqueness in the first axiom.

In this course I'll tell a bit of the story (spoiler : it is actually an axiom, not a theorem) and why it took so long.

I'm no historian so I'll probably get some facts wrong along the way and any correction is welcome. Part of the reason it took so long is that the problem is not very well stated, so I'll start by translating the axioms to our modern language.

First, we need a space (which I'll denote X). It could be an abstract space (historically this is how hyperbolic geometry was discovered, and we'll take that viewpoint in Section 6) but for now you may think of it as a surface in $\mathbb{R}^3$. This viewpoint has the advantage that it makes very obvious how to measure distances and angles on X, but beware that I'll always consider *intrinsic* distances, that is, distances as they would be measured by small beings who live on X and cannot fly nor dig into the surface of X. In this setting, the third axiom obviously translates to

3'. Between any two points in X, the distance can be measured.

In other words, it just says that distances are well-defined.

Now, a bit of vocabulary : we shall henceforth call **geodesic** a shortest path parametrized by time. Geodesics will henceforth play the role of straight lines in Euclidean geometry. In our language, the first axiom reads

1'. Between any two points in X, there exists a geodesic.

Note that it implies the connectedness of X. Also note that you could perfectly well have 3' without 1'.

The second axiom is a bit trickier. The ambiguity lies in the difference between a parametrized curve (a map $c : I \to X$ from some interval I to X), and a curve as a geometric object, where we forget about the parametrization and just consider the image $c(I)$.

In the geometric setting (which was probably the viewpoint of Euclid), extending to infinity means eventually leaving every compact set, that is,

2" Given a geodesic c between two points, and a compact set $K \subset X$, there exists a geodesic c' which contains c and is not contained in K. .

So for instance, if your curve is the circle $x^2 + y^2 = 1$, it does not extend to infinity. Note that 2" precludes X from being compact.

However, if we take the parametrized viewpoint, the second axiom could be interpreted as

2' Given a geodesic $c : I \to X$, there exists a geodesic $c' : \mathbb{R} \to X$, parametrized at unit speed, such that the restriction $c'_{|I} = c$.

In that case, the circle, viewed as the parametrized curve $[0, 2\pi] \to \mathbb{R}^2$ which maps t to $(\cos t, \sin t)$, does extend to infinity.

You might say "but then it is too easy : every geodesic may be extended to infinity" ! Can you think of an example where it may not ? In fact, this extension property is equivalent to the topological completeness of X, which is just a fancy way of saying X has no holes in its surface.

The parametrized viewpoint is not just a slick way of moving the goalposts, either : it comes naturally from the laws of nature, which, as Newton famously (is supposed to have) said, are expressed in differential equations, and this is often the only way to find the geodesics, although in the cases that we consider we shall rely on symmetry considerations instead.

For instance, if you are a physicist, you are probably thinking something like this : the shortest path is chosen by nature when there are no external forces (this is the Least Action, or Maupertuis, Principle). You know Newton's equation $F = m\gamma$, so your geodesic is a solution of a second order (since γ is the second derivative) differential equation.

Now we have the Cauchy-Lipschitz Theorem, which says that any (not too horrible) differential equation has a solution that lives for some time, but there is no guarantee that this time is infinite (for instance, the universe could explode this evening). Axiom 2' provides such a guarantee.

The modern formulation of the fourth axiom is in terms of transformation groups, in the spirit of Felix Klein : geometry is the study of group actions. It is a way of saying that the universe is homogenous (every two points look the same) and isotropic (every two directions look the same). What does it mean for two points $x, y \in X$ to look the same ?

First let us define what an **isometry** is : it is a map $f : X \to X$ such that for any two points $x, y \in X$, the distance $d(f(x), f(y)) = d(x, y)$.

Then we say that two points $x, y \in X$ look the same if it means there exists an isometry f such that $f(x) = y$. We say that X is homogenous, or that the group of isometries act transitively on X, if any two points look the same.

And what does it mean for two directions to look the same ? In this course I'll use the following convention : I'll say that X is isotropic, or any two directions look the same, if for any two geodesics $c, c' : \mathbb{R} \to X$, there exists an isometry f such that $f(c) = c'$. I'll say the group of isometries acts doubly transitively on X if X is homogenous and isotropic. So our fourth axiom is

4' The isometry group of X acts doubly transitively on X. .

Ok, those are the axioms we work with, now can we do anything meaningful with them ?

2 Spherical geometry

You may be wondering, what's this all about with those akward sounding axioms ? what's wrong with plain old plane geometry ? well, even when flying on an airplane, you usually don't go straight through space, do you ? The geometry you have to deal with in real life is not exactly plane geometry, even though plane geometry is a good approximation as long as you don't travel too far.

So, how do we measure distances on a sphere ? how can we find the geodesics ? Well, we know what the geodesics are. But how do we know ? And, come to think of it, how do we know that straight lines are geodesics in the Euclidean plane ?

For definiteness, we shall assume the sphere $\mathbb{S}^2$ is defined by the equation $x^2+y^2+z^2 = 1$ in $\mathbb{R}^3$.

2.1 Back to the Euclidean world for a minute

Before finding the geodesics of the sphere we should turn back to the good old plane and ask ourselves, how exactly do we know that the straight line is the Euclidean shortest path between A and B ?

Recall Newton's equation $F = m\gamma$, here there is no external force so $F = 0$, hence the acceleration γ is zero, that is, a shortest path c satisfies $\ddot{c} = 0$, which tells you that $c(t)$ is an affine function of t, that is, a rectilinear uniform movement.

If you are moving on a sphere instead of a plane, then there must be some external force (namely, gravity) that keeps you on the sphere and prevents you from flying into space, so the Least Action Principle says that your acceleration is perpendicular to the sphere.

Let $(x(t), y(t), z(t))$ be a geodesic in $\mathbb{S}^2$, parametrized by time, then its acceleration is perpendicular to the sphere, that is, it is proportional to $(x(t), y(t), z(t))$.

This could be translated into a system of differential equations, albeit not a particularly pleasant one to solve. But a physicist is not so easily discouraged : he calls his favourite concept to the rescue. And what is every physicist's favourite concept ? you guessed it, it's symmetry !

What do we know about the Euclidean plane's symmetries (i.e. isometries) ? We know that $\mathbb{R}^2$ is homogeneous (all points are the same), and isotropic (all directions are the same). From homogeneity we deduce that translations are isometries, and from isotropy, that rotations are isometries. In fact we have a third invariance property (from which the first two could be deduced) : left and right are just conventions, the laws of physics are the same across the looking-glass, though some authors [1] have claimed otherwise. From this we infer that reflections in straight lines are isometries.

Now, we know that geodesics are solutions of a second-order differential equation, Newton's equation. Even though we cannot (or, in the case of the sphere, are too lazy to) solve the differential equation, this still tells us something, thanks to Cauchy's Existence and Unicity Theorem : given an initial position (x, y), and an initial velocity (u, v), we know there is exactly one geodesic c such that $c(0) = (x, y)$ and $\dot{c}(0) = (u, v)$. Now, consider the reflection R in the straight line through (x, y) with direction (u, v). Since R

is an isometry, it must take geodesics to geodesics, so $R(c)$ is a geodesic. By definition of R, $R(c)$ starts at (x, y) with velocity (u, v). By the uniqueness part of Cauchy's theorem, this yields $R(c) = c$, which entails that c is contained in the invariant locus of R, and from there it is not hard to deduce that c is a straight line segment parametrized at constant speed.

Of course, in the case of the Euclidean plane, this is a very dumb way to find the geodesics. The purpose was to show you how knowing the group of isometries allows us to find the geodesics (Klein's viewpoint : geometry is the study of group actions). Now we shall apply this strategy to the sphere.

2.2 Isometries in spherical geometry

It is obvious that all linear isometries of $\mathbb{R}^3$ are isometries of the sphere, in particular reflections in planes through the origin are isometries of $\mathbb{S}^2$. That would be enough to find the geodesics, but just for fun let us find the full group of isometries.

The fundamental observation is that an isometry which fixes a point $x \in \mathbb{S}^2$, and whose derivative fixes some tangent direction at x, must be the identity.

Lemma 2.1. *Let ϕ be an orientation-preserving isometry of $\mathbb{S}^2$, and let us assume*

- *ϕ fixes the north pole $N = (0, 0, 1)$, that is, $\phi(N) = N$*
- *ϕ is differentiable so $\phi'(N)$ is a linear map $\mathbb{R}^2 \to \mathbb{R}^2$, where $\mathbb{R}^2$ is identified with the tangent plane $\{(x, y, 1) : (x, y) \in \mathbb{R}^2\}$ to $\mathbb{S}^2$ at N*
- $\phi'(N).\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}$.

Then ϕ is the identity.

Proof. First observe that an isometry must preserve geodesics : indeed, isometries preserve distances, and geodesics are defined by the property of minimizing distances. Then, recall that through a given point, in a given direction, goes a unique geodesic.

Since ϕ fixes N and ϕ' fixes $(0, 1)$, ϕ must send to itself the geodesic γ through N in the direction $(0, 1)$. This does not mean, *a priori*, that all points of γ are fixed by ϕ. But recall that ϕ preserves distances, and along γ, for every $t \in \mathbb{R}$, there is a unique point at (signed) distance t from N. So actually ϕ fixes γ pointwise, that is, $\forall t \in \mathbb{R}, \phi(\gamma(t)) = \gamma(t)$.

Now, observe that, ϕ being an isometry, and differentiable, it must preserve the norms of tangent vectors (because the length of a path is obtained by integrating the norm of the tangent vector along the path). So $\phi'(N)$ is an isometry of $\mathbb{R}^2$, preserving orientation, and fixing $(0, 1)$, therefore it must be the identity.

But then the argument we used to show γ is fixed pointwise can be applied to any geodesic through N. Now take any $z \in \mathbb{S}^2$. Let γ_z be a geodesic from N to z. Since γ_z is fixed pointwise, $\phi(z) = z$, whence ϕ is the identity map.

□

Remark 2.2. *I've been cheating a little bit in the previous proof. Can you spot where ?*

Lemma 2.3. *All (differentiable) isometries of $\mathbb{S}^2$ are restrictions of linear isometries of $\mathbb{R}^3$.*

Proof. Let ϕ be any (differentiable) isometry. Let P be the reflection in the perpendicular bissector of the segment $[N\phi(N)]$ (if $\phi(N) = N$, just take P to be the identity). Then $P \circ \phi$ fixes N. Now let $v = (P \circ \phi)'.(0,1)$. There exists a rotation R which fixes N and whose derivative takes v to $(0,1)$. Then by Lemma 2.2, $R \circ P \circ \phi$ is the identity of $\mathbb{S}^2$. □

From there, if you know a little bit about linear isometries of $\mathbb{R}^3$, it is not hard to see that all isometries of the sphere are rotations or plane reflections. What will be of interest to us is the fact that linear isometries of $\mathbb{R}^3$ are represented by 3x3 matrices M such that $M^tM = I_3$. This equality should be thought of as $M^tI_3M = I_3$, where the identity matrix I_3 should be viewed as the matrix of the quadratic form $x^2+y^2+z^2$, for reasons soon to be explained. The group of matrices satisfying this identity is usually denoted $O_3(\mathbb{R})$.

2.3 Geodesics in spherical geometry

Now we are equipped to find the geodesics of the sphere (or rather, to prove they are what we think they are). Take a point $x \in \mathbb{S}^2$ and a tangent vector v to $\mathbb{S}^2$ at x, and let us find the geodesic γ of $\mathbb{S}^2$ through x with direction v. Remember that geodesics are solutions of a second-order differential equation (which we can't solve, but we don't care), so a geodesic is uniquely determined by a point and a direction.

Together, the point x and the vector v determine a plane P through the origin of $\mathbb{R}^3$. Let σ_P be the reflection of $\mathbb{R}^3$ in the plane P. Since σ_P is an isometry of $\mathbb{S}^2$, it preserves the geodesics, in particular it sends γ to a geodesic γ'. But σ_P fixes x and v, so γ' is a geodesic through x with direction v, that is, $\gamma = \gamma'$, which to say, σ_P fixes γ.

Beware that we have only proved that γ is fixed as a set. We would like to prove that γ is fixed pointwise. But recall that ϕ preserves distances, and along γ, for every $t \in \mathbb{R}$, there is a unique point at (signed) distance t from x. So actually σ_P fixes γ pointwise, that is, $\forall t \in \mathbb{R}, \sigma_P(\gamma(t)) = \gamma(t)$.

Now observe that the set of points fixed by σ_P is the plane P, so γ must be the intersection of P with $\mathbb{S}^2$. We have proved the

Theorem 2.4. *The geodesics of $\mathbb{S}^2$ are the intersections of $\mathbb{S}^2$ with planes through the origin.*

2.4 Area of spherical triangles

The following theorem, although not directly relevant to the discussion of Euclid's postulates, is too cute to pass, and will be echoed by a similar property in hyperbolic geometry. It is known (to the French, at least) as Girard's theorem. It is about spherical triangles. Just like a Euclidean triangle is given by a triple of non-aligned points, a spherical triangle could be given by three points not contained in a geodesic, but since the geodesic between two antipodal points is not unique, the edges of the triangle may not be well-defined. I'll bypass this difficulty by defining a spherical triangle as a triple of non-aligned, non-pairwise antipodal points.

Theorem 2.5. *Let ABC be a spherical triangle. The area of ABC equals the sum of its interior angles, minus π.*

Sketch of proof (to be drawn) : consider the "lune" with vertex angle α. Its area is proportional to α and equals 4π when $\alpha = 2\pi$, so it must be 2α. Each angle of the triangle defines two such lunes, the six lunes cover the sphere and the intersection of two lunes is exactly the triangle. So the area of the sphere is exactly the sum of the areas of the six lunes, minus twice the area of the triangle and its antipodal, which are counted three times in the sum.

2.5 Spherical geometry vs Euclid's postulates

Now that we know a bit about spherical geometry, let's see how it relates to Euclid's axioms, both the ancient and the modern versions.

Axioms 3 and 3' are certainly verified : the distance between any two points is well-defined.

So are Axioms 4 and 4', since we have seen that the group of isometries is large enough.

Axiom 1' is also verified : between any two points, there is a geodesic (exercise : construct it). The status of Axiom 1 is trickier : the statement "between any two points there is a line" is true, but the uniqueness of the line, which is essential for many properties of Euclidean geometry, for instance the sum of the angles of a triangle being π, is not.

Axioms 2 and 2' are satisfied or not, depending on your viewpoint : if you take the "parametrized curves" viewpoint, then geodesics can be extended to infinity (being run over multiple times). Not if you take the geometric curve viewpoint, however.

As to Euclid's fifth postulate, it is trivially satisfied, because if two lines d_1, d_2 intersect a third line d, no matter what the angles are, d_1 and d_2 intersect on both sides of d. However, the more common formulation about the uniqueness of the parallel through a given point is not satisfied, because parallels simply do not exist in spherical geometry.

Remember, though, that the equivalence between the two formulations of the fifth postulate relies on the uniqueness in the first postulate, so the problem here has to do with the first postulate, not the fifth.

Also note that if we relax the uniqueness requirement in Axiom 1, then Axiom 5 is equivalent to "through a given point there is at most one parallel to a given line", and that is (trivially) true in spherical geometry.

So, while spherical geometry is interesting in itself, it does not tell us much about Euclid's fifth postulate. Still, it hints at the possibility of constructing other geometries, which we shall do in the next section.

3 The Lorentzian model of hyperbolic geometry

3.1 The "sphere" of the Lorentz metric

Part of the reason spherical geometry was a bit too far from Euclidean geometry to be relevant is that the sphere is compact, so in particular geodesics cannot go to infinity. Come to think of it, the compactness of the sphere is caused by the quadratic form $x^2 + y^2 + z^2$ being positive definite. So what if we change the quadratic form, say we consider the quadratic form $q(x, y, z) = x^2 + y^2 - z^2$, and the subset L of $\mathbb{R}^3$ defined by the equation $q(x, y, z) = -1$ (and $z > 0$, to ensure connectedness) ?

This is not as random as it looks : is is a 3-dimensional model of Einstein's Special Relativity, where x, y are the space coordinates, and z is the time coordinate. The surface L is the set of light-like unit vectors, so it's a kind of relativistic sphere. However this geometry was known, as a model of Lobacheveskian (i.e. hyperbolic) geometry, before Lorentz, at least as soon as the 1870's, to Weierstraß and Killing.

Now, wait, you're going to tell me, we're doing geometry here, so we need distances, and distances in Special Relativity are not true distances, since they may be negative ? Here a small miracle happens (the proof is left to you as an exercise) :

Lemma 3.1. *Let (x, y, z) be a point of L, and let (u, v, w) be a tangent vector to L at (x, y, z). Then the quadratic form q, evaluated at (u, v, w), is positive.*

This means that we can define a perfectly good distance on L, as follows. Take a smooth map c from an interval I to L. Then we may compute the length of c, with respect to the quadratic form q, by the following formula :

$$l_q(c) = \int_I \sqrt{q(\dot{c}(t))}dt.$$

Beware this is not the usual length, because the quadratic form q is not $x^2 + y^2 + z^2$!

Now that we can compute lengths, we can compute distances, simply by deciding that the distance between two points is the infimal length of all ways to go from one to the other. We are now going to try and find the geodesics for this new distance. As in the spherical case, we are going to use the symmetries of our space. Here the fact that our distance is defined by a quadratic form is going to help.

3.2 Isometries

Observe that any transformation of $\mathbb{R}^3$ which preserves the quadratic form q, preserves both L, and the distance we have defined on L. Recall that the transformations of $\mathbb{R}^3$ which preserves the quadratic form $x^2 + y^2 + z^2$ are given by the matrices which satisfy $M^t I_3 M = I_3$, where the identity matrix I_3 should be viewed as the matrix of the quadratic form $x^2 + y^2 + z^2$.

Likewise, the transformations which preserve q are the matrices which satisfy $M^t Q M =$

Q, where

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

is the matrix of q.

The group of matrices satisfying this identity is usually denoted $O_{(2,1)}(\mathbb{R})$ where $(2, 1)$ should be thought of as the signature of the quadratic form q (2 "+", 1 "-"). So far it is not obvious that this group contains any matrix, other than the identity ! Let us take a closer look at $O_{(2,1)}(\mathbb{R})$.

What we used most in the sphere case was orthogonal reflections, so maybe we should look for orthogonal reflections in $O_{(2,1)}(\mathbb{R})$. Let P be a plane through the origin in $\mathbb{R}^3$, and assume it intersects the light cone $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = z^2\}$ in a straight line. Assume P satisfies the equation $ax + by + cz = 0$, so $(a, b, c)^\perp = P$ where $^\perp$ means orthogonality with respect to the usual, Euclidean metric.

Exercise 3.2. *Show that the q-orthogonal to P is the straight line P^q generated by the vector $(a, b, -c)$.*

Exercise 3.3. *Show that the reflection σ_P in P, parallel to P^q is a q-isometry, that is, it lies in $O_{(2,1)}(\mathbb{R})$.*

3.3 Geodesics

Now we are equipped to find the geodesics of L with respect to the Lorentz distance. Take a point $x \in L$ and a tangent vector v to L at x, and let us find the geodesic γ of L through x with direction v. Remember that geodesics are uniquely determined by a point and a direction.

Together, the point x and the vector v determine a plane P through the origin of $\mathbb{R}^3$. Let σ_P be the reflection of $\mathbb{R}^3$ in the plane P, parallel to P^q. Since σ_P is an isometry of L, it preserves the geodesics, in particular it sends γ to a geodesic γ'. But σ_P fixes x and v, so γ' is a geodesic through x with direction v, that is, $\gamma = \gamma'$, which to say, σ_P fixes γ.

Beware that we have only proved that γ is fixed as a set. We would like to prove that γ is fixed pointwise. But recall that ϕ preserves distances, and along γ, for every $t \in \mathbb{R}$, there is a unique point at (signed) distance t from x. So actually σ_P fixes γ pointwise, that is, $\forall t \in \mathbb{R}, \sigma_P(\gamma(t)) = \gamma(t)$.

Now observe that the set of points fixed by σ_P is the plane P, so γ must be the intersection of P with L. We have proved the

Theorem 3.4. *The geodesics of L are the intersections of L with planes through the origin.*

3.4 Distances

Let u, v be two points of L. We'd like to compute their hyperbolic distance. So far all we know is the geodesic between them. So what we'll do is find a unit speed parametrization of the geodesic, and then the distance is just the difference of the parameters of u and v.

Let P be the plane through the origin which contains u and v. Then P intersects the line cone $x^2 + y^2 = z^2$ in two straight lines D_u and D_v. The geodesic γ which contains u and v is $P \cap L$. Let u', v' be vectors of D_u, D_v respectively, normalized so that $Q(u', v') = -1/2$, where Q is the bilinear form associated to q, that is,

$$Q'((x, y, z), (x', y', z')) = xx' + yy' - zz'.$$

Exercise 3.5. *Show that the map $t \mapsto e^t u' + e^{-t} v'$ is a unit-speed parametrization of γ.*

Then, if $u = \gamma(t_1), v = \gamma(t_2)$, the distance between u and v is $|t_1 - t_2|$. This is not a very convenient expression of the distance because it assumes we know the parameters of u and v along the geodesic. But :

Exercise 3.6. *Show that $Q(u, v) = -\cosh(t_2 - t_1)$.*

We have proved the following very neat expression of the distance :

Theorem 3.7. *Given any two points u, v in L, their hyperbolic distance is* $\arg\cosh(-Q(u, v))$.

3.5 Hyperbolic geometry vs Euclid's postulates

Exercise 3.8. *Prove that Axiom 1 is satisfied, with uniqueness.*

Remark 3.9. *The difference with spherical geometry is that straight lines through the origin intersect L in at most one point : there are no antipodal points in hyperbolic geometry.*

It is obvious that Axiom 2 is satisfied : all geodesics escape to infinity (beware, though, that *a priori*, Euclidean infinity and hyperbolic infinity could be different).

Likewise for Axiom 3 which basically just says distances exist.

Axiom 4' requires a bit of work. The following exercise proves homogeneity :

Exercise 3.10. *Show that given any point $(x, y, y) \in L$, there exists a plane P such that the reflection σ_P sends (x, y, z) to $(0, 0, 1)$.*

And the next one proves isotropy :

Exercise 3.11. *Show that any rotation with vertical axis is a q-isometry.*

Since we have uniqueness in Axiom 1, Axiom 5 is equivalent to the uniqueness of the parallel through a given point. Now, we have

Exercise 3.12. *Show that given a plane P which intersects L, there exist infinitely many planes P' which intersect L but do not intersect P inside the light cone.*

Hint : choose a plane P, and choose a line l contained in P but not in the light cone...

Thus we have constructed a geometry which obeys all of Euclid's axioms but the fifth, proving that the fifth axiom cannot be a consequence of the first four. So, here we are : we have solved a 2000+ years old problem dating back to Euclid, with a bit of help from Lambert, Bolyai, Lobachevski, Gauß, Riemann, Killing, Weierstraß, Cayley, Klein, Beltrami, and Poincaré (apologies to anyone I've omitted).

4 The Klein disc model

However, maybe Euclid would have considered our solution with suspicion : after all, the original problem is about plane geometry. So, we'd like to transfer our new geometry to the plane. Furthermore it makes drawing (hence, thinking) easier. This is something we are used to doing in spherical geometry, whenever we create maps. The idea is very simple : we take any map $f(L)$ from L to the plane, which we call the projection, and instead of studying L directly, we study $f(L)$. Of course we want the map f to be reasonable, that is, injective, smooth, with smooth inverse. And we should keep in mind that the geometry of $f(L)$ is not its geometry as a subset of the Euclidean plane, but the geometry it inherits from L. You are already familiar with this fact in the case of spherical maps : for instance, you know that Finland is not bigger than Türkiye, nor Greenland than India. So, let's look for a good projection f.

4.1 Hyperbolic geometry and projective geometry

As we have observed, any line through the origin intersects L in at most one point, so we may consider L as a subset of the set of lines in $\mathbb{R}^3$. If you have already met projective geometry, this should sound familiar to you. So, here is our projection : we represent a point x in L by the line $\mathbb{R}x$ through the origin, and we represent the line $\mathbb{R}x$ by its unique intersection point with the plane $\{z = 1\}$. This sends L to the open disk of radius 1 centered at $(0, 0, 1)$. Note that the boundary of the disk is at infinity, in the inherited geometry, so Axiom 2 is still verified !

A particularly nice feature of the Klein model is that its geodesics are straight line segments. Indeed, recall that the geodesics of the Lorentz model L are the intersections of L with planes through the origin. But such a plane intersects the plane $\{z = 1\}$ in a straight line.

4.2 Hyperbolic distance and cross-ratio

In projective geometry the cross ratio of four aligned points A, B, C, D is usually defined as $\frac{AC.BD}{AD.BC}$, and denoted $[A, B, C, D]$. Since it may not be obvious what, e.g., AC means, let us say that A and D are the outermost points, so we may express B and C as convex combinations of A and D : $B = (1-\lambda)A+\lambda D$ and $C = (1-\mu)A+\mu D$ with $\lambda, \mu \in [0, 1]$.

Exercise 4.1. *Show that* $[B, C, A, D] = \frac{\lambda(1-\mu)}{(1-\lambda)\mu}$.

Now we are going to express the hyperbolic distance in terms of the cross-ratio. Let us re-use the notation from Subsection 3.4. Denote p the projection $(x,y,z) \mapsto (x,y)$.

Exercise 4.2. *Prove that the image in the Klein disc of u' and v' are $P' = \frac{p(u')}{\|p(u')\|}$ and $Q' = \frac{p(v')}{\|p(v')\|}$.*

Recall that we had $u = e^{t_1}u' + e^{-t_1}v'$ and $v = e^{t_2}u' + e^{-t_2}v'$.

Exercise 4.3. *Prove that the image in the Klein disc of u and v are $P = \lambda P' + (1-\lambda)Q'$ and $Q = \mu P' + (1-\mu)Q'$ where*

$$\lambda = \frac{e^{t_1}\|p(u')\|}{e^{t_1}\|p(u')\| + e^{-t_1}\|p(v')\|} \text{ and } \mu = \frac{e^{t_2}\|p(u')\|}{e^{t_2}\|p(u')\| + e^{-t_2}\|p(v')\|}.$$

Therefore P', P, Q, Q' play the role of A, B, C, D, in that order.

Exercise 4.4. *Prove that $[Q, P, P', Q'] = e^{2(t_2-t_1)}$.*

We have proved the

Theorem 4.5. *For any two points P, Q in the Klein disc, if P', Q' are the intersections of the straight line PQ with the unit circle, such that P', P, Q, Q' lie in that order on the segment $[P'Q']$, then the hyperbolic distance between P and Q is*

$$d(P,Q) = \frac{1}{2}\left|\log\left[Q, P, P', Q'\right]\right|.$$

A particularly interesting consequence of this theorem is that the distance is expressed purely in terms of the cross-ratio, which is a projective invariant, so any projective transformation which preserves the disc is a hyperbolic isometry. We could use this fact to determine the isometries (but we won't).

5 The Poincaré disc model

In this section and the next I'll finally look at the most popular models of the hyperbolic plane. The reason they are popular, most notably with number theorists, is that they are subsets of the complex plane (unlike the Klein disc which should be seen as a subset of the real projective plane), and as such, they allow the use of the full complex analytic artillery.

5.1 Angles

We have seen that in the Klein model the hyperbolic "plane" is a disk in the Euclidean plane, and its geodesics are the same as the plane's, that is, they are the Euclidean straight lines. However the angles are not the same as the Euclidean angles (fact : if both geodesics and angles were Euclidean, then the geometry would be Euclidean, up to scale).

For some practical purposes such as navigation, it is more useful to have a map which respects angles, than a map in which geodesics are straight : what you measure with your compass are angles, and you want to be able to transfer them to your map without having to do some complicated transformation. Now we are going to see another projection of L to the disk, which respects angles.

What projection do you know from the sphere to the plane which preserves angles ? there are several, but a famous one is the stereographic projection. Here is how it is done : take any point x in the sphere, other than the south pole $S = (0, 0, -1)$, and send it to the intersection of the straight line (Sx) with the plane $\{z = 0\}$.

We do exactly the same here : we send a point x of L to the intersection of the straight line between x and $S = (0, 0, -1)$, with the plane $\{z = 0\}$. Again, this sends L to the unit open disk. The question is, why are angles preserved ? and what becomes of the geodesics ?

Theorem 5.1. *The sterographic projection preserves angles.*

I could give an ugly computational proof using the following two exercises :

Exercise 5.2. *Show that the stereographic projection Ψ sends a point $(x, y, z) \in L$ to $(\frac{x}{z+1}, \frac{y}{z+1}, 0)$.*

To show that angles are preserved, we are going to compute the derivative of Ψ, in true physicist (or engineer) fashion.

Exercise 5.3. *Show that*

$$\Psi(x+h, y+k, z+l) = \Psi(x, y, z) + \begin{pmatrix} \frac{1}{z+1} & 0 & \frac{-x}{(z+1)^2} \\ 0 & \frac{1}{z+1} & \frac{-y}{(z+1)^2} \end{pmatrix} \cdot \begin{pmatrix} h \\ k \\ l \end{pmatrix} + o(h, k, l).$$

But I'll give a nice geometric proof instead, inspired by the spherical case.

Let M be a point in L and let u, v be two tangent vectors to L at M. Let D_u (resp. D_v) be the straight line tangent to L at M directed by the vector u (resp. v). Let P_u (resp. P_v) be the plane in $\mathbb{R}^3$ containing the straight line D_u (resp. D_v) and the point $S = (0, 0, -1)$. Let D'_u (resp. D'_v) be the intersection of P_u (resp. P_v) with the plane $\{z = 0\}$.

Then the image by the stereographic projection of the angle $\theta(u, v)$ between D_u and D_v is the angle between D'_u and D'_v, and we want to show that they are equal. Note that since the tangent plane to $-L$ at S is parallel to $\{z = 0\}$, its intersections D''_u and D''_v with P_u and P_v are parallel to D_u and D_v, respectively, so we want to prove that $\theta(u, v)$ equals the angle $\theta''(u, v)$ between D''_u and D''_v.

Now remember we have computed the isometry group of L, in particular we know there exists an isometry r of L which is also a linear (possibly not orthogonal) reflection r of $\mathbb{R}^3$, and takes M to $N = (0, 0, 1)$. Then r takes D_u and D_v to two tangent lines to L at N, which make an angle $-\theta(u, v)$, since r is an isometry of L. Let $-r$ be the composite of r and $-I_3$, then $-r$ takes D_u and D_v to two tangent lines D'''_u and D'''_v to $-L$ at S, which make an angle $\theta(u, v)$. We want to prove that $D'''_u = D_u$ and $D'''_v = D_v$.

Observe that r is the composite of a reflection and $-I_3$, so it is a linear involution of $\mathbb{R}^3$.

Exercise 5.4. *If j is a linear involution of $\mathbb{R}^3$ which swaps two points a and b, and P is an affine plane which contains a and b, then P is invariant under j.*

Hint : since j is a linear involution, it is diagonalisable with eigenvalues ± 1. Since it swaps a and b, the vector $b - a$ lies in the -1-eigenspace. Then the linear plane underlying P is spanned by $b - a$ and the intersection of P with the 1-eigenspace, so it is invariant under j. Therefore j sends P to a parallel plane, but since $j(P)$ contains the straight line between a and b, $j(P) = P$.

Thus we know that $-r(P_u) = P_u$ and $-r(P_v) = P_v$. Therefore, $D_u''' = -r(D_u)$ (resp. $D_v''' = -r(D_v)$) is a straight line contained in P_u (resp. P_v) and in the plane $\{z = -1\}$, so it is D_u (resp. D_v). This concludes the proof of Theorem 5.1. □

5.2 Geodesics

Theorem 5.5. *The stereographic projection takes geodesics of L which do not contain N to circular arcs orthogonal to the boundary of the disc, and geodesics which contain N to diameters of the disc.*

Proof. The case of geodesics which contain N is obvious since in that case the straight lines from $(x, y, z) \in L$ to the origin, and to S, are co-planar.

For the remaining case, sorry I couldn't come up with a geometric argument, so we'll have to compute.

Since rotations around the vertical axis are isometries of L, it might be smart to use polar coordinates, so we'll denote a point $(x, y, z) \in L$ by (R, θ, z). Observe that $z = \sqrt{1 + R^2}$.

Then the projection to the Klein disc takes the point (x, y, z) to $(\frac{x}{z}, \frac{y}{z}, 0)$, which I'll denote in polar coordinates by (r, θ) with $r = \frac{R}{\sqrt{1+R^2}}$.

We have seen that geodesics in L are taken to straight line segments in the Klein disc. And we know (if not, it's an **exercise**) that the equation of a straight line in polar coordinates is

$$r \cos(\theta - \theta_0) = c$$

where θ_0 and c are constants which tell us the direction of the straight line and how close it comes to the origin.

On the other hand, the projection to the Poincaré disc takes (x, y, z) to $(\frac{x}{z+1}, \frac{y}{z+1}, 0)$, which I'll denote in polar coordinates by (r', θ) with

$$r' = \frac{R}{1 + \sqrt{1 + R^2}} = \frac{r}{1 + \sqrt{1 - r^2}} = \frac{c}{\cos(\theta - \theta_0) + \sqrt{\cos(\theta - \theta_0)^2 - c^2}}$$

which leads to a polar coordinate equation for the geodesic :

$$r'^2 - \frac{2}{c} r' \cos(\theta - \theta_0) + 1 = 0.$$

By going back to x, y-coordinates you recognize the equation of a circle. Furthermore, if you consider the last equality as a degree 2 equation for r', you notice that the roots multiply to 1. This shows that the circle is invariant by the inversion in the unit circle, that is, the map $(r, \theta) \mapsto (\frac{1}{r}, \theta)$. Now here is an exercise for you :

Exercise 5.6. *Show that the circles which are preserved by the inversion in the unit circle are the unit circle itself, and all circles which cross the unit circle orthogonally.*

□

It is interesting to observe that if a geodesic α in the Poincaré disc, and a geodesic β in the Klein disc, are both projections of the same geodesic γ in the hyperboloid, then α and β have the same endpoints in the boundary circle (this is because the difference in the projections becomes negligible when seen from afar, that is, when γ goes to infinity). So, if you have a geodesic in the Poincaré disc and you want the corresponding geodesic in the Klein disc, you just pull it by its endpoints to make it straight.

5.3 Isometries

I'll give a complete computation of the isometry group in the next section, which is about the half-plane model, where the isometry group has a nicer description. For now I just want to observe the following :

Proposition 5.7. *The isometries of the Poincaré disc are homographic maps, that is, of the form* $z \mapsto \frac{az+b}{cz+d}$ *for some* $a, b, c, d \in \mathbb{C}$. *In particular they are complex projective maps, so they preserve the cross-ratio.*

Proof. We have seen that the isometries preserve angles, so they are holomorphic maps. Now I'll take an ace out of my sleeve (that's basically Schwarz' Lemma): holomorphic bijections of the unit disc to itself are homographies. □

5.4 Distances

First let us compute the distance between two points P, Q on a diameter of the disc. We may even assume that they lie on the real axis, so they are images in the disc of two points on the hyperbola $\left\{(x, 0, \sqrt{x^2+1}) : x \in \mathbb{R}\right\}$, so there exist $t_1, t_2 \in \mathbb{R}$ such that P is the image of $(e^{t_1}, 0, e^{-t_1})$ and Q is the image of $(e^{t_2}, 0, e^{-t_2})$. This amounts to taking $u' = \frac{1}{2}(1, 0, 1)$ and $v' = \frac{1}{2}(-1, 0, 1)$ in the notation of Subsection 3.4. So we know that $d(P, Q) = |t_2 - t_1|$.

Now let us compute the coordinates of P and Q in the disc : I'll let you do the

Exercise 5.8. *Setting* $P' = (-1, 0, 0)$ *and* $Q' = (-1, 0, 0)$, *we have* $P = \lambda P' + (1-\lambda)Q'$ *and* $Q = \mu P' + (1-\mu)Q'$ *where*

$$\lambda = \frac{e^{\frac{t_1}{2}}}{e^{\frac{t_1}{2}} + e^{-\frac{t_1}{2}}} \text{ and } \mu = \frac{e^{\frac{t_2}{2}}}{e^{\frac{t_2}{2}} + e^{-\frac{t_2}{2}}}.$$

Then, exactly the same computation as in Exercise 4.4 proves that $[Q, P, P', Q'] = e^{t_2 - t_1}$.

Now, recall that isometries preserve the cross-ratio, and that any two points may be taken to the real axis by an isometry. Thus we have proved the

Theorem 5.9. *For any two points P, Q in the Poincaré disc, if P', Q' are the intersections of the geodesic (PQ) with the unit circle, such that P', P, Q, Q' lie in that order on the segment $[P'Q']$, then the hyperbolic distance between P and Q is*

$$d(P, Q) = \left|\log\left[Q, P, P', Q'\right]\right|.$$

Remark 5.10. *Note that the cross-ratio appears twice, once in a real projective context with a factor of 2, and once in a complex projective context, without the factor of 2. I find that mind-boggling.*

6 The half-plane model

This section is meant to be self-contained or at least independant from the previous sections, so it contains some amount of repetition. We are going to describe a new, apparently unrelated, geometry, and eventually we'll see it is nothing but another model of hyperbolic geometry.

For seemingly no reason at all, we are going to break the beautiful symmetry of the Poincaré disc, and send it to the upper half-plane by the inverse of the map $z \mapsto \frac{z-i}{z+i}$.

Exercise 6.1. *Prove that it does just what I said it does.*

Since this map is holomorphic, it preserves angles, so the angles in our new geometry are the same as in the Poincaré disc, which are the same as in the Lorentz model.

Here is how Poincaré explained the hyperbolic world : imagine a world which is two-dimensional, like a plane, but instead it is only a half-plane, say, $\{x + iy \in \mathbb{C} : y > 0\}$. From now on I'll call this world the **hyperbolic plane**, and denote it $\mathbb{H}^2$. Don't worry about the overlap in terminology, we'll soon prove it is nothing but the hyperbolic plane we have seen in the last sections.

In $\mathbb{H}^2$ the sun is at infinity (that is, its y coordinate is $+\infty$), and the absolute temperature at $x+iy$ is a linear function of y, so when you approach the lower boundary of $\mathbb{H}^2$, you freeze, and you move slower. I'll make the *ad hoc* hypothesis that the slowing down is also a linear function of y, so if you have height $y/2$, it is twice as hard to go at speed v, as it is when you have height y. Of course the inhabitants of $\mathbb{H}^2$ don't realize they're going slower, to them it's more like distances expand when you go towards $y = 0$. In other words, they measure velocities by the following formula : if your position is $x + iy$, and your velocity, to an outside observer, is the vector (u, v), so your speed is $\sqrt{u^2 + v^2}$, in $\mathbb{H}^2$, your speed is $\frac{1}{y}\sqrt{u^2 + v^2}$.

Once we can measure speed, we can also measure the length of a path : if

$$\begin{array}{rcl} c : [0,1] & \longrightarrow & \mathbb{H}^2 \\ t & \longmapsto & x(t) + iy(t) \end{array}$$

is the trajectory of a traveller in $\mathbb{H}^2$, then, as Euclidean observers, we would measure its length by integrating the speed : the Euclidean length of the path c is

$$\int_0^1 \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt.$$

In $\mathbb{H}^2$ the length of c is measured by the same method, just remembering that speed is measured according to the hyperbolic rule, so the hyperbolic length is

$$\int_0^1 \frac{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}}{y(t)} dt.$$

Once we can measure length, we can also measure distance, just because the distance between to points is the smallest possible length of a path between them. In the next sections we shall ask ourselves the question "what is the shortest hyperbolic path from point A to point B ?". Intuitively we feel that it may be interesting to go north a little to gain speed, even if we have to run a slightly longer distance in the process, but how far north exactly ?

Let us try the differential equation approach that we used in the plane : let $(x(t), y(t))$ be a shortest path, parametrized by time, then its acceleration is zero, only we need to remember that its velocity is not $(\dot{x}(t), \dot{y}(t))$, but $\frac{1}{y(t)}(\dot{x}(t), \dot{y}(t))$, so the acceleration is something like

$$\frac{1}{y(t)^2}\left(\dot{x}(t)y(t) - x(t)\dot{y}(t), \ddot{y}(t)y(t) - \dot{y}(t)^2\right).$$

Again, I can't even solve the differential equation $\ddot{y}(t)y(t) - \dot{y}(t)^2 = 0$. So let us try the isometry strategy instead. For this we need to find the isometries of $\mathbb{H}^2$.

6.1 Isometries

The fundamental observation is that, although the hyperbolic world is very different from the Euclidean world, they share an important feature : angles are the same in both worlds. Remember how you measure the Euclidean angle between two vectors (u, v) and (u', v') ? You take the scalar product $uu' + vv'$, divide it by the product of the length, and you get the cosine of the angle, which gives you the angle up to sign ; then all you need to know is whether (u, v) is to the right or to the left of (u', v'), and you have the sign of the angle.

Hyperbolic angles are measured the same way, only the scalar product of (u, v) and (u', v') depends on the point of $\mathbb{H}^2$ you are standing at : if you are at $x + iy$, then the scalar product of (u, v) and (u', v') is $\frac{uu'+vv'}{y^2}$ (this is the only way to get a quadratic form which gives the correct lengths of (u, v) and (u', v'), knowing that the length of a vector is the square root of its scalar product with itself). But then the ratio of the scalar product to the product of lengths is the same as in the Euclidean case.

The bearing of this on the problem of finding isometries is that now we know that a hyperbolic isometry must preserve Euclidean angles (assuming it preserves orientation, that is, left and right). From now on I shall speak of angles to mean either Euclidean or hyperbolic angles.

Question : what transformations do you know that preserve angles ? isometries for sure, but apart from horizontal translations, no Euclidean isometry preserves the upper half-plane, so they are not good candidates for being hyperbolic isometries.

Those of you who have studied complex analysis (if you haven't, just bear with me) know that holomorphic transformations preserve both angles and orientation. Apart from preserving angles, an obvious condition an isometry must satisfy is injectivity. What injective holomorphic transformations do you know ? bingo, the homographic transformations. Do they preserve the upper half-plane ? Yes, they do :

Exercise 6.2. *Let a, b, c, d be real numbers such that $ad - bc \neq 0$, then the map*

$$\begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & \frac{az+b}{cz+d} \end{array}$$

leaves the upper half-plane invariant.

Now that homographies are good candidates for isometries, let us check that they actually are isometries :

Exercise 6.3. *The map*

$$\begin{array}{ccc} \mathbb{C} & \longrightarrow & \mathbb{C} \\ z & \longmapsto & \frac{az+b}{cz+d} \end{array}$$

is an isometry of $\mathbb{H}^2$.

So we have found a whole family of isometries of $\mathbb{H}^2$, indexed by 4-tuples of real numbers. Observe that if we multiply a, b, c, and d by the same number $\lambda \neq 0$, we get the same map, so we actually have 3 parameters rather than 4, and we might as well assume $ad - bc = 1$. The fact that $ad - bc$ looks like a determinant is not a coincidence, our family of isometries looks a lot like a matrix group. Let us denote by $\mathcal{I}$ the set of all orientation-preserving isometries of $\mathbb{H}^2$.

Proposition 6.4. *The map*

$$\begin{array}{ccc} \Psi : \mathrm{SL}_2(\mathbb{R}) & \longrightarrow & \mathcal{I} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \longmapsto & \left(\begin{array}{ccc} \mathbb{H}^2 & \to & \mathbb{H}^2 \\ z & \mapsto & \frac{az+b}{cz+d} \end{array} \right) \end{array}$$

is a group morphism, almost injective, that is, its kernel is $\{\pm I_2\}$.

Now the question is, is the morphism Ψ onto ? in other words, have we found all the orientation-preserving isometries ? To answer this question, we first need a

Lemma 6.5. *Let ϕ be an orientation-preserving isometry of $\mathbb{H}^2$, and let us assume*

- *ϕ fixes $i \in \mathbb{C}$, that is, $\phi(i) = i$*

- *ϕ is differentiable (not necessarily holomorphic, differentiable as a map $\mathbb{R}^2 \to \mathbb{R}^2$, so $\phi'(z)$ is a linear map $\mathbb{R}^2 \to \mathbb{R}^2$)*
- $\phi'(i).\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Then ϕ is the identity.

Proof. First observe that an isometry must preserve geodesics : indeed, isometries preserve distances, and geodesics are defined by the property of minimizing distances. Then, recall that through a given point, in a given direction, goes a unique geodesic.

Since ϕ fixes i and ϕ' fixes $(0,1)$, ϕ must send to itself the geodesic γ through i in the direction $(0,1)$. This does not mean, *a priori*, that all points of γ are fixed by ϕ. But recall that ϕ preserves distances, and along γ, for every $t \in \mathbb{R}$, there is a unique point at (signed) distance t from i. So actually ϕ fixes γ pointwise, that is, $\forall t \in \mathbb{R}, \phi(\gamma(t)) = \gamma(t)$.

Now, observe that, ϕ being an isometry, and differentiable, it must preserve the norms of tangent vectors (because the length of a path is obtained by integrating the norm of the tangent vector along the path). So $\phi'(i)$ is an isometry of $\mathbb{R}^2$, preserving orientation, and fixing $(0,1)$, therefore it must be the identity.

But then the argument we used to show γ is fixed pointwise can be applied to any geodesic through i. Now take any $z \in \mathbb{H}^2$. Let γ_z be a geodesic from i to z. Since γ_z is fixed pointwise, $\phi(z) = z$, whence ϕ is the identity map. $\square$

Now we can answer our question : we have actually found all orientation-preserving isometries of $\mathbb{H}^2$.

Proposition 6.6. *The group morphism Ψ of Proposition 6.4 is onto.*

Proof. Let ϕ be any orientation preserving isometry of $\mathbb{H}^2$. Again I'll cheat a little bit by assuming that ϕ is differentiable ; see below how to deal with the differentiability assumption.

Let $\phi(i) = a + ib$, then, denoting $\phi_1(z) = az + b$, we have $\phi_1^{-1}(\phi(i)) = i$, so $\phi_1^{-1} \circ \phi$ is an orientation-preserving, differentiable isometry of $\mathbb{H}^2$, which fixes i. Then the vector $(\phi_1^{-1} \circ \phi)'.(0,1)$ has norm 1, because $\phi_1^{-1} \circ \phi$ is an isometry, so it may be written $(\sin\theta, \cos\theta) = \exp(\frac{\pi}{2} - \theta)$. Let

$$\begin{array}{rcl} \phi_2 : \mathbb{H}^2 & \longrightarrow & \mathbb{H}^2 \\ z & \longmapsto & \frac{z\cos\frac{\theta}{2} + \sin\frac{\theta}{2}}{z\sin\frac{\theta}{2} - \cos\frac{\theta}{2}}, \end{array}$$

then (exercise !) $\phi_2(i) = i$, and $\phi_2'.(\sin\theta, \cos\theta) = (0,1)$. Therefore $\phi_2 \circ \phi_1^{-1} \circ \phi$ is an isometry of $\mathbb{H}^2$, which fixes i, and whose derivative fixes $(0,1)$, so by Lemma 6.5, it is the identity, whence $\phi = \phi_2^{-1} \circ \phi_1$, which is in the image of $\mathrm{SL}_2(\mathbb{R})$ under Ψ. $\square$

If you are worried about the differentiability assumption, here is how to get rid of it, provided you know a bit of advanced analysis : observe that isometries are Lipschitz, and Lipschitz maps are differentiable almost everywhere by Rademacher's theorem.

Now that we have found the group of all isometries, it will be easy to find the geodesics (remember Klein's philosophy : look for the group first !).

To finish the job of finding all isometries, not just the orientation-preserving ones, just observe that the reflection in the y axis, that is, $z \mapsto -\bar{z}$, is an orientation-reversing isometry of $\mathbb{H}^2$.

Remark 6.7. *Once we prove that the hyperbolic half-plane is isometric with the Poincaré disc, which itself is isometric with the hyperboloid model of hyperbolic space, we'll have obtained the unexpected result that the matrix group $O_{2,1}(\mathbb{R})$, which consists of 3×3 matrices, is isomorphic to the group $\mathrm{PSL}_2(\mathbb{R})$ which is the quotient of $\mathrm{SL}_2(\mathbb{R})$ by $\pm I_2$. This is not something you would guess spontaneously, and even once you know it, it is not easy to prove directly (ok, now prove me wrong) !*

6.2 Geodesics

Lemma 6.8. *Vertical straight lines are geodesics.*

Proof. Reflections in vertical straight lines are (orientation-reversing) isometries. So the reflection ϕ in the vertical axis Oy sends the geodesic γ through i in the direction $(0,1)$ (which is unique because geodesics are solutions to a second-order differential equation) to the geodesic through $\phi(i) = i$, in the direction $\phi'(i).(0,1) = (0,1)$, that is, to γ itself. This means that γ is contained in the invariant locus of ϕ, which is the vertical axis Oy.

With a bit more work which I'll leave to you as an exercise, we conclude that γ is actually Oy itself, which is therefore geodesic. Since horizontal translations are isometries, it follows that all vertical straight lines are geodesic. $\square$

Proposition 6.9. *The geodesics of $\mathbb{H}^2$ are the vertical straight lines and the half-circles with center on the real axis Ox.*

Proof. Let $z = x + iy$ be some complex number, and $(u,v) \in \mathbb{R}^2$ be a vector, we are looking for the geodesic γ through z in the direction (u,v). Recall the proof of Proposition 6.6 : using the same method we construct an element, let us call it ϕ, of $\mathrm{SL}_2(\mathbb{R})$ which sends i to z and $(0,1)$ to (u,v). Since ϕ is an isometry, it sends Oy to γ. Now let ϕ, as an element of $\mathrm{SL}_2(\mathbb{R})$, be represented by the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Observe that the map $z \mapsto (az+b)/cz+d$ may be decomposed as follows (assuming $c \neq 0$, I'll leave the case $c = 0$ to you as an exercise) :

1. first, multiply z by c (homothety) $z \mapsto cz$
2. then, add d (horizontal translation) $z \mapsto cz + d$
3. apply the inversion $z \mapsto 1/z$, so we have done $z \mapsto 1/(cz+d)$

4. multiply by $-ad/c + b$, so we have done $z \mapsto \frac{-ad/c+b}{cz+d}$
5. add a/c, which yields

$$z \mapsto \frac{-ad/c + b}{cz + d} + \frac{a}{c} = \frac{a(cz + d)/c - ad/c + b}{cz + d} = \frac{az + b}{cz + d}.$$

Now, observe that homotheties and horizontal translations preserve both vertical straight lines and half-circles with real center, and the inversion either preserves or exchanges them, so the geodesic γ must be either a vertical straight line or a half-circle with real center.

Conversely, since through any point, in any non-vertical direction, there exists a unique half-circle with real center (exercise for you : construct it like real men, with ungraduated ruler and compass), we know that all vertical straight lines, and all half-circles with real center, are geodesics. $\square$

6.3 Distances

First let us compute the distance between i and λi, for $\lambda > 0$. We know that the vertical line through i is a geodesic, so the distance between i and λi is the length of the path $[1, \lambda] \to \mathbb{H}^2$, $t \mapsto it$, which is the absolute value of the integral

$$\int_1^\lambda \frac{dt}{t} = \log \lambda.$$

Now observe that

- $|\log \lambda| = |\log [i, \lambda i, 0, \infty]\,|$,
- any two points P and Q may be taken to i and λi, for some $\lambda > 0$, by an isometry of $\mathbb{H}^2$,
- isometries of $\mathbb{H}^2$ are complex projective maps, so they preserve the cross-ratio.

Therefore, for any two points $P, Q \in \mathbb{H}^2$, we have $d(P, Q) = |\log [P, Q, P', Q'],$ where P' and Q' are the endpoints of the geodesic between P and Q. Observe this is the same formula as in the Poincaré disc.

6.4 Isometry with the Poincaré disk

So, let Ψ be the map from H^2 to the Poincaré disk which takes z to $\frac{z-i}{z+i}$. We want to show that Ψ is an isometry, that is, for any two points x, y in $\mathbb{H}^2$, $d(x, y) = d(\Psi)$. But Ψ is an homography, so it preserves the cross-ratio, and we have seen that distances are expressed as the same function of the cross-ratio, both in the Poincaré disc and in the upper half-plane.

6.5 Area of hyperbolic triangles

So far I have carefully avoided mentioning any hyperbolic version of Girard's theorem. The reason is that it is easier in the half-plane.

Theorem 6.10 (Gauß-Bonnet formula)**.** *The (hyperbolic) area of a hyperbolic triangle with interior angles are α, β, γ is $\pi - \alpha - \beta - \gamma$.*

Notice the symmetry with Girard's theorem.

Proof. Let us begin with the seemingly most exotic case : when all three vertices are at infinity. Note that since geodesics are perpendicular to the boundary of $\mathbb{H}^2$, all three angles are zero.

First observe that up to isometry, there is only one ideal triangle, that with vertices $0, 1, \infty$. Indeed, let us assume that the vertices are $a, b, c \in \mathbb{R} \cup \{\infty\}$. Then the isometry $z \mapsto \frac{b-c}{b-a}\frac{z-a}{z-c}$ sends a to 0, c to ∞, and b to 1. So we just have to compute the area of the $(0, 1, \infty)$ triangle. But since the computation is a bit prettier, I'm going to compute that of the $(-1, 1, \infty)$ triangle, which is

$$\int_{x=-1}^{x=1} \int_{y=\sqrt{1-x^2}}^{+\infty} \frac{dxdy}{y^2} = \int_{x=-1}^{x=1} \frac{dx}{\sqrt{1-x^2}} = \int_{-\pi}^{0} d\theta = \pi$$

where

- the reason for the y^2 in the denominator is that hyperbolic areas are not euclidean areas : the area of a small square of side ϵ centered at (x, y) is roughly $\frac{\epsilon^2}{y^2}$
- we have used the change of variable $x = \cos\theta$, $\theta \in [-\pi, 0]$.

So, the area of ideal triangles is π, and since all angles are zero, so far our formula is valid.

Next let us do the case of a triangle with two vertices at infinity. Again, up to isometry, we may assume our triangle has vertices $1, \infty, e^{i\theta}$, with $\theta \in [0, \pi]$. The same computation as before yields

$$\int_{x=\cos(\pi-\theta)}^{x=1} \frac{dx}{\sqrt{1-x^2}} = \int_{\theta}^{\pi} d\theta = \pi - \theta.$$

Now, let us look at a triangle with only one vertex at infinity : $e^{i\alpha}, \infty, e^{i\beta}$ and we use the previous computation with the triangles $1, \infty, e^{i\beta}$ and $1, \infty, e^{i\alpha}$.

Finally, let us consider a triangle without a vertex at infinity.

□

6.6 Algebraic viewpoint

So far we have seen geometric properties of the hyperbolic plane, but how about algebraic properties ? After all, you can add points of the Euclidean plane, and even mutiply them, as complex numbers.

The first idea that comes to mind is that since the hyperbolic plane is a subset of the complex plane, maybe we could add or multiply points of $\mathbb{H}^2$, as complex numbers ? However this doesn't work very well, because the difference of two points of $\mathbb{H}^2$ may not be in $\mathbb{H}^2$, and likewise, the product of two points of $\mathbb{H}^2$ may not be in $\mathbb{H}^2$.

Ok, next idea : we have seen that the matrix group $\mathrm{SL}_2(\mathbb{R})$ plays a part in this story, and matrices can be added and multiplied, can't they ? What if we could just associate a matrix to every point of $\mathbb{H}^2$, and add or multiply hyperbolic points as matrices ?

For instance, assume we decide (why not ?) to associate the point i with the identity matrix. We know that any hyperbolic point z is the image of i under some isometry of $\mathbb{H}^2$, which may be viewed as a matrix $M \in \mathrm{SL}_2(\mathbb{R})$, so why not identify the point z with the matrix M ?

The problem is that the matrix M is not uniquely defined : for instance, there are many matrices, other than the identity, that send z to itself, as shown by the

Exercise 6.11. *1. Show that the matrices of* $\mathrm{SL}_2(\mathbb{R})$ *which send i to itself are precisely the rotation matrices, that is, $M_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ for some $\theta \in \mathbb{R}$.*

2. Show that the matrices of $\mathrm{SL}_2(\mathbb{R})$ *which send z to itself are precisely the matrices $M_z.M_\theta.M_z^{-1}$, for $\theta \in \mathbb{R}$, M_z being any matrix which send i to z.*

So, we would need, for each $z \in \mathbb{H}^2$, to choose a matrix M_z such that $M_z.i = z$. Note that the set $\mathcal{M}$ of all matrices M_z, for $z \in \mathbb{H}^2$, would need to satisfy a special property : if z, z' are two hyperbolic points, and we can add and multiply hyperbolic points, then we want $M_{z+z'} = M_z + M_{z'}$, and $M_{z.z'} = M_z.M_{z'}$, that is, the set $\mathcal{M}$ must be stable under both addition and multiplication. In Subsection 7.2 we shall see how to do that.

An alternative idea : what if we take a leap into abstraction and, instead of choosing a matrix M_z among the many that send i to z, we represent z as the set of all matrices that send i to z ? Then we would need to multiply, not just matrices, but sets of matrices (didn't you hear me say we would take a leap into abstraction ?).

Let $\mathcal{M}_z$ be the set of all matrices that send i to z. Then, for the multiplication of $\mathcal{M}_z$ and $\mathcal{M}_{z'}$ to make sense, we want that for any $M_z \in \mathcal{M}_z$ and $M_{z'} \in \mathcal{M}_{z'}$, the product matrix $M_z.M_{z'}$ lies in $\mathcal{M}_{z.z'}$. In particular, this entails that for any $M_z \in \mathcal{M}_z$ and $M_{z'} \in \mathcal{M}_{z'}$, the hyperbolic point $M_z.M_{z'}.i$ is the same (i.e. it does not depend on the choice of $M_z \in \mathcal{M}_z$ and $M_{z'} \in \mathcal{M}_{z'}$).

However, it is easy to see that this cannot work : for instance, taking $z = -1 + i$, $z' = 1 + i$, we could take $M_z = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ and $M_{z'} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, or we could take $M_z = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$ and $M_{z'} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$, and I'll leave it for you to check that $M_z.M_{z'}.i$ is not the same in both cases.

For those of you who know group theory, this is the quotient group construction, and the reason it doesn't work here is that the group of rotation matrices is not normal in $\mathrm{SL}_2(\mathbb{R})$. The quotient construction is a powerful tool, and not just in group theory, even though it doesn't work in that particular case. It is interesting to note that there is no (sensible) way that the sphere is a group, but that's a story for another time.

6.7 Geodesic flow

6.8 Circles and horocycles

6.9 Horocycle flow

7 Questions of uniqueness, and richness of hyperbolic geometry, and why

7.1 Quotients and tilings

' recollements (4m et 4m+2 gones) groupes

Lemma 7.1. *Given $n \in \mathbb{N}^*$, for every $\theta \in \left[0, \frac{n-2}{n}\pi\right]$, there exists a regular hyperbolic n-gon with interior angle θ.*

Theorem 7.2. *Given $n \in \mathbb{N}$, $n \geq 3$, the hyperbolic plane may be tiled by regular n-gons with angle $\frac{2\pi}{k}$, for any $k \in \mathbb{N} \cup \{\infty\}$, $k > n$.*

This is a particular case of a more general theorem by Poincaré.

7.2 Affine functions

In middle school you learnt about affine functions : $f_{a,b} : x \mapsto ax + b$, and probably found them very boring. I'll grant you that each affine function by itself is boring, but how about the space of all affine functions ? Let us denote $\mathcal{A}$ the set of all affine functions. We'll be particularly interested in a smaller set of functions : the set $\mathcal{A}^+$ of affine increasing functions, that is, we require that $a > 0$. Incidentally, the set $\mathcal{A}^+$ enjoys some interesting algebraic properties :

Exercise 7.3. *Show that $\mathcal{A}^+$ is actually a group under composition, that is, the composite of two non-decreasing affine maps is a non-decreasing affine map.*

Exercise 7.4. *Show that the set $\mathcal{AM}$ of matrices of the form $M_{a,b} = \begin{pmatrix} \sqrt{a} & \frac{b}{\sqrt{a}} \\ 0 & \frac{1}{\sqrt{a}} \end{pmatrix}$ with $a > 0$ and $b \in \mathbb{R}$ is a group under matrix multiplication. Note that is is actually a subgroup of $\mathrm{SL}_2(\mathbb{R})$.*

Exercise 7.5. *Show that the map $\Psi : f_{a,b} \mapsto M_{a,b}$ is a group isomorphism, that is, it is a bijection and for any $(a, b), (a', b')$, $\Psi(f_{a,b} \circ f'_a, b') = M_{a,b}.M_{a',b'}$.*

Ok, but how about the geometric properties ? first, note that we have two parameters, $a > 0$ and $b \in \mathbb{R}$, so if we set $z = b + ia$, we have a natural bijection $\mathcal{A}^+ \longrightarrow \mathbb{H}^2$. But a bijection is not a geometric description, it's just a description of $\mathcal{A}^+$ as a set. And as a set, the hyperbolic plane is not particularly interesting, in fact as a set it is just the same as $\mathbb{R}^2$, or $\mathbb{R}$, or $\mathbb{R}^+ \setminus \mathbb{Q}$, and plenty more.

If we want a geometric description, we need to be able to measure distances on $\mathcal{A}^+$. Well, we have two parameters a and b, so why don't we set the distance between $f_{a,b}$ and $f_{a',b'}$ to be $\sqrt{(a-a')^2 + (b-b')^2}$ or some other similar formula ?

The reason we don't do that is, the set $\mathcal{A}$ does not come naked, it comes with a group structure (the operation of composition), so we want our distance to be compatible with the composition operation. To take an analogy, the reason $\sqrt{(x-x')^2 + (y-y')^2}$ is a good way to measure the distance between points (x, y) and (x', y') in the plane is that it is invariant under the operation of vector addition in $\mathbb{R}^2$:

$$d((x, y), (x', y')) = d((x, y) + (x'', y''), (x', y') + (x'', y''))$$

for any three points $(x, y), (x', y'), (x'', y'')$. So, we would like a distance function d on $\mathcal{A}$ such that for any $(a, b), (a', b'), (a'', b'')$,

$$d(f_{a,b}, f_{a',b'}) = d(f_{a'',b''} \circ f_{a,b}, f_{a'',b''} \circ f_{a',b'}) \tag{1}$$

Note that order matters because map composition is not commutative.

Does such a thing even exist ?

We are going to use Exercices 7.3, 7.4, 7.5, to prove that the hyperbolic metric is indeed such a metric.

7.3 Elliptic curves

An elliptic curve is the set of solutions in $\mathbb{C}^2$ of the equation $y^2 = x^3 + ax + b$, where a, b are complex parameters. Now let me ask the question "what is the set of all elliptic curves ?" Easy, you say, there are two parameters, so it must be $\mathbb{C}^2$? Wait, how do you know that two different equations won't give you the same curve ? and what does it even mean for two curves to be the same ?

First, there is a kind of dumb transformation where you replace a with $\lambda^2 a$ and b with $\lambda^3 b$, for some complex number λ, and the curve you get is just the image of the original curve under the affine map $(x, y) \mapsto (\lambda x, \lambda^{3/2} y)$ so it is not really different (in technical terms, there is a holomorphic map from the first to the second). Thus we actually have only one complex parameter, for instance we may assume $b = 1$. But could there be more subtle, hidden transformations which do not actually change the curve ?

Now, by some black magic which I won't even attempt to describe, we know that an elliptic curve is actually a torus, which may be described as the quotient of $\mathbb{C}$ by a lattice $Zu \oplus \mathbb{Z}v$, where u and v are non-collinear vectors.

7.4 Ellipses

Now we are talking about actual, good old ellipses in the plane.

7.5 Why hyperbolic geometry is ubiquitous

There is a good reason why hyperbolic geometry keeps popping up in places where you least expect it, such as affine functions, or elliptic curves.

Part II
Higher-dimensional hyperbolic geometry

8 Three-dimensional hyperbolic geometry

9 Complex hyperbolic geometry

Acknowledgements I thank Ivan Babenko, François Fillastre, and Gaël Meignan for their useful input.

References

[1] Carroll, Lewis *Through the looking-glass, and what Alice found there*

[2] For Euclid's Elements, a useful ressource is the website http://aleph0.clarku.edu/ djoyce/java/elements/elements.html

[3] Ghys, Etienne *Poincaré et son disque* available online at http://perso.ens-lyon.fr/ghys/articles/disque-poincare.pdf

[4] Milnor, John *Hyperbolic geometry: the first 150 years* Bull. Amer. Math. Soc. (N.S.) 6 (1982), no. 1, 9–24.

Adress:
Daniel Massart, Institut Montpelliérain Alexander Grothendieck, Université de Montpellier, France
email: daniel.massart@umontpellier.fr