Mathematical Visualization

Jan Techter

February 16, 2024

Lecture notes for Mathematical Visualization 1, Winter 2023/24.

Contents

1 Projective spaces

1.1 Some motivation: Incidences between points and lines

> The elementary figures of projective geometry are points, straight lines, and planes. The elementary results of projective geometry deal with the simplest possible relations between these entities, namely their *incidence*. The word incidence covers all the following relations: A point lying on a straight line, a point lying in a plane, a straight line lying in a plane. Clearly, the three statements that a straight line passes through a point, that a plane passes through a point, that a plane passes through a straight line, are respectively equivalent to the first three. The term incidence was introduced to give these three pairs of statements symmetrical form: a straight line is incident with a point, a plane is incident with a point, a plane is incident with a straight line. (Geometry and the Imagination – Hilbert, Cohn-Vossen)

In projective geometry, we are interested in statements and configurations that are invariant under *projective transformations*. E.g., the incidence of a point lying on a line is invariant under projection from one plane to another (from some point). Let us take a closer look at this incidence in the plane.

A point in the Euclidean plane $\mathbb{R}^2$ can be described by two Cartesian coordinates

$$p = (p_1, p_2) \in \mathbb{R}^2,$$

and a line by

$$\ell = \left\{ p = (p_1, p_2) \in \mathbb{R}^2 \;\middle|\; \langle n, p \rangle + h = 0 \right\}$$

with some $n = (n_1, n_2) \in \mathbb{S}^1 \backslash \{0\}$ and $h \in \mathbb{R}$, where n can be interpreted as the unit normal vector of ℓ and h as the oriented distance of the origin to ℓ.

Note that the equation for ℓ can be multiplied by any scalar $\lambda \in \mathbb{R}$, $\lambda \neq 0$ without changing the line. Thus, we can replace (n_1, n_2, h) by

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \lambda \begin{pmatrix} n_1 \\ n_2 \\ h \end{pmatrix}, \qquad \text{with some } \lambda \in \mathbb{R}, \lambda \neq 0,$$

and write the equation for the line as

$$a_1 p_1 + a_2 p_2 + a_3 = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ 1 \end{pmatrix} = 0$$

Similarly, we can replace $(p_1, p_2, 1)$ by any non-zero scalar multiple

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mu \begin{pmatrix} p_1 \\ p_2 \\ 1 \end{pmatrix}, \qquad \text{with some } \mu \in \mathbb{R}, \mu \neq 0,$$

from which the Cartesian coordinates of p can be recovered by

$$p_1 = \frac{x_1}{x_3}, \quad p_2 = \frac{x_2}{x_3}.$$

The triple (x_1, x_2, x_3), and in particular $(p_1, p_2, 1)$, are called *homogeneous coordinates* of p.

Now the equation of the *incidence* of the point p lying on the line ℓ ($p \in \ell$), or equivalently, the line ℓ passing through the point p ($\ell \ni p$) has the symmetric form

$$a_1x_1 + a_2x_2 + a_3x_3 = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = 0 \tag{1}$$

Example 1.1. How to determine if three points $p, q, r \in \mathbb{R}^2$ lie on a line?

Equation (1) is a linear homogeneous equation in (a_1, a_2, a_3). Thus, there exists a line passing through these three points if and only if the linear homogeneous system

$$\begin{pmatrix} p_1 & p_2 & 1 \\ q_1 & q_2 & 1 \\ r_1 & r_2 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = 0$$

has a non-trivial solution, which is equivalent to

$$\det \begin{pmatrix} p_1 & p_2 & 1 \\ q_1 & q_2 & 1 \\ r_1 & r_2 & 1 \end{pmatrix} = 0.$$

Example 1.2. How to compute the intersection point of two lines?

$$\ell = \left\{ p \in \mathbb{R}^2 \;\middle|\; a_1p_1 + a_2p_2 + a_3 = 0 \right\}$$
$$\tilde{\ell} = \left\{ p \in \mathbb{R}^2 \;\middle|\; \tilde{a}_1p_1 + \tilde{a}_2p_2 + \tilde{a}_3 = 0 \right\}$$

Its homogeneous coordinates are given by a solution of the linear homogeneous system

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ \tilde{a}_1 & \tilde{a}_2 & \tilde{a}_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0. \tag{2}$$

If we assume that the two lines are distinct, i.e., the two rows are independent, then the solution space is one-dimensional

$$\operatorname{span}\{x\} = \{\lambda x \mid \lambda \in \mathbb{R}\} \qquad \text{with some } x \in \mathbb{R}^3, x \neq 0,$$

and we obtain the intersection point $p \in \mathbb{R}^2$ with

$$p_1 = \frac{x_1}{x_3}, \quad p_2 = \frac{x_2}{x_3}.$$

What if $x_3 = 0$? Then

$$\det \begin{pmatrix} a_1 & a_2 \\ \tilde{a}_1 & \tilde{a}_2 \end{pmatrix} = 0$$

and thus ℓ and $\tilde{\ell}$ are parallel.

The linear homogeneous system (2) always has a solution. Thus, in homogeneous coordinates of the plane two lines always intersect. In particular, for two parallel lines, the point of intersection has homogeneous coordinates of the form $(x_1, x_2, 0)$ which represents a point not in $\mathbb{R}^2$, but "at infinity".

1.2 Definition of projective spaces

Let V be a vector space of dimension $n+1$ over a field $\mathbb{F}$. Then the *projective space* of V is the set

$$\mathrm{P}(V) := \{\text{1-dimensional subspaces of } V\}$$

Its dimension is given by

$$\dim \mathrm{P}(V) := \dim V - 1 = n.$$

For $x \in V\backslash\{0\}$ we write $[x] := \operatorname{span}\{x\}$. Then $[x]$ is a point in $\mathrm{P}(V)$, and x is called a *representative vector* for this point.

If $\lambda \in \mathbb{F}\backslash\{0\}$ then $[\lambda x] = [x]$, and λx is another representative vector for the same point. This defines an equivalence relation on $V\backslash\{0\}$

$$x \sim y \quad \Leftrightarrow \quad x = \lambda y, \quad \text{for some } \lambda \in \mathbb{F}\backslash\{0\},$$

and we can identify

$$\mathrm{P}(V) \cong {}^{(V\backslash\{0\})}\!/_{\sim}.$$

For now we will only consider the real projective space

$$\mathbb{R}\mathrm{P}^n := \mathrm{P}(\mathbb{R}^{n+1}).$$

1.3 Homogeneous coordinates on $\mathbb{R}\mathrm{P}^n$

For a point $[x_1, \ldots, x_{n+1}] \in \mathbb{R}\mathrm{P}^n$ the coordinates of a representative vector $(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1}$ are called *homogeneous coordinates*. They are unique up to a common scalar multiple

$$[x_1, \ldots, x_{n+1}] = [\lambda x_1, \ldots, \lambda x_{n+1}]$$

for $\lambda \in \mathbb{R}\backslash\{0\}$.

If $x_{n+1} \neq 0$ then

$$[x_1, \ldots, x_{n+1}] = \left[\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}}, 1\right] = [y_1, \ldots, y_n, 1],$$

and $(y_1, \ldots, y_n)$ are called *affine coordinates* of the point $[x]$. This yields a decomposition of $\mathbb{R}\mathrm{P}^n$ into an affine part and a *hyperplane at infinity*

$$\mathbb{R}\mathrm{P}^n = \underbrace{\{[x_1, \ldots, x_{n+1}] \mid x_{n+1} \neq 0\}}_{\simeq \mathbb{R}^n} \cup \underbrace{\{[x_1, \ldots, x_{n+1}] \mid x_{n+1} = 0\}}_{\simeq \mathbb{R}\mathrm{P}^{n-1}}.$$

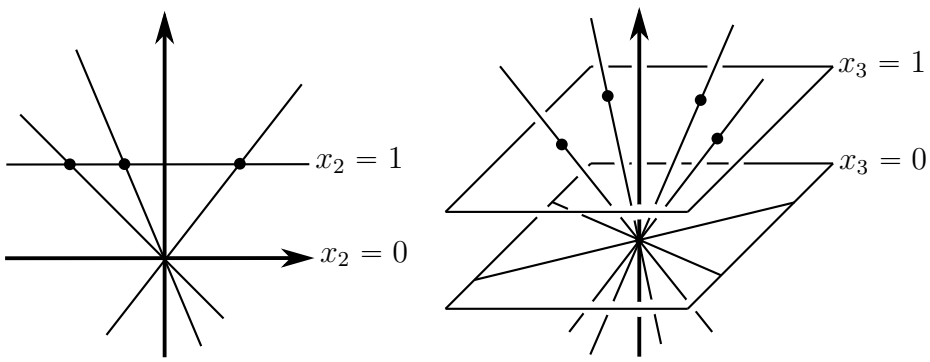


Figure 1. Affine coordinates for $\mathbb{R}\mathrm{P}^1$ and $\mathbb{R}\mathrm{P}^2$.

Example 1.3 (The real projective line $\mathbb{R}\mathrm{P}^1$)**.** For the real projective line this decomposition is given by

$$\mathbb{R}\mathrm{P}^1 \cong \mathbb{R} \cup \mathbb{R}\mathrm{P}^0 = \mathbb{R} \cup \{\infty\},$$

where $\mathbb{R}\mathrm{P}^0$ consists of only one point $[1, 0]$, which is usually denoted by ∞, and allowed as an "admissible" affine coordinate.

Example 1.4 (The real projective plane $\mathbb{R}\mathrm{P}^2$)**.** For the real projective plane this decomposition is given by

$$\mathbb{R}\mathrm{P}^2 \cong \mathbb{R}^2 \cup \mathbb{R}\mathrm{P}^1.$$

Thus, we obtain the Euclidean plane compactified by a (projective) line at infinity.

Example 1.5 (The real projective 3-space $\mathbb{R}\mathrm{P}^3$)**.** For the real projective plane this decomposition is given by

$$\mathbb{R}\mathrm{P}^3 \cong \mathbb{R}^3 \cup \mathbb{R}\mathrm{P}^2.$$

Thus, we obtain the Euclidean 3-space compactified by a (projective) plane at infinity.

More generally, let $b_1, \ldots, b_{n+1}$ be a basis of $\mathbb{R}^{n+1}$. For $x \in \mathbb{R}^{n+1}$ let $x_1, \ldots, x_{n+1} \in \mathbb{R}$ such that

$$x = \sum_{i=1}^{n+1} x_i b_i.$$

Then $(x_1, \ldots, x_{n+1})$ are called *homogeneous coordinates* of the point $[x] \in \mathbb{R}\mathrm{P}^n$ (with respect to $b_1, \ldots, b_{n+1}$). They depend on the chosen basis and are unique up to a common scalar multiple. We then identify

$$[x] \cong [x_1, \ldots, x_{n+1}].$$

A change of basis acts on the homogeneous coordinates as a general linear transformation

$$\begin{bmatrix} x_1 \\ \vdots \\ x_{n+1} \end{bmatrix} \mapsto \left[A \begin{pmatrix} x_1 \\ \vdots \\ x_{n+1} \end{pmatrix} \right]$$

with $A \in \mathrm{GL}(\mathbb{R}^{n+1})$.

1.4 Projective subspaces

For a $(k+1)$-dimensional linear subspace $U \subset \mathbb{R}^{n+1}$ its projective space

$$\mathrm{P}(U) \subset \mathbb{R}\mathrm{P}^n$$

is called a k-dimensional *projective subspace* of $\mathbb{R}\mathrm{P}^n$.

$\dim \mathrm{P}(U)$	**name**
0	point
1	line
2	plane
k	k-plane
$n-1$	hyperplane

Table 1. Naming conventions for projective (sub)spaces.

1.5 Meet and join

Let $\mathrm{P}(U_1), \mathrm{P}(U_2) \subset \mathbb{R}\mathrm{P}^n$ be two projective subspaces. Then their intersection, or *meet*, is given by

$$\mathrm{P}(U_1) \cap \mathrm{P}(U_2) = \mathrm{P}(U_1 \cap U_2),$$

and their span, or *join*, is given by

$$\mathrm{P}(U_1) \vee \mathrm{P}(U_2) = \mathrm{P}(U_1 + U_2).$$

The dimension formula for linear subspaces carries over to projective subspaces:

$$\dim\left(\mathrm{P}(U_1) \vee \mathrm{P}(U_2)\right) + \dim\left(\mathrm{P}(U_1) \cap \mathrm{P}(U_2)\right) = \dim \mathrm{P}(U_1) + \dim \mathrm{P}(U_2).$$

In particular, a k_1-plane and a k_2-plane in an n-dimensional projective space with $k_1 + k_2 \geqslant n$ always intersect in an at least $(k_1 + k_2 - n)$-dimensional projective subspace. Thus, certain incidences are always guaranteed in a projective space.

Example 1.6 ($\mathbb{R}\mathrm{P}^2$)**.** In $\mathbb{R}\mathrm{P}^2$ two (distinct) lines always intersect in a point. In affine coordinates, the two lines are parallel if and only if the intersection point lies on the line at infinity.

Example 1.7 ($\mathbb{R}\mathrm{P}^3$)**.** In $\mathbb{R}\mathrm{P}^3$ two (distinct) planes always intersect in a line. In affine coordinates, the two planes are parallel if and only if the intersection line lies in the plane at infinity.

However, in $\mathbb{R}\mathrm{P}^3$, two lines do not always intersect. They intersect if and only if they lie in a plane. In affine coordinates, two lines are parallel if and only if the intersection point lies in the plane at infinity.

1.6 Desargues' theorem

An *incidence theorem* is a statement about a projective configuration (of e.g. projective subspaces) where a certain set of incidences implies another set of incidences. As an example we state the theorem of Desargues. First in $\mathbb{R}\mathrm{P}^3$ where it is very easy to verify, and then in $\mathbb{R}\mathrm{P}^2$.

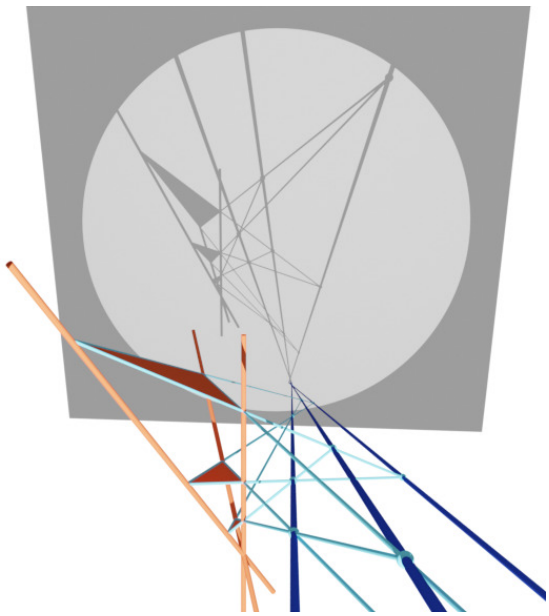


Figure 2. Three triangles in perspective and their shadow.

Theorem 1.1 (Desargues)**.** *Let A, A', B, B', C, C' be six points in $\mathbb{R}\mathrm{P}^3$, such that A, B, C span a plane, and A', B', C' span another plane.*

If the three lines AA', BB', and CC' pass through a common point, then the three points $A'' = BC \cap B'C'$, $B'' = CA \cap C'A'$, and $C'' = AB \cap A'B'$ lie on a common line.

Proof. First, the statement contains the implicit claim, that, e.g., the lines BC and $B'C'$ intersect in a point. Indeed, the four points B, C, B', C' lie in a plane since BB' and CC' are concurrent. Thus, the point $A'' = BC \cap B'C'$ exists.

The two planes

$$E = A \vee B \vee C, \qquad E' = A' \vee B' \vee C'$$

intersect in a line $\ell = E \cap E'$. Since $BC \in E$ and $B'C' \in E'$, their intersection point A'' lies in ℓ. Similarly, $B'', C'' \in \ell$. □

Consider what happens if we project such a configuration in $\mathbb{R}\mathrm{P}^3$ from a point into a plane, and denote the image points by $\tilde{A}, \tilde{B}, \tilde{C}, \ldots$. Then we obtain again six points $\tilde{A}$, $\tilde{A}'$, $\tilde{B}$, $\tilde{B}'$, $\tilde{C}$, $\tilde{C}'$ that satisfy that the lines $\tilde{A}\tilde{A}'$, $\tilde{B}\tilde{B}'$, and $\tilde{C}\tilde{C}'$ are concurrent and that the points $\tilde{A}'' = \tilde{B}\tilde{C} \cap \tilde{B}'\tilde{C}'$, $\tilde{B}'' = \tilde{C}\tilde{A} \cap \tilde{C}'\tilde{A}'$, and $\tilde{C}'' = \tilde{A}\tilde{B} \cap \tilde{A}'\tilde{B}'$ are collinear.

Indeed, Desargues theorem also holds in $\mathbb{R}\mathrm{P}^2$ which can be shown by lifting it to $\mathbb{R}\mathrm{P}^3$.

Theorem 1.2. *Let A, A', B, B', C, C' be six points in $\mathbb{R}\mathrm{P}^2$, such that no three lie on a line.*

If the three lines AA', BB', and CC' pass through a common point, then the three points $A'' = BC \cap B'C'$, $B'' = CA \cap C'A'$, and $C'' = AB \cap A'B'$ lie on a common line.

Proof. We embed $\mathbb{R}\mathrm{P}^2$ into $\mathbb{R}\mathrm{P}^3$ as the plane $\mathbb{R}\mathrm{P}^2 \cong E \subset \mathbb{R}\mathrm{P}^3$. Thus, E is the plane which contains the two triangles ABC, $A'B'C'$, and the point P which is incident with the three lines AA', BB', and CC'.

Choose a line through P which is not in E and two points X and Y on it.

The lines XA and YA' lie in a plane, so they intersect in a point $\tilde{A}$. Thus,

$$\tilde{A} = XA \cap YA',$$

and similarly

$$\tilde{B} = XB \cap YB',$$
$$\tilde{C} = XC \cap YC'.$$

Now A, B, C span E and $\tilde{A}, \tilde{B}, \tilde{C}$ span another plane $\tilde{E}$. The three lines $A\tilde{A}$, $B\tilde{B}$, and $C\tilde{C}$ pass through a common point (namely X). Thus, we can apply Theorem 1.1 to the six points $A, \tilde{A}, B, \tilde{B}, C, \tilde{C}$, and find that the line of intersection $E \cap \tilde{E}$ contains

$$A'' = BC \cap \tilde{B}\tilde{C} = BC \cap B'C',$$

and similarly B'' and C''. □

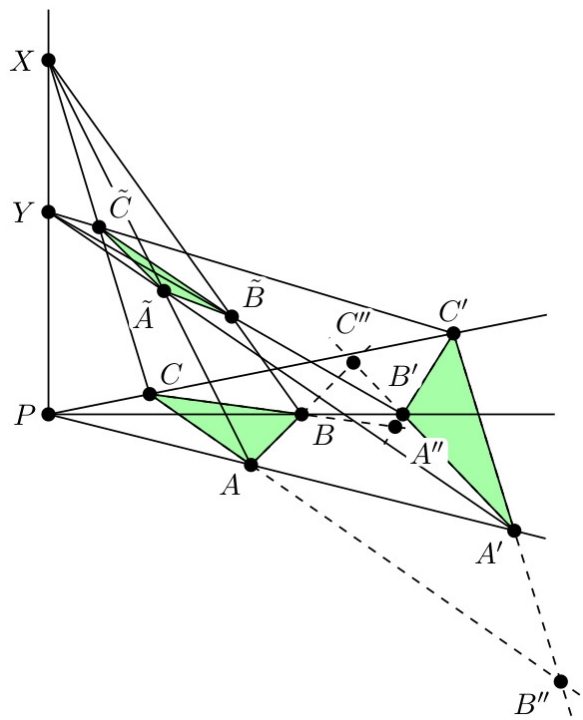


Figure 3. Desargues' theorem in $\mathbb{R}\mathrm{P}^2$ from Desargues' theorem in $\mathbb{R}\mathrm{P}^3$.

2 Duality

As we have seen in Section 1.1, in homogeneous coordinates x_1, x_2, x_3, the equation for a line in a projective plane is

$$a_1x_1 + a_2x_2 + a_3x_3 = 0,$$

where not all coefficients a_i are zero. The coefficients a_1, a_2, a_3 can be seen as homogeneous coordinates for the line, because if we replace in the equation a_i by λa_i for some $\lambda \neq 0$ we get an equivalent equation for the same line. Thus, the set of lines in a projective plane is itself a projective plane, the *dual plane*. Points in the dual plane correspond to lines in the original plane. Moreover, if we consider in the above equation the x_i as fixed and the a_i as variables, we get an equation for a line in the dual plane. Points on this line correspond to lines in the original plane that contain $[x]$. Thus, a the points on a line in the dual plane correspond to lines in the original plane through a point.

It makes sense to look at this phenomenon in a basis independent way and for arbitrary dimension. It boils down to the duality of vector spaces.

2.1 Dual space

The *dual vector space* of $\mathbb{R}^{n+1}$ is the space of linear functionals $\mathbb{R}^{n+1} \to \mathbb{R}$

$$(\mathbb{R}^{n+1})^* := \left\{ a \;\middle|\; a : \mathbb{R}^{n+1} \to \mathbb{R} \text{ linear} \right\}.$$

The *dual projective space* of $\mathbb{R}\mathrm{P}^n$ is correspondingly defined by

$$(\mathbb{R}\mathrm{P}^n)^* := \mathrm{P}((\mathbb{R}^{n+1})^*).$$

The natural identification $(\mathbb{R}^{n+1})^{**} = \mathbb{R}^{n+1}$ carries over to the projective setting $(\mathbb{R}\mathrm{P}^n)^{**} = \mathbb{R}\mathrm{P}^n$.

2.2 Dual subspaces

For a projective subspace $\mathrm{P}(U) \subset \mathbb{R}\mathrm{P}^n$ its *dual projective subspace* $\mathrm{P}(U)^\star \subset (\mathbb{R}\mathrm{P}^n)^*$ is defined by

$$\mathrm{P}(U)^\star := \{[a] \in (\mathbb{R}\mathrm{P}^n)^* \mid a(x) = 0 \text{ for all } x \in U\}.$$

The dimensions of a projective subspace and its dual projective subspace are related by

$$\dim \mathrm{P}(U) + \dim \mathrm{P}(U)^\star = n - 1.$$

Incidences are reversed by duality

$$\mathrm{P}(U_1) \subset \mathrm{P}(U_2) \quad \Leftrightarrow \quad \mathrm{P}(U_2)^\star \subset \mathrm{P}(U_1)^\star.$$

and meet and join are interchanged

$$\begin{aligned}(\mathrm{P}(U_1) \vee \mathrm{P}(U_2))^\star &= \mathrm{P}(U_1)^\star \cap \mathrm{P}(U_2)^\star,\\ (\mathrm{P}(U_1) \cap \mathrm{P}(U_2))^\star &= \mathrm{P}(U_1)^\star \vee \mathrm{P}(U_2)^\star.\end{aligned}$$

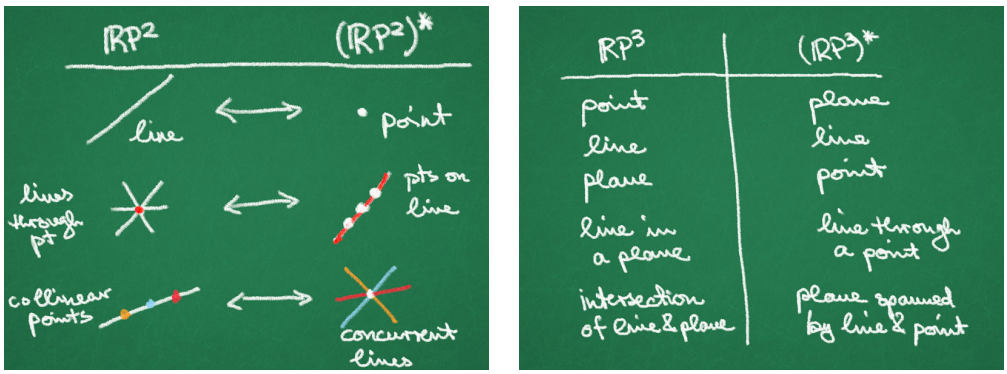


Figure 4. Duality in $\mathbb{R}\mathrm{P}^2$ and $\mathbb{R}\mathrm{P}^3$.

2.3 Duality in coordinates

Let $b_1, \ldots, b_{n+1}$ be a basis of $\mathbb{R}^{n+1}$ and $b_1^*, \ldots, b_{n+1}^*$ the corresponding dual basis of $(\mathbb{R}^{n+1})^*$, i.e.,

$$b_i^*(b_j) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j. \end{cases}$$

In homogeneous coordinates with respect to those bases the duality of two points

$$[x_1, \ldots, x_{n+1}] \cong [x] \in \mathbb{R}\mathrm{P}^n, \qquad [a_1, \ldots, a_{n+1}] \cong [a] \in (\mathbb{R}\mathrm{P}^n)^*$$

is expressed by

$$a(x) = (a_1 \ldots a_{n+1}) \begin{pmatrix} x_1 \\ \vdots \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_{n+1} \end{pmatrix}^\mathsf{T} \begin{pmatrix} x_1 \\ \vdots \\ x_{n+1} \end{pmatrix} = 0.$$

Thus, duality in linear algebra as well as in projective geometry expresses in a formal way that a subspace can either be expressed as the span of points or the solutions to a set of linear equations.

If a change of basis acts on the homogeneous coordinates of $\mathbb{R}\mathrm{P}^n$ as

$$\begin{bmatrix} x_1 \\ \vdots \\ x_{n+1} \end{bmatrix} \mapsto \left[A \begin{pmatrix} x_1 \\ \vdots \\ x_{n+1} \end{pmatrix} \right]$$

with $A \in \mathrm{GL}(\mathbb{R}^{n+1})$, it acts on the homogeneous coordinates of the dual space $(\mathbb{R}\mathrm{P}^n)^*$ as

$$\begin{bmatrix} a_1 \\ \vdots \\ a_{n+1} \end{bmatrix} \mapsto \left[A^{-\intercal} \begin{pmatrix} a_1 \\ \vdots \\ a_{n+1} \end{pmatrix} \right].$$

2.4 The dual of Desargues' theorem

> The interchangeability of points and lines is called the principle of *duality* in the projective plane. According to this principle, there belongs to every theorem a second theorem that corresponds to it dually, and to every figure a second figure that corresponds to it dually. (Geometry and the Imagination – Hilbert, Cohn-Vossen)

As an example consider the theorem of Desargues in in $\mathbb{R}\mathrm{P}^2$ (Theorem 1.2). Then its dual turns out to be the converse statement, which therefore also holds.

3 Projective transformations

Let $F \in \mathrm{GL}(\mathbb{R}^{n+1})$ an invertible linear transformation. Then the map

$$[F] : \mathbb{R}\mathrm{P}^n \to \mathbb{R}\mathrm{P}^n, \quad [v] \mapsto [F(v)]$$

is called a *projective transformation.*

Proposition 3.1.

(i) Projective transformations are well-defined maps (do not depend on the representative vectors of points).

(ii) For $F, G \in \mathrm{GL}(\mathbb{R}^{n+1})$

$$[F] = [G] \quad \Leftrightarrow \quad G = \lambda F \text{ with some } \lambda \in \mathbb{R}, \lambda \neq 0.$$

(iii) Projective transformations map projective subspaces to projective subspaces, while preserving their dimension and incidences.

(iv) Vice versa, any bijective map on $\mathbb{R}\mathrm{P}^n$, $n \geqslant 2$, that maps lines to lines is a projective transformation.

(v) Let $A_1, \ldots, A_{n+2} \in \mathbb{R}\mathrm{P}^n$ be $n+2$ points in general position, and let $B_1, \ldots, B_{n+2} \in \mathbb{R}\mathrm{P}^n$ be $n+2$ points in general position. Then there exists a unique projective transformation

$$f : \mathbb{R}\mathrm{P}^n \to \mathbb{R}\mathrm{P}^n \quad \text{with} \quad f(A_i) = B_i \text{ for } i = 1, \ldots, n+2.$$

(vi) Projective transformations preserve the cross-ratio of four points on a line.

3.1 Projective transformations in homogeneous coordinates

In homogeneous coordinates a projective transformation $[F] : \mathbb{R}\mathrm{P}^n \to \mathbb{R}\mathrm{P}^n$ is represented by a non-singular matrix $F \in \mathbb{R}^{(n+1)\times(n+1)}$ (up to non-zero scalar multiples).

For representative vectors $x = (u_1, \ldots, u_n, 1)$ and with

$$F = \left(\begin{array}{c|c} A & b \\ \hline c^\intercal & d \end{array}\right) \qquad \text{where } A \in \mathbb{R}^{n\times n}, b, c \in \mathbb{R}^n, d \in \mathbb{R}$$

we obtain

$$F(x) = \left(\begin{array}{c|c} A & b \\ \hline c^\intercal & d \end{array}\right) \begin{pmatrix} u \\ 1 \end{pmatrix} = \begin{pmatrix} Au + b \\ c^\intercal u + d \end{pmatrix} \sim \begin{pmatrix} \frac{Au+b}{c^\intercal u+d} \\ 1 \end{pmatrix}$$

if $c^\intercal u + d \neq 0$. Thus, in affine coordinates, projective transformations are *fractional linear transformations*:

$$\mathbb{R}^n \to \mathbb{R}^n \qquad u \mapsto \tfrac{Au+b}{c^\intercal u+d}$$

3.2 Affine transformations

If we choose a representative matrix of the form

$$F = \left(\begin{array}{c|c} A & b \\ \hline 0 & 1 \end{array}\right) \qquad \text{where } A \in \mathrm{GL}(\mathbb{R}^n), b \in \mathbb{R}^n,$$

we obtain

$$\left(\begin{array}{c|c} A & b \\ \hline 0 & 1 \end{array}\right) \begin{pmatrix} u \\ 1 \end{pmatrix} = \begin{pmatrix} Au + b \\ 1 \end{pmatrix}$$

In affine coordinates, this in an *affine transformation*

$$\mathbb{R}^n \to \mathbb{R}^n \qquad u \mapsto Au + b$$

Thus, affine transformations are projective transformations.

Note that affine transformations map the hyperplane at infinity $\{[x] \in \mathbb{R}\mathrm{P}^n \mid x_{n+1} = 0\}$ to itself:

$$\left(\begin{array}{c|c} A & b \\ \hline 0 & 1 \end{array}\right) \begin{pmatrix} u \\ 0 \end{pmatrix} = \begin{pmatrix} Au + b \\ 0 \end{pmatrix}$$

In fact, affine transformations are characterized by this property among the projective transformations.

Proposition 3.2. *A projective transformation $f : \mathbb{R}\mathrm{P}^n \to \mathbb{R}\mathrm{P}^n$ is an affine transformation if and only if f maps the hyperplane at infinity $\{[x] \in \mathbb{R}\mathrm{P}^n \mid x_{n+1} = 0\}$ to itself.*

3.3 Euclidean transformations

Euclidean transformations are affine transformations, and thus, projective transformations. Indeed, if we choose a representative matrix of the form

$$F = \left(\begin{array}{c|c} A & b \\ \hline 0 & 1 \end{array}\right) \qquad \text{where } A \in \mathrm{O}(n), b \in \mathbb{R}^n,$$

in affine coordinates, this is a Euclidean transformation.

Example 3.1 (reflection in a line)**.** Consider a line with unit normal $n = (n_1, n_2) \in \mathbb{S}^1$ through the point $q \in \mathbb{R}^2$

$$\ell = \left\{ u = (u_1, u_2) \in \mathbb{R}^2 \;\middle|\; \langle n, u - q \rangle = 0 \right\}$$

Then the (Euclidean) reflection $\hat{\sigma} : \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$\hat{\sigma}(u) = u - 2\langle u - q, n \rangle n$$

With $h := -\langle q, n \rangle$ the equation for the line becomes

$$\langle n, u \rangle + h = 0$$

and the reflection can be rewritten as

$$\hat{\sigma}(u) = u - 2\langle u, n \rangle n - 2hn = (I - 2nn^\intercal)u - 2hn$$

Thus, in homogeneous coordinates we can write

$$\begin{pmatrix} \hat{\sigma}(u) \\ 1 \end{pmatrix} = \underbrace{\left(\begin{array}{c|c} I - 2nn^\intercal & -2hn \\ \hline 0 & 1 \end{array}\right)}_{=:F} \begin{pmatrix} u \\ 1 \end{pmatrix},$$

where, indeed, $I - 2nn^\intercal \in \mathrm{O}(2)$. As an extension of $\hat{\sigma}$, we can now define a projective transformation $\sigma : \mathbb{R}\mathrm{P}^2 \to \mathbb{R}\mathrm{P}^2$ by $\sigma([x]) = [Fx]$. Note that $F^2 = I$ and thus σ is an involution: $\sigma \circ \sigma = \mathrm{id}$.

Let us also derive the matrix F for the reflection in the case that the line is given in homogeneous coordinates

$$\ell = \left\{ [x] \in \mathbb{R}\mathrm{P}^2 \;\middle|\; a^\intercal x = a_1 x_1 + a_2 x_2 + a_3 x_3 = 0 \right\} = [a]^\star \qquad \text{with some } a \in \mathbb{R}^3 \backslash \{0\}$$

With $\hat{a} := (a_1, a_2)$ and $|\hat{a}| \neq 0$ it relates to the Euclidean equation by

$$n = \frac{\hat{a}}{|\hat{a}|}, \quad h = \frac{a_3}{|\hat{a}|}.$$

Thus,

$$F = \left(\begin{array}{c|c} I - 2\frac{\hat{a}\hat{a}^\intercal}{|\hat{a}|^2} & -2\frac{a_3\hat{a}}{|\hat{a}|^2} \\ \hline 0 & 1 \end{array}\right) \sim \left(\begin{array}{c|c} |\hat{a}|^2 I - 2\hat{a}\hat{a}^\intercal & -2a_3\hat{a} \\ \hline 0 & |\hat{a}|^2 \end{array}\right)$$

Note that this formula easily generalizes to the (Euclidean) reflection in a hyperplane in $\mathbb{R}^n \subset \mathbb{R}\mathrm{P}^n$ given by

$$L = \{ [x] \in \mathbb{R}\mathrm{P}^n \mid a^\intercal x = 0 \} = [a]^\star,$$

which yields

$$F = \left(\begin{array}{c|c} |\hat{a}|^2 I - 2\hat{a}\hat{a}^\intercal & -2a_{n+1}\hat{a} \\ \hline 0 & |\hat{a}|^2 \end{array}\right),$$

where $\hat{a} = (a_1, \ldots, a_n)$.

3.4 Central projections

Another important class of projective transformations are projections.

Example 3.2 (orthogonal projection to a line)**.** Consider a line

$$\ell = \left\{ u = (u_1, u_2) \in \mathbb{R}^2 \;\middle|\; \langle n, u - q \rangle = \langle n, u \rangle + h = 0 \right\},$$

with some $n \in \mathbb{S}^1$, $q \in \mathbb{R}^2$, and $h = -\langle n, q \rangle$. Then the orthogonal projection $\hat{\pi} : \mathbb{R}^2 \to \ell$ is given by

$$\hat{\pi}(u) = u - \langle u - q, n \rangle n = u - \langle u, n \rangle n - hn = (I - nn^\intercal)u - hn$$

Thus, in homogeneous coordinates we can write

$$\begin{pmatrix} \hat{\sigma}(u) \\ 1 \end{pmatrix} = \underbrace{\left(\begin{array}{c|c} I - nn^\intercal & -hn \\ \hline 0 & 1 \end{array}\right)}_{=:F} \begin{pmatrix} u \\ 1 \end{pmatrix}.$$

Note that here F is not invertible, since in particular $F\left(\begin{smallmatrix} n \\ 0 \end{smallmatrix}\right) = 0$. Thus, we can be extend $\hat{\pi}$ to a map

$$\pi : \mathbb{R}\mathrm{P}^2 \backslash \{[\begin{smallmatrix} n \\ 0 \end{smallmatrix}]\} \to \ell$$

by $\pi([x]) = [Fx]$. Since π is not invertible, it does not constitute a projective transformation. But the restriction of π to any line (that does not contain $[\begin{smallmatrix} n \\ 0 \end{smallmatrix}]$ is.

Similar, to Example 3.1, this can easily be generalized to the orthogonal projection onto a hyperplane in $\mathbb{R}^n \subset \mathbb{R}\mathrm{P}^n$ given by

$$L = \{[x] \in \mathbb{R}\mathrm{P}^n \mid a^\intercal x = 0\} = [a]^\star,$$

which yields

$$F = \left(\begin{array}{c|c} |\hat{a}|^2 I - \hat{a}\hat{a}^\intercal & -a_{n+1}\hat{a} \\ \hline 0 & |\hat{a}|^2 \end{array}\right), \tag{3}$$

where $\hat{a} = (a_1, \ldots, a_n)$.

More generally, let $L \subset \mathbb{R}\mathrm{P}^n$ be a hyperplane and $P \in \mathbb{R}\mathrm{P}^n$ a point $P \notin L$. Then the *central projection* to L with center P is given by

$$\pi : \mathbb{R}\mathrm{P}^n \backslash \{P\} \to L, \qquad X \mapsto (P \vee X) \cap L$$

P and X span a line, since $X \neq P$. This line intersects L in exactly one point, since $P \notin L$. Thus, this map is well-defined.

Let us show that π is indeed a given by a linear map on the representative vectors. Let the hyperplane L be given by

$$L = \{[x] \in \mathbb{R}\mathrm{P}^n \mid a(x) = 0\} = [a]^\star \qquad \text{with some } a \in (\mathbb{R}^{n+1})^* \backslash \{0\}.$$

The image of a point $X = [x] \neq P = [p]$ lies on the line

$$X \vee P = \mathrm{P}(\left\{\lambda x + \mu p \;\middle|\; \lambda, \mu \in \mathbb{R}^2\right\}).$$

Thus, the intersection $(X \vee P) \cap L$ is determined by the condition

$$a(\lambda x + \mu p) = \lambda a(x) + \mu a(p) = 0$$

With $\lambda = a(p)$ and $\mu = -a(x)$, we obtain

$$\pi([x]) = [a(p)x - a(x)p],$$

which is indeed linear in x.

Again, this linear map is not invertible, since p is in its kernel. Furthermore, $\dim \mathbb{RP}^n = n > \dim L = n - 1$. Yet the map becomes a projective transformation once we restrict it to another hyperplane K with $P \notin K$:

$$\pi : K \to L \qquad X = [x] \mapsto (P \vee X) \cap L = [a(p)x - a(x)p]$$

To see that now it is invertible, note that $\dim K = \dim L$. Further $a(p)x - a(x)p = 0$ implies $x = 0$, otherwise we would have $[x] = [p]$, which contradicts $P \notin K$.

In homogeneous coordinates, we can write the representative matrix for the central projection as

$$F = a^\intercal p I - p a^\intercal.$$

Example 3.3 (orthogonal projection as central projection)**.** Let us recover the orthogonal projection from Example 3.2 as central projection with center at infinity.

Consider the hyperplane

$$L = \{[x] \in \mathbb{RP}^n \mid a^\intercal x = 0\} = [a]^\star \qquad \text{with some } a \in (\mathbb{R}^{n+1})^*\backslash\{0\}.$$

and $P = [p] = [\hat{a}, 0] = [a_1, \ldots, a_n, 0]$. Then

$$\begin{aligned} F &= a^\intercal p I - p a^\intercal \\ &= \begin{pmatrix} \hat{a}^\intercal & a_{n+1} \end{pmatrix} \begin{pmatrix} \hat{a} \\ 0 \end{pmatrix} I - \begin{pmatrix} \hat{a} \\ 0 \end{pmatrix} \begin{pmatrix} \hat{a}^\intercal & a_{n+1} \end{pmatrix} \\ &= |\hat{a}|^2 I - \left(\begin{array}{c|c} \hat{a}\hat{a}^\intercal & a_{n+1}\hat{a} \\ \hline 0 & 0 \end{array}\right), \end{aligned}$$

which indeed coincides with (3).

The definition for central projections can be generalized further by decreasing the dimension of the image space which at the same time increasing the dimension of the center.

Let $L, C \subset \mathbb{RP}^n$ be projective subspaces with

$$C \cap L = \varnothing, \quad C \vee L = \mathbb{RP}^n$$

Then the map

$$\pi : \mathbb{RP}^n \backslash C \to L, \quad X \mapsto (C \vee X) \cap L$$

is called *(generalized) central projection* onto L with center C. Indeed, this map is well-defined, since $\dim(C \vee X) = \dim C + 1$ and $\dim L + \dim C = n - 1$ and therefore, $C \vee X$ and L intersect in exactly one point.

Again, the map π becomes invertible and in particular a projective transpormation,

$$\pi : K \to L, \quad X \mapsto (C \vee X) \cap L$$

once restricted to any subspace $K \subset \mathbb{RP}^n$ with

$$\dim K = \dim L, \quad C \cap K = \varnothing$$

Example 3.4 (central projection)**.** If L is a hyperplanes, i.e. $\dim L = n - 1$, the center C is a point, and the generalized central projection becomes the standard central projection.

Example 3.5 (three skew lines)**.** If $n = 3$ and K, L are two non-intersecting lines, then the center C is another line, and we obtain three skew lines.

4 Conics and quadrics

While projective subspaces are described by linear homogeneous equations, we now add the objects that are described by quadratic homogeneous equations.

Conics or *conic sections* are planar sections of a cone of revolution (or a cylinder)

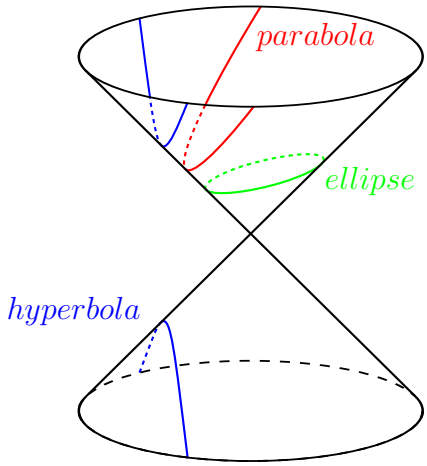


Figure 5. Ellipse, hyperbola, and parabola as a planar section of a cone.

It can be shown that conic sections correspond exactly to the sets of solutions of quadratic equations

$$\left\{(x,y) \in \mathbb{R}^2 \;\middle|\; q_{11}x^2 + 2q_{12}xy + q_{22}y^2 + 2q_{13}x + 2q_{23}y + q_{33} = 0.\right\}$$

Introducing homogeneous coordinates $x = \frac{x_1}{x_3}$, $y = \frac{x_2}{x_3}$, the (non-homogeneous) quadratic equation in 2 variables can be written as a homogeneous quadratic equation in 3 variables

$$q_{11}x_1^2 + 2q_{12}x_1x_2 + q_{22}x_2^2 + 2q_{13}x_1x_3 + 2q_{23}x_2x_3 + q_{33}x_3^2 = 0,$$

or equivalently,

$$b(x,x) := \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \underbrace{\begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{pmatrix}}_{=:Q} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

where Q is a symmetrice matrix, i.e. $Q^\intercal = Q$, and b is a symmetric bilinear form on $\mathbb{R}^3$

$$b : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}.$$

Example 4.1. An *ellipse* is a conic section. In normal form in $\mathbb{R}^2$ (up to a Euclidean transformation) it is given by

$$\left\{(x,y) \in \mathbb{R}^2 \;\middle|\; \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1\right\}.$$

Introducing homogeneous coordinates $x = \frac{x_1}{x_3}$, $y = \frac{x_2}{x_3}$, we can write its equation as a homogeneous quadratic equation

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - x_3^2 = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} \frac{1}{a^2} & & \\ & \frac{1}{b^2} & \\ & & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0.$$

4.1 Bilinear forms

Let V be a vector space over $\mathbb{R}$ of dimension $n+1$.

A *bilinear form* on V is a map

$$b : V \times V \to \mathbb{R}$$

which is linear in both arguments.

Let $e_1, \dots, e_{n+1}$ be a basis of V. Then the matrix $Q = (q_{ij}) \in \mathbb{R}^{(n+1)\times(n+1)}$

$$q_{ij} := b(e_i, e_j) \qquad \text{for } i, j = 1, \dots, n+1$$

is called the *representative matrix*, or *Gram matrix*, of the bilinear form b.

For two coordinate vectors $x = \sum_i x_i e_i, y = \sum_i y_i e_i \in V$ we have

$$b(x, y) = x^\intercal Q y.$$

A change of coordinates $\tilde{x} = Ax$ with $A \in \mathrm{GL}(n+1)$ acts on the representative matrix as

$$\tilde{Q} = A^{-\intercal} Q A^{-1}.$$

Symmetric bilinear forms and quadratic forms

A bilinear form is called *symmetric* if

$$b(x, y) = b(y, x) \qquad \text{for } x, y \in V,$$

or equivalenty, if its representative matrix is symmetric

$$Q^\intercal = Q.$$

The space of symmetric bilinear forms $\mathrm{Sym}(V)$ is a linear subspace of dimension

$$\dim \mathrm{Sym}(V) = \frac{(n+1)(n+2)}{2}.$$

A symmetric bilinear form $b(\cdot, \cdot)$ defines a corresponding *quadratic form* $b(\cdot)$

$$b(x) := b(x, x) \qquad \text{for } x \in V.$$

Vice versa, a quadratic form uniquely determines its bilinear form (polarization identity)

$$2b(x, y) = b(x + y) - b(x) - b(y),$$

and thus, the vector spaces of symmetric bilinear forms on V and quadratic forms on V are isomorphic.

4.2 Quadrics

The zero set of a non-zero quadratic form defines a *quadric* in $\mathrm{P}(V)$

$$\mathcal{Q}_b := \{[x] \in \mathrm{P}(V) \mid b(x) = 0\}.$$

Example 4.2. The quadratic form

$$b(x) = x_1^2 + x_2^2 - x_3^2$$

defines a quadric (conic) in $\mathbb{R}\mathrm{P}^2$

$$\left\{[x] \in \mathbb{R}\mathrm{P}^2 \;\middle|\; b(x) = x_1^2 + x_2^2 - x_3^2 = 0\right\}$$

In affine coordinates $x_3 = 1$ this is a circle

$$x_1^2 + x_2^2 = 1.$$

A non-zero scalar multiple of b defines the same quadric:

$$\mathcal{Q}_b = \mathcal{Q}_{\lambda b} \qquad \text{for } \lambda \neq 0.$$

Remark 4.1. For some very degenerate images, e.g. if $\mathcal{Q}_b$ is empty, the reverse statement is not true over $\mathbb{R}$. However, if we either exclude these cases, or consider the complexification of real quadrics, it holds that

$$\mathcal{Q}_b^{\mathbb{C}} = \mathcal{Q}_{\tilde{b}}^{\mathbb{C}} \qquad \Leftrightarrow \qquad b = \lambda \tilde{b} \quad \text{for some } \lambda \neq 0.$$

Example 4.3. The quadratic forms

$$b(x) = x_1^2 + x_2^2 + x_3^2, \qquad \tilde{b}(x) = x_1^2 + 4x_2^2 + x_3^2,$$

both define empty conics in $\mathbb{R}\mathrm{P}^2$

$$\mathcal{Q}_b = \mathcal{Q}_{\tilde{b}} = \varnothing$$

even though $b \neq \lambda \tilde{b}$ for all $\lambda \neq 0$. However, the point $[1, i, 0]$ is contained in $\mathcal{Q}_b^{\mathbb{C}}$, but not in $\mathcal{Q}_{\tilde{b}}^{\mathbb{C}}$. Thus,

$$\mathcal{Q}_b^{\mathbb{C}} \neq \mathcal{Q}_{\tilde{b}}^{\mathbb{C}}.$$

Thus, we can identify the space of quadrics with the projective space $\mathrm{P}\,\mathrm{Sym}(V)$. Its dimension is given by

$$\dim \mathrm{P}\,\mathrm{Sym}(V) = \dim \mathrm{Sym}(V) - 1 = \frac{(n+1)(n+2)}{2} - 1 = \frac{n(n+3)}{2}.$$

and the coefficients

$$q_{ij} = b(e_i, e_j), \qquad \text{for } j \leqslant i$$

can be taken as homogeneous coordinates on the space of quadrics.

4.3 Projective classification of quadrics in $\mathbb{R}\mathrm{P}^n$

Two quadrics $\mathcal{Q}, \tilde{\mathcal{Q}} \subset \mathbb{R}\mathrm{P}^n$ are called projectively equivalent if there exists a projective transformation $f : \mathbb{R}\mathrm{P}^n \to \mathbb{R}\mathrm{P}^n$ such that

$$f(\mathcal{Q}) = \tilde{\mathcal{Q}}$$

or equivalently, if there exists $F \in \mathrm{GL}(n+1)$ and $\lambda \in \mathbb{R}$, $\lambda \neq 0$, such that

$$\tilde{Q} = \lambda F^{\intercal} Q F,$$

where Q and $\tilde{Q}$ are representative matrices for $\mathcal{Q}$ and $\tilde{\mathcal{Q}}$, respectively. Note, that $f = [F^{-1}]$.

By *Sylvester's law of inertia*, there exists an $F \in \mathrm{O}(n+1)$ such that

$$\tilde{Q} = F^{\intercal} Q F = \mathrm{diag}(\lambda_1, \ldots, \lambda_r, \mu_1, \ldots, \mu_s, \underbrace{0, \ldots, 0}_{t})$$

where,

$$\lambda_i > 0, \quad \mu_i < 0, \quad r + s + t = n.$$

Thus, after applying this transformation the equation for the quadric is of the form

$$\lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + \mu_1 x_{r+1}^2 + \ldots \mu_s x_{r+s}^2 = 0$$

By applying a second transformation

$$F = \mathrm{diag}(\tfrac{1}{\sqrt{\lambda_1}}, \ldots, \tfrac{1}{\sqrt{\lambda_r}}, \tfrac{1}{\sqrt{-\mu_1}}, \ldots, \tfrac{1}{\sqrt{-\mu_s}}, \underbrace{1, \ldots, 1}_{t})$$

we obtain

$$\tilde{Q} = \mathrm{diag}(\underbrace{1, \ldots, 1}_{r}, \underbrace{-1, \ldots, -1}_{s}, \underbrace{0, \ldots, 0}_{t}),$$

or as an equation for the quadric

$$x_1^2 + \ldots + x_r^2 - x_{r+1}^2 + \ldots - x_{r+s}^2 = 0.$$

The tuple (r, s, t), also written as

$$(\underbrace{+ \cdots +}_{r} \underbrace{- \cdots -}_{s} \underbrace{0 \cdots 0}_{t}),$$

is called the *signature* of the quadric. We define the signature up to the following equivalence

$$(r, s, t) \quad \sim \quad (s, r, t),$$

and obtain the following classification result.

Theorem 4.1. *Two quadrics in $\mathbb{R}\mathrm{P}^n$ are projectively equivalent if and only if they have the same signature.*

Quadrics in $\mathbb{R}\mathrm{P}^1$

- $(++)$ *empty quadric.* By complexification these are two complex conjugate points.
- $(+-)$ *two points.*
- $(+0)$ *one (double) point.*

Quadrics in $\mathbb{R}\mathrm{P}^2$ (conics)

- $(+\,+\,+)$ *empty conic.* By complexification this is an imaginary conic.
- $(+\,+\,-)$ *oval conic.* Its normal form is given by

$$x_1^2 + x_2^2 - x_3^2 = 0$$

In affine coordinates this conic is an ellipse, a hyperbola, or a parabola. Indeed, if we choose $x_3 = 1$, the equation becomes the equation for a circle

$$x_1^2 + x_2^2 = 1.$$

If we choose coordinates $y_1 = x_1, y_2 = x_3, y_3 = x_2$ and $y_3 = 1$, the equation becomes the equation for a hyperbola

$$y_1^2 - y_2^2 = 1$$

If we choose coordinates $y_1 = x_1, y_2 = x_2 + x_3, y_3 = x_3 - x_2$ and $y_3 = 1$, the equation becomes the equation for a parabola

$$y_1^2 = y_2.$$

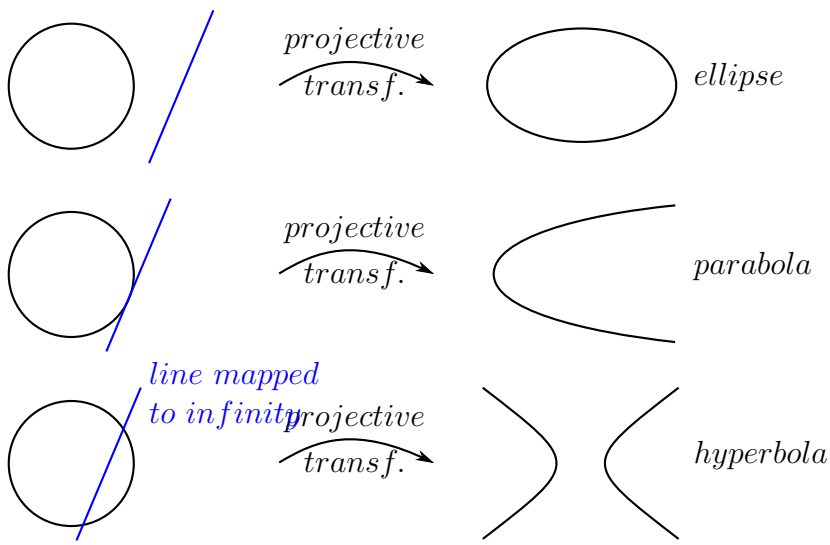


Figure 6. Projective transformations mapping a circle onto an ellipse, a parabola, or a hyperbola.

- $(+\,+\,0)$ *point.* By complexification these are two imaginary lines that intersect in a real point.
- $(+\,-\,0)$ *pair of lines.*
- $(+\,0\,0)$ *one (double) line.*

Quadrics in $\mathbb{R}\mathrm{P}^3$

non-degenerate quadrics:

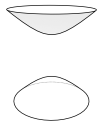
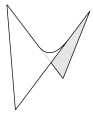
affine type	**affine signature** **affine normal form**	**picture**	**signature** **projective normal form**
ellipsoid	$(+++-)_-$ $x^2+y^2+z^2=1$		$(+++-)$ $x_1^2+x_2^2+x_3^2-x_4^2=0$
2-sheeted hyperboloid	$(+++-)_+$ $x^2+y^2-z^2=-1$		
elliptic paraboloid	$(+++-)_\mathrm{p}$ $z=x^2+y^2$		
1-sheeted hyperboloid	$(++--)_-$ $x^2+y^2-z^2=1$		$(++--)$ $x_1^2+x_2^2-x_3^2-x_4^2=0$
hyperbolic paraboloid	$(++--)_\mathrm{p}$ $z=x^2-y^2$		
empty (imaginary)	$(+++++)_+$ $x^2+y^2+z^2=-1$		$(+++++)$ $x_1^2+x_2^2+x_3^2+x_4^2=0$

Table 2. Affine types of non-degenerate quadrics in $\mathbb{R}^3$ and the corresponding projective types in $\mathbb{R}\mathrm{P}^3$.

degenerate quadrics:

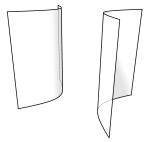
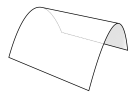
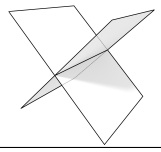
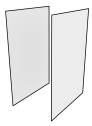
affine type	affine signature affine normal form	picture	signature projective normal form
cone	$(++-0)_0$ $x^2+y^2-z^2=0$		$(++-0)$ $x_1^2+x_2^2-x_3^2=0$
elliptic cylinder	$(++-0)_-$ $x^2+y^2=1$		
hyperbolic cylinder	$(++-0)_+$ $x^2-y^2=1$		
parabolic cylinder	$(++-0)_\mathrm{p}$ $z=x^2$		
one point (imaginary cone)	$(+++0)_0$ $x^2+y^2+z^2=0$	.	$(+++0)$ $x_1^2+x_2^2+x_3^2=0$
empty (imaginary cylinder)	$(+++0)_+$ $x^2+y^2=-1$		
two intersecting planes	$(+-00)_0$ $x^2-z^2=0$		$(+-00)$ $x_1^2-x_2^2=0$
two parallel planes	$(+-00)_-$ $x^2=1$		
one plane (and one at infinity)	$(+-00)_\mathrm{p}$ $x=0$		
one line (two intersecting imaginary planes)	$(++00)_0$ $x^2+z^2=0$		$(++00)$ $x_1^2+x_2^2=0$
empty (two parallel imaginary planes)	$(++00)_+$ $x^2=-1$		
one "double" plane	$(+000)_0$ $x^2=0$		$(+000)$ $x_1^2=0$
empty (one "double" plane at infinity)	$(+000)_+$ $1=0$		

Table 3. Affine types of degenerate quadrics in $\mathbb{R}^3$ and the corresponding projective types in $\mathbb{R}\mathrm{P}^3$.

4.4 Affine classification of quadrics in $\mathbb{R}^n \subset \mathbb{R}\mathrm{P}^n$

Two quadrics $\mathcal{Q}, \tilde{\mathcal{Q}} \subset \mathbb{R}\mathrm{P}^n$ are called affine equivalent if there exists an affine transformation $f : \mathbb{R}\mathrm{P}^n \to \mathbb{R}\mathrm{P}^n$ such that

$$f(\mathcal{Q}) = \tilde{\mathcal{Q}}$$

or equivalently, if there exists $F \in \mathrm{GL}(n+1)$ with

$$F = \left(\begin{array}{c|c} A & b \\ \hline 0 & 1 \end{array}\right), \qquad A \in \mathrm{GL}(n), b \in \mathbb{R}^n,$$

and a $\lambda \in \mathbb{R}$, $\lambda \neq 0$, such that

$$\tilde{Q} = \lambda F^\intercal Q F.$$

With

$$Q = \left(\begin{array}{c|c} S & q \\ \hline q^\intercal & \sigma \end{array}\right), \qquad S \in \mathrm{Sym}(n), q \in \mathbb{R}^n, \sigma \in \mathbb{R}.$$

we obtain

$$F^\intercal Q F = \left(\begin{array}{c|c} A^\intercal S A & A^\intercal(Sb+q) \\ \hline (b^\intercal S + q^\intercal)A & b^\intercal S b + 2q^\intercal b + \sigma \end{array}\right),$$

Thus, in a first step, we can use A to bring S to the form

$$S = \mathrm{diag}(\underbrace{1, \ldots, 1, -1, \ldots, -1}_{k}, 0, \ldots, 0).$$

Case 1: There exists $b \in \mathbb{R}^n$ such that $Sb + q = 0$. Then Q can be brought to the form

$$Q = \left(\begin{array}{c|c} S & 0 \\ \hline 0 & \sigma \end{array}\right), \qquad S = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0), \sigma = 0, 1, -1.$$

Here $\sigma = 0, 1, -1$ can be achieved by rescaling Q and then using A to rescale S. If (r, s, t) is the projective signature of $\mathcal{Q}$, we write the *affine signature* in this case as

$$(r, s, t)_\sigma$$

with

$$(r, s, t)_\sigma \sim (s, r, t)_{-\sigma}$$

Case 2: There exists no $b \in \mathbb{R}^n$ such that $Sb + q = 0$. Then S must be singular, i.e., $k < n$. Now we apply the following steps:

- We choose $b \in \mathbb{R}^n$ such that the first k components of $Sb + q$ vanish.
- We choose A such that $A^\intercal(Sb + q) = e_n$ without changing S.
- We choose $b = -\frac{\sigma}{2} e_n$ to eliminate σ.

Thus, Q can be brought to the form

$$Q = \left(\begin{array}{c|cc} \hat{S} & \multicolumn{2}{c}{0} \\ \hline 0 & 0 & 1 \\ & 1 & 0 \end{array}\right), \qquad \hat{S} = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0)$$

If (r, s, t) is the projective signature of $\mathcal{Q}$, we write the *affine signature* in this case as

$$(r, s, t)_{\mathrm{p}}$$

with

$$(r, s, t)_{\mathrm{p}} \sim (s, r, t)_{\mathrm{p}}.$$

Note that the block $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ corresponds to a projective signature of $(+-)$. Thus, an affine signature $(r, s, t)_{\mathrm{p}}$ is only possible with $r > 0$ and $s > 0$.

Theorem 4.2. *Two quadrics in $\mathbb{R}\mathrm{P}^n$ are affine equivalent if and only if they have the same affine signature.*

4.5 Signature of subspaces

Let $\mathcal{Q} \subset \mathbb{R}\mathrm{P}^n$ be a quadric, and $K = \mathrm{P}(U) \subset \mathbb{R}\mathrm{P}^n$ a projective subspace. Then the *signature of K* (with respect to $\mathcal{Q}$) is the signature of $\mathcal{Q}$ restricted to K:

$$\{[x] \in K \mid b(x) = 0\}$$

Thus, it is determined by the restriction of the symmetrice bilinear form b to U.

Signature of a point A quadric $\mathcal{Q} \subset \mathbb{R}\mathrm{P}^n$ separates $\mathbb{R}\mathrm{P}^n$ into two connected components. For point $[x] \in \mathbb{R}\mathrm{P}^n$ the signature can take 3 possible values:

- `(+)` if $b(x) > 0$. The point lies on one side of $\mathcal{Q}$.
- `(-)` if $b(x) < 0$. The point lies on the other side of $\mathcal{Q}$.
- `(0)` if $b(x) = 0$. The point lies on $\mathcal{Q}$.

Signature of a line A line $\ell \subset \mathbb{R}\mathrm{P}^n$ can have the following possible signatures:

- `(++)` The line does not intersect $\mathcal{Q}$.
- `(+-)` The line intersects $\mathcal{Q}$ in two points.
- `(+0)` The line intersects $\mathcal{Q}$ in one point.
- `(00)` The line is contained in $\mathcal{Q}$.

If the line is given as the span of two points $\ell = [x] \vee [y]$, the representative matrix for b on the corresponding subspace is given by

$$Q = \begin{pmatrix} b(x,x) & b(x,y) \\ b(x,y) & b(y,y) \end{pmatrix}.$$

Note that its determinant

$$\det Q = b(x,x)b(y,y) - b(x,y)^2$$

is the product of its eigenvalues. Thus, if we exclude the case (00), which corresponds to $Q = 0$, the other three cases can be distinguished by the sign of the determinant. The line ℓ has signature

$$\begin{aligned} \texttt{(+-)} &\Leftrightarrow \det Q < 0, \\ \texttt{(++)} &\Leftrightarrow \det Q > 0, \\ \texttt{(+0)} &\Leftrightarrow \det Q = 0. \end{aligned}$$

4.6 Tangent lines and tangent cones

Let $\mathcal{Q} \subset \mathbb{R}\mathrm{P}^n$ be a quadric.

A *tangent line* of $\mathcal{Q}$ is a line that intersects $\mathcal{Q}$ in exactly one point. We have established that these are the lines of signature `(+0)`, and can be characterized in the following way.

Lemma 4.3. *A line $[x] \vee [y]$ not contained in $\mathcal{Q}$ is a tangent line of $\mathcal{Q}$, if and only if*

$$b(x,x)b(y,y) - b(x,y)^2 = 0.$$

Let $X = [x] \subset \mathbb{R}\mathrm{P}^n \backslash \mathcal{Q}$ a point not on $\mathcal{Q}$. Then the *tangent cone* to $\mathcal{Q}$ from P is defined as the union of all tangent lines to $\mathcal{Q}$ that contain the point P:

$$\mathcal{C}_X = \bigcup_{\substack{\ell \ni X, \\ \ell \text{ tangent of } \mathcal{Q}}} \ell = \left\{ [y] \in \mathbb{R}\mathrm{P}^n \;\middle|\; c(y) := b(x,x)b(y,y) - b(x,y)^2 = 0 \right\}.$$

Note that c defines a quadratic form, and thus $\mathcal{C}_X$ is a quadric itself.

By definition, every tangent line has a point on $\mathcal{Q}$, which we call the *point of tangency*. Thus, to obtain the tangent cone it is sufficient to join X with all points of tangency. By Lemma 4.3, for a point $[y] \in \mathcal{Q}$ on $\mathcal{Q}$, the line $[x] \vee [y]$ is a tangent line if and only if

$$b(x,y) = 0.$$

Thus, the points of tangency of all tangent lines through X lie in a hyperplane,

$$X^\perp = \{[y] \in \mathbb{R}\mathrm{P}^n \mid b(x,y) = 0\}$$

called the *polar hyperplane* of X (with respect to $\mathcal{Q}$). Thus, we can write the tangent cone in the following way

$$\mathcal{C}_X = \bigcup_{Y \in \mathcal{Q} \cap X^\perp} X \vee Y.$$

Example 4.4 (Shadow of an ellipsoid)**.**

What form does the shadow of an ellipsoid have?

Consider an ellipsoid $\mathcal{E} \subset \mathbb{R}^3 \subset \mathbb{R}\mathrm{P}^3$ (an affine image of a sphere). Let X be a point outside $\mathcal{E}$, and K a plane. The shadow of the ellipsoid cast onto K by a light source in X is bounded by the intersection with (one half of) the tangent cone $\mathcal{C}_X$. Thus it is a conic section.

Which type of conic section can we obtain? Can it be a hyperbola?

The type of conic section (ellipse, parabola, hyperbola) depends on how many points of intersection (0, 1, 2) it has with the line at infinity on K, or equivalently, how many generators of $\mathcal{C}_X$ intersect K in the line at infinity.

Generally, a line intersects the plane K in the line at infinity, if it is parallel to K. Thus, consider the plane K_X through X parallel to K. Then the number of generators of $\mathcal{C}_X$ in K_K is the number of intersection points of $\mathcal{C}_X \cap K$ with infinity.

Consider the two planes K_1, K_2 parallel to K touching $\mathcal{E}$ in one point. This separates $\mathbb{R}\mathrm{P}^n$ into two regions, one containing $\mathcal{E}$, and one not containing $\mathcal{E}$.

- If X is in the region not containing $\mathcal{E}$, then $\mathcal{C}_X \cap K$ is an ellipse.

- If X is in the region containing $\mathcal{E}$, then $\mathcal{C}_X \cap K$ is a hyperbola.
- If X lies in K_1 or K_2, then $\mathcal{C}_X \cap K$ is a parabola.

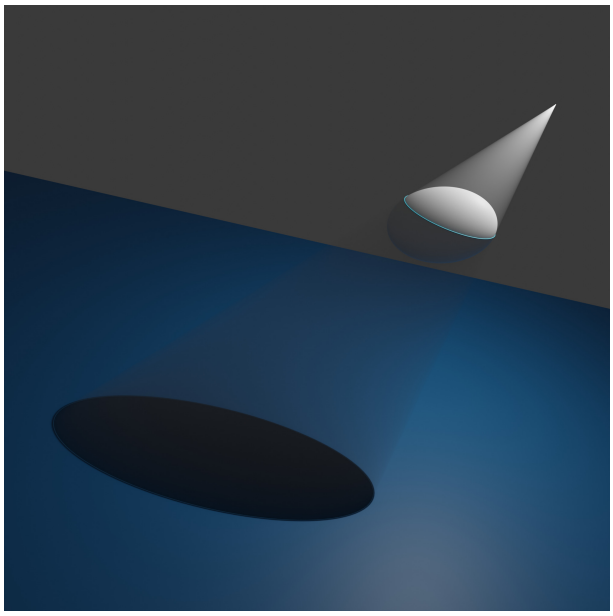


Figure 7. Shadow of an ellipsoid.

4.7 Polarity and tangent planes

Let $\mathcal{Q} \subset \mathbb{R}\mathrm{P}^n$ be a quadric of signature (r, s, t).

For a point $X = [x]$, its *polar hyperplane* (with respect to $\mathcal{Q}$) is given by

$$X^\perp = \{[y] \in \mathbb{R}\mathrm{P}^n \mid b(x, y) = 0\}.$$

If the point X has signature

- `(+)`, then $X^\perp$ has signature $(r-1, s, t)$.
- `(-)`, then $X^\perp$ has signature $(r, s-1, t)$.
- `(0)`, then $X^\perp$ has signature $(r-1, s-1, t+1)$.

For the cases `(+)` and `(-)`, we have established, that the intersection of $X^\perp$ with $\mathcal{Q}$ consists of all points common with the cone of contact $\mathcal{C}_X$.

In the case `(0)`, every point $Y \in X^\perp$ that does not lie on the quadric is a tangent line of $\mathcal{Q}$. Thus for a point $X \in \mathcal{Q}$ on the quadric, the polar hyperplane is the plane containing (and spanned by) all tangent lines though X, which we call the *tangent plane* of $\mathcal{Q}$ in the point X.

Example 4.5 (Tangent planes of a hyperboloid)**.** Consider a one-sheeted hyperboloid $\mathcal{H} \subset \mathbb{R}\mathrm{P}^3$, i.e. a quadric of signature `(++--)`. Then a tangent plane $X^\perp$ in any point $X \in \mathcal{H}$ has signature `(+-0)`. Thus, the restriction of $\mathcal{H}$ to $X^\perp$ consists of two lines.

In particular this means, that a one-sheeted hyperboloid, contains two lines through every point. In fact, it is a doubly ruled surface, and contains two families of lines, called its *generators*.

Example 4.6 (Projection of a generator)**.**

What is the shadow of a generator of a hyperboloid?

Consider a one-sheeted hyperboloid $\mathcal{H} \subset \mathbb{RP}^3$, a generator $\ell \subset \mathcal{H}$, and a center of projection X not on $\mathcal{H}$. We consider the projection to $X^\perp$.

The projection of $\mathcal{H}$ to $X^\perp$ is given by a conic section

$$\mathcal{D} := \mathcal{C}_X \cap X^\perp = \mathcal{H} \cap X^\perp$$

of signature `(++-)`. Its affine type can be determined in a similar way to Example 4.4.

Denote the central projection of ℓ to $X^\perp$ by $\tilde{\ell}$. The line ℓ intersects $X^\perp$ in some point $A \in \mathcal{D}$, which is fixed under the projection to $X^\perp$. Thus, $A \in \tilde{\ell}$.

Assume there exists another point $B \in \ell$ such that its projection $\tilde{B}$ lies on $\mathcal{D}$. Then the line $X \vee \tilde{B}$ is a tangent line of $\mathcal{H}$. On the other hand, this line intersects $\mathcal{H}$ in the two distinct points B and $\tilde{B}$, which is a contradiction. Thus, the projection ℓ only intersects $\mathcal{D}$ in A, and therefore is a tangent line of $\mathcal{D}$.

Note that projection to any other plane preserves this property.

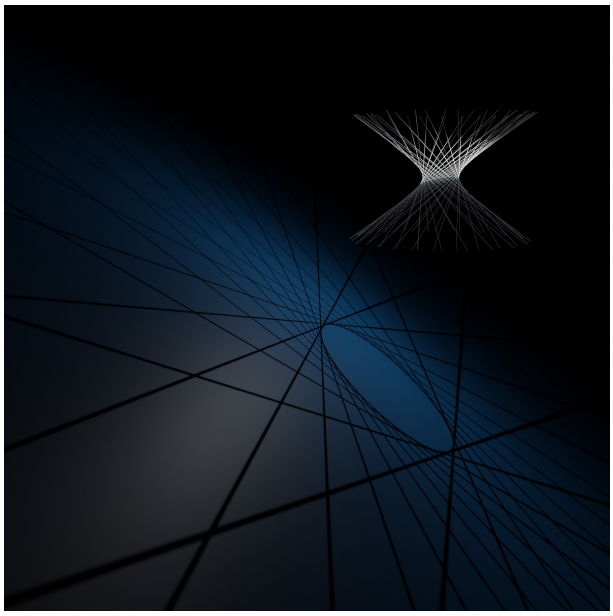


Figure 8. Shadow of the generators of a hyperboloid.

Differential geometric tangent plane Let us compare the notion of *tangent plane* that we have introduced for quadrics to the corresponding notion from Differential Geometry. In affine coordinates, we can view a quadric as a submanifold of $\mathbb{R}^n$ given as a level set of the function

$$0 = x^\intercal Q x = \begin{pmatrix} u^\intercal & 1 \end{pmatrix} \left(\begin{array}{c|c} S & q \\ \hline q^\intercal & \sigma \end{array}\right) \begin{pmatrix} u \\ 1 \end{pmatrix} = u^\intercal S u + 2q^\intercal u + \sigma =: f(u)$$

Then the normal vector of the tangent plane at some point $u_0 \in \mathbb{R}^n$ with $f(u_0) = 0$ is given by the gradient

$$\nabla_u f(u_0) = 2Su_0 + 2q.$$

Thus, the tangent plane at $u_0 \in \mathbb{R}$ is given by

$$\{u \in \mathbb{R}^n \mid \langle Su_0 + q, u - u_0 \rangle = 0\}$$

With

$$\langle Su_0 + q, u - u_0\rangle = u_0{}^\intercal Su + q^\intercal u - u_0{}^\intercal Su_0 - q^\intercal u_0 = u_0{}^\intercal Su + q^\intercal u + q^\intercal u_0 + \sigma$$

this coincides with the polar plane at u_0 in affine coordinates.

5 Plane curves and envelopes of lines

5.1 Plane curves

Definition 5.1.

(i) A *(plane) curve* is a smooth map

$$\gamma : I \to \mathbb{R}^2$$

with some interval $I \subset \mathbb{R}$.

(ii) Let γ be a curve.

- The vectors
$$\dot\gamma(t)$$
are called the *velocity* or *tangent vectors* of γ.
- The function
$$v(t) := \|\dot\gamma(t)\|$$
is called the *speed* of γ.
- The function
$$s(t) := \int_{t_1}^{t} v(t)\mathrm{d}t$$
is called the *arc-length* of γ, here $I = [t_1, t_2]$.
- If $v(t) = 1$ for all $t \in I$, then γ is called *arc-length parametrized.*

(iii) A curve γ is called *regular* if

$$\dot\gamma(t) \neq 0 \quad \text{for all } t \in I.$$

(iv) Let γ be a regular curve and $t \in I$.

- Any non-zero scalar multiple of $\dot\gamma(t)$ is called a *tangent vector* at $t \in I$.
- The line
$$T(t) := \{\gamma(t) + \alpha\dot\gamma(t) \mid \alpha \in \mathbb{R}\}$$
is called the *tangent line* at $t \in I$.
- Any vector $n(t)$ orthogonal to $\dot\gamma(t)$, i.e.,
$$\langle n(t), \dot\gamma(t)\rangle = 0,$$
is called a *normal vector* at $t \in I$. In particular one can choose.
$$n(t) = \frac{1}{v(t)} J\dot\gamma(t), \qquad J = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right),$$
which is called the *unit normal vector* at $t \in I$.

- The line
$$\begin{aligned} N(t) &:= \left\{x \in \mathbb{R}^2 \;\middle|\; \langle \dot\gamma(t), x - \gamma(t)\rangle = 0\right\} \\ &= \{\gamma(t) + \alpha n(t) \mid \alpha \in \mathbb{R}\}. \end{aligned}$$
is called the *normal line* at $t \in I$.

Note that the derivative of the arc-length is the speed

$$\dot s(t) = v(t).$$

For a regular curve γ the arc-length $s(\cdot)$ is monotonically increasing, and thus invertible. We call its inverse function $t(\cdot) = s^{-1}(\cdot)$ and thus write

$$\gamma(s) = (\gamma \circ t)(s).$$

For the derivative w.r.t. arc-length we write

$$\gamma' = \frac{\mathrm{d}}{\mathrm{d}s}\gamma = \frac{\mathrm{d}t}{\mathrm{d}s}\frac{\mathrm{d}}{\mathrm{d}t}\gamma = \frac{1}{v}\dot\gamma.$$

In particular, the parametrization of γ w.r.t. arc-length has unit speed

$$\|\gamma'\| = 1,$$

which implies

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\|\gamma'\|^2 = \frac{\mathrm{d}}{\mathrm{d}s}\langle\gamma', \gamma'\rangle = 2\langle\gamma'', \gamma'\rangle.$$

Thus γ'' always points in normal direction.

Definition 5.2. Let γ be a regular curve, and let n be the unit normal vector field of γ. Then

$$\kappa(s) = \langle\gamma''(s), n(s)\rangle$$

is called the *(signed) curvature* of γ at s, i.e.

$$\gamma''(s) = \kappa(s)n(s).$$

In terms of an arbitrary parametrization, and with unit tangent vector

$$\tau(t) := \frac{\dot\gamma(t)}{v(t)}$$

the curvature can be written as

$$\begin{aligned} \kappa(t) &= \frac{1}{v(t)}\langle\dot\tau(t), n(t)\rangle = \frac{1}{v(t)^2}\langle\ddot\gamma(t), n(t)\rangle \\ &= \frac{1}{v(t)^3}\langle\ddot\gamma(t), J\dot\gamma(t)\rangle = \frac{1}{v(t)^3}\det(\dot\gamma(t), \ddot\gamma(t)). \end{aligned}$$

Example 5.1. Consider a parametrized circle of radius $r > 0$

$$\gamma(t) = r\begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}, \qquad t \in [0, 2\pi].$$

Then

$$\dot\gamma(t) = r\begin{pmatrix}-\sin(t)\\ \cos(t)\end{pmatrix}, \quad v(t) = \|\dot\gamma(t)\| = r, \quad \tau(t) = \frac{\dot\gamma(t)}{v(t)} = \begin{pmatrix}-\sin(t)\\ \cos(t)\end{pmatrix},$$
$$n(t) = J\tau(t) = \begin{pmatrix}-\cos(t)\\ -\sin(t)\end{pmatrix}, \quad \dot\tau(t) = \begin{pmatrix}-\cos(t)\\ -\sin(t)\end{pmatrix}.$$

Thus, the curvature of γ is

$$\kappa(t) = \frac{1}{v(t)}\langle\dot\tau(t), n(t)\rangle = \frac{1}{r}.$$

Definition 5.3. Let $\gamma : I \to \mathbb{R}^2$ be a regular curve, and n its unit normal vector field. If $\kappa(t) \neq 0$, then the *osculating circle* at $t \in I$ is the circle with center

$$c(t) = \gamma(t) + \frac{1}{\kappa(t)}n(t)$$

and radius

$$r(t) = \frac{1}{|\kappa(t)|}.$$

If $\kappa(t) = 0$, then we consider the tangent line at $t \in I$ to be the osculating circle.

The osculating circle touches its curve the corresponding point. Furthermore, if parametrized in the same direction as the curve, it has the same (signed) curvature.

It can also be shown that it is the best approximating circle in the following sense. Consider the circle through three points of the curve $\gamma(t)$, $\gamma(t-\epsilon)$, and $\gamma(t+\epsilon)$. Then in the limit $\epsilon \to 0$, this circle converges to the osculating circle.

5.2 Discrete plane curves

Definition 5.4.

(i) A *discrete (plane) curve* is a map

$$\gamma : I \to \mathbb{R}^2$$

with some interval $I \subset \mathbb{Z}$. We denote its *vertices* by

$$\gamma_k = \gamma(k) \qquad \text{for } k \in I.$$

(ii) Let γ be a discrete curve.

- The vectors
$$\Delta\gamma_k := \gamma_{k+1} - \gamma_k$$
are called *discrete velocity vectors*, *vertex difference vectors*, or *edge tangent vectors*. They are naturally defined on edges $(k, k+1)$.
- We define the *turning angle* at a vertex $k \in I$ by
$$\varphi_k := \sphericalangle(\Delta\gamma_k, \Delta\gamma_{k-1}) \in [-\pi, \pi].$$

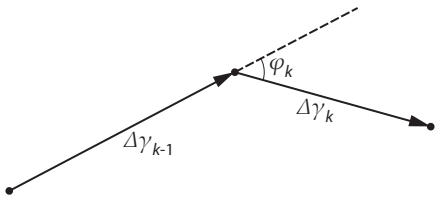


Figure 9. Turning angle at a vertex of a discrete curve.

- If
$$\|\Delta\gamma_k\| = \|\gamma_{k+1} - \gamma_k\| = 1$$
then γ is called *discrete arc-length parametrized curve.*

(iii) A discrete curve γ is called *regular* if any three successive points $\gamma_{k-1}, \gamma_k, \gamma_{k+1}$ are distinct, or equivalently, if any two successive edge tangent vectors are not anti-parallel.

(iv) Let γ be a discrete curve, $k \in I$.

- The line
$$T_k := \gamma_k \vee \gamma_{k+1}$$
is called the *edge tangent line* at the edge $(k, k+1)$.
- The perpendicular bisector N_k of γ_k and γ_{k+1} is called the *edge normal line* at the edge $(k, k+1)$.

We now introduce two types of discrete osculating circles.

Definition 5.5. Let $\gamma : I \to \mathbb{R}^2$ be a regular discrete curve. Then the circle C_k through three successive points $\gamma_{k-1}, \gamma_k, \gamma_{k+1}$ is called the *vertex osculating circle* at $k \in I$.

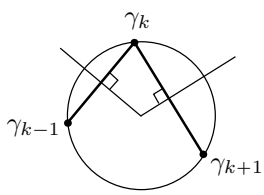


Figure 10. Vertex osculating circle.

- Note that the two involved edge normals N_{k-1} and N_k both contain the center of C_k.
- The discrete curvature at vertex k can now be defined by the radius of the vertex osculating circle. The radius is given by $\|\gamma_{k+1} - \gamma_{k-1}\| = 2R_k \sin\varphi_k$ which leads to the curvature
$$\kappa_k = \frac{2\sin\varphi_k}{\|\gamma_{k+1} - \gamma_{k-1}\|}.$$
- The vertex osculating circle inherits an orientation from the order of the three points on it. This can be used to also associate a sign to the discrete curvature, which corresponds to the sign in the formula above.

- The vertex osculating circle can also be used to define *vertex tangent lines* as the line tangent to C_k in the point γ_k.

Definition 5.6. Let $\gamma : I \to \mathbb{R}^2$ be a regular discrete curve. Then the circle C_k that touches three consecutive edge tangent lines T_{k-1}, T_k, T_{k+1} is called the *edge osculating circle* at $(k, k+1) \in I$.

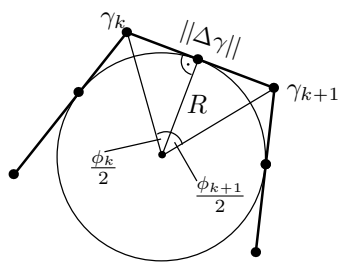


Figure 11. Edge osculating circle.

- For three (non-concurrent) lines in $\mathbb{R}^3$ there are four circles touching them. By endowing the tangent lines with the orientation coming from the order of the points of the curve on them, this choice can be made unique.

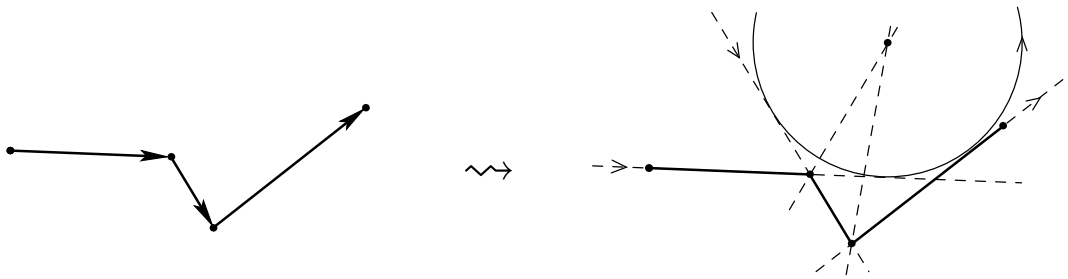


Figure 12. Edge osculating circle from oriented tangent lines.

- Note that the (correctly chosen) angle bisectors of successive edge tangent lines contain the center of the edge osculating circle. Thus, the edge osculating circle can be used to define *edge normal lines*.

- The (oriented) edge osculating circle can be used to define a (signed) discrete curvature at the edge $(k, k+1)$. The radius is given by $\|\Delta\gamma_k\| = R_k(\tan\frac{\varphi_k}{2} + \tan\frac{\varphi_{k+1}}{2})$. This leads to the curvature
$$\kappa_k = \frac{\tan\frac{\varphi_k}{2} + \tan\frac{\varphi_{k+1}}{2}}{\|\Delta\gamma_k\|}.$$

Computing angle bisectors Consider two oriented lines
$$\ell = \left\{x \in \mathbb{R}^2 \;\middle|\; \langle n, x\rangle + h = 0\right\}, \qquad \tilde{\ell} = \left\{x \in \mathbb{R}^2 \;\middle|\; \langle \tilde{n}, x\rangle + \tilde{h} = 0\right\}$$
with $n, \tilde{n} \in \mathbb{S}^1$, $h, \tilde{h} \in \mathbb{R}$, and orientation coming from the normal vectors $n, \tilde{n}$.

Then the two *angle bisectors* of ℓ and $\tilde{\ell}$ are given by
$$m_+ = \left\langle x \in \mathbb{R}^2, \langle n + \tilde{n}, x\rangle + h + \tilde{h} = 0\right\rangle,$$
$$m_- = \left\langle x \in \mathbb{R}^2, \langle n - \tilde{n}, x\rangle + h - \tilde{h} = 0\right\rangle.$$

Reflection in m_- maps ℓ to $\tilde{\ell}$, but with opposite orientation, while reflection in m_+ maps ℓ to $\tilde{\ell}$ with the same orientation.

Thus, for two adjacent edge tangent lines T_k, T_{k+1} the orientation reversing angle bisector m_- is the desired *vertex normal line*.

5.3 Envelopes

Consider a one-parameter family of curves C (implicitly) given by

$$C(t) = \left\{x \in \mathbb{R}^2 \;\middle|\; F(t,x) = 0\right\}, \qquad t \in I$$

with some smooth map $F : I \times \mathbb{R}^2 \to \mathbb{R}$.

Definition 5.7. A curve $\gamma : I \to \mathbb{R}^2$ is called *envelope* of the one-parameter family C if γ is tangent to $C(t)$ in the point $\gamma(t)$, i.e.

$$F(t,\gamma(t)) = 0 \qquad (\gamma(t) \text{ lies on } C(t)) \tag{4}$$

$$\langle \nabla_x F(t,\gamma(t)), \dot{\gamma}(t)\rangle = 0 \qquad (\gamma \text{ in tangent direction of } C(t) \text{ at } \gamma(t)) \tag{5}$$

This is a differential equation for γ. But we can reformulate this in the following way. Equation (4) implies

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t} F(t,\gamma(t)) = \mathrm{D}F(t,\gamma(t)) \begin{pmatrix} 1 \\ \dot{\gamma}(t) \end{pmatrix} = \begin{pmatrix} \partial_t F & \partial_{x_1} F & \partial_{x_2} F \end{pmatrix} \begin{pmatrix} 1 \\ \dot{\gamma}(t) \end{pmatrix} \\ &= \partial_t F(t,\gamma(t)) + \langle \nabla_x F(t,\gamma(t)), \dot{\gamma}(t)\rangle . \end{aligned}$$

Thus, equations (4) and (5) are equivalent to

$$\begin{aligned} F(t,\gamma(t)) &= 0 \\ \partial_t F(t,\gamma(t)) &= 0, \end{aligned}$$

which is not a differential equation in γ anymore.

In particular, if C is a family of lines, then the equations for the envelope are two linear equations in γ.

Example 5.2. For a regular curve $\gamma : I \to \mathbb{R}^2$ the envelope of its tangent lines is the curve γ itself,

Example 5.3. Consider

$$F(t,x) = x_1 - 2tx_2 + t^2.$$

Then

$$\partial_t F(t,x) = -2x_2 + 2t.$$

implies $x_2 = t$. Substituting this into $F(t,x) = 0$ we obtain $x_1 = t^2$. Thus the envelope is given by

$$\gamma(t) = \begin{pmatrix} t^2 \\ t \end{pmatrix}$$

which is a parabola.

Note, that, in homogeneous coordinates, the equation for the lines is given by

$$x_1 - 2tx_2 + t^2 x_3 = \begin{pmatrix} 1 & -2t & t^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0$$

which describes a curve $t \mapsto [1, -2t, t^2]$ in $(\mathbb{R}\mathrm{P}^2)^*$. This curve is implicitly given by

$$x_2^2 - 4x_1x_3 = 0,$$

which is a conic in $(\mathbb{R}\mathrm{P}^2)^*$. This is an example of the general fact, that (the envelope) of the dual of a conic is a conic.

Discrete envelope of a family of lines Let $C : \mathbb{Z} \supset I \to \mathrm{Lines}(\mathbb{R}^2)$ be a discrete one-parameter family of lines, such that no adjacent lines are equal or parallel.

Then we can define the *discrete envelope* as the discrete curve given by intersections of adjacent lines

$$\gamma_k := C_k \cap C_{k+1}.$$

In this way the edge tangent lines of γ_k coincide with the lines of C,

$$T_k = C_{k+1}.$$

5.4 Evolute

Definition 5.8. The *evolute* of a regular curve γ is the envelope of its normal lines N.

The envelope of the family of normal lines is described by the equations

$$\begin{aligned} F(t,x) &:= \langle \dot\gamma(t), x - \gamma(t)\rangle = 0 \\ \partial_t F(t,x) &= \langle \ddot\gamma, x - \gamma(t)\rangle - \|\dot\gamma(t)\|^2 = 0 \end{aligned}$$

With unit normal field n of γ, the first equation is equivalent to

$$x = e(t) = \gamma(t) + \alpha(t)n(t)$$

with some function α. Then, $\alpha(t)$ can be determined by the second equation

$$\langle \ddot\gamma, e(t) - \gamma(t)\rangle - \|\dot\gamma(t)\|^2 = \alpha(t)\langle \ddot\gamma(t), n(t)\rangle - \|\dot\gamma(t)\|^2 = 0$$

to be

$$\alpha(t) = \frac{\|\dot\gamma\|^2}{\langle \ddot\gamma(t), n(t)\rangle} = \frac{1}{\kappa(t)},$$

which is well-defined as long as $\langle \ddot\gamma(t), n(t)\rangle \neq 0$, i.e., $\kappa(t) \neq 0$. Thus, the evolute of γ is given by

$$e(t) = \gamma(t) + \frac{1}{\kappa(t)} n(t),$$

and we find

Proposition 5.1. *The evolute of a regular curve consists of the centers of its osculating circles.*

Proposition 5.2. *Let $\gamma : I \to \mathbb{R}^2$ be a regular curve. Then its evolute e is non-regular in $t \in I$ if and only if the curvature κ of γ has a local extremum in $t \in I$, i.e.,*

$$\dot e(t) = 0 \quad \Leftrightarrow \quad \dot\kappa(t) = 0$$

Proof. Let γ be arc-length parametrized. Then

$$e'(s) = \gamma'(s) + \left(\frac{1}{\kappa(s)}\right)' n(s) + \frac{1}{\kappa(s)} n'(s).$$

For the normal vector we have $0 = \frac{\mathrm{d}}{\mathrm{d}s}\langle n(s), n(s)\rangle = 2\langle n'(s), n(s)\rangle$, thus $n'(s) = \alpha(s)\gamma'(s)$ where

$$\alpha(s) = \langle n'(s), \gamma'(s)\rangle = -\langle n(s), \gamma''(s)\rangle = -\kappa(s).$$

So,

$$n'(s) = -\kappa(s)\gamma'(s)$$

Thus,

$$e'(s) = \left(\frac{1}{\kappa(s)}\right)' n(s) = -\frac{\kappa'(s)}{\kappa(s)^2} n(s)$$

□

Definition 5.9. A *parallel curve* of γ is a curve of the form

$$\gamma_r(t) := \gamma(t) + rn(t), \qquad r \in \mathbb{R}.$$

where n is the unit normal vector field of γ

Proposition 5.3. *Parallel curves have the same evolutes.*

Proof. We show that parallel curves have the same normal lines.

$$\langle \dot{\gamma}_r(t), n(t)\rangle = \langle \dot{\gamma}(t) + r\dot{n}(t), n(t)\rangle = 0.$$

□

Example 5.4. Consider a parabola

$$\gamma(t) := \begin{pmatrix} t \\ t^2 \end{pmatrix}$$

Then

$$\dot{\gamma}(t) = \begin{pmatrix} 1 \\ 2t \end{pmatrix}, \quad \ddot{\gamma}(t) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad n(t) = J\dot{\gamma}(t) \begin{pmatrix} -2t \\ 1 \end{pmatrix}$$

and

$$\langle \ddot{\gamma}(t), n(t)\rangle = 2, \quad \|\dot{\gamma}(t)\|^2 = 1 + 4t^2.$$

Therefore, the evolute is given by

$$e(t) = \gamma(t) + \frac{\|\dot{\gamma}\|^2}{\langle \ddot{\gamma}(t), n(t)\rangle} n(t) = \begin{pmatrix} -4t^3 \\ \frac{1}{2} + 3t^2 \end{pmatrix},$$

which is a semicubic parabola.

Note that it has a cusp at the point where the parabola has maximal curvature.

Discrete evolutes Let $\gamma : \mathbb{Z} \supset I \to \mathbb{R}^2$ be a regular discrete curve.

- We can define its *vertex evolute* as the discrete envelope of adjacent edge normal lines. The vertex evolute consists of the centers of the vertex osculating circles.
- Alternatively, we can define its *edge evolute* as the discrete envelope of adjacent vertex normal lines. The edge evolute consists of the centers of the edge osculating circles.

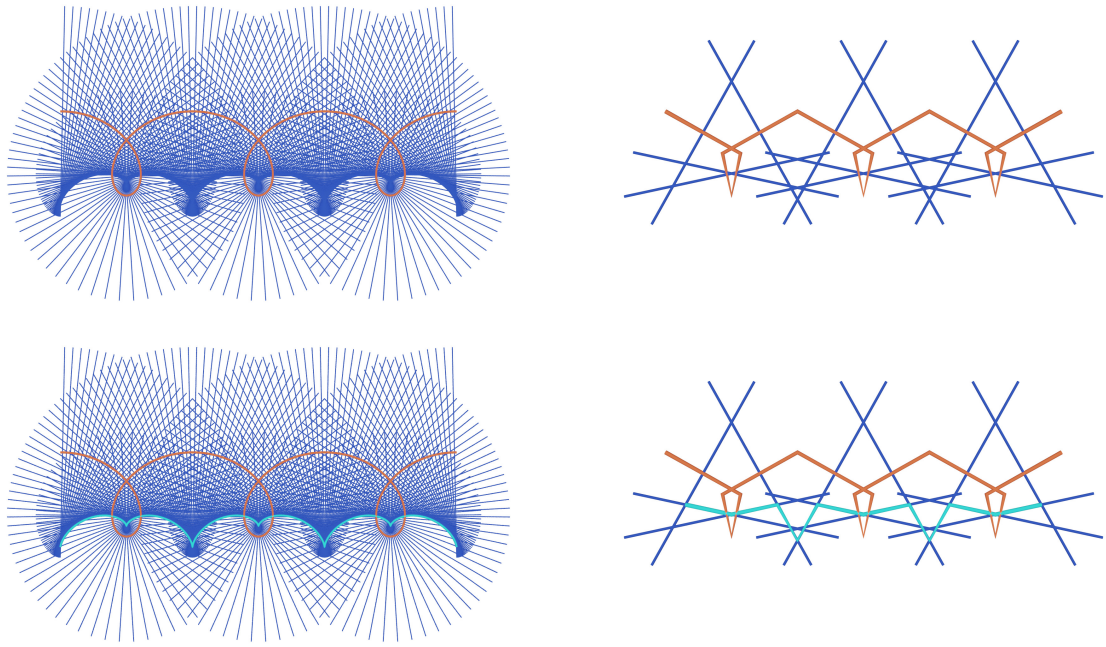


Figure 13. Top: Smooth and discrete curve and its tangent lines. Bottom: Smooth and discrete curve and its evolute.

5.5 Involute

Definition 5.10. An *involute* of a regular curve γ is a curve orthogonal to the tangent lines.

Thus, an involute $\Gamma : I \to \mathbb{R}^2$ must satisfy

$$\Gamma(t) = \gamma(t) + \alpha(t)\tau(t), \qquad \tau(t) = \frac{\dot\gamma(t)}{\|\dot\gamma(t)\|}$$

with some $\alpha : I \to \mathbb{R}$ and

$$0 = \left\langle \dot\Gamma(t), \dot\gamma(t) \right\rangle = \langle \dot\gamma(t), \dot\gamma(t) + \dot\alpha(t)\tau(t) + \alpha(t)\dot\tau(t) \rangle = \|\dot\gamma(t)\|^2 + \dot\alpha(t)\,\|\dot\gamma\|\,.$$

Thus,

$$\dot\alpha(t) = -\,\|\dot\gamma(t)\|$$

We obtain

$$\Gamma_a(t) = \gamma(t) - \frac{\dot\gamma(t)}{\|\dot\gamma(t)\|}\int_a^t \|\dot\gamma(t)\|\,\mathrm{d}t = \gamma(t) - \frac{\dot\gamma(t)}{\|\dot\gamma(t)\|}(s(t) - s(a)),$$

where s is the arc-length of γ.

Thus, in terms of arc-length parametrization the involute is given by

$$\Gamma_a(s) = \gamma(s) - \gamma'(s)(s-a).$$

The distance of the involute to the corresponding curve (along the tangent line) satisfies

$$\|\Gamma_a(s) - \gamma(s)\| = |s-a|\,.$$

- Thus, the involute is the locus of a point on a piece of taut string as the string is either unwrapped from or wrapped around the curve starting at the point $\gamma(a)$.
- Equivalently, it is the locus of the point on a straight line as it rolls without slipping along the curve.

Proposition 5.4. *Let γ be a regular curve.*

(i) The involute is regular at points where $\kappa(t) \neq 0$ and $t \neq a$.

(ii) The normal lines of the involute are the tangents of γ.

(iii) The evolute of the involute is γ.

(iv) The involutes are parallel curves.

Proof.

(i) $\Gamma_a'(s) = \gamma'(s) - \gamma''(s)(s-a) - \gamma'(s) = -(s-a)\kappa(s)n(s)$.

(ii) By definition of the involute $\langle \Gamma_a'(s), \gamma'(s)\rangle = 0$.

(iii) Follows from (ii).

(iv) $\Gamma_a(s) = \Gamma_0(s) + a\gamma'(s)$, where $\gamma'(s)$ is the unit normal at $\Gamma_0(s)$.

□

Remark 5.1. The one-parameter family of tangent lines of a curve together with its one-parameter family of involutes form an orthogonal coordinate system.

Example 5.5 (Involutes of a circle)**.** Consider a parametrized circle of radius $r > 0$

$$\gamma(t) = r\begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}, \qquad t \in [0, 2\pi].$$

Then

$$\dot\gamma(t) = r\begin{pmatrix}-\sin(t)\\ \cos(t)\end{pmatrix}, \quad v(t) = \|\dot\gamma(t)\| = r, \quad s(t) - s(a) = r(t-a).$$

Thus, the involutes of γ are given by

$$\Gamma_a(t) = r\begin{pmatrix}\cos(t) - (t-a)\sin(t)\\ \sin(t) + (t-a)\cos(t)\end{pmatrix}$$

This is a common shape for the teeth of gears, the so called "involute gears".

Example 5.6 (Involute of a semi-cubic)**.** Consider the semicubic parabola, we obtained as the evolute of a parabola. We reconstruct the parabola as one involute of semicubic parabola.

$$\gamma(t) = \begin{pmatrix} -4t^3 \\ \frac{1}{2} + 3t^2 \end{pmatrix}, \qquad t > 0.$$

Then

$$\dot{\gamma}(t) = \begin{pmatrix} -12t^2 \\ 6t \end{pmatrix}, \quad \|\dot{\gamma}(t)\| = 6t\sqrt{1+4t^2}, \quad \int_0^t \|\dot{\gamma}(t)\| \,\mathrm{d}t = \frac{1}{2}(1+4t^2)^{\frac{3}{2}} - \frac{1}{2}.$$

For simplicity, we add a constant of integration $\frac{1}{2}$ and obtain

$$\Gamma_0(t) = \begin{pmatrix} -4t^3 \\ \frac{1}{2} + 3t^2 \end{pmatrix} + \frac{1}{6t\sqrt{1+4t^2}} \begin{pmatrix} -12t^2 \\ 6t \end{pmatrix} \frac{1}{2}(1+4t^2)^{\frac{3}{2}} = \begin{pmatrix} t \\ t^2 \end{pmatrix},$$

which is a parabola.

Note that the other involutes of the semicubic parabola are not parabolas.

Discrete involutes We can derive constructions for discrete involutes from the property that evolute of the involute should be the original curve, i.e., the tangent lines of the original curve should be the normal lines of the evolute.

Let $\gamma : \mathbb{Z} \subset I \to$ be a regular discrete curve.

- Choose some starting point $\Gamma_0 \in \mathbb{R}^2$
- Obtain Γ_{k+1} from Γ_k by reflection in tangent line T_k of γ.

Then T_k is the edge normal line of Γ at the edge $(k, k+1)$.

Alternatively:

- Choose some starting edge tangent line $\tilde{T}_0 = \Gamma_0 \vee \Gamma_1$.
- Obtain $\tilde{T}_{k+1}$ from $\tilde{T}_k$ by reflection in tangent line T_k of γ, and thus, $\Gamma_{k+1} = \tilde{T}_k \cap T_{k+1}$.

Then T_k is the vertex normal line of Γ at the vertex $k+1$.

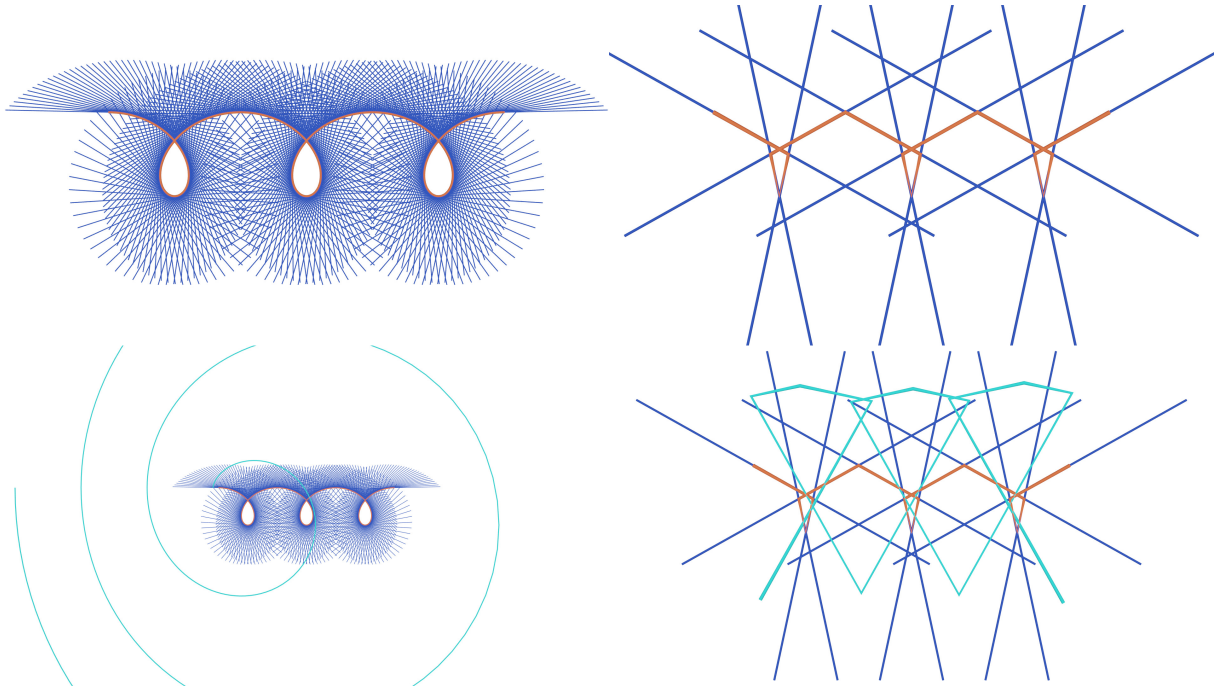


Figure 14. Top: Smooth and discrete curve and its normal lines. Bottom: Smooth and discrete curve and one of its involutes.

6 Möbius geometry

6.1 The elementary model of Möbius geometry

Consider the n-dimensional Euclidean space $\mathbb{R}^n$. The inversion in a hypersphere with center $c \in \mathbb{R}^n$ and radius $r > 0$ can be described in the following way: The point x and its image x' lie on the same ray emanating from c and the distances to c satisfy the relation

$$\| x - c \| \cdot \| x' - c \| \;=\; r^2.$$

This gives rise to an involution on $\mathbb{R}^n$, except that the center c has no image and no preimage. To fix this, we add one extra point to $\mathbb{R}^n$, called ∞, and obtain the *extended Euclidean space*

$$\widehat{\mathbb{R}^n} := \mathbb{R}^n \cup \{\infty\}.$$

Definition 6.1. The *(sphere) inversion* in the hypersphere with center $c \in \mathbb{R}^n$ and radius $r > 0$ is the map defined by

$$\begin{aligned} \widehat{\mathbb{R}^n} \to \widehat{\mathbb{R}^n}, \qquad & x \longmapsto x' = c + \frac{r^2}{\|x - c\|^2}\,(x - c) \qquad \text{for } x \neq c, \\ & c \longmapsto \infty \\ & \infty \longmapsto c \end{aligned}$$

Sphere inversions preserve angles and map hyperspheres and hyperplanes to hyperspheres and hyperplanes. This statement becomes simpler and more specific at the same time if we consider hyperplanes as special cases of hyperspheres through the point ∞. More precisely, let us adopt the following convention:

Definition 6.2. A *sphere in* $\widehat{\mathbb{R}^n}$ is either a sphere in $\mathbb{R}^n$ or the union of a plane in $\mathbb{R}^n$ with $\{\infty\}$.

Then we can simply say:

Theorem 6.1. *Sphere inversions preserve angles and map hyperspheres in* $\widehat{\mathbb{R}^n}$ *to hyperspheres in* $\widehat{\mathbb{R}^n}$.

Since circles and, more generally, k-dimensional spheres for $1 \leqslant k < n$ are intersections of $n-k$ hyperspheres, sphere inversions preserve spheres of any dimension:

Corollary 6.2. *Sphere inversions map k-spheres in* $\widehat{\mathbb{R}^n}$ *to k-spheres in* $\widehat{\mathbb{R}^n}$.

Just as hyperplanes are limiting cases of hyperspheres, reflections in hyperplanes are limiting cases of sphere inversions. The reflection in the hyperplane with equation $\langle x-a, v\rangle = 0$ is the map

$$x \longmapsto x' = x - 2\frac{\langle x-a, v\rangle}{\langle v, v\rangle}\, v,$$

which we extend from $\mathbb{R}^n$ to $\widehat{\mathbb{R}^n}$ by declaring that reflections in hyperplanes map ∞ to ∞.

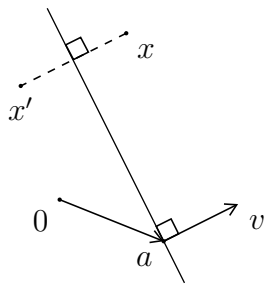


Figure 15. Reflection in a hyperplane

Definition 6.3. A *Möbius transformation* of $\mathbb{R}^n \cup \{\infty\}$ is a composition of sphere inversions and reflections in hyperplanes. The Möbius transformations form a group called the *Möbius group* and denoted by $\mathrm{Möb}(n)$.

Remark 6.1. A Möbius transformation is orientation reversing or preserving depending on whether it is the composition of an odd or even number of reflections. The subgroup of orientation preserving Möbius transformations is called the *special Möbius group* and denoted by $\mathrm{SMöb}(n)$.

Because reflections preserve angles and map spheres to spheres, Theorem 6.1 extends to Möbius transformations:

Theorem 6.3. *Möbius transformations preserve angles and map spheres in* $\widehat{\mathbb{R}^n}$ *to spheres in* $\widehat{\mathbb{R}^n}$.

Similarity transformations on $\mathbb{R}^n$ are the transformations of the form $x \mapsto \lambda Ax + b$ with $\lambda > 0$, $A \in O(n)$, and $b \in \mathbb{R}^n$. Reflections in hyperplanes are a special case, and like reflections in hyperplanes we extend all similarity transformations from $\mathbb{R}^n$ to $\widehat{\mathbb{R}^n}$ by declaring that ∞ maps to ∞.

Proposition 6.4. *The Möbius group contains all similarity transformations.*

Proof. The group of similarity transformations is generated by translations, orthogonal transformations, and scalings.

- A translation $x \mapsto x + v$ is the composition of two reflections in parallel hyperplanes.
- An orthogonal transformation $x \mapsto Ax$ with $A \in O(n)$ is the composition of at most n reflections in hyperplanes through the origin.
- A scaling transformation $x \mapsto \lambda x$ with $\lambda > 0$ is the composition of a reflection in the unit sphere followed by a reflection in a sphere with center 0 and radius $\sqrt{\lambda}$. □

Conversely, one only needs to add one sphere inversion to the group of similarity transformations to generate the Möbius group:

Proposition 6.5. *Every Möbius transformation is a composition of similarity transformations and inversions in the unit sphere.*

By Theorem 6.3, Möbius transformations map hyperspheres to hypersphers. This property already characterizes all Möbius transformations.

Theorem 6.6 (Fundamental theorem of Möbius geometry)**.** *Any bijective map $f : \widehat{\mathbb{R}^n} \to \widehat{\mathbb{R}^n}$ which maps hyperspheres in $\widehat{\mathbb{R}^n}$ to hyperspheres in $\widehat{\mathbb{R}^n}$ is a Möbius transformation.*

6.2 The projective model of Möbius geometry

The idea is to transfer Möbius geometry from $\widehat{\mathbb{R}^n}$ to the n-dimensional sphere S^n via stereographic projection

$$\sigma : S^n \to \widehat{\mathbb{R}^n}.$$

Via the standard embedding

$$\mathbb{R}^{n+1} \hookrightarrow \mathbb{R}\mathrm{P}^{n+1}, \qquad u \longmapsto \begin{bmatrix} u \\ 1 \end{bmatrix} = [x]$$

we identify the unit sphere $S^n \subset \mathbb{R}^{n+1}$ with a quadric in $\mathbb{R}\mathrm{P}^{n+1}$,

$$S^n = \left\{ [x] \in \mathbb{R}\mathrm{P}^{n+1} \;\middle|\; \langle x, x \rangle_{n+1,1} = 0 \right\}, \tag{6}$$

where

$$\langle x, \tilde{x} \rangle_{n+1,1} = x_1 \tilde{x}_1 + \ldots + x_{n+1} \tilde{x}_{n+1} - x_{n+2} \tilde{x}_{n+2}$$

is the Lorentz product. Thus, indeed, in affine coordinates $x_{n+2} = 1$, the quadric S^n is the unit sphere

$$x_1^2 + \ldots + x_{n+1}^2 = 1.$$

Upon identifying the affine part of $\widehat{\mathbb{R}^n}$ with the affine part of the hyperplane

$$E = \left\{ [x] \in \mathbb{R}\mathrm{P}^{n+1} \;\middle|\; x_{n+1} = 0 \right\}.$$

the stereographic projection coincides with the central projection of S^n to the hyperplane E from the "north pole" $[e_{n+1} + e_{n+2}]$. A point $[x] \in S^n \backslash [e_{n+1} + e_{n+2}]$ on the quadric is mapped to

$$([x] \vee [e_{n+1} + e_{n+2}]) \cap E = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 0 \\ x_{n+2} - x_{n+1} \end{bmatrix}.$$

Inversely, a point $[u, 0, 1] \in E$, $u \in \mathbb{R}^n$, is mapped to

$$\left(\begin{bmatrix} u \\ 0 \\ 1 \end{bmatrix} \vee [e_{n+1} + e_{n+2}] \right) \cap S^n = \begin{bmatrix} 2u \\ \|u\|^2 - 1 \\ \|u\|^2 + 1 \end{bmatrix}.$$

Thus, we obtain the following equations for stereographic projection and its inverse:

$$\begin{array}{llll}
\sigma : & S^n & \longrightarrow & \widehat{\mathbb{R}^n} \\
 & \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ x_{n+2} \end{bmatrix} & \longmapsto & \dfrac{1}{x_{n+2} - x_{n+1}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{if} \quad x_{n+1} \neq x_{n+2} \\
 & [x] & \longmapsto & \infty \qquad\qquad\qquad\qquad \text{if} \quad x_{n+1} = x_{n+2} \\
\\
\sigma^{-1} : & \widehat{\mathbb{R}^n} & \longrightarrow & S^n \\
 & u & \longmapsto & \begin{bmatrix} 2u \\ \|u\|^2 - 1 \\ \|u\|^2 + 1 \end{bmatrix} \qquad \text{if} \quad u \in \mathbb{R}^n \\
 & \infty & \longmapsto & [e_{n+1} + e_{n+2}]
\end{array}$$

The stereographic projection σ is the restriction to $S^n \subset \mathbb{R}^{n+1}$ of the inversion in the hypersphere with center e_{n+1} (the north pole of S^n) and radius $\sqrt{2}$ (so that it contains the equatorial sphere $\{x \in S^n \mid x_{n+1}\} = 0$). Thus, Theorem 6.1 implies, that the stereographic projection preserves angles and maps spheres to spheres.

It is convenient to additionally employ another basis for the projective model, one that contains the center of projection. We change the basis for our homogeneous coordinates from the standard basis $(e_1, \ldots, e_{n+2})$ of $\mathbb{R}^{n+2}$ to the new basis

$$B = (e_1, \ldots, e_n, e_\infty, e_0)$$

with

$$e_\infty \;=\; \tfrac{1}{2}\,(e_{n+1} + e_{n+2}), \qquad e_0 \;=\; \tfrac{1}{2}\,(-e_{n+1} + e_{n+2}).$$

The points $[e_0]$ and $[e_\infty]$ are the "south pole" and the "north pole" of S^n. We choose the subscripts 0 and ∞ for the new basis vectors because

$$\sigma^{-1}(0) \;=\; [e_0] \quad \text{and} \quad \sigma^{-1}(\infty) \;=\; [e_\infty].$$

Note that B is not an orthonormal basis of $\mathbb{R}^{n+1,1}$. Rather, we have

$$\langle e_0, e_0 \rangle_{n+1,1} \;=\; \langle e_\infty, e_\infty \rangle_{n+1,1} \;=\; 0, \quad \langle e_0, e_\infty \rangle \;=\; -\frac{1}{2}\,.$$

Now the change-of-basis-transformations between coordinate vectors with respect to the standard basis and the new basis B are

$$x \;=\; \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ x_{n+2} \end{pmatrix} \;=\; \sum_{k=1}^{n} x_k e_k \;+\; x_\infty e_\infty \;+\; x_0 e_0 \;=\; \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ \frac{1}{2}(x_\infty - x_0) \\ \frac{1}{2}(x_\infty + x_0) \end{pmatrix},$$

and inversely

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_\infty \\ x_0 \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} + x_{n+2} \\ -x_{n+1} + x_{n+2} \end{pmatrix} = \underbrace{\left(\begin{array}{c|cc} I_n & \multicolumn{2}{c}{0} \\ \hline 0 & 1 & 1 \\ & -1 & 1 \end{array}\right)}_{F} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ x_{n+2} \end{pmatrix}$$

We may interpret $[x] = [x_1, \ldots, x_n, x_{n+1}, x_{n+2}]$ and $[x_1, \ldots, x_n, x_\infty, x_0]$ as representing the same point in $\mathbb{R}\mathrm{P}^{n+1}$ with respect to different bases, or we may interpret $[x_1, \ldots, x_n, x_\infty, x_0]$ as the image of $[x]$ under the projective transformation $f : \mathbb{R}\mathrm{P}^{n+1} \to \mathbb{R}\mathrm{P}^{n+1}, [x] \mapsto [Fx]$.

In coordinates of the basis B, the Lorentz scalar product is

$$\langle x, \tilde{x} \rangle_{n+1,1} = x_1 \tilde{x}_1 + \ldots + x_n \tilde{x}_n - \frac{1}{2} x_\infty \tilde{x}_0 - \frac{1}{2} x_0 \tilde{x}_\infty .$$

In affine coordinates $x_0 = 1$, the "north pole" $[e_\infty]$ lies at infinity and the quadric S^n becomes a paraboloid

$$x_\infty = x_1^2 + \ldots + x_n^2.$$

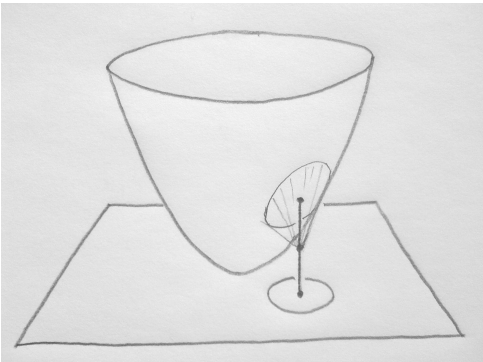


Figure 16. Stereographic projection becomes vertical projection in the paraboloid model

Correspondingly the stereographic projection (and its inverse) becomes vertical orthogonal projection from (and onto) this paraboloid (see Fig. 16).

$$\sigma^{-1}(u) = \begin{bmatrix} 2u \\ \|u\|^2 - 1 \\ \|u\|^2 + 1 \end{bmatrix} = [u + \|u\|^2 e_\infty + e_0].$$

Thus, the projective model in homogeneous coordinates with respect to B is sometimes called the *paraboloid model of Möbius geometry.*

6.3 Spheres in the projective model

Hyperspheres in S^n are intersections of S^n with hyperplanes in $\mathbb{R}^{n+1}$, or, since we view S^n as a quadric in $\mathbb{R}\mathrm{P}^{n+1}$ (see equation (6)) intersections with hyperplanes in $\mathbb{R}\mathrm{P}^{n+1}$. The pole-polar relationship provides a bijection between hyperplanes that intersect S^n and points outside S^n.

On the other hand, hyperspheres in S^n correspond to hyperspheres in $\widehat{\mathbb{R}^n}$ via stereographic projection. Thus, we have the following bijections:

$$\text{hyperspheres in } \widehat{\mathbb{R}^n} \quad \xleftrightarrow[\text{projection}]{\text{stereogr.}} \quad \text{hyperspheres in } S^n \quad \xleftrightarrow{\text{polarity}} \quad \text{points outside } S^n.$$

The following proposition provides explicit formulas.

Proposition 6.7. *Let $[y] \in \mathbb{R}\mathrm{P}^{n+1}$ be a point outside S^n, i.e.,*

$$\langle y, y\rangle_{n+1,1} > 0.$$

Then the polar plane

$$[y]^\perp = \left\{\, [x] \in \mathbb{R}\mathrm{P}^{n+1} \,\middle|\, \langle y, x\rangle_{n+1,1} \;=\; 0 \,\right\}$$

intersects S^n in a hypersphere

$$S_{[y]} \;=\; S^n \cap [y]^\perp$$

whose image $\sigma(S_{[y]}) \subset \widehat{\mathbb{R}^n}$ under stereographic projection is a hypersphere in $\widehat{\mathbb{R}^n}$:

- *If $y_{n+1} \neq y_{n+2}$, or equivalently, $y_0 \neq 0$, then $\sigma(S_{[y]})$ is the sphere in $\mathbb{R}^n$ with center c and radius r, where*

$$c = \frac{1}{y_0}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \frac{1}{y_{n+2} - y_{n+1}}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \qquad r = \frac{\sqrt{\langle y, y\rangle_{n+1,1}}}{|y_0|} = \frac{\sqrt{\langle y, y\rangle_{n+1,1}}}{|y_{n+2} - y_{n+1}|}\,.$$

- *If $y_{n+1} = y_{n+2}$, or equivalently, $y_0 = 0$, then $\sigma(S_{[y]})$ is the union of $\{\infty\}$ and the hyperplane $\{u \in \mathbb{R}^n \,|\, \langle v, u\rangle = d\,\}$,*

$$v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \qquad d = \frac{y_\infty}{2} = y_{n+1} = y_{n+2}.$$

Conversely, the inverse projection σ^{-1} maps

- *the sphere in $\mathbb{R}^n$ with center c and radius r to the hypersphere $S_{[y]} \subset S^n$ with*

$$[y] = \left[c + (\|c\|^2 - r^2)e_\infty + e_0\right] = \begin{bmatrix} 2c \\ \|c\|^2 - r^2 - 1 \\ \|c\|^2 - r^2 + 1 \end{bmatrix}.$$

- *the union of $\{\infty\}$ and the hyperplane $\{u \in \mathbb{R}^n \,|\, \langle v, u\rangle = d\}$ in $\mathbb{R}^n$ to the hypersphere $S_{[y]} \subset S^n$ with*

$$[y] = [v + 2de_\infty] = \begin{bmatrix} v \\ d \\ d \end{bmatrix}.$$

Remark 6.2. Note that the center of the sphere is obtained by the central projection of the point $[y]$ that represents the sphere.

Proof. As an exemplary case, we show that a point $[x] = [u + \|u\|^2 e_\infty + e_0] \in S^n \backslash \{[e_\infty]\}$ in the polar hyperplane of $[y] = \left[c + (\|c\|^2 - r^2)e_\infty + e_0\right]$ is projected to a point $u = \sigma([x]) \neq \infty$ that lies on the sphere in $\mathbb{R}^n$ with center c and radius r.

$$\begin{aligned} 0 = \langle y, x \rangle_{n+1,1} &= \left\langle c + (\|c\|^2 - r^2)e_\infty + e_0, u + \|u\|^2 e_\infty + e_0 \right\rangle_{n+1,1} \\ &= \|c - u\|^2 - r^2. \end{aligned}$$

□

Proposition 6.8. *The hyperspheres in S^n corresponding to two points $[y_1]$ and $[y_2]$ outside the sphere S^n intersect if and only if*

$$\langle y_1, y_2 \rangle_{n+1,1}^2 \leqslant \langle y_1, y_1 \rangle_{n+1,1} \langle y_2, y_2 \rangle_{n+1,1}$$

In this case the intersection angle θ is determined up to $\theta \mapsto \pi - \theta$ by the equation

$$\cos^2 \theta = \frac{\langle y_1, y_2 \rangle_{n+1,1}^2}{\langle y_1, y_1 \rangle_{n+1,1} \langle y_2, y_2 \rangle_{n+1,1}}.$$

Corollary 6.9. *The hyperspheres corresponding to two points $[y_1]$ and $[y_2]$ outside S^n intersect orthogonally if and only if*

$$\langle y_1, y_2 \rangle_{n+1,1} = 0.$$

6.4 Pencils of spheres

Definition 6.4. In Möbius geometry, a *pencil of spheres* is the family of hyperspheres corresponding to the points of a line in the projective model of Möbius geometry.

There are three different types of sphere pencils, depending on the signature of the corresponding line $\ell \subset \mathbb{R}\mathrm{P}^{n+1}$:

`(++)` The line ℓ does not intersect S^n. Then the sphere pencil consists of all hyperspheres that contain a common fixed $(n-2)$-sphere. Indeed, the polar $(n-1)$-plane $\ell^\perp$ has signature `(n-1,1)` and intersects S^n in a $(n-2)$-sphere. The polar planes $Y^\perp$ of points $Y \in \ell$ are precisely the hyperplanes containing $\ell^\perp$. Hence, the pencil of ℓ consists precisely of all spheres containing $\ell^\perp \cap S^n$.

For $n = 2$, these are all circles through two fixed points. For $n = 3$, these are all spheres through a fixed circle.

`(+-)` The line ℓ intersects S^n in two points. Then the polar $(n-1)$-plane $\ell^\perp$ has signature `(n,0)` and does not intersect S^n. By Corollary 6.9, the pencil of ℓ consists precisely of all hyperspheres that intersect all hyperspheres corresponding to points in $\ell^\perp$ orthogonally.

For $n = 2$, these are all circles orthogonal two circles that intersect in two points, and thus all circles from a pencil of type `(++)`. For $n = 3$, these are all spheres orthogonal two three spheres that span a plane of signature `(+++)`.

`(+0)` The line ℓ is a tangent to S^n. Then the polar $(n-1)$-plane $\ell^\perp$ has signature `(n-1,0,1)` and is also tangent to S^n in the same point P. Furthermore, $\ell^\perp$ is the

common tangent plane for all spheres in ℓ. Thus, the pencil ℓ consists of hyperspheres that are tangent to each other in a fixed point.

For $n = 2$, these are all circles tangent in a common point, while $\ell^\perp$ corresponds to the orthogonal pencil of the same type. For $n = 2$, these are all spheres tangent in a common point.

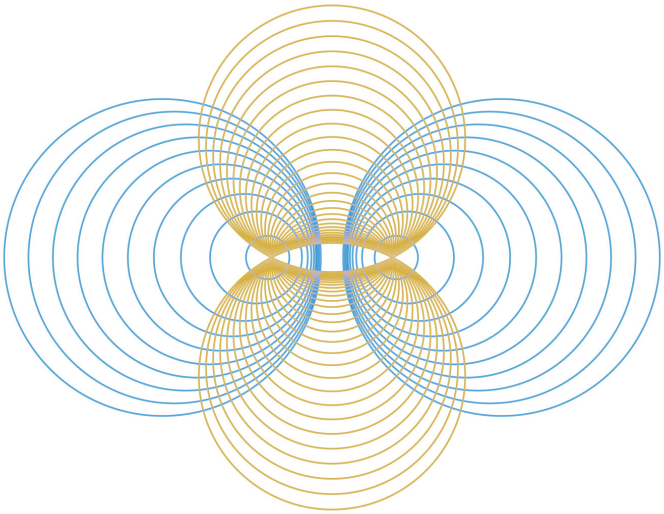


Figure 17. Pencil of circles of type `(++)` (yellow) and its orthogonal pencil of circles of type `(+-)` (blue).

Proposition 6.10. *Möbius transformations map sphere pencils to sphere pencils and preserve their type.*

Proof. This follows directly from Theorem 6.3. In the case `(+-)` consider spheres orthogonal to the pencil. □

6.5 Möbius transformations in the projective model

In the elementary model, Möbius geometry is the geometry of the Möbius group $\text{Möb}(n)$ generated by sphere inversions acting on the extended n-dimensional space $\widehat{\mathbb{R}^n}$. The Möbius group $\text{Möb}(n)$ is geometrically characterized as the group of all transformations that map spheres in $\widehat{\mathbb{R}^n}$ to spheres in $\widehat{\mathbb{R}^n}$. What is the corresponding group of transformations in the projective model of Möbius geometry?

First, let us consider the subgroup of projective transformations of $\mathbb{R}\mathrm{P}^{n+1}$ that map S^n to S^n. This is the projective orthogonal group

$$\mathrm{PO}(n+1,1) \subset \mathrm{PGL}(n+2,\mathbb{R}).$$

Since spheres in S^n are intersections of S^n with planes in $\mathbb{R}\mathrm{P}^{n+1}$ and projective transformations map planes to planes, a projective orthogonal transformation $g \in \mathrm{PO}(n+1,1)$ maps spheres in S^n to spheres in S^n. By the Fundamental Theorem of Möbius geometry (Theorem 6.6), the map

$$\sigma \circ g \circ \sigma^{-1} : \widehat{\mathbb{R}^n} \longrightarrow \widehat{\mathbb{R}^n}$$

is a Möbius transformation. Thus we have a group homomorphism

$$\begin{array}{ccc} \mathrm{PO}(n+1,1) & \longrightarrow & \text{Möb}(n), \\ g & \longmapsto & \sigma \circ g \circ \sigma^{-1}. \end{array} \tag{7}$$

It is injective because the identity on $\mathbb{R}\mathrm{P}^{n+1}$ is the only projective transformation that fixes the sphere S^n pointwise. However, it is also surjective, and thus:

Theorem 6.11. *The map* (7) *is a group isomorphism.*

With this, we complete the *projective model of Möbius geometry*, which is the space $S^n \subset \mathbb{R}\mathrm{P}^{n+1}$ with the action of $\mathrm{PO}(n+1,1)$. Stereographic projection translates between the elementary model and the projective model according to the following dictionary:

elementary model		projective model
$\widehat{\mathbb{R}^n}$	$\longleftrightarrow$	S^n
sphere in $\widehat{\mathbb{R}^n}$	$\longleftrightarrow$	sphere in S^n
Möb(n)	$\longleftrightarrow$	$PO(n+1,1)$

The Möbius group Möb(n) is generated by inversions in spheres and reflections in planes. In the projective model these transformations are given by projective inversions that preserve the quadric.

Proposition 6.12. *Let* $[y] \in \mathbb{R}\mathrm{P}^{n+1}$ *be a point outside* S^n*, i.e.,* $\langle y, y\rangle_{n+1,1} > 0$*. Then the map*

$$g : \mathbb{R}\mathrm{P}^{n+1} \to \mathbb{R}\mathrm{P}^{n+1}, \quad [x] \mapsto \left[x - 2\frac{\langle x, y\rangle_{n+1,1}}{\langle y, y\rangle_{n+1,1}} y \right] \tag{8}$$

is a projective orthogonal transformation, such that $\sigma \circ g \circ \sigma^{-1}$ *is the inversion in the hypersphere (or reflection in the hyperplane) corresponding to the point* $[y]$*.*

Proof. The map $G : x - 2\frac{\langle x,y\rangle_{n+1,1}}{\langle y,y\rangle_{n+1,1}} y$

- is linear,
- invertible since $G(x) = 0$ implies $x = 0$,
- orthogonal since $\langle G(x), G(x)\rangle_{n+1,1} = \langle x, x\rangle_{n+1,1}$.

Furthermore, $g = [G]$ fixes all points on $[y]^\perp$ and preserves all hyperplanes through $[y]$. □

Remark 6.3. The map (8) for a point $[y] \in \mathbb{R}\mathrm{P}^{n+1}$ inside the quadric corresponds to a fixed point free involution, which is still a Möbius transformation, but not a sphere inversion.

7 Curves in projective and Möbius geometry

7.1 Curves in $\mathbb{R}\mathrm{P}^n$

We can lift a curve $\gamma : I \to \mathbb{R}^n$ to $\mathbb{R}\mathrm{P}^n$ by

$$[\hat{\gamma}] : I \to \mathbb{R}\mathrm{P}^n, \quad \hat{\gamma}(t) := \begin{pmatrix} \gamma(t) \\ 1 \end{pmatrix}$$

Then

$$\dot{\hat{\gamma}}(t) = \begin{pmatrix} \dot{\gamma}(t) \\ 0 \end{pmatrix}$$

describes a point at infinity on the lift of the tangent line

$$T(t) = \left\{ [\alpha_1 \hat{\gamma}(t) + \alpha_2 \dot{\hat{\gamma}}(t)] \;\middle|\; \alpha_1, \alpha_2 \in \mathbb{R} \right\} = [\hat{\gamma}(t)] \vee [\dot{\hat{\gamma}}(t)]$$

What happens if we take different representative vectors for the lift of the curve?

More generally, consider two smooth maps

$$\hat{\gamma} : I \to \mathbb{R}^{n+1}, \qquad \lambda : I \to \mathbb{R}\backslash\{0\}.$$

Then $\hat{\gamma}$ and $\tilde{\gamma} := \lambda\hat{\gamma}$ define the same curve in $\mathbb{R}\mathrm{P}^n$

$$[\hat{\gamma}] = [\tilde{\gamma}].$$

But, the derivative changes in the following way

$$\dot{\tilde{\gamma}} = \dot{\lambda}\hat{\gamma} + \lambda\dot{\hat{\gamma}}.$$

Thus, in general, $[\dot{\tilde{\gamma}}] \neq [\dot{\hat{\gamma}}]$, but

$$[\tilde{\gamma}(t)] \vee [\dot{\tilde{\gamma}}(t)] = [\hat{\gamma}(t)] \vee [\dot{\hat{\gamma}}(t)].$$

Therefore, the tangent line stays invariant under the change of the lift (change of representative vectors) for the curve.

Similarly, for higher derivatives, in general, $[\ddot{\tilde{\gamma}}] \neq [\ddot{\hat{\gamma}}]$, but

$$[\tilde{\gamma}(t)] \vee [\dot{\tilde{\gamma}}(t)] \vee [\ddot{\tilde{\gamma}}(t)] = [\hat{\gamma}(t)] \vee [\dot{\hat{\gamma}}(t)] \vee [\ddot{\hat{\gamma}}(t)].$$

Definition 7.1.

(i) A *(projective) curve* is a map

$$[\hat{\gamma}] : I \to \mathbb{R}\mathrm{P}^n$$

with some interval $I \subset \mathbb{R}$ and a smooth map $\hat{\gamma} : I \to \mathbb{R}^{n+1}$

(ii) The curve $[\hat{\gamma}]$ is called *regular* if

$$[\hat{\gamma}(t)] \neq [\dot{\hat{\gamma}}(t)] \quad \text{for all } t \in I,$$

or equivalently, if $\hat{\gamma}(t)$ and $\dot{\hat{\gamma}}(t)$ are linearly independent.

(iii) Let $[\hat{\gamma}]$ be a regular curve.

- The line
$$T(t) = [\hat{\gamma}(t)] \vee [\dot{\hat{\gamma}}(t)]$$
is called the *tangent line* of $[\gamma]$ at $t \in I$.
- If additionally $[\ddot{\hat{\gamma}}(t)] \notin T(t)$, then the plane
$$[\hat{\gamma}(t)] \vee [\dot{\hat{\gamma}}(t)] \vee [\ddot{\hat{\gamma}}(t)]$$
is called the *osculating plane* of $[\hat{\gamma}]$ at $t \in I$.

Note that same as the definition of the tangent line and the osculating plane, the condition for the regularity of $[\hat{\gamma}]$ does not depend on the lift. In affine coordinates, it is equivalent to the regularity that we have introduced for curves in $\mathbb{R}^n$. Furthermore, all these definitions are invariant under projective transformations and under reparametrization of the curve.

7.2 Planar curves in Möbius geometry

Let
$$\gamma : I \to \mathbb{R}^2$$
be a regular planar curve. By inverse stereographic projection, we can map it to the sphere (Möbius lift)
$$[\hat{\gamma}] : I \to \mathbb{S}^2 \subset \mathbb{R}\mathrm{P}^3, \qquad \hat{\gamma}(t) := \gamma(t) + \|\gamma(t)\|^2 e_\infty + e_0.$$

Recall that the osculating circle of γ at $t \in I$ is the circle with center and radius
$$c(t) := \gamma(t) + \frac{1}{\kappa(t)} n(t), \qquad r(t) := \frac{1}{\kappa(t)},$$
where n is the unit normal vector field of γ and
$$\kappa(t) = \frac{\langle \ddot{\gamma}(t), n(t) \rangle}{\|\dot{\gamma}(t)\|^2}$$
is the curvature of γ. Its inverse stereographic projection (Möbius lift) to the sphere is given by
$$[\hat{c}(t)]^\perp \cap \mathbb{S}^2, \qquad \hat{c}(t) := c(t) + (\|c(t)\|^2 - r(t)^2) e_\infty + e_0$$

Proposition 7.1. *Let γ be a regular plane curve. Then the Möbius lift of the osculating circle of γ lies in the osculating plane of the Möbius lift of γ:*
$$[\hat{c}]^\perp = [\hat{\gamma}(t)] \vee [\dot{\hat{\gamma}}(t)] \vee [\ddot{\hat{\gamma}}(t)]$$

Proof. With
$$\hat{c} = \gamma + \tfrac{1}{\kappa} n + e_0 + \left(\|\gamma\|^2 + \tfrac{2}{\kappa} \langle \gamma, n \rangle \right) e_\infty$$
we obtain
$$\langle \hat{\gamma}, \hat{c} \rangle_{3,1} = \left\langle \gamma, \gamma + \tfrac{1}{\kappa} n \right\rangle - \tfrac{1}{2} \|\gamma\|^2 - \tfrac{1}{\kappa} \langle \gamma, n \rangle - \tfrac{1}{2} \|\gamma\|^2 = 0.$$

Now with

$$\dot{\hat{\gamma}} = \dot{\gamma} + 2\langle \gamma, \dot{\gamma} \rangle e_\infty$$

we obtain

$$\left\langle \dot{\hat{\gamma}}, \hat{c} \right\rangle_{3,1} = \left\langle \dot{\gamma}, \gamma + \tfrac{1}{\kappa} n \right\rangle - \langle \gamma, \dot{\gamma} \rangle = 0.$$

Finally, with

$$\ddot{\hat{\gamma}} = \ddot{\gamma} + 2\left(\|\dot{\gamma}\|^2 + \langle \dot{\gamma}, \ddot{\gamma} \rangle \right) e_\infty$$

we obtain

$$\left\langle \ddot{\hat{\gamma}}, \hat{c} \right\rangle_{3,1} = \left\langle \ddot{\gamma}, \gamma + \tfrac{1}{\kappa} n \right\rangle - \|\dot{\gamma}\|^2 - \langle \gamma, \ddot{\gamma} \rangle = 0.$$

□

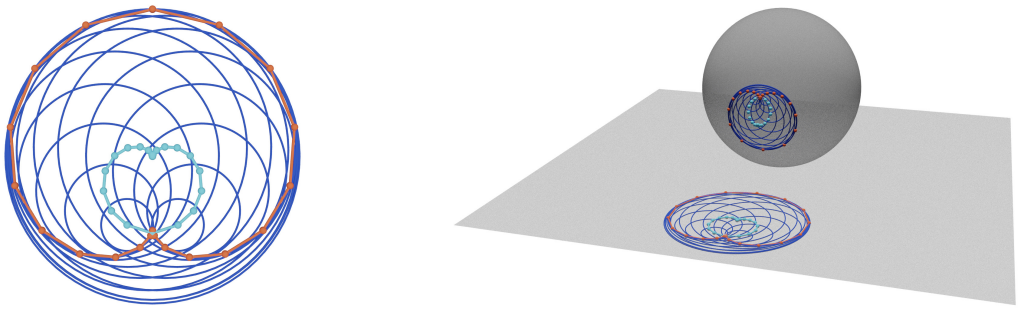


Figure 18. Osculating circles of a cardoid and the lift to Möbius geometry.

Corollary 7.2. *The osculating circle of a planar curve is Möbius invariant.*

Proof. In the Möbius lift osculating circles are given by osculating planes. On the other hand, Möbius transformations are given by projective transformations that preserve the quadric $\mathbb{S}^2$. But projective transformations map osculating planes to osculating planes. □

Remark 7.1. Note that the discrete vertex osculating circles we defined are also Möbius invariant.

Example 7.1. Recall that the evolute of a plane curve consists of the centers of the osculating circles. As an exercise, let we use the Möbius lift to determine the evolute of a parabola

$$\gamma(t) := \begin{pmatrix} t \\ t^2 \end{pmatrix}$$

Its Möbius lift is given by

$$\hat{\gamma}(t) = \begin{pmatrix} t \\ t^2 \end{pmatrix} + (t^2 + t^4) e_\infty + e_0,$$

and its first two derivatives by

$$\dot{\hat{\gamma}}(t) = \begin{pmatrix} 1 \\ 2t \end{pmatrix} + 2(t + 2t^3) e_\infty, \qquad \ddot{\hat{\gamma}}(t) = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + 2(1 + 6t^2) e_\infty.$$

We determine the polar point

$$\hat{c}(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} + c_\infty(t) e_\infty + e_0,$$

From

$$0 = \left\langle \hat{c}, \ddot{\hat{\gamma}} \right\rangle = 2c_2 - 1 - 6t^2$$

we obtain

$$c_2(t) = \frac{1}{2}(1 + 6t^2).$$

and from

$$0 = \left\langle \hat{c}, \dot{\hat{\gamma}} \right\rangle = c_1 + t + 6t^3 - t - 2t^3$$

we obtain

$$c_1(t) = -4t^3$$

Thus, the evolute of γ is given by

$$e(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} = \begin{pmatrix} -4t^3 \\ \frac{1}{2}(1 + 6t^2), \end{pmatrix}$$

which coincides with the solution from Example 5.4. Note, that we don't have to compute c_∞, if we are only interested in the evolute of γ, and not the osculating circles.

8 Roulettes and cycloidal pendulum

8.1 Interlude: complex numbers and geometry

Complex numbers can be a useful tool for computations in Euclidean, similarity, Möbius, and other geometries.

- Consider the action of complex multiplication on the complex plane $\mathbb{C} \cong \mathbb{R}^2$

  $$z \mapsto az, \qquad a \in \mathbb{C},\ a \neq 0.$$

  Geometrically this corresponds to scale-rotation

  $$re^{i\varphi} \mapsto Re^{i\theta} re^{i\varphi} = Rre^{i(\theta+\varphi)},$$

  where $a = Re^{i\theta}$ and $z = re^{i\varphi}$.

- The action of complex addition

  $$z \mapsto z + b, \qquad b \in \mathbb{C}$$

  corresponds to translation.

- Together we obtain all similarity transformations

  $$z \mapsto az + b, \qquad a, b \in \mathbb{C},\ a \neq 0.$$

- The action of complex conjugation

$$z = z_1 + iz_2 \mapsto \bar{z} = z_1 - iz_2$$

  corresponds to reflection in the real axis.

- The absolute value of a complex number $z = z_1 + iz_2$ recovers the Euclidean norm of the corresponding vector

$$|z|^2 = \bar{z}z = z_1^2 + z_2^2.$$

  It is also given by the product of z with its complex conjugate number $\bar{z}$.

- From the complex multiplication of the complex conjugate of a complex number $z = z_1 + iz_2$ with another complex number $w = w_1 + iw_2$

$$\bar{z}w = z_1w_1 + z_1w_2 + i(z_1w_2 - w_1z_2),$$

  we can recover the scalar product and determinant of the two corresponding vectors in $\mathbb{R}^2$:

$$\begin{aligned} \langle z, w\rangle = \Re(\bar{z}w) &= \frac{1}{2}(\bar{z}w + z\bar{w}), \\ \det(z, w) = \Im(\bar{z}w) &= \frac{1}{2i}(\bar{z}w - z\bar{w}). \end{aligned}$$

- Now we can write reflection in a vector $n \in \mathbb{C}$, $n \neq 0$ as

$$z \mapsto -\left(z - 2\frac{\langle z, n\rangle}{\|n\|^2}n\right) = -z + \frac{\bar{z}n + z\bar{n}}{n\bar{n}}n = \frac{n}{\bar{n}}\bar{z}.$$

- Inversion in the unit circle is given by

$$z \mapsto \frac{1}{\bar{z}},$$

  and inversion in the circle with center $c \in \mathbb{C}$ and radius $r > 0$ by

$$z \mapsto c + \frac{r^2}{\bar{z} - \bar{c}}.$$

- Orientation preserving Möbius transformations are given by

$$z \mapsto \frac{az + b}{cz + d}, \qquad a, b, c, d \in \mathbb{C},\ ad - bc \neq 0,$$

  and orientation reversing Möbius transformations are given by

$$z \mapsto \frac{a\bar{z} + b}{c\bar{z} + d}, \qquad a, b, c, d \in \mathbb{C},\ ad - bc \neq 0.$$

8.2 Euclidean motions and instant center of rotation

Let $I \subset \mathbb{R}$ be some interval, and consider a one-parameter family of Euclidean motions

$$A(t) : I \to \mathbb{C}, \quad z \mapsto a(t)z + b(t)$$

with two smooth functions $a, b : I \to \mathbb{C}$, $|a(t)| = 1$ for all $t \in I$.

Consider the trace of some initial point $z_0 \in \mathbb{C}$ under this motion

$$z(t) = A(t)z = a(t)z_0 + b(t).$$

An *instant center of rotation* is given by a point where the velocity vanishes. For every $t \in I$ with $\dot{a}(t) \neq 0$ there is exactly one such point:

$$\dot{z}(t) = \dot{a}(t)z_0 + \dot{b}(t) \quad \Leftrightarrow \quad z_0 = -\frac{\dot{b}(t)}{\dot{a}(t)}.$$

The point z_0 is the initial point which leads to zero velocity. Thus the point in $\mathbb{C}$ at the which zero velocity is attained at $t \in I$ is given by

$$A(t)z_0 = a(t)z_0 + b(t) = b(t) - \frac{a(t)\dot{b}(t)}{\dot{a}(t)}.$$

8.3 Roulettes of curves

A *roulette* is curve described by a point attached to a given curve as that curve rolls without slipping along another fixed curve. Let us denote the given fixed curve, and given rolling curve by

$$\gamma, \rho : I \leftarrow \mathbb{C}.$$

The motion of ρ as it rolls along γ is a Euclidean motion

$$A(t) : z \mapsto a(t)z + b(t), \qquad |a(t)| = 1.$$

At any given time $t \in I$ the point $A(t)\rho(t)$ should coincide with corresponding point on the fixed curve

$$A(t)\rho(t) = a(t)\rho(t) + b(t) = \gamma(t).$$

Thus,

$$b(t) = \gamma(t) - a(t)\rho(t)$$

and $A(t)$ is of the form

$$A(t)z = a(t)z + b(t) = a(t)(z - \rho(t)) + \gamma(t).$$

That ρ rolls along γ without slipping means that the instant center of rotation of the Euclidean motion $A(t)$ is always the point of contact $\gamma(t) = A(t)\rho(t)$ of the two curves. As we will see this will lead to the two curves being tangent remaining tangent at any given time if they were tangent at some time. The instant center of rotation is given by

$$\begin{aligned} b(t) - \frac{a(t)\dot{b}(t)}{\dot{a}(t)} &= \gamma(t) - a(t)\rho(t) - \frac{a(t)(\dot{\gamma}(t) - \dot{a}(t)\rho(t) - a(t)\dot{\rho}(t))}{\dot{a}(t)} \\ &= \gamma(t) - a(t)\frac{\dot{\gamma}(t) - a(t)\dot{\rho}(t)}{\dot{a}(t)}, \end{aligned}$$

which should coincide with $\gamma(t)$. So we obtain

$$a(t) = \frac{\dot{\gamma}(t)}{\dot{\rho}(t)}.$$

In particular, since $|a(t)| = 1$, we must have $|\dot{\gamma}(t)| = |\dot{\rho}(t)|$, i.e., the two curves must be parametrized with the same speed. Now $A(t)$ describes a rotation between the corresponding tangent vectors $\dot{\rho}(t)$ and $\dot{\gamma}(t)$, and thus, the two curves remain tangent if they were initially tangent.

Definition 8.1. Let $\gamma, \rho : I \leftarrow \mathbb{C}$ be two curves parametrized with the same speed $|\gamma(t)| = |\rho(t)|$ for all $t \in I$ and tangent for some initial value $t_0 \in I$. Let $z \in \mathbb{C}$ be a point.

Then the *roulette* tracing the point z attached to the curve ρ as this curve slides without slipping along the fixed curve γ is given by

$$\sigma(t) = \frac{\dot{\gamma}(t)}{\dot{\rho}(t)}(z - \rho(t)) + \gamma(t).$$

Proposition 8.1. *The roulette of a point on a straight line as it rolls along a regular curve is the involute of that curve.*

Proof. Let $\gamma : I \to \mathbb{C}$ be a regular curve, parametrized by arc-length, and ρ a tangent line, parametrized by the arc-length parameter of γ:

$$\rho(s) = \gamma(0) + s\gamma'(0)$$

and let z be some point on this tangent line

$$z = \gamma(0) + a\gamma'(0)$$

Then the roulette is given by

$$\sigma(t) = \frac{\gamma'(s)}{\rho(s)}(z - \rho(s)) + \gamma(s) = \frac{\gamma'(s)}{\gamma'(0)}(a - s)(\gamma'(0) + \gamma(s) = \gamma'(0)(a - s) + \gamma(s)$$

which coincides with the involute of γ. □

Let us compute some more roulette curves:

Example 8.1 (Cycloid)**.** Rolling a circle along a straight line, while following a point on the circle.

$$\gamma(t) = Rt, \qquad \rho(t) = Re^{i(t-\frac{\pi}{2})} + iR = iR(1 - e^{it})$$

Then

$$\dot{\gamma}(t) = R, \qquad \dot{\rho}(t) = Re^{it},$$

and in particular $|\dot{\gamma}(t)| = |\dot{\rho}(t)| = R$. With $z = 0$ the roulette is given by

$$\begin{aligned}\sigma(t) &= \frac{\dot{\gamma}(t)}{\dot{\rho}(t)}(z - \rho(t)) + \gamma(t) = e^{-it}iR(e^{it} - 1) + Rt \\ &= R(t + i - ie^{-it}),\end{aligned}$$

or as a map to $\mathbb{R}^2$ by

$$\sigma(t) = R\begin{pmatrix} t - \sin t \\ 1 - \cos t \end{pmatrix}.$$

One segment (from cusp to cusp) is given by $t \in [0, 2\pi]$.

Let us derive a slightly different normalization of this curve. Reflected at the real axis and translated such that the lowest point lies in the origin:

$$\begin{aligned}\tilde{\sigma}(\varphi) &= \bar{\sigma}(\varphi + \pi) + R(2i - \pi) = R(\varphi + \pi - i + ie^{i(\varphi+\pi)}) + R(2i - \pi) \\ &= R(\varphi + i - ie^{i\varphi}),\end{aligned}$$

or as a map to $\mathbb{R}^2$:

$$\tilde{\sigma}(\varphi) = R \begin{pmatrix} \varphi + \sin\varphi \\ 1 - \cos\varphi \end{pmatrix}.$$

Now one segment (from cusp to cusp with lowest point at the origin) is given by $\varphi \in [-\pi, \pi]$.

Example 8.2 (Cardioid)**.** Rolling a circle along another circle of equal radius, while following a point on the circle.

We start with two circles of radius $R > 0$ centered at $-R$ and R, such that the initial point of contact is the origin $z = 0$.

$$\gamma(t) = Re^{it} - R = R(e^{it} - 1), \qquad \rho(t) = -Re^{-it} + R = R(1 - e^{-it}).$$

Then

$$\dot{\gamma}(t) = iRe^{it}, \qquad \dot{\rho}(t) = iRe^{-it}$$

and $|\dot{\gamma}(t)| = |\dot{\rho}(t)| = R$. The roulette is given by

$$\begin{aligned}\sigma(t) &= \frac{\dot{\gamma}(t)}{\dot{\rho}(t)}(z - \rho(t)) + \gamma(t) = e^{2it}R(e^{-it} - 1) + R(e^{it} - 1) \\ &= R(2e^{it} - e^{2it} - 1),\end{aligned}$$

or alternatively as

$$\begin{aligned}\sigma(t) &= R(2\cos t + 2i\sin t - (\cos t + i\sin t)^2 - 1) \\ &= 2R(1 - \cos t)\cos t + 2iR(1 - \sin t)\cos t \\ &= 2R \begin{pmatrix} (1 - \cos t)\cos t \\ (1 - \sin t)\cos t \end{pmatrix}.\end{aligned}$$

Example 8.3 (Nephroid)**.** Rolling a circle along another circle of twice the radius, while following a point on the circle.

We start with a circle of radius $2R > 0$ centered at the origin, and a circle of radius R centered at $3R$, such that the initial point of contact is $z = 2R$.

$$\gamma(t) = 2Re^{it}, \qquad \rho(t) = -Re^{-2it} + 2R = R(3 - e^{-2it})$$

Then

$$\dot{\gamma}(t) = 2iRe^{it}, \qquad \dot{\rho}(t) = 2iRe^{-2it}$$

and $|\dot{\gamma}(t)| = |\dot{\rho}(t)| = 2R$. The roulette is given by

$$\begin{aligned}\sigma(t) &= \frac{\dot{\gamma}(t)}{\dot{\rho}(t)}(z - \rho(t)) + \gamma(t) = e^{3it}(2R - R(3 - e^{-2it})) + 2Re^{it} \\ &= R(3e^{it} - e^{3it}),\end{aligned}$$

or alternatively as

$$\sigma(t) = R \begin{pmatrix} 3\cos t - \cos 3t \\ 3\sin t - \sin 3t \end{pmatrix}.$$

8.4 Cycloidal pendulum

Harmonic oscillator Consider the harmonic oscillator

$$\ddot{s}(t) = -\omega^2 s(t)$$

with initial conditions $s(0) = s_0$ and $\dot{s}(0) = 0$ is solved by

$$s(t) = s_0 \cos \omega t.$$

It has the property that the period of one full swing

$$T = \frac{2\pi}{\omega}$$

does not depend on the amplitude s_0. Such an oscillator is called *isochronous*.

Isochronous pendulum The string pendulum, where a mass attached to a string swings freely under the influence of gravity does not have this property. However, if the mass is restricted to another path than a circle, can the pendulum become isochronous?

Let

$$\gamma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

be the trajectory of the mass. If s is the arc-length of γ, and θ the angle that the tangent line makes with the x-axis, we have

$$\gamma'(s) = \frac{\mathrm{d}\gamma}{\mathrm{d}s} = \frac{\mathrm{d}\gamma}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}s} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

and thus

$$\dot{x}(t) = \dot{s}(t)\cos\theta(t), \qquad \dot{y}(t) = \dot{s}(t)\sin\theta(t).$$

The force acting on the mass along the curve γ is given by

$$m\ddot{s}(t) = -mg\cos(\frac{\pi}{2} + \theta(t)) = -mg\sin\theta(t),$$

where m is the mass and g the gravitational acceleration. On the other hand, we want to choose γ such that $s(t)$ satisfies the harmonic oscillator equation, i.e.,

$$\ddot{s}(t) = -\omega^2 s(t)$$

with some ω. Thus, we must have

$$g\sin\theta(t) = \omega^2 s(t),$$

which implies

$$g\dot{\theta}(t)\cos\theta(t) = \omega^2 \dot{s}(t).$$

Replacing $\dot{s}$ by $\dot{x}$, we obtain

$$\dot{x}(t) = \frac{g}{\omega^2}\dot{\theta}(t)\cos^2\theta(t)$$

Expressing x as a function of θ we can eliminate the time dependence

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = \frac{g}{\omega^2}\cos^2\theta,$$

which is solved by

$$x(\theta) = \frac{g}{4\omega^2}(2\theta + \sin 2\theta) + c_1$$

with some constant of integration c_1.

Similarly, replacing $\dot{s}$ by $\dot{y}$, we obtain

$$\dot{y}(t) = \frac{g}{\omega^2}\dot{\theta}(t)\cos\theta(t)\sin\theta(t) = \frac{g}{2\omega^2}\dot{\theta}(t)\sin 2\theta(t)$$

Expressing y as a function of θ we can eliminate the time dependence

$$\frac{\mathrm{d}y}{\mathrm{d}\theta} = \frac{g}{2\omega^2}\sin 2\theta,$$

which is solved by

$$y(\theta) = -\frac{g}{4\omega^2}(\cos 2\theta) + c_2$$

with some constant of integration c_2.

If we set $\varphi = 2\theta$, $c_1 = 0$, and $c_2 = R = \frac{g}{4\omega^2}$, we obtain

$$\gamma(\varphi) = R\begin{pmatrix}\varphi + \sin\varphi \\ 1 - \cos\varphi\end{pmatrix}$$

which is the cycloid from Example 8.1. Thus, if we restrict the pendulum to this curve, the arc-length parameter

$$s(t) = \frac{g}{\omega^2}\sin\theta(t) = \frac{g}{\omega^2}\sin\frac{\varphi(t)}{2}$$

satisfies the harmonic oscillator equation, i.e.,

$$s(t) = s_0\cos\omega t.$$

Thus

$$\sin\frac{\varphi(t)}{2} = A\cos\omega t$$

with $A = \frac{g}{\omega^2 s_0} < 1$ describing the amplitude of the pendulum.

String construction Can we restrict the string of a pendulum such that the trajectory of the mass is restricted to the cycloid? If a string wraps around some curve, a point on the taut end of the string moves along the involute of that curve. Vice versa, if the involute is given, the curve we need to restrict the string is the evolute of that curve.

Thus, to realize the cycloidal pendulum with mass attached to a string, we need to compute the evolute of the cycloid.

With

$$\gamma(\varphi) = R(\varphi + i - ie^{i\varphi})$$

we obtain

$$\dot{\gamma}(\varphi) = R(1 + e^{i\varphi}), \quad n(\varphi) = i\dot{\gamma}(\varphi) = iR(1 + e^{i\varphi})\ddot{\gamma}(\varphi) = iRe^{i\varphi},$$

and

$$\begin{aligned}
|\dot\gamma(\varphi)|^2 &= R^2(1+e^{i\varphi})(1+e^{-i\varphi}) = R^2(2+e^{i\varphi}+e^{-i\varphi}) \\
&= 2R^2(1+\cos\varphi) \\
\langle n(\varphi), \ddot\gamma(\varphi)\rangle &= \frac{1}{2}(\bar n\ddot\gamma + n\bar{\ddot\gamma}) = \frac{R^2}{2}((1+e^{-i\varphi})e^{i\varphi} + (1+e^{i\varphi})e^{-i\varphi}) = R^2(2+e^{i\varphi}+e^{-i\varphi}) \\
&= R^2(1+\cos\varphi).
\end{aligned}$$

Thus, the evolute of the cycloid γ is given by

$$\begin{aligned}
\gamma(\varphi) + \frac{|\dot\gamma(t)|^2}{\langle n(t), \ddot\gamma(t)\rangle} n(t) &= R(\varphi + i - ie^{i\varphi}) + 2iR(1+e^{i\varphi}) \\
&= R(\varphi + 3i + ie^{i\varphi}) = \gamma(\varphi+\pi) + R(2i-\pi),
\end{aligned}$$

which is a translated cycloid.

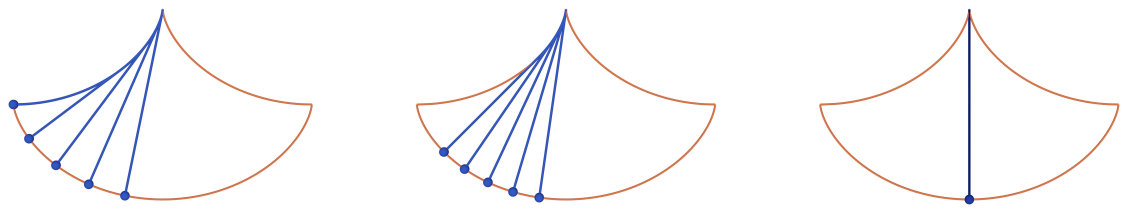


Figure 19. Cycloidal pendulum.

Result A mass attached to a string of length $L = 4R$ suspended from the origin, such that the string wraps around a cycloid of radius R with cusp in the origin, moves on a cycloid

$$\gamma(\varphi) = R(\varphi - 3i - ie^{i\varphi}) = R\begin{pmatrix} \varphi + \sin\varphi \\ -3 - \cos\varphi \end{pmatrix},$$

where the motion is given by

$$\sin\frac{\varphi(t)}{2} = A\cos\omega t$$

with $\omega^2 = \frac{g}{4R}$ and $A \leqslant 1$, which determines the amplitude of the pendulum motion. The period is given by

$$T = \frac{2\pi}{\omega} = 4\pi\sqrt{\frac{R}{g}}.$$

9 Billiards and caustics

Here we treat problems on reflecting rays in plane curves.

9.1 Optical properties of conics

We now present some optical properties of conic sections. Light rays are represented by straight lines. If the rays hit a reflective surface (a "mirror"), the *law of reflection* states that the incoming and outgoing ray have the same angle with the normal line of the surface (or equivalently the tangent line) at the point of reflection.

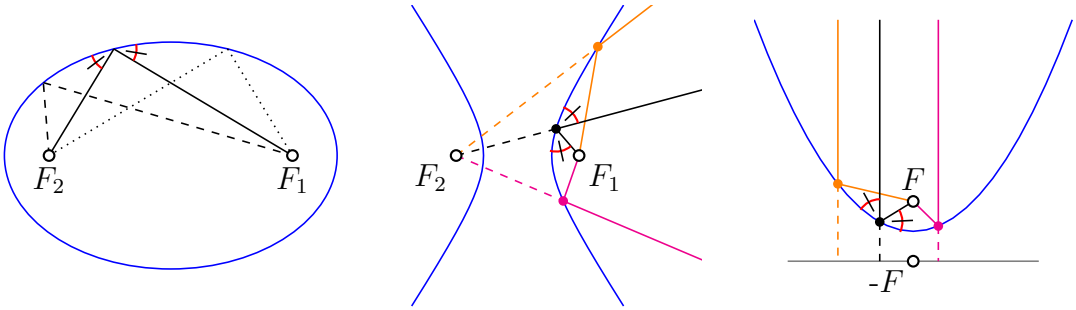


Figure 20. Reflection in conic mirrors.

Theorem 9.1. *Light rays emitted from one focus of an elliptic mirror after reflection go through the other focus (see Fig. 20, left). Light rays emitted from one focus of a hyperbolic mirror are reflected as if emitted from the other focus (see Fig. 20, middle). Light rays emitted from the focus of a parabolic mirror after reflection become parallel to the axis of the parabola (see Fig. 20, right).*

Proof. We give a proof of the elliptic case only. The other two can be proven similarly. Let P be a point on the ellipse with distances r_1 and r_2 from the foci F_1 and F_2, respectively (see Figure 21). Extend the line segment F_2P by a distance of r_1 beyond P. Call the new endpoint of the extended segment F_1'. Let ℓ be the perpendicular bisector of F_1F_1'. We will show that ℓ is the tangent line of the ellipse at P, and thus, F_1' is the reflection of F_1 in this tangent line. From this, the equality of the angles follows easily.

Indeed, P lies on ℓ because it has equal distance r_1 from F_1 and F_1'. Consider any other point $\tilde{P}$ on ℓ and let $\tilde{r}_1$ be its distance to both F_1 and F_1' and let $\tilde{r}_2$ be its distance to F_2. Then the triangle inequality for the triangle $F_2\tilde{P}F_1'$ reads

$$\tilde{r}_1 + \tilde{r}_2 > r_1 + r_2,$$

so $\tilde{P}$ does not lie on the ellipse. Hence, ℓ intersects the ellipse in precisely one point, P, and thus is tangent to P. □

9.2 Elliptic billiards

A *billiard trajectory* in an ellipse is a sequence of points on that ellipse and their connecting edges such that at every point the law of reflection is satisfied (see Figure 22).

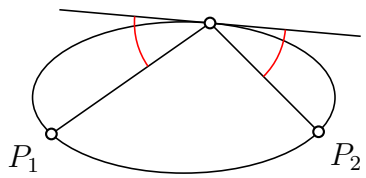


Figure 22. At every point on the ellipse a billiard trajectory satisfies the law of reflection (equal angles of the incoming and outgoing rays).

In order to state the next theorem on elliptic billiards we need to first introduce confocal conics.

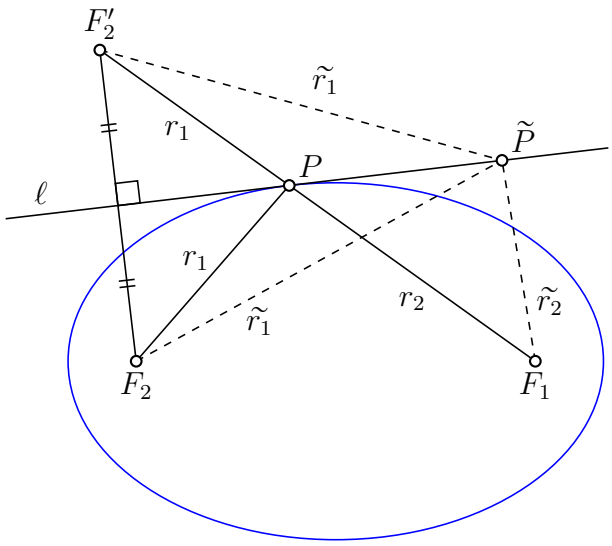


Figure 21. Illustration of the proof of the optical properties of an ellipse.

Definition 9.1 (confocal conics)**.** Let F_1, F_2 be two points in the Euclidean plane $\mathbb{R}^2$. Then the family of all conics with same foci F_1 and F_2 is called a family of *confocal conics*.

It is easy to show that up to Euclidean transformation a family of confocal conics is given by the formula

$$\mathcal{Q}_\lambda = \left\{ (x,y) \in \mathbb{R}^2 \;\middle|\; \frac{x^2}{a_1+\lambda} + \frac{y^2}{a_2+\lambda} = 1 \right\}, \qquad \lambda \in \mathbb{R},$$

for with some $a_1 > a_2$. It includes ellipses ($\lambda > -a_2$) and hyperbolas ($-a_1 < \lambda < -a_2$). The foci of this family are given by $(\pm f, 0) = (\pm\sqrt{a_1 - a_2}, 0)$.

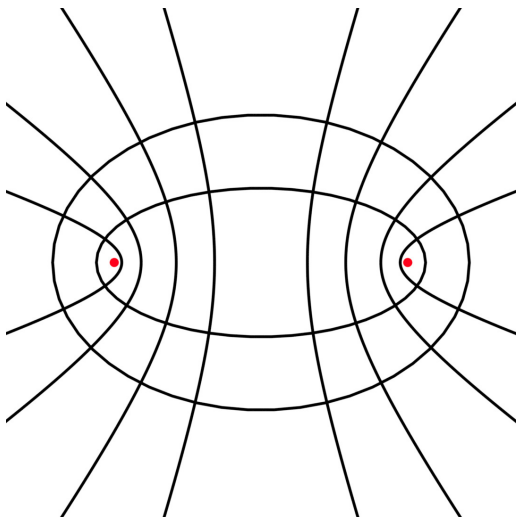


Figure 23. Confocal conics: Through every point in the plane goes exactly one ellipse and one hyperbola from a confocal family and intersect orthogonally in this point. Conversely, an ellipse and a hyperbola from a family of confocal conics always intersect in four points, while two confocal ellipses or two confocal hyperbolas do not intersect.

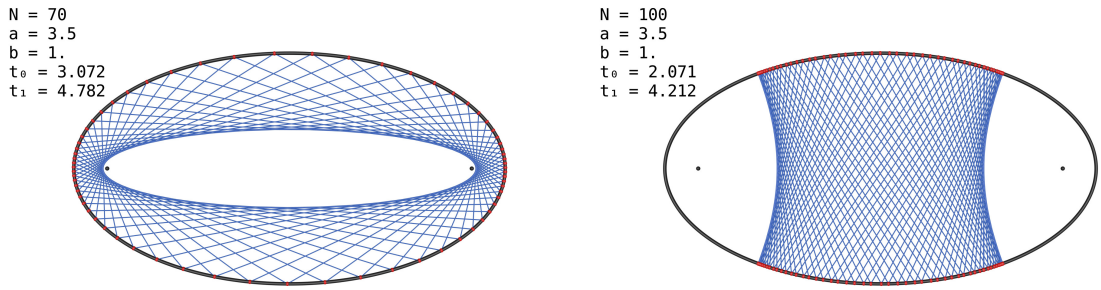


Figure 24. Elliptic billiard trajectories.

Theorem 9.2. *The lines of a billiard trajectory inside an ellipse are tangent to a fixed confocal conic.*

Proof. Let A_0A_1 and A_1A_2 be the two subsequentive lines of the trajectory, and assume that the line A_0A_1 does not intersect the segment $[F_1F_2]$. From the optical properties of the ellipses (see Theorem 9.1) we have $\angle A_0A_1F_1 = \angle A_2A_1F_2$. Reflect F_1 and F_2 in the lines A_0A_1 and A_1A_2 respectively, we obtain F_1' and F_2' (see Figure 25). Define $B = F_1'F_2 \cap A_0A_1$ and $C = F_2'F_1 \cap A_1A_2$. Let $\mathcal{Q}_1$ be the conic with foci F_1, F_2 (confocal) that is tangent to A_0A_1. From the optical properties of ellipses (equal reflection angles) we see that $\mathcal{Q}_1$ touches A_0A_1 at B. Similarly, the confocal conic $\mathcal{Q}_2$ touches the line A_1A_2 at C.

To prove that $\mathcal{Q}_1 = \mathcal{Q}_2$ it is enough to show that $|F_2F_1'| = |F_1F_2'|$. The triangles $F_1A_1F_2'$ and $F_1'A_1F_2$ are congruent, they have the same angle at A_1 and equal pairs of edges at this vertex. Their third edges must also coincide: $|F_2F_1'| = |F_1F_2'|$.

Thus, two consecutive and then all edges are tangent to the same confocal conic. □

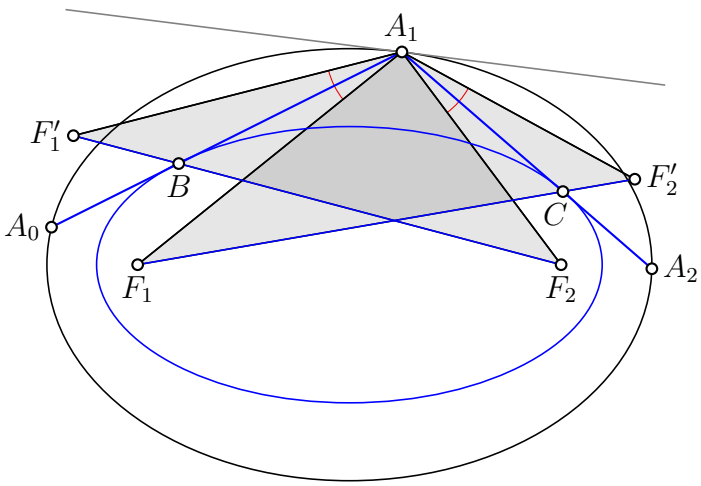


Figure 25. Proof of Theorem 9.2.

9.3 Caustics

A caustic is the envelope of rays reflected or refracted by an object represented by a curve in the plane. We will consider parallel incoming rays, and discuss this concept by looking at an example.

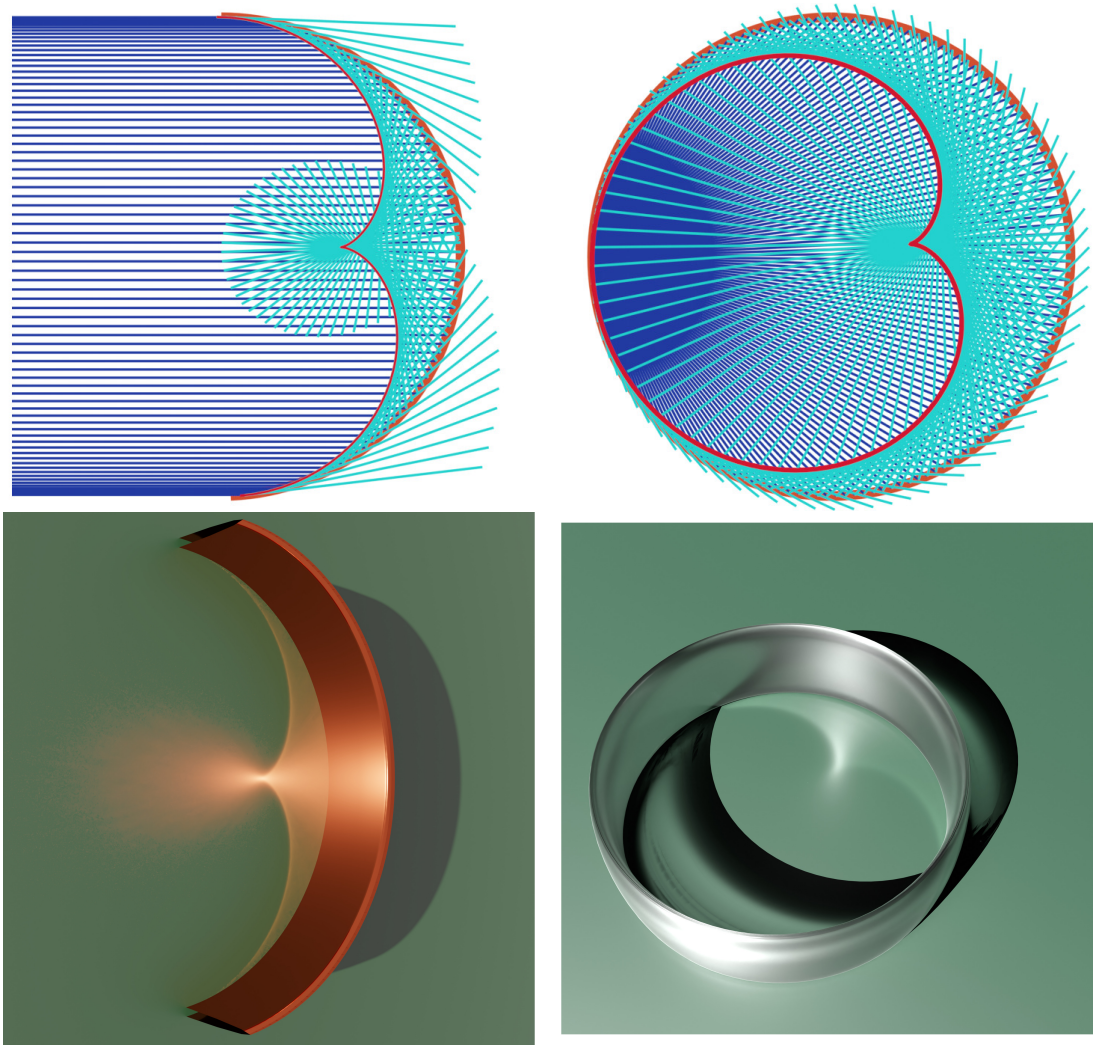


Figure 26. Caustics of a circle.

Example 9.1. Consider a circle in the plane of radius R, parametrized by

$$\gamma(t) := Re^{it}.$$

We compute the caustic of the circle for the reflection of incoming parallel rays in the direction of the real axis.

The normal vector of the incoming rays is given by $u = i$, A tangent vector at some point $\gamma(t)$ of the curve by

$$T(t) = \dot{\gamma}(t) = iRe^{it}$$

Thus, the normal vector of the reflected ray is given by

$$v(t) = \frac{T(t)}{\bar{T}(t)}\bar{u} = \frac{-iRe^{it}}{iRe^{-it}}(-i) = ie^{2it},$$

and the equation for the line of the reflected ray by

$$0 = \langle v(t), z - \gamma(t)\rangle = \left\langle ie^{2it}, z - Re^{it}\right\rangle = \left\langle ie^{2it}, z\right\rangle - R\left\langle ie^{2it}, e^{it}\right\rangle$$

The two scalar products can be written as

$$\begin{aligned}\left\langle ie^{2it}, z\right\rangle &= \frac{1}{2}(ie^{2it}\bar{z} - ie^{-2it}z),\\ \left\langle ie^{2it}, e^{it}\right\rangle &= \frac{1}{2}(ie^{2it}e^{-it} - ie^{-2it}e^{it}) = -\frac{1}{2i}(e^{it} - e^{-it}) = -\sin t.\end{aligned}$$

We arrive at the following equation for the reflected ray

$$ie^{-2it}z - ie^{2it}\bar{z} - 2R\sin t = 0.$$

To compute the envelope of this one-parameter family of lines, we combine this equation with its partial derivative w.r.t. t

$$2e^{-2it}z + 2e^{2it}\bar{z} - 2R\cos t = 0.$$

By eliminating $\bar{z}$ we obtain

$$4ie^{-2it}z - 4R\sin t - 2iR\cos t = 0.$$

Solving for z yields

$$z = R\left(\frac{1}{i}\sin t - \frac{1}{2}\cos t\right)e^{2it} = \frac{R}{4}(3e^{it} - e^{3it}),$$

which is the parametric representation of a nephroid.

10 Families of circles: envelopes and orthogonal trajectories

Evolutes and involutes are envelopes and orthogonal trajectories to a one-parameter family of lines, respectively. Here we treat the analogous problems for one-parameter families of circles, which are both Möbius invariant.

Let $I \subset \mathbb{R}$ be an open interval,

$$c : I \to \mathbb{R}^2, \qquad r : I \to \mathbb{R}_+,$$

smooth functions, and consider the one-parameter family of circles

$$C(t) = \left\{x \in \mathbb{R}^2 \;\middle|\; F(t,x) := \|x - c(t)\|^2 - r(t)^2 = 0\right\}$$

in the Euclidean plane $\mathbb{R}^2$.

10.1 Envelopes of one-parameter families of circles

An envelope of the one-parameter family of circles $C(t)$ is given by a curve $\gamma : I \to \mathbb{R}^2$ that satisfies

$$\begin{aligned}0 &= F(t,x) = \|x - c(t)\|^2 - r(t^2)\\ 0 &= \partial_t F(t,x) = -2\langle x - c(t), \dot{c}(t)\rangle - 2r(t)\dot{r}(t)\end{aligned}$$

with $x = \gamma(t)$. The first equation is satisfied by the general ansatz

$$x = \gamma(t) = c(t) + r(t)\begin{pmatrix}\cos\theta(t)\\ \sin\theta(t)\end{pmatrix}$$

with some function $\theta : I \to \mathbb{R}$, for which the second equation becomes

$$\dot{c}_1 \cos\theta + \dot{c}_2 \sin\theta + \dot{r} = 0.$$

Upon introducing the function

$$\tau(t) := \tan\frac{\theta(t)}{2},$$

this equation turns into a quadratic equation

$$(\dot{r} - \dot{c}_1)\tau^2 + 2\dot{c}_2\tau + \dot{r} + c_1 = 0, \tag{9}$$

where we use the trigonometric identities

$$\sin\theta = \frac{2\tau}{1+\tau^2}, \qquad \cos\theta = \frac{1-\tau^2}{1+\tau^2}.$$

Thus, generically the envelope consists of two curves.

Proposition 10.1. *For a one-parameter family of cirles $C(t)$ the following holds for any part of the interval I:*

- *If $\|\dot{c}(t)\|^2 > \dot{r}(t)^2$, the envelope consists of two curves.*
- *If $\|\dot{c}(t)\|^2 = \dot{r}(t)^2$, the envelope consists of one curve.*
- *If $\|\dot{c}(t)\|^2 < \dot{r}(t)^2$, the envelope does not exist.*

Proof. The discriminant of the quadratic equation (9) that describes the envelope is given by

$$\Delta = 4\dot{c}_2^2 - 2(\dot{r} - \dot{c}_1)(\dot{r} + \dot{c}_1) = 4\,\|\dot{c}(t)\|^2 - 4\dot{r}(t)^2$$

□

Remark 10.1. In the case $\|\dot{c}(t)\|^2 > \dot{r}(t)^2$ let $\gamma_1(t)$ and $\gamma_2(t)$ be the two envelopes. Let $\tilde{C}(t)$ be the circle through $\gamma_1(t)$ and $\gamma_2(t)$ orthogonal to $C(t)$. Then the inversion $\iota_{\tilde{C}(t)}$ in the circle $\tilde{C}(t)$ is a first-order symmetry for both envelopes, i.e.,

$$\iota_{\tilde{C}(t)} \circ \gamma_i(t) = \gamma_i(t), \qquad (\iota_{\tilde{C}(t)} \circ \gamma_i)^{\cdot}(t) = \alpha\dot{\gamma}_i(t) \quad \text{with some } \alpha \in \mathbb{R}, \alpha \neq 0.$$

Finding the envelopes of a one-parameter family of circles is a Möbius invariant problem. Thus, let us transfer it to the projective model of Möbius geometry. The one-parameter family of circles $C(t)$ can be represented by a curve outside the Möbius quadric given by

$$\hat{c}(t) = c(t) + \left(\|c(t)\|^2 - r(t)^2\right) e_\infty + e_0.$$

Its tangent line is the span $[\hat{c}(t)] \vee [\dot{\hat{c}}(t)]$, where

$$\dot{\hat{c}}(t) = \dot{c}(t) + 2\left(\langle \dot{c}(t), c(t)\rangle - \dot{r}(t)r(t)\right) e_\infty.$$

Now the three cases from Proposition 10.1 correspond to the three possible signatures of the tangent line:

Proposition 10.2. *The tangent line of the Möbius lift $[\hat{c}(t)]$ of the one-parameter family of circles $C(t)$*

- *has signature* `(++)` *if* $\|\dot{c}(t)\|^2 > \dot{r}(t)^2$,
- *has signature* `(+0)` *if* $\|\dot{c}(t)\|^2 = \dot{r}(t)^2$,
- *has signature* `(+-)` *if* $\|\dot{c}(t)\|^2 < \dot{r}(t)^2$.

Proof. We have

$$\langle \hat{c}(t), \hat{c}(t) \rangle_{3,1} = r(t)^2, \qquad \left\langle \dot{\hat{c}}(t), \dot{\hat{c}}(t) \right\rangle_{3,1} = \|c(t)\|^2, \qquad \left\langle \hat{c}(t), \dot{\hat{c}}(t) \right\rangle_{3,1} = r(t)\dot{r}(t).$$

Thus, the determinant of a Gram matrix for the tangent line is given by

$$\det \begin{pmatrix} r^2 & r\dot{r} \\ r\dot{r} & \|\dot{c}\|^2 \end{pmatrix} = r^2(\|\dot{c}\|^2 - \dot{r}^2).$$

□

Let us have a more direct look at envelope equation in Möbius geometry.

Proposition 10.3. *A curve* $\gamma : I \to \mathbb{R}^2$ *is the envelope of a one-parameter family of circles* $C(t)$ *if its Möbius lift*

$$\hat{\gamma} : I \to \mathbb{S}^2 \subset \mathbb{R}\mathrm{P}^3, \qquad \hat{\gamma}(t) = \gamma(t) + \|\gamma(t)\|^2 e_\infty + e_0$$

satisfies

$$\langle \hat{\gamma}(t), \hat{c}(t) \rangle_{3,1} = 0, \qquad \left\langle \hat{\gamma}(t), \dot{\hat{c}}(t) \right\rangle_{3,1} = 0.$$

Proof.

$$\left\langle \hat{\gamma}(t), \dot{\hat{c}}(t) \right\rangle_{3,1} = \langle \gamma(t) - c(t), \dot{c}(t) \rangle + r(t)\dot{r}(t).$$

□

Thus, the envelope lies on the polar line of the tangent line of $[\hat{c}]$.

Proposition 10.4. *If* $\|\dot{c}(t)\| > r(t)$ *the Möbius lift of the two envelopes of the one-parameter family of circles* $C(t)$ *is given by*

$$([\hat{c}(t)] \vee [\dot{\hat{c}}(t)])^\perp \cap \mathbb{S}^2.$$

Discrete envelope of a one-parameter family of circles Similar to the discretization of the evolute, we can obtain a simple discretization of the envelope for a discrete one-parameter family of circles by taking the points of intersection of consecutive circles.

In the projective model of Möbius geometry this corresponds to considering the points of intersection of the polar line of the edge tangent lines of the Möbius lift of the discrete one-parameter family of circles.

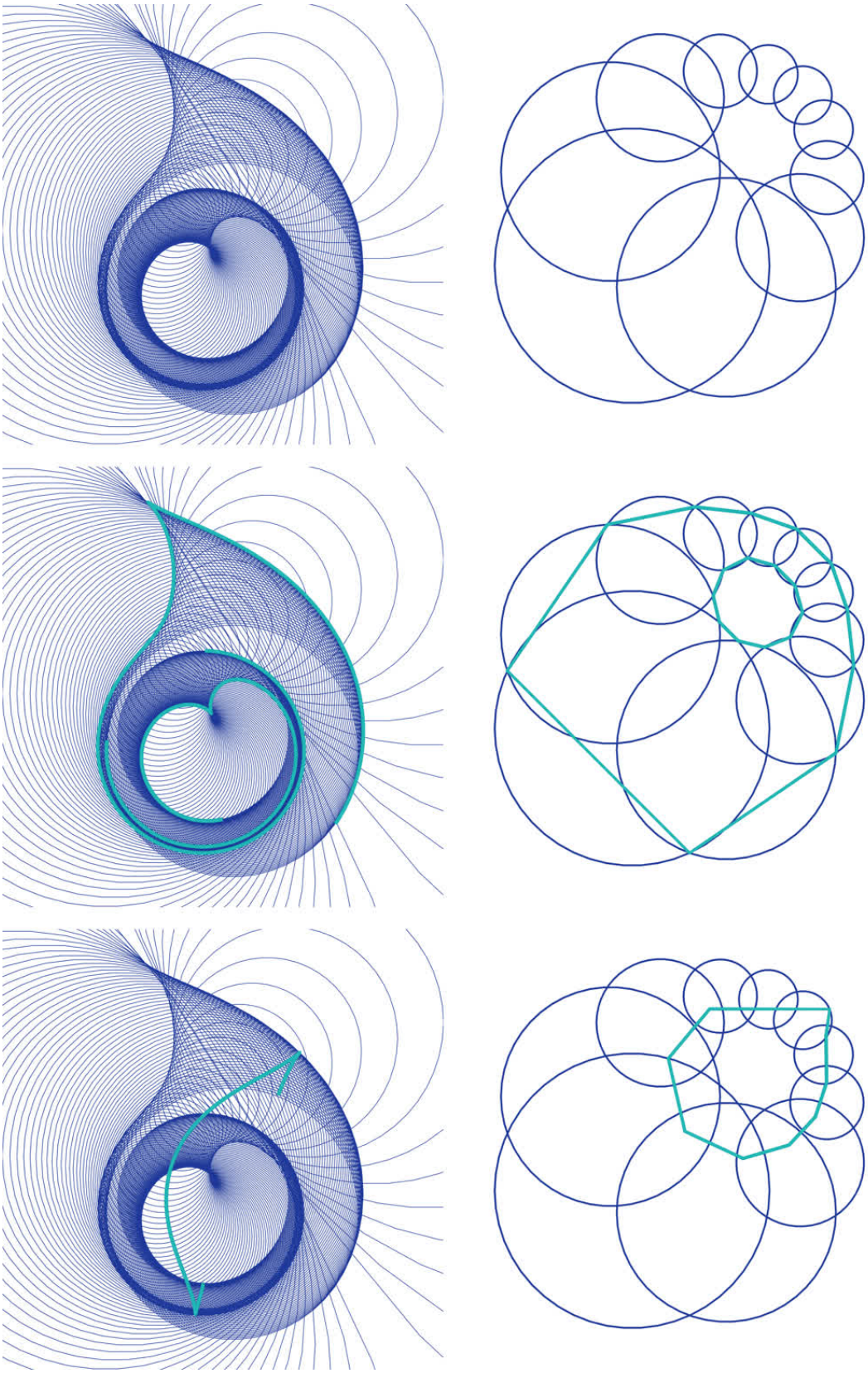


Figure 27. Smooth and discrete envelope and orthogonal trajectory of a one-parameter family of cirlces.

10.2 Orthogonal trajectories of one-parameter families of circles

An orthogonal trajectory of the one-parameter family of circles $C(t)$ is given by a curve $\gamma : I \to \mathbb{R}^2$ that satisfies

$$\begin{aligned} 0 &= F(t, \gamma(t)) = \|\gamma(t) - c(t)\|^2 - r(t^2) \\ 0 &= \det(\dot\gamma(t), \nabla_x F(t, \gamma(t))) = 2\det(\dot\gamma(t), \gamma(t) - c(t)) \end{aligned}$$

Employing the same ansatz as before

$$x = \gamma(t) = c(t) + r(t) \begin{pmatrix} \cos\theta(t) \\ \sin\theta(t) \end{pmatrix}$$

we obtain

$$\det(\gamma(t) - c(t), \dot\gamma(t)) = r \det \begin{pmatrix} \cos\theta & \dot c_1 \\ \sin\theta & \dot c_2 \end{pmatrix} + r^2 \dot\theta \det \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

and thus

$$c_1 \sin\theta - c_2 \cos\theta = r\dot\theta.$$

And with

$$\tau(t) := \tan\frac{\theta(t)}{2}, \qquad \dot\tau(t) = \frac{\dot\theta(t)(1 + \tau(t)^2)}{2}$$

we obtain

$$2r\dot\tau = 2\dot c_1 \tau - \dot c_2 (1 - \tau^2),$$

which is a *Riccati equation*.

The general solution of a Riccati equation is of the form

$$\tau(t) = \frac{a(t)\tau_0 + b(t)}{c(t)\tau_0 + d(t)}$$

with some coefficients $a(t), b(t), c(t), d(t)$ and initial value τ_0.

Remark 10.2. For an orthogonal trajectory the inversion $\iota_{C(t)}$ in the circle $C(t)$ is a first-order symmetry.

Discrete orthogonal trajectories of a one-parameter family of circles Similar to the discretization of the involute, we can obtain a simple discretization of orthogonal trajectories of a family of circles by taking an arbitrary initial point and reflecting it by inversions in the circles of the discrete one-parameter family of circles.

Same as for discrete involutes, this procedure has the disadvantage, that it highly depends on a suitable choice for the initial point, to even stay close to the circles of the family.

We will introduce another discretization, which is based on mapping an initial point on one of the circles of the family to the other circles by inversion, and thus ensuring, that we stay on the circles of the family.

10.3 Tractrix and Darboux transform

Assume that a point moves along a curve γ and pulls an interval $(\gamma, \hat{\gamma})$ so that the distance $\|\hat{\gamma} - \gamma\|$ ist constant, and the velocity vector $\hat{\gamma}'$ is parallel to $\gamma - \hat{\gamma}$. The curve $\hat{\gamma}$ can be thought of as a trajectory of the second wheel of a bicycle whose first wheel moves along the curve γ.

Definition 10.1 (tractrix)**.** Let $\gamma : I \to \mathbb{R}^2$ be a plane curve. A curve $\hat{\gamma} : I \to \mathbb{R}^2$ is called a *tractrix* of γ if

$$\|\hat{\gamma} - \gamma\| = \textit{const.} \qquad \text{and} \qquad \dot{\hat{\gamma}} \parallel (\hat{\gamma} - \gamma)$$

On the other hand, this is the same as saying that $\hat{\gamma}$ lies on a circle with center γ and constant radius $\|\hat{\gamma} - \gamma\|$, while moving orthogonal to that circle. Thus, a tractrix is the special case of an orthogonal trajectory to a one-parameter family of circles where the radius of the circles is constant.

Lemma 10.5. *Let $\gamma : I \to \mathbb{R}^2$ be a curve, and $\hat{\gamma}$ a tractrix of γ. Then the curve*

$$\tilde{\gamma} := \gamma + 2(\hat{\gamma} - \gamma) = 2\hat{\gamma} - \gamma.$$

is parametrized by the same speed as γ, i.e.,

$$\|\dot{\gamma}\| = \|\dot{\tilde{\gamma}}\|,$$

and $\hat{\gamma}$ is a tractrix of $\tilde{\gamma}$ as well.

Proof. Let $v := \hat{\gamma} - \gamma$. Then

$$\|\dot{\gamma}\|^2 - \|\dot{\tilde{\gamma}}\|^2, = \left\langle \dot{\tilde{\gamma}} + \dot{\gamma}, \dot{\tilde{\gamma}} - \dot{\gamma} \right\rangle = 4 \left\langle \dot{\hat{\gamma}}, \dot{\hat{\gamma}} - \dot{\gamma} \right\rangle \sim 2 \langle v, \dot{v} \rangle = \frac{\mathrm{d}}{\mathrm{d}t} \|v\|^2 = 0.$$

Furthemore,

$$\hat{\gamma} - \tilde{\gamma} = -v,$$

and thus $\hat{\gamma}$ is a tractrix of $\tilde{\gamma}$. □

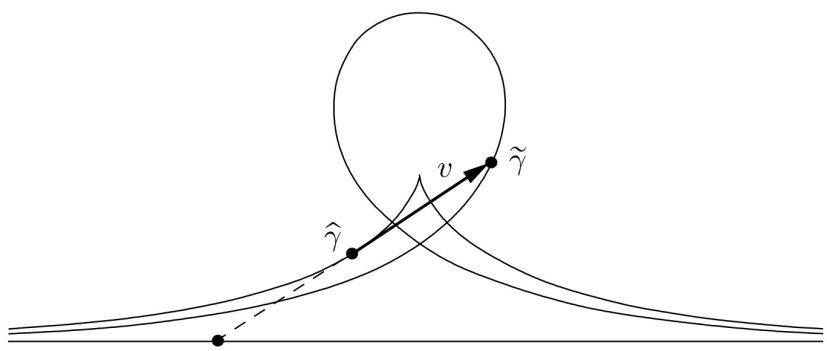


Figure 28. A traktrix and the corresponding Darboux transform of γ.

Definition 10.2 (Darboux transform)**.** Two curves $\gamma, \tilde{\gamma} : I \to \mathbb{R}^2$ parametrized by the same speed (in particular two arc-length parametrized curves) are called *Darboux transforms* of each other if

$$\|\tilde{\gamma} - \gamma\| = \textit{const.},$$

and $\tilde{\gamma}$ is not just a translate of γ.

Lemma 10.5 showed how to construct a Darboux transform from a tractrix. The verse construction also holds.

Theorem 10.6. *Let $\gamma : I \to \mathbb{R}^2$ be a curve. Then the following claims are equivalent:*

(i) $\tilde{\gamma}$ is a Darboux transform of γ

(ii) $\hat{\gamma} := \frac{1}{2}(\gamma + \tilde{\gamma})$ is a tractrix of γ (and of $\tilde{\gamma}$).

Proof.

($\Leftarrow$) Lemma 10.5.

($\Rightarrow$) It is clear that $v := \frac{1}{2}(\tilde{\gamma} - \gamma)$ is of constant length. It remains to show that $\dot{\hat{\gamma}} \parallel v$. Since $\dot{v} \perp v$ this is equivalent to $\dot{\hat{\gamma}} \perp \dot{v}$.

$$\left\langle \dot{\hat{\gamma}}, \dot{v} \right\rangle = \left\langle \frac{1}{2}(\dot{\gamma} + \dot{\tilde{\gamma}}), \frac{1}{2}(\dot{\tilde{\gamma}} - \dot{\gamma}) \right\rangle = \frac{1}{4}(\|\dot{\tilde{\gamma}}\|^2 - \|\dot{\gamma}\|^2) = 0$$

□

10.4 Midcircles

Definition 10.3. Let C_1, C_2 be two circles (or lines) in $\widehat{\mathbb{R}^2}$. A circle K such that the inversion in K maps C_1 to C_2 is called a *midcircle* of C_1 and C_2. Given C_1 and C_2, we want to find the midcicles K (if they exists).

In the projective model of Möbius geometry the circles C_1, C_2, K are represented by points $[x_1], [x_2], [y] \in \mathbb{RP}^3$ outside the Möbius quadric. We assume x_1, x_2 are normalized to satisfy

$$\langle x_1, x_1 \rangle_{3,1} = \langle x_2, x_2 \rangle_{3,1} = 1$$

and we are looking for y in dependence of x_1, x_2 (if it exists).

In the Möbius lift the inversion in K is represented by the linear map

$$\iota_y : \mathbb{R}^4 \to \mathbb{R}^4, \quad x \mapsto x - 2\frac{\langle x, y \rangle_{3,1}}{\langle y, y \rangle_{3,1}} y.$$

Note that for $x \in \mathbb{R}^4$

$$\langle \iota_y(x), \iota_y(x) \rangle_{3,1} = \langle x, x \rangle_{3,1},$$

The condition for y is given by $[\iota_y(x_1)] = [x_2]$, or equivalently,

$$\iota_y(x_1) = x_1 - 2\frac{\langle x_1, y \rangle_{3,1}}{\langle y, y \rangle_{3,1}} y = \pm x_2. \tag{10}$$

In particular, this means $[y] \in [x_1] \vee [x_2]$, i.e., if we exclude the case $K = C_1 = C_2$

$$y = x_1 + \sigma x_2 \qquad \text{for some } \sigma \in \mathbb{R}.$$

If we take the Lorentz product of (10) with y, we obtain the necessary condition

$$\left| \langle x_1, y \rangle_{3,1} \right| = \left| \langle x_2, y \rangle_{3,1} \right|.$$

Note that if C_1 and C_2 intersect in two points, this condition means, that K intersects C_1 in the same angle as it intersects C_2, i.e., K is the angle bisecting circle of C_1 and C_2. The derived condition is equivalent to

$$\begin{aligned} &\langle x_1, x_1 + \sigma x_2\rangle_{3,1}^2 = \langle x_2, x_1 + \sigma x_2\rangle_{3,1}^2 \\ \Leftrightarrow\ &(1 + \sigma \langle x_1, x_2\rangle_{3,1})^2 = (\langle x_1, x_2\rangle_{3,1} + \sigma)^2 \\ \Leftrightarrow\ &\sigma^2 = 1 \quad \text{if } \langle x_1, x_2\rangle_{3,1} \neq 1. \end{aligned}$$

Thus, two possible solutions for y are given by

$$y_\pm = x_1 \pm x_2,$$

and one easily checks that indeed

$$\iota_\pm(x_1) = \mp x_2.$$

On the question of existence of the midcircles, we still have to check whether $[y_\pm]$ lies outside the Möbius quadric, and thus indeed describes a (real) circle.

First, note that $[y_+]$ and $[y_-]$ are polar,

$$\langle y_+, y_-\rangle_{3,1} = \langle x_1 + x_2, x_1 - x_2\rangle_{3,1} = 0,$$

and thus cannot both lie inside the Möbius quadric. Therefore, there always exists at least one midcircle. Now

$$\langle y_\pm, y_\pm\rangle_{3,1} = 2(1 \pm \langle x_1, x_2\rangle_{3,1}).$$

On the other hand, the signature of the line $[x_1] \vee [x_2]$ is determined by the sign of

$$\begin{aligned} \det\begin{pmatrix} 1 & \langle x_1,x_2\rangle_{3,1} \\ \langle x_1,x_2\rangle_{3,1} & 1 \end{pmatrix} &= 1 - \langle x_1, x_2\rangle_{3,1}^2 = (1 + \langle x_1, x_2\rangle_{3,1})(1 - \langle x_1, x_2\rangle_{3,1}) \\ &= \frac{1}{4}\langle y_+, y_+\rangle_{3,1}\langle y_+, y_+\rangle_{3,1}. \end{aligned}$$

Thus, two midcircles exist if and only if the signature of $[x_1] \vee [x_2]$ is equal to `(++)`, or equivalently, if the two circles C_1 and C_2 intersect in two points. We summarize our findings in the following proposition.

Proposition 10.7. *Let C_1 and C_2 be two circles (or lines) in $\widehat{\mathbb{R}^2}$. Let $[x_1], [x_2] \in \mathbb{R}\mathrm{P}^4$ be their Möbius lifts, respectively, satisfying*

$$\langle x_1, x_1\rangle_{3,1} = \langle x_2, x_2\rangle_{3,1} = 1.$$

- *If C_1 and C_2 intersect in two points, they have exactly two midcircles, which are the two angle bisectors of C_1 and C_2, and given by $[x_1 + x_2]$ and $[x_1 - x_2]$.*
- *If C_1 and C_2 do not intersect or touch, they have exactly one midcircle, given by either $[x_1 + x_2]$ or $[x_1 - x_2]$.*

Since we can exchange $x_i \to -x_i$, there is no projectively invariant way to distinguish the two points $[x_1+x_2]$ and $[x_1-x_2]$. Thus, there is no Möbius invariant way to distinguish the two midcircles of two (non-oriented) circles C_1 and C_2. However, if we consider two oriented circles, the midcircles can be distinguished in the following way.

Definition 10.4. Let C_1, C_2 be two oriented circles (or lines) in $\widehat{\mathbb{R}^2}$, and let K be a midcircle of C_1 and C_2

- If $\imath|_K C_1 : C_1 \to C_2$ preserves orientation, K is called *circle of similtude*, and its center is called *center of similtude* or *internal similtude center*.
- If $\imath|_K C_1 : C_1 \to C_2$ reverses orientation, K is called *circle of anti-similtude*, and its center is called *center of anti-similtude* or *external similtude center*.

For two circles C_1 and C_2 in $\mathbb{R}^2$, we can encode the orientation in the sign of the radius. Thus, let $c_1, c_2 \in \mathbb{R}^2$ be the two centers and $r_1, r_2 \in \mathbb{R}$ two two signed radii. The lift

$$x_i = \frac{1}{r_i}(c_i + (\|c_i\|^2 - r_i^2)e_\infty + e_0), \qquad i = 1, 2$$

satisfies $\langle x_i, x_i \rangle_{3,1} = 1$ and the two possible signs $x_i \leftrightarrow -x_i$ uniquely corresponds to two possibly signs $r_i \leftrightarrow -r_i$, and thus the two possible orientations of the circle C_i.

With this specific lift, the two solutions $x_1 \pm x_2$ for the midcircles can be distinguished depending on the combination of orientations of C_1 and C_2. One may check, that $x_1 + x_2$ always corresponds to the circle of similtude, and $x_1 - x_2$ to the circle of anti-similtude. Their existence can now be read off from the sign of

$$\begin{aligned}\langle y_\pm, y_\pm \rangle_{3,1} &= 2(1 \pm \langle x_1, x_2 \rangle_{3,1}) = 2\left(1 \pm \frac{1}{2r_1r_2}(r_1^2 + r_2^2 - \|c_1 - c_2\|^2)\right) \\ &= \pm\frac{1}{r_1r_2}\left((r_1 \pm r_2)^2 - \|c_1 - c_2\|^2\right)\end{aligned}$$

and we may now derive explicit formulas for the center $c_\pm$ and radius $r_\pm$ of these two midcircles.

$$\begin{aligned}y_\pm = x_1 \pm x_2 &= \frac{c_1}{r_1} \pm \frac{c_2}{r_2} + \left(\frac{\|c_1\|^2 - r_1^2}{r_1} \pm \frac{\|c_2\|^2 - r_2^2}{r_2}\right)e_\infty + \left(\frac{1}{r_1} \pm \frac{1}{r_2}\right)e_0 \\ &\sim \frac{r_2c_1 \pm r_1c_2}{r_2 \pm r_1} + \frac{r_1r_2}{r_2 \pm r_1}\left(\frac{\|c_1\|^2 - r_1^2}{r_1} \pm \frac{\|c_2\|^2 - r_2^2}{r_2}\right)e_\infty + e_0 \\ &= c_\pm + (\|c_\pm\|^2 - r_\pm^2) + e_0\end{aligned}$$

by comparing coefficients (or using the formulas from Proposition 6.7):

$$c_\pm = \frac{r_2c_1 \pm r_1c_2}{r_2 \pm r_1}, \qquad r_\pm^2 = \pm r_1r_2\left(1 - \frac{\|c_1 - c_2\|^2}{(r_1 \pm r_2)^2}\right)$$

We summarize again:

Proposition 10.8. *Let C_1 and C_2 be two oriented circles in $\mathbb{R}^2$ with centers $c_1, c_2 \in \mathbb{R}^2$ and signed radii $r_1, r_2 \in \mathbb{R}$.*

- *The (unique) circle of similtude exists if and only if*

$$r_1r_2\left(1 - \frac{|c_1 - c_2|^2}{(r_1 + r_2)^2}\right) > 0,$$

and its center and radius are given by

$$c_+ = \frac{r_1 c_2 + r_2 c_1}{r_1 + r_2}, \qquad r_+ = \sqrt{r_1 r_2 \left(1 - \frac{|c_1 - c_2|^2}{(r_1 + r_2)^2}\right)}.$$

- *The (unique) circle of anti-similtude exists if and only if*

$$r_1 r_2 \left(\frac{|c_1 - c_2|^2}{(r_1 - r_2)^2} - 1\right) > 0,$$

and its center and radius are given by

$$c_- = \frac{r_1 c_2 - r_2 c_1}{r_1 - r_2}, \qquad r_- = \sqrt{r_1 r_2 \left(\frac{|c_1 - c_2|^2}{(r_1 - r_2)^2} - 1\right)}.$$

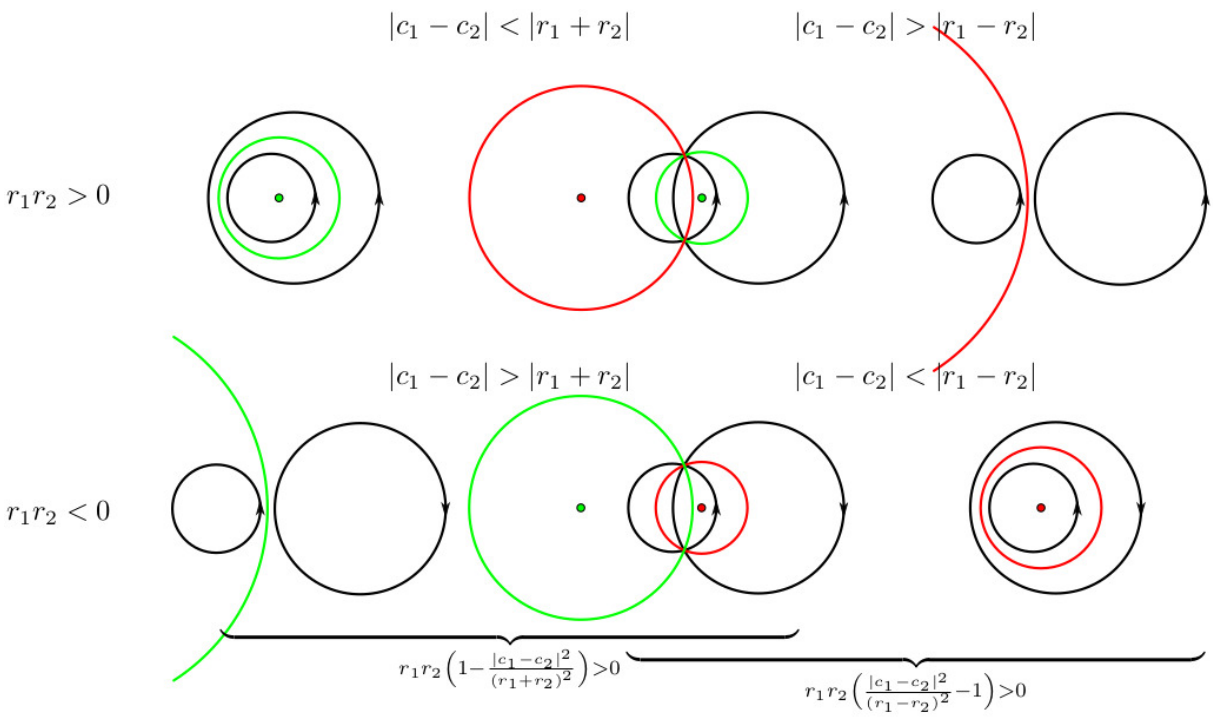


Figure 29. Midcircles of oriented circles. Circles of similtude (green) and circle of anti-similtude (red).

10.5 Discrete envelopes and orthogonal trajectories from midcircles

Consider two circles C_1 and C_2 with the same orientation in a pencil of circles through two points. Then in the limit $C_1 \to C_2$ the circle of similtude K_+ goes to $C_1 = C_2$, and the circle of anti-similtude goes to a circle orthogonal to $C_1 = C_2$. Recalling the symmetries of smooth envelopes in Remark 10.1 and of smooth orthogonal trajectories in Remark 10.2, this motivates the following alternative definitions for discrete envelopes and discrete orthogonal trajectories.

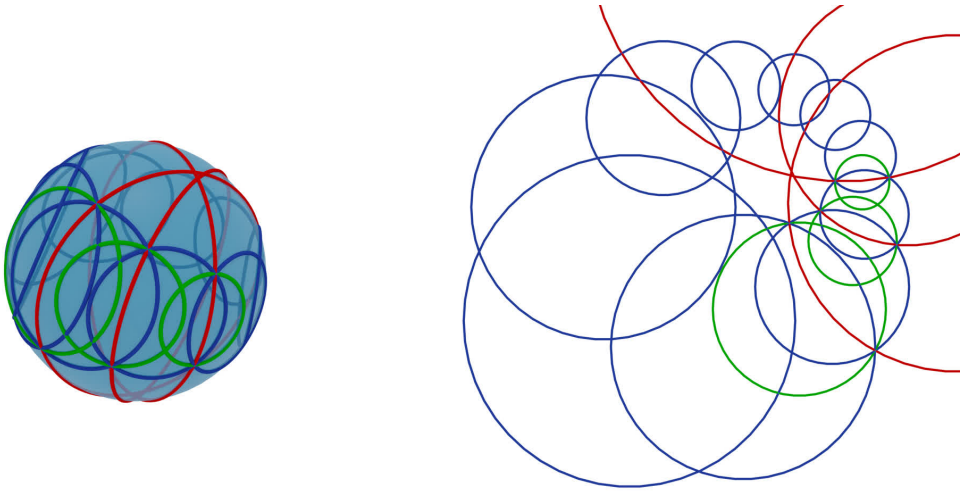


Figure 30. Midcricles of a discrete one-parameter family of circles.

Definition 10.5. Let $I \subset \mathbb{Z}$ be a discrete interval, and let $C : I \to \widehat{\mathbb{R}^2}$ be a discrete one-parameter family of oriented circles. Then $\gamma : I \to \widehat{\mathbb{R}^2}$ is called

(i) a *discrete envelope* of C if the circle of anti-similtude $K_-(n)$ of $C(n)$ and $C(n+1)$ exists for all $n, n+1 \in I$ and
$$\gamma_{n+1} = \iota_{K_-}(\gamma_n).$$

(ii) a *discrete orthogonal trajectory* of C if the circle of similtude $K_+(n)$ of $C(n)$ and $C(n+1)$ exists for all $n, n+1 \in I$ and
$$\gamma_{n+1} = \iota_{K_+}(\gamma_n).$$

Remark 10.3.

(i) The special case of discrete families of lines also leads to alternative definitions for discrete evolutes and involutes.

(ii) The discrete envelope and orthogonal trajectories defined in this way both have one degree of freedom. It can be fixed by one initial point on one of the circles.

(iii) The edge (γ_{n+1}, γ_n) of an orthogonal trajectory is orthogonal to the circle of similtude $K_+(n)$.

(iv) If we introduce coordinates
$$\gamma_n = c_n + r_n e^{i\theta_n}, \qquad \tau_n = \tan\frac{\theta_n}{2}$$
the discrete envelope satisfies the equation
$$(\Delta r_n - \Delta(c_1)_n)\tau_n\tau_{n+1} - \Delta(c_2)_n(\tau_n + \tau_{n+1}) + \Delta r_n + \Delta(c_1)_n = 0,$$
which is a discrete analogue of the quadratic equation satisfied by the smooth envelope. Similarly, the discrete orthogonal trajectory satisfies the equation
$$(r_n + r_{n+1})\Delta\tau_n = \Delta(c_1)_n(\tau_n + \tau_{n+1}) - \Delta(c_2)_n(1 - \tau_n\tau_{n+1}).$$
which is a discrete analogue of the Riccati equation satisfied by the smooth orthogonal trajectory.

10.6 Discrete tractrix and Darboux transform

We noted that in the smooth case a tractrix is a special case of an orthogonal trajectory of a one-parameter family of circles with constant radius.

Definition 10.6. Let $\gamma : \mathbb{Z} \supset I \to \mathbb{R}^2$ be a discrete curve. A discrete curve $\hat{\gamma} : I \to \mathbb{R}^2$ is called a *(discrete) tractrix* of γ if $\hat{\gamma}$ is a discrete orthogonal trajectory of discrete one-parameter family of circles with centers γ and constant radii.

We define discrete Darboux transforms in the following way.

Definition 10.7. Two discrete curves $\gamma, \tilde{\gamma} : \mathbb{Z} \supset I \to \mathbb{R}^2$ are called *(discrete) Darboux transforms* of each other if

$$\|\gamma_{n+1} - \gamma_n\| = \|\tilde{\gamma}_{n+1} - \tilde{\gamma}_n\|, \quad \text{and} \quad \|\tilde{\gamma}_n - \gamma_n\| = \text{const.} \tag{11}$$

and $\tilde{\gamma}$ is not a parallel translation of γ.

Given $\gamma_n, \gamma_{n+1}, \tilde{\gamma}_n$ there are two solutions $\tilde{\gamma}_{n+1}$ satisfying the conditions 11, leading to a parallelogram and a parallelogram folded in one diagonal, which is also called *Darboux butterfly.* The parallelgram is excluded by the condition that $\tilde{\gamma}$ is not a parallel translation of γ. Thus, the elementary quadrilaterals of the Darboux transformation consists of Darboux butterflies.

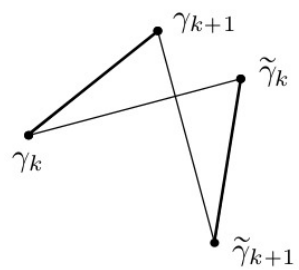


Figure 31. A Darboux butterfly.

Before we continue we give the following geometric characterization of Darboux butterflies.

Lemma 10.9. *A quadrilateral $x_1, x_2, x_3, x_4 \in \mathbb{R}^2$ is a Darboux butterfly if and only if the following three conditions are satisfied*

(i) The two diagonals $x_1 \vee x_3$ and $x_2 \vee x_4$ are parallel, or equivalently, three (and therefore all) of the edge midpoints $\frac{1}{2}(x_i + x_{i+1}), i = 1, 2, 3, 4$ are collinear.

(ii) $\|x_1 - x_2\| = \|x_3 - x_4\|$.

(iii) The two edges $x_1 \vee x_2$ and $x_3 \vee x_4$ are not parallel.

Proof. To see the equivalence of the two conditions in (i), note that the midpoinst of a quadrilateral form a parallelogram whose sides are parallel to the diagonals.

($\Leftarrow$) Clearly, a Darboux butterflies satisfies all three conditions.

($\Rightarrow$) Consider two parallel lines as the diagonals ℓ_1, ℓ_2 of the quadrilateral and two points $x_1 \in \ell_1, x_2 \in \ell_2$. Then up two translation there are only two choices for the points $x_3 \in \ell_1, x_4 \in \ell_4$, such that (ii) is satisfied. Only one these satisfies (iii). By symmetry, this choice leads to $\|x_1 - x_4\| = \|x_2 - x_3\|$.

□

With this observation we obtain a discrete analogue of Theorem 10.6.

Theorem 10.10. *Let $\gamma : \mathbb{Z} \supset I \to \mathbb{R}^2$ be a discrete curve. Then the following claims are equivalent:*

(i) $\tilde{\gamma}$ is a discrete Darboux transform of γ

(ii) $\hat{\gamma} := \frac{1}{2}(\gamma + \tilde{\gamma})$ is a disrcete tractrix of γ (and of $\tilde{\gamma}$).

Proof.

($\Leftarrow$) $\|\tilde{\gamma}_n - \gamma_n\| = \|\tilde{\gamma}_{n+1} - \gamma_{n+1}\|$ implies $\|\hat{\gamma}_n - \gamma_n\| = \|\hat{\gamma}_{n+1} - \gamma_{n+1}\| =: r$. Furthermore, $\frac{1}{2}(\gamma_n + \gamma_{n+1})$ is the center of similtude of the two circles with centers γ_n and γ_{n+1} and radius r. The three midpoints $\frac{1}{2}(\gamma_n + \gamma_{n+1}), \hat{\gamma}_n, \hat{\gamma}_{n+1}$ lie on a line. Thus, $\hat{\gamma}_n$ and $\hat{\gamma}_{n+1}$ are symmetric with respect to the circle of similtude.

($\Rightarrow$) $\|\hat{\gamma}_n - \gamma_n\| = \|\hat{\gamma}_{n+1} - \gamma_{n+1}\|$ implies $\|\tilde{\gamma}_n - \gamma_n\| = \|\tilde{\gamma}_{n+1} - \gamma_{n+1}\|$, while the two corresponding edges cannot be parallel. The three midpoints $\frac{1}{2}(\gamma_n + \gamma_{n+1}), \hat{\gamma}_n, \hat{\gamma}_{n+1}$ of the quadrilateral $\gamma_n, \gamma_{n+1}, \tilde{\gamma}_{n+1}, \tilde{\gamma}_n$ lie on a line. Thus, by Lemma 10.9, this quadrilateral is a Darboux butterfly.

□

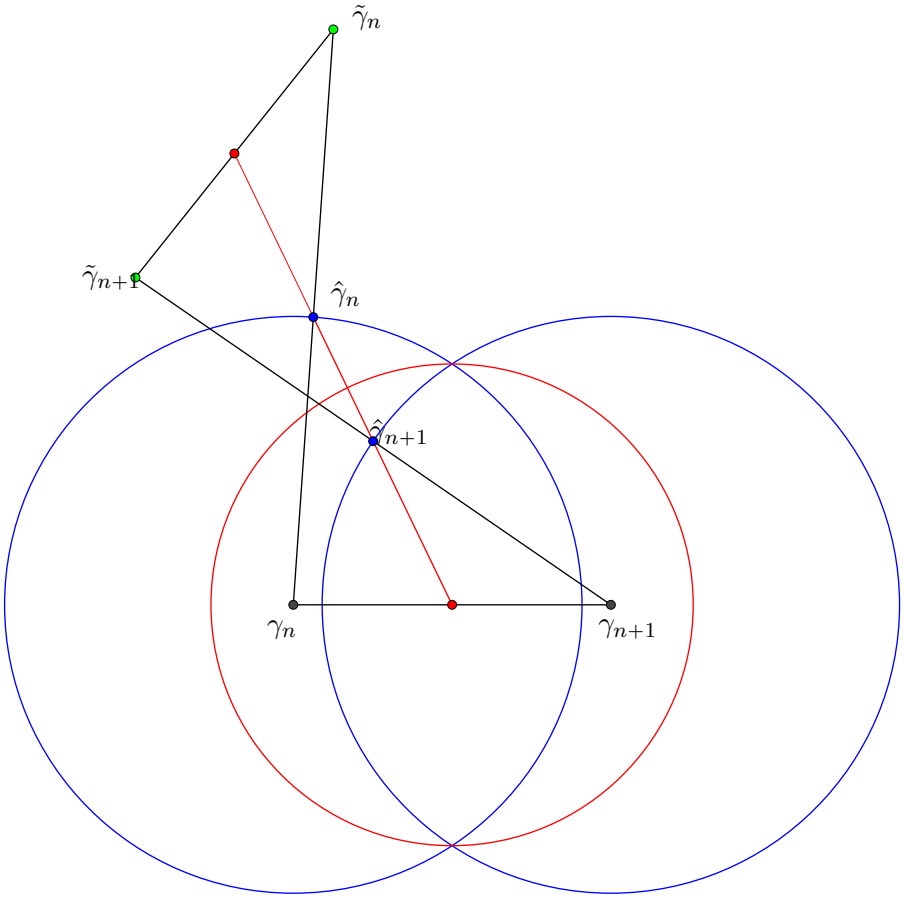


Figure 32. Discrete tractrix (blue points) as discrete orthogonal trajectory and relation to discrete Darboux transform (green points).

11 Surfaces and curvature line parametrizations

11.1 Parametrized surfaces

Definition 11.1. Let $U \subset \mathbb{R}^2$ be a open set. Then a smooth map

$$f : U \to \mathbb{R}^n, \qquad (u, v) \mapsto f(u, v)$$

is called a *(smooth parametrized) surface (patch)* in $\mathbb{R}^n$.

The curves

$$u \mapsto f(u, v), \qquad v \mapsto f(u, v)$$

are called *parameter lines* of f.

We usually denote the two parameters by u and v. and the partial derivatives with respect to u and v by

$$f_u := \frac{\partial f}{\partial u}, \quad f_v := \frac{\partial f}{\partial v}.$$

Regularity is defined for surface patches by the linear independence of the first partial derivatives.

Definition 11.2. A surface $f : U \to \mathbb{R}^n$ is called *regular* if $f_u(u,v)$ and $f_v(u,v)$ are linearly independent at every point $(u,v) \in U$.

For a regular surface the parameter lines are regular curves, and the tangent plane is well-defined at every point. It is the plane that best approximates the surface patch at some point up to first order.

Definition 11.3. Let $f : U \to \mathbb{R}^n$ be a regular surface. Then the plane

$$Tf(u,v) := \{f(u,v) + \alpha f_u(u,v) + \beta f_v(u,v) \mid \alpha, \beta \in \mathbb{R}\}$$

is called the *tangent plane* of f at $(u,v) \in U$.

11.2 Surfaces in projective geometry

Similar to curves, we can lift a surfaces $f : U \to \mathbb{R}^n$ to the projective space $\mathbb{R}\mathrm{P}^n$ by

$$[\hat{f}] : U \to \mathbb{R}\mathrm{P}^n, \qquad \hat{f}(u,v) := \begin{pmatrix} f(u,v) \\ 1 \end{pmatrix}.$$

If f is regular, the partial derivatives

$$\hat{f}_u(u,v) = \begin{pmatrix} f_u(u,v) \\ 0 \end{pmatrix}, \qquad \hat{f}_v(u,v) = \begin{pmatrix} f_v(u,v) \\ 0 \end{pmatrix},$$

describe points at infinity on the lift of the tangent plane

$$\begin{aligned} Tf(u,v) &= \left\{\alpha_1 \hat{f}(u,v) + \alpha_2 \hat{f}_u(u,v) + \alpha_3 f_v(u,v) \;\middle|\; \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\right\} \\ &= [f(u,v)] \vee [f_u(u,v)] \vee [f_v(u,v)]. \end{aligned}$$

Generally, we define projective surfaces in the following way.

Definition 11.4. Let $U \subset \mathbb{R}^2$ be a open set and $\hat{f} : U \to \mathbb{R}^{n+1}$ a smooth map Then

$$[\hat{f}] : U \to \mathbb{R}\mathrm{P}^n, \qquad (u,v) \mapsto [\hat{f}(u,v)]$$

is called a *(smooth parametrized) surface (patch)* in $\mathbb{R}\mathrm{P}^n$.

The curves

$$u \mapsto f(u,v), \qquad v \mapsto f(u,v)$$

are called *parameter lines* of $[\hat{f}]$.

Consider a surface $[\hat{f}] : U \to \mathbb{R}\mathrm{P}^n$ and a smooth function

$$\lambda : U \to \mathbb{R}\backslash\{0\}.$$

Then $\hat{f}$ and $\tilde{f} := \lambda \hat{f}$ define the same surface in $\mathbb{R}\mathrm{P}^n$,

$$[\hat{f}(u,v)] = [\tilde{f}(u,v)] \qquad \text{for all } (u,v) \in U.$$

Similar to the considerations for curves, the points described by the first partial derivatives may change, but the span

$$[\hat{f}(u,v)] \vee [\hat{f}_u(u,v)] \vee [\hat{f}_v(u,v)] = [\tilde{f}(u,v)] \vee [\tilde{f}_u(u,v)] \vee [\tilde{f}_v(u,v)]$$

remains the same. Thus, the following definition of regularity for surfaces in $\mathbb{R}\mathrm{P}^n$ is independent of the choice of representative vectors. Furthermore, in affine coordinates, it coincides with the corresponding definition for surfaces in $\mathbb{R}^n$.

Definition 11.5. A surface $[\hat{f}] : U \to \mathbb{R}\mathrm{P}^n$ is called *regular* if $[\hat{f}(u,v)]$, $[\hat{f}_u(u,v)]$, $[\hat{f}_v(u,v)]$ span a plane, or equivalently, if $\hat{f}(u,v)$, $\hat{f}_u(u,v)$, $\hat{f}_v(u,v)$ are linearly independent.

The same holds for the following definition of the tangent planes for surfaces in $\mathbb{R}\mathrm{P}^n$.

Definition 11.6. Let $[\hat{f}] : U \to \mathbb{R}\mathrm{P}^n$ be a regular surface. Then the plane

$$T[\hat{f}](u,v) := [\hat{f}(u,v)] \vee [\hat{f}_u(u,v)] \vee [\hat{f}_v(u,v)]$$

is called the *tangent plane* of $[\hat{f}]$ at $(u,v) \in U$.

Similar to the considerations for curves, one finds that the introduced notions are also invariant under reparametrization and under projective transformations. We summarize in the following proposition.

Proposition 11.1. *For a surface $[\hat{f}] : U \to \mathbb{R}\mathrm{P}^n$, regularity, and the tangent plane are invariant under*

(i) a change of representative vectors

$$\hat{f}(u,v) \to \lambda(u,v)\hat{f}(u,v)$$

with a smooth non-vanishing function λ.

(ii) reparametrization

$$\hat{f}(u,v) \to \hat{f} \circ \varphi(\tilde{u},\tilde{v})$$

with a smooth bijective map φ.

(iii) projective transformations

$$\hat{f}(u,v) \to F\hat{f}(u,v)$$

with $F \in \mathrm{GL}(n+1,\mathbb{R})$.

11.3 Dual representation of surfaces

Instead of describing a surface as a two-parameter family of points, we can equivalently describe it as the envelope of its two-parameter family of tangent planes. In particular, for a surface in $\mathbb{R}^3$, the tangent planes can be described in terms of a normal field.

Definition 11.7. Let $f : U \to \mathbb{R}^3$ be a regular surface. Then a smooth map

$$n : U \to \mathbb{R}^3 \backslash \{0\}$$

is called a *normal field* of f if

$$\begin{aligned} n \cdot f_u &= 0, \\ n \cdot f_v &= 0. \end{aligned}$$

The tangent plane of a surface f in $\mathbb{R}^3$ can be described in terms of a normal field

$$Tf(u,v) = \left\{ x \in \mathbb{R}^3 \;\middle|\; n(u,v) \cdot (x - f(u,v)) = n(u,v) \cdot (x + h(u,v) = 0 \right\}$$

and some function $h(u,v) = -n(u,v)\cdot f(u,v)$. Thus, the tangent planes of f are described by the tuple (n,h), which is unique up to a common scalar multiple, and determined by the equations

$$\begin{aligned} n\cdot f_u &= 0,\\ n\cdot f_v &= 0,\\ n\cdot f + h &= 0. \end{aligned} \tag{12}$$

Differentiating the last equation with respect to u and v, respectively, we find that (12) is equivalent to

$$\begin{aligned} f\cdot n_u + h_u &= 0,\\ f\cdot n_v + h_v &= 0,\\ f\cdot n + h &= 0. \end{aligned} \tag{13}$$

Note that if we consider the lifts

$$\begin{aligned} \hat{f} &:= (f,1),\\ \hat{n} &:= (n,h) \end{aligned}$$

to homogeneous coordinates of $\mathbb{R}\mathrm{P}^3$ and $(\mathbb{R}\mathrm{P}^3)^*$, respectively, then equations (12) and (13) become the duality relations for tangent planes of the respective surfaces $[\hat{f}]$ and $[\hat{n}]$.

Definition 11.8. Let $[\hat{f}] : U \to \mathbb{R}\mathrm{P}^3$ be a regular surface. Then

$$[\hat{n}] := ([\hat{f}] \vee [\hat{f}_u] \vee [\hat{f}_v])^\star : U \to (\mathbb{R}\mathrm{P}^3)^*$$

is called the *dual surface* of f.

In homogeneous coordinates the dual surface is determined by the three linearly independent equations

$$\begin{aligned} \hat{n}\cdot \hat{f}_u &= 0,\\ \hat{n}\cdot \hat{f}_v &= 0,\\ \hat{n}\cdot \hat{f} &= 0, \end{aligned} \tag{14}$$

and satisfies

$$\begin{aligned} \hat{f}\cdot \hat{n}_u &= 0,\\ \hat{f}\cdot \hat{n}_v &= 0,\\ \hat{f}\cdot \hat{n} &= 0. \end{aligned} \tag{15}$$

These equations are completely symmetric in $\hat{f}$ and $\hat{n}$.

Proposition 11.2. *If the dual surface of a regular surface $[\hat{f}]$ in $\mathbb{R}\mathrm{P}^3$ is itself regular, then the dual surface of a the dual surface is $[\hat{f}]$.*

Remark 11.1. The primal surface is regular if it is locally not a curve. The dual surface is regular if the primal surface is locally not developable (see next section).

11.4 Ruled surfaces and developable surfaces

Definition 11.9. A *ruled surface* is a surface traced out by the movement of a straight line through space. The lines on the resulting surface are called *rulings*.

Example 11.1. A one-sheeted hyperboloid is a doubly ruled surface.

It can be described by connecting corresponding points of two parametrized curves. Given $a, b : I \to \mathbb{R}^3$, we obtain a parametrized ruled surface $f : I \times \mathbb{R} \to \mathbb{R}^3$

$$f(u,v) = (1-v)a(u) + vb(u) = a(u) + vr(u), \qquad \text{with } r(u) := b(u) - a(u).$$

Definition 11.10. A *developable surface* is the envelope of a one-parameter family of planes.

Proposition 11.3. *Every developable surface is a ruled surface.*

Proof. Consider a one-parameter family of planes

$$T(u) = \left\{ x \in \mathbb{R}^3 \;\middle|\; n(u) \cdot x + h(u) = 0 \right\}.$$

Then the envelope is the solution of the two equations

$$\begin{aligned} n \cdot x + h &= 0, \\ n_u \cdot x + h_u &= 0. \end{aligned}$$

For each u these are two linear equations, and thus the solution is a line. □

Proposition 11.4. *A ruled surfaces*

$$f(u,v) = a(u) + vr(u)$$

is a developable surface if and only if

$$\det(r, a_u, r_u) = 0$$

Proof. A ruled surface is the envelope of a one-parameter family of lines if and only if it has a fixed tangent plane along each ruling, or equivalently, if all tangent vectors of f

$$f_u = a_u + vr_u, \qquad f_v = r$$

along a ruling lie in one plane:

$$0 = \det(r, a_u, b_u) = \det(r, a_u, r_u).$$

□

Infinitesimally, this means that adjacent lines of the rulings intersect, and thus envelope a curve in space.

Proposition 11.5. *The rulings of a developable surface envelope a curve called the* ***line of striction****.*

For the developable surface $f(u,v)$

$$f(u,v) = a(u) + vr(u),$$

and the two functions $\alpha(u)$ and $\beta(u)$ given by

$$r_u(u) = \alpha(u)a_u(u) + \beta(u)r(u)$$

the line of striction is given by

$$s(u) = a(u) - \frac{1}{\alpha(u)} r(u).$$

Proof. For the rulings to be tangent lines of the curve $s(u)$ it must satisfy

$$s(u) = a(u) + \lambda(u)r(u)s_u = a_u + \lambda_u r + \lambda r_u = (1 + \alpha\lambda)a_u + (\lambda_u + \beta\lambda)r \sim r$$

and thus

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

□

11.5 Conjugate line parametrizations

We now study special parametrizations, in the sense that the parameter lines satisfy some geometric condition. We start with conjugate line parametrizations, which we first introduce for surfaces in $\mathbb{R}^3$. Conjugate line paramtrizations are geometrically characterized by the following condition: Along each parameter line of the surface, the tangent planes rotate around the tangent line in the other coordinate direction. Put differently: The tangent planes along one parameter line envelop a surface that is ruled by the tangent lines in the other coordinate direction.

Definition 11.11. Let $f : U \to \mathbb{R}^3$ be a regular surface, and $n : U \to \mathbb{R}^3$ a normal field of f. Then f is a called a *conjugate line parameterization* if one and hence all of the following equivalent conditions hold:

(i) $n_v \cdot f_u = 0$

(ii) $n_u \cdot f_v = 0$

(iii) $n \cdot f_{uv} = 0$

(iv) $f_{uv} \in \operatorname{span}(f_u, f_v)$

(v) $f_{uv} = \alpha f_u + \beta f_v$ for smooth functions $\alpha, \beta : U \to \mathbb{R}$

Proof. Taking the v-derivative of $n \cdot f_u = 0$ and the u-derivative of $n \cdot f_v = 0$, we obtain

$$\begin{aligned} n_v \cdot f_u &= n \cdot f_{uv} \\ n_u \cdot f_v &= n \cdot f_{vu} \end{aligned}$$

and since $f_{uv} = f_{vu}$ by the symmetry of second derivatives, conditions (i), (ii), and (iii) are equivalent.

Condition (iii) implies (iv) because (f_u, f_v) is a basis for the orthogonal subspace to n. This also means that the equation of condition (v) determines the functions α and β uniquely. In fact, by Cramer's rule,

$$\alpha = \frac{\det(n, f_{uv}\ f_v)}{\det(n, f_u\ f_v)}, \quad \beta = \frac{\det(n, f_u\ f_{uv})}{\det(n, f_u\ f_v)},$$

which also shows that α and β are smooth because f is. Finally, condition (v) clearly implies (iii) and (iv). □

Conditions (iv) and (v) of Definition 11.11 do not mention the normal field n. We may use them to define conjugate line parametrizations in $\mathbb{R}^n$:

Definition 11.12. A regular surface $f : U \to \mathbb{R}^n$ is called a *conjugate line parameterization* if it satisfies one and hence both equivalent conditions (iv) and (v) of Definition 11.11.

The definition for conjugate line parametrizations translates as follows to surfaces in $\mathbb{R}\mathrm{P}^n$:

Proposition 11.6. *Let $f : U \to \mathbb{R}^n$ be a regular surface. Let*

$$\hat{f} := \lambda \cdot (f, 1) : U \to \mathbb{R}^{n+1}$$

be an arbitrary lift to homogeneous coordinates with a smooth function $\lambda : U \to \mathbb{R}\backslash\{0\}$.

Then f *is a conjugate line parametrization if and only if* $\hat{f}$ *satisfies*

$$\hat{f}_{uv} = \alpha \hat{f}_u + \beta \hat{f}_v + \gamma \hat{f} \tag{16}$$

with some smooth functions α, β, γ.

Equation (16) states the linear dependence of four representative vectors, or equivalently that four points lie in a plane. While the four points are not projectively well-defined (the points defined by the derivatives are not invariant under scaling $\hat{f}$) this property is.

Definition 11.13. Let $[\hat{f}] : U \to \mathbb{R}\mathrm{P}^n$ be a regular surface. Then $[\hat{f}]$ is called a *conjugate line parametrization* if the four points $[\hat{f}]$, $[\hat{f}_u]$, $[\hat{f}_v]$, $[\hat{f}_{uv}]$ lie in a plane for every $(u, v) \in U$.

We have seen that this property is projectively well-defined. Furthermore, it is a property of the coordinate lines. Thus, it is invariant under reparametrization of the surface along the coordinate lines. Finally, it is also invariant under applying a projective transformation to the surface. We summarize these properties in the following proposition.

Proposition 11.7. *A regular surface* $[\hat{f}] : U \to \mathbb{R}\mathrm{P}^n$ *being a conjugate line parametrization is invariant under*

(i) a change of representative vectors

$$\hat{f}(u, v) \to \lambda(u, v)\hat{f}(u, v)$$

with a smooth non-vanishing function λ.

(ii) reparametrization along the coordinate lines

$$\hat{f}(u, v) \to \hat{f}(\varphi(\tilde{u}), \chi(\tilde{v}))$$

with two smooth bijective functions φ, χ.

(iii) projective transformations

$$\hat{f}(u, v) \to F\hat{f}(u, v)$$

with $F \in \mathrm{GL}(n + 1, \mathbb{R})$.

For surfaces in $\mathbb{R}\mathrm{P}^3$ the property of being a conjugate line parametrization is also invariant under dualization.

Proposition 11.8. *A regular surface* $[\hat{f}] : \mathbb{R}^2 \supset U \to \mathbb{R}\mathrm{P}^3$ *is a conjugate line parametrization if and only if its dual surface* $[\hat{n}] : U \to (\mathbb{R}\mathrm{P}^3)^*$ *is a conjugate line parametrization.*

Proof. $[\hat{f}]$ is a conjugate line parametrization if $\hat{f}$ satisfies an equation of the form (16), which is equivalent to

$$\hat{f}_{uv} \cdot \hat{n} = 0.$$

From equations (14), or equivalently, equations (15), we find that this is equivalent to either of the three equations

$$\begin{aligned} \hat{f}_u \cdot \hat{n}_v &= 0, \\ \hat{f}_v \cdot \hat{n}_u &= 0, \\ \hat{f} \cdot \hat{n}_{uv} &= 0, \end{aligned} \tag{17}$$

and thus in turn to

$$\hat{n}_{uv} = \tilde{\alpha}\hat{n}_u + \tilde{\beta}\hat{n}_v + \tilde{\gamma}\hat{n},$$

□

Remark 11.2. The first two equations of (17) state, respectively, that

$$[\hat{f}] \vee [\hat{f}_u] = ([\hat{n}] \vee [\hat{n}_v])^\star,$$
$$[\hat{f}] \vee [\hat{f}_v] = ([\hat{n}] \vee [\hat{n}_u])^\star.$$

which capture the geometric description of conjugate line parametrizations given in the beginning of the section.

11.6 Curvature line parametrizations

Definition 11.14. Let

$$f : \mathbb{R}^2 \supset U \to \mathbb{R}^3$$

be a smooth regular parametrized surface patch.

(i) f is called *orthogonal* if

$$\langle f_u, f_v \rangle = 0$$

(ii) f is called *curvature line parametrization* if it is orthogonal and conjugate, i.e.,

$$\langle f_u, f_v \rangle = 0, \quad \text{and} \quad f_{uv} = \alpha f_u + \beta f_v.$$

Proposition 11.9. *The property of a parametrization to be orthogonal is Möbius invariant.*

Proof. Möbius transformations are conformal, i.e., preserve angles. □

Conjugate parametrizations, on the other hand, are not Möbius invariant. Are curvature line parametrizations?

Proposition 11.10. *Let $f : \mathbb{R}^2 \supset U \to \mathbb{R}^3$ be a parametrized surface and*

$$\hat{f} := f + e_0 + \|f\|^2 e_\infty$$

its lift to the Möbius quadric. Then f is a curvature line parametrization if and only if $[\hat{f}]$ is a conjugate parametrization.

Proof. For the derivatives of the lift we obtain

$$\begin{aligned}
\hat{f}_u &= f_u + 2\langle f, f_u \rangle e_\infty, \\
\hat{f}_v &= f_v + 2\langle f, f_v \rangle e_\infty, \\
\hat{f}_{uv} &= f_{uv} + 2\left(\langle f, f_{uv} \rangle + \langle f_u, f_v \rangle\right) e_\infty.
\end{aligned}$$

Let $\hat{f}$ be a curvature line parametrization. Then

$$\hat{f}_{uv} = f_{uv} + 2\langle f, f_{uv} \rangle e_\infty = \alpha f_u + \beta f_v + 2\left(\alpha \langle f, f_u \rangle + \beta \langle f, f_v \rangle\right) e_\infty = \alpha \hat{f}_u + \beta \hat{f}_v.$$

The reverse direction is shown similarly. □

Corollary 11.11. *Curvature line parametrizations are Möbius invariant.*

11.7 Focal surfaces and principal curvature spheres

Let $f : \mathbb{R}^2 \supset U \to \mathbb{R}^3$ be a regular parametrized surface with unit normal field

$$n(u,v) := \frac{f_u \times f_v}{\|f_u \times f_v\|}$$

Lemma 11.12. *If f is orthogonal and conjugate, then*

$$n_u = -\kappa_1 f_u, \qquad n_v = -\kappa_2 f_v$$

with two smooth functions $\kappa_1, \kappa_2 : U \to \mathbb{R}$.

Proof. Then

$$n_u = \alpha f_u + \beta f_v$$

with two functions α, β. Since f is conjugate and orthogonal, we obtain

$$0 = n_u \cdot f_v = \alpha f_u \cdot f_v + \beta f_v \cdot f_v = \beta f_v \cdot f_v$$

and thus, $\beta = 0$. Similar for n_v. □

The two values, κ_1 and κ_2 are called the *principal curvatures* of the surface f. Points with $\kappa_1 = \kappa_2$ are called *umbilic* points. Away from umbilic points the reverse is also true.

Lemma 11.13. *If*

$$n_u = -\kappa_1 f_u, \qquad n_v = -\kappa_2 f_v$$

with two smooth functions $\kappa_1, \kappa_2 : U \to \mathbb{R}$, then at points with $\kappa_1(u,v) \neq \kappa_2(u,v)$, the parametrization f is orthogonal and conjugate.

Proof. From $n_u \cdot f_v = n_v \cdot f_u$ we obtain

$$-\kappa_1 f_u \cdot f_v = n_u \cdot f_v = n_v \cdot f_u = -\kappa_2 f_v \cdot f_u.$$

Thus, if $\kappa_1 \neq \kappa_2$ this implies $f_u \cdot f_v = 0$, and further $n_u \cdot f_v = 0$. □

We can now characterize curvature line parametrizations by their normals.

Definition 11.15. Let $f : \mathbb{R}^2 \supset U \to \mathbb{R}^3$ be a regular parametrized surface. Then the two-parameter family of lines

$$\ell : U \to \mathrm{Lines}(\mathbb{R}^3), \qquad (u,v) \mapsto \ell(u,v) = \{f(u,v) + \lambda n(u,v) \mid \lambda \in \mathbb{R}\}$$

is called the *normal congruence* of the surface f.

Proposition 11.14. *Let $f : \mathbb{R}^2 \supset U \to \mathbb{R}^3$ be a regular parametrized surface. Then f is a curvature line parametrization if and only if the ruled surfaces*

$$u \mapsto \ell(u, v = const), \qquad v \mapsto \ell(u = const, v)$$

in the normal congruence are developable.

Proof. Since $n \cdot n = 1$, we have

$$n_u = \alpha f_u + \beta f_v$$

Now

$$0 = \det(n, f_u, n_u) = \beta \det(n, f_u, f_v)$$

if and only if $\beta = 0$. □

Thus, in particular, each of these developable surfaces in the normal congruence has a line of striction. For a developable surfaces in u-direction

$$u \mapsto \ell(u, v = \text{const})$$

it is given by

$$u \mapsto f(u,v) + \frac{1}{\kappa_1(u,v)} n(u,v)$$

Together these lines of striction in u-direction form a surface.

Definition 11.16. Let $f : \mathbb{R}^2 \supset U \to \mathbb{R}^3$ be a curvature line parametrization. Then the two surfaces given by

$$f^{(1)}(u,v) := f(u,v) + \frac{1}{\kappa_1(u,v)} n(u,v), \qquad f^{(2)}(u,v) := f(u,v) + \frac{1}{\kappa_1(u,v)} n(u,v)$$

are called the *focal surfaces* of the surface f.

Note that the focal surfaces are a generalization of the evolute of a plane curve to surfaces. Correspondingly, they each form the centers of a two-parameter family of sphere, called *curvature spheres*.

Definition 11.17. Let $f : \mathbb{R}^2 \supset U \to \mathbb{R}^3$ be a curvature line parametrization. Then the two two-parameter families of spheres $S^{(i)}$ with centers f^i and radii $\frac{1}{|\kappa_i|}$ are called *curvature spheres* of the surface f.

Remark 11.3. The (unique) midsphere of the two (touching) curvature spheres $S^{(1)}$ and $S^{(2)}$ at any given point $(u,v) \in U$ is the *mean curvature sphere* of the surface f.

In the Möbius lift the curvature spheres are represented by the points

$$\hat{s}^{(i)} := f^{(i)} + \left(\left\| f^{(i)} \right\|^2 - \frac{1}{\kappa_i^2} \right) e_\infty + e_0 = f + \frac{1}{\kappa_i} n + \left(\|f\|^2 + \frac{2}{\kappa_i} \langle f, n \rangle \right) e_\infty + e_0$$

Proposition 11.15. *Let $f : \mathbb{R}^2 \supset U \to \mathbb{R}^3$ be a curvature line parametrization, and*

$$\hat{f} = f + \|f\|^2 e_\infty + e_0$$

its Möbius lift. Then the Möbius lift $[\hat{s}^{(1)}]$ of the curvature spheres $S^{(1)}$ is given by

$$[\hat{s}^{(1)}] = \left([\hat{f}] \vee [\hat{f}_u] \vee [f_v] \vee [f_{uu}] \right)^\perp,$$

and, similarly, the Möbius lift $[\hat{s}^{(2)}]$ of the curvature spheres $S^{(2)}$ is given by

$$[\hat{s}^{(2)}] = \left([\hat{f}] \vee [\hat{f}_u] \vee [f_v] \vee [f_{vv}] \right)^\perp.$$

Proof. One easily checks

$$\left\langle \hat{s}^{(1)}, \hat{f} \right\rangle_{4,1} = \left\langle \hat{s}^{(1)}, \hat{f}_u \right\rangle_{4,1} = \left\langle \hat{s}^{(1)}, \hat{f}_v \right\rangle_{4,1} = 0.$$

And the Lorentz product with the second derivative

$$\hat{f}_{uu} = f_{uu} + 2(\|f_u\|^2 + \langle f, f_{uu} \rangle) e_\infty$$

is given by

$$\left\langle \hat{s}^{(1)}, \hat{f}_{uu} \right\rangle_{4,1} = \frac{1}{\kappa_1} \langle n, f_{uu} \rangle - \|f_u\|^2 = 0,$$

since

$$\langle n, f_{uu} \rangle = -\langle n_u, f_u \rangle = \kappa_1 \|f_u\|^2.$$

□

11.8 Channel surfaces and Dupin cyclides

Definition 11.18. The surfaces given by the envelope of a one-parameter family of spheres is called a *channel surface*.

Consider a one-parameter family of spheres

$$S : \mathbb{R} \supset I \to \mathbb{R}^3, \qquad S(u) := \left\{ x \in \mathbb{R}^3 \;\middle|\; F(x,u) := \|c(u) - x\|^2 - r(u)^2 = 0 \right\}$$

Then an envelope surface

$$f : \mathbb{R}^2 \supset U = I \times J \to \mathbb{R}^3$$

is given by

$$\begin{aligned} F(f(u,v),u) &= 0, \\ F_u(f(u,v),u) &= 2\langle c_u, c - f\rangle - 2r_u = 0. \end{aligned}$$

The first equation defines a sphere, while the second defines a plane. Thus, their intersection is given by a circle

$$C : I \to \mathbb{R}^3, \qquad C(u) = \left\{ x \in \mathbb{R}^3 \;\middle|\; F(x,u) = F_u(x,u) = 0 \right\}$$

and for each $u \in I$ the entire v parameter line lies on $C(u)$:

$$f(u,v) \in C(u) \quad \text{for all } v \in J.$$

The u parameter lines are not uniquely defined by the two envelope equations. To obtain unique u parameter lines we can add the condition of orthogonality

$$\langle f_u, f_v \rangle = 0.$$

This leads to a curvature line parametrization of the channel surface. Indeed, by symmetry note that all normal lines along the circular v parameter lines must go through one point on the axis of the circle $C(u)$, and thus

$$n_v \sim f_v.$$

Together with the orthogonality this implies

$$\langle n_v, f_u \rangle = 0.$$

In particular, that all normal lines along the v parameter lines go through a common point implies that the focal surfaces $f^{(2)}$ of the corresponding direction degenerates to a curve:

$$f_v^{(2)} = 0$$

In fact, this property characterizes channel surfaces.

Proposition 11.16. *Let $f : \mathbb{R}^2 \supset U \to \mathbb{R}^3$ be a curvature line parametrization. Then f is a channel surface (in u direction) if and only if its focal surface (in v direction) degenerates to a curve in the following way:*

$$f_v^{(2)} = 0$$

Proof. We have already demonstrated ($\Rightarrow$). Now ($\Leftarrow$):

$$0 = f_v^{(2)} = f_v + \left(\frac{1}{\kappa_2}\right)_v n + \frac{1}{\kappa_2} n_v = \frac{1}{\kappa_2}{}_v\, n \quad \Leftrightarrow \quad (\kappa_2)_v = 0$$

Thus the curvature spheres $S^{(2)}$ in have constant radius v direction, i.e., do not depend on v. (Equivalently the lift $[\hat{s}_v^{(2)}]$ degenerates to a curve as well $\hat{s}_v^{(2)} = 0$.) This means that the f is the envelope of this one parameter family of spheres $\square$

Remark 11.4. The Möbius lift of the parametrization $[\hat{f}]$ and of the one-parameter family of spheres $[\hat{s}]$ are related by

$$\left\langle \hat{f}, \hat{s} \right\rangle_{4,1} = \left\langle \hat{f}, \hat{s}_u \right\rangle_{4,1} = 0$$

Thus, the circles C are given by $([\hat{s}] \vee [\hat{s}_u])^{\perp}$. Furthermore, the lift of the curvature spheres in v direction $[\hat{s}^{(2)}]$ describes a curve, which coincides with $[\hat{s}]$.

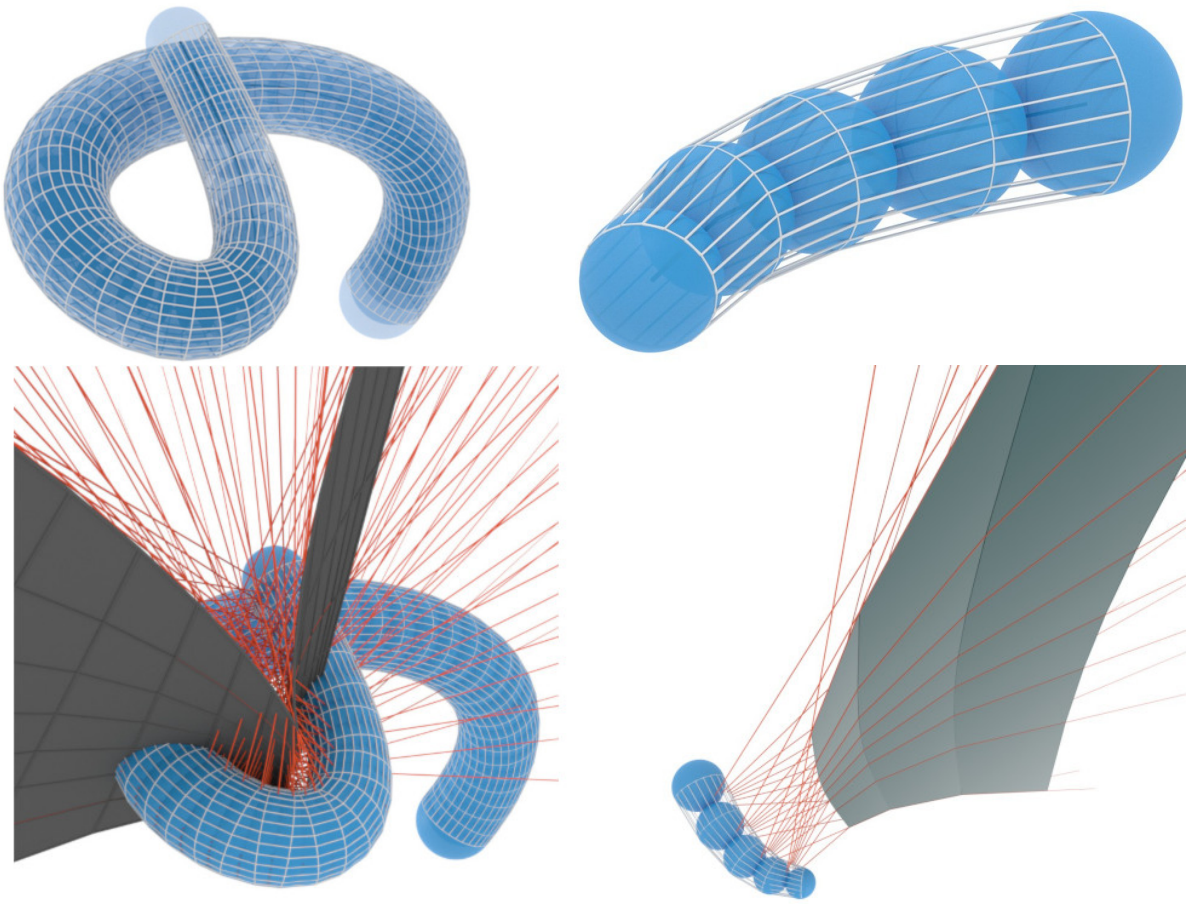


Figure 33. Smooth and discrete channel surfaces and their focal surfaces.

We now consider the special case of double channel surfaces .

Definition 11.19. A surface which is the envelope of two (distinct) one-parameter families of spheres are called *Dupin cyclides*.

By the consideration above these two one-parameter family of spheres must coincide with the curvature spheres $S^{(1)}(u)$ and $S^{(2)}$ of the surface. Furthermore, since the surface envelopes both families, each sphere from $S^{(1)}$ must touch each sphere of $S^{(2)}$. For the Möbius lifts $[\hat{s}^{(1)}]$, $[\hat{s}^{(2)}]$ this means that all lines $[\hat{s}^{(1)}] \vee [\hat{s}^{(2)}]$ are tangent to the Möbius quadric $\mathbb{S}^3 \subset \mathbb{R}\mathrm{P}^4$, or equivalently,

$$\left\langle \hat{s}^{(1)}, \hat{s}^{(2)} \right\rangle_{4,1}^2 - \left\| \hat{s}^{(1)} \right\|_{4,1}^2 \left\| \hat{s}^{(2)} \right\|_{4,1}^2 = 0$$

If we add an additional homogeneous coordinate

$$\tilde{s}^{(i)} := \left(\hat{s}^{(i)}, \left\| \hat{s}^{(i)} \right\|_{4,1} \right)$$

to define a lift $[\tilde{s}^{(i)}] \in \mathcal{L} \subset \mathbb{R}\mathrm{P}^5$ to a quadric with signature $(4,2)$ in $\mathbb{R}\mathrm{P}^5$ the condition above becomes polarity

$$\left\langle \tilde{s}^{(1)}, \tilde{s}^{(2)} \right\rangle_{4,2} = 0$$

with respect to this quadric, the so called *Lie quadric*. This implies that the two curves $[\tilde{s}^{(1)}]$ and $[\tilde{s}^{(2)}]$ must be planar sections of the Lie quadric, i.e, conic section. This means that the projection back to $\mathbb{R}\mathrm{P}^4$ is given by two conics, and so is the stereographic projection to $\mathbb{R}^3$, which describes the focal surfaces. Further investigation reveals that the two focal surfaces are given by two *focal conics*.

Example 11.2. A torus is a channel surface in two directions, and thus a Dupin cyclide. This property is invariant under Möbius transformations. So, all Möbius images of a torus are Dupin cyclides.

11.9 Q-nets, circular nets, and discrete channel surfaces

- We discretize parametrized surfaces in terms of discrete nets, i.e, maps

$$f : \mathbb{Z}^2 \supset U \to \mathbb{R}\mathrm{P}^n.$$

- A *Q-net* (*discrete conjugate nets*) is a net $f : U \to \mathbb{R}\mathrm{P}^n$, such that the four image points of each quad are contained in a plane, i.e.,

$$f(m,n), f(m+1,n), f(m+1,n+1), f(m,n+1) \text{ are coplanar for all } m,n.$$

  Equivalently, in affine coordinates, f satisfies an equation of the form

$$\Delta_1 \Delta_2 f = \alpha \Delta_1 f + \beta \Delta_2 f.$$

- A *circular net* (*discrete principal nets*) is a net $f : U \to \mathbb{R}^n$, such that the four image points of each quad are contained in a circle, i.e.,

$$f(m,n), f(m+1,n), f(m+1,n+1), f(m,n+1) \text{ lie on a circle for all } m,n.$$

  The axes of the circles can be interpreted as discrete normals (per face). Adjacent discrete normal lines intersect, and in this sense they form discrete developable surfaces.

- Starting from a discrete one-parameter family of spheres $S : \mathbb{Z} \supset I \to \mathbb{R}^3$, we can generalize the two definitions for envelopes of circles to this case:

- The sequence of intersection circles of adjacent spheres can be thought of as the discrete envelope.
- Start with one circle on one sphere, and propagate this circle to the other spheres by inversion in the midspheres of adjacent spheres. The sequence of the obtained circles can be thought of as the discrete envelope.

  To obtain a discrete net, sample the initial circle, and propagate the points by the same inversions. In this way, one obtains a circular net as the discrete envelope.