Spherical Geometry

Eric Lehman

february-april 2012

Table of content

Introduction

The aim of this course is to show different aspects of spherical geometry for itself, in relation to applications and in relation to other geometries and other parts of mathematics. The chapters will be (mostly) independant from each other.

To begin, we'l work on the sphere as Euclid did in the plane looking at triangles. Many things look alike, but there are some striking differences. The second viewpoint will be the introduction of coordinates and the application to basic astronomy. The theorem of Pythagoras has a very nice and simple shape in spherical geometry. To contemplate spherical trigonometry will give us respect for our ancestors and navigators, but we shall skip the computations! and let the GPS do them. The stereographic projection is a marvellous tool to understand the pencils of coaxial circles and many aspects of the relation between the spherical geometry, the euclidean affine plane, the complex projective line, the real projective plane, the Möbius strip and even the hyperbolic plane.

cf. http ://math.rice.edu/ pcmi/sphere/

Chapter 1

Spherical biangles and spherical triangles

§ 1. The unit sphere

We denote the usual Euclidean 3-dimensional space by P.

Definition. Given a point O in P and a real number r, we call sphere centered at O with radius r the subset S of P whose elements are the points M such that the distance MO is equal to r.

Let us take an orthonormal frame $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$. The point $M(x, y, z)$ is such that

$$\overrightarrow{OM} = x\vec{\imath} + y\vec{\jmath} + z\vec{k}$$

The distance OM is equal to $\|\overrightarrow{OM}\| = \sqrt{\overrightarrow{OM} \cdot \overrightarrow{OM}} = \sqrt{x^2 + y^2 + z^2}$, where $\cdot$ denotes the scalar product (we can also get that result using Pythagoras' theorem twice).

As a consequence, we see that M belongs to S if and only if

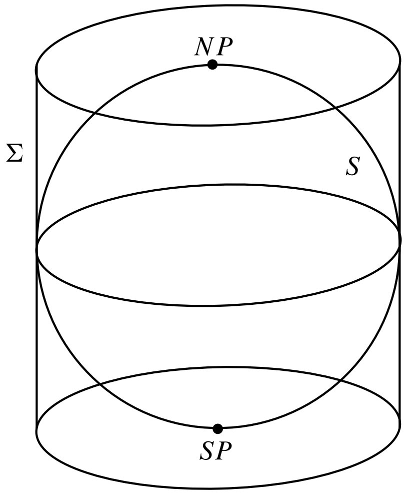
$$x^2 + y^2 + z^2 = r^2 \qquad (1)$$

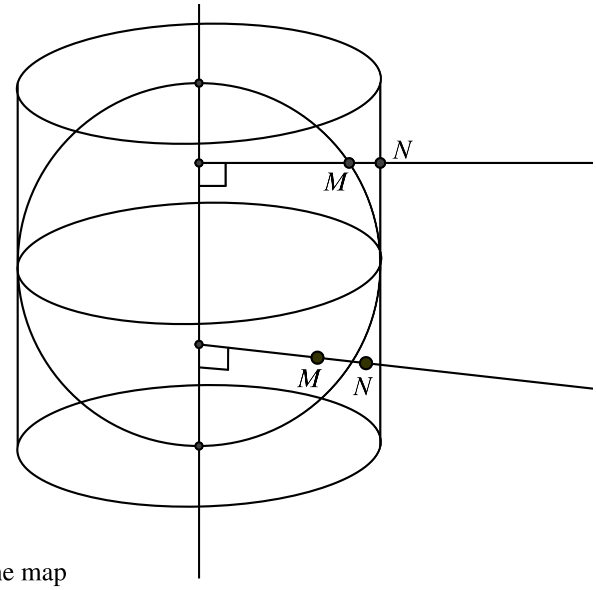
We say that (1) is the equation of the sphere. We call North Pole the point $NP(0, 0, r)$ and South Pole the point $SP(0, 0, -r)$.

1.1 Area of the sphere (Archimedes)

Definition. Given the sphere S of equation (1), we call *circumscribed cylinder* Σ the set of points $N(x, y, z)$ such that

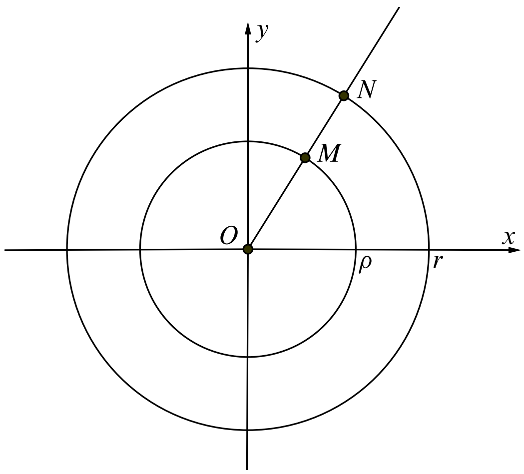
$$\begin{cases} x^2 + y^2 = r^2 \\ \text{and} \\ -r < z < r \end{cases}$$





Definition. The Lambert projection is the map

$$\Lambda : S \setminus \{NP, SP\} \longrightarrow \Sigma, M(x, y, z) \mapsto N\left(\frac{xr}{\sqrt{x^2 + y^2}}, \frac{yr}{\sqrt{x^2 + y^2}}, z\right)$$



$$\frac{x_N}{x_M} = \frac{y_N}{y_M} = \frac{r}{\rho}$$

$$\rho^2 = x_M^2 + y_M^2$$

Proposition. The Lambert projection is bijective and area-preserving.
Proof. To prove bijectivity, we have to prove that the system in (x_M, y_M, z_M)

$$\begin{cases} \dfrac{x_M r}{\sqrt{x_M^2+y_M^2}} = x_N \\ \dfrac{y_M r}{\sqrt{x_M^2+y_M^2}} = y_N \\ z_M = z_N \\ x_M^2 + y_M^2 + z_M^2 = r^2 \end{cases}$$

has a unique solution when $x_N^2 + y_N^2 = r^2$ and $-r < z_N < r$. The computation is simple and the geometrical proof is even simpler. The solution is

$$\begin{cases} x_M = \dfrac{\sqrt{r^2-z_N^2}}{r} x_N \\ y_M = \dfrac{\sqrt{r^2-z_N^2}}{r} y_N \\ z_M = z_N \end{cases}$$

To show the preservation of areas, we follow Archimedes (with modern notations !) : consider the circle with latitude θ and take a small piece $\mathrm{d}h_S \times \mathrm{d}\ell_S$ where $\mathrm{d}h_S$ is along that circle and $\mathrm{d}\ell_S$ is orthogonal to that circle. Let us make the Lambert projection : we get $\mathrm{d}h_\Sigma \times \mathrm{d}\ell_\Sigma$, where $\mathrm{d}h_S = \mathrm{d}h_\Sigma \cos\theta$ and $\mathrm{d}\ell_\Sigma = \mathrm{d}\ell_S \cos\theta$. Thus $\mathrm{d}h_\Sigma \times \mathrm{d}\ell_\Sigma = \mathrm{d}h_S \times \mathrm{d}\ell_S$. □

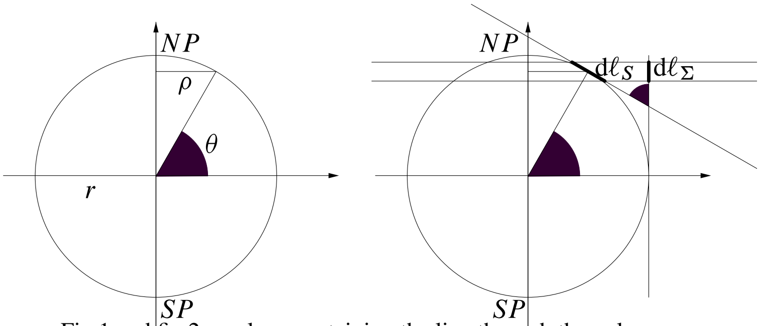


Fig 1 and fig 2 : a plane containing the line through the poles

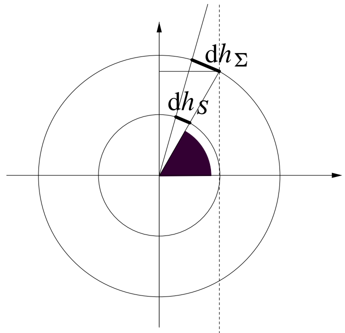


Fig 3 : the plane $z = z_M$ seen from above

Corollary. The area of the sphere is $4\pi r^2$.
Proof. It is possible to unfold the cylinder on a plane getting a rectangle. One side has length the height $2r$ of the cylinder ; the other side has the length of the circle at the basis of the cylinder, that is $2\pi r$. Thus the area of Σ is $4\pi r^2$ and since the Lambert projection preserves areas, the sphere S has the same area. □
Remark 1. From the area, it is easy to get the volume

$$V_S = \int_0^r 4\pi r^2 \,\mathrm{d}r = \frac{4}{3}\pi r^3$$

** **Exercise 1.** Show that it is not possible to cover completely a round hole whose diameter is **1** meter, with **9** (rectangular) planks having the same breadth **10** cm and any length greater than **1** meter.

Remark 2. Horizontal stripes of a sphere with equal differences of altitudes have same area. More precisely : let a stripe $S(z_1, z_2)$ with $-r \leqslant z_1 < z_2 \leqslant r$ denote the set of points (x, y, z) of the sphere (center O and radius r) such that $z_1 \leqslant z \leqslant z_2$. Then

$$z_2 - z_1 = z'_2 - z'_1 \implies \text{area of } S(z_1, z_2) = \text{area of } S(z'_1, z'_2)$$

1.2 Solid angles

Recall that to characterize a plane angle A you look at the arc a of a circle centered at the vertex O of the angle and of radius 1. The measure of A is the length of a. The corresponding unit for measuring angles is then the *radian* : 1 radian is the measure of an angle U such that the corresponding arc u has length 1. Thus a full turn measures 2π radians, a flat angle measures π radians and a right angle $\frac{\pi}{2}$ radians.

Definition. A *solid angle* (or *space angle*) Ω is the union of half-lines or rays having all the same initial point O (we require also $\Omega \smallsetminus \{O\}$ to be connected). The intersection ω of Ω with the sphere S with center O and radius 1, caracterize the solid angle Ω. The measure of Ω is the area of ω. The corresponding unit for measuring solid angles is then the *steradian* (denoted sr) : 1 steradian is the measure of a solid angle W such that the corresponding surface w has an area equal to 1. Thus the whole space measures 4π steradians, a half-space measures 2π steradians, a quarter of the whole space (between 2 orthogonal half-spaces with common edge going through O) measures π steradians and an octant (the first octant is the set of all points such that $x \geqslant 0$, $y \geqslant 0$ and $z \geqslant 0$) measures $\frac{\pi}{2}$ steradians.

1.3 Intersections of the sphere S with planes

Proposition. The intersection of a sphere with a plane is a circle (a point is a circle with radius 0) or empty. More precisely, let S be a sphere with center O and radius r, let P be a plane and C the orthogonal projection of O on P and put $d = OC$. Then if $d \leqslant r$, $\mathcal{C} = S \cap P$ is the circle in the plane P, with center C and radius $\sqrt{r^2 - d^2}$ and if $d > r$, $\mathcal{C} = S \cap P = \varnothing$.

Definitions. A *great circle* of a sphere S is an intersection of S with a plane containing the center O of the sphere.

Two points on the sphere are *antipodal* or *opposite relatively to the center* if the line joining them goes through the center of the sphere.

The great circles of a sphere are the circles included in the sphere which are centered at the center O of the sphere.

A circle of S which is not a great circle is called a *little circle*.

Exercise 2. Let $\mathcal{C}$ be a circle included in a sphere S. Can we know the plane of that circle even when the radius of the circle is 0 ?

Exercise 3. Let A and A' be two antipodal points on a sphere with center O. Show that O is the middle of the segment AA'.

Exercise 4. Show that if a circle in S contains two opposite points, then it is a great circle.

Exercise 5. Show that the intersection of two great circles is a pair of antipodal points.

§ 2. Biangles

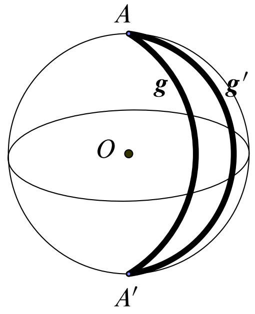
From now on, the sphere S has center O and radius 1 and all circles are circles included in the surface S (unless special specification).

Definition. Let A and A' be two antipodal points and g and g' two half-circles with end-points A and A'. The part B of S between g and g', included in one half-space, is called a *biangle*. The *edges* of the biangle B are g and g', its *vertices* are A and A'.

The biangle B has two angles : one angle α at the vertex A and one α' at the vertex A'. The angle α is the angle formed by the two half-lines with end point A tangent to g and g'. The angle α' is defined the same way at A'. Since the half-tangents to g at A and A' are parallel and with the same direction and since the same holds at A', we see that the angles α and α' are equal

$$\alpha = \alpha'$$

Therefor, this angle α is called the angle of the biangle.



Proposition. The area of a biangle is twice its angle.
Remark. A biangle B may be looked as the intersection of a solid angle Ω with the sphere S. The area of the biangle B is then the measure ω_B in steradians of the solid angle Ω. On the other hand, the angle α_B of B is a plane angle and if it is measured in radians, we have

$$\omega_B = 2\alpha_B$$

Notice that we could wright the formula : $\omega_B = \alpha_A + \alpha_{A'}$.

Exercise 6. Describe the biggest possible biangle (on S).

Exercise 7. Let M an N be two points on S. We suppose that there are 2 distinct arcs g and g' of great circles γ and γ', with endpoints M and N. Show that if M and N are not antipodal, then $\gamma = \gamma'$. What do you think of "biangles" that would not be "equilateral" ?

§ 3. Triangles

3.1 An attempt to formulate a definition

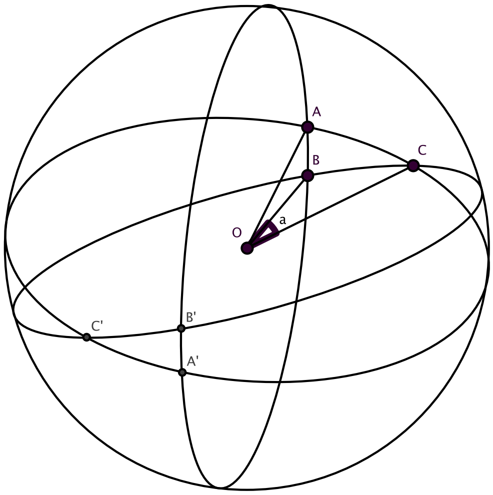
Definition. Let A, B and C be three points on S, which do not belong to a commen great circle. Let a be the smallest arc of a great circle with endpoints B and C, b be the smallest arc of a great circle with endpoints C and A and c be the smallest arc of a great circle with endpoints B and A. The closed curve $\ell = a \cup b \cup c$ divides S into three parts : ℓ, $\mathcal{T}_1$ and $\mathcal{T}_2$. Both subsets $T_1 = \ell \cup \mathcal{T}_1$ and $T_2 = \ell \cup \mathcal{T}_2$ are closed bounded subset of S. One of them is included in a hemisphere, we call it the *little spherical triangle ABC*, the other one contains a hemisphere, we call it the *big spherical triangle ABC*. The simplified denominations "triangle ABC" or "ABC" will be used for "little spherical triangle ABC".

The points A, B and C are the *vertices* or *vertexes* (singular : *vertex*) of the triangle ABC, the arcs a, b and c are the *sides* of ABC. We measure the lengths of the sides in radians since the sides are arcs of circles with radius 1. We keep the same letter for the side and for its measure, thus if we denote the center of S by O :

$$a = \widehat{BOC} \quad ; \quad b = \widehat{COA} \quad ; \quad c = \widehat{AOB}$$

We measure the area of ABC in steradians. Thus :

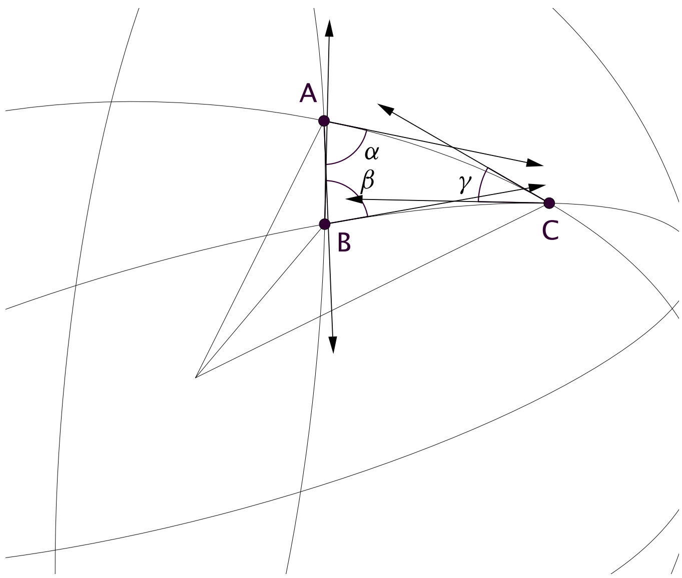
$$\text{area of little } ABC + \text{area of big } ABC = \text{area of } S = 4\pi$$



Remark. If, for instance B and C are antipodal, there is not one smallest arc of a great circle with endpoints B and C, but all such arcs have the same length π and there are infintely many of them. In that case, we have to choose the arc a among all the arcs of great circles with endpoints B and C.

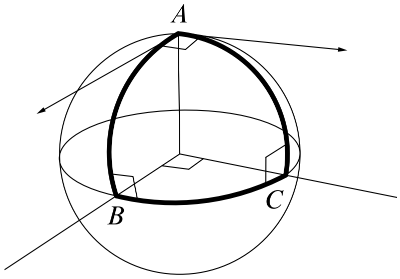
3.2 Angles of a spherical triangle

Definition. Let ABC be little spherical triangle. The plane angle $\widehat{A}$ is the angle between the half-lines tangent to the arcs b and c in A. We call α the measure in radians of $\widehat{A}$. We define in the same way the angles B and C and the measures in radians β and γ of these angles.



Proposition. Let T_2 be the big triangle ABC. The angles of T_2 are $\widehat{A_2}$, $\widehat{B_2}$ and $\widehat{C_2}$ such that $\widehat{A} \cup \widehat{A_2}$ is the tangent plane to S at A, $\widehat{B} \cup \widehat{B_2}$ is the tangent plane to S at B and $\widehat{C} \cup \widehat{C_2}$ is the tangent plane to S at C. The measure α_2 in radians of $\widehat{A_2}$ is $\alpha_2 = 2\pi - \alpha$, the measure β_2 in radians of $\widehat{B_2}$ is $\beta_2 = 2\pi - \beta$ and the measure γ_2 in radians of $\widehat{C_2}$ is $\gamma_2 = 2\pi - \gamma$.

Example. Let A be a point of S and let g be a great circle through A. The point A', antipodal to A, lies on g. Let B be a point on g such that $\widehat{AOB} = \widehat{BOA'}$. We call c the little arc of g joining A and B. Draw the great circles through A and B orthogonal to c. Let C be one of the intersection points of these two great circles. The little triangle ABC has three right angles. To draw the picture, imagine the point A at the North Pole.



Definition. Let $T = ABC$ be a spherical triangle and let α, β and γ be the measures in radians of the angles of the spherical triangle T. The number

$$E = \alpha + \beta + \gamma - \pi$$

is called the *excess* of the triangle T.

Example. In the previous example, the excess is equal to $\frac{\pi}{2}$.

3.3 Girard's theorem

Quotes from Wikipedia, Article by : J J O'Connor & E F Robertson :

Albert Girard (1595, Saint-Mihiel –8 December 1632, Leiden) was French but, being a member of the Reformed church, went as a religious refugee to the Netherlands.

He is the first person who understood the general doctrine of the formation of the coefficients of the powers, from the sums of their roots, and their products, etc.

He is the first who understood the use of negative roots in the solution of geometrical problems.

He is the first who spoke of the imaginary roots, and understood that every equation might have as many roots real and imaginary, and no more, as there are units in the index of the highest power. And he was the first who gave the whimsical name of quantities less than nothing to the negative.

He is the first who discovered the rules for summing the powers of the roots of any equation.

He is also famed for being the first to formulate the (now well-known) inductive definition fn+2 = fn+1 + fn for the Fibonacci sequence, and stating that the ratios of terms of the Fibonacci sequence tend to the golden ratio.

He is the first to have used the symbols sin for sinus, cos for cosine and tan for tangent.

He gave the formula for the area of a spherical triangle, but he was not the first ! It was discovered earlier but not published by the English mathematician Thomas Harriot.

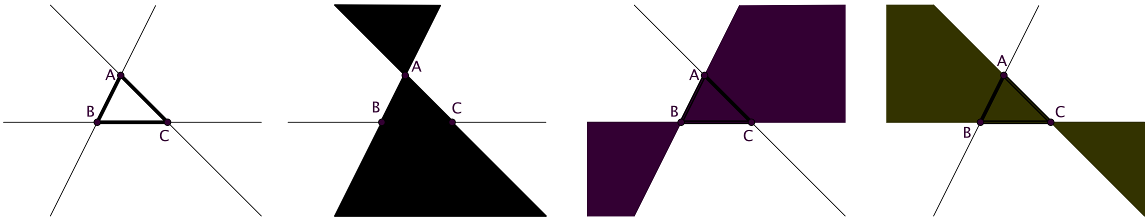
Theorem. The area of a spherical triangle is equal to its excess.

$$\boxed{\text{Area of the spherical triangle } ABC = \text{sum of the plane angles of the spherical triangle} ABC - \pi}$$

or

$$\boxed{S_{ABC} = \alpha + \beta + \gamma - \pi}$$

Proof.



Preliminary remark : we neglet all the boundaries when we use or reason about areas.

Let us call A', B' and C' the points antipodal respectively to A, B and C. We may first notice the triangle $\mathcal{T}' = A'B'C'$ is the image of the triangle $\mathcal{T} = ABC$ by the symmetry relatively to the center O of the sphere S. This symmetry is an isometry, thus the areas are equal

$$S_{\mathcal{T}'} = S_{\mathcal{T}}$$

Let us call g_a the great circle containing B and C (or the side a), g_b the great circle containing C and A (or the side b) and g_c the great circle containing A and B (or the side c). The great circles g_b and g_a determine two antipodal biangles with angle of measure α. Let us call $\mathcal{A}$ the union of these two biangles. Each biangle has an area equal to 2α. The total area of $\mathcal{A}$ is then

$$S_{\mathcal{A}} = 4\alpha$$

In the same way, we define $\mathcal{B}$ as the union of the two biangles with angle β defined by g_c and g_a and $\mathcal{C}$ as the union of the two biangles with angle γ defined by g_a and g_b. We have

$$S_{\mathcal{B}} = 4\beta \qquad \text{and} \qquad S_{\mathcal{C}} = 4\gamma$$

If we neglect parts of area 0, we have

$$\begin{aligned} \mathcal{T} \cap \mathcal{T}' &= \varnothing \\ \mathcal{B} \cap \mathcal{C} &= \mathcal{T} \cup \mathcal{T}' \\ \mathcal{C} \cap \mathcal{A} &= \mathcal{T} \cup \mathcal{T}' \\ \mathcal{A} \cap \mathcal{B} &= \mathcal{T} \cup \mathcal{T}' \\ \mathcal{A} \cup \mathcal{B} \cup \mathcal{C} &= \text{sphere } S \end{aligned}$$

Thus

$$S_{\mathcal{A}} + S_{\mathcal{B}} + S_{\mathcal{C}} - 2(S_{\mathcal{T}} + S_{\mathcal{T}'}) = 4\pi$$

or

$$4\alpha + 4\beta + 4\gamma - 2(S_{\mathcal{T}} + S_{\mathcal{T}}) = 4\pi$$

Finally

$$\alpha + \beta + \gamma = \pi + S_{\mathcal{T}} \qquad \square$$

Example. If the triangle ABC is trirectangle, we have indeed

$$\frac{\pi}{2} + \frac{\pi}{2} + \frac{\pi}{2} = \pi + \frac{1}{8}(4\pi)$$

Exercise 8. Let ABC be a little spherical triangle on S, with sides a, b and c. Let a' be the arc of great circle with endpoints B and C such that $a \cup a'$ is a great circle. We denote by $\mathcal{T}_{1a}$ and $\mathcal{T}_{2a}$ the spherical triangles with vertices A, B and C and with sides a', b and c. What are the angles of these two triangles ? Does the Girard's theorem apply to $\mathcal{T}_{1a}$ and $\mathcal{T}_{2a}$?

Exercise 9. Let Q be a convex spherical quadrilateral with angles α, β, γ and δ. Determine the area S_Q of that quadrileral. What happens with a "complex" or "self-intersecting" quadrilateral ? Generalize.

Exercise 10. Show that it is not possible to draw a "spherical rectangle", that is a quadrilateral with 4 right angles (*Hint.* Compute the area of such a quadrangle, using exercise 9). How can we define a "spherical square" ?

Let $\mathcal{K}$ be a cube such that all vertices of $\mathcal{K}$ belong to S. Project centrally, with O as center of projection, each point M_K belonging to $\mathcal{K}$ on the sphere S and denote by M_S the image, that is the point in S such that O, M_K and M_S are on a line. The images of the faces of $\mathcal{K}$ are 6 isometric quadrilaterals, which can be called "squares". What are the angles of these "squares" ? What are the areas of these squares ?

3.4 Application : Euler's formula

(We did it last year). Legendre's beautiful proof of Euler's formula is based on Girard's theorem.

Suppose that F triangles make a triangulation of the sphere S ; denote by E the number of edges ; and by V the number of vertices. Then, summing all angles in all triangles, the total angle sum is $2\pi V$ (as all of the angles occur at a vertex without overlap, and the angle sum at any one of the V vertices is exactly 2π). Also, the sum of the areas of the triangles is the area of the sphere, namely 4π ; thus we see that

$$2\pi V = 4\pi + F\pi$$

or

$$V - \frac{F}{2} = 2$$

Now (by counting edges of each triangle and noting that this counts each edge twice), we obtain $3F = 2E$, or $E = \frac{3F}{2}$. Thus we have $F - E + V = F - \frac{3F}{2} + (2 + \frac{F}{2}) = 2$ which is Euler's formula :

$$\boxed{F - E + V = 2}$$

The formula is valid for any convex polyhedron : you project centrally the polyhedron on a circumscribed sphere.

Indications, answers or solutions

Ex 1. Imagine a sphere having as great circle the circle surrounding the given round hole. Let us think the hole as done in a horizontal plane. If a plank covers a part of the hole, consider the part of space between the two vertical planes going through the long sides of the plank. We call *induced* by the plank the part of the sphere between these two planes.
If we had 10 planks, we could put them side by side and cover the hole completely (the whole hole !). But following the remark 2, the parts of the sphere induced by the planks are all of the same area. That means that the area of the sphere induced by one plank is at most one 10th of the total area of the sphere. Thus with 9 planks we get a total induced part with an area equal at most to the $\frac{9}{10}$th of the total sphere. And the orthogonal projection on the plane of the points of the sphere which are not in an induced part will not be covered by any plank.
Ex 2. The plane tangent to the sphere at that point.
Ex 3. Since O is the center of the sphere $OA = OA'$.
Ex 4. The plane containing the circle contains these two antipodal points and thus also there midpoint, that is the center O of the sphere.
Ex 5. The 2 planes containing the great circles contain the center O of the sphere. Therefore theire intersection is a line containing O. This line intersects S in two antipodal points on S.
Ex 6. A hemisphere (measure : 2π).
Ex 7. If the points M and N are not antipodal, then there is only one great circle containing both M and N, because there is only one plane going through the three points O, M and N when they are not on one line.
If a biangle is not equilateral it has to be a hemisphere : take two points M and N which are not antipodal, the great circle through M and N is cut into two arcs of different lengths, defining a biangle with plane angles which are flat angles.
Ex 8. The shapes of $\mathcal{T}_{1a}$ and $\mathcal{T}_{2a}$ are the hemisphere containing A with boundary the great circle containing B and C from which is taken away the little triangle ABC OR the hemisphere not containing A, with boundary the great circle containing B and C, to which is added the little triangle ABC. The angles of $\mathcal{T}_{1a}$ are

$$\alpha_1 = 2\pi - \alpha \quad ; \quad \beta_1 = \pi - \beta \quad ; \quad \gamma_1 = \pi - \gamma$$

Thus the sum is
$$\begin{aligned}\alpha_1 + \beta_1 + \gamma_1 &= 4\pi - (\alpha + \beta + \gamma) = 4\pi - (\pi + S_{\mathcal{T}})\\ &= \pi + (2\pi - S_{\mathcal{T}})\end{aligned}$$
The area of $\mathcal{T}_{1a}$ is $\frac{1}{2}(4\pi) - S_{\mathcal{T}} = 2\pi - S_{\mathcal{T}}$. Thus we see that Girard's theorem is still valid.
The angles of $\mathcal{T}_{2a}$ are

$$\alpha_2 = \alpha \quad ; \quad \beta_2 = \pi + \beta \quad ; \quad \gamma_2 = \pi + \gamma$$

Thus the sum is
$$\begin{aligned}\alpha_2 + \beta_2 + \gamma_2 &= 2\pi + (\alpha + \beta + \gamma) = 2\pi + (\pi + S_{\mathcal{T}})\\ &= \pi + (2\pi + S_{\mathcal{T}})\end{aligned}$$
The area of $\mathcal{T}_{2a}$ is $\frac{1}{2}(4\pi) + S_{\mathcal{T}} = 2\pi + S_{\mathcal{T}}$. Thus we see that Girard's theorem is still valid.
Ex 9. Draw a "spherical" diagonal and apply Girards theorem to each of them and add, you get

$$\alpha + \beta + \gamma + \delta = 2\pi + S_Q$$

If you have a "spherical" polygon P with n sides that do not cross each other, the area S_P of P is related to the sum of its angles by the relation

$$\text{The sum of the (interior) angles of } P = (n-2)\pi + S_P$$

The formula is still valid for a self-intersecting quadrilateral if you introduce the algebraic area.
Ex 10. Suppose there is a "rectangle", quadrilateral with 4 right angles. The area of this rectangle would be $\frac{\pi}{2} + \frac{\pi}{2} + \frac{\pi}{2} + \frac{\pi}{2} - 2\pi = 0$.
$120°$ or $\alpha = \frac{2\pi}{3}$, since at each vertex $3\alpha = 2\pi$(=1complete turn).
Area $= 4\pi : 6$, or Area $= 4\frac{2\pi}{3} - 2\pi = \frac{2\pi}{3}$.

Chapter 2

Coordinates on a sphere

§ 1. Cartesian versus spherical coordinates

There are mainly two coordinate systems used on a sphere : the cartesian coordinates and the spherical coordinates. We take as unit length the radius of the sphere S.

1.1 Cartesian coordinates

We suppose given an orthonormal frame $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$ of the space. The coordinates (x, y, z) of a point M in space are the such that

$$\overrightarrow{OM} = x\vec{\imath} + y\vec{\jmath} + z\vec{k}$$

The point $M(x, y, z)$ belongs to sphere S if and only if

$$x^2 + y^2 + z^2 = 1 \tag{2.1}$$

Advantages. These coordinates are convenient to compute scalar products and norms.

Example : let $M_1(x_1, y_1, z_1)$ and $M_2(x_2, y_2, z_2)$ be two points on S. The "length" $\ell_{M_1 M_2}$ of the arc $\overset{\frown}{M_1 M_2}$ is equal to the angle θ between the vectors $\overrightarrow{OM_1}$ and $\overrightarrow{OM_2}$

$$\cos \ell_{M_1 M_2} = \cos\theta = \overrightarrow{OM_1} \cdot \overrightarrow{OM_2} = x_1 x_2 + y_1 y_2 + z_1 z_2$$

Disadvantages. The coordinates are not independent on the sphere. You must always use the relation 2.1 as a constraint.

1.2 Spherical coordinates

Let $M(x, y, z)$ be a point in space such that $(x, y) \neq (0, 0)$, the spherical coordinates of M are (r, θ, φ) defined in the following way

- $r = OM$ or $\|\overrightarrow{OM}\|$
- θ is the oriented angle $(\overrightarrow{k}, \overrightarrow{OM})$.
- φ is the oriented angle $(\overrightarrow{i}, \overrightarrow{Om})$, where m is the orthogonal projection of M on the plane xOy.

If the point M has coordinates $(0, 0, z)$, where $z > 0$, φ is not defined and $\theta = 0$.

If the point M has coordinates $(0, 0, z)$, where $z < 0$, φ is not defined and $\theta = \pi$.

If the point M has coordinates $(0, 0, 0)$, that is $M = O$, φ and θ are not defined.

The sphere S is characterized by the equation

$$r = 1 \tag{2.2}$$

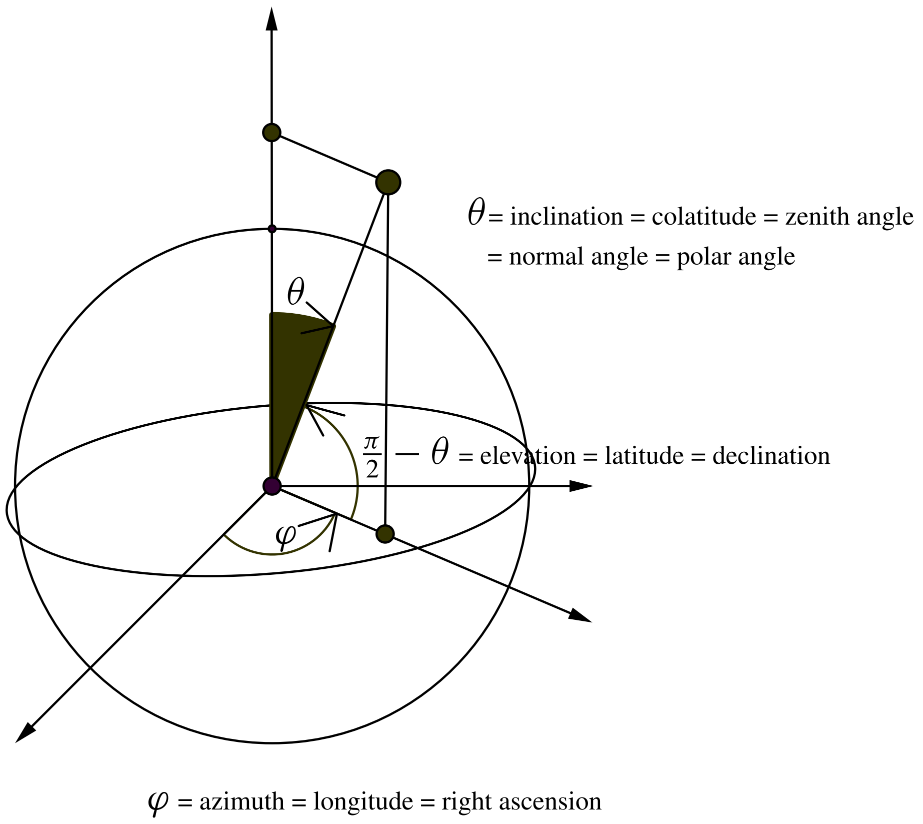
A point M on the sphere and distinct from the points $(0, 0, 1)$ and $(0, 0, -1)$ is characterized by two independent coordinates θ and φ.

The curve with $\theta = \text{Constant} = \theta_0$ is a circle, intersection of S with the plane parallel to xOy which has the equation $z = \cos\theta_0$. The angle θ is calles the *inclination angle*. The *elevation* of M is the angle $\frac{\pi}{2} - \theta$. The inclination may also be called *colatitude*, *zenith angle*, *normal angle*, or *polar angle*.

The curve with $\varphi = \text{Constant} = \varphi_0$ is a half great circle. The great circle is the intersection of S with the plane which has the equation $x\sin\varphi_0 - y\cos\varphi_0 = 0$. The half is such that $y\sin\varphi_0 > 0$ if $\sin\varphi_0 \neq 0$ and $x\cos\varphi_0 > 0$ if $\sin\varphi_0 = 0$. The angle φ is called the *azimuth angle*.

In geography, the elevation and azimuth are called the *latitude* and *longitude*.

In astronomy, the elevation and azimuth are called the *declination* and *right ascension*.



1.3 Change of coordinates

From spherical coordinates to cartesian coordinates.

$$\begin{cases} x = \sin\theta\ \cos\varphi \\ y = \sin\theta\ \sin\varphi \\ z = \cos\theta \end{cases}$$

From cartesian coordinates to spherical coordinates.

$$\begin{cases} \tan\varphi = \frac{y}{x} \\ \qquad \theta = \arccos z \end{cases}$$

But the first relation is not enough to determine φ completely. More precisely : if $y > 0$, then $0 < \varphi < \pi$ (modulo 2π) ; if $y < 0$, then $\pi < \varphi < 2\pi$ (modulo 2π) ; if $y = 0$ and $x > 0$, then $\varphi = 0$ (modulo 2π) and if $y = 0$ and $x < 0$, then $\varphi = \pi$ (modulo 2π).

Exercise 1. Let us define the spherical quadrilateral $Q = ABCD$ in the following way : $A(\theta = \frac{\pi}{4}, \varphi = 0)$, $B(\theta = \frac{\pi}{4}, \varphi = \frac{\pi}{2})$, $C(\theta = \frac{3\pi}{4}, \varphi = \frac{\pi}{2})$, $D(\theta = \frac{3\pi}{4}, \varphi = 0)$ and the sides AB and CD are on the circles $\theta = \frac{\pi}{4}$ and $\theta = \frac{3\pi}{4}$, with $0 \leqslant \varphi \leqslant \frac{\pi}{2}$ and the sides BC and DA are on the half-crcles $\varphi = 0$ and $\varphi = \frac{\pi}{2}$ with $\frac{\pi}{4} \leqslant \theta \leqslant \frac{3\pi}{4}$.

Show that the four angles of Q are right angles. Is Q a spherical square or a spherical rectangular ?

Exercise 2. Let us make the assumption that the Earth is spherical with radius $6\,400$ km. Compute 1°) the distance between Joensuu (Lattitude : 62°35'North, Longitude : 29°45'East) and Caen (49°10'N ; 0°22'West) ; 2°) the distance between Paris(48°49'N,2°19'E) and Christchuch(43°32'S,172°40'E) ; 3°) the distance from one pole to the other. Does this last value have anything to do with the french revolution ?

§ 2. Basic astronomy

Vocabulary :
– Celestial sphere
– Zenith = Strait above your head
– Nadir = Antipodal point to Zenith on the celestial sphere
– Horizon = Great circle through the plane tangent to the Earth at Joensuu
– NCP = North celestial pole intersetion of the axis of rotation of the Earth with the celestial sphere
– SCP = South celestial pole = antipodal to NCP
– Celestial equator = plane through the Earth's equator $\cap$ the celestial sphere
– Polaris = North star ; lies near the NCP
– Celsetial meridian = meridian of the celestial sphere through NCP, Zenith, SCP and Nadir
– Ecliptic = anual path of the sun
– NEP = North ecliptic pole
– SEP = South ecliptic pole
– Vernal equinox = March 20
– Autumnal equinox = September 23
– Summer solstice = June 21
– Winter solstice = December 22

Nice pictures and explanations at the following adress :

http://stars.astro.illinois.edu/celsph.html

Remark. Why are the hands of a watch going round in the order North-East-South-West-North ? Because we see the sun going round in this way and if we look at the shadows of fixed objects they are turning the same way. But in fact it is not the Sun that moves around the Earth, but the Earth is moving around its axis in the other direction. Therefore the astronomers and mathematicians have chosen as *positive* the movement going North-West-South-East-North. But all this has been thought of in the northern hemisphere. South of the equator the sun and the shadows are moving in the "positive" way.

§ 3. Geodesics

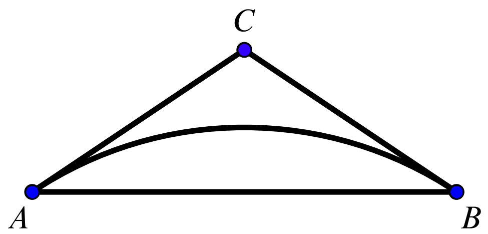
A geodesic is the shortest path between two points. On the sphere S we have to consider a curve on the sphere.

Theorem. Let A and B be two points on the sphere S. The shortest path from A to B on the sphere is the arc of great circle through A and B.

We give two proofs of this theorem. More precisely, the first proof will prove that the shortest arc of a circle with endpoints A and B is an arc of the great circle going through A and B.

3.1 Method based on Archimedes' intuition

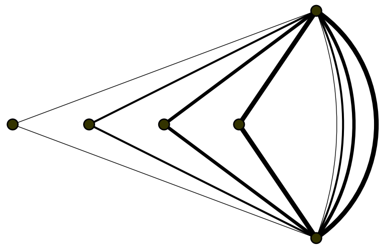
When Archimedes tries to find the length of a circle (that is the value of π), he makes the following drawing



and he argues that if two convex curves from A to B lie inside each other, the shortest one is the one nearest the straight line AB. Thus

$$AB < \overset{\frown}{AB} < AC + CB$$

Proof. We look at all the arcs of circles with endpoints A and B. Following Archimedes, we see that the shortest one will be the one with maximum radius.



The circles are the intersections of the planes containing the line AB and the sphere S. To make the computations, we take a frame such that A is at the "north pole" and B on the main great circle. Thus A has coordinates $(0, 0, 1)$. Let a be the abscissa of the intersection of AB with the axis Ox. The plane through AB will intersect Oy in a point with coordinates $(0, b, 0)$. The equation of the plane P is

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{1} = 1$$

if $a \neq 0$ and $b \neq 0$.

When the parameter a is equal to 0, B is at the south pole $(0, 0, -1)$ and all the circles through A and B are great circles, all with the same length.

The intersection of the plane with Oy do not exist when the plane is parallel to Oy and then the equation is $\frac{x}{a} + \frac{z}{1} = 1$. The parameter b takes the value 0 when the plane is the plane xOz which has the equation $y = 0$. When $b \longrightarrow 0$, the plane through AB tends to the plane xOz.

Let H be the orthogonal projection of O on the plane P. The radius of the circle intersection of P and S is R such that

$$R^2 = 1 - OH^2$$

Let $M(x_0, y_0, z_0)$ be any point in space. The distance from that point to the plane with equation $\alpha x + \beta y + \gamma z + \delta = 0$ is given by

$$\mathrm{dist}(M_0, P) = \frac{|\alpha x_0 + \beta y_0 + \gamma z_0 + \delta|}{\sqrt{\alpha^2 + \beta^2 + \gamma^2}}$$

The length OH is then

$$OH = \frac{1}{\sqrt{1 + \frac{1}{a^2} + \frac{1}{b^2}}}$$

and

$$R^2 = 1 - \frac{1}{1 + \frac{1}{a^2} + \frac{1}{b^2}}$$

The maximum value is obtained for $b = 0$, which means that the circle is a great circle. □

3.2 Method based on differential geometry and variational computation

The theorem is the same. The proof is different. The distance between two infinitely near points $\mathrm{d}\ell$ is such that

$$\mathrm{d}\ell^2 = \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2$$

Thus if we have a curve from $A(\theta_1, \varphi_1)$ to $B(\theta_2, \varphi_2)$ with parametric equation

$$\begin{cases} \theta = \theta(t) \\ \varphi = \varphi(t) \\ t_1 < t < t_2 \end{cases}$$

The length of this curve is

$$\mathcal{L} = \int_{t_1}^{t_2} \sqrt{\theta'(t)^2 + \sin^2\theta(t)\varphi'(t)^2}\ \mathrm{d}t$$

Let us put

$$L = L(\theta, \theta', \varphi, \varphi') = \sqrt{\theta'^2 + \sin^2\theta\varphi'^2}$$

The Euler-Lagrange equations tell us that the $\mathcal{L}$ is extremal when

$$\begin{cases} \dfrac{\partial L}{\partial \theta} - \dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\dfrac{\partial L}{\partial \theta'}\Big) = 0 \\ \dfrac{\partial L}{\partial \varphi} - \dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\dfrac{\partial L}{\partial \varphi'}\Big) = 0 \end{cases}$$

We choose our frame so that A is at the north pole, thus $\theta(t_1) = 0$. The second equation is

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\sin^2\theta\ \varphi'}{L}\Big) = 0$$

Thus

$$\frac{\sin^2\theta(t)\,\varphi'(t)}{L(t)} = \text{Constant}$$

With our choice for A, we get Constant $= 0$, and thus

$$\varphi'(t) = 0$$

or

$$\varphi(t) = \text{Constant}$$

This Constant has to be the longitude of B. Notice that the first equation reduces to the identity $0 = 0$. □

3.3 Proof of Euler-Lagrange equations

Theorem. Let L be a function of four variables. Let $A(x_A, y_A)$ and $B(x_B, y_B)$ be two points in $\mathbb{R}^2$. For any given path or parametrized curve

$$\begin{cases} x = X(t) \\ y = Y(t) \\ t \in [t_1, t_2] \end{cases}$$

such that $X(t_1) = x_A$, $Y(t_1) = y_A$, $X(t_2) = x_B$ and $Y(t_2) = y_B$, we can compute

$$S = \int_{t_1}^{t_2} L(X(t), X'(t), Y(t), Y'(t))\,\mathrm{d}t$$

We suppose all the functions regular enough, at least two times continuously derivable. If a path is such that S is maximal, then

$$\begin{cases} \dfrac{\partial L}{\partial X} - \dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\dfrac{\partial L}{\partial X'}\Big) = 0 \\ \dfrac{\partial L}{\partial Y} - \dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\dfrac{\partial L}{\partial Y'}\Big) = 0 \end{cases}$$

Proof. Let us consider a family of paths from A to B, parametrized by a parameter ϵ.

$$\begin{cases} x = X(t,\epsilon) \\ y = Y(t,\epsilon) \\ t \in [t_1, t_2] \end{cases} \quad \text{and} \quad \begin{cases} X(t_1,\epsilon) = x_A \\ Y(t_1,\epsilon) = y_A \\ X(t_2,\epsilon) = x_B \\ Y(t_2,\epsilon) = y_B \end{cases}$$

Then S becomes a function of ϵ

$$S(\epsilon) = \int_{t_1}^{t_2} L\Big(X(t,\epsilon), \frac{\partial X}{\partial t}(t,\epsilon), Y(t,\epsilon), \frac{\partial Y}{\partial t}(t,\epsilon)\Big)\,\mathrm{d}t$$

The extrema of S are such that $\frac{\mathrm{d}S}{\mathrm{d}\epsilon} = 0$. Regularity of the functions makes it possible to derive under the integration sign. Thus

$$\frac{\mathrm{d}S}{\mathrm{d}\epsilon} = \int_{t_1}^{t_2} \Big\{\frac{\partial L}{\partial X}\frac{\partial X}{\partial \epsilon} + \frac{\partial L}{\partial X'}\frac{\partial X'}{\partial \epsilon} + \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial \epsilon} + \frac{\partial L}{\partial Y'}\frac{\partial Y'}{\partial \epsilon}\Big\}\,\mathrm{d}t \qquad (*)$$

We make the assumption that X and Y have continuous second derivates, thus

$$\frac{\partial^2 X}{\partial\epsilon\partial t} = \frac{\partial^2 X}{\partial t\partial\epsilon} \quad \text{and thus} \quad \frac{\partial X'}{\partial\epsilon} = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial X}{\partial\epsilon}\Big)$$

Integrating by parts, we have

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial X'}\frac{\partial X'}{\partial\epsilon}\,\mathrm{d}t = \int_{t_1}^{t_2} \frac{\partial L}{\partial X'}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial X}{\partial\epsilon}\Big)\,\mathrm{d}t = \Big[\frac{\partial L}{\partial X'}\frac{\partial X}{\partial\epsilon}\Big]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2}\frac{\partial X}{\partial\epsilon}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial X'}\Big)\,\mathrm{d}t$$

But the functions $\epsilon \longmapsto X(t_1,\epsilon)$ and $\epsilon \longmapsto X(t_2,\epsilon)$ are constant functions (equal respectively to x_A and x_B). Thus

$$\frac{\partial X}{\partial\epsilon}(t_1,\epsilon) = \frac{\partial X}{\partial\epsilon}(t_2,\epsilon) = 0$$

and then

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial X'}\frac{\partial X'}{\partial\epsilon}\,\mathrm{d}t = -\int_{t_1}^{t_2}\frac{\partial X}{\partial\epsilon}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial X'}\Big)\,\mathrm{d}t$$

In the same way, we have

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial Y'}\frac{\partial Y'}{\partial\epsilon}\,\mathrm{d}t = -\int_{t_1}^{t_2}\frac{\partial Y}{\partial\epsilon}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial Y'}\Big)\,\mathrm{d}t$$

Then equation $(*)$ becomes

$$\frac{\mathrm{d}S}{\mathrm{d}\epsilon} = \int_{t_1}^{t_2}\Big\{\frac{\partial L}{\partial X}\frac{\partial X}{\partial\epsilon} - \frac{\partial X}{\partial\epsilon}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial X'}\Big) + \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial\epsilon} - \frac{\partial Y}{\partial\epsilon}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial Y'}\Big)\Big\}\,\mathrm{d}t$$

or

$$\frac{\mathrm{d}S}{\mathrm{d}\epsilon} = \int_{t_1}^{t_2}\Big\{\Big[\frac{\partial L}{\partial X} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial X'}\Big)\Big]\frac{\partial X}{\partial\epsilon} + \Big[\frac{\partial L}{\partial Y} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial Y'}\Big)\Big]\frac{\partial Y}{\partial\epsilon}\Big\}\,\mathrm{d}t$$

We want this derivative to be 0 for any choice of the functions $\frac{\partial X}{\partial\epsilon}$ and $\frac{\partial Y}{\partial\epsilon}$. One nice choice would be $\frac{\partial X}{\partial\epsilon} = \frac{\partial L}{\partial X} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial X'}\Big)$ and $\frac{\partial Y}{\partial\epsilon} = \frac{\partial L}{\partial Y} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial Y'}\Big)$. Thus

$$\frac{\partial L}{\partial X} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial X'}\Big) = 0 \quad \text{and} \quad \frac{\partial L}{\partial Y} - \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial Y'}\Big) = 0 \quad .$$

Exercise 3. Show with the help of Euler-Lagrange equations that the (straight) line segment is the shortest path between two points.

Indications, answers or solutions

Ex 1. The tangents to the sides at the vertices are orthogonal since in the direction of parallels or meridians. But Q is not a spherical square or even a spherical rectangular since the sides on the "parallels" are not geodesics.

Ex 2. 1°) The cartesian coordinates of Joensuu

$$\begin{cases} x = \cos(\frac{\pi}{180}(62+\frac{35}{60}))\cos(\frac{\pi}{180}(29+\frac{45}{60})) = 0{,}3999769 \\ y = \cos(\frac{\pi}{180}(62+\frac{35}{60}))\sin(\frac{\pi}{180}(29+\frac{45}{60})) = 0{,}2284869 \\ z = \sin(\frac{\pi}{180}(62+\frac{35}{60})) = 0{,}8876814 \end{cases}$$

The cartesian coordinates of Caen

$$\begin{cases} x = \cos(\frac{\pi}{180}(49+\frac{10}{60}))\cos(-\frac{\pi}{180}(\frac{22}{60})) = 0{,}653847 \\ y = \cos(\frac{\pi}{180}(49+\frac{10}{60}))\sin(-\frac{\pi}{180}(\frac{22}{60})) = -0{,}0041843808 \\ z = \sin(\frac{\pi}{180}(49+\frac{10}{60})) = 0{,}75661478 \end{cases}$$

The scalar product is

$$0{,}3999769 \times 0{,}653847 - 0{,}2284869 \times 0{,}0041843808 + \\ +0{,}8876814 \times 0{,}75661478 = 0{,}932202$$

The corresponding angle is $\arccos(0{,}932202) = 0{,}37$ radians. The corresponding distance is then $0{,}37 \times 6\,400 = 2\,368$ km.

Remark. The precise distance between Joensuu and Caen is $2\,364{,}54$ km.

2°) The precise distance from Paris to Christchuch is $19\,090{,}5$ km.

3°) The precise distance from tne North Pole and the South Pole is $20\,004$ km. At the French revolution, it was decided to give a length unit which would be universal for all humans. The idea was to take the Earth as a basis. Since the Earth has the shape of an ellipsoid with cylindrical symmetry, all the shortest lines from one pole to the other have the same length. The "mètre" would be such that the quarter of a meridian would be $10\,000\,000$m. The measure was not so easy. With the final meter, the length of a quarter of a meridian is about $10\,002$m.

Quoted from Wikipedia :

For the WGS84 ellipsoid the distance from equator to pole is given (in meters) by $10\,001\,965{,}729$*m.*

The World Geodetic System is a standard for use in cartography, geodesy, and navigation. It comprises a standard coordinate frame for the Earth, a standard spheroidal reference surface (the datum or reference ellipsoid) for raw altitude data, and a gravitational equipotential surface (the geoid) that defines the nominal sea level. The latest revision is WGS 84 (dating from 1984 and last revised in 2004), which will be valid up to about 2010. Earlier schemes included WGS 72, WGS 66, and WGS 60. WGS 84 is the reference coordinate system used by the Global Positioning System.

Ex 3. The parametric curve

$$\begin{cases} x = x(t) \\ y = y(t) \\ t \in [t_1, t_2] \end{cases} \quad \text{and} \quad \begin{cases} x(t_1) = x_A \\ y(t_1) = y_A \\ x(t_2) = x_B \\ y(t_2) = y_B \end{cases}$$

describes a path from A to B. The length of this curve is

$$\ell = \int_{t_1}^{t_2} \sqrt{x'^2 + y'^2}\,\mathrm{d}t$$

Let us put $L = \sqrt{x'^2 + y'^2}$. Since L is independant of x and y, the Euler-Lagrange equations become

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial x'}\Big) = 0 \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial y'}\Big) = 0$$

These equations are easy to solve into

$$\frac{x'}{\sqrt{x'^2 + y'^2}} = a \quad \text{and} \quad \frac{y'}{\sqrt{x'^2 + y'^2}} = b$$

where a and b are constants.

If $a = 0$, we get $x' = 0$ and thus $x = \text{Constant} = x_A = x_B$. The function $y(t)$ is arbitrary with the only constraint signy' is constant.

If $b = 0$ same thing as above exchanging x and y.

If $a \neq 0$ and $b \neq 0$, we get

$$\frac{y'}{x'} = \frac{b}{a} \quad \text{thus} \quad y(t) = \frac{b}{a}x(t) + \text{Constant}$$

Notice that it follows from the equations that the sign of $\frac{y'}{x'}$ is constant. Thus the segment is described only once. Notice that the "speed" is not determined.

Chapter 3

Spherical trigonometry

§ 1. Pythagoras' theorem

1.1 Distance

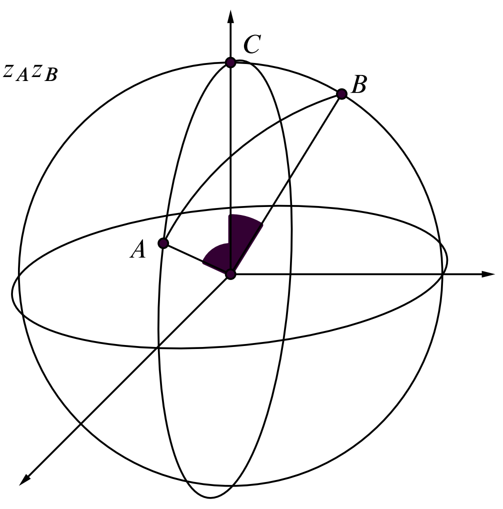
Definition. Let S be a sphere with radius 1, A and B two points belonging to S. The distance (on the the sphere) from A to B denoted $\widehat{AB}$ is the length of the shortest arc of great circle with endpoints A and B.
Remark. If A and B are antipodal, there are infinitely many arcs of great circles with endpoints A and B, but all have the same length which is π, that is the length of a half-circle with radius 1.

Theorem. Let O be the center of a sphere S of radius 1. The distance $c = \widehat{AB}$ between two points A and B is such that

$$\cos c = \overrightarrow{OA} \cdot \overrightarrow{OB} \qquad 0 \leqslant c \leqslant \pi$$

Let $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$ be an orthonormal frame. The sphere S is the surface with equation $x^2 + y^2 + z^2 = 1$. The distance $c = \widehat{AB}$ between two points $A(x_A, y_A, z_A)$ and $B(x_B, y_B, z_B)$ belonging to S is such that

$$\cos c = x_A x_B + y_A y_B + z_A z_B$$



1.2 Basic version

Pythagoras' theorem. Let ABC be a spherical triangle on the sphere S with radius 1. We put

$$c = \overset{\frown}{AB} \qquad a = \overset{\frown}{BC} \qquad b = \overset{\frown}{CA}$$

We suppose that the triangle is orthogonal in C. Then

$$\boxed{\cos c = \cos a \ \cos b}$$

Proof. Choose the frame $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$ orthonormal such that $\vec{k} = \overrightarrow{OC}$, $A \in$half-plane xOz with $x > 0$ and $B \in$half-plane yOz with $y > 0$.

The coordinates of C are $(0, 0, 1)$, those of A are $(\sin a, 0, \cos a)$ and those of B are $(0, \sin b, \cos b)$. Then

$$\cos c = \sin a \times 0 + 0 \times \sin b + \cos a \times \cos b \qquad \square$$

1.3 The case of small rectangular triangles

Let us explain why the theorem above is Pythagoras' theorem. If the triangle is small, that is a, b and c are small, then $\cos a \approx 1 - \frac{1}{2}a^2$ and $\cos b \approx 1 - \frac{1}{2}b^2$ and thus

$$\cos a \ \cos b \approx (1 - \frac{1}{2}a^2)(1 - \frac{1}{2}b^2) \approx 1 - \frac{1}{2}a^2 - \frac{1}{2}b^2$$

Using $\cos c \approx 1 - \frac{1}{2}c^2$, we then get

$$c^2 = a^2 + b^2$$

which is the usual aspect of Pythagoras' theorem.

§ 2. The three laws of spherical trigonometry

Let ABC be a spherical triangle on the sphere S with center O and radius 1. We suppose that the sides are all between 0 and π.

$$\begin{aligned} 0 < a = \overset{\frown}{BC} < \pi \\ 0 < b = \overset{\frown}{CA} < \pi \\ 0 < c = \overset{\frown}{AB} < \pi \end{aligned}$$

We have the same inequalities for the three angles of the spherical triangle ABC

$$\begin{aligned} 0 < \alpha = \text{angle}(\text{plane}OAB, \text{plane}OAC) < \pi \\ 0 < \beta = \text{angle}(\text{plane}OBC, \text{plane}OBA) < \pi \\ 0 < \gamma = \text{angle}(\text{plane}OCA, \text{plane}OCB) < \pi \end{aligned}$$

We are going to show the following three formulae

The cosinus law

$$\boxed{\cos a = \sin b \sin c \cos \alpha + \cos b \cos c}$$

and the two other relations deduced from the former one by circular permutation.

The sinus law

$$\boxed{\frac{\sin \alpha}{\sin a} = \frac{\sin \beta}{\sin b} = \frac{\sin \gamma}{\sin c} = \frac{\text{Volume}(\text{Tetrahedron}\, OABC)}{\sin a \sin b \sin c}}$$

The dual cosinus law

$$\boxed{\cos \alpha = \sin \beta \sin \gamma \cos a - \cos \beta \cos \gamma}$$

2.1 The cosinus law

Let us choose an orthonormal frame $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$ such that $\vec{\imath} = \overrightarrow{OA}$, $\vec{\jmath}$ orthogonal to $\vec{\imath}$, is in the plane OAB with $\vec{\jmath} \cdot \overrightarrow{OB} > 0$ and $\vec{k}$, orthogonal to both $\vec{\imath}$ and $\vec{\jmath}$, such that $\vec{k} \cdot \overrightarrow{OC} > 0$.

Exercise 1. Draw the picture.

The coordinates of A, B and C are then

$$A \left| \begin{array}{l} 1 \\ 0 \\ 0 \end{array} \right. \qquad B \left| \begin{array}{l} \cos c \\ \sin c \\ 0 \end{array} \right. \qquad C \left| \begin{array}{l} \cos b \\ \sin b \cos \alpha \\ \sin b \sin \alpha \end{array} \right.$$

Then the cosinus of the arc $a = \overset{\frown}{BC}$ is then $\overrightarrow{OB} \cdot \overrightarrow{OC}$ thus

$$\cos a = \sin b \sin c \cos \alpha + \cos b \cos c$$

2.2 The sinus law

From the cosinus law, we get

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

and thus

$$\sin^2 \alpha = 1 - \cos^2 \alpha = \frac{\sin^2 b \sin^2 c - \cos^2 a - \cos^2 b \cos^2 c + 2 \cos a \cos b \cos c}{\sin^2 b \sin^2 c}$$

using $\sin^2 b = 1 - \cos^2 b$ and $\sin^2 c = 1 - \cos^2 c$, we get

$$\sin^2 \alpha = \frac{1 - \cos^2 a - \cos^2 b - \cos^2 c + 2 \cos a \cos b \cos c}{\sin^2 b \sin^2 c}$$

or

$$\frac{\sin^2 \alpha}{\sin^2 a} = \frac{1 - \cos^2 a - \cos^2 b - \cos^2 c + 2 \cos a \cos b \cos c}{\sin^2 a \sin^2 b \sin^2 c}$$

and we see that the right hand side of the last equality is invariant by the permutations of the couples (a,α), (b,β) and (c,γ). Thus

$$\frac{\sin^2\alpha}{\sin^2 a} = \frac{\sin^2\beta}{\sin^2 b} = \frac{\sin^2\gamma}{\sin^2 c}$$

and

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

How can we interpret the right hand side ? Let us compute the volume V of the tetrahedron $OABC$

$$V = \frac{1}{6}\det(\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}) = \frac{1}{6}\begin{vmatrix} 1 & \cos c & \cos b \\ 0 & \sin c & \sin b\cos\alpha \\ 0 & 0 & \sin b\sin\alpha \end{vmatrix} = \frac{1}{6}\sin b\sin c\sin\alpha$$

and dividing by $\sin a$, we get

$$\frac{\sin\alpha}{\sin a} = \frac{6V}{\sin a\sin b\sin c}$$

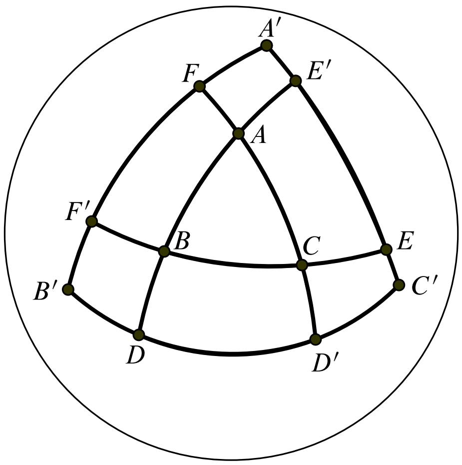
2.3 The dual cosinus law

The dual triangle of a spherical triangle

Theorem and definition. If a, b, c, α, β and γ are the measures of the sides and the angles of a spherical triangle ABC (on S) then there are triangles with measures of the sides and the angles a', b', c', α', β' and γ' such that

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

Such a triangle is called a *dual triangle* of ABC.
Proof. We shall construct a triangle $A'B'C'$ having the required properties.



Construct the arcs of great circles ABD and ACD' both with measure $\frac{\pi}{2}$. The great circle through D and D' is an equator if we take A as a pole.

Construct in the same way the points E, E', F and F'.

Draw the great circles through D and D', through E and E' and through F and F'. Choose the intersection points A', B' and C' in such a way that $A'B'C'$ is a "small" triangle. On the picture we have choosen the triangle that overlaps the triangle ABC.

Since $DD' = B'DD'C'$ is an equator with respect to A, we have $\overset{\frown}{AB'} = \frac{\pi}{2}$ and for the same reason $\overset{\frown}{CB'} = \frac{\pi}{2}$, thus $AC = FACD'$ is an equator with respect to B' and then $\overset{\frown}{B'D'} = \frac{\pi}{2}$. By the same reasoning, we get $\overset{\frown}{DC'} = \frac{\pi}{2}$. Notice that $\overset{\frown}{DD'} = \alpha$, thus $a' = \overset{\frown}{B'C'} = \frac{\pi}{2} + \frac{\pi}{2} - \alpha = \pi - \alpha$.

As above we see that $F'BCE$ is the equator with respect to A', thus $\alpha' = \overset{\frown}{F'E} = \overset{\frown}{F'C} + \overset{\frown}{BE} - \overset{\frown}{BC} = \frac{\pi}{2} + \frac{\pi}{2} - a = \pi - a$.

Proof of the dual cosinus law

From the cosinus law applied to the dual triangle, we get

$$\cos a' = \sin b' \sin c' \cos \alpha' + \cos b' \cos c'$$

or

$$\cos(\pi - \alpha) = \sin(\pi - \beta) \sin(\pi - \gamma) \cos(\pi - a) + \cos(\pi - \beta) \cos(\pi - \gamma)$$

that is

$$-\cos\alpha = -\sin\beta\ \sin\gamma \cos a + \cos\beta\ \cos\gamma$$

Finally

$$\cos\alpha = \sin\beta\ \sin\gamma \cos a - \cos\beta\ \cos\gamma$$

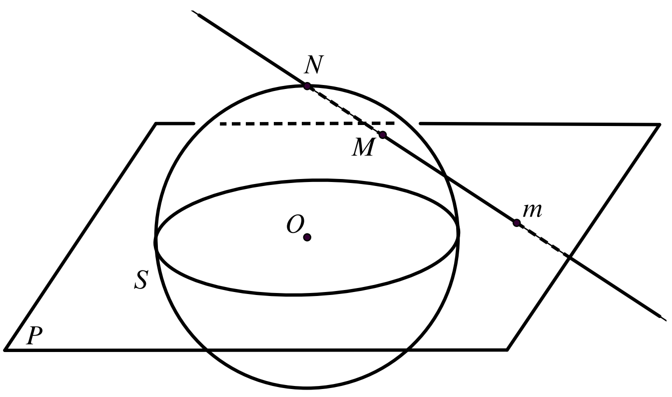
Chapter 4

Stereographic projection

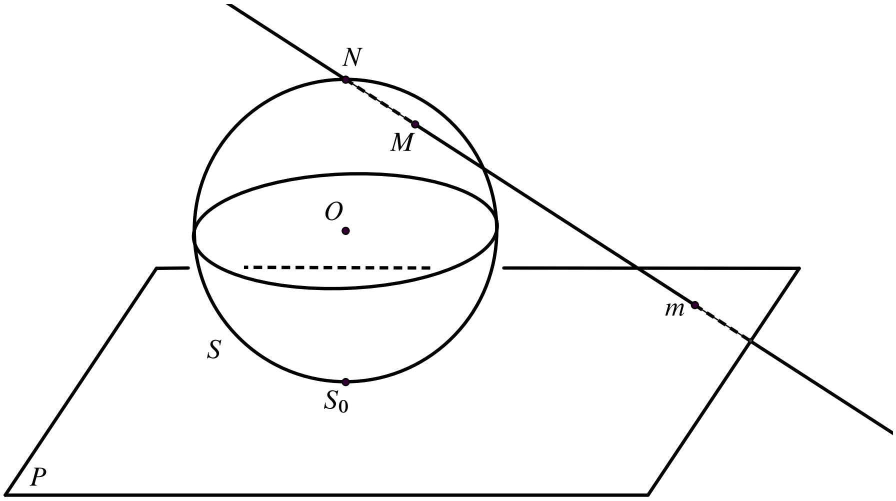
§ 1. Definition

The stereographic projection is a central projection of a sphere onto a plane when the center of projection is on the sphere. The tradition is to take the plane horizontal and to take the North Pole as center of projection. The two most common choices for P are the plane through the center O and the plane tangent to the sphere ar the South Pole.

Choice N° 1 :



Choice N° 2 :

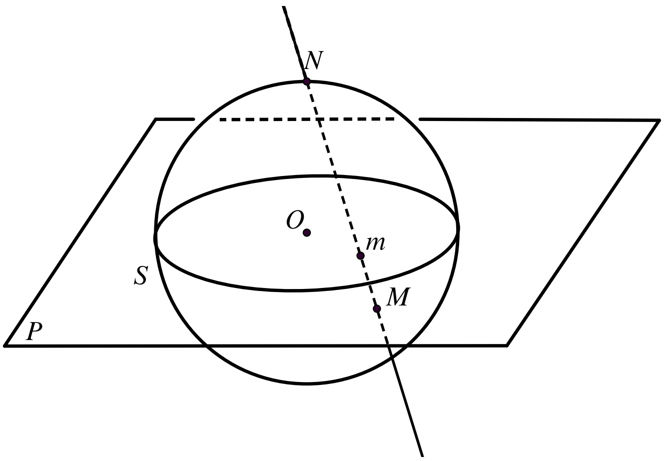


Definition. Given a sphere S with center O, a plane P and a point N such that $N \in S$, that $N \notin P$ and that the line NO is orthogonal to P. The stereographic prjection of S on P, is the map

$$\sigma : S \setminus \{N\} \longrightarrow P, M \longmapsto m \text{ such that the points } N,\ M \text{ and } m \text{ are aligned}$$

Question 1. Why is the map σ well-defined ?
Question 2. Is the map σ a bijection ?
Question 3. How does the image of S change when the plane P is moving ? More precisely, let P_1 be the plane orthogonal to NO and containing the point O. Let us call σ_1 the stereographic projection of $S \setminus \{N\}$ on P_1 and put $m_1 = \sigma_1(M)$. Similarly, let P_2 be the plane parallel to P_1 containing the south pole S_0 and let us call σ_2 the stereographic projection of $S \setminus \{N\}$ on P_2 and put $m_2 = \sigma_2(M)$. What is the relation between $\overrightarrow{Om_1}$ and $\overrightarrow{S_0m_2}$?

1.1 Choice N° 1 with coordinates



Let us take an orthonormal frame $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$. The point $M(X, Y, Z)$ is such that

$$\overrightarrow{OM} = X\vec{\imath} + Y\vec{\jmath} + Z\vec{k}$$

The distance OM is such that $\|\overrightarrow{OM}\|^2 = \overrightarrow{OM} \cdot \overrightarrow{OM} = X^2 + Y^2 + Z^2$, where $\cdot$ denotes the scalar product. We suppose that the radius of the sphere is 1. The point M belongs to S if and only if

$$X^2 + Y^2 + Z^2 = 1 \qquad (1)$$

The North Pole is the point $N(0, 0, 1)$.

The image $m = \sigma(M)$ is in the plane $Z = 0$. Let us denote its coordinates by (x, y) in the frame $(O, \vec{\imath}, \vec{\jmath})$. To say that m is on the line NM, we may say that m is a barycenter of N and M, that is that there is a real number α such that

$$\overrightarrow{Om} = (1 - \alpha)\overrightarrow{ON} + \alpha\overrightarrow{OM}$$

or

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = (1-\alpha)\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \alpha\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

that is

$$\begin{cases} x = (1-\alpha)\cdot 0 + \alpha X \\ y = (1-\alpha)\cdot 0 + \alpha Y \\ 0 = (1-\alpha)\cdot 1 + \alpha Z \end{cases}$$

The last equality gives us α :

$$\alpha = \frac{1}{1-Z}$$

and thus

$$\begin{cases} x = \dfrac{X}{1-Z} \\ y = \dfrac{Y}{1-Z} \end{cases}$$

Knowing that $X^2 + Y^2 + Z^2 = 1$, we can compute the other way round : first we may write :

$$X = x(1-Z) \quad \text{and} \quad Y = y(1-Z)$$

Thus Z has to be solution of the equation :

$$x^2(1-Z)^2 + y^2(1-Z)^2 + Z^2 = 1$$

or

$$(x^2 + y^2 + 1)Z^2 - 2(x^2 + y^2)Z + x^2 + y^2 - 1 = 0$$

This second degree equation has reduced discriminant[1] :

$$\Delta' = (x^2 + y^2)^2 - (x^2 + y^2 + 1)(x^2 + y^2 - 1) = 1$$

[1] The reduced discriminant of the equation $ax + 2b'x + c = 0$ is $\Delta' = b'^2 - ac$ and the solutions are $x_1 = \frac{-b' + \sqrt{\Delta'}}{a}$ and $x_1 = \frac{-b' - \sqrt{\Delta'}}{a}$

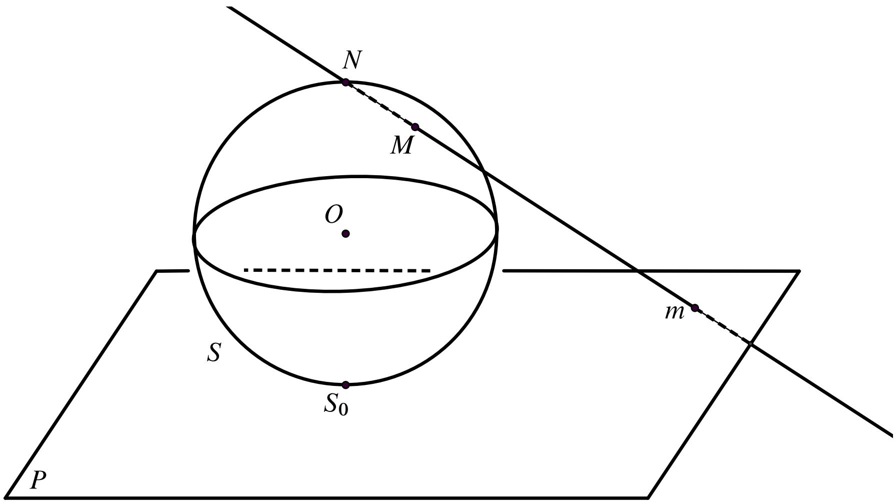
and thus the two solutions are :

$$Z_1 = \frac{x^2+y^2+1}{x^2+y^2+1} = 1 \quad \text{and} \quad Z_2 = \frac{x^2+y^2-1}{x^2+y^2+1}$$

The solution $Z = Z_1$ corresponds to the point N and should be excluded. Thus $Z = Z_2$, and then we get X and Y and get

$$\begin{cases} X = \dfrac{2x}{x^2+y^2+1} \\ Y = \dfrac{2y}{x^2+y^2+1} \\ Z = \dfrac{x^2+y^2-1}{x^2+y^2+1} \end{cases}$$

1.2 Choice N° 2 with coordinates



Let us take as orthonormal frame $(S_0, \vec{\imath}, \vec{\jmath}, \vec{k})$. The point $M(X,Y,Z)$ is such that

$$\overrightarrow{S_0M} = X\vec{\imath} + Y\vec{\jmath} + Z\vec{k}$$

The distance S_0M is such that $\|\overrightarrow{S_0M}\|^2 = \overrightarrow{S_0M} \cdot \overrightarrow{S_0M} = X^2+Y^2+Z^2$, where $\cdot$ denotes the scalar product. We suppose that the radius of the sphere is 1. The point M belongs to S if and only if

$$X^2 + Y^2 + (Z-1)^2 = 1 \qquad (1)$$

The North Pole is the point $N(0,0,2)$.

The image $m = \sigma(M)$ is in the plane $Z = 0$. Let us denote its coordinates by (x, y) in the frame $(S_0, \vec{\imath}, \vec{\jmath})$. To say that m is on the line NM, we may say that m is a barycenter of N and M, that is that there is a real number α such that

$$\overrightarrow{S_0m} = (1-\alpha)\overrightarrow{S_0N} + \alpha\overrightarrow{S_0M}$$

or

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = (1-\alpha)\begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} + \alpha \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

that is

$$\begin{cases} x = (1-\alpha)\cdot 0 + \alpha X \\ y = (1-\alpha)\cdot 0 + \alpha Y \\ 0 = (1-\alpha)\cdot 2 + \alpha Z \end{cases}$$

The last equality gives us α :

$$\alpha = \frac{2}{2-Z}$$

and thus

$$\begin{cases} x = \dfrac{2X}{2-Z} \\ y = \dfrac{2Y}{2-Z} \end{cases}$$

Knowing that $X^2 + Y^2 + Z^2 = 2Z$, we can compute the other way round and get

$$(2-Z)^2 \frac{x^2+y^2}{4} + Z^2 - 2Z = 0$$

or

$$(Z^2 - 4Z + 4)(x^2 + y^2) + 4Z^2 - 8Z = 0$$

that is

$$(x^2 + y^2 + 4)Z^2 - 4(x^2 + y^2 + 2)Z + 4(x^2 + y^2) = 0$$

This is second degree equation in Z, with

$$\Delta' = 4(x^2 + y^2 + 2)^2 - 4(x^2 + y^2 + 4)(x^2 + y^2) = 4^2$$

The roots are

$$Z_1 = \frac{2(x^2+y^2+2)+4}{x^2+y^2+4} = 2 \quad \text{and} \quad Z_2 = \frac{2(x^2+y^2+2)-4}{x^2+y^2+4} = \frac{2(x^2+y^2)}{x^2+y^2+4}$$

The root Z_1 corresponds to the point N, thus $Z = Z_2$, and

$$2 - Z = \frac{8}{x^2+y^2+4}$$

$$\begin{cases} X = \dfrac{4x}{x^2+y^2+4} \\ Y = \dfrac{4y}{x^2+y^2+4} \\ Z = \dfrac{2(x^2+y^2)}{x^2+y^2+4} \end{cases}$$

1.3 Lifting

We have seen that the stereographic projection is bijective. Thus we can define the inverse map.

Definition. We call *lifting* the map from the plane P to the set $S \setminus \{N\}$ which is the inverse of the stereographic projection of S onto P.

§ 2. The stereographic projection preserves the angles

A transformation is *conformal* if it preserves the angles.

Theorem. The stereographic projection is conformal.

Let $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$ be a frame of the usual euclidean space E. The coordinates of a point M are denoted (X, Y, Z). The coordinates of a point m in the plane P, which has equation $Z = 0$, with frame $(O, \vec{\imath}, \vec{\jmath})$, are denoted (x, y). Let S be the sphere with center O and radius 1.

2.1 Geometric proof

Let H be the halfspace with equation $Z < 1$. The stereographic map s from $S \setminus \{N\}$ on the plane P equatorial relatively to the pole $N(0, 0, 1)$, can be extended to a map also denoted by s from H onto P, using the formulae (see §1.1)

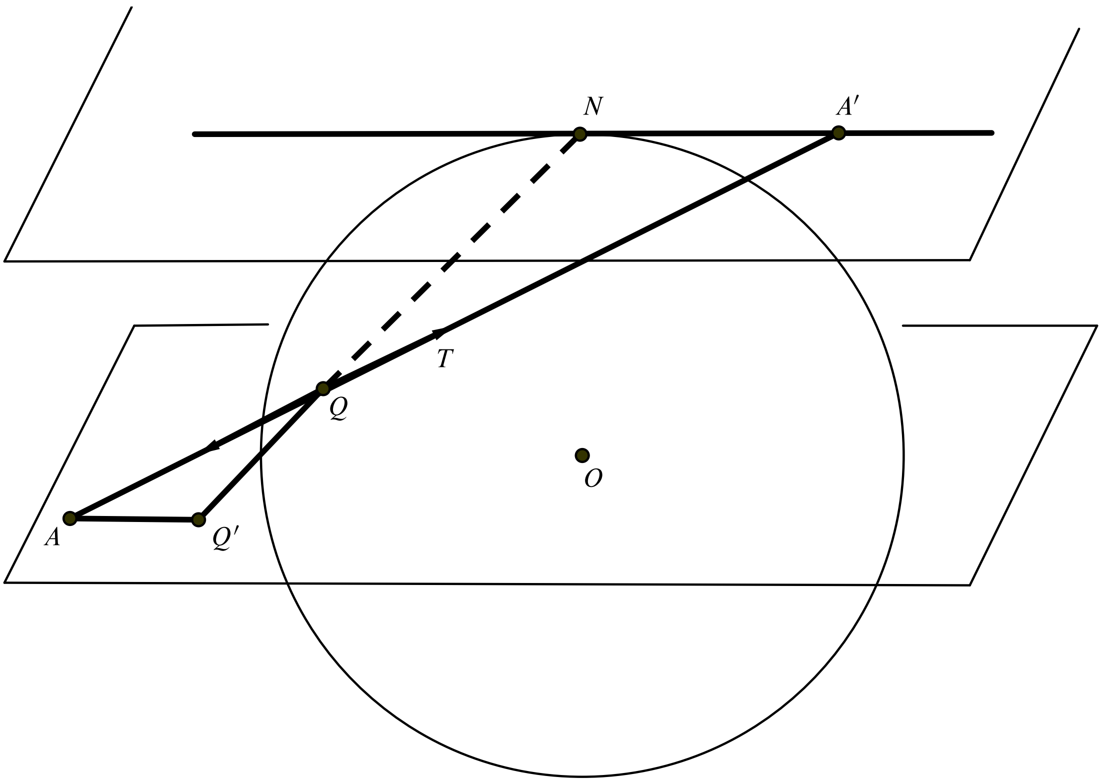
$$\begin{cases} x = \dfrac{X}{1 - Z} \\ y = \dfrac{Y}{1 - Z} \end{cases} \qquad (*)$$

These functions are C^1 (in fact, they are C^∞ and even analytic). Here also, the image of a point M is the intersection m of the line NM with the plane P. Let Δ be any line in E such that $\Delta \cap H \neq \varnothing$. If $N \notin \Delta$, the image by s of $\Delta \cap H$ is a line in P. If $N \in \Delta$, then $s(\Delta \cap H)$ is only a point, in fact the point $\Delta \cap P$.

Let us suppose that two curves γ_1 and γ_2 drawn on the sphere S intersect each other in a point Q and that these two curves have tangents T_1 and T_2 at that point Q. Notice that T_1 and T_2 are tangent to S and thus do not contain N. The curves $s(\gamma_1)$ and $s(\gamma_2)$ intersect in the point $Q' = s(Q)$ and have tangents T_1' and T_2' in Q'. Since the map $(*)$ is C^1, we have $s(T_1 \cap H) = T_1'$ and $s(T_2 \cap H) = T_2'$. What we have to prove is that the angle (T_1, T_2) is equal to the angle (T_1', T_2').

First step.

We denote by Π the median plane of the segment QQ'. Let T be a tangent to the sphere at the point Q. Either T is parallel to the plane P or T intersects P in a point A. If $T /\!/ P$, then the image $T' = s(T)$ is parallel to T and both these lines are orthogonal to the line QQ' and thus symetrical relatively to the plane Π. If T intersects P in A, the image $T' = s(T)$ will be AQ' since $s(Q) = Q'$ and $s(A) = A$. The tangent T intersects also the plane tangent to S in N. Call A' the intersection point of T and the plane tangent to S at N. Since $A'N$ and $A'Q$ are tangent to the same sphere S, the segments $A'N$ and $A'Q$ have same length. But AQ' is parallel to $A'N$ and therefore the triangles AQQ' and $A'QN$ are similar, from what we get $AQ = AQ'$ and thus A belongs to Π.



Second step.
Let T_1 and T_2 be the two tangents to the curves γ_1 and γ_2. They are projected in two lines in P, T_1' and T_2' which are symmetrical to T_1 and T_2 with respect to the plane Π and therefore we have the equality of angles

$$(T_1', T_2') = (T_1, T_2)$$

□

2.2 Differential proof

To say that s is conformal is to say that the linear tangent map $\mathrm{d}s_Q$ is a similarity between euclidean spaces. If we use orthonormal bases and express the linear map $\mathrm{d}s_Q$ by a matrix A relatively to these bases, the condition is that there is a real number λ such that $A\,{}^t\!A = \lambda I$. Since we have $\mathrm{d}\ell^2 = \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2$, we take as an orthonormal basis of the tangent space to S at Q : $(\mathrm{d}\theta, \sin\theta \mathrm{d}\varphi)$. Since

$$\begin{cases} x = \dfrac{X}{1-Z} = \dfrac{\sin\theta\cos\varphi}{1-\cos\theta} = a(\theta)\cos\varphi \\ y = \dfrac{Y}{1-Z} = \dfrac{\sin\theta\sin\varphi}{1-\cos\theta} = a(\theta)\sin\varphi \end{cases}$$

where $a(\theta) = \dfrac{\sin\theta}{1-\cos\theta}$, we have

$$\begin{cases} \mathrm{d}x = a'(\theta)\cos\varphi \mathrm{d}\theta - a(\theta)\sin\varphi \mathrm{d}\varphi \\ \mathrm{d}y = a'(\theta)\sin\varphi \mathrm{d}\theta + a(\theta)\cos\varphi \mathrm{d}\varphi \end{cases}$$

or

$$\begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix} = \begin{bmatrix} a'(\theta)\cos\varphi & -\frac{a(\theta)}{\sin\theta}\sin\varphi \\ a'(\theta)\sin\varphi & \frac{a(\theta)}{\sin\theta}\cos\varphi \end{bmatrix} \begin{bmatrix} \mathrm{d}\theta \\ \sin\theta \mathrm{d}\varphi \end{bmatrix} = A \begin{bmatrix} \mathrm{d}\theta \\ \sin\theta \mathrm{d}\varphi \end{bmatrix}$$

Let us compute

$$A\,A^{\mathrm{T}} = \begin{bmatrix} a'(\theta)\cos\varphi & -\frac{a(\theta)}{\sin\theta}\sin\varphi \\ a'(\theta)\sin\varphi & \frac{a(\theta)}{\sin\theta}\cos\varphi \end{bmatrix} \begin{bmatrix} a'(\theta)\cos\varphi & a'(\theta)\sin\varphi \\ -\frac{a(\theta)}{\sin\theta}\sin\varphi & \frac{a(\theta)}{\sin\theta}\cos\varphi \end{bmatrix}$$

$$= \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} a'(\theta) & 0 \\ 0 & \frac{a(\theta)}{\sin\theta} \end{bmatrix}^2 \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}$$

But

$$a'(\theta) = \frac{(1-\cos\theta)\cos\theta - \sin\theta\ \sin\theta}{(1-\cos\theta)^2} = \frac{\cos\theta - 1}{(1-\cos\theta)^2} = -\frac{1}{1-\cos\theta} = -\frac{a(\theta)}{\sin\theta}$$

and

$$a'(\theta)^2 = \left(\frac{a(\theta)}{\sin\theta}\right)^2$$

thus, finally $A\,A^{\mathrm{T}} = \left(\frac{a(\theta)}{\sin\theta}\right)^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Thus A is a similarity matrix and preserves the ratios of scalar products, that is preserves the angles. The transformation s is conformal. ☐

§ 3. Images of circles

3.1 Image of one circle

Theorem. The image of a circle C on the sphere S by a stereographic projection with pole N is

- **a line** if $N \in C$
- **a circle** if $N \notin C$.

Proof. We choose the frame $(O, \vec{\imath}, \vec{\jmath}, \vec{k})$ such that O is the center of S and $\overrightarrow{ON} = \vec{k}$. A circle C may be defined as the intersection of S with a plane P, that is the set of points (X, Y, Z) such that

$$\begin{cases} X^2 + Y^2 + Z^2 = 1 \\ \alpha X + \beta Y + \gamma Z + \delta = 0 \end{cases}$$

where $(\alpha, \beta, \gamma) \neq (0, 0, 0)$.

Let $M(X, Y, Z)$ be a point belonging to C and let $m(x, y)$ be its image by the stereographic projection on the plane xOy. Using the relations obtained in section 1.1.
$X = \frac{2x}{x^2+y^2+1}$, $Y = \frac{2y}{x^2+y^2+1}$ and $Z = \frac{x^2+y^2-1}{x^2+y^2+1}$,
we get $\alpha \frac{2x}{x^2+y^2+1} + \beta \frac{2y}{x^2+y^2+1} + \gamma \frac{x^2+y^2-1}{x^2+y^2+1} + \delta = 0$ or

$$(\gamma + \delta)(x^2 + y^2) + 2\alpha x + 2\beta y + (\delta - \gamma) = 0 \qquad (2)$$

If $\gamma + \delta = 0$, it means that the plane P contains the point $N(0, 0, 1)$ since $\alpha 0 + \beta 0 + \gamma 1 + \delta = 0$.

If $(\alpha, \beta) = (0, 0)$, the plane P is the plane through N parallel to the plane xOy and C is the point-circle N. We'l understand later how to interpret the image of this point-circle.

If $(\alpha, \beta) \neq (0, 0)$, the image of C is a line.

If $\gamma + \delta \neq 0$, (2) is the equation of a circle with center $(-\frac{\alpha}{\gamma+\delta}, -\frac{\beta}{\gamma+\delta})$ and a radius such that

$$R^2 = \frac{\alpha^2 + \beta^2 + \gamma^2 - \delta^2}{(\gamma + \delta)^2}$$

We see on this expression that the circle is real if and only if $\delta^2 \leqslant \alpha^2 + \beta^2 + \gamma^2$. This corresponds to a plane P which intersects the sphere S in real points.

Reciprocally, every point of the circle or line (2) is the image of a point on C.

It is also easy to see that the lifting of any circle or line in P is a circle on S. □

3.2 Pencils of curves in the plane

Definition. Let C_0 and $\widetilde{C}$ be two curves with equations

$$f_0(x, y) = 0 \quad \text{and} \quad \widetilde{f}(x, y) = 0$$

the *pencil of curves* determined by C_0 and $\widetilde{C}$ is the set of curves with equations

$$\lambda f_0(x, y) + \mu \widetilde{f}(x, y) = 0$$

where $(\lambda, \mu) \neq (0, 0)$.

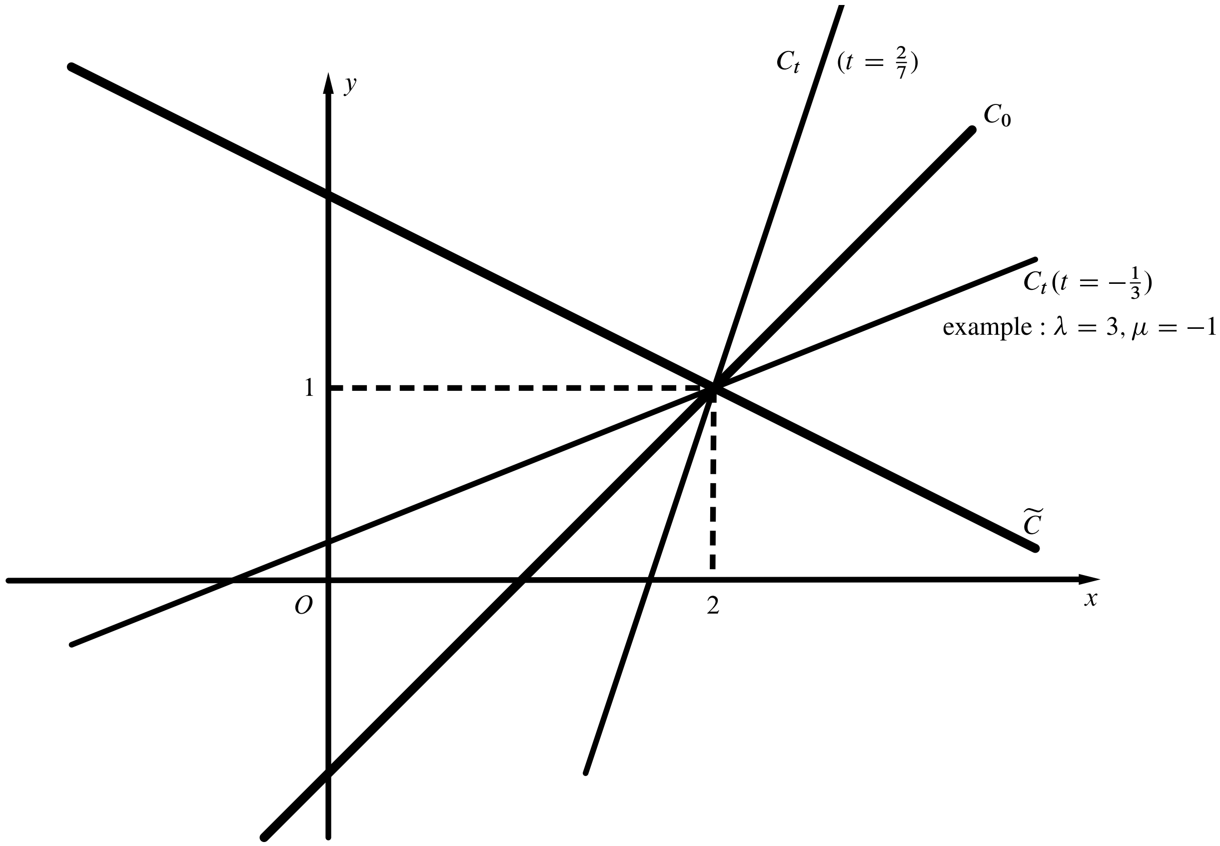
Example. Let C_0 and $\widetilde{C}$ be the lines with equations

$$f_0(x, y) \equiv x - y - 1 = 0 \quad \text{and} \quad \widetilde{f}(x, y) \equiv x + 2y - 4 = 0$$

We note that C_0 and $\widetilde{C}$ intersect each other in the point $(2, 1)$. Then for any choice of (λ, μ), distinct from $(0, 0)$,

$$(\lambda f_0 + \mu \widetilde{f})(x, y) \equiv (\lambda + \mu)x + (-\lambda + 2\mu)y + (-\lambda - 4\mu) = 0$$

is the equation of a line C containing the point $(2, 1)$. Reciprocally, given any line C containing the point $(2, 1)$ there are couples (λ, μ) such that the equation of C may be written $(\lambda f_0 + \mu \widetilde{f})(x, y) = 0$: the pencil induced by C_0 and $\widetilde{C}$ is the set of all the lines through the point $(2, 1)$.



Exercise 1. Show that if a line D is going through the point $(2, 1)$ then there is a couple (λ, μ) of real numbers such that D has the equation

$$(\lambda f_0 + \mu \widetilde{f})(x, y) = (\lambda + \mu)x + (-\lambda + 2\mu)y + (-\lambda + 4\mu) = 0$$

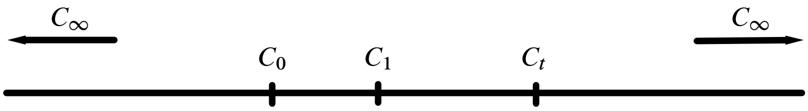
Comments. Let us denote $C_{\begin{pmatrix}\lambda\\ \mu\end{pmatrix}}(x, y) = 0$ the curve with equation

$$f_{\begin{pmatrix}\lambda\\ \mu\end{pmatrix}}(x, y) \equiv \lambda f_0(x, y) + \mu \widetilde{f}(x, y) = 0$$

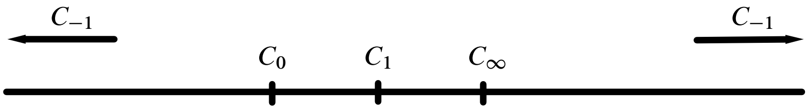
There is only one curve in the pencil with $\lambda = 0$, it is the curve $\widetilde{C}$. For all other curves in the pencil $\lambda \neq 0$ and we have

$$C_{\begin{pmatrix}\lambda\\ \mu\end{pmatrix}} = C_{\begin{pmatrix}1\\ \frac{\mu}{\lambda}\end{pmatrix}}$$

We put $t = \frac{\mu}{\lambda}$ and we denote simply by C_t the curve $C_{\begin{pmatrix}1\\ t\end{pmatrix}}$. For every real t we have one curve $C_t = C_0 + t\widetilde{C}$. As a consequence we see that if we exclude the curve $\widetilde{C}$, we have a bijection of the pencil of curves with the real line. To get a bijection with the complete pencil, we have to add one element to $\mathbb{R}$. It is a habit to denote that extra element by ∞ and call it "point at infinity" of the real line. Then we denote $\widetilde{C}$ by C_∞.



But we can as well decide to draw C_∞ in our page. Thus we get the following picture of the set of curves in the pencil

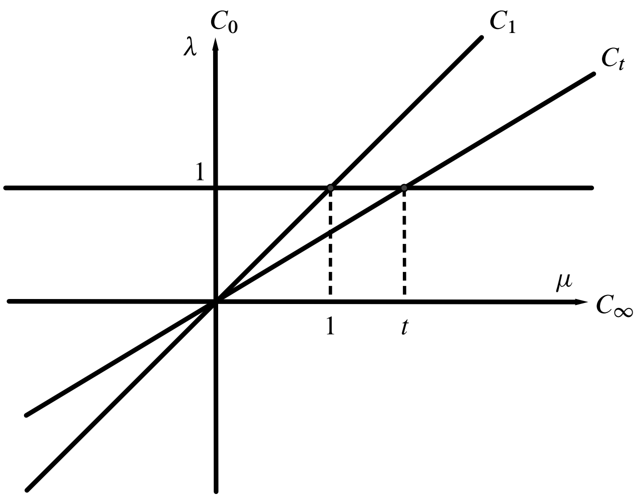


The following definition of the projective line will make everything clear.

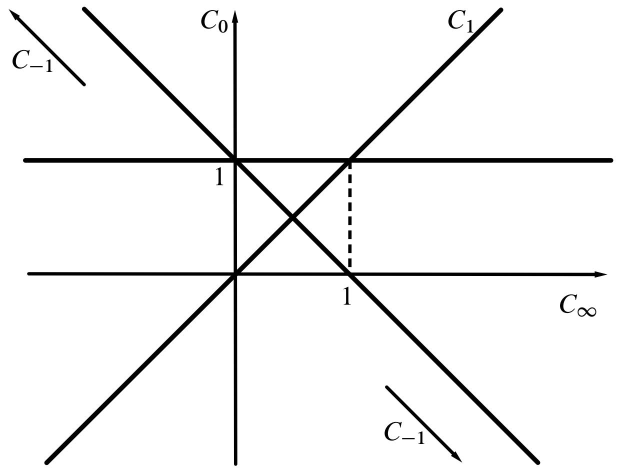
Definition. In $\dot{P} = \mathbb{R}^2 \setminus \{(0,0)\}$ the relation $\sim$ defined by

$$(\lambda, \mu) \sim (\lambda', \mu') \quad \overset{\text{definition}}{\Longleftrightarrow} \quad \lambda\mu' = \lambda'\mu$$

is an equivalence relation. The quotient of $\dot{P}$ by $\sim$ is the real projective line, denoted by $\hat{\mathbb{R}}$, by $P^1(\mathbb{R})$ or by $\mathbb{R}P^1$.



If we want to show the fact that C_0 and C_∞ play symmetrical roles, we cut the lines by the line $\lambda + \mu = 1$



3.3 Examples of pencils

Pencils of lines in a plane

1. Pencil of intersecting lines.

See the picture page 38.

Theorem. Given two distinct lines L_1 and L_2 with equations

$$a_1x+b_1y+c_1 = 0,\ (a_1,b_1) \neq (0,0) \qquad \text{and} \qquad a_2x+b_2y+c_2 = 0,\ (a_2,b_2) \neq (0,0)$$

intersecting in the point $M_0(x_0, y_0)$, that is such that $a_1b_2 - b_1a_2 \neq 0$, $c_1 = -a_1x_0 - b_1y_0$ and $c_2 = -a_2x_0 - b_2y_0$, the set of lines through M_0 is the pencil of lines determined by L_1 and L_2.

Proof. Let $(\lambda, \mu) \neq (0,0)$ and define $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ by

$$f(x, y) = \lambda(a_1x + b_1y + c_1) + \mu(a_2x + b_2y + c_2)$$

We have $f(x, y) = ax + by + c$ where

$$a = \lambda a_1 + \mu a_2,\ b = \lambda b_1 + \mu b_2 \quad \text{and} \quad c = \lambda c_1 + \mu c_2$$

We have to show that $(a, b) \neq (0,0)$ and $c = -ax_0 - by_0$.

If $(a, b) = (0,0)$, since $a_1b_2 - b_1a_2 \neq 0$, we would have $(\lambda, \mu) = (0,0)$ which is contrary to the hypothesis.

By direct computation we have $c = \lambda c_1 + \mu c_2 = \lambda(-a_1x_0 - b_1y_0) + \mu(-a_2x_0 - b_2y_0) = -(\lambda a_1 + \mu a_2)x_0 - (\lambda b_1 + \mu b_2)y_0 = -ax_0 - by_0$.

Reciprocally, let L be a line through M_0. The equation of L may be written

$$ax + by + c = 0 \text{ with } (a, b) \neq (0,0) \text{ and } c = -ax_0 - by_0$$

We have to show the existence of a couple (λ, μ) such that

$$(\lambda, \mu) \neq (0,0) \text{ and } a = \lambda a_1 + \mu a_2,\ b = \lambda b_1 + \mu b_2 \quad \text{and} \quad c = \lambda c_1 + \mu c_2$$

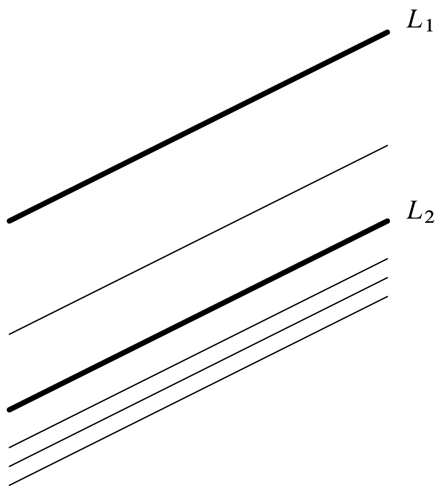
The linear system in (λ, μ)

$$\begin{cases} a_1\lambda + a_2\mu = a \\ b_1\lambda + b_2\mu = b \end{cases}$$

has a unique solution since $a_1b_2 - b_1a_2 \neq 0$. This solution cannot be $(0,0)$ because $(a,b) \neq (0,0)$. Finally we check

$$c = -(a_1\lambda + a_2\mu)x_0 - (b_1\lambda + b_2\mu)y_0 = \lambda(-a_1x_0 - b_1y_0) + \mu(-a_2x_0 - b_2y_0) = \lambda c_1 + \mu c_2$$

2. Pencil of parallel lines.



Theorem. Given two distinct parallel lines L_1 and L_2 with equations

$$a_1x + b_1y + c_1 = 0,\ (a_1, b_1) \neq (0,0) \qquad \text{and} \qquad a_2x + b_2y + c_2 = 0,\ (a_2, b_2) \neq (0,0)$$

that is such that $a_1b_2 - b_1a_2 = 0$ and $a_1c_2 - c_1a_2 \neq 0$ or $c_1b_2 - b_1c_2 \neq 0$, the pencil of lines determined by L_1 and L_2 is the set of lines parallel to L_1 (and L_2) union an extra line "the line at infinity of the plane".

Exercise 2. Why is the set of lines parallel to L_1 (and L_2) not a complete pencil ?

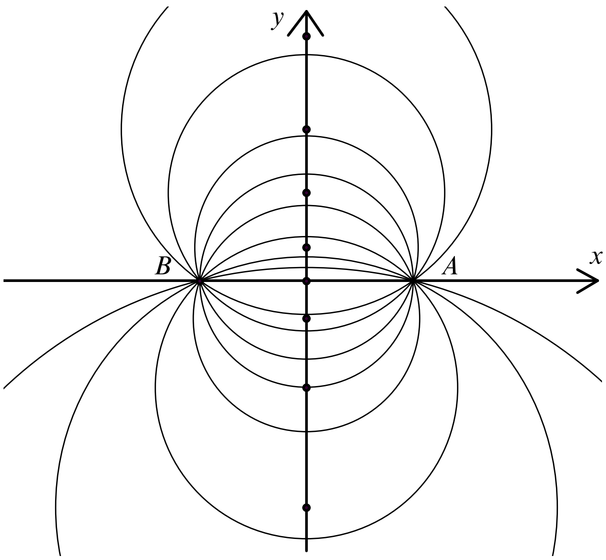
Pencils of planes in space

Exercise 3. Define pencils of intersecting planes along a line and pencils of parallel planes.

Pencils of circles in a plane

1. Elliptic pencil of circles or pencil of circles with base points.

Theorem and definition. Let A and B be two points in an affine euclidean plane, the set of circles containing A and B is the pencil defined by any two of these circles. The points A and B are called the *base points* of the pencil of circles. This pencil is called *elliptic*.



Equations of the circles of an elliptic pencil of circles.

Let us choose the middle of the segment AB as origin O and Ox such that the coordinates of A are $(a, 0)$ and then $B(-a, 0)$. A circle containing A and B has its center Ω on Oy. Write $(0, t)$ the coordinates of Ω, the equation of the circle with center Ω and containing A and B is $x^2 + (y - t)^2 - (a^2 + t^2) = 0$ or

$$x^2 + y^2 - 2ty - a^2 = 0$$

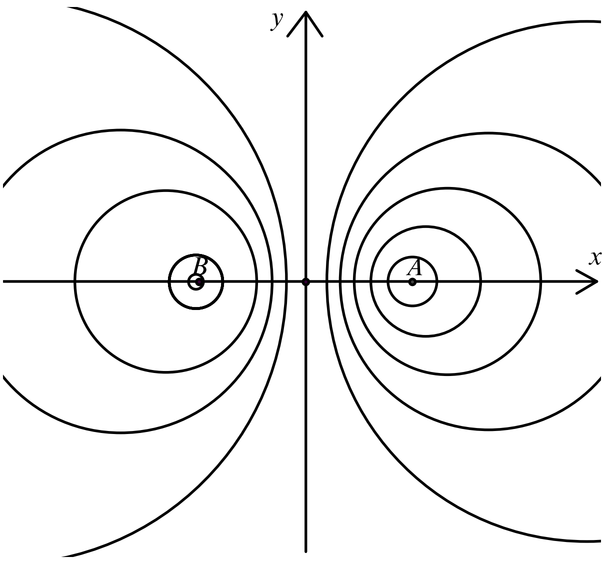
Proof of the theorem. The pencil of curves determined by any two distinct circles containing A and B is the set of all curves with equation

$$\lambda(x^2 + y^2 - 2t_1 y - a^2) + \mu(x^2 + y^2 - 2t_2 y - a^2) = 0$$

where $t_1 \neq t_2$ and $(\lambda, \mu) \neq (0, 0)$, that is the line $y = 0$ if $\lambda + \mu = 0$ and the circle containing A and B with center $(0, \frac{\lambda t_1 + \mu t_2}{\lambda + \mu})$ if $\lambda + \mu \neq 0$. $\square$

2. Hyperbolic pencil of circles or pencil of circles with limit points.

Theorem et definition. Let C_1 and C_2 be two disjoint circles with distinct centers in an affine euclidean plane. The pencil defined by these two circles contains two circles with radius 0 : let us call the centers of these point-circles A and B. The points A and B are called the *limit points* of the pencil of circles. This pencil is called *hyperbolic*.



Proof.

Let Ω_1 and Ω_2 be the centers of the circles C_1 and C_2. Let us choose the line $\Omega_1\Omega_2$ as the line $x'Ox$. The equations of C_1 and C_2 may be written

$$(x-\omega_1)^2+y^2-r_1^2=0 \qquad \text{and} \qquad (x-\omega_2)^2+y^2-r_2^2=0$$

where ω_1 et ω_2 are the abscissae (or abscissas) of Ω_1 and Ω_2 and r_1 and r_2 are the radii (or radiuses) of C_1 and C_2.

The circles C_1 and C_2 have common point(s) if and only if

$$|r_2-r_1| \leqslant |\omega_2-\omega_1| \leqslant r_1+r_2$$

This condition may be written

$$(r_2-r_1)^2 \leqslant (\omega_2-\omega_1)^2 \leqslant (r_2+r_1)^2$$

or

$$-2r_1r_2 \leqslant (\omega_2-\omega_1)^2-(r_1^2+r_2^2) \leqslant 2r_1r_2$$

or even

$$\big((\omega_2-\omega_1)^2-(r_1^2+r_2^2)\big)^2 \leqslant 4r_1^2r_2^2$$

Thus C_1 and C_2 have no common point if and only if

$$\big((\omega_2-\omega_1)^2-(r_1^2+r_2^2)\big)^2 > 4r_1^2r_2^2 \qquad (*)$$

The circles in the pencil determined by C_1 and C_2 have the equations

$$\lambda\big((x-\omega_1)^2+y^2-r_1^2\big)+(1-\lambda)\big((x-\omega_2)^2+y^2-r_2^2\big)=0$$

or

$$x^2+y^2-2(\lambda\omega_1+\omega_2-\lambda\omega_2)x+\omega_2^2-r_2^2+\lambda(\omega_1^2-r_1^2-\omega_2^2+r_2^2)=0$$

or even

$$(x-(\lambda\omega_1+\omega_2-\lambda\omega_2))^2+y^2-(\lambda\omega_1+\omega_2-\lambda\omega_2)^2+\omega_2^2-r_2^2+\lambda(\omega_1^2-r_1^2-\omega_2^2+r_2^2)=0$$

It is possible to find two circles with radii equal to 0 if and only if there are two real values of λ such that

$$-(\lambda\omega_1+\omega_2-\lambda\omega_2)^2+\omega_2^2-r_2^2+\lambda(\omega_1^2-r_1^2-\omega_2^2+r_2^2)=0$$

or

$$(\omega_1-\omega_2)^2\lambda^2+\big(-(\omega_1-\omega_2)^2+(r_1^2-r_2^2)\big)\lambda+r_2^2=0$$

The condition is that the discriminant Δ of that equation of the second degree in λ is such that $\Delta>0$, where

$$\Delta=[(\omega_1-\omega_2)^2-(r_1^2+r_2^2)]2-4r_1^2r_2^2$$

We see that $\Delta>0$ is equivalent to $(*)$. □
Consequence. Instead of taking the two circles C_1 and C_2 to determine a hyperbolic pencil of circles, one may take the two distinct point-circles, the limit points of the pencil.

Equations of the circles of an hyperbolic pencil of circles

Let us take as limit-points the points $A(a,0)$ and $B(-a,0)$. The two point-circles have the equations

$$(x-a)^2+y^2=0 \qquad \text{and} \qquad (x+a)^2+y^2=0$$

or

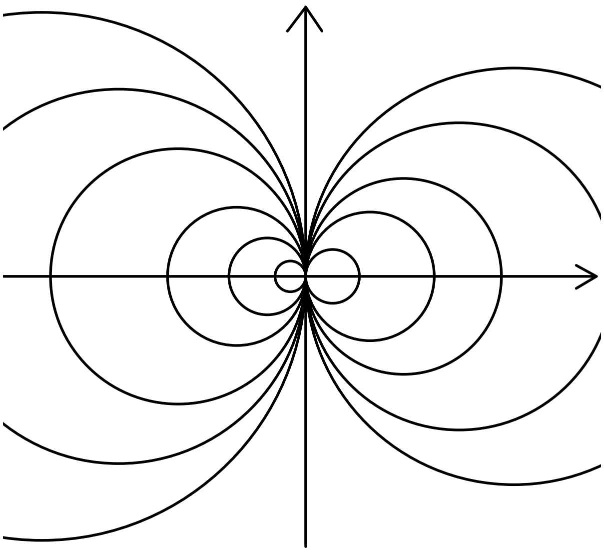
$$x^2+y^2-2ax+a^2=0 \qquad \text{and} \qquad x^2+y^2+2ax+a^2=0$$

Thus the pencil contains the line $x=0$ and the circles with equations

$$x^2+y^2-2tx+a^2=0 \quad \text{where } t\in\mathbb{R}$$

3. Parabolic pencil of circles or pencil of tangent circles.

Theorem et definition. Let C_1 and C_2 be two disjoint circles tangent in a point T ; the pencil defined by these two circles contains the line Λ tangent at T to these two circles and all the circles containing T and tangent to two the line Λ. This pencil is called *parabolic*.



4. What have we forgotten ?

There are three other types of "circle"-pencils ; the notation "circle" is to say "circle or line" :

1. Pencil of concentric circles
2. Pencil of convergent lines
3. Pencil of parallel lines

We'l see later how these pencils can be included in the former three types of pencils.

3.4 Orthogonal circles and lines ; orthogonal pencils of circles in the plane

Definition. Two circles are *orthogonal* if the tangents to the circles at their common points are orthogonal. The definition may be extended to "circles" in the meaning of "circles or lines".

Theorem. Two circles C_1 and C_2 with centers Ω_1 and Ω_2 and radii R_1 and R_2 are orthogonal if and only if

$$d^2 = R_1^2 + R_2^2$$

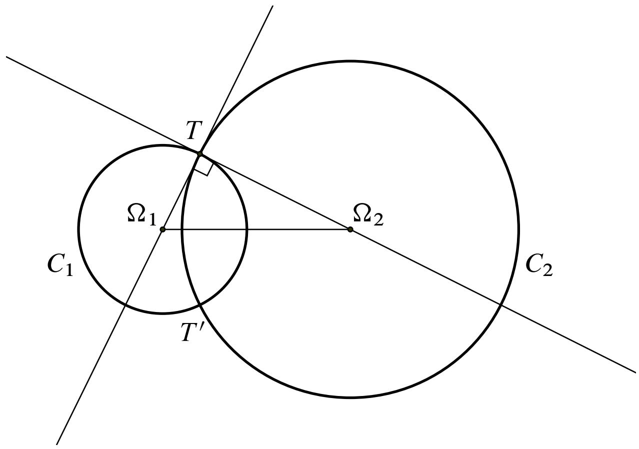
where $d = \Omega_1\Omega_2$ is the distance between the centers of the circles C_1 and C_2.

A circle C with center Ω and a line Λ are orthogonal if and only if $\Omega \in \Lambda$, that is if and only if the line is a diameter of the circle.

Two lines are orthogonal if they are orthogonal in the usual meaning.

Proof. Let T and T' be the common points to the circles C_1 and C_2, if there are any. By symmetry relatively to the line $\Omega_1\Omega_2$ the angles of the tangents to the circles at T and at T' are equal. Since the tangent to a circle at a point T belonging to the circle is the line orthogonal to the radius ending at T, the circles will be orthogonal if and only if the triangle $\Omega_1 T \Omega_2$ is a right triangle or by Pythagoras' theorem

$$d^2 = R_1^2 + R_2^2$$



Theorem. Two circles or lines with equations

$$\alpha_1(x^2+y^2) - 2a_1x - 2b_1y + \gamma_1 = 0 \qquad \text{and} \qquad \alpha_2(x^2+y^2) - 2a_2x - 2b_2y + \gamma_2 = 0$$

where $(\alpha_1, a_1, b_1) \neq (0,0,0)$ and $(\alpha_2, a_2, b_2) \neq (0,0,0)$ are orthogonal if and only if

$$\boxed{\alpha_1\gamma_2 + \gamma_1\alpha_2 - 2a_1a_2 - 2b_1b_2 = 0}$$

Proof. We first consider the situation when the curves are circles with centers Ω_1 and Ω_2. These centers have the coordinates (a_1, b_1) and (a_2, b_2), thus

$$d^2 = \left(\Omega_1\Omega_2\right)^2 = \left(\frac{a_1}{\alpha_1} - \frac{a_2}{\alpha_2}\right)^2 + \left(\frac{b_1}{\alpha_1} - \frac{b_2}{\alpha_2}\right)^2$$

On the other hand, the radius R_1 of C_1 is such that the equation of C_1 is $(x - \frac{a_1}{\alpha_1})^2 + (y - \frac{b_1}{\alpha_1})^2 = R_1^2$, thus

$$R_1^2 = \frac{-\alpha_1\gamma_1 + a_1^2 + b_1^2}{\alpha_1^2}$$

and similarly $R_2^2 = \frac{-\alpha_2\gamma_2 + a_2^2 + b_2^2}{\alpha_2^2}$. The condition $d^2 = R_1^2 + R_2^2$ becomes

$$\left(\frac{a_1}{\alpha_1}\right)^2 + \left(\frac{b_1}{\alpha_1}\right)^2 + \left(\frac{a_2}{\alpha_2}\right)^2 + \left(\frac{b_2}{\alpha_2}\right)^2 - 2\frac{a_1a_2}{\alpha_1\alpha_2} - 2\frac{b_1b_2}{\alpha_1\alpha_2} = \left(\frac{a_1}{\alpha_1}\right)^2 + \left(\frac{b_1}{\alpha_1}\right)^2 + \left(\frac{a_2}{\alpha_2}\right)^2 + \left(\frac{b_2}{\alpha_2}\right)^2 - \frac{\gamma_1}{\alpha_1} - \frac{\gamma_2}{\alpha_2}$$

Simplyfying and reducing to the same denominator, we get the result.

If at least one of the ”circles” is a line, it is easy to check that the relation is still valid. □

Definition. Two pencils of circles are *orthogonal* if every circle of one pencil is orthogonal to every circle in the other pencil.

Theorem. The hyperbolic pencil of circles with limit points A and B is orthogonal to the pencil of circles with base points A and B.

The pencil of concentric circles with center Ω is orthogonal to the pencil of lines through Ω.

The elliptic pencil of circles tangent to the line Λ at point T is orthogonal to the elliptic pencil of circles tangent to the line Λ' at point T, where Λ' is the line containing T and orthogonal to Λ.

The pencil of lines parallel to a line Λ is orthogonal to the pencil of lines parallel to any line Λ' orthogonal to Λ.

Proof. Let us choose coordinates in such a way that $A(0, a)$ and $B(0, -a)$. The circles of the pencil with limit-points A and B have equations

$$x = 0 \qquad \text{or} \qquad x^2 + y^2 - 2tx + a^2 = 0$$

that is

$$\alpha_1 = 0, a_1 = 1, b_1 = 0, \gamma_1 = 0 \qquad \text{or} \qquad \alpha_1 = 1, a_1 = t, b_1 = 0, \gamma_1 = a^2$$

The circles of the pencil with base-points A and B have equations

$$y = 0 \qquad \text{or} \qquad x^2 + y^2 - 2t'y - a^2 = 0$$

that is

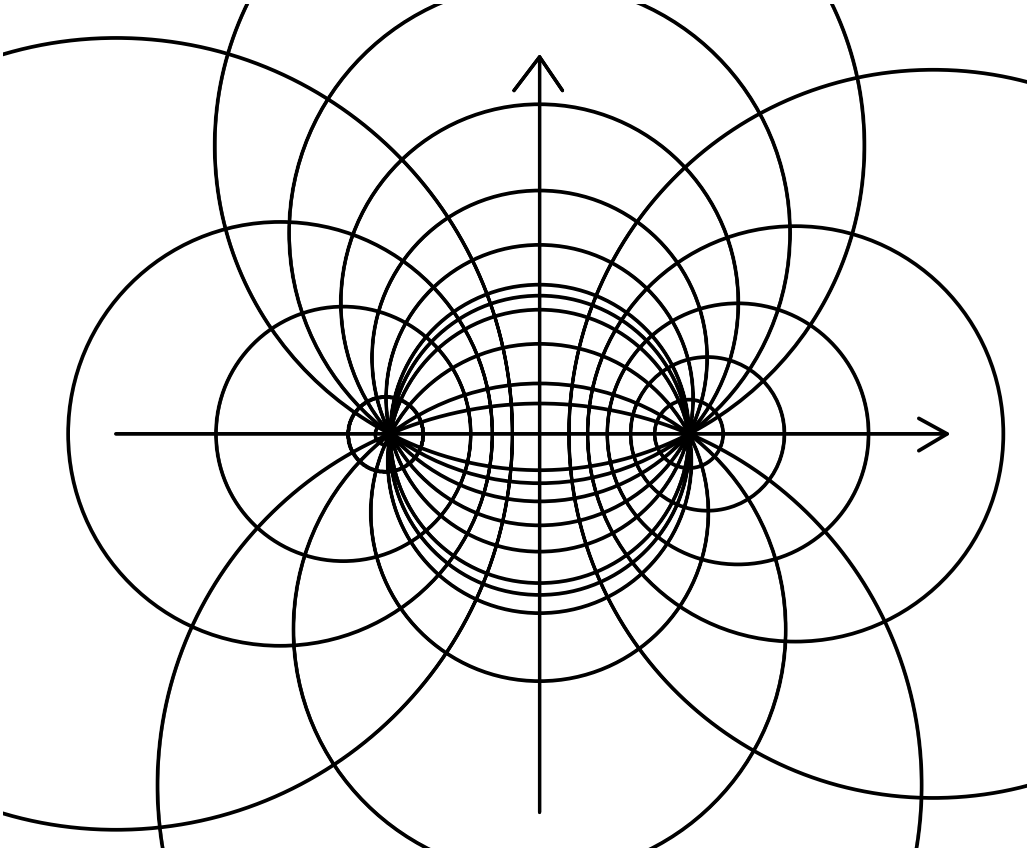
$$\alpha_2 = 0, a_2 = 0, b_2 = 1, \gamma_2 = 0 \qquad \text{or} \qquad \alpha_2 = 1, a_2 = 0, b_2 = t', \gamma_2 = -a^2$$

Using the previous theorem, the four identites we have to verify are thus

$$\begin{aligned} 0 \times 0 + 0 \times 0 - 2 \times 1 \times 0 - 2 \times 0 \times 1 &= 0 \\ 0 \times (-a^2) + 1 \times 0 - 2 \times 1 \times 0 - 2 \times 0 \times t &= 0 \\ 1 \times 0 + a^2 \times 0 - 2 \times t \times 0 - 2 \times 0 \times 1 &= 0 \\ 1 \times (-a^2) + a^2 \times 1 - 2 \times t \times 0 - 2 \times 0 \times t' &= 0 \end{aligned}$$

which are trivial.

The cases of other pencils are even simpler to verify. □



3.5 Orthogonal circles and lines and orthogonal pencils on the sphere

Definition. A pencil of circles on the sphere is the set of intersections of the sphere with the planes of a pencil of planes in space.

The pencil of planes is characterized either by a common line Λ or by a common direction of planes (when all the planes are parrallel to each other). The parallel case may be considered as a special case of the general case, when the common line is ”at infinty”.

There are 3 possibilities : the line Λ cuts the sphere S in two points A and B : the corresponding pencil of circles is the set of all circles on S containing A and B. If Λ is tangent to S, we gat the set of circles on S tangent to Λ. If Λ does not cut the sphere, we can take the two planes in the pencil of planes which are tangent to S : the contact points are point-circles...

By the lifting from the plane to the sphere, two orthogonal pencils of circles have images which are orthogonal pencils on the sphere associated to two lines Λ_1 and Λ_2 orthogonal to each other and such that the center O of the sphere belongs to the common orthogonal line cutting Λ_1 and Λ_2 in points H_1 and H_2 such that $OH_1 \cdot OH_2 = 1 (= R^2)$.

Now we have to draw a nice picture with ”cabri 3D”.

Chapter 5

Projective geometry

§ 1. First description of the real projective plane

The projective geometry is beautiful, it is easy when you get rid of some prejudices, it is effective for solving many problems and it unifies many different aspects of mathematics. We start with "concrete" way to look at this geometry.

1.1 How to describe a line in the plane ?

Let $(O, \vec{\imath}, \vec{\jmath})$ be a frame of a plane P.

There are two usual way to describe a line in the plane P :

$$y = mx + p \quad (1) \qquad \text{and} \qquad ax + by + c = 0 \quad (2)$$

The formula (1) is nice to describe a function (an ”affine function”), but it is not convenient for geometry because you do not get the ”vertical” lines, the lines parallel to the vector $\vec{\jmath}$.

The formula (2) is nice for geometry because it is able to describe all the lines in the plane P. But it has (at least) two drawbacks :

1. all triplets (a, b, c) are not convenient. We have to exclude those where $(a, b) = (0, 0)$ i.e. $a = 0$ and $b = 0$.
2. two triplets (a, b, c) and(a', b', c') describe the same line if there are proportional

$$\frac{a'}{a} = \frac{b'}{b} = \frac{c'}{c} \qquad (3)$$

Note that the relation (3) is not so convenient when some of the coefficients are 0. It is better to say that the vectors (a, b, c) and (a', b', c') are colinear or linearly dependant.

These drawbacks are inherent to the affine geometry we are using. They will loose their acuity in the projective geometry. So we'll go on with the formula (2).

1.2 A first prejudice to get rid of

Prejudice 1. A line is a set of points.

It is easier to think of the set of points as a set $\mathcal{P}$ and the set of lines as another set $\mathcal{L}$. Then one has to define a relation between these two sets.

Let us stick to formula (2) : we say that a point M described by (x, y) *is on* the line d described by (a, b, c) if and only if (2) holds. The other way round we say that the line d *goes through* the point M.

We see here the duality between points and lines. This duality is not perfect, there are exceptions but the dual of the statement : "There is one and only one line going through two given distinct points" would be : "There is one and only one point lying on two given distinct lines". We have to add :"non parallel". With the projective geometry we'll get rid of that adendum. The first step will be to modify the presentation of our formula to enhance this duality.

Dead ends. One idea could be : since proportional triplets describe the same line, let us choose a specific one. If we take $b = -1$ or any non zero constant, we are back to the "bad" formula (1). But let us try $c = 1$, then the equation (2) becomes

$$ax + by + 1 = 0 \quad \text{or} \quad ax + by = -1 \quad \text{or} \quad \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -1 \quad \text{or} \quad \begin{bmatrix} a & b & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0$$

It is indeed nice and symmetrical, but we have lost all the lines going through the origin O. The dual relation of parallelisme for lines would be for points : the points (x, y) and (x', y') such that $xy' - yx' = 0$ are kind of "parallel points". We have indeed the symmetry between points and lines, but what we have obtained is to transfer to the points the drawbacks of not intersecting distinct lines. The optimal solution is the other way round to get even parallel lines to intersect. Instead of trying to get the lines look as points, we'll try to get the points look like lines. That will be done introducing homogenous cordinates.

1.3 The homogenous coordinates

Definition. Let (x, y) be the coordinates of a point M in P. We call *homogenous coordinates* of M all triplets (X, Y, Z) such that :

$$x = \frac{X}{Z} \qquad \text{and} \qquad y = \frac{Y}{Z} \qquad \text{and} \qquad Z \neq 0$$

Theorem. A point M with homogenous coordinates (X, Y, Z) is on the line d described by the triplet (a, b, c) if and only if

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = 0 \qquad (4)$$

Proof. The relation $ax + by + c = 0$ becomes $a\frac{X}{Z} + b\frac{Y}{Z} + c = 0$ or $aX + bY + cZ = 0$. We can go the way back since $Z \neq 0$. $\square$

Remark. We'l take care of the case $Z = 0$ later on.

Let us take two distinct points, that is (X, Y, Z) and (X', Y', Z') non proportional (or linearly independant), then there is one and only one triplet up to a multiplicative constant (a, b, c) such that (4). Explicitly

$$\begin{cases} a = \lambda(YZ' - ZY') \\ b = \lambda(ZX' - XZ') \\ c = \lambda(XY' - YX') \end{cases} \quad \text{where } \lambda \text{ is any number different from } 0$$

How do we get that result ? We have to solve the following system of 2 equations with 3 unknowns a, b and c :

$$\begin{cases} Xa + Yb + Zc = 0 \\ X'a + Y'b + Z'c = 0 \end{cases} \text{where } XY' - YX' \neq 0 \text{ or } YZ' - ZY' \neq 0 \text{ or } ZX' - XZ' \neq 0$$

You may solve the system the way you like. An easy way to do it is to think of (X, Y, Z), (X', Y', Z') and (a, b, c) as vectors $\overrightarrow{U}$, $\overrightarrow{U'}$, where $\overrightarrow{U}$ and $\overrightarrow{U'}$ are not colinear, and $\overrightarrow{v}$ in a 3-dimensional euclidean space and look for $\overrightarrow{v}$ orthogonal to both $\overrightarrow{U}$ and $\overrightarrow{U'}$. Then the solution is colinear to the vectorial product, that is $\lambda \overrightarrow{U} \wedge \overrightarrow{U'}$. Thus we have proved

Theorem. There is one and only one line which goes through two distinct points with homogenous coordinates (X, Y, Z) and (X', Y', Z') with $Z \neq 0$ and $Z' \neq 0$.

We may look at the dual problem. Let us take 2 distinct lines (a, b, c) and (a', b', c') with $bc' - cb' \neq 0$ or $ca' - ac' \neq 0$ or $ab' - ba' \neq 0$. If the lines are not parallel, we have $ab' - ba' \neq 0$ and the system of 2 equations with 3 unknowns X, Y and Z :

$$\begin{cases} aX + bY + cZ = 0 \\ a'X + b'Y + c'Z = 0 \end{cases} \quad \text{where } ab' - ba' \neq 0$$

You may solve this system using the computations done above and get

$$\begin{cases} X = \lambda(bc' - cb') \\ Y = \lambda(ca' - ac') \\ Z = \lambda(ab' - ba') \end{cases} \quad \text{where } \lambda \text{ is any number different from } 0$$

You even get from the inequality that $Z \neq 0$, thus

Corollary. There is one and only one point which is on two distinct non parallel lines.

1.4 Inventing a new line and new points

The triplet $(a, b, c) = (0, 0, 0)$ does not describe any line, since for any (X, Y, Z) we have $aX + bY + cZ = 0$. So we have to exclude these triplets. In the same way we have to exclude the triplet $(0, 0, 0)$ for points. But we accept all the other triplets.

Prejudice 2. One knows what goes on at infinity.

Many people are so used with the standard geometry that they feel they know what happens at infinity and that it is a kind of experimental knowledge. But try to travel to infinity and come back and tell us the truth ! Infinity is a mathematic concept, certainly not an experimental one.

Definition. We call *line at infinity* the line described by any triplet $(0, 0, c)$ where $c \neq 0$.

Definition. We call *points at infinity* the points with homogenous coordinates (X, Y, Z) such that $Z = 0$ and $(X, Y) \neq (0, 0)$.

Theorem. A point is at infinity if and only if it is on the line at infinity.
Proof. The relation $aX + bY + cZ = 0$ for $a = b = 0$ and $c \neq$ has the solution $Z = 0$ and since we want $(X, Y, Z) \neq (0, 0, 0)$, we have $(X, Y) \neq (0, 0)$. $\square$

Provisional definition. A *projective plane* is caractrized by the three following sets :

1. the set of points $\mathcal{P}$ with homogenous coordinates (X, Y, Z) such that $(X, Y, Z) \neq (0, 0, 0)$
2. the set of lines $\mathcal{L}$ with homogeneous coordinates (a, b, c) such that $(a, b, c) \neq (0, 0, 0)$
3. the relation M is on d or d goes through M is defined by

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = 0$$

In a projective plane we have the following dual theorems :

Theorem. Given two distinct points, there is one and only one line going through these two points.

Theorem. Given two distinct lines, there is one and only one point on these two line.
Proof. We may now solve the same systems as before but without exception. $\square$

§ 2. The real projective plane

2.1 Definition

Let E be a 3-dimensional real vector space. We denote by 0 the null vector and we put $E^* = E \smallsetminus \{0\}$.

Definition. Two vectors u and v belonging to E are colinear if there is a real number k such that $u = kv$ or if there is a real number h such that $v = hu$.

Proposition. For any vector u in E, the vectors u and 0 are colinear.
Proof. We may write[1] $0 = 0\,u$. $\square$

Proposition. For any vectors u and v in E^*, the following assertions are equivalent

(i). the vectors u and v are colinear

(ii). there is a real number k such that $u = kv$

(iii). there is a real number h such that $v = hu$.

Proof. By definition of colinearity (i)$\Longrightarrow$(ii) and (i)$\Longrightarrow$(ii). Suppose (ii) ; since $u \in E^*$, $u \neq 0$ and thus $k \neq 0$ and we may multiply both sides of the equality by k^{-1} and thus (iii) is verified with $h = k^{-1}$. In the same way (iii)$\Longrightarrow$(ii). Finally (ii)$\Longrightarrow$(i) and (iii)$\Longrightarrow$(i) by definition. $\square$

[1] The symbol **0** has two different meanings. We should write : "null vector = real number zero $\boldsymbol{u}$".

Proposition and notation. The relation ”to be colinear” is an equivalence relation in E^*. We denote that equivalence relation by $\propto$.

Proof. It is trivial to check that the relation ”to be colinear” is reflexive (since $u = 1u$) and symmetrical. Let us show that the relation is transitive in E^* : suppose that u, v and w belong to E^* and that u and v are colinear and v and w are colinear. From the previous proposition, we have real numbers k and k' such that $v = ku$ and $w = k'v$, thus $w = (kk')u$ and u and w are colinear. □

Definition 1. The projective plane $\mathcal{P}$ associated to the 3-dimensional real vectorspace E is the set $E^0/\propto$. The elements of $\mathcal{P}$ are called points.

Thus a point in $\mathcal{P}$ is a set of colinear vectors different from 0. If M is a point in $\mathcal{P}$ and if u is a vector belonging to M then

$$M = \mathbb{R}^* u \qquad \text{where } \mathbb{R}^* = \mathbb{R} \setminus \{0\}$$

(Homogeneous) coordinates of a point in a projective plane

Definition. Let $(\Omega, \vec{\imath}, \vec{\jmath}, \vec{k})$ be a basis in E. Let $M \in \mathcal{P}$. We call *homogeneous coordinates* the coordinates (X, Y, Z) of any vector $u = X\vec{\imath} + Y\vec{\jmath} + Z\vec{k}$ belonging to M.

2.2 Equivalent definitions

The following definition is the simplest and the best one. It is clear that it is eqivalent to the definition 1.

Definition 2. The projective plane $\mathcal{P}$ associated to the 3-dimensional real vectorspace E is the set of 1-dimensional subspaces. The elements of $\mathcal{P}$ are called points.

Our third definition will be convenient to try to visualize the projective plane. It supposes that we have put a euclidean structure on the space E and to show the equivalence of that definition with the previous ones, one would have to show that the result is independant of the choice of the scalar product on E. We'll skip that proof and suppose it is evident

Definitions. Let E be a 3-dimensional euclidean linear space. The unit sphere Σ is the set of vectors with norm 1

$$\Sigma = \{u \in E \mid \|u\| = 1\}$$

If we choose an orthonormal basis in E, the sphere Σ has the equation

$$X^2 + Y^2 + Z^2 = 1$$

Two vectors u and v belonging to Σ are opposite if $u = -v$.

Proposition. The relation ”to be opposite” is an equivalence relation on the sphere. We'll denote it by $\leftrightarrow$. The classes of equivalence are the pairs of opposite vectors.

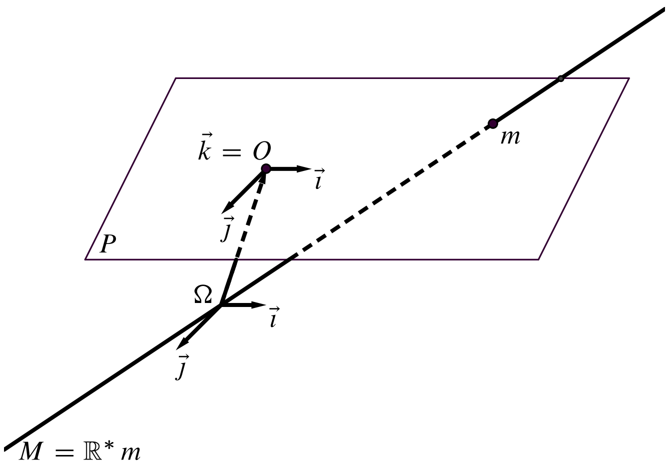
Definition 3. The projective plane $\mathcal{P}$ associated to the 3-dimensional real vectorspace E is the set $\Sigma/\leftrightarrow$. The elements of $\mathcal{P}$ are called points. Thus the points of $\mathcal{P}$ are (in this definition) pair of opposite vectors belonging to the unit sphere.

2.3 The affine plane associated to a frame of a projective plane

Let $(\Omega, \vec{\imath}, \vec{\jmath}, \vec{k})$ be a basis in E and let $\mathcal{P}$ be the projective plane associated to E. Let P be the plane in E which has the equation $Z = 1$. An element m in P is a vector in E and a point in P. For each point $m \in P$, there is one and only point M in $\mathcal{P}$ such that $m \in M$, that is the point $M = \mathbb{R}^* m$. If the coordinates of m are $(x, y, 1)$, the homogenous coordinates of M are $(X, Y, Z) = (kx, ky, k)$, where $k \in \mathbb{R}^*$.

The other way round is not so simple : given a point M in $\mathcal{P}$, is it possible to find a point m in the plane P such that $M = \mathbb{R}^* m$? If (X, Y, Z) are the homogenous coordinates of M, the coordinates of m has to be $(x, y, 1)$ such that $(X, Y, Z) = k(x, y, 1)$. That is possible if and only if $Z \neq 0$, and then $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$.

Thus we may consider the affine plane P as a subset of the projective plane $\mathcal{P}$.



Starting from the affine plane P to get the projective plane $\mathcal{P}$ we have to add all the projective points $\mathbb{R}^* u$ where the vector u is parallel to the plane P. That means that we have to add a projective line. We define projective lines in next paragraph.

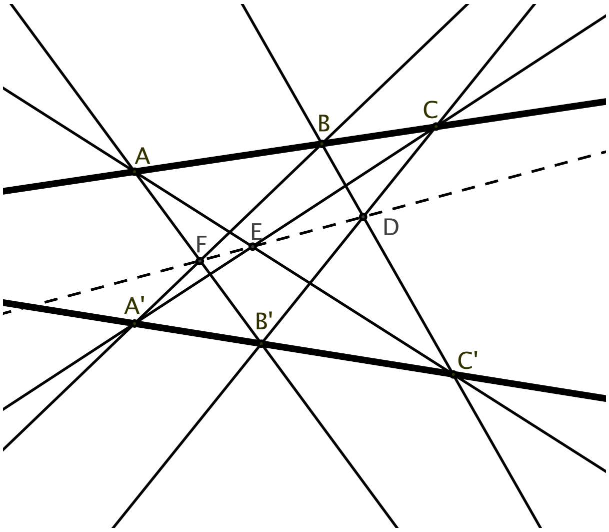
The general method to get results about the affine plane using the projective plane

Method. Task : to solve a problem in the affine plane.

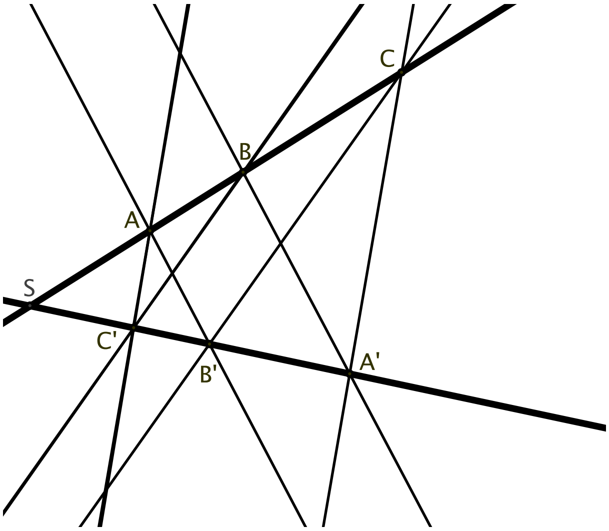
We add to the affine plane the line at infinity, getting a projective plane. Then we change the basis in the 3-dimensional space and we look at the new affine plane. The problem might be much simpler to solve in this new affine plane. Once it is solved, we move back to the original basis and the original affine plane.

Example. Some aspects of the following example of proof using projective geometry may seem awkward but it will be clearer later on. The purpose of having this example here is to show the efficiency of the theory as soon as possible

Pappus's theorem. . Let A, B and C be 3 aligned distinct points, A', B' and C' be 3 distinct points aligned on an other line. We call D the intersection of the lines BC' and $B'C$, E the intersection of the lines CA' and $C'A$ and F the intersection of the lines AB' and $A'B$. Then the points D, E and F are aligned.

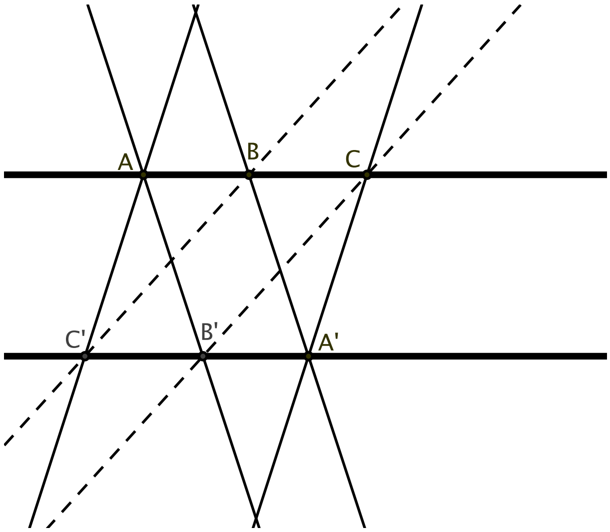


Proof. Take as line at infinity the line FE, we have to show that D is at infinity, that is that the lines BC' and $B'C$ are parallel. Just draw the picture : since E is at infinity, the lines CA' and AC' are parallel and since F is at infinity the lines AB' and BA' are parallel. We do not know if the lines AB and $A'B'$ are intersecting in a point colinear with F and E or not. So we have to look at two different cases. Let us first suppose that their intersection point S is not on the line FE, thus not at infinity. We get the following drawing :



Our problem has become : points S, A, B and C aligned, points S, A', B' and C' aligned. Knowing that $AC' \parallel CA'$ and $AB' \parallel BA'$, show that $BC' \parallel CB'$. The dilation δ_1 with center S that transforms B in A, transfroms A' into B' ; the dilation δ_2 with center S that transforms A in C, transfroms C' en A'. Since the dilations with same center commute, the dilation $\delta_2 \circ \delta_1$ is a dilation that transforms B into C and C' into B' and the line BC' into a parallel line CB'.

If the point S is on the line, that is S is at infinity, then the lines AB and $A'B'$ are parallel and we have the following picture :



The proof is the same as the previous one : you just have to change dilations into translations. □

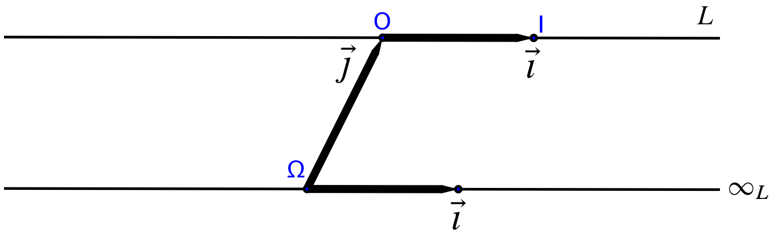
2.4 The real projective lines in the projective plane

We have allready used lines in the projective plane, but it was more or less intuitive, since we have'nt given any definition yet. First we'll define a real projectif line in itself and then we'll define the lines in the projective plane $\mathcal{P}$. For that purpose the definition 2 might be the easiest one.

Definition. Let F be a 2-dimensional real vector space (or linear space). The projective line $\mathcal{L}$ associated with F is the set of 1-dimensional subspaces.

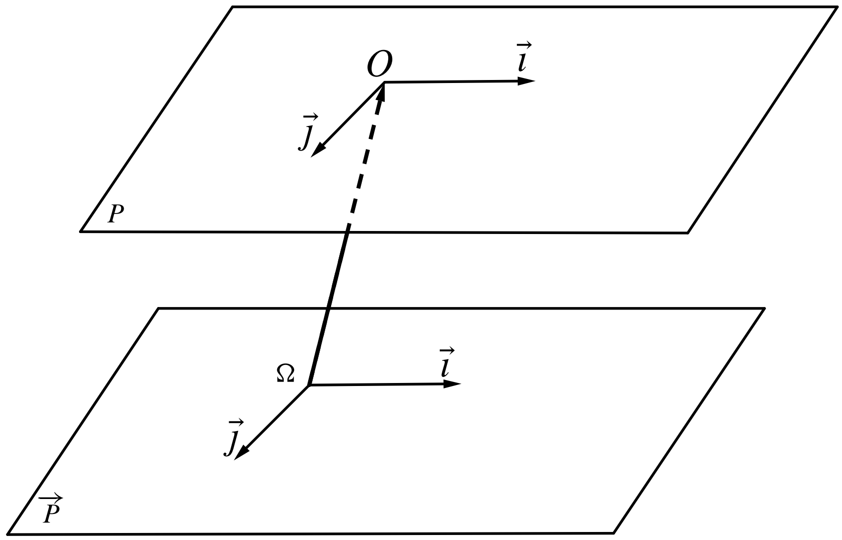
Let $(\vec{\imath}, \vec{\jmath})$ be a basis of F. The set of vectors with coordinates $(X, 1)$ is the set of points of an affine line L. Each point of L belongs to one and only point in $\mathcal{L}$. But there is one point in $\mathcal{L}$ that does not contain any elemet of L ; it is the subspace $\mathbb{R}\vec{\imath}$. Thus we may consider $\mathcal{L}$ to be L and one extra point. We call that extra point the point at infinity of L. Let us denote it by ∞_L, we have

$$\mathcal{L} = L \cup \{\infty_L\}$$



Definition. Let E be a 3-dimensional real vector space and let $\mathcal{P}$ be the projective plane associated with E. We call projective lines in $\mathcal{P}$ the projective lines $\mathcal{L}$ associated with the 2-dimesional subspaces F of E.

Notice that $\mathcal{P} = P \cup \mathcal{L}$, where $\mathcal{L}$ is the projective line associated with the 2-dimensional linear subspace $\overrightarrow{P}$ of E parallel to P. When you have the projective plane $\mathcal{P}$ you can choose any line as line "at infinity" and if you take away that line, what is left is an affine plane.



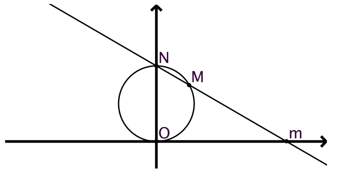
2.5 Desargues's theorem

Definitions. Let ABC and $A'B'C'$ be two triangles with sides which are lines denoted $a = BC$, $b = CA$, $c = AB$, $a' = B'C'$, $b' = C'A'$ and $c' = A'B'$. We say that the two triangles are *centraly perspective* if the lines AA', BB' and CC' are concurrent in a point S or parallel (that means concurrent at a point S at infinity). The two triangles are *linely perspective* if the intersection points of a and a', b and b' and c and c' are aligned on a line s or if $a /\!/ a'$, $b /\!/ b'$ and $c /\!/ c'$.

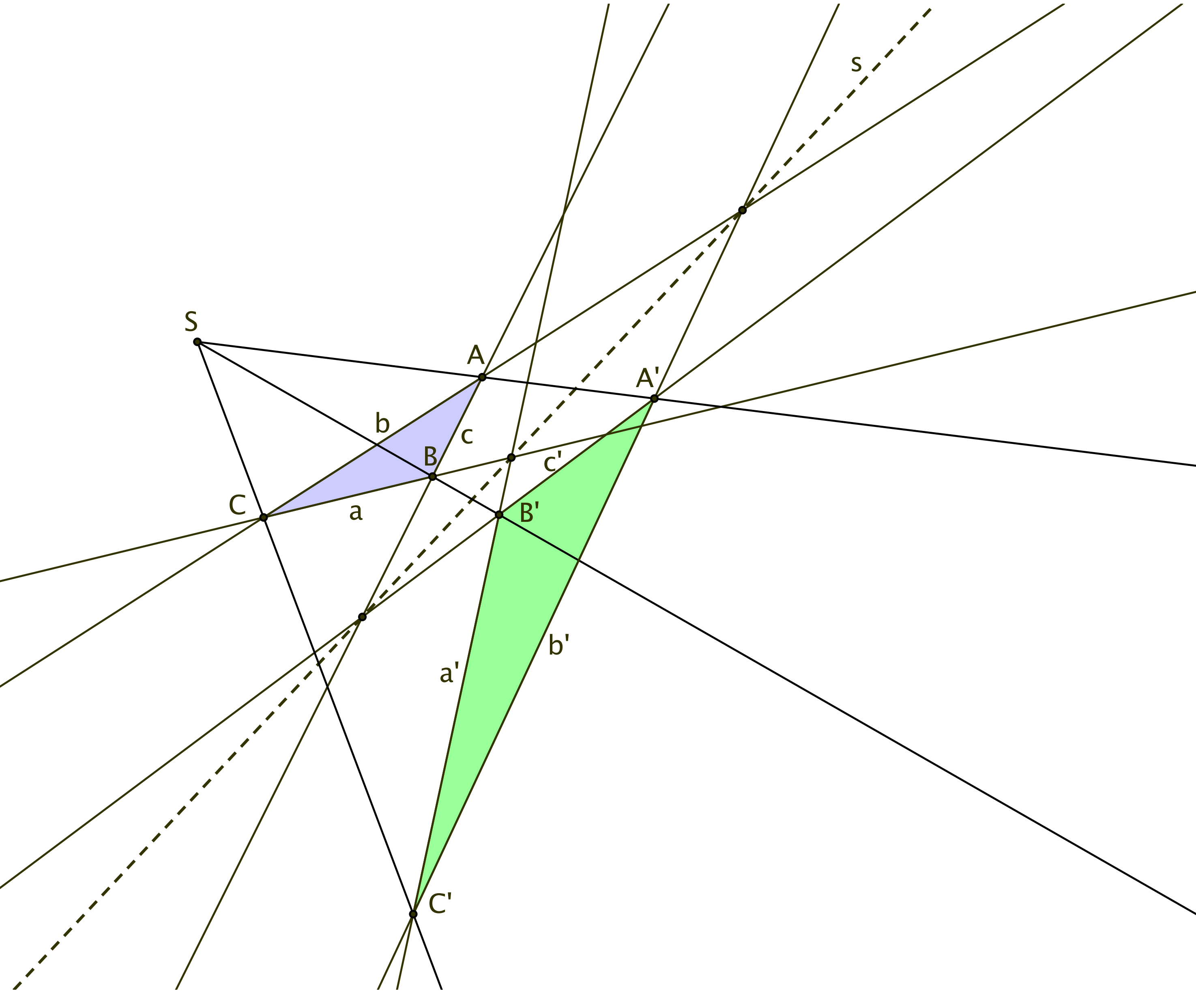
Theorem. Two triangles are centraly perspective if and only if they are linely perspective.

The proof is left as an exercise. Use the same method as for the proof of Pappus's theorem.

2.6 Topology of the real projective line and of the real projective plane

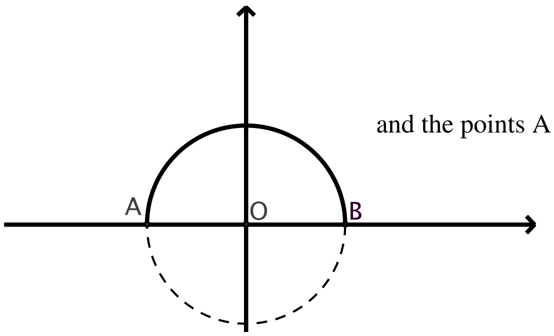


Let L be an affine line. Let us put it in an affine euclidean plane with orthonormal frame $(O, \vec{\imath}, \vec{\jmath})$ as the x-axis and draw the circle Γ with diameter the points $O(0,0)$ and $N(0,1)$. To each point m of L, we associate the point M on Γ such that N, M and m are aligned. We have a bijection. If we want to get a bijection with $\mathcal{L}$ we have to add the point N and thus we "see" that the real projective line "is" a circle.


s
S
A
A'
b
c
B
c'
C
a
B'
b'
a'
C'

Topology of the real projective line as a half circle

We may also consider the real projective line as a circle divided by the equivalence relation ”to be opposite to”. Thus the line becomes a half circle where we have to identify the two ends :



and the points A and B are the same point.

Topology of the real projective plane

We follow the same procedure as for the line, but now we start with a half sphere. But now the border is a circle and we have to glue together opposite points : that is not so easy because yo have to progress in the same direction. If you could go in opposite directions you would get a ”topological” sphere.

Another way to do it is to cut the sphere into three parts : a narrow belt arround the equator and two nearly halfspheres which are topologically equivalent to discs. These discs are opposite, so we can keep just one. It is a bit harder with the belt : you have to keep only half of it. Let us cut it along two opposite ”vertical” cuts. You get to vertical cuts and now you have to glue them one up one down getting a ... Möbius strip. Finally the projective plane is a Möbius strip glued to a disc allong its unique side. Not easy to see, but what is easy to understand is that the real projective line and the real projective plane are compact because closed and bounded in some space of sitable dimension.

§ 3. Generalisations

3.1 Projective spaces

Change the dimensions : linear space of dimension $n + 1$. The set of 1-dimensional subspaces is a projective space of dimension n.

3.2 Projective complex spaces

Change $\mathbb{R}$ into $\mathbb{C}$. For instance the projective complex line is the ”complex plane” to which is added ONE point : THE point at infinity. By stereographic projection it is the image of the North Pole.