CHAPTER 5

Affine morphisms

Contents

Let $\phi : V \to W$ be a linear map between the $\mathbb{R}$-vector spaces V and W. Let $\boldsymbol{v} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a basis for V and let $\boldsymbol{w} = \{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ be a basis for W. In your Algebra course, you used the notation $[\phi]_{\boldsymbol{v},\boldsymbol{w}}$ for the matrix of the linear map ϕ with respect to the bases $\boldsymbol{v}$ and $\boldsymbol{w}$. We will use the notation

$$\mathrm{M}_{\boldsymbol{w},\boldsymbol{v}}(\phi) = [\phi]_{\boldsymbol{v},\boldsymbol{w}}$$

Notice that the indices $\boldsymbol{v}, \boldsymbol{w}$ are *reversed*.

You have also learned that if $\psi : W \to U$ is another linear map, to some vector space U with basis $\boldsymbol{u} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$, then

$$\mathrm{M}_{\boldsymbol{u},\boldsymbol{v}}(\psi \circ \phi) = \mathrm{M}_{\boldsymbol{u},\boldsymbol{w}}(\psi) \cdot \mathrm{M}_{\boldsymbol{w},\boldsymbol{v}}(\phi).$$

In particular if $V = W$ then

$$\mathrm{M}_{\boldsymbol{w},\boldsymbol{w}}(\phi) = \mathrm{M}_{\boldsymbol{w},\boldsymbol{v}}(\mathrm{Id}_V) \cdot \mathrm{M}_{\boldsymbol{v},\boldsymbol{v}}(\phi) \cdot \mathrm{M}_{\boldsymbol{v},\boldsymbol{w}}(\mathrm{Id}_V) = \mathrm{M}_{\boldsymbol{v},\boldsymbol{w}}(\mathrm{Id}_V)^{-1} \cdot \mathrm{M}_{\boldsymbol{v},\boldsymbol{v}}(\phi) \cdot \mathrm{M}_{\boldsymbol{v},\boldsymbol{w}}(\mathrm{Id}_V).$$

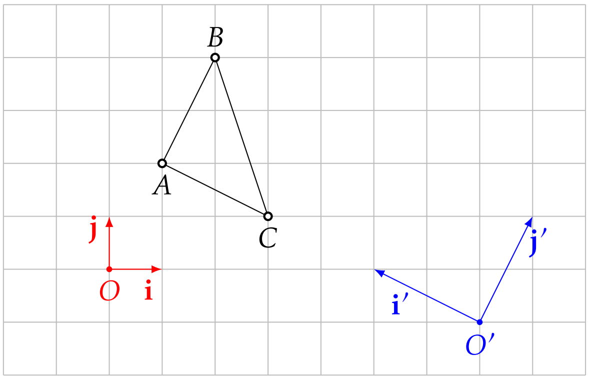
So, the matrix of ϕ with respect to the basis $\boldsymbol{w}$ is obtained from the matrix of ϕ with respect to the basis $\boldsymbol{v}$ by conjugating with the matrix $\mathrm{M}_{\boldsymbol{v},\boldsymbol{w}}(\mathrm{Id}_V)$. We call the matrix $\mathrm{M}_{\boldsymbol{v},\boldsymbol{w}} := \mathrm{M}_{\boldsymbol{v},\boldsymbol{w}}(\mathrm{Id}_V)$ the *change of basis matrix from the basis $\boldsymbol{w}$ to the basis $\boldsymbol{v}$*. If $\mathcal{K}$ and $\mathcal{K}'$ are two coordinate systems, we write $\mathrm{M}_{\mathcal{K},\mathcal{K}'}$ for the base change matrix w.r.t. the corresponding bases.

5.1 Changing coordinate systems

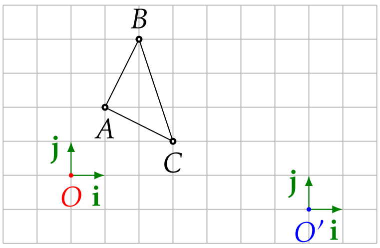
Example (In dimension 2)**.** Let $\mathcal{K} = O\mathbf{ij}$ and $\mathcal{K}' = O'\mathbf{i}'\mathbf{j}'$ be two coordinate systems (reference frames) of $\mathbb{E}^2$. Suppose that we know O', $\mathbf{i}'$ and $\mathbf{j}'$ relative to $\mathcal{K}$:

$$[O']_{\mathcal{K}} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}_{\mathcal{K}}, \quad \mathbf{i}' = -2\mathbf{i} + \mathbf{j} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}_{\mathcal{K}}, \quad \mathbf{j}' = \mathbf{i} + 2\mathbf{j} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}_{\mathcal{K}}.$$

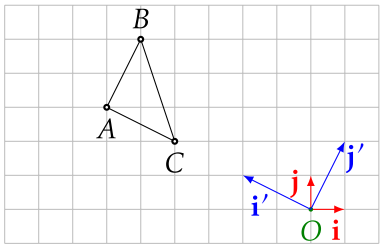
How can we translate the coordinates of points from $\mathcal{K}$ to $\mathcal{K}'$?



We can do this in two steps: (a) first we change the origin, i.e. we go from $O\mathbf{ij}$ to $O'\mathbf{ij}$ and (b) we change the directions of the coordinate axes, i.e. we go from $O'\mathbf{ij}$ to $O'\mathbf{i}'\mathbf{j}'$. The first step is just a translation and the second step corresponds to the usual base change from linear algebra.



(a) Change the origin.



(b) Change the direction of the axes.

For the first step

$$[\overrightarrow{O'A}]_{\mathcal{K}'} = [\overrightarrow{O'A}]_{\mathcal{K}} = [\overrightarrow{OA}]_{\mathcal{K}} - [\overrightarrow{OO'}]_{\mathcal{K}}.$$

For the second step let $\mathrm{M}_{\mathcal{K},\mathcal{K}'}$ denote the base change matrix from the basis in $\mathcal{K}'$ to the basis in $\mathcal{K}$. Then

$$[\overrightarrow{OA}]_{\mathcal{K}} = \mathrm{M}_{\mathcal{K},\mathcal{K}'}[\overrightarrow{OA}]_{\mathcal{K}'} \quad \text{and} \quad [\overrightarrow{OA}]_{\mathcal{K}'} = \mathrm{M}_{\mathcal{K}',\mathcal{K}}[\overrightarrow{OA}]_{\mathcal{K}}.$$

Hence composing the two operations, (a) and (b), we obtain

$$[\overrightarrow{OA}]_{\mathcal{K}'} = \mathrm{M}_{\mathcal{K}',\mathcal{K}}([\overrightarrow{OA}]_{\mathcal{K}} - [\overrightarrow{OO'}]_{\mathcal{K}}) = \mathrm{M}_{\mathcal{K}',\mathcal{K}}[\overrightarrow{OA}]_{\mathcal{K}} - \mathrm{M}_{\mathcal{K}',\mathcal{K}}[\overrightarrow{OO'}]_{\mathcal{K}} = \mathrm{M}_{\mathcal{K}',\mathcal{K}}[\overrightarrow{OA}]_{\mathcal{K}} - [\overrightarrow{OO'}]_{\mathcal{K}'}$$

Hence, the formula for changing coordinates from the system $\mathcal{K}$ to the system $\mathcal{K}'$ is

$$[A]_{\mathcal{K}'} = \mathrm{M}_{\mathcal{K}',\mathcal{K}}([A]_{\mathcal{K}} - [O']_{\mathcal{K}}) = \mathrm{M}_{\mathcal{K}',\mathcal{K}}[A]_{\mathcal{K}} + [O]_{\mathcal{K}'} \tag{5.1}$$

Suppose now that a point A is given and that the coordinates of A in the frame $\mathcal{K}$ (relative to the coordinate frame/coordinate system $\mathcal{K}$) are $(1,2)$. Then the coordinates of A relative to $\mathcal{K}'$ are

$$[A]_{\mathcal{K}'} = \mathrm{M}_{\mathcal{K}',\mathcal{K}}\begin{bmatrix}1\\2\end{bmatrix}_{\mathcal{K}} + [O]_{\mathcal{K}'}$$

Since we know $\mathbf{i}'$ and $\mathbf{j}'$ with respect to $\mathbf{i}$ and $\mathbf{j}$, we can write down the matrix $\mathrm{M}_{\mathcal{K},\mathcal{K}'}$ and then $\mathrm{M}_{\mathcal{K}',\mathcal{K}} = \mathrm{M}^{-1}_{\mathcal{K},\mathcal{K}'}$. Since we know the coordinates of O' with respect to $\mathcal{K}$, it is more convenient to use the first equality in (5.1)

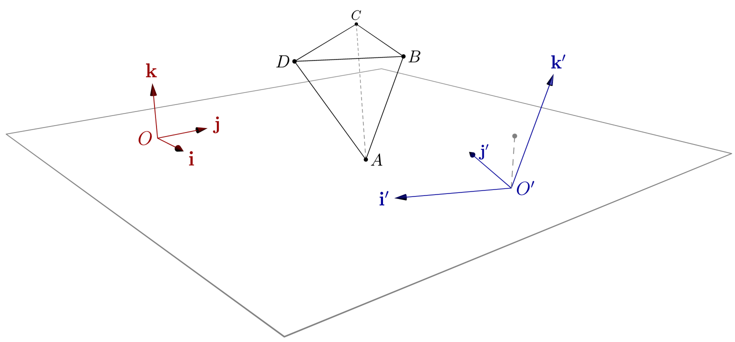
$$[A]_{\mathcal{K}'} = \begin{bmatrix}-2 & 1\\ 1 & 2\end{bmatrix}^{-1} \cdot \left(\begin{bmatrix}1\\2\end{bmatrix}_{\mathcal{K}} - \begin{bmatrix}7\\-1\end{bmatrix}_{\mathcal{K}}\right) = \frac{1}{-5}\begin{bmatrix}2 & -1\\ -1 & -2\end{bmatrix} \cdot \left(\begin{bmatrix}1\\2\end{bmatrix}_{\mathcal{K}} + \begin{bmatrix}7\\-1\end{bmatrix}_{\mathcal{K}}\right) = \begin{bmatrix}3\\0\end{bmatrix}.$$

Theorem 5.1. Let $\mathcal{K} = O\mathbf{e}_1 \dots \mathbf{e}_n$ and $\mathcal{K}' = O'\mathbf{f}_1, \dots, \mathbf{f}_n$ be two coordinate systems of $\mathbb{E}^n$. Denote by $\boldsymbol{e} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and by $\boldsymbol{f} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ the two bases of $\mathbb{V}^n$. Let $\mathrm{M}_{\mathcal{K}',\mathcal{K}}$ be the change of basis matrix from the basis $\boldsymbol{e}$ to the basis $\boldsymbol{f}$ and let $\mathrm{M}_{\mathcal{K},\mathcal{K}'}$ be the change of basis matrix from the basis $\boldsymbol{f}$ to the basis $\boldsymbol{e}$. For any point $P \in \mathbb{E}^n$ we have

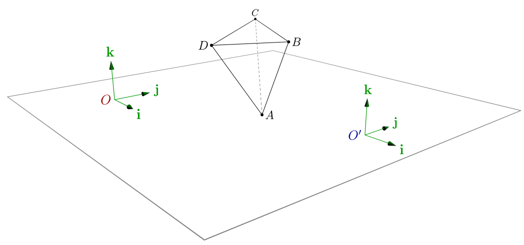
$$[P]_{\mathcal{K}'} = \mathrm{M}_{\mathcal{K}',\mathcal{K}} \cdot ([P]_{\mathcal{K}} - [O']_{\mathcal{K}}) = \mathrm{M}^{-1}_{\mathcal{K},\mathcal{K}'} \cdot ([P]_{\mathcal{K}} - [O']_{\mathcal{K}}) = \mathrm{M}_{\mathcal{K}',\mathcal{K}} \cdot [P]_{\mathcal{K}} + [O]_{\mathcal{K}'}. \tag{5.2}$$

Example (In dimension 3)**.** Let $\mathcal{K} = O\mathbf{ijk}$ and $\mathcal{K}' = O'\mathbf{i}'\mathbf{j}'\mathbf{k}'$ be two coordinate systems (reference frames) of $\mathbb{E}^3$. Suppose that we know O', $\mathbf{i}'$, $\mathbf{j}'$ and $\mathbf{k}'$ relative to $\mathcal{K}$:

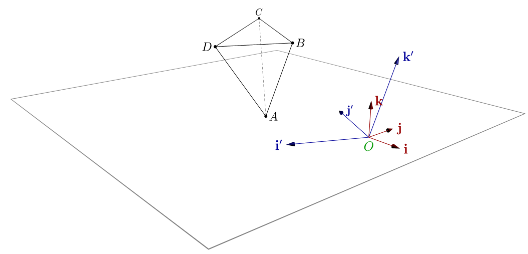
$$[O']_{\mathcal{K}} = \begin{bmatrix}4\\5\\-1\end{bmatrix}_{\mathcal{K}}, \quad \mathbf{i}' = -\mathbf{i} - 2\mathbf{j} = \begin{bmatrix}-1\\-2\\0\end{bmatrix}_{\mathcal{K}}, \quad \mathbf{j}' = -2\mathbf{i} + \mathbf{j} = \begin{bmatrix}-2\\1\\0\end{bmatrix}_{\mathcal{K}}, \quad \mathbf{k}' = \mathbf{j} + 2\mathbf{k} = \begin{bmatrix}0\\1\\2\end{bmatrix}_{\mathcal{K}}.$$



The coordinates with respect to $\mathcal{K}'$ can be obtained from the coordinates with respect to $\mathcal{K}$ in two steps:



(a) Change the origin.



(b) Change the direction of the axes.

If B is the point with coordinates $(1,5,1)$ with respect to $\mathcal{K}$, then

$$[B]_{\mathcal{K}'} = \mathrm{M}_{\mathcal{K},\mathcal{K}'}^{-1} \cdot ([B]_{\mathcal{K}} - [O']_{\mathcal{K}}) = \begin{bmatrix} -1 & -2 & 0 \\ -2 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}^{-1} \left(\begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix} - \begin{bmatrix} 4 \\ 5 \\ -1 \end{bmatrix} \right) = \frac{1}{10} \begin{bmatrix} -2 & -4 & 2 \\ -4 & 2 & -1 \\ 0 & 0 & 5 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix} - \begin{bmatrix} 4 \\ 5 \\ -1 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

5.2 Affine morphisms

Definition. An affine morphism $\phi : \mathbb{E}^n \to \mathbb{E}^m$ is a map which relative to some reference frame $\mathcal{K}$ of $\mathbb{E}^n$ and some reference frame $\mathcal{K}'$ of $\mathbb{E}^m$ can be expressed as

$$[\phi(P)]_{\mathcal{K}'} = A \cdot [P]_{\mathcal{K}} + b \tag{5.3}$$

with $A \in \mathrm{Mat}_{m \times n}(\mathbb{R})$ and $b \in \mathrm{Mat}_{m \times 1}(\mathbb{R})$.

- In (5.3) both A and b depend on the choice of the coordinate systems $\mathcal{K}$ and $\mathcal{K}'$.
- If $n = m$, so, if $\phi : \mathbb{E}^n \to \mathbb{E}^n$ then ϕ is called an *affine endomorphisms*. The set of all such endomorphisms is denoted by $\mathrm{End}_{\mathrm{aff}}(\mathbb{E}^n)$.
- The morphism ϕ is invertible if and only if the matrix A is invertible.
- If $n = m$ and ϕ is invertible then ϕ is called an *affine automorphism* or *affine transformation*. The set of all affine transformations of $\mathbb{E}^n$ is denoted by $\mathrm{AGL}(\mathbb{E}^n)$
- If $\mathcal{K} = \mathcal{K}'$ in (5.3) and b is the zero vector, then ϕ is a linear map. In particular

$$\mathrm{GL}(\mathbb{V}^n) \subseteq \mathrm{AGL}(\mathbb{E}^n).$$

5.3 Projections and reflections in a hyperplane

Definition. A hyperplane in $\mathbb{E}^n$ is a subset of points H satisfying a linear equation,

$$H : a_1 x_1 + \cdots + a_n x_n + a_{n+1} = 0$$

with respect to some coordinate system $O\mathbf{e}_1 \ldots \mathbf{e}_n$.

- Hyperplanes in $\mathbb{E}^2$ are lines.
- Hyperplanes in $\mathbb{E}^3$ are planes.
- Consider a line $\ell \subseteq \mathbb{E}^n$ passing through a point $Q(q_1,\dots,q_n)$ and having $\mathbf{v}(v_1,\dots,v_n)$ as direction vector:
$$\ell = \{Q + t\mathbf{v} : t \in \mathbb{R}\}. \tag{5.4}$$
Consider a hyperplane $H \subseteq \mathbb{E}^n$ given by the Cartesian equation
$$H : a_1x_1 + \cdots + a_nx_n + a_{n+1} = 0 \tag{5.5}$$
The intersection $\ell \cap H$ can be described as follows
$$Q + t\mathbf{v} \in \ell \cap H \quad \Leftrightarrow \quad a_1(q_1 + tv_1) + \cdots + a_n(q_n + tv_n) + a_{n+1} = 0$$
So, the intersection point (if it exists) is
$$Q' = Q - \frac{a_1q_1 + \cdots + a_nq_n + a_{n+1}}{v_1q_1 + \cdots + v_nq_n}\mathbf{v}. \tag{5.6}$$

5.3.1 Tensor products

Definition. Let $\mathbf{v}(v_1,\dots,v_n)$ and $\mathbf{w}(w_1,\dots,w_n)$ be two vectors. The *tensor product* $\mathbf{v} \otimes \mathbf{w}$ is the $n \times n$ matrix defined by $(\mathbf{v} \times \mathbf{w})_{ij} = v_iw_j$. In other words

$$\mathbf{v} \otimes \mathbf{w} = \mathbf{v} \cdot \mathbf{w}^t = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \cdot \begin{bmatrix} w_1, \dots, w_n \end{bmatrix} = \begin{bmatrix} v_1w_1 & \dots & v_1w_n \\ \vdots & & \vdots \\ v_nw_1 & \dots & v_nw_n \end{bmatrix}.$$

Proposition 5.2. The map $\mathbb{R}^n \times \mathbb{R}^n \to \mathrm{Mat}_{n\times n}(\mathbb{R})$ given by $(\mathbf{v},\mathbf{w}) \mapsto \mathbf{v} \otimes \mathbf{w}$ has the following properties:

1. It is linear in both arguments,
2. $(\mathbf{v} \otimes \mathbf{w})^t = \mathbf{w} \otimes \mathbf{v}$.

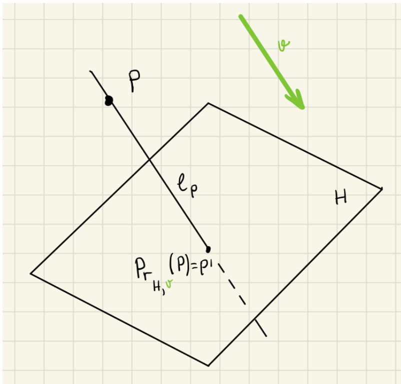
- The tensor product of two vectors is also called outer product - to be compared with the inner product which is
$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^t \cdot \mathbf{w} = \begin{bmatrix} v_1, \dots, v_n \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} = v_1w_n + \cdots + v_nw_n.$$
- It is easy to show that for three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{V}^n$ we have
$$(\mathbf{u} \otimes \mathbf{v})\mathbf{w} = (\mathbf{v} \cdot \mathbf{w})\mathbf{u}.$$

5.3.2 Parallel projection on a hyperplane

Definition. Let H be a hyperplane and let $\mathbf{v}$ be a vector in $\mathbb{V}^n$ which is not parallel to H. For any point $P \in \mathbb{E}^n$ there is a unique line ℓ_P passing through P and having $\mathbf{v}$ as direction vector. The line ℓ_P is not parallel to H, hence, it intersects H in a unique point P'. We denote P' by $\mathrm{Pr}_{H,\mathbf{v}}(P)$ and call it the *projection of the point P on the hyperplane H parallel to* $\mathbf{v}$. This gives a map

$$\mathrm{Pr}_{H,\mathbf{v}} : \mathbb{E}^n \to \mathbb{E}^n$$

called, the *projection on the hyperplane H parallel to* $\mathbf{v}$.



- With respect to the reference frame $\mathcal{K}$, the hyperplane H has an equation as in (5.5).
- By (5.6), $\mathrm{Pr}_{H,\mathbf{v}}(P) = P - \frac{\varphi(P)}{\mathrm{lin}\,\varphi(\mathbf{v})}\mathbf{v}$ where $\phi(P(p_1,\dots,p_n)) = a_1p_1 + \cdots + a_np_n + p_{n+1}$.
- Hence, if we denote by $p'_1,\dots,p'_n$ the coordinates of the projected point $\mathrm{Pr}_{H,\mathbf{v}}(P)$ then

$$\begin{cases} p'_1 = p_1 + \mu v_1 \\ \quad\vdots \\ p'_n = p_n + \mu v_n \end{cases} \quad \text{where} \quad \mu = -\frac{a_1p_1 + \cdots + a_np_n + a_{n+1}}{a_1v_1 + \cdots + a_nv_n}.$$

- In matrix form, we can rearrange this as follows

$$\begin{bmatrix} p'_1 \\ p'_2 \\ \vdots \\ p'_n \end{bmatrix} = \underbrace{\frac{1}{\mathrm{lin}\,\varphi(\mathbf{v})}\begin{bmatrix} \sum_{i=1}^{n,i\neq 1} a_iv_i & -a_2v_1 & -a_3v_1 & \dots & -a_nv_1 \\ -a_1v_2 & \sum_{i=1}^{n,i\neq 2} a_iv_i & -a_3v_2 & \dots & -a_nv_2 \\ \vdots & & & & \vdots \\ -a_1v_n & -a_2v_n & \dots & -a_{n-1}v_n & \sum_{i=1}^{n,i\neq n} a_iv_i \end{bmatrix}}_{\mathrm{lin}\,\mathrm{Pr}_{H,\mathbf{v}}}\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{bmatrix} - \frac{a_{n+1}}{\mathrm{lin}\,\varphi(\mathbf{v})}\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

- It is possible to give a more compact description of the above matrix form if we use tensor products:

$$[\mathrm{Pr}_{H,\mathbf{v}}(P)]_{\mathcal{K}} = \frac{1}{\operatorname{lin}\varphi(\mathbf{v})}\left((\mathbf{v}^t\cdot\mathbf{a})\,\mathrm{Id}_n - \underbrace{\mathbf{v}\cdot\mathbf{a}^t}_{\mathbf{v}\otimes\mathbf{a}}\right)\cdot[P]_{\mathcal{K}} - \frac{a_{n+1}}{\operatorname{lin}\varphi(\mathbf{v})}[\mathbf{v}]_{\mathcal{K}}$$

  where $\mathbf{a} = (a_1,\dots,a_n)^t$.

- If we further notice that $\operatorname{lin}\varphi(\mathbf{v}) = \mathbf{v}^t\cdot\mathbf{a}$ then

$$[\mathrm{Pr}_{H,\mathbf{v}}(P)]_{\mathcal{K}} = \left(\mathrm{Id}_n - \frac{\mathbf{v}\cdot\mathbf{a}^t}{\mathbf{v}^t\cdot\mathbf{a}}\right)\cdot[P]_{\mathcal{K}} - \frac{a_{n+1}}{\mathbf{v}^t\cdot\mathbf{a}}[\mathbf{v}]_{\mathcal{K}}$$

- Parallel projections on hyperplanes are affine morphisms. Obviously, they are not bijective, so

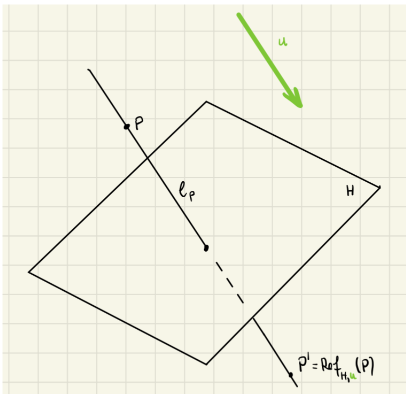
$$\mathrm{Pr}_{H,\mathbf{v}} \in \mathrm{End}_{\mathrm{aff}}(\mathbb{E}^n) \quad \text{but} \quad \mathrm{Pr}_{H,\mathbf{v}} \notin \mathrm{AGL}(\mathbb{E}^n).$$

5.3.3 Parallel reflection in a hyperplane

Definition. Let H be a hyperplane and let $\mathbf{v}$ be a vector in $\mathbb{V}^n$ which is not parallel to H. For any point $P\in\mathbb{E}^n$ there is a unique point P' such that $\mathrm{Pr}_{H,\mathbf{v}}(P)$ is the midpoint of the segment $[PP']$. We denote P' by $\mathrm{Ref}_{H,\mathbf{v}}(P)$ and call it the *reflection of the point P in the hyperplane H parallel to* $\mathbf{v}$. This gives a map

$$\mathrm{Ref}_{H,\mathbf{v}} : \mathbb{E}^n \to \mathbb{E}^n$$

called, the *reflection in the hyperplane H parallel to* $\mathbf{v}$.



- With respect to the reference frame $\mathcal{K}$, the hyperplane H has an equation as in (5.5).
- Since $\mathrm{Pr}_{H,\mathbf{v}}(P)$ is the midpoint of the segment $[PP']$, with respect to $\mathcal{K}$ we have

$$[P]_{\mathcal{K}} - \frac{\varphi(P)}{\operatorname{lin}\varphi(\mathbf{v})}[\mathbf{v}]_{\mathcal{K}} = \frac{[P]_{\mathcal{K}}+[P']_{\mathcal{K}}}{2} \quad\Leftrightarrow\quad 2[P]_{\mathcal{K}} - 2\frac{\varphi(P)}{\operatorname{lin}\varphi(\mathbf{v})}[\mathbf{v}]_{\mathcal{K}} = [P]_{\mathcal{K}}+[P']_{\mathcal{K}}.$$

Therefore

$$[\mathrm{Ref}_{H,\mathbf{v}}(P)]_{\mathcal{K}} = [P]_{\mathcal{K}} - 2\frac{\varphi(P)}{\mathrm{lin}\,\varphi(\mathbf{v})}[\mathbf{v}]_{\mathcal{K}}.$$

- As in the case of $\mathrm{Pr}_{H,\mathbf{v}}$, it is possible to give a compact description of the matrix form if we use tensor products:

$$[\mathrm{Pr}_{H,\mathbf{v}}(P)]_{\mathcal{K}} = \left(\mathrm{Id}_n - 2\frac{\mathbf{v}\otimes\mathbf{a}}{\mathbf{v}^t\cdot\mathbf{a}}\right)\cdot[P]_{\mathcal{K}} - 2\frac{a_{n+1}}{\mathbf{v}^t\cdot\mathbf{a}}[\mathbf{v}]_{\mathcal{K}}$$

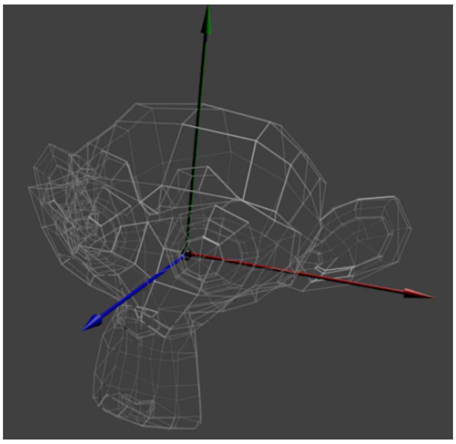
where $\mathbf{a} = (a_1,\dots,a_n)^t$.

- Parallel reflections in hyperplanes are affine morphisms. Obviously, they are bijective, so

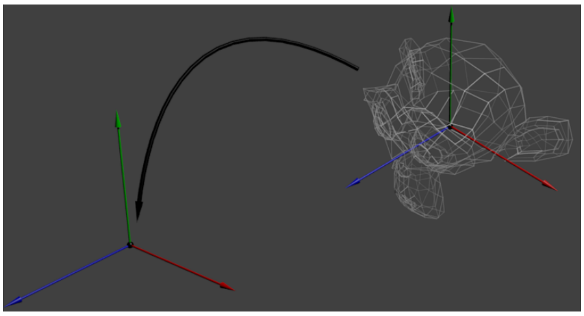
$$\mathrm{Ref}_{H,\mathbf{v}} \in \mathrm{AGL}(\mathbb{E}^n) \subseteq \mathrm{End}_{\mathrm{aff}}(\mathbb{E}^n).$$

5.4 Applications

- [Changing reference frames] Changing coordinates is so common in real life applications that it is impossible to do any meaningful modeling without coordinate changes. Take for example the OpenGL-pipeline (a process under which objects are rendered on the screen in order to simulate a 3D scene). In this process there are two dimensional changes of coordinates for instance when a PNG-image is transformed into an OpenGL-texture, or when an OpenGL texture is translated into device coordinates before being displayed on the screen.



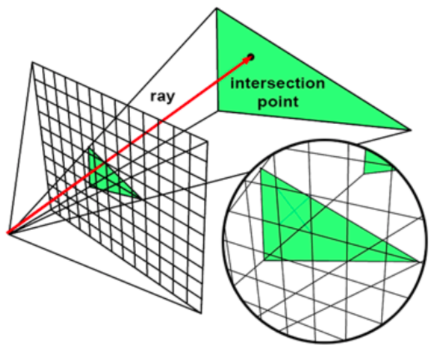
(a) Model space.



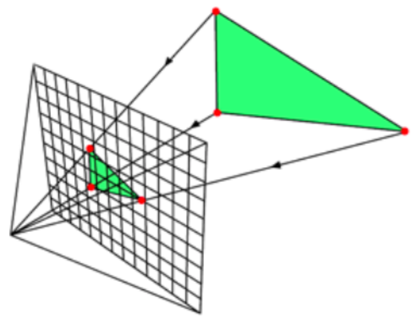
(b) World space.

Three dimensional changes of coordinates are needed in the OpenGL-pipeline as well as in any modeling software (like Blender or Maya). Complex 3D-models are constructed by several people, even several teams. The different pieces are constructed separately, each in its own coordinate system. When they are ready, the whole object is assembled by putting all pieces in a common coordinate system. In this context one calls the coordinate system of a single piece 'model space' and the coordinate system where the whole object is placed in, is called the 'world space'. This is important when modeling machines such as spacecrafts or cars. It is also important when constructing computer games. As an example, above is Suzanne, the Blender monkey.

- [Projecting] There are two main types of algorithms which project a 3D scene in computer graphics: ray-tracing algorithms and rasterization algorithms. Ray-tracing algorithms intersect rays with planes determined by the triangles in the scene while rasterization algorithms try to project the triangles on the display screen.



(a) Ray-tracing.



(b) Rasterization.

It is clearly much more efficient to construct a projection map like $\mathrm{Pr}_{H,\ell}$ and project all the triangles, however, this only works for parallel projections. In order to simulate perspective, rasterization algorithms use a projective transformation before using a projection like the one described in the previous paragraphs. Ray-tracing algorithms are conceptually much simpler but they are much more resource intensive.

CHAPTER 6

Isometries

Contents

In these notes we assume working knowledge of eigenvalues and eigenvectors. Have a look at your Algebra course from last semester.

6.1 Isometries

Definition. An *isometry* is a map $\phi : \mathbb{E}^n \to \mathbb{E}^n$ which preserves distances, i.e.

$$d(\phi(P), \phi(Q)) = d(P, Q)$$

for any points $P, Q \in \mathbb{E}^n$.

- One can show that isometries are affine transformations, i.e. they are elements in $\mathrm{AGL}(\mathbb{E}^n)$.

Proposition 6.1. Let $\phi \in \mathrm{AGL}(\mathbb{E}^n)$ be an affine transformation given by $\phi(\mathbf{x}) = A\mathbf{x} + b$ with respect to some orthonormal coordinate system. The following are equivalent:

1. ϕ is an isometry
2. $A^{-1} = A^t$.

Proposition 6.2. Let $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ be a matrix such that $A^tA = I_n$. Then $\det(A) \in \{\pm 1\}$.

Definition. A matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ such that $A^tA = I_n$ is called *orthogonal*. The set of all such matrices are denoted by $O(n)$. The set of matrices in $O(n)$ with determinant 1 is denoted by $SO(n)$. Such matrices are called *special orthogonal*.

The set $O(n)$ is a subgroup of $\mathrm{AGL}(\mathbb{R}^n)$ and $SO(n)$ is a normal subgroup of $O(n)$:

$$SO(n) \trianglelefteq O(n) \leq \mathrm{AGL}(\mathbb{R}^n).$$

Let $\phi \in \mathrm{AGL}(\mathbb{E}^n)$ be given by $\phi(\mathbf{x}) = A\mathbf{x}+b$ with respect to some orthonormal coordinate system. Then ϕ is called a *displacement*, or a *direct isometry*, if $A \in SO(n)$.

6.1.1 Rotations in dimension 2

Proposition 6.3. A matrix A is in $SO(2)$ if and only if A has the form

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

for some $\theta \in \mathbb{R}$.

Corollary 6.4. A direct isometry ϕ of $\mathbb{E}^2$ that fixes a point is either the identity or a rotation. Moreover, the angle θ of the rotation is such that

$$\cos(\theta) = \frac{\mathrm{tr}(\mathrm{lin}(\phi))}{2}.$$

6.1.2 Rotations in dimension 3

Theorem 6.5 (Euler)**.** A direct isometry ϕ of $\mathbb{E}^3$ that fixes a point is either the identity or a rotation around an axis that passes through that point. Moreover, the angle θ of the rotation is such that

$$\cos(\theta) = \frac{\mathrm{tr}(\mathrm{lin}(\phi)) - 1}{2}.$$

6.1.3 Classification of isometries

Theorem 6.6 (Chasles)**.** An isometry of the plane $\mathbb{E}^2$ is either a direct isometry, in which case it is

- the identity, or
- a translation, or
- a rotation;

or an indirect isometry, in which case it is

- a reflection, or

- a glidereflection.

Theorem 6.7 (Euler)**.** Any isometry of the 3-dimensional Euclidean space $\mathbb{E}^3$ is either a direct isometry, in which case it is

- the identity, or
- a translation, or
- a rotation around an axis, or
- a gliderotation (also called helical displacement);

or an indirect isometry, in which case it is

- a reflection, or
- a glidereflection, or
- a rotation-reflection.

6.2 Spectral theorem

Proposition 6.8. Any set of mutually orthogonal vectors is linearly independent. In particular, any set of n mutually orthogonal vectors is a basis of $\mathbb{V}^n$.

Proposition 6.9. Let $\boldsymbol{e} = \{\mathbf{e}_1,\dots,\mathbf{e}_n\}$ and $\boldsymbol{f} = \{\mathbf{f}_1,\dots,\mathbf{f}_n\}$ be two bases of the vector space $\mathbb{V}^n$, and suppose that $\boldsymbol{e}$ is orthonormal. The basis $\boldsymbol{f}$ is orthonormal if and only if the base change matrix $\mathrm{M}_{\boldsymbol{e},\boldsymbol{f}}$ is orthogonal.

- Proposition 6.9, says that if we change the coordinate system from an orthonormal basis $\boldsymbol{e}$ to an orthonormal basis $\boldsymbol{f}$ then the base change matrix $\mathrm{M}_{\boldsymbol{e},\boldsymbol{f}}$ is in $O(n)$, i.e. the base change is an isometry.
- For a matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ and a matrix $M \in O(n)$, since $M^{-1} = M^t$ we have
$$M^{-1}AM = M^tAM.$$
- The above observation is of funamental importance for the isometric classification of quadrics. Suppose you have a quadratic equation
$$2x^2 - 6xy - 1y^2 = c \tag{6.1}$$
for some constant $c \in \mathbb{R}$. Then, you may write it in matrix form like this:
$$\begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} 2 & -3 \\ -3 & -1 \end{bmatrix}}_{=:A} \begin{bmatrix} x \\ y \end{bmatrix} = c$$

and if we change coordinates with a change of basis matrix M then

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = M \begin{bmatrix} x \\ y \end{bmatrix}$$

and thus, Equation (6.1) becomes

$$\begin{bmatrix} x' & y' \end{bmatrix} \left(M^t A M \right) \begin{bmatrix} x' \\ y' \end{bmatrix} = c$$

Now if M is orthogonal, i.e. $M \in O(n)$ (which by Proposition 6.9 happens when we go from one orthonormal coordinate system to another), then the above equation becomes

$$\begin{bmatrix} x' & y' \end{bmatrix} \left(M^{-1} A M \right) \begin{bmatrix} x \\ y \end{bmatrix} = c$$

The matrix A is the matrix of a linear map ϕ_A, and $M^{-1}AM$ is the matrix of the same linear map with respect to a different (orthonormal) basis. Thus, if we find an orthonormal basis with respect to which the matrix of ϕ_A is 'nice' (spoiler: diagonal), then with respect to that basis Equation (6.1) becomes

$$\lambda_1 x^2 + \lambda_2 y^2 = c$$

for some values $\lambda_1, \lambda_2 \in \mathbb{R}$ which are the eigenvalues of ϕ_A and of A.

- Notice that if we change the order of two vectors in the basis $\boldsymbol{f}$ then we change the sign of $\det(M_{\boldsymbol{e},\boldsymbol{f}})$. Notice also that if we change the sign of one vector in $\boldsymbol{f}$ we change the sign of $\det(M_{\boldsymbol{e},\boldsymbol{f}})$. So, changing orthonormal coordinate systems can be performed with matrices in $SO(n)$ if we give ourselves the freedom of interchanging two axes or of changing the direction of one axis.

- The above observation is of interest because coordinate changes performed with matrices in $SO(n)$ are displacements. Such transformations are close to our intuition since they correspond to the usual movements that we do/see in our surroundings.

- The next statements show how a symmetric operator (a linear map with a symmetric matrix) can be diagonalized with matrices in $SO(n)$.

Lemma 6.10. The characteristic polynomial of a symmetric matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ has only real roots.

Theorem 6.11. (Spectral Theorem) Let $T : \mathbb{V}^n \to \mathbb{V}^n$ be a symmetric operator. There is an orthonormal basis of $\mathbb{V}^n$ with respect to which the matrix of T is diagonal.

Theorem 6.12. For every real symmetric matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ there is an orthogonal matrix $M \in O(n)$ such that $M^{-1}AM$ is diagonal.

Proposition 6.13. Let $T : \mathbb{V}^n \to \mathbb{V}^n$ be a symmetric operator on a Euclidean vector space. If λ and μ are two distinct eigenvalues of T then every eigenvector with eigenvalue λ is orthogonal to every eigenvector with eigenvalue μ.

CHAPTER 7

Quadratic curves (conics)

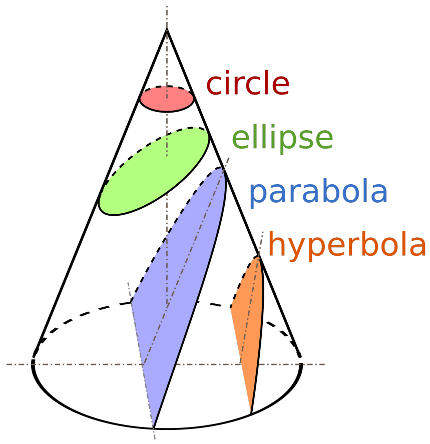
Contents

Here we consider objects in $\mathbb{E}^2$ with respect to an orthonormal coordinate system $Oxy = (O, \mathbf{i}, \mathbf{j})$.

Definition. A *quadratic curve* (or *conic*) in $\mathbb{E}^2$ is a curve described by a quadratic equations

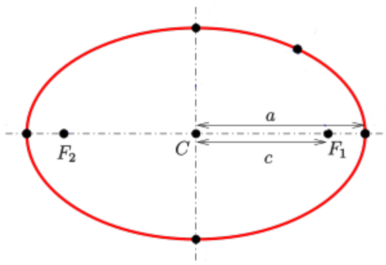
$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

for some $a, b, c, d, e, f \in \mathbb{R}$.



7.1 Ellipse

7.1.1 Geometric description



Definition. An *ellipse* is the geometric locus of points in $\mathbb{E}^2$ for which the sum of the distances from two given points, the *focal points*, is constant.

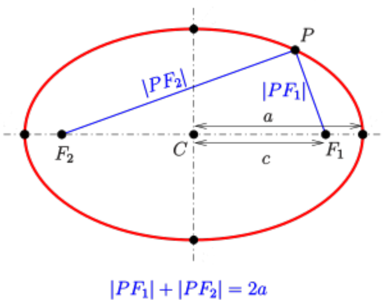
7.1.2 Canonical equation - global description

In general we can describe a conic with a so-called canonical equation. Such an equation is with respect to a well chosen coordinate system.

Proposition 7.1. Let F_1 and F_2 be two points in $\mathbb{E}^2$ and let a be a positive real scalar. Choose the coordinate system $Oxy = (O, \mathbf{i}, \mathbf{j})$ such that F_1 and F_2 are on the Ox axis, such that $\overrightarrow{F_2F_1}$ has the same direction as $\mathbf{i}$ and such that O is the midpoint of $[F_1F_2]$. With these choices, the ellipse with focal points F_1 and F_2 for which the sum of distances from the focal points is $2a$ has an equation of the form

$$\mathcal{E}_{a,b} : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{7.1}$$

for some positive scalar $b \in \mathbb{R}$. We denote this ellipse by $\mathcal{E}_{a,b}$.



- The equation (7.1) is called the *canonical equation of the ellipse* $\mathcal{E}_{a,b}$. Clearly, with respect to some other coordinate system, the same ellipse will have a different equation.
- If $2c$ denotes the distance between F_1 and F_2 then $b^2 = a^2 - c^2$.
- The intersections of $\mathcal{E}_{a,b}$ with the coordinate axes are the points $(\pm a, 0)$ and $(0, \pm b)$.
- The numerical quantity
$$\varepsilon = \frac{c}{a} = \sqrt{1 - \frac{b^2}{a^2}} \in [0,1)$$
is called the *eccentricity* of the ellipse $\mathcal{E}_{a,b}$. It measures how flat or how round the ellipse is.
- The canonical equation shows that $M(x_M, y_M) \in \mathcal{E}_{a,b}$ if and only if $(\pm x_M, \pm y_M) \in \mathcal{E}_{a,b}$.

7.1.3 Parametric equations - local description

Parametric equations are never unique. Depending on what your intentions are you may prefer one over the other.

The equation (7.1) allows us to express y in terms of x:

$$y(x) = \pm \frac{b}{a}\sqrt{a^2 - x^2}.$$

This gives a partial parametrization of $\mathcal{E}_{a,b}$. For the 'northern part' we have the parametrization

$$\phi : [-a, a] \to \mathbb{E}^2 \quad \text{given by} \quad \phi(x) = \big(x, y(x)\big) = \Big(x, \frac{b}{a}\sqrt{a^2 - x^2}\Big).$$

This is the graph of the function

$$y(x) = \pm\frac{b}{a}\sqrt{a^2 - x^2} \quad \text{for which} \quad y'(x) = \frac{-bx}{a\sqrt{a^2 - x^2}} \quad \text{and} \quad y''(x) = \frac{ab}{(x-a)(x+a)\sqrt{a^2 - x^2}}.$$

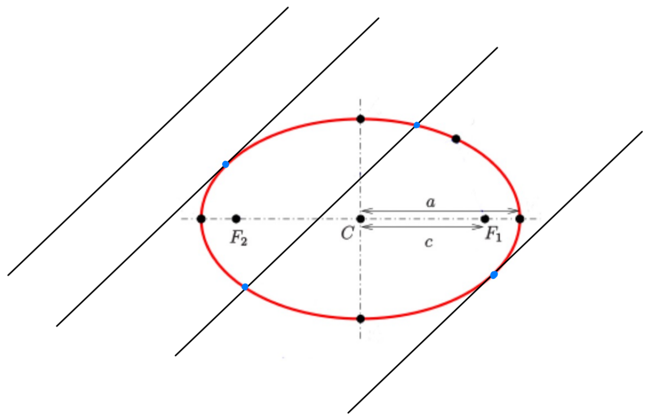
Thus, we can use the known methods to verify the monotony and the convexity of $y(x)$ which describes this part of the ellipse.

A second way of parametrizing the ellipse $\mathcal{E}_{a,b}$ is with

$$\phi : \mathbb{R} \to \mathbb{E}^2 \quad \text{defined by} \quad \phi(t) = (a\cos(t), b\sin(t)).$$

It is easy to check using equation (7.1) that this is a parametrization.

7.1.4 Relative position of a line



Consider the canonical equation of the ellipse $\mathcal{E}_{a,b}$. Let ℓ be a line with equation $y = kx + m$ (relative to the same coordinate system). The intersection of the two objects is the set of points with coordinates solutions to the system

$$\begin{cases} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \\ y = kx + m \end{cases} \Leftrightarrow \begin{cases} \frac{x^2}{a^2} + \frac{(kx+m)^2}{b^2} = 1 \\ y = kx + m \end{cases}.$$

The solutions to this system are $(x, y) = (x, kx + m)$ where x is a solution to the first equation. So let us discus that equation:

$$(b^2 + a^2k^2)x^2 + 2kma^2x + a^2(m^2 - b^2) = 0.$$

This is a quadratic equation in x since a, b, k, m are fixed. The discriminant of this equation is

$$\Delta = 4k^2m^2a^4 - 4a^2(m^2 - b^2)(b^2 + a^2k^2) = 4a^2b^2(a^2k^2 + b^2 - m^2).$$

So, the number of solutions is controlled by $a^2k^2 + b^2 - m^2$:

- $-\sqrt{a^2k^2+b^2} < m < \sqrt{a^2k^2+b^2}$ in which case ℓ intersects $\mathcal{E}_{a,b}$ in two distinct points.
- $m = \pm\sqrt{a^2k^2+b^2}$ in which case ℓ intersects $\mathcal{E}_{a,b}$ in a unique point. Such a point is a *double intersection point* because it is obtained as a double solution to the algebraic equation. For these two values of m, the line $\ell : y = kx + m$ is tangent to the ellipse. Therefore, if a slope k is given, there are two tangent lines to the ellipse:
$$y = kx \pm \sqrt{a^2k^2+b^2}.$$
- $m < \sqrt{a^2k^2+b^2}$ or $m > \sqrt{a^2k^2+b^2}$ in which case there is no intersection point between ℓ and $\mathcal{E}_{a,b}$.

7.1.5 Tangent line in a given point - algebraic

Consider an ellipse and a line:
$$\mathcal{E}_{a,b} : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \text{and} \quad \ell : \begin{cases} x = x_0 + tv_x \\ y = y_0 + tv_y \end{cases}.$$
which have the point (x_0, y_0) in common. When is ℓ tangent to the ellipse? If the intersection $\mathcal{E}_{a,b} \cap \ell$ has a unique point. In order to determine when this is the case, we check which points on ℓ satisfy the equaion of the ellipse:
$$\frac{(x_0 + v_x t)^2}{a^2} + \frac{(y_0 + v_y t)^2}{b^2} = 1.$$
The parameters t satisfying the above equations corrspond to points on ℓ which lie on $\mathcal{E}_{a,b}$. The equation is equivalent to
$$\left(\frac{v_x^2}{a^2} + \frac{v_y^2}{b^2}\right)t^2 + 2\left(\frac{x_0v_x}{a^2} + \frac{y_0v_y}{b^2}\right)t + \underbrace{\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} - 1}_{=0} = 0$$
In order for ℓ to be tangent to $\mathcal{E}_{a,b}$, there needs to be a unique solution t to the above equation. Since $t = 0$ is obviously a solution, this needs to be the *only* solution. In other words, $t = 0$ should be a double solution. For this to happen we must have
$$\frac{x_0v_x}{a^2} + \frac{y_0v_y}{b^2} = 0 \quad \Leftrightarrow \quad \mathbf{n} \cdot \mathbf{v} = 0$$
where $\mathbf{n} = \mathbf{n}(\frac{x_0}{a^2}, \frac{y_0}{b^2})$. Thus, ℓ is tangent to the ellipse if and only if the the vector $\mathbf{n}$ is orthogonal to ℓ, i.e. if and only if $\mathbf{n}$ is a normal vector for ℓ. It follows that ℓ is tangent to $\mathcal{E}_{a,b}$ in the point $(x_0, y_0) \in \mathcal{E}_{a,b}$ if and only if it satisfies the Cartesian equation:
$$\ell : \frac{x_0}{a^2}(x - x_0) + \frac{y_0}{b^2}(y - y_0) = 0.$$
We call this line the *tangent line to $\mathcal{E}_{a,b}$ at the point $(x_0, y_0) \in \mathcal{E}_{a,b}$* and denote it by $T_{(x_0,y_0)}\mathcal{E}_{a,b}$. Rearranging the above equation we see that:
$$T_{(x_0,y_0)}\mathcal{E}_{a,b} : \frac{x_0x}{a^2} + \frac{y_0y}{b^2} = 1. \tag{7.2}$$

7.1.6 Tangent line in a given point - via gradients

It is possible to describe the tangent line $T_{(x_0,y_0)}\mathcal{E}_{a,b}$ to $\mathcal{E}_{a,b}$ at the point $(x_0,y_0)\in\mathcal{E}_{a,b}$ using gradients. For this consider the map

$$\psi:\mathbb{E}^2\to\mathbb{R}\quad\text{defined by}\quad \psi(x,y)=\frac{x^2}{a^2}+\frac{y^2}{b^2}$$

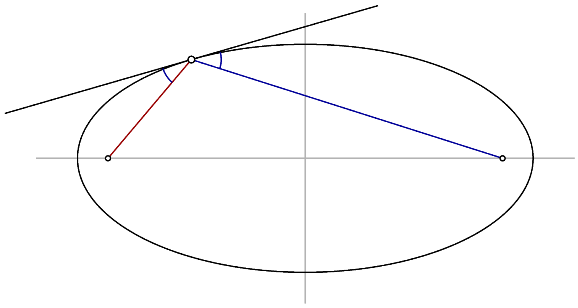
and notice that $\mathcal{E}_{a,b}=\psi^{-1}(1)$. The gradient in a point $(x_0,y_0)\in\mathcal{E}_{a,b}$ is

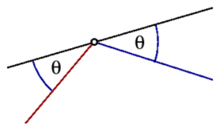
$$\nabla_{(x_0,y_0)}(\psi)=\left(2\frac{x}{a^2},2\frac{y}{b^2}\right)_{(x_0,y_0)}=2\left(\frac{x_0}{a^2},\frac{y_0}{b^2}\right).$$

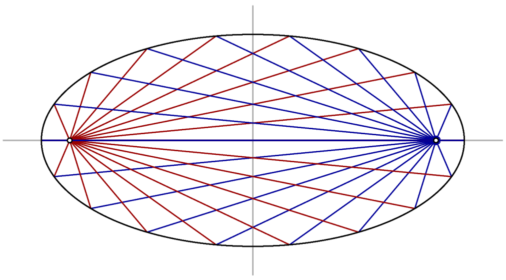
Using a parametrization $\phi:I\to\mathbb{E}^2$ of $\mathcal{E}_{a,b}$ and calculating $\partial_t\psi(\phi(t))$ with the chain rule, one shows that $\nabla_{(x_0,y_0)}(\psi)$ is orthogonal to the tangent vectors at the point (x_0,y_0). In other words, $\nabla_{(x_0,y_0)}(\psi)$ is a normal vector at the point $(x_0,y_0)\in\mathcal{E}_{a,b}$. This gives a different way of obtaining the equation (7.2).

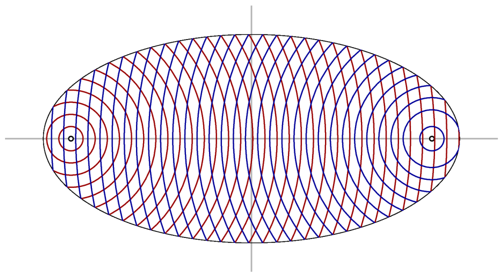
7.1.7 Applications

An ellipse has reflective properties:



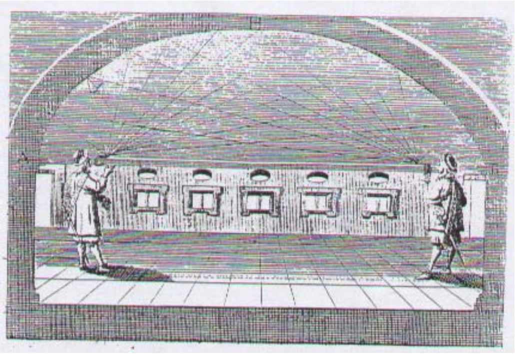






These properties are exploited in praxis.

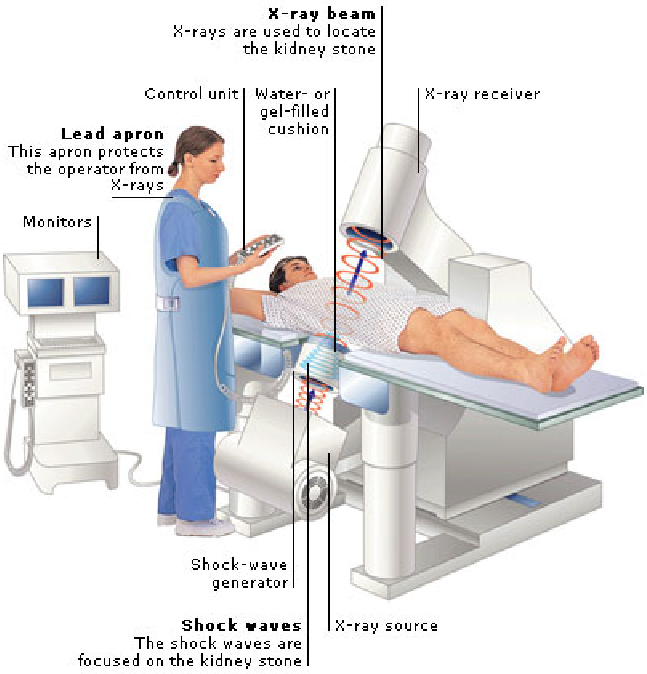
- [Elliptical rooms]



(a) Elliptical room.

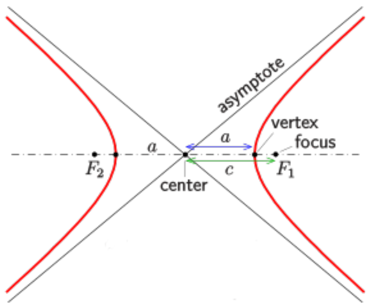
(b) The National Statuary Hall in Washington, D.C.

- [Lithotripsy]



7.2 Hyperbola

7.2.1 Geometric description



Definition. A *hyperbola* is the geometric locus of points in $\mathbb{E}^2$ for which the difference of the distances from two given points, the *focal points*, is constant.

7.2.2 Canonical equation - global description

Proposition 7.2. Let F_1 and F_2 be two points in $\mathbb{E}^2$ and let a be a positive real scalar. Choose the coordinate system $Oxy = (O, \mathbf{i}, \mathbf{j})$ such that F_1 and F_2 are on the Ox axis, such that $\overrightarrow{F_2F_1}$ has the same direction as $\mathbf{i}$ and such that O is the midpoint of $[F_1F_2]$. With these choices, the hyperbola with focal points F_1 and F_2 for which the absolute value of the difference of distances from the focal points is $2a$ has an equation of the form

$$\mathcal{H}_{a,b} : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{7.3}$$

for some positive scalar $b \in \mathbb{R}$. We denote this hyperbola by $\mathcal{H}_{a,b}$.

- The equation (7.3) is called the *canonical equation of the hyperbola* $\mathcal{H}_{a,b}$. Clearly, with respect to some other coordinate system, the same hyperbola will have a different equation.
- If $2c$ denotes the distance between F_1 and F_2 then $b^2 = c^2 - a^2$.
- The intersections of $\mathcal{H}_{a,b}$ with the coordinate axes are the points $(\pm a, 0)$.
- The numerical quantity
$$\varepsilon = \frac{c}{a} = \sqrt{1 + \frac{b^2}{a^2}} \in (1, \infty)$$
is called the *eccentricity* of the hyperbola $\mathcal{H}_{a,b}$. It measures how open or how closed the two branches of the hyperbola are.
- The canonical equation shows that $M(x_M, y_M) \in \mathcal{H}_{a,b}$ if and only if $(\pm x_M, \pm y_M) \in \mathcal{H}_{a,b}$.

7.2.3 Parametric equations - local description

Parametric equations are never unique. Depending on what your intentions are you may prefer one over the other.

The equation (7.3) allows us to express y in terms of x:

$$y(x) = \pm\frac{b}{a}\sqrt{x^2 - a^2}.$$

This gives a partial parametrization of $\mathcal{H}_{a,b}$. For the 'northern part' we have the parametrization

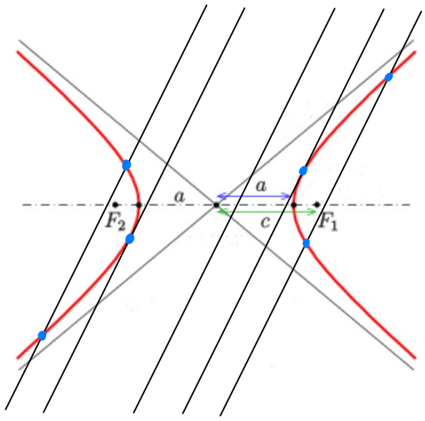
$$\phi : (-\infty, -a] \cup [a, \infty) \to \mathbb{E}^2 \quad \text{given by} \quad \phi(x) = \big(x, y(x)\big) = \Big(x, \frac{b}{a}\sqrt{x^2 - a^2}\Big).$$

This is the graph of the function

$$y(x) = \pm\frac{b}{a}\sqrt{x^2 - a^2} \quad \text{for which} \quad y'(x) = \frac{bx}{a\sqrt{x^2 - a^2}} \quad \text{and} \quad y''(x) = \frac{ab}{(a-x)(x+a)\sqrt{x^2 - a^2}}$$

Thus, we can use the known methods to verify the monotony and the convexity of $y(x)$ which describes this part of the hyperbola.

7.2.4 Relative position of a line



Consider the canonical equation of the hyperbola $\mathcal{H}_{a,b}$. Let ℓ be a line with equation $y = kx + m$ (relative to the same coordinate system). The intersection of the two objects is the set of points with coordinates solutions to the system

$$\begin{cases} \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \\ y = kx + m \end{cases} \quad \Leftrightarrow \quad \begin{cases} \frac{x^2}{a^2} - \frac{(kx+m)^2}{b^2} = 1 \\ y = kx + m \end{cases}.$$

The solutions to this system are $(x, y) = (x, kx + m)$ where x is a solution to the first equation. So let us discus that equation:

$$(b^2 - a^2k^2)x^2 - 2kma^2x - a^2(m^2 + b^2) = 0 \tag{7.4}$$

This is a quadratic equation in x since a, b, k, m are fixed. The discriminant of this equation is

$$\Delta = 4k^2m^2a^4 + 4a^2(m^2 + b^2)(b^2 - a^2k^2) = 4a^2b^2(m^2 + b^2 - a^2k^2).$$

So, the number of solutions is controlled by $m^2 + b^2 - a^2k^2$ ... *if* the equation is quadratic. Suppose equation (7.4) is quadratic, i.e. $b^2 - a^2k^2 \neq 0$.

- $m < \sqrt{a^2k^2 - b^2}$ or $m > \sqrt{a^2k^2 - b^2}$ in which case ℓ intersects $\mathcal{H}_{a,b}$ in two distinct points.

- $m = \pm\sqrt{a^2k^2 - b^2}$ in which case ℓ intersects $\mathcal{H}_{a,b}$ in a unique point. Such a point is a *double intersection point* because it is obtained as a double solution to the algebraic equation. For these two values of m, the line $\ell : y = kx + m$ is tangent to the hyperbola. Therefore, if a slope k is given such that $b^2 - a^2k^2 \neq 0 \Leftrightarrow k \neq \pm\frac{b}{a}$, there are two tangent lines to the hyperbola:

$$y = kx \pm \sqrt{a^2k^2 - b^2}.$$

- $-\sqrt{a^2k^2 - b^2} < m < \sqrt{a^2k^2 - b^2}$ in which case there is no intersection point between ℓ and $\mathcal{H}_{a,b}$.

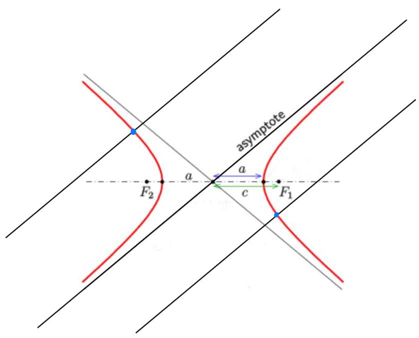
Suppose equation (7.4) is not quadratic, i.e. $b^2 - a^2k^2 = 0$ and the equation is

$$-2kma^2x - a^2(m^2 + b^2) = 0$$

Notice that $k \neq 0$ and $a \neq 0$ and that $k = \pm\frac{b}{a}$. We have two cases:

- $m \neq 0$, hence the unique solution $x = -\frac{m^2+b^2}{2a^2}$ which corresponds to a unique intersection point. In this cases it is a *simple intersection point*, it corresponds to a simple solution of an algebraic equation (not a double solution).

- $m = 0$ in which case there is no intersection point and ℓ is either $y = \frac{b}{a}x$ or $y = -\frac{b}{a}x$. These are the two asymptotes of the hyperbola $\mathcal{H}_{a,b}$. One can check with the parametrization in the previous section that these two lines really are asymptotes.



7.2.5 Tangent line in a given point

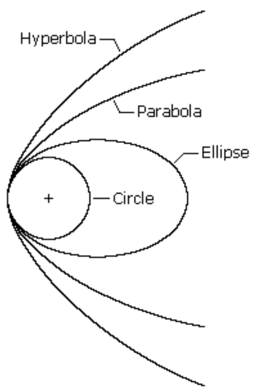
The *tangent line to $\mathcal{H}_{a,b}$ at the point* $(x_0, y_0) \in \mathcal{H}_{a,b}$ has an equation of the form

$$T_{(x_0,y_0)}\mathcal{H}_{a,b} : \frac{x_0 x}{a^2} - \frac{y_0 y}{b^2} = 1. \tag{7.5}$$

This can be deduced either with the algebraic method or via the gradient as in the case of the ellipse.

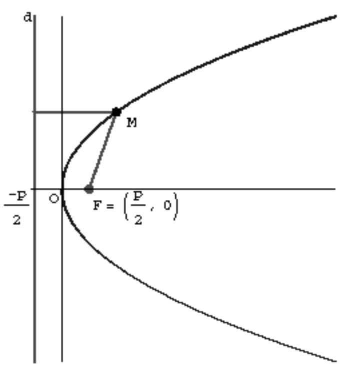
7.2.6 Applications

- [Two body problem] Newton generalized Kepler's laws to apply to any two bodies orbiting each other. The shape of an orbit is a conic section with the center of mass at one focus (first law of orbital motion).



7.3 Parabola

7.3.1 Geometric description



Definition. A *parabola* is the geometric locus of points in $\mathbb{E}^2$ for which the the distances from a given point, the *focal point*, equals the distance to a given line, the *directrix*.

7.3.2 Canonical equation - global description

Proposition 7.3. Let F be a point, let d be a line in $\mathbb{E}^2$ and let p be a positie real scalar. Choose the coordinate system $Oxy = (O, \mathbf{i}, \mathbf{j})$ such that F lies on the Ox axis, such that the Ox axis is orthogonal to d, the origin is at equal distance from d and F and the vector $\mathbf{i}$ has the same direction as $\overrightarrow{OF}$. With these choices, the parabola with focal point F and directrix d for which the $d(F,d) = p$ has an equation of the form

$$\mathcal{P}_p : y^2 = 2px \tag{7.6}$$

We denote this parabola by $\mathcal{P}_p$.

- The equation (7.6) is called the *canonical equation of the parabola* $\mathcal{P}_p$. Clearly, with respect to some other coordinate system, the same parabola will have a different equation.
- The focal point is $F(\frac{p}{2}, 0)$ and the directrix has equation $d : x = -\frac{p}{2}$.
- The intersections of $\mathcal{P}_p$ with the coordinate axes is the point $(0,0)$.
- The canonical equation shows that $M(x_M, y_M) \in \mathcal{P}_p$ if and only if $(x_M, \pm y_M) \in \mathcal{P}_p$.

7.3.3 Parametric equations - local description

Parametric equations are never unique. Depending on what your intentions are you may prefer one over the other.

The equation (7.6) allows us to express y in terms of x:

$$y(x) = \pm\sqrt{2px}.$$

This gives a partial parametrization of $\mathcal{P}_p$. For the 'northern part' we have the parametrization

$$\phi : [0, \infty) \to \mathbb{E}^2 \quad \text{given by} \quad \phi(x) = \big(x, y(x)\big) = \big(x, \sqrt{2px}\big).$$

This is the graph of the function

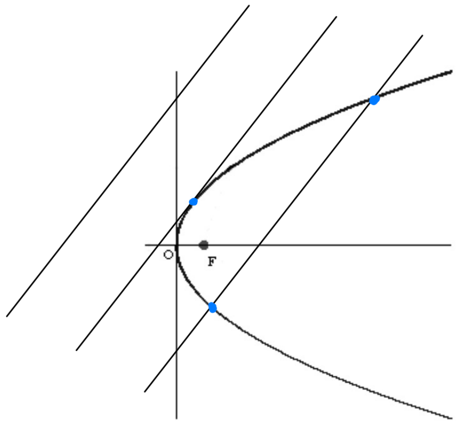
$$y(x) = \sqrt{2px} \quad \text{for which} \quad y'(x) = \frac{\sqrt{2p}}{2\sqrt{x}} \quad \text{and} \quad y''(x) = -\frac{\sqrt{2p}}{4x^{3/2}}$$

Thus, we can use the known methods to verify the monotony and the convexity of $y(x)$ which describes this part of the parabola.

We can in fact parametrize the whole parabola if we express x in terms of y, which is another way of reading equation (7.6). We then have the parametrization

$$\phi : \mathbb{R} \to \mathbb{E}^2 \quad \text{given by} \quad \phi(x) = \big(x(y), y\big) = \Big(\frac{y^2}{2p}, y\Big).$$

7.3.4 Relative position of a line



Consider the canonical equation of the parabola $\mathcal{P}_p$. Let ℓ be a line with equation $y = kx + m$ (relative to the same coordinate system). The intersection of the two objects is the set of points with coordinates solutions to the system

$$\begin{cases} y^2 = 2px \\ y = kx + m \end{cases} \Leftrightarrow \begin{cases} (kx+m)^2 = 2px \\ y = kx + m \end{cases}.$$

The solutions to this system are $(x, y) = (x, kx + m)$ where x is a solution to the first equation. So let us discus that equation:

$$k^2x^2 + 2(km - p)x + m^2 = 0 \tag{7.7}$$

This is a quadratic equation in x since p, k, m are fixed. The discriminant of this equation is

$$\Delta = 4p(p - 2km).$$

So, the number of solutions is controlled by $p - 2km$:

- $km < p/2$ in which case ℓ intersects $\mathcal{P}_p$ in two distinct points.
- $km = p/2$ in which case ℓ intersects $\mathcal{P}_p$ in a unique point. Such a point is a *double intersection point* because it is obtained as a double solution to the algebraic equation. For this value of m, the line $\ell : y = kx + m$ is tangent to the parabola. Therefore, if a slope k is given, there is one tangent line to the parabola:
$$y = kx + \frac{p}{2k}.$$
- $km > p/2$ in which case there is no intersection point between ℓ and $\mathcal{P}_p$.

7.3.5 Tangent line in a given point

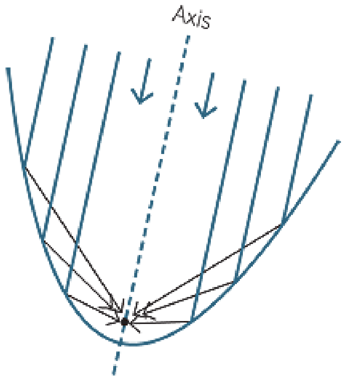
The *tangent line to $\mathcal{P}_p$ at the point* $(x_0, y_0) \in \mathcal{P}_p$ has an equation of the form

$$T_{(x_0,y_0)}\mathcal{P}_p : yy_0 = p(x + x_0) \tag{7.8}$$

This can be deduced either with the algebraic method or via the gradient as in the case of the ellipse.

7.3.6 Applications

A parabola has reflective properties:



These properties are used for lenses, parabolic reflectors, satellite dishes, etc.

CHAPTER 8

Hyperquadrics

Contents

All equations are given with respect to an orthonormal coordinate system $\mathcal{K} = (O, \mathbf{e}_1, \dots, \mathbf{e}_n)$.

8.1 Hyperquadrics

Definition. A *hyperquadric* $\mathcal{Q}$ in $\mathbb{E}^n$ is a the set of points whose coordinates satisfy a quadratic equation, i.e.

$$\mathcal{Q} : \sum_{i,j=1}^{n} a_{ij} x_i x_j + \sum_{i}^{n} b_i x_i + c = 0, \tag{8.1}$$

with respect to some coordinate system.

- Hyperquadrics in $\mathbb{E}^2$ are called *conic sections* (see previous chapter).

- The following lectures are dedicated to hyperquadrics in $\mathbb{E}^3$. In dimension 3 we generally refer to them as *quadrics*.
- Notice that Equation (8.1) is equivalent to

$$\mathcal{Q}: \sum_{i,j=1}^{n} q_{ij}x_i x_j + \sum_{i}^{n} b_i x_i + c = 0, \tag{8.2}$$

where $q_{ii} = a_{ii}$ and $q_{ij} = q_{ji} = \frac{a_{ij}+a_{ji}}{2}$. The matrix $Q = (q_{ij})$ is symmetric and we call it the *symmetric matrix associated to Equation* (8.2) *of the quadric* $\mathcal{Q}$.

- The matrix Q defines a homogeneous polynomial of degree 2 in the above equation.
- Notice that Equation (8.2) can be rearranged in matrix form as follows

$$\mathcal{Q}: \mathbf{x}^t \cdot Q \cdot \mathbf{x} + \mathbf{b}^t \cdot \mathbf{x} + c = 0 \tag{8.3}$$

where $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$.

8.2 Reducing to the canonical form

- Let $\mathcal{Q}$ be a hyperquadric described by Equation (8.2) with respect to the coordinate system $\mathcal{K} = (O, \boldsymbol{e})$ where $\boldsymbol{e} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$.
- Let Q be the matrix associated to Equation (8.2) of $\mathcal{Q}$.
- [Step 1 - Rotation] By the Spectral Theorem, there is an orthonormal basis $\boldsymbol{e}'$ of eigenvectors for Q which diagonalizes Q. Changing the coordinate system from $\mathcal{K} = (O, \boldsymbol{e})$ to $\mathcal{K}(O, \boldsymbol{e}')$, the equation of the hyperquadric becomes

$$\mathcal{Q}: \mathbf{y}^t \cdot D \cdot \mathbf{y} + \mathbf{v}^t \cdot \mathbf{y} + c = 0 \quad \text{where} \quad D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \tag{8.4}$$

and where $\mathbf{y} = (y_1, \dots, y_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$. This change of coordinates consists in replacing $\mathbf{x}$ by $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}\mathbf{y}$ since $\mathbf{x} = \mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}\mathbf{y}$.

- Since the two bases $\boldsymbol{e}$ and $\boldsymbol{e}'$ are orthonormal, the matrix $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}$ is an orthogonal matrix, i.e. $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} \in O(n)$.
- Since Q is symmetric, the eigenvalues $\lambda_1, \dots, \lambda_n$ are real. Hence, eventually after permuting the basis vectors in $\boldsymbol{e}'$ we may assume that $\lambda_1, \dots, \lambda_p > 0$, $\lambda_{p+1}, \dots, \lambda_r < 0$ and $\lambda_{r+1}, \dots, \lambda_n = 0$ where r is the rank of Q. This permutation corresponds to changing $\boldsymbol{e}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ to $\boldsymbol{e}'' = (\mathbf{e}_{\pi(1)}, \dots, \mathbf{e}_{\pi(n)})$ for some permutation π of $\{1, \dots, n\}$. The base change matrix $\mathrm{M}_{\boldsymbol{e}',\boldsymbol{e}''}$ is again orthogonal.

- So far, we have a change of coordinates from $\mathcal{K} = (O, \boldsymbol{e})$ to $\mathcal{K} = (O, \boldsymbol{e}'')$ given by the base change matrix $M_{\boldsymbol{e},\boldsymbol{e}''} = \mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}\mathrm{M}_{\boldsymbol{e}',\boldsymbol{e}''} \in O(n)$.
- Notice that $\det(\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}''}) = 1$ if $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}''} \in SO(n)$ and $\det(\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}''}) = -1$ if $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}''}$ is not special orthogonal. In the latter case, the matrix $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}''}$ changes the orientation of the basis $\boldsymbol{e}''$. So, if we replace one vector in $\boldsymbol{e}''$ by minus that vector, for example $\mathbf{e}''_1 \leftrightarrow -\mathbf{e}''_1$, then $\boldsymbol{e}''$ has the same orientation as $\boldsymbol{e}$ and $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} \in SO(n)$ (See Examples below). We conclude that we may choose $\boldsymbol{e}''$ such that $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}''} \in SO(n)$.
- Writing out the equation of $\mathcal{Q}$ in $\mathcal{K}''$ we have:
$$\mathcal{Q} : \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + v_1 y_1 + v_2 y_2 + \cdots + v_n y_n + c = 0. \tag{8.5}$$
- [Step 2 - translation] Notice that $r > 0$, i.e. there is at least one eigenvalue distinct from 0 otherwise $Q = 0_n$ and (8.2) is not a quadratic polynomial.
- If $v_1 \neq 0$ then
$$\lambda_1 y_1^2 + v_1 y_1 = \lambda_1\Big(y_1^2 + 2\frac{v_1}{2\lambda_1}y_1\Big) = \lambda_1\Big(y_1 + \frac{v_1}{2\lambda_1}\Big)^2 - \frac{v_1^2}{4\lambda_1^2}.$$
- Thus, if for $i \in \{1, \ldots, r\}$ we let $z_i = y_i + \frac{v_i}{2\lambda_i}$ and $z_i = y_i$ for $i > r$ then Equation (8.5) becomes
$$\mathcal{Q} : \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_r z_r^2 + v_{r+1} z_{r+1} + \cdots + v_n z_n = k \tag{8.6}$$
for some $k \in \mathbb{R}$. This change of variables corresponds to a change of coordinates from $\mathcal{K}'' = (O, \boldsymbol{e}'')$ to $\mathcal{K}''' = (O', \boldsymbol{e}'')$ given by the translation
$$\begin{bmatrix} z_1 \\ \vdots \\ z_r \\ z_{r+1} \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} \frac{v_1}{2\lambda_1} \\ \vdots \\ \frac{v_r}{2\lambda_r} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$
- Notice that so far we changed the coordinates from $\mathcal{K}$ to $\mathcal{K}''$ with an orthogonal transformation in $SO(n)$ and from $\mathcal{K}''$ to $\mathcal{K}'''$ with a translation. These are isometries, so the composition of these two transformations is an isometry.
- It is possible to customize and simplify the equation (8.6) of $\mathcal{Q}$ further, by making other changes of coordinates. However these will in general not correspond to isometries (See the discussion in Section 8.3.5).
- We *reduced* Equation (8.2) to Equation (8.6). Equation (8.6) is not yet the canonical form but it is an important step towards the canonical form. We may refer to Equation (8.6) as the *intermediate canonical form*.
- In what follows we look at what more can be done if we restrict to conics and quadrics, i.e. if we look at hyperquadrics in $\mathbb{E}^2$ and $\mathbb{E}^3$.

8.3 Classification of conics - dimension 2

8.3.1 Isometric classification

- A hyperquadric in $\mathbb{E}^2$ is a curve given by an equation of the form

$$\mathcal{C} : q_{11}x^2 + 2q_{12}xy + q_{22}y^2 + b_1x + b_2y + c = 0. \tag{8.7}$$

- From the above discussion, we may apply a rotation and a translation to change the coordinate system such that Equation (8.4) becomes

$$\mathcal{C} : \lambda_1 x^2 + \lambda_2 y^2 = k \quad \text{or} \quad \mathcal{C} : \lambda_1 x^2 + v_2 y = k. \tag{8.8}$$

  where $\lambda_1 > 0$ (if $\lambda_1, \lambda_2 < 0$, multiply the whole equation by -1).

- If $\lambda_2 > 0$ and $k = 0$, then the equation has only $(0,0)$ as solution, in this case the curve is degenerate to a point, the origin.

- If $\lambda_2 > 0$ and $k < 0$, then there are no real solutions to the equation.

- If $\lambda_2 > 0$ and $k > 0$, after dividing by k the equation becomes

$$\frac{x^2}{\frac{k}{\lambda_1}} + \frac{y^2}{\frac{k}{\lambda_2}} = 1$$

  which is the equation of an ellipse. The ellipse is in canonical form if $\frac{k}{\lambda_1} > \frac{k}{\lambda_2}$. If this is not the case, we need to do one additional change of coordinates by interchanging x with y. This is again on orthogonal transformation (see Examples below).

- If $\lambda_2 < 0$ and $k = 0$, then the equation becomes

$$(\sqrt{\lambda_1}x - \sqrt{-\lambda_2}y)(\sqrt{\lambda_1}x + \sqrt{-\lambda_2}y) = 0.$$

  This is the union of two lines.

- If $\lambda_2 < 0$ and $k < 0$, we may multiply the whole equation by -1 and interchange the role of λ_1 and λ_2. Doing so, we need to do one additional change of coordinates by interchanging x with y. This is again on orthogonal transformation (see Examples below).

- If $\lambda_2 < 0$ and $k > 0$, after dividing by k the equation becomes

$$\frac{x^2}{\frac{k}{\lambda_1}} - \frac{y^2}{\frac{k}{-\lambda_2}} = 1$$

  which is a the equation of a hyperbola.

- If $\lambda_2 = 0$ we have the equation

$$\mathcal{C} : \lambda_1 x^2 + v_2 y = k.$$

  This is the case if and only if $\operatorname{rank}(Q) = 1$.

- If $v_2 = 0$ and $k \geq 0$ then we have two lines described by the equation $(\sqrt{\lambda_1}x - \sqrt{k})(\sqrt{\lambda_1}x + \sqrt{k}) = 0$. If $k = 0$ this is a double-line, $x^2 = 0$.
- If $v_2 = 0$ and $k < 0$ then we have no solutions.
- If $v_2 \neq 0$ then, dividing by v_2, the equation becomes

$$\frac{\lambda_1}{|v_1|}x^2 = \frac{k}{|v_1|} - y$$

and we may change the coordinates $(x, \frac{k}{|v_1|} - y) \to (x, y)$ such that the equation becomes

$$x^2 = \frac{|v_1|}{\lambda_1}y.$$

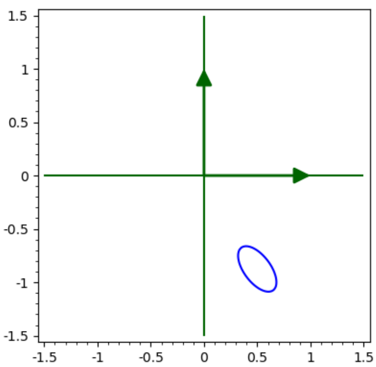
This change of coordinates corresponds to a reflection and a translation (again an isometric change of coordinates) and we recognize the equation of a parabola with parameter $p = \frac{|v_1|}{2\lambda_1}$. However, here again, we don't yet have the canonical form of a parabola. In order to obtain $\mathcal{P}_p$ we need to interchange the x-axis with the y-axis.

- [Conclusion] Starting with an equation of the form (8.7), we may use rotations, translations and reflections to recognize that we obtain either
  - degenerate cases: two lines, double lines, points, or
  - non-degenerate cases:

$$\mathcal{E}_{a,b} : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \text{or} \quad \mathcal{H}_{a,b} : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad \text{or} \quad \mathcal{P}_p : y^2 = 2px.$$

  So, the curves described by an equation of the form (8.7) are conic sections.

8.3.2 Example - ellipse



Consider the curve with equation

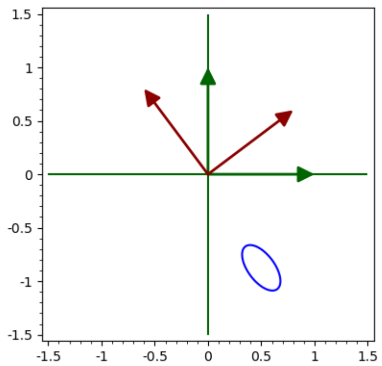
$$\mathcal{C} : 73x^2 + 72xy + 52y^2 - 10x + 55y + 25 = 0. \tag{8.9}$$

It is the ellipse in the above image, however, a priori it is not at all clear that $\mathcal{C}$ is an ellipse. The symmetric matrix associated to this equation is

$$Q = \begin{bmatrix} 73 & 36 \\ 36 & 52 \end{bmatrix}$$

and in matrix form Equation (8.9) becomes

$$\mathcal{C} : \begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} 73 & 36 \\ 36 & 52 \end{bmatrix}}_{Q} \begin{bmatrix} x \\ y \end{bmatrix} + \underbrace{\begin{bmatrix} -10 & 55 \end{bmatrix}}_{b} \begin{bmatrix} x \\ y \end{bmatrix} + 25 = 0. \tag{8.10}$$



The eigenvalues of Q are $\lambda_1 = 100$ and $\lambda_2 = 25$. An eigenvector for the eigenvalue λ_1 is $(4,3)$ and an eigenvector for the eigenvalue λ_2 is $(3,-4)$. These two vectors form an orthogonal basis, thus, an orthonormal basis is $\boldsymbol{e}' = (\mathbf{e}'_1(4/5,3/5), \mathbf{e}'_2(3/5,-4/5))$. The base change matrix from $\boldsymbol{e}'$ to $\boldsymbol{e}$ is

$$\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}.$$

We know that this matrix is orthogonal, which in this example is easy to check directly. In particular $\mathrm{M}^{-1}_{\boldsymbol{e},\boldsymbol{e}'} = \mathrm{M}^{t}_{\boldsymbol{e},\boldsymbol{e}'}$. Moreover, the determinant is -1 which means that $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}$ is not a direct isometrie, i.e. $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} \in O(2) \setminus SO(2)$. If we change the direction of the second eigenvector, we still have an eigenvector for the eigenvalue λ_2 but with respect to the basis $\boldsymbol{e}' = (\mathbf{e}'_1(4/5,3/5), \mathbf{e}'_2(-3/5,4/5))$ we have

$$\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} = \frac{1}{5}\begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix} \in SO(2).$$

Changing the coordinate system from $\mathcal{K}$ to $\mathcal{K}' = (O, \boldsymbol{e}')$, Equation (8.9) becomes

$$\mathcal{C} : \begin{bmatrix} x' & y' \end{bmatrix} M^{t}_{\boldsymbol{e},\boldsymbol{e}'} Q M_{\boldsymbol{e},\boldsymbol{e}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + b M_{\boldsymbol{e},\boldsymbol{e}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + 25 = 0$$

which one calculates to be

$$\mathcal{C}:\begin{bmatrix}x' & y'\end{bmatrix}\underbrace{\begin{bmatrix}100 & 0\\ 0 & 25\end{bmatrix}}_{Q'}\begin{bmatrix}x'\\ y'\end{bmatrix}+\underbrace{\begin{bmatrix}-25 & 50\end{bmatrix}}_{b'}\begin{bmatrix}x'\\ y'\end{bmatrix}+25=0$$

and we have

$$\mathcal{C}: 100x'^2+25y'^2+25x'+50y'+25=0$$

equivalently

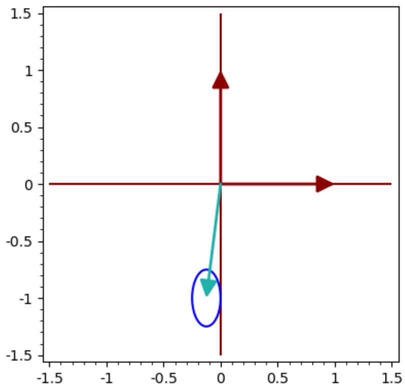
$$\mathcal{C}: 4x'^2+y'^2+x'+2y'+1=0$$

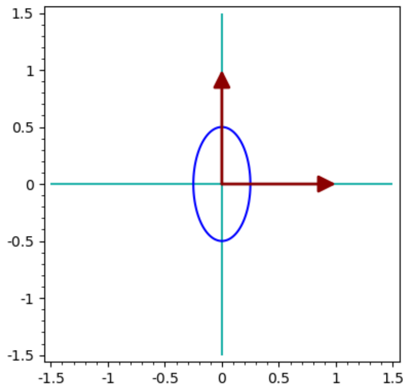
equivalently

$$\mathcal{C}: 4\Big(x'^2+\frac{x'}{4}+\frac{1}{64}\Big)-\frac{1}{16}+\Big(y'^2+2y'+1\Big)-1+1=0$$

equivalently

$$\mathcal{C}: 4\Big(x'+\frac{1}{8}\Big)^2+\Big(y'+1\Big)^2-\frac{1}{16}=0.$$





Now, let us change the coordinate system again, using a translation of vector $(-\frac{1}{8},-1)$. The new coordinate system is $\mathcal{K}''=(O'',\boldsymbol{e}'')$ where $O''=(-\frac{1}{8},-1)$ and the basis $\boldsymbol{e}''=\boldsymbol{e}'$ doesn't change. In $\mathcal{K}''$ the equation of $\mathcal{C}$ becomes

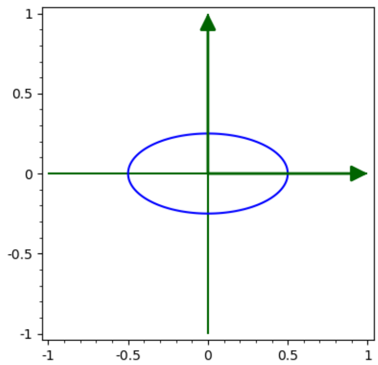
$$\mathcal{C}: 4x''^2+y''^2-\frac{1}{16}=0 \quad\Leftrightarrow\quad \frac{x''^2}{\frac{1}{64}}+\frac{y''^2}{1}-1=0.$$

Clearly, this is the equation of an ellipse. But it is not yet in canonical form because the focal points are on the y-axes. So, in order to obtain the canonical form we need to do a final coordinate change and permute the coordinate axes, for instance with

$$\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\in SO(2) \quad \text{or with} \quad \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\in O(2).$$

With any of the two transformations we obtain

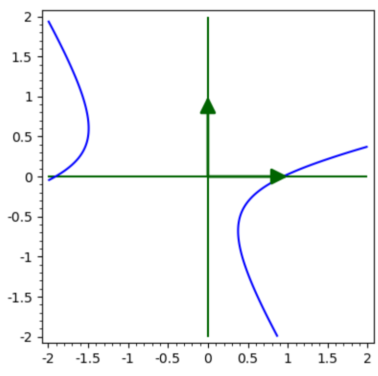
$$\mathcal{C} = \mathcal{E}_{1,\frac{1}{8}} : \frac{y'''^2}{1} + \frac{x'''^2}{\frac{1}{64}} - 1 = 0.$$



To recap:

- We changed the coordinates from $\mathcal{K}$ to $\mathcal{K}'$ with the rotation $M_{e,e'}$ of angle θ where $\cos(\theta) = \frac{4}{5}$.
- We changed the coordinates from $\mathcal{K}'$ to $\mathcal{K}''$ with a translation of vector $(-\frac{1}{8}, -1)$.
- We changed the coordinates from $\mathcal{K}''$ to $\mathcal{K}'''$ in order to interchange the variables. And we obtained $\mathcal{E}_{1,\frac{1}{8}}$.

8.3.3 Example - hyperbola



Consider the curve with equation

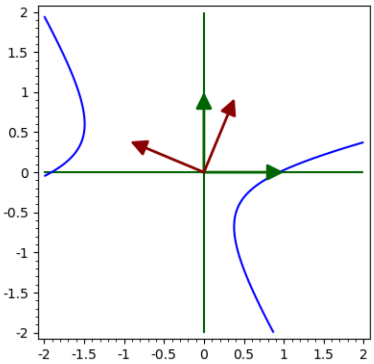
$$\mathcal{C} : -94x^2 + 360xy + 263y^2 - 91x + 221y + 169 = 0. \tag{8.11}$$

It is the hyperbola in the above image, however, a priori it is not at all clear that $\mathcal{C}$ is a hyperbola. The symmetric matrix associated to this equation is

$$Q = \begin{bmatrix} -94 & 180 \\ 180 & 263 \end{bmatrix}$$

and in matrix form Equation (8.11) becomes

$$\mathcal{C} : \begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} -94 & 180 \\ 180 & 263 \end{bmatrix}}_{Q} \begin{bmatrix} x \\ y \end{bmatrix} + \underbrace{\begin{bmatrix} -91 & 221 \end{bmatrix}}_{b} \begin{bmatrix} x \\ y \end{bmatrix} + 25 = 0. \tag{8.12}$$



The eigenvalues of Q are $\lambda_1 = 338$ and $\lambda_2 = -169$. An eigenvector for the eigenvalue λ_1 is $(5,12)$ and an eigenvector for the eigenvalue λ_2 is $(12,-5)$. These two vectors form an orthogonal basis, thus, an orthonormal basis is $\boldsymbol{e}' = (\mathbf{e}_1'(5/13, 12/13), \mathbf{e}_2'(12/13, -5/13))$. The base change matrix from $\boldsymbol{e}'$ to $\boldsymbol{e}$ is

$$\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} = \frac{1}{13}\begin{bmatrix} 5 & 12 \\ 12 & -5 \end{bmatrix}.$$

We know that this matrix is orthogonal, which in this example is easy to check directly. In particular $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}^{-1} = \mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}^{t}$. Moreover, the determinant is -1 which means that $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}$ is not a direct isometrie, i.e. $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} \in O(2) \setminus SO(2)$. If we change the direction of the second eigenvector, we still have an eigenvector for the eigenvalue λ_2 but with respect to the basis $\boldsymbol{e}' = (\mathbf{e}_1'(5/13, 12/13), \mathbf{e}_2'(-12/13, 5/13))$ we have

$$\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} = \frac{1}{13}\begin{bmatrix} 5 & -12 \\ 12 & 5 \end{bmatrix} \in SO(2).$$

Changing the coordinate system from $\mathcal{K}$ to $\mathcal{K}' = (O, \boldsymbol{e}')$, Equation (8.11) becomes

$$\mathcal{C} : \begin{bmatrix} x' & y' \end{bmatrix} M_{\boldsymbol{e},\boldsymbol{e}'}^{t} Q M_{\boldsymbol{e},\boldsymbol{e}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + b M_{\boldsymbol{e},\boldsymbol{e}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + 169 = 0$$

which one calculates to be

$$\mathcal{C}:\begin{bmatrix}x' & y'\end{bmatrix}\underbrace{\begin{bmatrix}338 & 0\\ 0 & -169\end{bmatrix}}_{Q'}\begin{bmatrix}x'\\ y'\end{bmatrix}+\underbrace{\begin{bmatrix}169 & 169\end{bmatrix}}_{b'}\begin{bmatrix}x'\\ y'\end{bmatrix}+169=0$$

and we have

$$\mathcal{C}: 338x'^2-169y'^2+169x'+169y'+169=0$$

equivalently

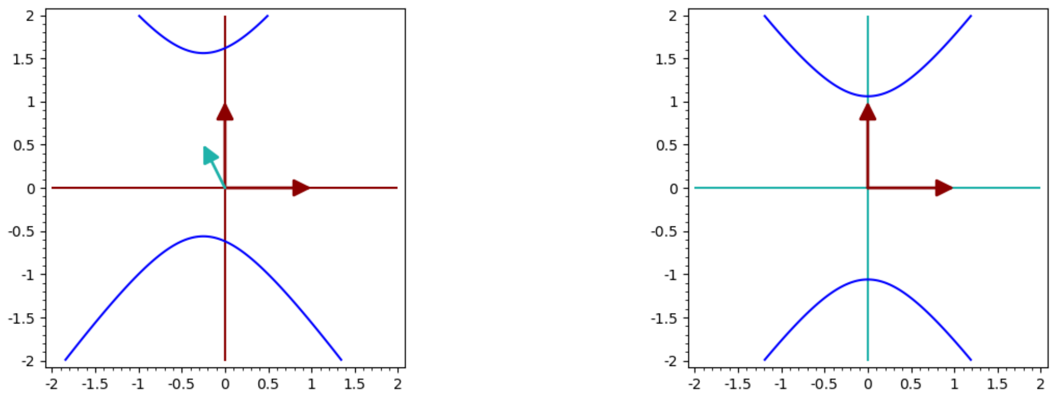
$$\mathcal{C}: 2x'^2-y'^2+x'+y'+1=0$$

equivalently

$$\mathcal{C}: 2\Big(x'^2+\frac{x'}{2}+\frac{1}{16}\Big)-\frac{1}{8}-\Big(y'^2-y'+\frac{1}{4}\Big)+\frac{1}{4}+1=0$$

equivalently

$$\mathcal{C}: 2\Big(x'+\frac{1}{4}\Big)^2+\Big(y'-\frac{1}{2}\Big)^2+\frac{9}{8}=0.$$



Now, let us change the coordinate system again, using a translation of vector $(-\frac{1}{4},\frac{1}{2})$. The new coordinate system is $\mathcal{K}''=(O'',\boldsymbol{e}'')$ where $O''=(-\frac{1}{4},\frac{1}{2})$ and the basis $\boldsymbol{e}''=\boldsymbol{e}'$ doesn't change. In $\mathcal{K}''$ the equation of $\mathcal{C}$ becomes

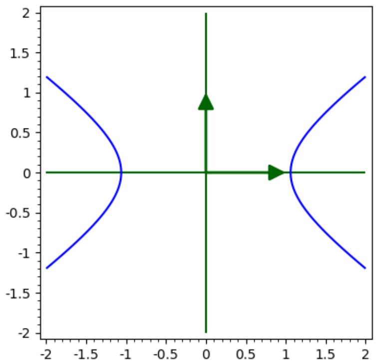
$$\mathcal{C}: 2x''^2-y''^2+\frac{9}{8}=0 \quad\Leftrightarrow\quad -\frac{x''^2}{\frac{9}{16}}+\frac{y''^2}{\frac{9}{8}}-1=0.$$

Clearly, this is the equation of a hyperbola. But it is not yet in canonical form because the focal points are on the y-axes. So, in order to obtain the canonical form we need to do a final coordinate change and permute the coordinate axes, for instance with

$$\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\in SO(2)\quad\text{or with}\quad\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\in O(2).$$

With any of the two transformations we obtain

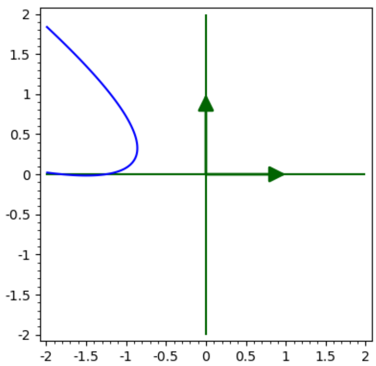
$$\mathcal{C} = \mathcal{H}_{\frac{3}{4},\frac{3}{\sqrt{2}}} : \frac{x''^2}{\frac{9}{16}} - \frac{y''^2}{\frac{9}{8}} - 1 = 0.$$



To recap:

- We changed the coordinates from $\mathcal{K}$ to $\mathcal{K}'$ with the rotation $\mathrm{M}_{e,e'}$ of angle θ where $\cos(\theta) = \frac{5}{13}$.
- We changed the coordinates from $\mathcal{K}'$ to $\mathcal{K}''$ with a translation of vector $(-\frac{1}{4}, \frac{1}{2})$.
- We changed the coordinates from $\mathcal{K}''$ to $\mathcal{K}'''$ in order to interchange the variables. And we obtained $\mathcal{H}_{\frac{3}{4},\frac{3}{\sqrt{2}}}$.

8.3.4 Example - parabola



Consider the curve with equation

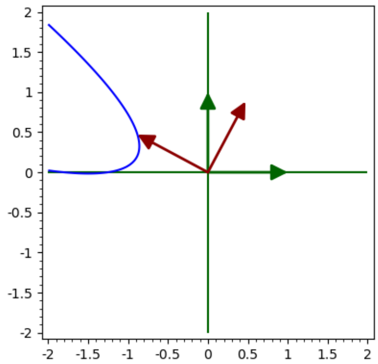
$$\mathcal{C} : 128x^2 + 480xy + 450y^2 + 391x + 119y + 289 = 0. \tag{8.13}$$

It is the parabola in the above image, however, a priori it is not at all clear that $\mathcal{C}$ is a parabola. The symmetric matrix associated to this equation is

$$Q = \begin{bmatrix} 128 & 240 \\ 240 & 225 \end{bmatrix}$$

and in matrix form Equation (8.13) becomes

$$\mathcal{C} : \begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} 128 & 240 \\ 240 & 225 \end{bmatrix}}_{Q} \begin{bmatrix} x \\ y \end{bmatrix} + \underbrace{\begin{bmatrix} 391 & 119 \end{bmatrix}}_{b} \begin{bmatrix} x \\ y \end{bmatrix} + 25 = 0. \tag{8.14}$$



The eigenvalues of Q are $\lambda_1 = 578$ and $\lambda_2 = 0$. An eigenvector for the eigenvalue λ_1 is $(8, 15)$ and an eigenvector for the eigenvalue λ_2 is $(15, -8)$. These two vectors form an orthogonal basis, thus, an orthonormal basis is $\boldsymbol{e}' = (\mathbf{e}_1'(8/17, 15/17), \mathbf{e}_2'(15/17, -8/17))$. The base change matrix from $\boldsymbol{e}'$ to $\boldsymbol{e}$ is

$$\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} = \frac{1}{17}\begin{bmatrix} 8 & 15 \\ 15 & -8 \end{bmatrix}.$$

We know that this matrix is orthogonal, which in this example is easy to check directly. In particular $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}^{-1} = \mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}^{t}$. Moreover, the determinant is -1 which means that $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}$ is not a direct isometrie, i.e. $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} \in O(2) \setminus SO(2)$. If we change the direction of the second eigenvector, we still have an eigenvector for the eigenvalue λ_2 but with respect to the basis $\boldsymbol{e}' = (\mathbf{e}_1'(8/17, 15/17), \mathbf{e}_2'(-15/17, 8/17))$ we have

$$\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'} = \frac{1}{17}\begin{bmatrix} 8 & -15 \\ 15 & 8 \end{bmatrix} \in SO(2).$$

Changing the coordinate system from $\mathcal{K}$ to $\mathcal{K}' = (O, \boldsymbol{e}')$, Equation (8.13) becomes

$$\mathcal{C} : \begin{bmatrix} x' & y' \end{bmatrix} M_{\boldsymbol{e},\boldsymbol{e}'}^{t} Q M_{\boldsymbol{e},\boldsymbol{e}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + b M_{\boldsymbol{e},\boldsymbol{e}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + 289 = 0$$

which one calculates to be

$$\mathcal{C}:\begin{bmatrix}x' & y'\end{bmatrix}\underbrace{\begin{bmatrix}578 & 0\\ 0 & 0\end{bmatrix}}_{Q'}\begin{bmatrix}x'\\ y'\end{bmatrix}+\underbrace{\begin{bmatrix}289 & -289\end{bmatrix}}_{b'}\begin{bmatrix}x'\\ y'\end{bmatrix}+289=0$$

and we have

$$\mathcal{C}: 578x'^2+289x'-289y'+289=0$$

equivalently

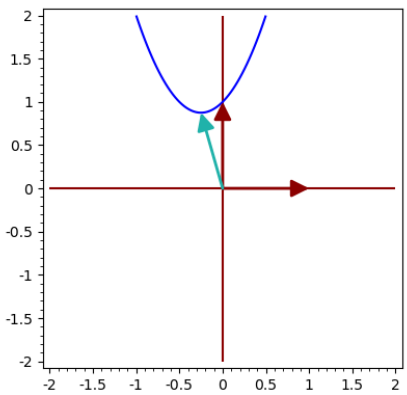
$$\mathcal{C}: 2x'^2+x'-y'+1=0$$

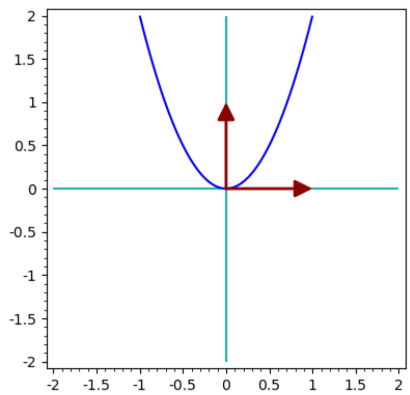
equivalently

$$\mathcal{C}: 2\Big(x'^2+\frac{x'}{2}+\frac{1}{16}\Big)-\frac{1}{8}-y'+1=0$$

equivalently

$$\mathcal{C}: 2\Big(x'+\frac{1}{4}\Big)^2-\Big(y'-\frac{7}{8}\Big)=0.$$





Now, let us change the coordinate system again, using a translation of vector $(-\frac{1}{4},\frac{7}{8})$. The new coordinate system is $\mathcal{K}''=(O'',\boldsymbol{e}'')$ where $O''=(-\frac{1}{4},\frac{7}{8})$ and the basis $\boldsymbol{e}''=\boldsymbol{e}'$ doesn't change. In $\mathcal{K}''$ the equation of $\mathcal{C}$ becomes

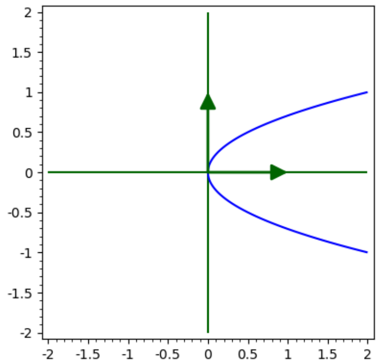
$$\mathcal{C}: 2x''^2-y''=0 \quad\Leftrightarrow\quad x''^2=\frac{1}{2}y''.$$

Clearly, this is the equation of a parabola. But it is not yet in canonical form because the focal point is on the y-axes. So, in order to obtain the canonical form we need to do a final coordinate change and permute the coordinate axes, for instance with

$$\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\in SO(2)\quad\text{or with}\quad\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\in O(2).$$

With any of the two transformations we obtain

$$\mathcal{C}=\mathcal{P}_{\frac{1}{4}}: y''^2=2\frac{1}{4}x.$$



To recap:

- We changed the coordinates from $\mathcal{K}$ to $\mathcal{K}'$ with the rotation $\mathrm{M}_{\boldsymbol{e},\boldsymbol{e}'}$ of angle θ where $\cos(\theta) = \frac{8}{17}$.
- We changed the coordinates from $\mathcal{K}'$ to $\mathcal{K}''$ with a translation of vector $(-\frac{1}{4}, \frac{7}{8})$.
- We changed the coordinates from $\mathcal{K}''$ to $\mathcal{K}'''$ in order to interchange the variables. And we obtained $\mathcal{P}_{\frac{1}{4}}$.

8.3.5 Affine classification

- In the isometric classification, we allowed only changes of coordinates which are isometries. The non-degenerate curves that we obtained are:
$$\mathcal{E}_{a,b} : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \text{or} \quad \mathcal{H}_{a,b} : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \quad \text{or} \quad \mathcal{P}_p : y^2 = 2px.$$
- It is possible to change the coordinates further and rescale the basis vectors of the coordinate axes. Rescalings are clearly not isometries.
- As an example let us consider the case of $\mathcal{E}_{a,b}$, if we replace x with ax and y with by we are doing the following coordinate change
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
and then
$$\mathcal{E}_{a,b} \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1 \quad \Leftrightarrow \quad \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = 1.$$
So, after changing coordinates the equation of $\mathcal{E}_{a,b}$ becomes
$$x'^2 + y'^2 = 1$$

- Such affine transformations can be applied to Equation (8.6) in general for hypersurfaces.
- If we restrict attention to $\mathbb{E}^2$, after inspecting the possible cases, one shows that the solutions to a quadratic equation is one of the possibilities indicated in the following table.

Case	$r = \operatorname{rank} Q$	$(p, r-p)$	k	equation	name
(a)	2	$(0,2)$	-1	$-x^2 - y^2 - 1 = 0$	imaginary ellipse (IE)
(a)	2	$(1,1)$	-1	$x^2 - y^2 - 1 = 0$	**hyperbola** (H)
(a)	2	$(2,0)$	-1	$x^2 + y^2 - 1 = 0$	**ellipse** (E)
(a)	2	$(0,2)$ or $(2,0)$	0	$-x^2 - y^2 = 0$	two complex lines
(a)	2	$(1,1)$	0	$x^2 - y^2 = 0$	**two real lines**
(a)	1	$(0,1)$ or $(1,0)$	1	$x^2 + 1 = 0$	two complex lines
(a)	1	$(1,1)$	-1	$x^2 - 1 = 0$	**two real lines**
(a)	1	$(1,1)$	0	$x^2 = 0$	**a real double-line**
Case	$r = \operatorname{rank} Q$	$(p, r-p)$	k'	equation	name
(b)	1	$(0,1)$ or $(1,0)$	1	$x^2 - y = 0$	**parabola** (P)

Table 8.1: Affine classification in dimension 2.

8.3.6 Algorithm 1: Isometric invariants

The discussion in Subsections 8.3.1 and 8.3.5 is an algebraic case-by-case analysis which may appear lengthy. However, it is not difficult to extract a recipe which allows us to decide what type of curve a given quadratic equation describes.

Fix a quadratic curve, i.e. a hyperquadric in dimension 2:

$$\mathcal{C} : q_{11}x^2 + 2q_{12}xy + q_{22}y^2 + b_1x + b_2y + c = 0. \tag{8.15}$$

(Step 1) Read off the symmetric matrix associated to Equation 8.15:

$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{12} & q_{22} \end{bmatrix}.$$

(Step 2) Calculate the eigenvalues λ_1 and λ_2 by finding the roots of the characteristic polynomial

$$P_Q(\lambda) = \begin{vmatrix} \lambda - q_{11} & q_{12} \\ q_{12} & \lambda - q_{22} \end{vmatrix}.$$

(Step 3) Calculate the constant term (see [Step 2 - translation] in the discussion of Section 8.2).

$$k = c - \frac{b_1^2}{4\lambda_1^2} - \frac{b_2^2}{4\lambda_2^2}.$$

(Step 4) Then, with respect to a coordinate system, the curve $\mathcal{C}$ has the equation

$$\lambda_1 x^2 + \lambda_2 y^2 + k = 0.$$

Now it is easy to see what type of curve this is (compare to the equations in Table 8.1).

8.3.7 Algorithm 2: Lagrange's method

One reason the discussion in Subsections 8.3.1 and 8.3.5 is lengthy is the interpretation that we added to each step (rotation, translation, eigenvectors give the direction of the new axes,...etc). If we focus just on the algebra that is behind the calculations, then the method of bringing in canonical form is particularly simple. This point of view is sometimes attributed to Lagrange. It works in any dimension.

Fix a quadratic curve, i.e. a hyperquadric in dimension 2:

$$\mathcal{C} : q_{11}x^2 + 2q_{12}xy + q_{22}y^2 + b_1x + b_2y + c = 0. \tag{8.16}$$

(Step 1) Eliminate the mixed terms by completing the squares.

(Step 2) Eliminate the linear terms by completing the squares.

(Step 3) Then, with respect to a coordinate system, the curve $\mathcal{C}$ has the equation

$$ax^2 + by^2 + c = 0.$$

where a, b, c are obtained in Steps 1 and 2. Now it is easy to see what type of curve this is (compare to the equations in Table 8.1).

8.4 Classification of quadrics - dimension 3

- A hyperquadric in $\mathbb{E}^3$ is a curve given by an equation of the form

$$\mathcal{C} : q_{11}x^2 + q_{22}y^2 + q_{33}z^2 + 2q_{12}xy + 2q_{12}yz + 2q_{12}xz + b_1x + b_2y + b_3z + c = 0. \tag{8.17}$$

- From the above discussion, we may apply an orthogonal change of coordinates and a translation to change the coordinate system such that Equation (8.4) becomes

$$\mathcal{Q} : \lambda_1x_1^2 + \lambda_2x_2^2 + \cdots + \lambda_rx_r^2 + v_{r+1}x_{r+1} + \cdots + v_nx_n = k. \tag{8.18}$$

- The classification in this case is similar: we work out all possible cases to see what we obtain.
- One important remark is that the base change matrix in Step 1, the matrix $\mathrm{M}_{e,e''}$ used to obtain Equation (8.18), is an element of the group $SO(3)$. So, by Euler's theorem, this is indeed a rotation around an axis.
- Furthermore, one can 'stretch' the coordinate axes with affine transformations which are not isometries in order to show that one may change the coordinate system such that Equation is one of the possibilities listed in the following table.

Case	$r = \operatorname{rank} Q$	$(p, r-p)$	k	equation	name
(a)	3	$(3,0)$	-1	$x^2+y^2+z^2-1=0$	**ellipsoid** (E)
(a)	3	$(2,1)$	-1	$x^2+y^2-z^2-1=0$	**hyperboloid of one sheet** (H1)
(a)	3	$(1,2)$	-1	$x^2-y^2-z^2-1=0$	**hyperboloid of two sheets** (H2)
(a)	3	$(0,3)$	-1	$-x^2-y^2-z^2-1=0$	imaginary ellipsoid (IE)
(a)	3	$(3,0)$	0	$x^2+y^2+z^2=0$	imaginary cone
(a)	3	$(2,1)$	0	$x^2+y^2-z^2=0$	(real, elliptic) **cone**
(a)	2	$(0,2)$	-1	$-x^2-y^2-1=0$	cylinder on imaginary ellipse
(a)	2	$(1,1)$	-1	$x^2-y^2-1=0$	**cylinder** on hyperbola
(a)	2	$(2,0)$	-1	$x^2+y^2-1=0$	**cylinder** on ellipse
(a)	2	$(0,2)$ or $(2,0)$	0	$-x^2-y^2=0$	cylinder on two complex lines
(a)	2	$(1,1)$	0	$x^2-y^2=0$	**cylinder** on two real lines
(a)	1	$(1,0)$	1	$x^2+1=0$	two complex planes
(a)	1	$(1,0)$	-1	$x^2-1=0$	**two real planes**
(a)	1	$(1,0)$	0	$x^2=0$	**a double plane**
(a)	1	$(0,1)$	-1	$x^2+1=0$	two complex planes
(a)	1	$(0,1)$	1	$x^2-1=0$	**two real planes**
(a)	1	$(0,1)$	0	$x^2=0$	**a double plane**
Case	$r = \operatorname{rank} Q$	$(p, r-p)$	k'	equation	name
(b)	2	$(2,0)$ or $(0,2)$	-1	$x^2+y^2-z=0$	**elliptic paraboloid** (EP)
(b)	2	$(1,1)$	-1	$x^2-y^2-z=0$	**hyperbolic paraboloid** (HP)
(b)	1	$(1,0)$	1	$x^2+y=0$	**cylinder** on parabola

CHAPTER 9

Canonical equations of real quadrics

Contents

Here we look at the main examples of *quadrics* in $\mathbb{E}^3$, sometimes called *quadratic surfaces*. These are surfaces which satisfy a quadratic equation of the form

$$\mathcal{S} : q_{11}x^2 + q_{22}y^2 + q_{33}z^2 + 2q_{12}xy + 2q_{13}xz + 2q_{23}yz + a_1x + a_2y + a_3z + c = 0. \tag{9.1}$$

Notice that an equation as above may not define a surface. It could happen that there are no solutions or that the coordinates of only one point satisfy a given quadratic equation.

9.1 Ellipsoid

9.1.1 Canonical equation - global description

An *ellipsoid* is a surface which (in some coordinate system) satisfies an equation of the form

$$\mathcal{E}_{a,b,c} : \underbrace{\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}}_{\varphi(x,y,z)} = 1 \quad \text{so} \quad \mathcal{E}_{a,b,c} = \varphi^{-1}(1)$$

for some positive constants $a, b, c \in \mathbb{R}$.

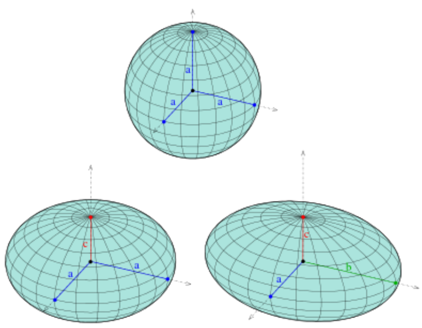


Figure 9.1: Ellipsoid[1]

If we fix one variable, we obtain intersections with planes parallel to the coordinate axes. For this surface the intersections are either ellipses, points or the empty set. Check this for $z = h$ and deduce the axes of the ellipses that you obtain.

[1]Image source: Wikipedia

9.1.2 Tangent planes

Using the gradient or the algebraic method we obtain the tangent plane at a point $p = (x_p, y_p, z_p) \in \mathcal{E}_{a,b,c}$ to be

$$T_p\mathcal{E}_{a,b,c} : \frac{x_p x}{a^2} + \frac{y_p y}{b^2} + \frac{z_p z}{c^2} = 1.$$

We will check this with the algebraic method. Let us consider a line passing through p in parametric form

$$l : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \underbrace{\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}}_{=v} \quad \Leftrightarrow \quad l = p + \langle v \rangle.$$

Recall that the tangent plane $T_p\mathcal{E}_{a,b,c}$ is the union of all lines intersecting the quadric $\mathcal{E}_{a,b,c}$ at p in a special way, i.e. in a double point. This is why we investigate the intersection of our quadric with l. How do we obtain the intersection $\mathcal{E}_{a,b,c} \cap l$? We look at those points of l which satisfy the equation of the ellipsoid, i.e. we look for solutions t for the equation

$$\varphi(\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}) = 1 \quad \Leftrightarrow \quad \frac{(x_p + tv_x)^2}{a^2} + \frac{(y_p + tv_y)^2}{b^2} + \frac{(z_p + tv_z)^2}{c^2} - 1 = 0.$$

If we rearrange the left-hand side as a polynomial in t we get

$$\Big(\frac{v_x^2}{a^2} + \frac{v_y^2}{b^2} + \frac{v_z^2}{c^2}\Big)t^2 + 2\Big(\frac{x_p v_x}{a^2} + \frac{y_p v_y}{b^2} + \frac{z_p v_z}{c^2}\Big)t + \underbrace{\frac{x_p^2}{a^2} + \frac{y_p^2}{b^2} + \frac{z_p^2}{c^2}}_{=1} - 1 = 0$$

$$\Leftrightarrow \quad \underbrace{\Big(\frac{v_x^2}{a^2} + \frac{v_y^2}{b^2} + \frac{v_z^2}{c^2}\Big)}_{\neq 0}t^2 + 2\Big(\frac{x_p v_x}{a^2} + \frac{y_p v_y}{b^2} + \frac{z_p v_z}{c^2}\Big)t = 0.$$

This equation in t admits the solution $t = 0$. That is clear, the point on l corresponding to $t = 0$ is the point p which, by assumption, lies on $\mathcal{E}_{a,b,c}$. Furthermore, the second solution will correspond to a second point of intersection. However, the line l is tangent to our quadric if p is a *double point of intersection*. This happens if and only if the above equation has $t = 0$ as double solution, i.e. if and only if

$$\frac{x_p v_x}{a^2} + \frac{y_p v_y}{b^2} + \frac{z_p v_z}{c^2} = 0.$$

How do we interpret this? In the Euclidean setting this can be interpreted as saying that $(\frac{x_p}{a^2}, \frac{y_p}{b^2}, \frac{z_p}{c^2})$ is perpendicular to the direction vector v of the line. All lines which are tangent to the surface and contain p, need to satisfy this condition, so

$$T_p\mathcal{E}_{a,b,c} : \begin{bmatrix} \frac{x_p}{a^2} \\ \frac{y_p}{b^2} \\ \frac{z_p}{c^2} \end{bmatrix} \cdot (\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}) = 0 \quad \Leftrightarrow \quad \frac{x_p x}{a^2} + \frac{y_p y}{b^2} + \frac{z_p z}{c^2} - 1 = 0.$$

Deduce this equation also with the gradient.

9.1.3 Parametrizations - local description

A parametrization of this surface is

$$\begin{cases} x(\theta_1,\theta_2) = a\cos(\theta_1)\cos(\theta_2) \\ y(\theta_1,\theta_2) = b\sin(\theta_1)\cos(\theta_2) \\ z(\theta_1,\theta_2) = c\sin(\theta_2) \end{cases} \qquad \theta_1 \in [0,2\pi[\quad \theta_2 \in [-\frac{\pi}{2},\frac{\pi}{2}[$$

Why? Check this and deduce a parametrization of the tangent plane $T_p\mathcal{E}_{a,b,c}$ at the point

$$p = \begin{bmatrix} x(\theta_{1,p},\theta_{2,p}) \\ y(\theta_{1,p},\theta_{2,p}) \\ z(\theta_{1,p},\theta_{2,p}) \end{bmatrix} \in \mathcal{E}_{a,b,c}.$$

9.2 Elliptic Cone

9.2.1 Canonical equation - global description

An *elliptic cone* is a surface which (in some coordinate system) satisfies an equation of the form

$$\mathcal{C}_{a,b,c} : \underbrace{\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}}_{\varphi(x,y,z)} = 0 \quad \text{so} \quad \mathcal{E}_{a,b,c} = \varphi^{-1}(0)$$

for some positive constants $a,b,c \in \mathbb{R}$.

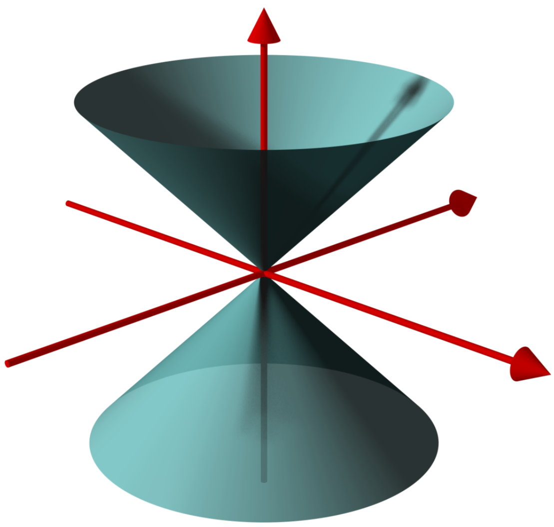


Figure 9.2: Elliptic cone[2]

If we fix one variable, we obtain intersections with planes parallel to the coordinate axes. For this surface the intersections are either ellipses, hyperbolas or a point. Check this for $z = h$ and deduce the axes of the ellipses that you obtain.

[2]Image source: Wikipedia

9.2.2 Conic sections

Above we noticed that the intersection of an elliptic cone with planes parallel to the coordinate axes are quadratic surfaces (possibly degenerate). In fact we have the following result.

Proposition 9.1. The intersection of an elliptic cone with an arbitrary plane is a (possibly degenerate) quadratic curve.

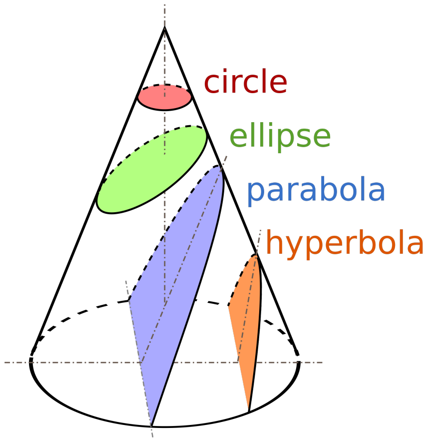


Figure 9.3: Conic sections[3]

9.2.3 Tangent planes

As in the case of the ellipsoid, using the gradient or the algebraic method we obtain the tangent plane at a point $p = (x_p, y_p, z_p) \in \mathcal{C}_{a,b,c}$ to be

$$T_p\mathcal{C}_{a,b,c} : \frac{x_p x}{a^2} + \frac{y_p y}{b^2} - \frac{z_p z}{c^2} = 0.$$

9.2.4 $\mathcal{C}_{a,b,c}$ as ruled surface

A *ruled surface* is a surface $\mathcal{S}$ such that for any point p on the surface S there is a line l which containes p and which is contained in S:

$$\forall p \in \mathcal{S}, \exists \text{ a line } l \text{ such that } p \in l \text{ and } l \subseteq S.$$

If we denote by $\mathcal{L}$ the family of all these lines, it is easy to see that the surface is the union of them:

$$\mathcal{S} = \bigcup_{l \in \mathcal{L}} l.$$

[3]Image source: Wikipedia

The lines in $\mathcal{L}$ are called *rectilinear generators of the surface* $\mathcal{S}$. We refer to them as *generators* since we don't consider here non-rectilinear generators.

So, how is the cone a ruled surface? Fix a point $(x_0, y_0, z_0) \in \mathcal{C}_{a,b,c}$ and notice that for any $t \in \mathbb{R}$ we have

$$\frac{(tx_0)^2}{a^2} + \frac{(ty_0)^2}{b^2} - \frac{(tz_0)^2}{c^2} = t^2\left(\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} - \frac{z_0^2}{c^2}\right) = 0.$$

Thus, the line $\{(tx_0, ty_0, tz_0) : t \in \mathbb{R}\}$, which passes through the given point and the origin is contained in $\mathcal{C}_{a,b,c}$. The set of all lines obtained in this way form the generators $\mathcal{L}$ of the cone.

9.2.5 Parametrizations - local description

Describing a parametrization of an elliptic cone can be generalized to any planar curve. Suppose you have a parametrization of a curve in the plane Oxy. In our case, for the ellipse

$$\begin{cases} x(\theta) = a\cos(\theta) \\ y(\theta) = b\sin(\theta) \end{cases} \quad \theta \in [0, 2\pi[.$$

You want to rescale this curve with the height such that when the height $z = 0$ you have a point, and for all other values of z you have a rescaled versions of your curve:

$$\begin{cases} x(\theta, h) = ha\cos(\theta) \\ y(\theta, h) = hb\sin(\theta) \\ z(\theta, h) = hc \end{cases} \quad \theta \in [0, 2\pi[\quad h \in \mathbb{R}.$$

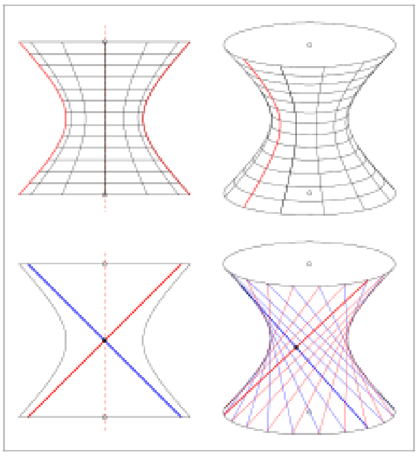
9.3 Hyperboloid of one sheet

9.3.1 Canonical equation - global description

A *hyperboloid of one sheet* is a surface which (in some coordinate system) satisfies an equation of the form

$$\mathcal{H}^1_{a,b,c} : \underbrace{\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}}_{\varphi(x,y,z)} = 1 \quad \text{so} \quad \mathcal{H}^1_{a,b,c} = \varphi^{-1}(1) \tag{9.2}$$

for some positive constants $a, b, c \in \mathbb{R}$.



(a) Hyperboloid of one sheet[4]

(b) Kobe Port Tower

If we fix one variable, we obtain intersections with planes parallel to the coordinate axes. For this surface the intersections are either ellipses, hyperbolas or two lines. Check this for $y = h$ and deduce the axes of the hyperbolas that you obtain.

9.3.2 Tangent planes

Using the gradient or the algebraic method we obtain the tangent plane at a point $p = (x_p, y_p, z_p) \in \mathcal{H}^1_{a,b,c}$ to be

$$T_p\mathcal{H}^1_{a,b,c} : \frac{x_p x}{a^2} + \frac{y_p y}{b^2} - \frac{z_p z}{c^2} = 1.$$

We will check this with the algebraic method. Let us consider a line passing through p in parametric form

$$l : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \underbrace{\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}}_{=v} \quad \Leftrightarrow \quad l = p + \langle v \rangle.$$

Recall that the tangent plane $T_p\mathcal{H}^1_{a,b,c}$ is the union of all lines intersecting the quadric $\mathcal{H}^1_{a,b,c}$ at p in a special way, i.e. in a double point. This is why we investigate the intersection of our quadric with l. How do we obtain the intersection $\mathcal{H}^1_{a,b,c} \cap l$? We look at those points of l which satisfy the equation of our hyperboloid, i.e. we look for solutions t for the equation

$$\varphi\left(\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}\right) = 1 \quad \Leftrightarrow \quad \frac{(x_p + tv_x)^2}{a^2} + \frac{(y_p + tv_y)^2}{b^2} - \frac{(z_p + tv_z)^2}{c^2} - 1 = 0. \tag{9.3}$$

[4] Image source: Wikipedia

If we rearrange the left-hand side as a polynomial in t we get

$$\Big(\frac{v_x^2}{a^2}+\frac{v_y^2}{b^2}-\frac{v_y^2}{c^2}\Big)t^2+2\Big(\frac{x_pv_x}{a^2}+\frac{y_pv_y}{b^2}-\frac{z_pv_z}{c^2}\Big)t+\underbrace{\frac{x_p^2}{a^2}+\frac{y_p^2}{b^2}-\frac{z_p^2}{c^2}}_{=1}-1=0$$

$$\Leftrightarrow\quad \underbrace{\Big(\frac{v_x^2}{a^2}+\frac{v_y^2}{b^2}-\frac{v_y^2}{c^2}\Big)}_{=0?}t^2+2\Big(\frac{x_pv_x}{a^2}+\frac{y_pv_y}{b^2}-\frac{z_pv_z}{c^2}\Big)t=0. \tag{9.4}$$

This equation in t admits the solution $t=0$. That is clear, the point on l corresponding to $t=0$ is the point p which, by assumption, lies on $\mathcal{H}^1_{a,b,c}$. Furthermore, a second solution will correspond to a second point of intersection. However, the line l is tangent to our quadric if p is a *double point of intersection*. This happens if and only if the above equation has $t=0$ as double solution, i.e. if and only if

$$\frac{v_x^2}{a^2}+\frac{v_y^2}{b^2}-\frac{v_y^2}{c^2}\neq 0\quad\text{and}\quad\frac{x_pv_x}{a^2}+\frac{y_pv_y}{b^2}-\frac{z_pv_z}{c^2}=0.$$

How do we interpret the second condition? In the Euclidean setting this can also be interpreted as saying that $(\frac{x_p}{a^2},\frac{y_p}{b^2},-\frac{z_p}{c^2})$ is perpendicular to the direction vector v of the line. In both cases, all lines which are tangent to the surface and contain p, need to satisfy this equations, so the tangent plane is

$$T_p\mathcal{H}^1_{a,b,c}:\begin{bmatrix}\frac{x_p}{a^2}\\ \frac{y_p}{b^2}\\ -\frac{z_p}{c^2}\end{bmatrix}\cdot\left(\begin{bmatrix}x\\ y\\ z\end{bmatrix}-\begin{bmatrix}x_p\\ y_p\\ z_p\end{bmatrix}\right)=0\quad\Leftrightarrow\quad\frac{x_px}{a^2}+\frac{y_py}{b^2}-\frac{z_pz}{c^2}-1=0.$$

Deduce this equation also with the gradient.

How do we interpret the first condition? If

$$\frac{v_x^2}{a^2}+\frac{v_y^2}{b^2}-\frac{v_y^2}{c^2}=0 \tag{9.5}$$

then equation (9.4) is linear, and has one simple solution $t=0$. This means that l intersects $\mathcal{H}^1_{a,b,c}$ only once (it punctures the surface in one point). Such lines are not tangent to the surface. How can we visualize this? Let us start with what we know: the vector $v=(v_x,v_y,v_z)$ satisfies the equation (9.5), so we can think of it as the position vector of some point on the cone $\mathcal{C}_{a,b,c}$. How does this cone relate to our hyperboloid? Our surface $\mathcal{H}^1_{a,b,c}$ is the union of hyperbolas (revolving on ellipses around the z-axis) and if we take the union of all the asymptotes to these hyperbolas we get $\mathcal{C}_{a,b,c}$ (see Figure 9.5).

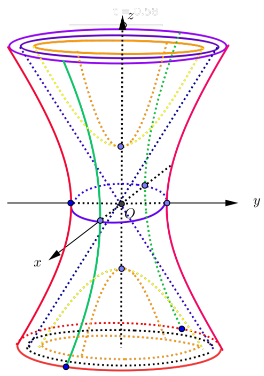


Figure 9.5: Hyperboloid and asymptotic cone[5]

This should help to see that, when the vector $v = (v_x, v_y, v_z)$ satisfies equation (9.5), the line l is parallel to a line contained in the cone $\mathcal{C}_{a,b,c}$. It will therefore intersect $\mathcal{H}^1_{a,b,c}$ in at most one point.

Notice also that if $l \subseteq \mathcal{C}_{a,b,c}$, it will not intersect $\mathcal{H}^1_{a,b,c}$ at all, but this cannot happen in our setting because we chose the point p such that it lies both on l and on our quadric. In fact, the related question 'does a given line l intersect $\mathcal{H}^1_{a,b,c}$?' can be answered by investigating equation (9.3) without the assumption that p lies on the surface $\mathcal{H}^1_{a,b,c}$. How would you do this?

9.3.3 $\mathcal{H}^1_{a,b,c}$ as ruled surface

Here is a fact: the hyperboloid with one sheet is a *doubly ruled surface* (this is visible in Figure 9.4a). Hmm.. I know what a ruled surface is, because cones and cylinders are ruled surfaces, but, 'doubly ruled'? *Doubly ruled* just means that it is a ruled surface in two ways. So, there are two distinct families of lines, $\mathcal{L}_1$ and $\mathcal{L}_2$ respectively, such that

$$\mathcal{H}^1_{a,b,c} = \bigcup_{l\in\mathcal{L}_1} l \quad \text{and} \quad \mathcal{H}^1_{a,b,c} = \bigcup_{l\in\mathcal{L}_2} l.$$

One way to see where the two families of lines come from is to rearrange Equation (9.2):

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad\Leftrightarrow\quad \frac{x^2}{a^2}-\frac{z^2}{c^2}=1-\frac{y^2}{b^2} \quad\Leftrightarrow\quad \left(\frac{x}{a}-\frac{z}{c}\right)\left(\frac{x}{a}+\frac{z}{c}\right)=\left(1-\frac{y}{b}\right)\left(1+\frac{y}{b}\right) \tag{9.6}$$

Now, assume that the factors in the last equation are not 0, then we can divide to obtain

$$\Leftrightarrow\quad \frac{\frac{x}{a}-\frac{z}{c}}{1-\frac{y}{b}}=\frac{1+\frac{y}{b}}{\frac{x}{a}+\frac{z}{c}}=\frac{\mu}{\lambda}$$

for some parameters λ and μ. We introduced these parameters in order to separate the above equation:

$$\Leftrightarrow\quad l_{\lambda,\mu}:\begin{cases}\lambda\left(\frac{x}{a}-\frac{z}{c}\right)=\mu\left(1-\frac{y}{b}\right)\\ \mu\left(\frac{x}{a}+\frac{z}{c}\right)=\lambda\left(1+\frac{y}{b}\right)\end{cases}.$$

[5]Prof. C. Pintea - lecture notes

What we end up with is a system of two equations, which are linear in x, y, z and which depend on the parameters λ and μ. For each fixed pair of parameters, λ and μ, we get a line which we denote with $l_{\lambda,\mu}$. Reading the above deduction backwards it is easy to see that all points on such a line satisfy the equation of $\mathcal{H}^1_{a,b,c}$. So, we have a family of lines contained in your hyperboloid.

We assumed that the factors in (9.6) are not zero. In fact, you only divide by two of them, so .. if one of those two is zero, you can flip the above fraction and divide by the other two. That will lead to the same family of lines $\mathcal{L}_1 = \{l_{\lambda,\mu} : \lambda, \mu \in \mathbb{R}, \lambda^2 + \mu^2 \neq 0\}$.

OK, what about $\mathcal{L}_2$? The second family of generators (these lines are called generators), is obtained if you group the terms differently:

$$\frac{\frac{x}{a} - \frac{z}{c}}{1 + \frac{y}{b}} = \frac{1 - \frac{y}{b}}{\frac{x}{a} + \frac{z}{c}} = \frac{\mu}{\lambda}.$$

Then, you obtain:

$$\tilde{l}_{\lambda,\mu} : \begin{cases} \lambda\left(\frac{x}{a} - \frac{z}{c}\right) = \mu\left(1 + \frac{y}{b}\right) \\ \mu\left(\frac{x}{a} + \frac{z}{c}\right) = \lambda\left(1 - \frac{y}{b}\right) \end{cases} .$$

As above, one can check that points on these lines satisfy the equation of our hyperboloid.

One important thing to notice is that, although we write down two parameters, λ and μ, we don't necessarily get distinct lines for distinct parameters: $l_{\lambda,\mu} = l_{t\lambda,t\mu}$ for any nonzero scalar t. So, in fact, $\mathcal{L}_1$ depends on one parameter. More concretely

$$\mathcal{L}_1 = \left\{ l_\alpha : \begin{cases} \left(\frac{x}{a} - \frac{z}{c}\right) = \alpha\left(1 - \frac{y}{b}\right) \\ \alpha\left(\frac{x}{a} + \frac{z}{c}\right) = \left(1 + \frac{y}{b}\right) \end{cases} \right\} \bigcup \left\{ l_\infty : \begin{cases} 0 = \left(1 - \frac{y}{b}\right) \\ \left(\frac{x}{a} + \frac{z}{c}\right) = 0 \end{cases} \right\}$$

and similarly for $\mathcal{L}_2$.

9.3.4 Parametrizations - local description

Two parametrizations of this surface are

$$\sigma_1(\theta_1, \theta_2) = \begin{bmatrix} a\sqrt{1+\theta_2^2}\cos(\theta_1) \\ b\sqrt{1+\theta_2^2}\sin(\theta_1) \\ c\theta_2 \end{bmatrix} \quad \text{and} \quad \sigma_2(\theta_1, \theta_2) = \begin{bmatrix} a\cosh(\theta_2)\cos(\theta_1) \\ b\cosh(\theta_2)\sin(\theta_1) \\ c\sinh(\theta_2) \end{bmatrix}$$

for $\theta_1 \in [0, 2\pi[$ and $\theta_2 \in \mathbb{R}$. Why? The parameter θ_1 is used to rotate on ellipses the curve obtained for $\theta_1 = 0$. What is this curve that we 'rotate'? Check this and deduce a parametrization of the tangent plane $T_p\mathcal{H}^1_{a,b,c}$ at the point

$$p = \begin{bmatrix} x(\theta_{1,p}, \theta_{2,p}) \\ y(\theta_{1,p}, \theta_{2,p}) \\ z(\theta_{1,p}, \theta_{2,p}) \end{bmatrix} \in \mathcal{H}^1_{a,b,c}$$

where $\theta_{1,p}$ and $\theta_{2,p}$ are the parameters of the point p, i.e. $p = p(x(\theta_{1,p}, \theta_{2,p}), y(\theta_{1,p}, \theta_{2,p}), z(\theta_{1,p}, \theta_{2,p}))$.

9.4 Hyperboloid of two sheets

9.4.1 Canonical equation - global description

A *hyperboloid of two sheets* is a surface which (in some coordinate system) satisfies an equation of the form

$$\mathcal{H}^2_{a,b,c} : \underbrace{\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}}_{\varphi(x,y,z)} = -1 \quad \text{so} \quad \mathcal{H}^2_{a,b,c} = \varphi^{-1}(-1)$$

for some positive constants $a, b, c \in \mathbb{R}$.

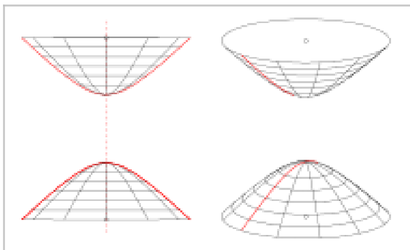


Figure 9.6: Hyperboloid of two sheets[6]

If we fix one variable, we obtain intersections with planes parallel to the coordinate axes. For this surface the intersections are either ellipses, hyperbolas or the empty set. Check this for $y = h$ and deduce the axes of the hyperbolas that you obtain.

9.4.2 Tangent planes

Using the gradient or the algebraic method we obtain the tangent plane at a point $p = (x_p, y_p, z_p) \in \mathcal{H}^2_{a,b,c}$ to be

$$T_p\mathcal{H}^2_{a,b,c} : \frac{x_p x}{a^2}+\frac{y_p y}{b^2}-\frac{z_p z}{c^2} = -1.$$

We will check this with the algebraic method. Let us consider a line passing through p in parametric form

$$l : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \underbrace{\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}}_{=v} \quad \Leftrightarrow \quad l = p + \langle v \rangle.$$

Recall that the tangent plane $T_p\mathcal{H}^2_{a,b,c}$ is the union of all lines intersecting the quadric $\mathcal{H}^2_{a,b,c}$ at p in a special way, i.e. in a double point. This is why we investigate the intersection of our quadric with l. How do we obtain the intersection $\mathcal{H}^2_{a,b,c} \cap l$? We look at those points of l which satisfy the equation

[6]Image source: Wikipedia

of our hyperboloid, i.e. we look for solutions t for the equation

$$\varphi(\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}) = -1 \quad \Leftrightarrow \quad \frac{(x_p+tv_x)^2}{a^2} + \frac{(y_p+tv_y)^2}{b^2} - \frac{(z_p+tv_z)^2}{c^2} + 1 = 0.$$

If we rearrange the left-hand side as a polynomial in t we get

$$\Big(\frac{v_x^2}{a^2} + \frac{v_y^2}{b^2} - \frac{v_y^2}{c^2}\Big)t^2 + 2\Big(\frac{x_p v_x}{a^2} + \frac{y_p v_y}{b^2} - \frac{z_p v_z}{c^2}\Big)t + \underbrace{\frac{x_p^2}{a^2} + \frac{y_p^2}{b^2} - \frac{z_p^2}{c^2}}_{=-1} + 1 = 0$$

$$\Leftrightarrow \quad \underbrace{\Big(\frac{v_x^2}{a^2} + \frac{v_y^2}{b^2} - \frac{v_y^2}{c^2}\Big)}_{=0?}t^2 + 2\Big(\frac{x_p v_x}{a^2} + \frac{y_p v_y}{b^2} - \frac{z_p v_z}{c^2}\Big)t = 0. \tag{9.7}$$

This equation in t admits the solution $t = 0$. That is clear, the point on l corresponding to $t = 0$ is the point p which, by assumption, lies on $\mathcal{H}^2_{a,b,c}$. Furthermore, a second solution will correspond to a second point of intersection. However, the line l is tangent to our quadric if p is a *double point of intersection*. This happens if and only if the above equation has $t = 0$ as double solution, i.e. if and only if

$$\frac{v_x^2}{a^2} + \frac{v_y^2}{b^2} - \frac{v_y^2}{c^2} \neq 0 \quad \text{and} \quad \frac{x_p v_x}{a^2} + \frac{y_p v_y}{b^2} - \frac{z_p v_z}{c^2} = 0.$$

How do we interpret the second condition? Similar to the case of the hyperboloid of one sheet:

$$T_p\mathcal{H}^2_{a,b,c} : \begin{bmatrix} \frac{v_x}{a^2} \\ \frac{v_y}{b^2} \\ -\frac{v_x}{c^2} \end{bmatrix} \cdot (\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}) = 0 \quad \Leftrightarrow \quad \frac{x_p x}{a^2} + \frac{y_p y}{b^2} - \frac{z_p z}{z^2} + 1 = 0.$$

Deduce this equation also with the gradient.

How do we interpret the first condition? If

$$\frac{v_x^2}{a^2} + \frac{v_y^2}{b^2} - \frac{v_y^2}{c^2} = 0 \tag{9.8}$$

then equation (9.7) is linear, and has one simple solution $t = 0$. This means that l intersects $\mathcal{H}^2_{a,b,c}$ only once (it punctures the surface in one point). Such lines are not tangent to the surface. How can we visualize this? Let us start with what we know: the vector $v = (v_x, v_y, v_z)$ satisfies the equation (9.8), so we can think of it as the position vector of some point on the cone $\mathcal{C}_{a,b,c}$. How does this cone relate to our hyperboloid? Our surface $\mathcal{H}^2_{a,b,c}$ is the union of hyperbolas and if we take the union of all the asymptotes to these hyperbolas we get $\mathcal{C}_{a,b,c}$ (see Figure 9.5). So, when l is parallel to a line contained in $\mathcal{C}_{a,b,c}$, it will intersect $\mathcal{H}^2_{a,b,c}$ in at most one point. Notice also that if $l \subseteq \mathcal{C}_{a,b,c}$, it will not intersect $\mathcal{H}^2_{a,b,c}$ at all, but this cannot happen because we chose the point p such that it lies both on l and on our quadric.

9.4.3 Parametrizations - local description

Two parametrizations of this surface are

$$\sigma_1(\theta_1,\theta_2)=\begin{bmatrix} a\sqrt{\theta_2^2-1}\cos(\theta_1)\\ b\sqrt{\theta_2^2-1}\sin(\theta_1)\\ c\theta_2\end{bmatrix}\quad\text{and}\quad \sigma_2(\theta_1,\theta_2)=\begin{bmatrix} a\sinh(\theta_2)\cos(\theta_1)\\ b\sinh(\theta_2)\sin(\theta_1)\\ \varepsilon c\cosh(\theta_2)\end{bmatrix}$$

for $\theta_1\in[0,2\pi[$, $\theta_2\in\mathbb{R}$ and $\varepsilon\in\{\pm 1\}$. Why? The parameter θ_1 is used to 'rotate' on ellipses the curve obtained for $\theta_1=0$. What is this curve? Check this and deduce a parametrization of the tangent plane $T_p\mathcal{H}^2_{a,b,c}$ at the point

$$p=\begin{bmatrix} x(\theta_{1,p},\theta_{2,p})\\ y(\theta_{1,p},\theta_{2,p})\\ z(\theta_{1,p},\theta_{2,p})\end{bmatrix}\in\mathcal{H}^2_{a,b,c}$$

where $\theta_{1,p}$ and $\theta_{2,p}$ are the parameters of the point p, i.e. $p=p(x(\theta_{1,p},\theta_{2,p}),y(\theta_{1,p},\theta_{2,p}),z(\theta_{1,p},\theta_{2,p}))$. Notice also that with σ_2 we have a parametrization for each sheet of this hyperboloid, with $\varepsilon=1$ we get one sheet and with $\varepsilon=-1$ we get the other sheet. One should also be careful with where the parameters live: for σ_1 you want to choose θ_2 in $]-\infty,-1]\cup[1,\infty[$ so that the square root is defined.

9.5 Elliptic paraboloid

9.5.1 Canonical equation - global description

An *elliptic paraboloid* is a surface which (in some coordinate system) satisfies an equation of the form

$$\mathcal{P}^e_{a,b}:\underbrace{\frac{x^2}{a}+\frac{y^2}{b}-2z}_{\varphi(x,y,z)}=0\quad\text{so}\quad\mathcal{P}^e_{a,b}=\varphi^{-1}(0)$$

for some positive constants $a,b\in\mathbb{R}$.

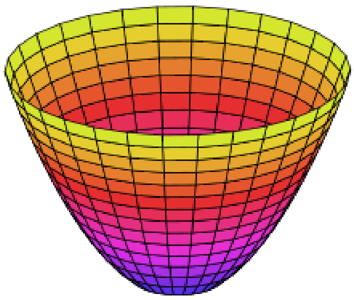


Figure 9.7: Elliptic paraboloid[7]

[7] Image source: Wikipedia

If we fix one variable, we obtain intersections with planes parallel to the coordinate axes. For this surface the intersections are either ellipses, parabolas or the empty set. Check this for $y = h$ and see what parabolas you obtain.

9.5.2 Tangent planes

Using the gradient or the algebraic method we obtain the tangent plane at a point $p = (x_p, y_p, z_p) \in \mathcal{P}^e_{a,b}$ to be

$$T_p\mathcal{P}^e_{a,b} : \frac{x_p x}{a} + \frac{y_p y}{b} - z_p - z = 0.$$

We will check this with the algebraic method. Let us consider a line passing through p in parametric form

$$l : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \underbrace{\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}}_{=v} \quad \Leftrightarrow \quad l = p + \langle v \rangle.$$

Recall that the tangent plane $T_p\mathcal{P}^e_{a,b}$ is the union of all lines intersecting the quadric $\mathcal{P}^e_{a,b}$ at p in a special way, i.e. in a double point. This is why we investigate the intersection of our quadric with l. How do we obtain the intersection $\mathcal{P}^e_{a,b} \cap l$? We look at those points of l which satisfy the equation of our paraboloid, i.e. we look for solutions t for the equation

$$\varphi(\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}) = 0 \quad \Leftrightarrow \quad \frac{(x_p + tv_x)^2}{a} + \frac{(y_p + tv_y)^2}{b} - 2(z_p + tv_z) = 0.$$

If we rearrange the left-hand side as a polynomial in t we get

$$\Big(\frac{v_x^2}{a} + \frac{v_y^2}{b}\Big)t^2 + 2\Big(\frac{x_p v_x}{a} + \frac{y_p v_y}{b} - v_z\Big)t + \underbrace{\frac{x_p^2}{a} + \frac{y_p^2}{b} - 2z_p}_{=0} = 0$$

$$\Leftrightarrow \quad \underbrace{\Big(\frac{v_x^2}{a} + \frac{v_y^2}{b}\Big)}_{\neq 0}t^2 + 2\Big(\frac{x_p v_x}{a} + \frac{y_p v_y}{b} - v_z\Big)t = 0. \tag{9.9}$$

This equation in t admits the solution $t = 0$. That is clear, the point on l corresponding to $t = 0$ is the point p which, by assumption, lies on $\mathcal{P}^e_{a,b}$. Furthermore, a second solution will correspond to a second point of intersection. However, the line l is tangent to our quadric if p is a *double point of intersection*. This happens if and only if the above equation has $t = 0$ as double solution, i.e. if and only if

$$\frac{x_p v_x}{a} + \frac{y_p v_y}{b} - v_z = 0.$$

How do we interpret this condition? Similar to the quadrics treated in the previous sections, so

$$T_p\mathcal{P}^e_{p,q} : \begin{bmatrix} \frac{x_p}{a} \\ \frac{y_p}{b} \\ -1 \end{bmatrix} (\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}) = 0 \quad \Leftrightarrow \quad \frac{x_p x}{a} + \frac{y_p y}{b} - z_p - z = 0.$$

Deduce this equation also with the gradient.

9.5.3 Parametrizations - local description

A parametrization of this surface is

$$\sigma(\theta_1, \theta_2) = \begin{bmatrix} \sqrt{a\theta_2}\cos(\theta_1) \\ \sqrt{b\theta_2}\sin(\theta_1) \\ \theta_2/2 \end{bmatrix} \quad \theta_1 \in [0, 2\pi[\quad \theta_2 \in [0, \infty[$$

Why? Check this and deduce a parametrization of the tangent plane $T_p\mathcal{P}^e_{a,b}$ at the point

$$p = \begin{bmatrix} x(\theta_{1,p}, \theta_{2,p}) \\ y(\theta_{1,p}, \theta_{2,p}) \\ z(\theta_{1,p}, \theta_{2,p}) \end{bmatrix} \in \mathcal{P}^e_{a,b}.$$

where $\theta_{1,p}$ and $\theta_{2,p}$ are the parameters of the point p, i.e. $p = p(x(\theta_{1,p}, \theta_{2,p}), y(\theta_{1,p}, \theta_{2,p}), z(\theta_{1,p}, \theta_{2,p}))$.

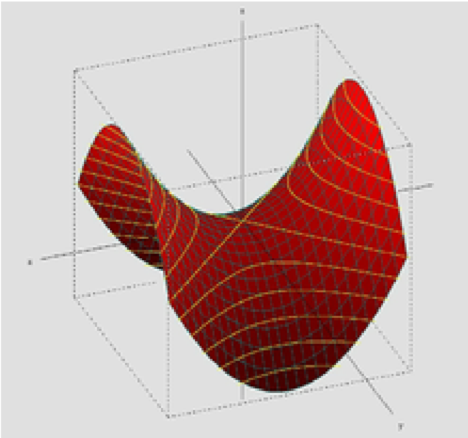
9.6 Hyperbolic paraboloid

9.6.1 Canonical equation - global description

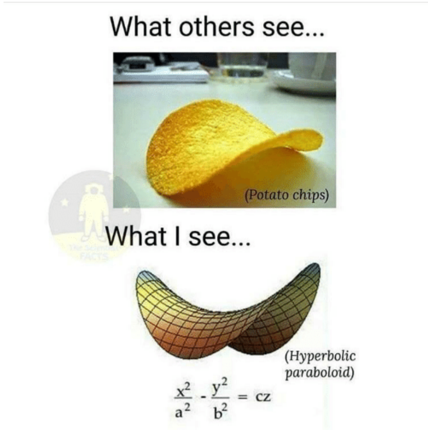
A *hyperbolic paraboloid* is a surface which (in some coordinate system) satisfies an equation of the form

$$\mathcal{P}^h_{a,b} : \underbrace{\frac{x^2}{a} - \frac{y^2}{b} - 2z}_{\varphi(x,y,z)} = 0 \quad \text{so} \quad \mathcal{P}^h_{a,b} = \varphi^{-1}(0) \tag{9.10}$$

for some positive constants $a, b \in \mathbb{R}$.



(a) Hyperbolic paraboloid[8]



(b) Potato chips[9]

If we fix one variable, we obtain intersections with planes parallel to the coordinate axes. For this surface the intersections are either parabolas, hyperbolas or two lines. Check this for $y = h$ and deduce the parabolas that you obtain.

9.6.2 Tangent planes

Using the gradient or the algebraic method we obtain the tangent plane at a point $p = (x_p, y_p, z_p) \in \mathcal{P}^h_{a,b}$ to be

$$T_p\mathcal{P}^h_{a,b} : \frac{x_p x}{a} - \frac{y_p y}{b} - z_p - z = 0.$$

We will check this with the algebraic method. Let us consider a line passing through p in parametric form

$$l : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \underbrace{\begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}}_{=v} \quad \Leftrightarrow \quad l = p + \langle v \rangle.$$

Recall that the tangent plane $T_p\mathcal{P}^h_{a,b}$ is the union of all lines intersecting the quadric $\mathcal{P}^h_{a,b}$ at p in a special way, i.e. in a double point. This is why we investigate the intersection of our quadric with l. How do we obtain the intersection $\mathcal{P}^h_{a,b} \cap l$? We look at those points of l which satisfy the equation of our paraboloid, i.e. we look for solutions t for the equation

$$\varphi(\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}) = 0 \quad \Leftrightarrow \quad \frac{(x_p + tv_x)^2}{a} - \frac{(y_p + tv_y)^2}{b} - 2(z_p + tv_z) = 0.$$

[6]Image source: Wikipedia

[9]Image source: https://me.me/

If we rearrange the left-hand side as a polynomial in t we get

$$\Big(\frac{v_x^2}{a}-\frac{v_y^2}{b}\Big)t^2+2\Big(\frac{x_pv_x}{a}-\frac{y_pv_y}{b}-v_z\Big)t+\underbrace{\frac{x_p^2}{a}-\frac{y_p^2}{b}-2z_p}_{=0}=0$$

$$\Leftrightarrow\quad \underbrace{\Big(\frac{v_x^2}{a}-\frac{v_y^2}{b}\Big)}_{=0?}t^2+2\Big(\frac{x_pv_x}{a}-\frac{y_pv_y}{b}-v_z\Big)t=0. \tag{9.11}$$

This equation in t admits the solution $t=0$. That is clear, the point on l corresponding to $t=0$ is the point p which, by assumption, lies on $\mathcal{P}^h_{a,b}$. Furthermore, a second solution will correspond to a second point of intersection. However, the line l is tangent to our quadric if p is a *double point of intersection*. This happens if and only if the above equation has $t=0$ as double solution, i.e. if and only if

$$\frac{v_x^2}{a}-\frac{v_y^2}{b}\neq 0 \quad\text{and}\quad \frac{x_pv_x}{a}-\frac{y_pv_y}{b}-v_z=0.$$

How do we interpret the second condition? Similar to the quadrics treated in the previous sections, so

$$T_p\mathcal{P}^h_{p,q}:\begin{bmatrix}\frac{x_p}{a}\\ -\frac{y_p}{b}\\ -1\end{bmatrix}\cdot\left(\begin{bmatrix}x\\ y\\ z\end{bmatrix}-\begin{bmatrix}x_p\\ y_p\\ z_p\end{bmatrix}\right)=0\quad\Leftrightarrow\quad \frac{x_px}{a}-\frac{y_py}{b}-z_p-z=0.$$

Deduce this equation also with the gradient.

How do we interpret the first condition? If

$$\frac{v_x^2}{a}-\frac{v_y^2}{b}=0\quad\Leftrightarrow\quad\Big(\frac{v_x}{\sqrt{a}}-\frac{v_y}{\sqrt{b}}\Big)\Big(\frac{v_x}{\sqrt{a}}+\frac{v_y}{\sqrt{b}}\Big)=0 \tag{9.12}$$

then equation (9.11) is linear, and has one simple solution $t=0$. This means that l intersects $\mathcal{P}^h_{a,b}$ only once (it punctures the surface in one point). Such lines are not tangent to the surface. How can we visualize this? Let us start with what we know: the vector $v=(v_x,v_y,v_z)$ satisfies the equation (9.12), so we can think of it as the position vector of some point on the cylinder $\mathrm{Cyl}(\mathcal{A},\mathbf{k})$ where $\mathcal{A}$ is the union of two lines given by the equation (9.12). These two lines are the intersection of our quadric $\mathcal{P}^h_{a,b}$ with the coordinate plane $z=0$ (they are visible in Figure 9.8a). Three things can happen here:

1. l is one of the lines in $\mathcal{A}$, then all points of l lie in $\mathcal{P}^h_{a,b}$, so we have infinitely many solutions t for equation (9.11), or
2. l is one of the lines in $\mathcal{A}$ translated in the positive direction of the z-axis, in which case l will not intersect the surface $\mathcal{P}^h_{a,b}$ (this cannot happen for our choice of l because we assume that $p\in l\cap\mathcal{P}^h_{a,b}$), or
3. l is parallel to one of the lines in $\mathcal{A}$ (excluding the previous two cases), in which case l will puncture the surface $\mathcal{P}^h_{a,b}$ in a simple point (it will not be tangent to the surface).

9.6.3 $\mathcal{P}^h_{a,b}$ as ruled surface

Here is another fact: the hyperbolic paraboloid is a *doubly ruled surface* (like the hyperboloid of one sheet). In other words, there are two families of lines, $\mathcal{L}_1$ and $\mathcal{L}_2$ respectively, such that

$$\mathcal{P}^h_{a,b} = \bigcup_{l\in\mathcal{L}_1} l \quad \text{and} \quad \mathcal{P}^h_{a,b} = \bigcup_{l\in\mathcal{L}_2} l.$$

Every point on this surface lies on one line in $\mathcal{L}_1$ and on one line in $\mathcal{L}_2$. The generators containing the *saddle point* (with our equation, this point is the origin of the coordinate system) are visible in Figure 9.8a.

Again, one way to see where the two families of lines come from is to rearrange (9.10)

$$\frac{x^2}{a} - \frac{y^2}{b} - 2z = 0 \quad \Leftrightarrow \quad \frac{x^2}{a} - \frac{y^2}{b} = 2z \quad \Leftrightarrow \quad \left(\frac{x}{\sqrt{a}} - \frac{y}{\sqrt{b}}\right)\left(\frac{x}{\sqrt{a}} + \frac{y}{\sqrt{b}}\right) = 2z \tag{9.13}$$

Similar to the case of the hyperboloid of one sheet, we can introduce two parameters λ and μ, in order to separate the above equation:

$$\Leftrightarrow \quad l_{\lambda,\mu} : \begin{cases} \lambda\left(\frac{x}{\sqrt{a}} - \frac{y}{\sqrt{b}}\right) = 2\mu z \\ \mu\left(\frac{x}{\sqrt{a}} + \frac{y}{\sqrt{b}}\right) = \lambda \end{cases}.$$

What we end up with is a system of two equations, which are linear in x, y, z and which depend on the parameters λ and μ. For each fixed pair of parameters, λ and μ, we get a line which we denote with $l_{\lambda,\mu}$. It is easy to check that all points on such a line satisfy the equation of $\mathcal{P}^h_{a,b}$. This is the first family of lines $\mathcal{L}_1 = \{l_{\lambda,\mu} : \lambda, \mu \text{ not both zero}\}$.

The second family of generators, $\mathcal{L}_2$, is obtained if you group the terms differently:

$$\tilde{l}_{\lambda,\mu} : \begin{cases} \lambda\left(\frac{x}{\sqrt{a}} + \frac{y}{\sqrt{b}}\right) = 2\mu z \\ \mu\left(\frac{x}{\sqrt{a}} - \frac{y}{\sqrt{b}}\right) = \lambda \end{cases}.$$

As above, one can check that points on these lines satisfy the equation of our paraboloid.

Again, one important thing to notice is that, although we write down two parameters, λ and μ, we don't necessarily get distinct lines for distinct parameters: $l_{\lambda,\mu} = l_{t\lambda,t\mu}$ for any nonzero scalar t. So, in fact, $\mathcal{L}_1$ depends on one parameter. More concretely

$$\mathcal{L}_1 := \left\{ l_\alpha : \begin{cases} \left(\frac{x}{\sqrt{a}} - \frac{y}{\sqrt{b}}\right) = 2\alpha z \\ \alpha\left(\frac{x}{\sqrt{a}} + \frac{y}{\sqrt{b}}\right) = 1 \end{cases} \right\} \bigcup \left\{ l_\infty : \begin{cases} 0 = 2z \\ \left(\frac{x}{\sqrt{a}} + \frac{y}{\sqrt{b}}\right) = 0 \end{cases} \right\}$$

and similarly for $\mathcal{L}_2$. You might have noticed that l_∞ is one of the lines visible in Figure 9.8a, since it lies in the plane $z = 0$. The other one, $\tilde{l}_\infty$, belongs to the family $\mathcal{L}_2$.

9.6.4 Parametrizations - local description

A parametrization of this surface is

$$\sigma_2(\theta_1, \theta_2) = \begin{bmatrix} \sqrt{a}\theta_1 \\ \sqrt{b}\theta_2 \\ \frac{1}{2}(\theta_1^2 - \theta_2^2) \end{bmatrix} \quad \theta_1, \theta_2 \in \mathbb{R}$$

Check this and deduce a parametrization of the tangent plane $T_p\mathcal{P}^h_{a,b}$ at the point

$$p = \begin{bmatrix} x(\theta_{1,p}, \theta_{2,p}) \\ y(\theta_{1,p}, \theta_{2,p}) \\ z(\theta_{1,p}, \theta_{2,p}) \end{bmatrix} \in \mathcal{P}^h_{a,b}.$$

where $\theta_{1,p}$ and $\theta_{2,p}$ are the parameters of the point p, i.e. $p = p(x(\theta_{1,p}, \theta_{2,p}), y(\theta_{1,p}, \theta_{2,p}), z(\theta_{1,p}, \theta_{2,p}))$.