

Mathematics and Algorithms for the Hyperbolic Tree Visualization

Benjamin Bergé¹, Christophe Bouthier²

Abstract

In this paper, we give mathematical foundations and algorithms necessary to create the hyperbolic tree visualization (*hypertree*). This visualization uses the Poincaré disk as a model of the (2 dimensional)-hyperbolic space in order to obtain a *focus+context* representation of a tree. We also describe how classical transformations of the disk enable implementation of dynamical interactions.

Keywords : Hyperbolic geometry, conformal model, hyperbolic tree, visualization of hierarchical data

1 Introduction

Large sets of structured data appear in many areas like climate studies, human genome, taxonomy, exploration of a hard drive file system, Web searches ... Navigation into such data sets, either to understand their inner structure or to look for a specific item, is a part of the research field of information visualization. We refer to Herman, Melançon and Marshall [4] for a good survey on the subject. Following Shneiderman's motto [13]: "Overview first, zoom and filter, then detail on demand," a good way to help the navigation is to provide a detailed view on a particular part of the data, while still keeping a global view of the structure of the set. A visualization with such a property is called a *focus+context* visualization.

The hyperbolic tree visualization is a *focus+context* visualization, made for the navigation in large tree-structured data. It uses hyperbolic geometry to provide its detailed and global views.

The parallelism concept is based on the 5th Euclidean postulate ("By a point outside of a given line, goes one and only one line parallel to the given line [10]"). Bolyai [2] and Lobachevsky [8], while trying to demonstrate it by supposing that one can draw an infinity of parallels to a given line through a point outside of this line, separately discovered the hyperbolic geometry (consistent in the sense of the later *Erlangen Program* [3]). By supposing that no such parallel can be drawn, Riemann has unveiled the elliptic geometry [11]. An intuitive representation of the elliptic geometry in the 2-dimensional Euclidian space is given by the unit sphere. At the end of the 19th century, Poincaré suggested a representation of the hyperbolic space (of dimension 2) in the Euclidian space, using the unit disk of the plane [3, 10].

Applications of the non-Euclidean geometries have grown up, particularly in domains like information visualization, thanks to computer advance. Looking at Escher's drawings (see [3, page 356]),

¹Institut de Mathématiques, École Polytechnique Fédérale de Lausanne,
CH-1015 Lausanne (Switzerland)
e-mail: benjamin.berge@epfl.ch

²Laboratoire Lorrain de Recherche en Informatique et ses Applications, UMR 7503, B.P. 239
F-54506 Vandœuvre-lès-Nancy Cedex (France)
e-mail: bouthier@loria.fr

Lamping and Rao [6, 7] had the idea to create the hyperbolic tree visualization. Unfortunately, their works, and later Munzner's [9], do not contain enough details to understand the hyperbolic layout of the tree in order to be able to reproduce the visualization. For instance, Herman and *al.* [4] complain that "hyperbolic views are also surrounded by a sort of mystery (...). Unfortunately, none of the papers are didactic enough to reveal the mystery."

We adopt this point of view to write this paper. In the next part, we put the stress on the didactic while describing mathematical background of non-Euclidean geometries (in particular the hyperbolic one), and transformations and formulas of Poincaré's model. In the third section, we focus on the algorithms used to layout, draw and interact with the tree. We always link up the algorithms with the formulas of the mathematical section. This leads to substantial improvements, concerning for example geodesic drawing and display speed.

2 Mathematical foundations

The purpose of this section is to describe as precisely as possible the Poincaré model, a 2-dimensional representation of the hyperbolic space.

In our mind, the best way to understand the hyperbolic geometry is to proceed by duality with the (spherical) elliptic geometry which is more intuitive than the hyperbolic one. This constitutes the first step towards understanding non-Euclidean geometries. First, we recall some definitions in a 3-dimensional Euclidean context. Then we give some generalities on the 2-dimensional hyperbolic space (for which an intuitive idea is given by a subspace of \mathbb{R}^3) in order to build the Poincaré model using the so-called *stereographic* projection. Essential properties ensue.

2.1 Generalities on the Euclidean Geometry

2.1.1 Metric Spaces

This paragraph can be skipped if the 3D-Euclidean space is well known. However, the fundamental ideas of the Euclidean geometry are easily understood in this case and lead to complete and general definitions for general geometries.

Definition 2.1 A metric on a non-empty set E is a function $d : E \times E \rightarrow \mathbb{R}$ satisfying the following four properties, for all $x, y, z \in E$:

- (i) $d(x, y) \geq 0$ (positivity);
- (ii) $d(x, y) = 0$ if and only if $x = y$ (nondegeneracy);
- (iii) $d(x, y) = d(y, x)$ (symmetry);
- (iv) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

A metric space (E, d) is a set together with a metric d .

Example 2.1 The real line $(\mathbb{R}, |\cdot|)$ is a metric space.

In the view of the applications, we focus on the 3D-Euclidean space. Let us equip \mathbb{R}^3 with its natural structure of a linear space. Let $\langle \cdot, \cdot \rangle_E$ be the inner product defined on $\mathbb{R}^3 \times \mathbb{R}^3$ by

$$\langle x, y \rangle_E = \sum_{i=1}^3 x_i y_i$$

for $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ and $y = (y_1, y_2, y_3) \in \mathbb{R}^3$. The associated (Euclidean) norm is $|x|_E = \sqrt{\langle x, x \rangle_E}$. The map d_E defined on $\mathbb{R}^3 \times \mathbb{R}^3$ by $d_E(x, y) = |x - y|_E$ is a metric and (E, d_E) becomes a metric space.

2.1.2 Isometries

We now put the stress on an important class of transformations between two metric spaces, namely isometries. Let (E, d) and (F, δ) be two metric spaces, and ϕ a mapping from E to F .

Definition 2.2 (Isometry) *We say that ϕ is an isometry if ϕ is a distance preserving bijection, in the sense that ϕ is bijective and, for all $x, y \in E$,*

$$\delta(\phi(x), \phi(y)) = d(x, y). \quad (2.1)$$

The set of all isometries defined on a metric space (E, d) into itself is a group for the composition of mappings. We denote it by $I(E)$. Of course, this notion is closely related to the metric defined on the set E . It enables us to investigate the notion of *shortest path*.

2.1.3 Geodesics

We now introduce the notion of *geodesic*, or *shortest path* (or distance) between two different points in a metric space E . We choose to do it as rigourously as possible without dealing with difficulties that exceed the scope of this paper. However we need a succession of definitions, leading to various interesting results.

Definition 2.3 *A curve in a metric space (E, d) is a continuous function $\gamma : [a, b] \rightarrow E$ where $[a, b]$ is a non-empty interval of \mathbb{R} .*

$\gamma(a)$ is the initial point and $\gamma(b)$ the end point (final point). We say that γ is a curve joining $\gamma(a)$ to $\gamma(b)$.

Definition 2.4 *A geodesic arc in a metric space (E, d) is a distance preserving function $\gamma : [a, b] \rightarrow E$ (cf. Definition 2.2). Of course, this is an injective continuous function (and in particular a curve).*

Definition 2.5 *A geodesic segment joining a point x to a point y in a metric space (E, d) is the image of a geodesic arc $\gamma : [a, b] \rightarrow E$ whose initial point is x and final point is y . In other words, $\gamma(a) = x$, $\gamma(b) = y$ and $d(\gamma(t_1), \gamma(t_2)) = |t_1 - t_2|$ for all $t_1, t_2 \in [a, b]$.*

In order to manipulate these definitions, we have the following theorem.

Theorem 2.1 ([10]) *The geodesic segments of the Euclidean space \mathbb{R}^3 are the straight segments.*

Definition 2.6 *A geodesic line in a metric space (E, d) is a locally distance preserving function $\lambda : \mathbb{R} \rightarrow E$. In other words, for all $t_0 \in \mathbb{R}$, there is $\eta \in \mathbb{R}_*^+$ such that if $t_1, t_2 \in \mathbb{R}$ satisfy $|t_1 - t_0| \leq \eta$, $|t_2 - t_0| \leq \eta$, then $d(\lambda(t_1), \lambda(t_2)) = |t_1 - t_2|$.*

Definition 2.7 *A geodesic in a metric space (E, d) is the image of a geodesic line $\lambda : \mathbb{R} \rightarrow E$.*

Finally, we have the following well known result.

Theorem 2.2 *The geodesics of the Euclidean space \mathbb{R}^3 are its straight lines.*

Remark 2.1 For other geometries (spherical and hyperbolic ones), the geodesics play the same role as the straight lines in the Euclidean space.

2.1.4 Reflections

We now focus on the study of a specific class of Euclidean isometries. Rotations and translations are known to be isometries of the Euclidean space. However, there is a specific set of isometries that can span all the others isometries: the reflections with respect to a hyperplane.

Let $a \in \mathbb{R}^3$ be a unit vector i.e. $|a|_E = 1$ and $t \in \mathbb{R}$. Let $P(a, t)$ be the hyperplane in \mathbb{R}^3 defined by

$$\begin{aligned} P(a, t) &= \{x \in \mathbb{R}^3 : \langle a, x \rangle_E = t\} \\ &= \{x \in \mathbb{R}^3 : \langle a, x - ta \rangle_E = 0\}. \end{aligned}$$

We note that a is a vector normal to the hyperplane $P(a, t)$. In addition, $P(a, t)$ is a *linear* plane if $t = 0$ and an *affine* plane if $t \neq 0$ (the term *hyperplane* is used in geometry in higher dimensions; In 3 dimensions, the hyperplanes are the planes we are used to).

Definition 2.8 (Reflection) *The reflection ρ (in \mathbb{R}^3) with respect to the hyperplane $P(a, t)$ is defined by*

$$\rho(x) = x + sa$$

where s is a real such that $\left(x + \frac{1}{2}sa\right)$ belongs to $P(a, t)$.

It is easy to see that the image of x , i.e. $\rho(x)$, is such that $P(a, t)$ is the middle plane of the segment $[x, \rho(x)]$. The vector $\rho(x) - x$ is then collinear to a . We are looking for s such that $\rho(x) - x = sa$, with the added fact that the middle of the segment $[x, \rho(x)]$ has to belong to the plane $P(a, t)$. This leads to

$$\left\langle a, \frac{x + \rho(x)}{2} - ta \right\rangle_E = 0.$$

In other words, by substituting $x + sa$ for $\rho(x)$, we obtain

$$\left\langle a, \frac{2x + sa}{2} \right\rangle_E = t$$

which is equivalent to

$$s = 2(t - \langle a, x \rangle_E).$$

Consequently, $\rho(x) = x + 2(t - \langle a, x \rangle_E)a$. It is easy to check that ρ is an isometry, and is its own inverse.

We have the following general theorem that specifies the structure of $I(\mathbb{R}^3)$. Actually, this is the group generated by the reflections. This result justifies their importance.

Theorem 2.3 ([10]) *Let σ be an isometry in the Euclidean space \mathbb{R}^3 . Then σ is the composition of at most 4 reflections.*

2.1.5 Inversions

There is a transformation of the plane which plays an essential role in the construction of the reflections in hyperbolic geometry: the inversion with respect to a sphere (a circle in 2 dimensions).

Let $S^2(a, r) = \{x \in \mathbb{R}^3 : |a - x|_E = r\}$ be the (Euclidean) sphere centered at $a \in \mathbb{R}^3$ of radius $r \in \mathbb{R}_*^+$.

Definition 2.9 (Inversion) *The inversion σ in \mathbb{R}^3 with respect to the sphere $S^2(a, r)$ is the function σ defined by $\sigma(x) = a + s(x - a)$ where s is a positive real such that $|\sigma(x) - a||x - a| = r^2$. This relation leads to $\sigma(x) = a + \frac{r^2}{|x - a|^2}(x - a)$.*

It is easy to check that $\sigma^2(x) = x$ and that

$$|\sigma(x) - \sigma(y)| = \frac{r^2|x - y|}{|x - a||y - a|}.$$

A property common to reflections and inversions is *conformity*.

Definition 2.10 *A mapping $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is conformal if it preserves the angles between regular curves.*

Theorem 2.4 ([10]) *In \mathbb{R}^3 , all the reflections with respect to a hyperplane and all the inversions with respect to a sphere are conformal.*

Next, we will need the notion of *orthogonal spheres*. The following result is a characteristic property. The result can be taken for the definition of two orthogonal spheres.

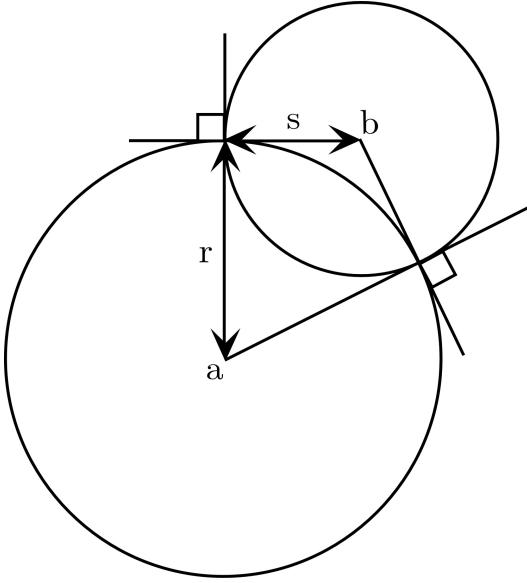


Figure 1: Two orthogonal circles

Theorem 2.5 ([10]) *Two spheres $S(a, r)$ and $S(b, s)$ are orthogonal if and only if*

$$|a - b|^2 = r^2 + s^2 \quad (2.2)$$

(see fig. 1 for the dimension 2).

2.2 Hyperbolic geometry

2.2.1 Spherical geometry

We are going to introduce the hyperbolic geometry by analogy with the spherical one. Indeed, the two geometries are dual of each other. Thus, we begin with generalities on the spherical geometry.

Consider $S^2(r) = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = r\}$ the sphere of center O, the origin, and of radius $r > 0$. We call *great circle* the intersection of a linear plane with $S^2(r)$. We can show that geodesics of the sphere are the great circles (*cf.* [10]). The curvature of the sphere of radius $r > 0$ is $\frac{1}{r^2}$. Then, the sphere of radius 1 provides a model for the spherical geometry with curvature 1. The duality with the hyperbolic geometry suggests that the curvature of the hyperbolic space must be negative. This is possible only if we allow the radius to be purely imaginary. Hence the generalisation of the inner product to the lorentzian one.

2.2.2 3D-Lorentz space

We start to construct a different metric space based on \mathbb{R}^3 . At the first sight, the resulting metric space could be disconcerting because distances will be complex numbers. It is in this metric space that we will be able to deal with hyperbolic space.

Definition 2.11 *Let $(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3$, $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$.*

(i) The lorentzian product of x and y is

$$x \circ y = x_1 y_1 + x_2 y_2 - x_3 y_3. \quad (2.3)$$

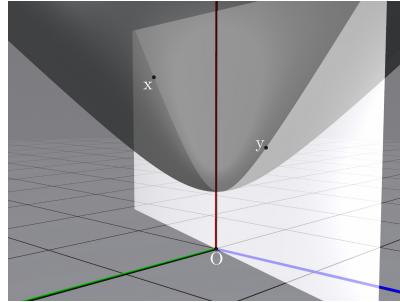


Figure 2: A geodesic in the case where x and y are in a plane perpendicular to the horizontal one

The space \mathbb{R}^3 equipped with the Lorentz product is the 3-dimensional Lorentz space and is denoted by $\mathbb{R}^{2,1}$ or \mathbb{R}^3 if the context is clear.

- (ii) The lorentzian norm of $x \in \mathbb{R}^{2,1}$ is the complex number $\|x\|_L = (x \circ x)^{1/2}$. The lorentzian norm of x is positive, 0 or pure imaginary with positive imaginary part.
- (iii) The lorentzian distance on \mathbb{R}^3 is the function $d_L : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{C}$ defined by $d_L(x, y) = \|x - y\|_L$.
- (iv) We say that the vector is time-like if $\|x\|_L$ is pure imaginary. The vector x is time-like if and only if $x_3^2 > x_1^2 + x_2^2$. A time-like vector is positive (resp. negative) if x_3 is positive (resp. negative).

The following theorem is a first step in the construction of a metric on the hyperbolic space.

Theorem 2.6 ([10]) Let x and y be two time-like positive (negative) vectors. Then $x \circ y \leq \|x\|_L \|y\|_L < 0$, with equality if x and y are linearly dependent.

Let x and y be two time-like positive (negative) vectors. Thanks to Theorem 2.6, there is a unique real positive number $\eta(x, y)$, called *time-like angle*, such that

$$x \circ y = \|x\|_L \|y\|_L \operatorname{ch} \eta(x, y). \quad (2.4)$$

2.2.3 The 2-dimensional hyperbolic space

Let $F^2 = \{x \in \mathbb{R}^{2,1} : \|x\|_L^2 = -1\}$. The problem is that the set F^2 is disconnected. The set F^2 is the hyperboloid with two sheets defined by the equation $x_1^2 + x_2^2 - x_3^2 = -1$. The set $H^2 = \{x \in \mathbb{R}^{2,1} : \|x\|_L^2 = -1, x_3 > 0\}$ is called the *positive sheet of F^2* . The set H^2 is a model of the 2-dimensional hyperbolic space. Two elements x and y belonging to H^2 are time-like positive vectors of $\mathbb{R}^{2,1}$ (cf. Definition 2.11 (iv)). By Theorem 2.6, there is a unique positive real number $\eta(x, y)$ such that

$$x \circ y = \|x\|_L \|y\|_L \operatorname{ch} \eta(x, y) = ii \operatorname{ch} \eta(x, y) = -\operatorname{ch} \eta(x, y). \quad (2.5)$$

The *hyperbolic distance* between x and y is (by definition) $d_H(x, y) = \eta(x, y)$. Of course, we have $\operatorname{ch} d_H(x, y) = -x \circ y$.

Theorem 2.7 ([10]) The hyperbolic distance d_H defined on H^2 is a metric on H^2 .

The next goal is to find the geodesics (the shortest paths) in the 2-dimensional hyperbolic space.

Definition 2.12 A hyperbolic line in H^2 is the intersection between H^2 and a linear plane in \mathbb{R}^3 which contains at least one time-like vector. We use the term *time-like linear subspace*.

Let $x, y \in H^2$ be distinct. It is easy to check that x and y are not collinear. They span a time-like linear plane $V(x, y)$ in \mathbb{R}^3 . Let $L(x, y) = H^2 \cap V(x, y)$. Then $L(x, y)$ is the only hyperbolic line in H^2 containing x and y , composed by a branch of a hyperbola (see fig. 2 and fig. 3).

Theorem 2.8 ([10]) The geodesics in H^2 are the hyperbolic lines.

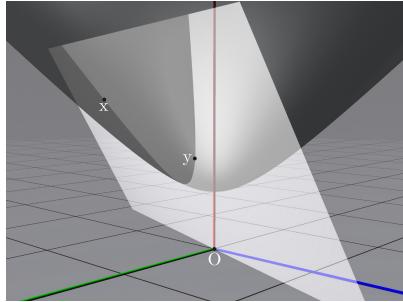
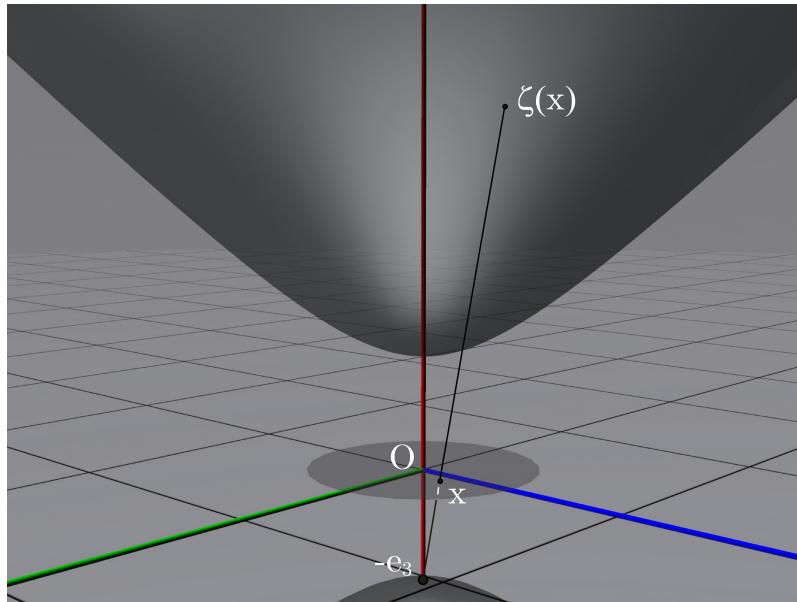
Figure 3: A geodesic passing through x and y in the general case

Figure 4: Stereographic projection

2.3 The conformal disk model

2.3.1 Stereographic projection

Let $H^2 = \{x \in \mathbb{R}^{2,1} : \|x\|_H^2 = -1, x_3 > 0\}$ be the model of the 2-dimensional hyperbolic space. Let $B^2 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 < 1, x_3 = 0\}$ and $\{e_1, e_2, e_3\}$ be the canonical basis of \mathbb{R}^3 .

Definition 2.13 (Stereographic Projection) *The stereographic projection of a point $x \in B^2$ on the hyperbolic space H^2 is the intersection point $\zeta(x)$ between H^2 and the line passing through the point x and the point $(0, 0, -1) = -e_3$ (see fig. 4).*

Clearly, $\zeta(x) - x$ is collinear to $x - (-e_3)$. Thus there is $s \in \mathbb{R}^*$ such that $\zeta(x) = x + s(x + e_3)$. As $\zeta(x) \in H^2$, we have $\|\zeta(x)\|_L^2 = -1$. This leads immediately to $s = \frac{1 + |x|_E^2}{1 - |x|_E^2}$. The fact that we can use the Euclidean inner product comes from the fact that $x \in B^2 \subset \mathbb{R}^2$. The space \mathbb{R}^2 is embedded in the Lorentz space $\mathbb{R}^{2,1}$ by adding a third coordinate fixed to 0.

Then, $\zeta(x) = \left(\frac{2x_1}{1 - |x|_E^2}, \frac{2x_2}{1 - |x|_E^2}, \frac{1 + |x|_E^2}{1 - |x|_E} \right)$. The stereographic projection ζ is a bijection with inverse is $\zeta^{-1}(y) = \left(\frac{y_1}{1 + y_3}, \frac{y_2}{1 + y_3} \right)$.

Now, we have to equip B^2 with a metric making it a hyperbolic space.

Definition 2.14 On B^2 , we define the distance $d_B(x, y) = d_H(\zeta(x), \zeta(y))$. By definition, the projection ζ is then an isometry from (B^2, d_B) onto (H^2, d_H) . The space (B^2, d_B) is called the conformal disk model of the 2-dimensional hyperbolic space, or Poincaré disk.

We give an explicit expression of d_B in the following result.

Theorem 2.9 The metric d_B on B^2 is given by

$$\operatorname{ch} d_B(x, y) = 1 + \frac{2|x - y|_E^2}{(1 - |x|_E^2)(1 - |y|_E^2)} \quad (2.6)$$

for all $x, y \in B^2$.

Using the fact that $\operatorname{argch} x = \ln(x + \sqrt{x^2 - 1})$ and $\operatorname{argth} x = \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right)$, we can establish the next result.

Corollary 2.1 For all $x \in B^2$,

$$d_B(O, x) = 2 \operatorname{argth} |x|_E \quad (2.7)$$

As $d_B(O, x) = 2 \operatorname{argth} |x|_E$, the Euclidean disk centered at O and of radius $\frac{1}{2}$ is then the hyperbolic disk centered at O and of radius $\ln 3 \simeq 1.097$, hence the zoom effect (*focus*). The fact that $d_B(O, 1) = +\infty$ proves also that the whole space is contained in the Euclidean unit disk, hence the *context* idea.

The fact that H^2 presents a rotation invariance around the axis (O, e_3) suggests that rotations centered at O in B^2 equipped with d_B are the same as those of B^2 equipped with d_E .

2.3.2 Geodesics, angles

Since the stereographic projection is an isometry, the same is true of its inverse. Thus the geodesics in the hyperbolic space (H^2, d_H) are projected into geodesics of (B^2, d_B) . Indeed, the geodesics in (B^2, d_B) are on the one hand the diameters of the unit circle (corresponding to the projections of the hyperbolic lines resulting from the intersection between H^2 with the hyperplane $V(\zeta(x), \zeta(y))$ where x and y belong to the same diameter) and on the other hand the circle arcs orthogonal to the unit circle (corresponding to the projections of the branches of the hyperbola resulting from the intersection between H^2 with the hyperplanes $V(\zeta(x), \zeta(y))$ where x and y are not on the same diameter (see fig. 5)). It is easy to convince oneself that a geodesic in B^2 passing through O is necessarily a diameter.

It is possible to show that angles between two regular curves in B^2 are the same as angles between these curves in E^2 but this is not our purpose here (*cf.* [10]). However, we will use this fact in a fundamental way.

2.3.3 Hyperbolic transformations

Rotations In view of the preceding paragraph, there is a family of simple transformations in the Poincaré model: Rotations centered at the origin O . Euclidean rotations centered at O are spanned by two Euclidean reflections with axis passing through O . As the Euclidean geodesics passing through O are also hyperbolic geodesics (passing through O), Euclidean and hyperbolic reflections with axes passing through O are the same. Consequently, Euclidean and hyperbolic rotations centered at O are also the same. Of course hyperbolic rotations are isometries of the unit disk, thanks to the rotation symmetry of the Poincaré disk.

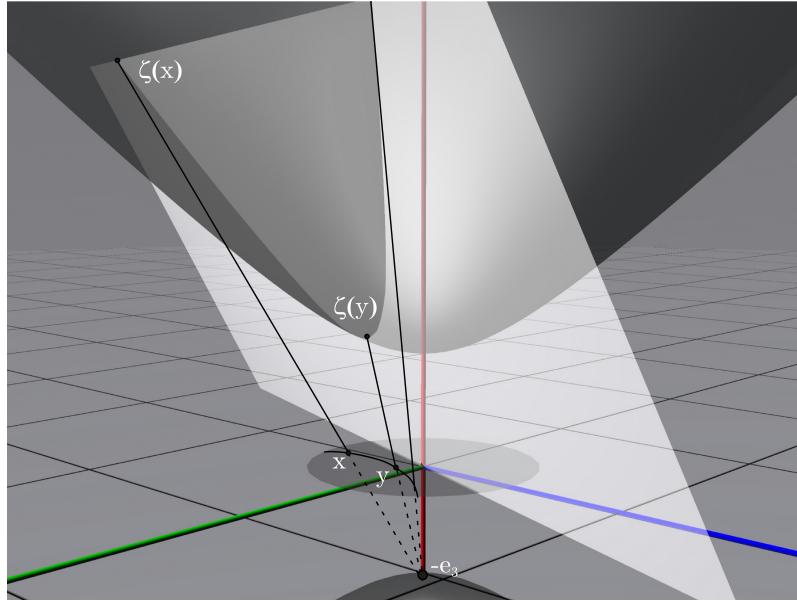


Figure 5: A geodesic in the hyperbolic plane, and its projection in the Poincaré disk

Translations Following the model of the Euclidean plane, a translation is the composition of two reflections with parallel axes. In hyperbolic geometry, the problem comes from “parallel”, since if we consider that two “lines” (*i.e.* two geodesics) are parallel if they do not intersect, then there are an infinity of lines parallel to a given one and passing through a given point outside of the given line (*cf.* Introduction).

We start with particular reflections with respect to an axis, and then move to particular inversions. The translation will be the map resulting from the composition of these two kinds of transformations.

Let ρ be a reflection with respect to a linear hyperplane, that is to say a reflection with respect to a diameter of the unit circle in \mathbb{R}^2 ($= S(O, 1)$). Of course, it leaves B^2 globally invariant. Let us calculate $\operatorname{ch} d_B(\rho(x), \rho(y))$ for any x and any y in B^2 . Using (2.6), the fact that $\rho(O) = O$ and that reflections are Euclidean isometries, we obtain

$$\begin{aligned}\operatorname{ch} d_B(\rho(x), \rho(y)) &= 1 + \frac{2|\rho(x) - \rho(y)|_E^2}{(1 - |\rho(x)|_E^2)(1 - |\rho(y)|_E^2)} \\ &= 1 + \frac{2|x - y|_E^2}{(1 - |x|_E^2)(1 - |y|_E^2)} \\ &= \operatorname{ch} d_B(x, y),\end{aligned}$$

therefore ρ is a hyperbolic isometry.

Concerning inversions, we note that if $S(a, r)$ is a circle orthogonal to S^2 , then Theorem 2.5 implies that $r^2 = |a|_E^2 - 1$. Therefore r is function of a . Moreover, the center of such a sphere cannot belong to B^2 . Let σ_a be the inversion with respect to such a sphere (*cf.* Definition 2.9). For $x \in B^2$, we have

$$\begin{aligned}|\sigma_a(x)|_E^2 &= \left| a + \frac{|a|_E^2 - 1}{|x - a|_E^2}(x - a) \right|^2 \\ &= \frac{|x|_E^2 |a|_E^2 - 2\langle a, x \rangle_E + 1}{|x - a|_E^2}.\end{aligned}$$

But

$$\begin{aligned} x \in B^2 &\Leftrightarrow |x|_E^2 \leq 1 \\ &\Leftrightarrow (|a|_E^2 - 1)|x|_E^2 \leq |a|_E^2 - 1 \\ &\Leftrightarrow -2\langle a, x \rangle_E + |a|_E^2|x|_E^2 + 1 \leq |x|_E^2 + |a|_E^2 - 2\langle a, x \rangle_E \\ &\Leftrightarrow |\sigma_a(x)|^2 \leq 1 \\ &\Leftrightarrow |\sigma_a(x)| \in B^2. \end{aligned}$$

This calculation proves that inversions with respect to a sphere which is orthogonal to the unit circle leaves B^2 invariant.

Let us now show that such an inversion is a hyperbolic isometry. Since $\sigma_a \left(\frac{a}{|a|_E^2} \right) = 0$, we have

$$\begin{aligned} 1 - |\sigma_a(x)|_E^2 &= 1 - \left| \sigma_a(x) - \sigma_a \left(\frac{a}{|a|_E^2} \right) \right|_E^2 \\ &= 1 - \frac{(|a|_E^2 - 1)^2 \left| x - \frac{a}{|a|_E^2} \right|_E^2}{|x - a|_E^2 \left| \frac{a}{|a|_E^2} - a \right|_E^2} \\ &= \frac{|x - a|_E^2 |a|_E^2 - ||a|_E^2 x - a|_E^2}{|x - a|_E^2 |a|_E^2} \\ &= \frac{(|a|_E^2 - 1)(1 - |x|_E^2)}{|x - a|_E^2}. \end{aligned}$$

Then, using the relation (2.6), we have

$$\begin{aligned} \operatorname{ch} d_B(\sigma_a(x), \sigma_a(y)) &= 1 + \frac{2|\sigma_a(x) - \sigma_a(y)|_E^2}{(1 - |\sigma_a(x)|_E^2)(1 - |\sigma_a(y)|_E^2)} \\ &= 1 + 2 \frac{(|a|_E^2 - 1)^2 |x - y|_E^2}{|x - a|_E^2 |y - a|_E^2} \frac{|x - a|_E^2}{(|a|_E^2 - 1)(1 - |x|_E^2)} \frac{|y - a|_E^2}{(|a|_E^2 - 1)(1 - |y|_E^2)} \\ &= \operatorname{ch} d_B(x, y) \end{aligned}$$

and σ_a is a hyperbolic isometry.

If ρ_a is the reflection with respect to the hyperplane (the linear line in \mathbb{R}^2) for which a is a normal vector, then the isometry $\rho_a \circ \sigma_a$ leaves B^2 globally invariant. Moreover, analytically, we have

$$\rho_a \circ \sigma_a(x) = \frac{|a|_E^2 - 1}{|x - a|_E^2} x - \frac{|x|_E^2 - \langle x, \frac{a}{|a|_E^2} \rangle_E + 1}{|x - a|_E^2} a. \quad (2.8)$$

In particular, $\rho_a \circ \sigma_a(O) = -\frac{a}{|a|_E^2}$. This calculation suggests a change from a to b' in order to obtain $\rho_{b'} \circ \sigma_{b'}(O) = b$ where b is given. Then, if $b \in B^2$ is given, the circle $S \left(\frac{-b}{|b|_E^2}, \sqrt{\frac{1}{|b|_E^2} - 1} \right)$ is orthogonal to S^2 . We can define a hyperbolic isometry $t_b = \rho_{-b/|b|_E^2} \circ \sigma_{-b/|b|_E^2}$. More explicitly, we have

$$t_b(x) = \frac{1 - |b|_E^2}{|b|_E^2 |x|_E^2 + 2\langle x, b \rangle_E + 1} x + \frac{|x|_E^2 + 2\langle x, b \rangle_E + 1}{|b|_E^2 |x|_E^2 + 2\langle x, b \rangle_E + 1} b. \quad (2.9)$$

The isometry t_b results from the composition of an inversion and a reflection with “parallel” axes (the sphere $S \left(\frac{-b}{|b|_E^2}, \sqrt{\frac{1}{|b|_E^2} - 1} \right)$ and the axis of the reflection $\rho_{-b/|b|_E^2}$ do not intersect). Moreover,

they are both orthogonal to the (Euclidean) segment $\left[\frac{-b}{|b|_E^2}, \frac{b}{|b|_E^2} \right]$. The transformation t_b acts like a translation along this segment (in the sense that this segment is globally invariant). Also, we have $t_b(O) = b$. The transformation t_b is called the *hyperbolic translation of vector $b \in B^2$* (see fig. 6).

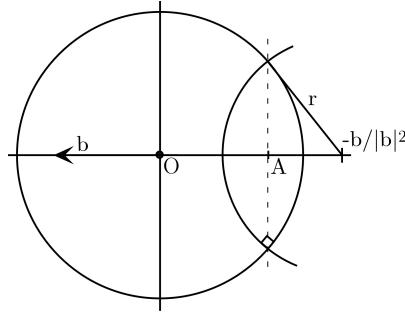


Figure 6: Translation in the Poincaré disk; $r = \sqrt{\frac{1}{|b|^2} - 1}$ and $A = \sigma_{-b/|b|^2}(O)$

Remark 2.2 It is worth noting that reflections and inversions preserve the measures of angles, but not the orientation. The composition of two such transformations are both measure angle and orientation preserving. Hyperbolic rotations and hyperbolic translations are conformal transformations (*i.e.* preserve the orientation).

Remark 2.3 This construction is also valid, modulo minor changes, in the case of higher dimensions. However, 2 dimensions lead us naturally to a complex interpretation. In the sequel, we identify B^2 with the unit disk of the complex plane. If $x = (x_1, x_2) \in B^2$, we denote $z = x_1 + ix_2$, $\bar{z} = x_1 - ix_2$, $x_1 = \Re(z)$ and $x_2 = \Im(z)$. In this case, rotations centered at O with angle $\Theta \in [0, 2\pi[$ are of the analytic form $r_\Theta(z) = ze^{i\Theta} = \theta z$ where $\theta = e^{i\Theta}$. The goal is to give a complex number interpretation of the translation constructed in the last paragraph. With complex number notations, we find, if $x = (x_1, x_2)$ et $b = (b_1, b_2)$,

$$\begin{aligned}
 t_b(x) &= \frac{1 - b_1^2 - b_2^2}{(b_1^2 + b_2^2)(x_1^2 + x_2^2) + 2(x_1 b_1 + x_2 b_2) + 1} (x_1 + ix_2) \\
 &\quad + \frac{x_1^2 + x_2^2 + 2(x_1 b_1 + x_2 b_2) + 1}{(b_1^2 + b_2^2)(x_1^2 + x_2^2) + 2(x_1 b_1 + x_2 b_2) + 1} (b_1 + ib_2) \\
 &= \frac{x_1 - x_1 b_1^2 - x_1 b_2^2 + x_1^2 b_1 + x_2^2 b_1 + 2x_1 b_1^2 + 2x_2 b_1 b_2 + b_1}{(b_1^2 + b_2^2)(x_1^2 + x_2^2) + 2(x_1 b_1 + x_2 b_2) + 1} \\
 &\quad + i \frac{x_2 - x_2 b_1^2 - x_2 b_2^2 + x_1^2 b_2 + x_2^2 b_2 + 2x_1 b_1 b_2 + 2x_2 b_2^2 + b_2}{(b_1^2 + b_2^2)(x_1^2 + x_2^2) + 2(x_1 b_1 + x_2 b_2) + 1} \\
 &= \frac{(x_1 + b_1 + i(x_2 + b_2))(1 + x_1 b_1 + x_2 b_2 - i(x_2 b_1 - x_1 b_2))}{(1 + x_1 b_1 + x_2 b_2)^2 + (x_2 b_1 - x_1 b_2)^2} \\
 &= \frac{x_1 + ix_2 + b_1 + ib_2}{1 + x_1 b_1 + x_2 b_2 + i(x_2 b_1 - x_1 b_2)} \\
 &= \frac{z + b}{1 + \bar{b}z}.
 \end{aligned}$$

It will be in this synthetic form

$$t_b(z) = \frac{z + b}{1 + \bar{b}z}, \tag{2.10}$$

given by Lampang and *al.* [6], that we will consider hyperbolic transformations.

Remark 2.4 We have $|t_b(O)|_E = |b|_E$ and $|t_b(-b)|_E = 0$. These are the only points in B^2 that are effectively translated by a (Euclidean) length $|b|_E$. The points of the diameter passing through the point B with affix b are also translated, that is to say that the diameter is globally invariant under t_b . On the other hand, the other points in B^2 are subject to a rotation. The farther from

the origin they are, the more rotated they are.

It is easy to see that if b and b' are two non-collinear complex numbers with modulus smaller than 1, then, for all $z \in \mathbb{C}$,

$$\begin{aligned} t_{b'} \circ t_b(z) &= \frac{\frac{z+b}{1+\bar{b}z} + b'}{1 + \bar{b}' \frac{z+b}{1+\bar{b}z}} \\ &= \frac{(1+\bar{b})z + b + b'}{1 + b\bar{b}' + (\bar{b} + \bar{b}')z} \end{aligned} \quad (2.11)$$

is not a hyperbolic translation. The set of hyperbolic translations is not a group under the composition of mappings. We have to add another set of transformations in order to obtain a group, namely, the rotations centered at O .

Let $\tau_{\theta,b} = t_b \circ r_\theta$ be the hyperbolic transformation composed of a rotation of an angle equal to $\arg \theta$ ($|\theta| = 1$) followed by a translation of a vector $b \in B^2$. For all $z \in B^2$, we have

$$\tau_{\theta,b}(z) = \frac{\theta z + b}{1 + \bar{b}\theta z}. \quad (2.12)$$

We have the following theorem.

Theorem 2.10 *The set $\mathcal{M}(B^2) = \{\tau_{\theta,b} \in I(B^2) : |\theta| = 1, |b| < 1\}$ is a group for the composition of mappings. Moreover, every element in $\mathcal{M}(B^2)$ leaves the unit circle invariant.*

Proof: First we show that the composition of mappings is a binary operation.

First $Id_{B^2} = \tau_{1,0}$.

Let θ and θ' be two complex of modulus 1 and $b, b' \in B^2$. We have to show that $\tau_{\theta',b'} \circ \tau_{\theta,b} = \tau_{\theta_f, b_f}$ for suitable θ_f with modulus 1 and b_f satisfying $|b_f| < 1$. For all $z \in B^2$, we have

$$\begin{aligned} \tau_{\theta',b'} \circ \tau_{\theta,b}(z) &= \frac{\theta' \frac{\theta z + b}{1 + \bar{b}\theta z} + b'}{1 + \bar{b}' \theta' \frac{\theta z + b}{1 + \bar{b}\theta z}} \\ &= \frac{\frac{\theta\theta' + \theta\bar{b}b'}{1 + \theta'\bar{b}b'} z + \frac{\theta'b + b'}{1 + \theta'\bar{b}b'}}{1 + \frac{\theta\bar{b} + \theta'\bar{b}b'}{1 + \theta'\bar{b}b'} z}. \end{aligned}$$

In order to show that the law is a binary one, it remains to show that:

- $\theta_f = \frac{\theta\theta' + \theta\bar{b}b'}{1 + \theta'\bar{b}b'}$ has modulus 1;
- $b_f = \frac{\theta'b + b'}{1 + \theta'\bar{b}b'}$ belongs to B^2 ;
- $\bar{b}_f \theta_f = \frac{\theta\bar{b} + \theta'\bar{b}b'}{1 + \theta'\bar{b}b'}$.

But the fact that $|\theta| = 1$ implies $\bar{\theta} = \frac{1}{\theta}$. Moreover, $|z| = |\bar{z}|$. Then, for all $z \in \mathbb{C}$,

$$\begin{aligned} |\theta_f| &= \frac{|\theta\theta' + \theta\bar{b}b'|}{|1 + \theta'\bar{b}b'|} \\ &= \frac{|1 + \frac{\bar{b}b'}{\theta'}|}{|1 + \theta'\bar{b}b'|} \\ &= \frac{|1 + \overline{\theta'\bar{b}b'}|}{|1 + \theta'\bar{b}b'|} = 1 \end{aligned}$$

In addition, we have

$$|b_f|^2 = b_f \bar{b}_f = \frac{|b|^2 + |b'|^2 + 2\Re(\theta'\bar{b}b')}{|b|^2|b'|^2 + 2\Re(\theta'\bar{b}b') + 1}$$

Now $(|b|^2 - 1)(1 - |b'|^2) < 0$, therefore, $|b|^2 + |b'|^2 - |b|^2|b'|^2 - 1 < 0$ and consequently, adding the missing term,

$$|b|^2 + |b'|^2 + 2\Re(\theta'bb') < |b|^2|b'|^2 + 1 + 2\Re(\theta'bb')$$

and thus $|b_f| < 1$.

Finally,

$$\begin{aligned} \bar{b}_f \theta_f i &= \frac{\bar{\theta}'\bar{b} + \bar{b}'}{1 + \bar{\theta}'\bar{b}\bar{b}'} \frac{\theta\theta' + \theta\bar{b}b'}{1 + \theta'\bar{b}\bar{b}'} \\ &= \frac{\theta\theta'\bar{\theta}'\bar{b} + \theta\bar{\theta}'\bar{b}^2 b' + \theta\theta'\bar{b}'}{(1 + \bar{\theta}'\bar{b}\bar{b}')(1 + \theta'\bar{b}\bar{b}')} \\ &= \frac{\bar{\theta}'\bar{b}\bar{b}'(\theta\bar{b} + \theta\theta'\bar{b}') + \theta\bar{b} + \theta\theta'\bar{b}'}{(1 + \bar{\theta}'\bar{b}\bar{b}')(1 + \theta'\bar{b}\bar{b}')} \\ &= \frac{\theta\bar{b} + \theta\theta'\bar{b}'}{1 + \theta'\bar{b}\bar{b}'}. \end{aligned}$$

Therefore, $\mathcal{M}(B^2)$ is a group. Indeed, since $t_p^{-1} = t_{-p}$ and $r_\theta^{-1} = r_{-\theta}$, the inverse of $\tau_{\theta,p} \in \mathcal{M}(B^2)$ is

$$\begin{aligned} \tau_{\theta,p}^{-1} &= (t_p \circ r_\theta)^{-1} \\ &= r_\theta^{-1} \circ t_p^{-1} \\ &= r_{-\theta} \circ t_{-p} \\ &= t_{-p\bar{\theta}} \circ r_{-\bar{\theta}} \in \mathcal{M}(B^2). \end{aligned}$$

It remains to show the second statement of the Theorem, namely that the image of a point belonging to the circle by an arbitrary transformation of the form $\tau_{\theta,b}$ is a point belonging to the circle. Let z be the affix of such a point. We have $|z| = 1$ and thus $\bar{z} = \frac{1}{z}$. Then, the sequence of following inequalities is valid:

$$\begin{aligned} |\tau_{\theta,b}(z)| &= \left| \frac{\theta z + b}{1 + \bar{b}\theta z} \right| \\ &= |\theta z| \frac{|1 + \frac{b}{\theta z}|}{|1 + \bar{b}\theta z|} \\ &= \frac{|1 + \bar{\theta}zb|}{|1 + \bar{b}\theta z|} \\ &= 1, \end{aligned}$$

which proves that $\tau_{\theta,b}(z)$ is a point belonging to the unit circle, as claimed. ■

We will use these formulas to visualize arborescent data with hyperbolic geometry.

3 Implementation

We describe in this section the formulas and algorithms used in the hyperbolic tree visualization. Those formulas and algorithms are directly based on the mathematical definitions and theorems of the preceding section. We also explain why and how we use them and the logic behind them. In the following, the hyperbolic tree visualization is called a *hypertree*.

3.1 Hypertree description

The hypertree is a way to visualize structured data as a tree. The tree is drawn radially — the root is centered, the nodes are placed around it, and the children are set on a circle centered on their father — in the hyperbolic plane. Then, once the layout of the tree in the hyperbolic space

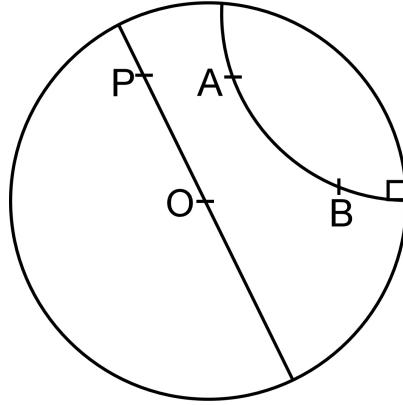


Figure 7: Two kinds of geodesics: A diameter going through O and P , and a circle arc \hat{AB} orthogonal to the unit disk

is done, the whole tree is projected (*cf.* paragraph 2.3.1) in the Poincaré disk. The root of the tree is projected on the center of the unit disk. The branches of the tree turn into geodesics of the Poincaré model, that is either diameters or circle arcs (*cf.* fig. 7 and paragraph 2.3.2).

The main advantage of such a visualization is that the length contraction for nodes farther from the center of the disk allows users to keep a global view on the tree structure, even if the tree is really large. Moreover, the center of the disk offers a detailed view of the items in it. Such a view is called a “*focus+context*” visualization: A detailed view (*focus*) associated with a global view (*context*).

The main functionalities of the *hypertree* are:

- the tree is drawn in the unit disk (*i.e.* in a bounded area (*context*)), as if the tree has been drawn in the hyperbolic space and then projected into the Poincaré model;
- the whole tree can be dragged with the mouse to put any part of it in the middle of the disk (*focus*);
- clicking on a node moves it automatically at the center of the disk.

We will now see each functionality in turn, in order to understand the necessary formulas, their logic, and their implementation.

3.2 Tree layout

The first functionality of the *hypertree* is to layout a tree of structured data in the 2-dimensional hyperbolic space, and then to project it in the Poincaré model in order to draw it.

The goal of the layout algorithm is to optimize the space used while setting up the nodes. In our case, we use a radial layout. This means that the root of the tree is put at the center O and the remaining nodes (P_1, P_2, P_3, \dots and their children) are spread in every direction from the root (*cf.* fig. 8).

The first difficulty is to lay out the tree in the hyperbolic plane. But before dealing with the non-usual hyperbolic geometry, we first describe the radial layout algorithm in the Euclidean plane.

3.2.1 Radial layout in the Euclidean plane

The main idea of the algorithm is the following: For each generation composed of a father node and his children nodes, the children nodes are put on a circle centered on the father. Let us suppose

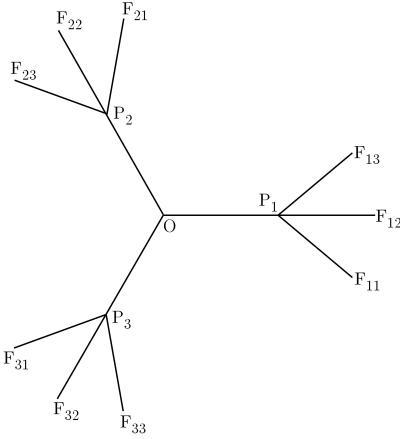


Figure 8: An example of a tree radial layout in the Euclidean plane

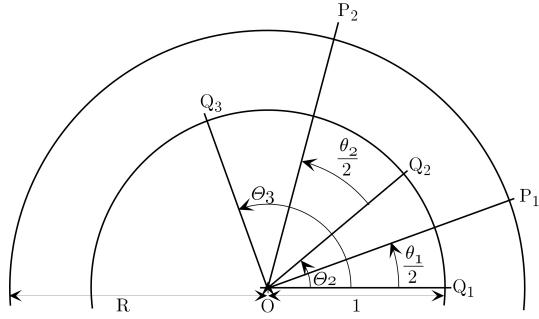


Figure 9: Layout of the first generation nodes

that, before optimisation, the father-child length is the same for every generation. Let R be this length.

The algorithm starts by putting the root of the tree at the origin O . Let n be the number of first generation nodes (the root children). Those n points P_1, \dots, P_n are put on a circle centered at O and of radius R . To have an optimal use of the space available, the 2π radians of the circle are set up between the n children. An angular sector θ_j is given to each of the node P_j of the first generation. Put:

- $\Theta_1 = 0$;
- $\Theta_j = \Theta_{j-1} + \theta_{j-1}$, $j = 2, \dots, n$;
- Q_j the point of coordinates $(\cos \Theta_j, \sin \Theta_j)$.

Each of the n nodes P_j is then put on the line bisecting the angular sector $Q_j \widehat{OQ}_{j+1}$, at a distance R of O (*cf.* fig. 9). We have

$$P_j : \left(R \cos \left(\Theta_j + \frac{\theta_j}{2} \right), R \sin \left(\Theta_j + \frac{\theta_j}{2} \right) \right). \quad (3.1)$$

Once P_j has been layed out, we focus on its children. The algorithm to layout the children of P_j is the same as the one to layout the children of O . The only difference is that the angular sector

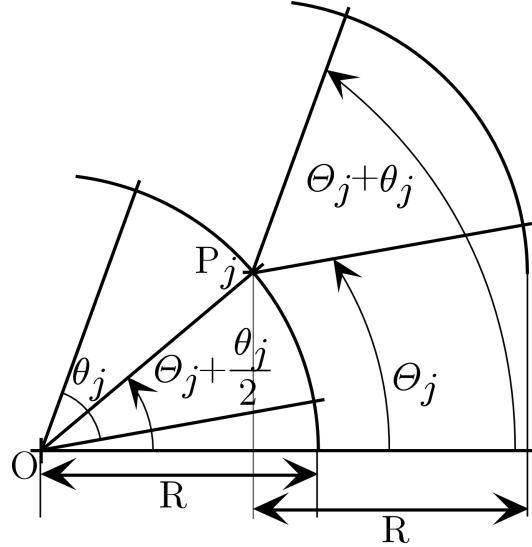


Figure 10: The angular sector to share between the children of P_j

to share is not 2π anymore. A simple algorithm gives to the children of P_j the same angular sector as the one given to P_j , namely θ_j . This angular sector will be shared symmetrically around the line (OP_j) . There is an angle $\Theta_j + \frac{\theta_j}{2}$ between this line and the abscissa axis. Thus, the angular sector spreads from $(\Theta_j + \frac{\theta_j}{2}) - \frac{\theta_j}{2} = \Theta_j$ to $(\Theta_j + \frac{\theta_j}{2}) + \frac{\theta_j}{2} = \Theta_{j+1}$ (cf. fig. 10).

Let:

- n_j the number of children of the node P_j ;
- $F_{j,k}$ the k^{th} child of P_j ;
- $\theta_{j,k}$ the angular sector allocated to $F_{j,k}$ and bisected by the line $(P_j F_{j,k})$;
- $\Theta_{j,1} = 0$;
- $\Theta_{j,k} = \Theta_{j,k-1} + \theta_{j,k-1}$, $k = 2, \dots, n_j$.

Then the nodes $F_{j,k}$ are put at a length R of P_j , on the line bisecting their allocated angular sector (as before). Let $(\cdot, \cdot)_Q$ be the relative coordinates in the system centered at Q . To find the cartesian coordinates of $F_{j,k}$, we switch by translation to the coordinate system centered at P_j . In this new system, relative coordinates of $F_{j,k}$ are

$$\left(R \cos \left(\Theta_j + \Theta_{j,k} + \frac{\theta_{j,k}}{2} \right), R \sin \left(\Theta_j + \Theta_{j,k} + \frac{\theta_{j,k}}{2} \right) \right)_{P_j}.$$

We need to apply the translation of vector $\overrightarrow{OP_j}$ to those relative coordinates to find the absolute coordinates of the $F_{j,k}$ in the system centered at O . If $P_j = (X_{P_j}, Y_{P_j})$, then the coordinates of the vector $\overrightarrow{OP_j}$ are (X_{P_j}, Y_{P_j}) . Thus, the absolute coordinates of $F_{j,k}$ are

$$\left(R \cos \left(\Theta_j + \Theta_{j,k} + \frac{\theta_{j,k}}{2} \right) + X_{P_j}, R \sin \left(\Theta_j + \Theta_{j,k} + \frac{\theta_{j,k}}{2} \right) + Y_{P_j} \right).$$

For the third generation, we proceed in a similar manner, with the only difference that the angular sector should be shared around the line $(P_j F_{j,k})$. There is an angle $\alpha = \Theta_j + \Theta_{j,k} + \frac{\theta_{j,k}}{2}$

between this line and the abscissa axis. Thus in the system centered at $F_{j,k}$, the relative coordinates of the children $G_{j,k,l}$ of $F_{j,k}$ are (with $\Theta_{j,k,l}$ and $\theta_{j,k,l}$ defined by analogy)

$$\left(R \cos \left(\left(\alpha - \frac{\theta_{j,k}}{2} \right) + \left(\Theta_{j,k,l} + \frac{\theta_{j,k,l}}{2} \right) \right), R \sin \left(\left(\alpha - \frac{\theta_{j,k}}{2} \right) + \left(\Theta_{j,k,l} + \frac{\theta_{j,k,l}}{2} \right) \right) \right)_{F_{j,k}},$$

or, after simplification,

$$\left(R \cos \left(\Theta_j + \Theta_{j,k} + \Theta_{j,k,l} + \frac{\theta_{j,k,l}}{2} \right), R \sin \left(\Theta_j + \Theta_{j,k} + \Theta_{j,k,l} + \frac{\theta_{j,k,l}}{2} \right) \right)_{F_{j,k}}.$$

Eventually we apply the algorithm recursively to all generations to find the coordinates of every tree node:

- the root of the tree is put at the origin $O : (0, 0)$;
- for the nodes P_j of the first generation:
 - we divide out the 2π radians between the children. Each child receives an angular sector θ_j ,
 - $\Theta_1 = 0$,
 - $\Theta_j = \Theta_{j-1} + \theta_{j-1}$, $j = 2, \dots, n$,
 - $P_j : \left(R \cos \left(\Theta_j + \frac{\theta_j}{2} \right), R \sin \left(\Theta_j + \frac{\theta_j}{2} \right) \right)$,
 - the vector $\overrightarrow{OP_j}$ makes an angle $\alpha_j = \Theta_j + \frac{\theta_j}{2}$ with the abscissa axis;
- for the next generations :
 - let P , with coordinates (X_P, Y_P) , be the father that received the angular sector θ ,
 - let F_1, \dots, F_n be the children of P ,
 - let M be the father of P ,
 - let Θ be the angle between the vector \overrightarrow{MP} and the abscissa axis,
 - we divide out θ between the children F_1, \dots, F_n . Each child owns an angular sector θ_j ,
 - $\Theta_1 = 0$,
 - $\Theta_j = \Theta_{j-1} + \theta_{j-1}$, $j = 2, \dots, n$,
 - $F_j : \left(R \cos \left(\left(\Theta - \frac{\theta}{2} \right) + \left(\Theta_j + \frac{\theta_j}{2} \right) \right) + X_P, R \sin \left(\left(\Theta - \frac{\theta}{2} \right) + \left(\Theta_j + \frac{\theta_j}{2} \right) \right) + Y_P \right) := (X_{F_j}, Y_{F_j})$
 - the vector $\overrightarrow{PF_j}$ makes an angle $\Theta' = \left(\left(\Theta - \frac{\theta}{2} \right) + \left(\Theta_j + \frac{\theta_j}{2} \right) \right)$ with the horizontal axis.

3.2.2 Radial layout in the hyperbolic plane

The tree layout algorithm is the same in the hyperbolic plane as in the Euclidean one. But the formulas are different, since the classical trigonometry formulas cannot be used anymore in the hyperbolic space.

The first step of the layout is the same in the hyperbolic space and in the Euclidean space: The root of the tree is put at the origin O . Then, its n children P_1, \dots, P_n are put on the hyperbolic circle centered at O and of radius R . Indeed, as told in the paragraph 2.3.3, circles centered at O in the hyperbolic space remain circles in the Poincaré model. As in the Euclidean space, the 2π radians of the circle are shared between the n children, and each node P_j of the first generation receives an angular sector θ_j .

As the root of the tree, which is the father node, is at the origin, every geodesic linking the father node to its children is a diameter (*i.e.* a “real” line, in the Euclidean sense). Therefore we can apply the classic trigonometric formulas. Thus, the coordinates of first generation children P_j are

$$\left(R \cos \left(\Theta_j + \frac{\theta_j}{2} \right), R \sin \left(\Theta_j + \frac{\theta_j}{2} \right) \right)$$

like in the Euclidean space (*cf.* relation (3.1)), with

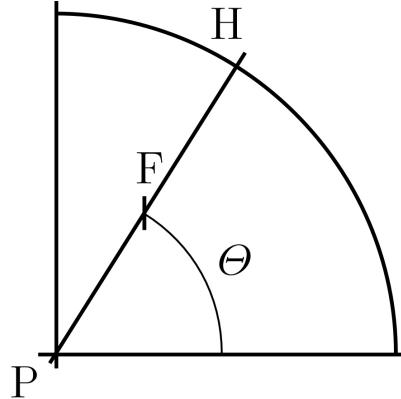


Figure 11: Layout of the child node F when the father node P is at the center of the circle

- θ_j the angular sector allocated to P_j ;
- $\Theta_1 = 0$;
- $\Theta_j = \Theta_{j-1} + \theta_{j-1}$, $j = 2, \dots, n$.

However, for the next generation, the computation is different. Indeed, as the nodes of the first generation are not at the unit disk center O anymore, the geodesics linking those nodes to their children are not lines anymore but circle arcs orthogonal to the unit circle (*cf.* fig. 7). In this case, classical trigonometric formulas are no longer valid. As those formulas are only valid when the father node is at the origin, we will switch to another coordinate system: We will lay out the children nodes with respect to a father node centered at the origin. This will be done with the formulas described in the preceding section. Then, we will apply to the resulting points the translation that puts back the father node into its real position. However, there is a problem with this method: When we translate a child, the diameter (PF) linking the father to the child becomes a circle arc. Now, it is around this geodesic that the children of F should be laid out. This means that the angle around which the children of F should be spread is not the Euclidean angle between the diameter (PF) and the abscissa axis, measured before the translation. More exactly, let H be the intersection point between the diameter (PF) and the unit circle when P is at the center of the circle. If Θ is the angle between (PF) and the abscissa axis, the affix of H is $e^{i\Theta}$ (*cf.* fig. 11).

When we bring back P to its real position by the hyperbolic translation t_{z_P} of vector \overrightarrow{OP} , the diameter (PF) becomes a circle arc orthogonal to the unit circle. It is around this geodesic that the children of F should be laid out. So we need to find the hyperbolic angle between this geodesic and the abscissa axis. Let H' be the image of H by the hyperbolic translation t_{z_P} (*cf.* fig. 12). Thanks to Theorem 2.10, H' is on the unit circle.

As we need the angle from the point of view of F , we apply the hyperbolic translation t_{-z_F} of vector $-\overrightarrow{OF}$. This translation moves F to the origin. Let H'' be the image of H' by this translation. As H' is on the unit circle, thanks to Theorem 2.10, H'' is also on the unit circle. Its affix is $e^{i\Theta''}$, with Θ'' the searched angle.

It is around this angle that the children of F will be laid out. So, if:

- P (of affix z_P) is the father of F (of affix z_F);
- there is an angle Θ between the diameter (PF) and the abscissa axis when P is at the origin;
- Θ'' is the angle around which the children of F should be laid out;

then we have

$$e^{i\Theta''} = t_{-z_F} \circ t_{z_P}(e^{i\Theta}). \quad (3.2)$$

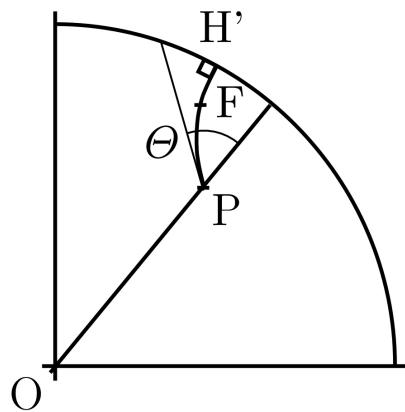


Figure 12: Image of the figure 11 after the translation t_{z_P}

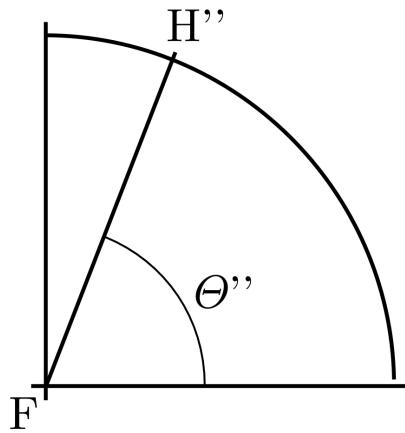


Figure 13: Image of the figure 12 after the translation t_{-z_F}

Once the geodesic angle has been found, we can apply the formulas that were valid in the Euclidean space. Thus, the relative coordinates of the children $F_{j,k}$ are

$$\left(R \cos \left(\left(\Theta_j'' - \frac{\theta_j}{2} \right) + \left(\Theta_{j,k} + \frac{\theta_{j,k}}{2} \right) \right), \sin \left(\left(\Theta_j'' - \frac{\theta_j}{2} \right) + \left(\Theta_{j,k} + \frac{\theta_{j,k}}{2} \right) \right) \right)_{P_j}$$

$$:= (X_{F_{j,k}}, Y_{F_{j,k}})_{P_j} \text{ with}$$

- M the father of P_j ;
- Θ_j the angle between the diameter (MP_j) and the abscissa axis when M is at the origin;
- Θ_j'' the angle calculated from Θ_j in the formula (3.2);
- θ_j the angular sector to share between the children $F_{j,k}$ of P_j ;
- $\theta_{j,k}$ the angular sector allocated to $F_{j,k}$;
- $\Theta_{j,1} = 0$;
- $\Theta_{j,k} = \Theta_{j,k-1} + \theta_{j,k-1}$, $k = 2, \dots, n$.

For the children of $F_{j,k}$, the analog of Θ_j is $\left(\Theta_j'' - \frac{\theta_j}{2} \right) + \left(\Theta_{j,k-1} + \frac{\theta_{j,k}}{2} \right)$.

The last step is to apply the hyperbolic translation of vector $\overrightarrow{OP_j}$. The formula (2.10) tells us that the analytic expression of an hyperbolic translation of vector b is, for all complex z such that $|z| < 1$,

$$t_b(z) = \frac{z + b}{1 + \bar{b}z}. \quad (3.3)$$

Let $p_j = X_{P_j} + iY_{P_j}$ (respectively $f_{j,k} = X_{F_{j,k}} + iY_{F_{j,k}}$) be the affix of the father P_j (for each j) in the system centered at O (respectively be the affix of the child $F_{j,k}$ in the system centered at P_j). Then, the affix $z_{F_{j,k}}$ of the child $F_{j,k}$ in the system centered at O is given by

$$z_{F_{j,k}} = \frac{f_{j,k} + p_j}{1 + \bar{p}_j f_{j,k}}. \quad (3.4)$$

Therefore, recursively, we can lay out every node of the tree in the hyperbolic plane represented by the Poincaré disk.

3.2.3 Angular sector optimization

Such a layout does not take advantage of the possible space optimization suggested by the use of the hyperbolic space. More precisely, the inherent properties of the hyperbolic space make it possible to optimize the angular sector allocated to the father node and to be shared between its children. Indeed, the allocation and the share of the angular sector are directed by a specific constraint: Ensure that two subtrees coming from two different branches will never cross. To guarantee this constraint in Euclidean geometry, we use the property that two parallel lines never do cross. This property is used to clearly delimit different angular sectors (*cf.* fig. 14). The angular sector to share between the children of P is thus the same as the angular sector given to P .

In the Poincaré model, the space is limited to the unit disk. This limitation gives us a first optimization for the angle. Instead of taking the dashed line as in Euclidean geometry, we could take the angular sector in full line (with vertice P_1) (*cf.* fig. 15). Moreover, the geodesics that do not contain O are circle arcs orthogonal to the unit circle in hyperbolic geometry (*cf.* fig. 16). We will calculate the optimal angular sector θ' to give to the children of P , which own the angular sector θ .

Let :

- l the Euclidean length between the origin O and P ($l < 1$);
- Q the intersection between (OP) and the unit circle;

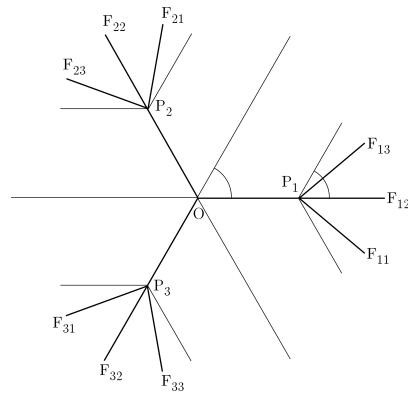


Figure 14: Shared sectors in Euclidean space

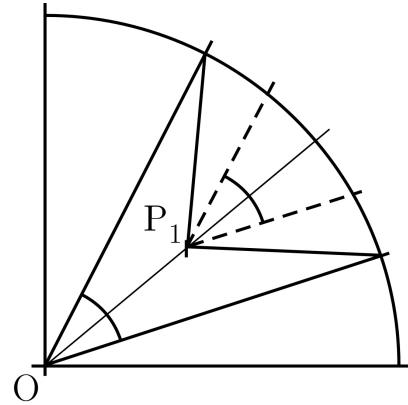
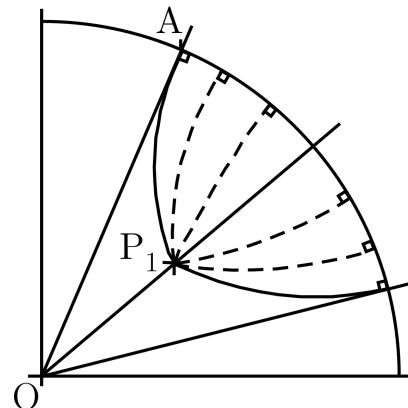


Figure 15: Optimization of the angle in a bounded space

Figure 16: Several different geodesics parallel to (OA) in the hyperbolic space

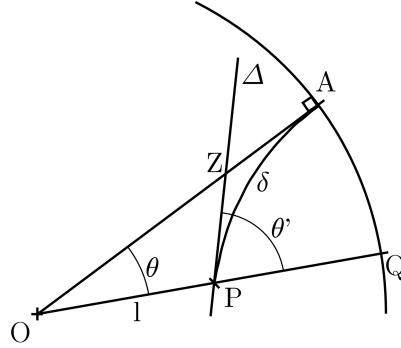


Figure 17: Optimisation of the angular sector

- $[OA]$ the ray of the unit circle that has angle θ with the ray $[OP]$, A belonging to the unit circle;
- δ the hyperbolic geodesic going through A and P ;
- Δ the tangent to δ in P ;
- Z the intersection between $[OA]$ and Δ .

We can see that the desired angle θ' is also the angle between the ray (OP) and the line (PZ) (cf. fig. 17). We can also see that since the geodesic δ is a circle arc orthogonal to the unit circle (cf. paragraph 2.3.2), the ray (OA) is tangent to δ . Thus the lines (ZA) and (ZP) are tangent to the circle arc δ , respectively in A and P . Then we have $|ZP| = |ZA|$. Let $|OZ| = \lambda$. We have $|ZA| = |OA| - |OZ| = 1 - \lambda = |ZP|$.

We are going to use several times Al Kashi's formula [5] in the triangle OPZ . First of all, we want to find the angle θ' . We obtain

$$OZ^2 = OP^2 + PZ^2 - 2|OP||PZ|\cos(\pi - \theta').$$

Since $|OZ| = \lambda$, $|OP| = l$, $|PZ| = 1 - \lambda$ and $\cos(\pi - \theta') = -\cos\theta'$. Then, by substituting in the last formula,

$$\lambda^2 = l^2 + (1 - \lambda)^2 + 2l(1 - \lambda)\cos\theta',$$

which gives, isolating $\cos\theta'$,

$$\cos\theta' = \frac{2\lambda - l^2 - 1}{2l(1 - \lambda)} \quad (3.5)$$

We need now to express λ as a function of l and $\cos\theta$. Applying Al Kashi's formula [5] in the triangle OPZ , we obtain

$$PZ^2 = OP^2 + OZ^2 - 2|OP||OZ|\cos\theta.$$

Similarly, substituting with notations above, we obtain

$$(1 - \lambda)^2 = l^2 + \lambda^2 - 2l\lambda\cos\theta.$$

Thus, resolving in λ ,

$$\lambda = \frac{1 - l^2}{2(1 - l\cos\theta)}.$$

By substituting in (3.5), we obtain

$$\cos\theta' = \frac{(1 + l^2)\cos\theta - 2l}{(1 + l^2) - 2l\cos\theta}. \quad (3.6)$$

Remark 3.1 There is another way to calculate the new angular sector. We can see that the measure of θ' does not depend of the angle between (OP) and the abscissa axis. Thus, we can consider that P is on the abscissa axis, at a length l of O . Let (OA) be a ray of the unit circle doing an angle θ with the abscissa axis. The point A has an affix equal to $e^{i\theta}$. To delimit the angular sector of P , we can consider a geodesic which is parallel to (OA) , passing through P . But, following the rules of the hyperbolic geometry (*cf.* section 1), there is an infinity of such geodesics. In the limit, the geodesic parallel to (OA) which maximizes the angle allocated to P crosses (OA) at infinity, that is, in the Poincaré model, on the unit circle (in A). This geodesic delimits the maximal angular sector for the children of P (*cf.* fig. 16). The hyperbolic translation of coordinates $(-l, 0)$ translates P at the origin. Let A' be the image of A by this translation. Thanks to the Theorem 2.10, A' is on the unit circle. The image of the geodesic (PA) is the diameter (OA') . Then, since the hyperbolic translation is a conformal transformation (*cf.* Remark 2.2), the hyperbolic angle between the abscissa axis and the geodesic (PA) is the same as between the abscissa axis and the geodesic (OA') . This angle is the angle θ' we are seeking. Since (OA') is a diameter, A' has an affix equals to $e^{i\theta'}$. Thus we obtain

$$e^{i\theta'} = t_{-l}(e^{i\theta}) = \frac{e^{i\theta} - l}{1 - le^{i\theta}}. \quad (3.7)$$

A straightforward calculation shows that the two formulas (3.6) and (3.7) are the same.

3.2.4 Optimization relative to the number of children

In general, the trees that we consider are seldom balanced. It is possible, for example, that some nodes have more children than others. Then, it may be interesting to adapt the layout of tree nodes to the number of children, especially the allocated angle and the father-child length. Indeed, if a father node has more children than its neighbours, it will need more room to layout all its children than a node with few or no child. In the same way, if a father node has many children, a bigger father-child length leaves more space between the children.

The first adaptation concerns the angular sector allocated. The main idea is to give a weight to each node. This weight is used to allocate to each node an angular sector proportional to the quotient of its weight by the sum of its brothers's weights. For example, let us suppose that the node P has an angular sector θ to share between its n children F_1, \dots, F_n . Each child F_j has a weight w_j . The angular sector θ_j allocated to the child F_j is

$$\theta_j = \theta \frac{w_j}{\sum_{k=1}^n w_k}.$$

Concerning the calculation of the weight w of the node P , we can remark that a node with only one child should receive the same angular sector as a node without child. This leads us to the following formula:

- $w = 1$ for leaf nodes;
- $w_P = 1 + \ln \left(\sum_{F \text{ child of } P} w_F \right)$ for others.

We can see that this formula is recursive.

Concerning the length optimization, we use an *ad-hoc* optimization. This father-child length $R(n)$, function of the number n of children, has to be between 0 and 1, since it is the Euclidean length between the father and its children when the father is at the center of the unit disk. Arbitrarily, this length for a unique child is equals to $R(1) = 0.3$. The optimization of this length $R(n)$ should answer to two opposite constraints. On the one hand, increasing the space between the children allows for a better legibility. On the other hand, a length too big squashes the data on the disk border. We have to find a trade-off between those two constraints.

For a large number of children (for example, more than 50), the first constraint is stronger: The legibility between the children is crucial. So, for n big enough, $R(n)$ should be as big as

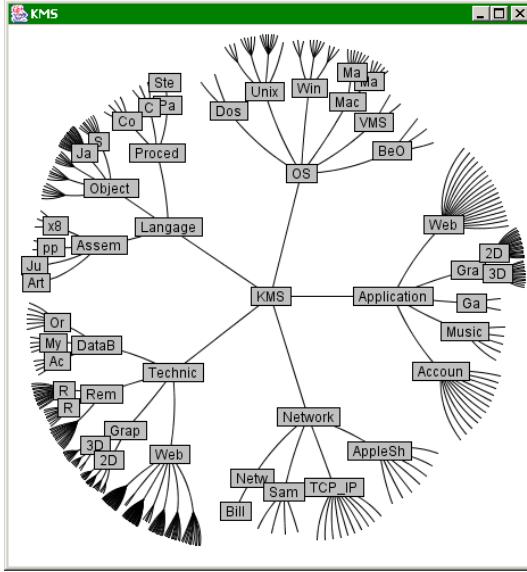


Figure 18: Hypertree in initial position

possible. However, the second constraint implies a length smaller than 1 for $R(\infty)$. For example, $R(\infty) = 0.95$.

For a small number of children, the second constraint wins. The legibility mainly results from the zoom effect of the disk center area. Consequently, $R(n)$ should increase progressively with n , starting from $R(1)$. We give here a function R that satisfies those conditions.

$$R(n) = R(1) + (R(\infty) - R(1)) \cos \frac{10\pi}{n+19}.$$

We could have chosen a linear interpolation between $R(1)$ and $R(\infty) \cong R(50)$, but instead of choosing a peculiar value to represent the infinity, we have preferred to use a trigonometric function which spaces out the length.

3.3 Interactions

The visualization of arborescent data in the Poincaré model induces a zoomed view on a portion of the data, while still keeping visible the remainder of the tree. In order for such a visualization to be really useful, the user should be able to select any part of the data to be zoomed. The user should be able to navigate into the data.

Hypertrees allow two kinds of user interactions. First, the user can select a node to zoom it. The tree will automatically move in order to put the selected node in the center of the unit disk. The user can also drag dynamically the whole tree to zoom to any part of it. Those two kinds of user interactions allow for complementary ways of navigation. The user can successively select the nodes he is interested in, letting the animation move from node to node. Or he can choose a looser navigation, by moving the tree himself (*cf. fig. 18 and fig. 19*).

In the two cases, the motion algorithm is the same. The point to move is given, only the destination point changes: The center of the disk in the first case, the user specified spot in the second case. As the first case can clearly and simply be reduced to the second case, we explain only the second case here.

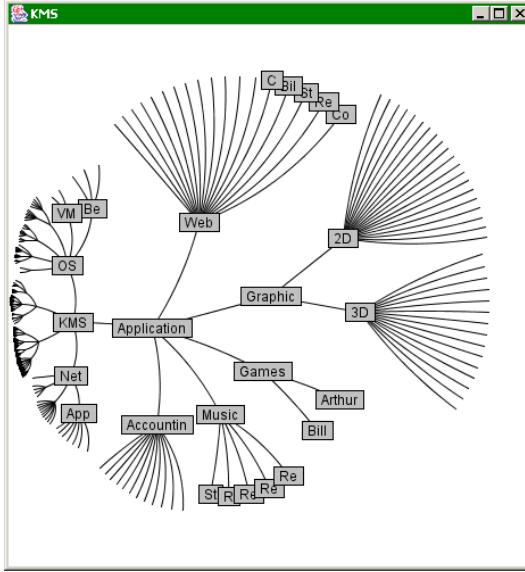


Figure 19: Hypertree after a move

We should find the hyperbolic transformation which brings the chosen point to the specific spot. Unfortunately, a mere hyperbolic translation is not suitable. Indeed, during the navigation, the motion algorithm is used many times. But we have seen that the composition of two hyperbolic translations is not in general an hyperbolic translation (*cf.* Remark 2.4) but a hyperbolic transformation that can be seen as the composition of a hyperbolic rotation centered at O (eventually of a null angle) followed by a hyperbolic translation. This means that the composition of several hyperbolic translations could well rotate the tree around the center of the disk. The user could be confused by such an effect, since rotating something only with translations is impossible in Euclidean geometry. This rotation effect is against user's common sense, and can hinder its navigation.

To avoid this rotation, every move is calculated from the original position of the tree, when its root is at the origin. In this way, there is always only one translation applied. This also restrains the propagation of calculation imprecisions.

The motion algorithm used in user interactions is the following. Let us suppose that the tree is in whatever position, with:

- z_R the affix of R , the root of the tree;
- z_S the affix of S , the user start point;
- z_E the affix of E , the spot selected as the end point.

We are looking for the hyperbolic transformation $\tau_{\theta,p}$ (*cf.* relation (2.12)) such that $\tau_{\theta,p}(z_S) = z_E$. This hyperbolic transformation should be equivalent to a single hyperbolic translation t_b applied when the tree is in the original position, when R is at the center of the disk. Let t_{-z_R} (*cf.* relation (2.10)) be the hyperbolic translation that put back R at the center of the disk, and z'_S the affix of S in this position: $z'_S = t_{-z_R}(z_S)$. We should find t_b so that $t_b(z'_S) = z_E$. Since $t_b(z) = \frac{z+b}{1+\bar{b}z}$, thanks to (2.10), we have to find b so that

$$z_E = \frac{z'_S + b}{1 + \bar{b}z'_S}.$$

After simplification, we obtain

$$b = \frac{z_E (1 - |z'_S|^2) - z'_S (1 - |z_E|^2)}{1 - |z_E|^2 |z'_S|^2}.$$

The researched hyperbolic transformation $\tau_{\theta,p}$ is thus the composition of two hyperbolic translations t_{-z_R} and t_b . Thanks to the proof of Theorem 2.10, we have, for $\tau_{\theta,p} = t_b \circ t_{-z_R}$,

$$\theta = \frac{1 - \bar{z}_R b}{1 - z_R \bar{b}}$$

and

$$p = \frac{b - z_R}{1 - z_R \bar{b}},$$

thus, for all $z \in B^2$,

$$\tau_{\theta,p}(z) = \frac{(1 - \bar{z}_R b) z + (b - z_R)}{(1 - \bar{z}_R b) + (b - z_R) z},$$

with

$$b = \frac{z_E (1 - |z'_S|^2) - z'_S (1 - |z_E|^2)}{1 - |z_E|^2 |z'_S|^2}.$$

4 Discussion

In this paper we have seen how the hyperbolic geometry, and especially the Poincaré model, can be used to create an interactive *focus+context* visualization for arborescent data: The *hypertree* visualization.

This visualization, invented by Lamping et Rao [6, 7], was only available as a commercial product. Our goal in this article was to give the adequate tools to recreate such a visualization. We have followed a strict scientific path, starting from the mathematical foundations to obtain the formulas and algorithms used to implement the visualization. The result of this article is a free and open-source *hypertree* application available on the web: <http://hypertree.sf.net>.

The *hypertree* visualization is not perfect. Compared to others kind of visualization like the *treemap* [12], a lot of space is lost in the *hypertree*, mainly around the disk and in the center, especially when the tree is not well balanced. Moreover, on the verge of the disk, even if the global structure of the tree is visible, the nodes themselves cannot be distinguished. Thus, it is impossible to compare in one sight every node of the tree without moving it. Such a thing is possible with the *treemap*. On the contrary, the *hypertree* is a very good visualization when it comes to hierarchy navigation and data mining, things that are hard to do with the *treemap*. The *hypertree* and the *treemap* visualizations are complementary. This is the reason why we have also made a free and open-source *treemap* visualization available on the web: <http://treemap.sf.net>.

The *hypertree* visualization is based on the 2-dimensional hyperbolic space. Munzner [9] has worked on extending the visualization to the 3-dimensional hyperbolic space. The main problem was then to layout the children of a node. In 2D, the children are points, placed on a circle. In 3D, she has chosen circles for children, placed on the surface of a sphere; But packing circles on a sphere is an eminently hard problem. Munzner has circumvented this with a non-optimal layered approach. Moreover, she has used the less intuitive projective model, and the projecting sphere (equivalent to the 2D unit disk) is seen from the outside.

We are working on another way of using the 3D hyperbolic space. To really extend the work we made in 2D, we use the more intuitive conformal model. Moreover, the user is immersed in the hyperbolic space, seeing it from the center of the projective sphere. Finally, we use the *treemap* slicing algorithm to solve the children layout problem. This work will be described in [1].

References

- [1] BERGÉ B. and BOUTHIER C. 3D-Hyperbolic Tree Visualization: Mathematical Challenges and Algorithmic Solutions. In preparation, 2003.
- [2] BOLYAI J. Appendix, scientiam spatii absolute veram exhibens. In *Tentamen juventutem studiosam in elementa matheseos purae (...) introducenti*, 1832.
- [3] HENLE M. *Modern Geometries*. Prentice Hall, 2001.
- [4] HERMAN I., MELANÇON G. and MARSHALL M.S. Graph Visualization and Navigation in Information Visualization: a Survey. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):24–43, 2000.
- [5] LALESKO T. *La géométrie du triangle, 2^e édition*. Jacques Gabay Éd., 1987.
- [6] LAMPING J. and RAO R. The Hyperbolic Browser: A *focus+context* Technique for Visualizing Large Hierarchies. *Journal of Visual Languages and Computing*, 7(1):33–55, 1996.
- [7] LAMPING J., RAO R. and PIROLI P. A *focus+context* Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI-95)*, pages 401–408, Denver, 1995. ACM Press.
- [8] LOBATCHEWSKY N. Géométrie imaginaire. *J. Reine Angew. Math.*, 17:295–320, 1837. originally published at Kazan in 1836.
- [9] MUNZNER T. H3: Laying Out Large Directed Graphs in 3D Hyperbolic Space. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis-97)*, Phoenix, 1997. IEEE.
- [10] RATCLIFFE J. G. *Foundations of Hyperbolic Manifolds*, volume 149 of *Graduate Texts in Mathematics*. Springer, 1994.
- [11] RIEMANN B. Über die Hypothesen, welche der Geometrie zu Grunde liegen. *Abh. Ges. Wiss. Göttingen*, 13:133–152, 1867.
- [12] SHNEIDERMAN B. Tree Visualization with Tree-Maps: A 2-D Space-Filling Approach. *ACM Transactions on Computer-Human Interaction*, 11(1):92–99, 1992.
- [13] SHNEIDERMAN B. *Designing the User Interface, 3rd edition*. Addison Wesley, 1997.