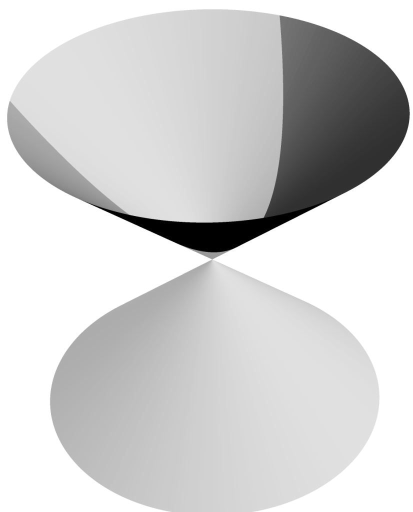
Geometry



Contents

CHAPTER 1

Affine space

Contents

1.1 Geometric Vectors

We denote by $\mathbb{E}^2$ the Euclidean plane, the usual geometric plane in which you did planar geometry. We let $\mathbb{E}^3$ denote the Euclidean space. If you didn't do geometry in dimension 3 before, don't worry we will define the n-dimensional Euclidean space $\mathbb{E}^n$ for each n. For this, we start by reviewing what vectors are.

Two points A and B in $\mathbb{E}^2$ or $\mathbb{E}^3$ can be assembled in an ordered pair (A,B). Such a pair contains the following geometric information:

- (distance) the distance from A to B,
- (direction) the direction from A to B,
- (location) the line segment $[A,B]$.

While the third bit of geometric information depends on the points A and B, the first two do not depend on these points. They can be abstracted with the following notion.

Definition 1.1. Two ordered pairs of points (A,B) and (C,D) are called *equipollent*, and we write $(A,B) \sim (C,D)$, if the line segments $[A,D]$ and $[B,C]$ have the same midpoints.

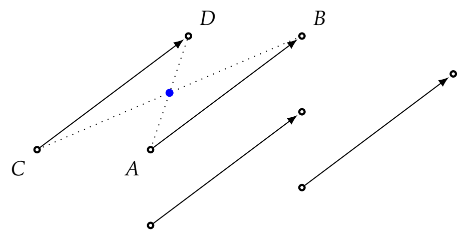


Figure 1.1: Equipolent pairs of points.

Proposition 1.2. For two ordered pairs of points (A,B) and (C,D) the following statements are equivalent:

1. $(A,B) \sim (C,D)$.
2. $ABDC$ is a parallelogram, possibly degenerate.
3. (A,B) and (C,D) define the same distance and direction.

Proposition 1.3. For any ordered pair of points (A,B) and any point O, there is a unique point X such that $(A,B) \sim (O,X)$.

Proposition 1.4. The equipollence relation is an equivalence relation.

Definition 1.5. The equivalence classes of the equipollence relation are called *vectors*. The vector containing the ordered pair (A,B) is denoted by $\overrightarrow{AB}$:

$$\overrightarrow{AB} = \left\{\text{ordered pairs } (X,Y) \text{ such that } (X,Y) \sim (A,B)\right\}.$$

We say that $\overrightarrow{AB}$ *is represented by the pair* (A,B) or that (A,B) *is a representative of the vector* $\overrightarrow{AB}$. Notice that, by definition, $\overrightarrow{AB} = \overrightarrow{CD}$ if and only if $(A,B) \sim (C,D)$. Thus, all representatives of the vector $\overrightarrow{AB}$ define the same distance and the same direction.

The set of vectors defined with points in $\mathbb{E}^2$ and $\mathbb{E}^3$ are the equivalence classes

$$\mathbb{V}^2 = \left\{\overrightarrow{AB} : (A,B) \in \mathbb{E}^2 \times \mathbb{E}^2\right\} = \mathbb{E}^2 \times \mathbb{E}^2/\sim \quad \text{respectively} \quad \mathbb{V}^3 = \left\{\overrightarrow{AB} : (A,B) \in \mathbb{E}^3 \times \mathbb{E}^3\right\} = \mathbb{E}^3 \times \mathbb{E}^3/\sim.$$

Proposition 1.6. For any point O, the map $\phi^2_O : \mathbb{E}^2 \to \mathbb{V}^2$ defined by $\phi^2_O(A) = \overrightarrow{OA}$ is a bijection between points and vectors. Similarly, we also obtain a bijective map $\phi^3_O : \mathbb{E}^3 \to \mathbb{V}^3$.

1.2 Affine space structure

In what follows we describe the vector space structure of geometric vectors. In view of Proposition 1.6, this will provide a way of defining lines, planes and higher dimensional analogues as affine spaces using real vector spaces. In doing so we don't need the notion of 'length' or 'distance'. We only make use of the arithmetic of vectors.

1.2.1 Vector space structure of geometric vectors

Definition 1.7. Consider two vectors $\mathbf{a}$ and $\mathbf{b}$. If we fix a point O then, by Proposition 1.3, there is a unique point A such that $\mathbf{a} = \overrightarrow{OA}$ and for the point A there exists a unique point X such that $\mathbf{b} = \overrightarrow{AX}$. The *sum* of $\mathbf{a}$ and $\mathbf{b}$ is by definition the vector $\overrightarrow{OX}$ and we denote the sum by $\mathbf{a}+\mathbf{b}$.

Equivalently, for a fixed point O there are unique points A and B such that $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$ and for the points O, A and B there is a unique point Y such that $OAYB$ is a parallelogram. It follows that $X = Y$ and therefore $\mathbf{a}+\mathbf{b} = \overrightarrow{OY} = \overrightarrow{OX}$.

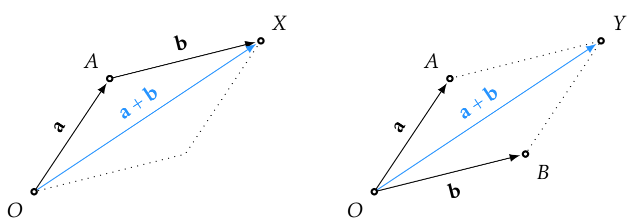


Figure 1.2: Addition of vectors.

Proposition 1.8. The addition of vectors is well defined.

Proposition 1.9. The set of vectors $\mathbb{V}^2$ with addition is a commutative group. Similarly, $(\mathbb{V}^3,+)$ is a commutative group.

Definition 1.10. Consider a vector $\mathbf{a}$ and a scalar $x \in \mathbb{R}$. If we fix a point O then there is a unique point A such that $\mathbf{a} = \overrightarrow{OA}$. If $x > 0$ then there is a unique point X on the half-line $(OA$ such that $|OB| = x|OA|$. The *multiplication* of the vector $\mathbf{a}$ *by the scalar* x, denoted $x \cdot \mathbf{a}$ (or simply $x\mathbf{a}$), is by definition

$$x \cdot \mathbf{a} = \begin{cases} \overrightarrow{OX} & \text{for } \mathbf{a} \neq 0,\ x > 0 \text{ and } X \text{ as above,} \\ -(|x|\mathbf{a}) & \text{for } \mathbf{a} \neq 0,\ x < 0, \\ \overrightarrow{0} & \text{for } \mathbf{a} = 0 \text{ or } x = 0. \end{cases}$$

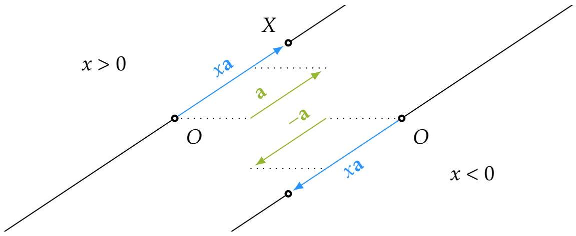


Figure 1.3: Multiplication of vectors with scalars.

Proposition 1.11. The multiplication of vectors with scalars is well defined.

Proposition 1.12. For $\mathbf{a},\mathbf{b} \in \mathbb{V}^3$ and $x,y \in \mathbb{R}$ we have

1. $(x+y)\cdot\mathbf{a} = x\cdot\mathbf{a} + y\cdot\mathbf{a}$
2. $x\cdot(\mathbf{a}+\mathbf{b}) = x\cdot\mathbf{a} + x\cdot\mathbf{b}$
3. $x\cdot(y\cdot\mathbf{a}) = (xy)\cdot\mathbf{a}$
4. $1\cdot\mathbf{a} = \mathbf{a}$.

Theorem 1.13. The set of vectors $\mathbb{V}^2$ with vector addition and scalar multiplication is a vector space. Similarly, $(\mathbb{V}^3,+,\cdot)$ is a vector space.

1.2.2 n-dimensional real affine space

The vector space structure deduced in Theorem 1.13 consists in particular of two maps

$$+ : \mathbb{V}^2 \times \mathbb{V}^2 \to \mathbb{V}^2 \quad \text{and} \quad \cdot : \mathbb{R} \times \mathbb{V}^2 \to \mathbb{V}^2$$

(similarly for $\mathbb{V}^3$). Moreover, with Proposition 1.6 we can define an ‘addition’ of vectors with points. For a vector $\mathbf{a}$ and a point O there is a unique point X such that $\mathbf{a} = \overrightarrow{OX}$, i.e. we have a map

$$+ : \mathbb{V}^2 \times \mathbb{E}^2 \to \mathbb{E}^2 \quad \text{given by} \quad \mathbf{a} + O = X.$$

We say that *the vectors in* $\mathbb{V}^2$ ***act*** *on the set of points* $\mathbb{E}^2$ *by translations*. It is this observation which allows us to use real vector spaces as an underlying model for Euclidean spaces. The underlying model is the following.

Definition 1.14. Let $\mathbb{V}^n$ be an n-dimensional real vector space. An *n-dimensional real affine space over* $\mathbb{V}^n$ is a non-empty set $\mathbb{A}^n$, whose elements are called *points* of $\mathbb{A}^n$, together with a map

$$t : \mathbb{V}^n \times \mathbb{A}^n \to \mathbb{A}^n$$

which satisfies the following two axioms:

(AS1) For every $A, B \in \mathbb{A}^n$ there is a unique $\mathbf{a} \in \mathbb{V}^n$ such that

$$B = t(\mathbf{a}, A).$$

(AS2) For every $A \in \mathbb{A}^n$ and $\mathbf{a}, \mathbf{b} \in \mathbb{V}^n$ we have

$$t(\mathbf{a}, t(\mathbf{b}, A)) = t(\mathbf{a} + \mathbf{b}, A).$$

Notice that if we fix a point $O \in \mathbb{A}^n$, by (AS1), for each point $P \in \mathbb{A}^n$ there is a unique vector $\mathbf{v}$ such that $P = t(\mathbf{v}, O)$. This vector is denoted by $\overrightarrow{OP}$ and gives a bijection $\phi_O^n : \mathbb{A}^n \to \mathbb{V}^n$.

We know that up to isomorphism there is a unique real vector space of dimension n. We write $\mathbb{V}^n$ for such a vector space and we know that $\mathbb{V}^n \cong \mathbb{R}^n$. However, since $\mathbb{R}^n$ has a standard basis, we use the notation $\mathbb{V}^n$ in order to ignore the standard basis. The main, and in fact the only example for this notion of affine space is the following: every real vector space $\mathbb{V}$ is a real affine space over itself. Indeed, we may take the set of points $\mathbb{A}$ to be $\mathbb{V}$ and the map

$$t : \mathbb{V} \times \mathbb{A} \to \mathbb{V}, \quad \text{defined by} \quad t(\mathbf{v}, P) = \mathbf{v} + P.$$

The whole point of this notion of affine space is to view the elements of the vector space $\mathbb{V}$ in two distinct ways: as points and as vectors which act on points. The reason for doing so is simple: a vector space has an origin, the zero vector, but in a plane all points are equal. One way of phrasing this is by saying that 'an affine space is nothing more than a vector space whose origin we try to forget about' [2, Chapter 2].

The geometric plane and the geometric space can be bridged with linear algebra through the following theorem which gives the basis of considering the n-dimensional affine space $\mathbb{A}^n$.

Theorem 1.15. Let S be a subset of $\mathbb{E}^2$ or $\mathbb{E}^3$ and let O be a point in S. Write ϕ_O for either ϕ_O^2 or ϕ_O^3. Then

1. the set S is a line if and only if $\phi_O(S)$ is a 1-dimensional vector subspace;
2. the vectors $\overrightarrow{OA}$, $\overrightarrow{OB}$ are linearly dependent if and only if the points O, A, B are colinear;
3. the set S is a plane if and only if $\phi_O(S)$ is a 2-dimensional vector subspace;
4. the vectors $\overrightarrow{OA}$, $\overrightarrow{OB}$, $\overrightarrow{OC}$ are linearly dependent if and only if the points O, A, B, C are coplanar;
5. if S is a line or a plane then the vector subspace $\phi_O(S)$ is independent of the choice of O in S.

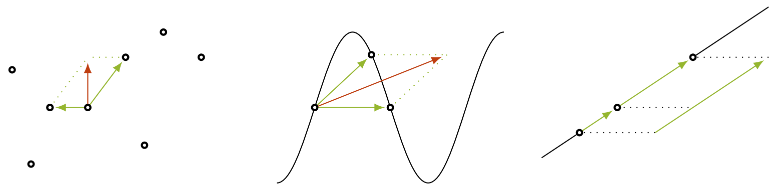


Figure 1.4: Two non-saturated sets and a line.

1.3 Cartesian coordinate systems

Coordinates are a way of associating tuples of numbers to points in an object/space. Given an object/space S we want to have a subset $C \subseteq \mathbb{R}^n$ and a bijective map $C \to S$ which sets up a correspondence $(x_1,\dots,x_n) \leftrightarrow P$ between coordinates and points. In general, there are many choices of such maps. The choice of such a map determines the control that you have over the object/space S. This in turn may depend on what you want to do with S. This semester we only consider Cartesian coordinate systems.

Cartesian coordinate systems in dimension 2. Fix two lines ℓ_1 and ℓ_2 in $\mathbb{E}^2$ which intersect in exactly one point O. We can describe any point P in $\mathbb{E}^2$ as follows. By Theorem 1.15, $\phi_O(\ell_1)$ and $\phi_O(\ell_2)$ are 1-dimensional linearly independent vector subspaces of the two dimensional vector space $\mathbb{V}^2$. Thus, if we choose $\mathbf{i} \in \phi_O(\ell_1)$ and $\mathbf{j} \in \phi_O(\ell_2)$ two non-zero vectors, then $(\mathbf{i},\mathbf{j})$ is a basis of $\mathbb{V}^2$. Hence there are *unique* scalars $x, y \in \mathbb{R}$ such that

$$\phi_O^2(P) = x\mathbf{i} + y\mathbf{j}.$$

This gives a bijection between points P in $\mathbb{E}^2$ and pairs of real numbers (x,y) in $\mathbb{R}^2$. This bijection is the composition of two bijections:

1. the map $\phi_O^2 : \mathbb{E}^2 \to \mathbb{V}^2$ giving the identification $\mathbb{E}^2 \cong \mathbb{V}^2$ and
2. the decomposition of vectors with respect to a basis giving the identification $\mathbb{V}^2 \cong \mathbb{R}^2$.

Thus, for this bijection we made two choices:

1. we chose the point O for the map ϕ_O^2, and
2. we chose two vectors $\mathbf{i}$ and $\mathbf{j}$ which form a basis of $\mathbb{V}^2$.

These choices identify points with tuples of real numbers, giving us coordinates in $\mathbb{E}^2$.

Definition 1.16. A *coordinate system for* $\mathbb{E}^2$ (or *reference frame*) is a tuple $\mathcal{K} = (O, \mathcal{B})$ where O is a point in $\mathbb{E}^2$ and $\mathcal{B} = (\mathbf{i},\mathbf{j})$ is a basis of $\mathbb{V}^2$. Given such a coordinate system, for any point $P \in \mathbb{E}^2$ there is a unique pair of scalars (x_P, y_P) such that

$$\overrightarrow{OP} = x_P\mathbf{i} + y_P\mathbf{j}.$$

The pair (x_P, y_P) is called *coordinates of P with respect to the coordinate system* $\mathcal{K}$ and we write $P_{\mathcal{K}}(x_P, y_P)$ when we want to indicate the coordinates. If it is clear from the context what $\mathcal{K}$ is we drop the subscript $\mathcal{K}$ and simply write $P(x_P, y_P)$. Most of the time, we use the letter x for the first coordinate and y for the second coordinate. We call ℓ_1 the x-axis and denote it by Ox. Similarly, ℓ_2 is the y-axis denoted by Oy. The point O is called the *origin* of the coordinate system $\mathcal{K}$.

For computational purposes, it is common to identify points and vectors with column matrices when we work with their coordinates respectively with their components:

$$[P]_{\mathcal{K}} = P(x_P, y_P) = \begin{bmatrix} x_P \\ y_P \end{bmatrix} \in \mathbb{E}^2 \quad \text{and} \quad [\mathbf{a}]_{\mathcal{B}} = a_x\mathbf{i} + a_y\mathbf{j} = \begin{bmatrix} a_x \\ a_y \end{bmatrix} \in \mathbb{V}^2.$$

With the subscript $\mathcal{K}$ we indicate the coordinate system with respect to which P has the indicated coordinates. For vectors, the subscript $\mathcal{B}$ indicates the basis with respect to which the vector $\mathbf{a}$ has components a_x and a_y.

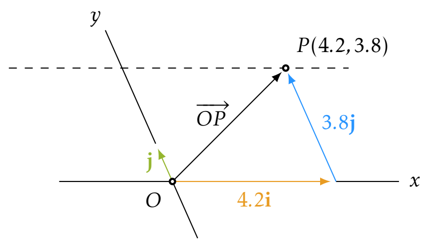


Figure 1.5: Coordinates in $\mathbb{E}^2$.

Remark. In $\mathbb{E}^2$, when we fix a coordinate system $\mathcal{K} = (O, (\mathbf{i}, \mathbf{j}))$ with $\mathbf{i}$ and $\mathbf{j}$ of length 1, we can interpret the coordinates of a point P as follows. For each such point P there is a unique parallelogram $OXPY$ with $X \in Ox$ and $Y \in Oy$ and the coordinates (x_P, y_P) of the point P with respect to $\mathcal{K}$ are the lengths of the sides of this parallelogram. Notice however that in an affine space $\mathbb{A}^n$ we don't have a notion of 'length' or 'distance'. These notions will be introduced later in Chapter 3.

Cartesian coordinate systems in dimension 3**.** Similarly, if we fix three non-coplanar lines ℓ_1, ℓ_2 and ℓ_3 in $\mathbb{E}^3$ which intersect in exactly one point O, we can describe any point P in $\mathbb{E}^3$ as follows. By Theorem 1.15, $\phi_O^3(\ell_1)$, $\phi_O^3(\ell_2)$ and $\phi_O^3(\ell_3)$ are 1-dimensional linearly independent vector subspaces of the three-dimensional vector space $\mathbb{V}^3$. Thus, if we choose $\mathbf{i} \in \phi_O^3(\ell_1)$, $\mathbf{j} \in \phi_O^3(\ell_2)$ and $\mathbf{k} \in \phi_O^3(\ell_3)$ three non-zero vectors, then $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ is a basis of $\mathbb{V}^3$. Hence there are *unique* scalars $x, y, z \in \mathbb{R}$ such that

$$\phi_O^3(P) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

This gives a bijection between points P in $\mathbb{E}^3$ and triples of real numbers (x, y, z) in $\mathbb{R}^3$. As in the case of $\mathbb{E}^2$, this bijection is determined by the choice of O, $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ and it is the composition of two bijections.

Definition 1.17. A *coordinate system for* $\mathbb{E}^3$ (or *reference frame*) is a tuple $\mathcal{K} = (O, \mathcal{B})$ where O is a point in $\mathbb{E}^3$ and $\mathcal{B} = (\mathbf{i}, \mathbf{j}, \mathbf{k})$ is a basis of $\mathbb{V}^3$. Given such a coordinate system, for any point $P \in \mathbb{E}^3$ there is a unique triple of scalars (x_P, y_P, z_P) such that

$$\overrightarrow{OP} = x_P\mathbf{i} + y_P\mathbf{j} + z_P\mathbf{k}.$$

The triple (x_P, y_P, z_P) is called *coordinates of P with respect to the coordinate system $\mathcal{K}$* and we write $P_{\mathcal{K}}(x_P, y_P, z_P)$ when we want to indicate the coordinates, or simply $P(x_P, y_P, z_P)$ if it is clear from the context what $\mathcal{K}$ is. Most of the time, we use the letter x for the first coordinate, y for the second coordinate and z for the third coordinate. We call ℓ_1 the x-axis and denote it by Ox. Similarly, ℓ_2 is the y-axis denoted by Oy and ℓ_3 is the z-axis denoted by Oz. The point O is called the *origin* of the coordinate system $\mathcal{K}$. The planes containing two coordinate axes are called *coordinate planes*. The plane containing Ox and Oy is denoted by Oxy and similarly for the other two.

Here again, for computational purposes, we identify points and vectors with column matrices when we work with their coordinates and with their components respectively:

$$[P]_{\mathcal{K}} = P(x_P, y_P, z_p) = \begin{bmatrix} x_P \\ y_P \\ z_P \end{bmatrix} \in \mathbb{E}^3 \quad \text{and} \quad [\mathbf{a}]_{\mathcal{B}} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \in \mathbb{V}^3.$$

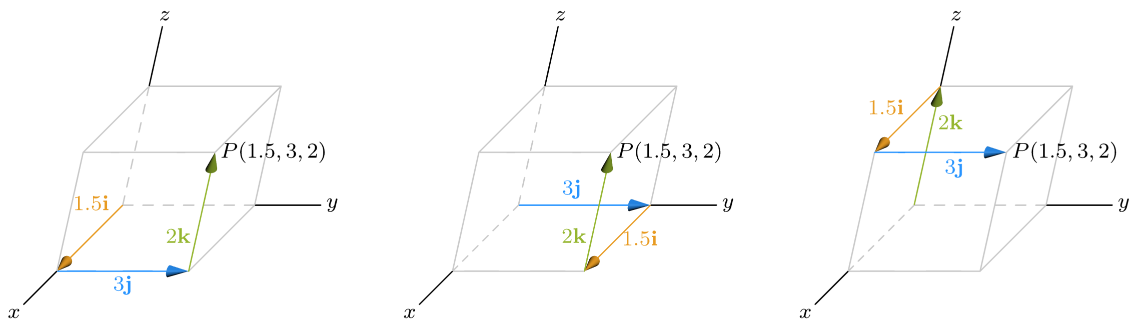


Figure 1.6: Viewing coordinates in $\mathbb{E}^3$ via the sum of rescaled basis vectors.

Remark. In $\mathbb{E}^3$, when we fix a coordinate system $\mathcal{K} = (O, (\mathbf{i}, \mathbf{j}, \mathbf{k}))$ with $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ of length 1, we can interpret the coordinates of a point P as follows. For each such point P there is a unique parallelepiped with vertices O, $X \in Ox$, $Y \in Oy$, $Z \in Oz$ such that P is opposite to O. Then the coordinates (x_P, y_P, z_P) of the point P with respect to $\mathcal{K}$ are the lengths of the sides of this parallelepiped. 'Length' and 'distance' will be introduced later in Chapter 3.

Cartesian coordinate systems in dimension n. We generalize from dimension 2 and 3 to higher dimensions with the following definition.

Definition 1.18. A *coordinate system for $\mathbb{A}^n$* (or *reference frame*) is a tuple $\mathcal{K} = (O, \mathcal{B})$ where O is a point in $\mathbb{A}^n$ and $\mathcal{B} = (\mathbf{i}_1, \dots, \mathbf{i}_n)$ is a basis of $\mathbb{V}^n$. Given such a coordinate system $\mathcal{K}$, for any point $P \in \mathbb{A}^n$ there is a unique tuple of scalars $(x_1, \dots, x_n)$ such that

$$\overrightarrow{OP} = x_1\mathbf{i}_1 + \cdots + x_n\mathbf{i}_n.$$

The n-tuple $(x_1, \dots, x_n)$ is called *coordinates of P with respect to the coordinate system $\mathcal{K}$* and we write $P_{\mathcal{K}}(x_1, \dots, x_n)$ when we want to indicate the coordinates, or simply $P(x_1, \dots, x_n)$ if it is clear from the context what $\mathcal{K}$ is. The point O is called the *origin* of the coordinate system $\mathcal{K}$. Here again, for computational purposes, we identify points and vectors with column matrices when we work with their coordinates and with their components respectively:

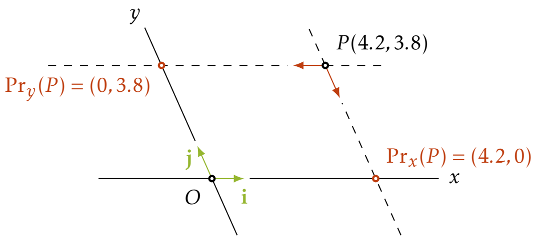
$$[P]_{\mathcal{K}} = P(x_1,\dots,x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{A}^n \quad \text{and} \quad [\mathbf{a}]_{\mathcal{B}} = a_1\mathbf{i}_1 + \cdots + a_n\mathbf{i}_n = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{V}^n.$$

1.3.1 Cartesian coordinates as projections

We deal with projections in larger generality in Chapter 5. Here we notice that coordinates may be viewed as projections on coordinate axes. Let $\mathcal{K} = (O,\mathcal{B})$ be a coordinate system of $\mathbb{E}^2$ where $\mathcal{B} = (\mathbf{i},\mathbf{j})$ is a basis of $\mathbb{V}^2$. The correspondence between points and coordinates in the frame $\mathcal{K}$

$$P \leftrightarrow \begin{bmatrix} x_P \\ y_P \end{bmatrix}_{\mathcal{K}}$$

automatically gives a map $\mathrm{Pr}_x : \mathbb{E}^2 \to Ox$ defined by $\mathrm{Pr}_x(P) = (x_P, 0)$ and a map $\mathrm{Pr}_y : \mathbb{E}^2 \to Oy$ defined by $\mathrm{Pr}_y(P) = (0, y_P)$. The map Pr_x is called *the projection on Ox along Oy* and Pr_y is called *the projection on Oy along Ox*.



Similarly for vectors: considering the basis $\mathcal{B} = (\mathbf{i},\mathbf{j})$ of $\mathbb{V}^2$, the correspondence

$$\mathbf{a} \leftrightarrow \begin{bmatrix} a_x \\ a_y \end{bmatrix}_{\mathcal{B}}$$

gives similar maps $\mathrm{pr}_x : \mathbb{V}^2 \to \mathbb{R}$ defined by $\mathrm{pr}_x(\mathbf{a}) = a_x$ and $\mathrm{pr}_y : \mathbb{V}^2 \to \mathbb{R}$ defined by $\mathrm{pr}_y(\mathbf{a}) = a_y$. These maps are related by

$$\overrightarrow{O\mathrm{Pr}_x(P)} = \mathrm{pr}_x(\overrightarrow{OP})\mathbf{i}, \quad \overrightarrow{O\mathrm{Pr}_y(P)} = \mathrm{pr}_y(\overrightarrow{OP})\mathbf{j} \quad \text{and by} \quad \overrightarrow{OP} = \mathrm{pr}_x(\overrightarrow{OP})\mathbf{i} + \mathrm{pr}_y(\overrightarrow{OP})\mathbf{j}.$$

Thus

$$[P]_{\mathcal{K}} = P(x_P, y_P) = \begin{bmatrix} x_P \\ y_P \end{bmatrix} = \begin{bmatrix} \mathrm{pr}_x(\overrightarrow{OP}) \\ \mathrm{pr}_y(\overrightarrow{OP}) \end{bmatrix}.$$

In dimension 3 we have maps defined by $\mathrm{Pr}_x(P) = (x_P, 0, 0)$, $\mathrm{Pr}_y(P) = (0, y_p, 0)$ and $\mathrm{Pr}_z(P) = (0, 0, z_P)$. The map Pr_x is called *the projection on Ox along the plane Oyz* and similarly for the other two.

These are projections on coordinate axes along coordinate planes. For vectors in $\mathbb{V}^3$ we have maps $\mathrm{pr}_x, \mathrm{pr}_y, \mathrm{pr}_z : \mathbb{V}^3 \to \mathbb{R}$ defined by $\mathrm{pr}_x(\mathbf{a}) = a_x$, $\mathrm{pr}_y(\mathbf{a}) = a_y$, $\mathrm{pr}_z(\mathbf{a}) = a_z$. They are related by

$$\overrightarrow{O\mathrm{Pr}_x(P)} = \mathrm{pr}_x(\overrightarrow{OP})\mathbf{i}, \quad \overrightarrow{O\mathrm{Pr}_y(P)} = \mathrm{pr}_y(\overrightarrow{OP})\mathbf{j}, \quad \overrightarrow{O\mathrm{Pr}_z(P)} = \mathrm{pr}_z(\overrightarrow{OP})\mathbf{k}$$

$$\text{and} \quad \overrightarrow{OP} = \mathrm{pr}_x(\overrightarrow{OP})\mathbf{i} + \mathrm{pr}_y(\overrightarrow{OP})\mathbf{j} + \mathrm{pr}_z(\overrightarrow{OP})\mathbf{k}.$$

Thus

$$[P]_{\mathcal{K}} = P(x_P, y_P, z_P) = \begin{bmatrix} x_P \\ y_P \\ z_P \end{bmatrix} = \begin{bmatrix} \mathrm{pr}_x(\overrightarrow{OP}) \\ \mathrm{pr}_y(\overrightarrow{OP}) \\ \mathrm{pr}_z(\overrightarrow{OP}) \end{bmatrix}.$$

This easily generalizes to dimension n where we have

$$[P]_{\mathcal{K}} = P(x_1, \ldots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \mathrm{pr}_1(\overrightarrow{OP}) \\ \vdots \\ \mathrm{pr}_n(\overrightarrow{OP}) \end{bmatrix}.$$

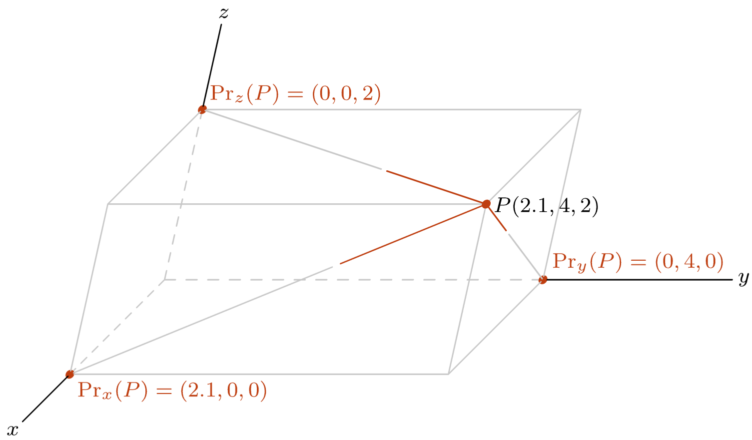


Figure 1.7: Viewing coordinates in $\mathbb{E}^3$ via projections on coordinate axes.

Remark. With a fixed coordinate system $\mathcal{K}$ of $\mathbb{E}^2$, the correspondence

$$P \leftrightarrow \begin{bmatrix} x_P \\ y_P \end{bmatrix}$$

identifies points with tuples of real numbers. A pair of numbers (x_P, y_P) has no geometric meaning in the absence of $\mathcal{K}$. Moreover, in a different coordinate system $\mathcal{K}'$ the point P will correspond to some other pair (x_P', y_P'):

$$[P]_{\mathcal{K}} = \begin{bmatrix} x_P \\ y_P \end{bmatrix} \quad \text{and} \quad [P]_{\mathcal{K}'} = \begin{bmatrix} x_P' \\ y_P' \end{bmatrix}.$$

So, when there is more than one coordinate system around, we need to understand how to translate the coordinates from one coordinate system to the other coordinate system.

1.3.2 Changing the basis in a vector space

Let $\phi : V \to W$ be a linear map between the vector spaces V and W. Let $\mathcal{E} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be a basis for V and let $\mathcal{F}$ be a basis for W. In your Algebra course, you used the notation $[\phi]_{\mathcal{E},\mathcal{F}}$ for the matrix of the linear map ϕ with respect to the bases $\mathcal{E}$ and $\mathcal{F}$ [4, Definition 3.4.1]. We will use the notation

$$\mathrm{M}_{\mathcal{F},\mathcal{E}}(\phi) = [\phi]_{\mathcal{E},\mathcal{F}}.$$

Notice that the indices $\mathcal{E}, \mathcal{F}$ are *reversed*. Recall that this is the matrix whose columns are the coordinates of the $\phi(\mathbf{e}_i)$'s with respect to the basis $\mathcal{F}$:

$$\mathrm{M}_{\mathcal{F},\mathcal{E}}(\phi) = \begin{bmatrix} \uparrow & \ldots & \uparrow \\ [\phi(\mathbf{e}_1)]_{\mathcal{F}} & \ldots & [\phi(\mathbf{e}_n)]_{\mathcal{F}} \\ \downarrow & \ldots & \downarrow \end{bmatrix}$$

You have also learned [4, Theorem 3.4.8] that if $\psi : W \to U$ is another linear map, to some vector space U with basis $\mathcal{G}$, then

$$\mathrm{M}_{\mathcal{G},\mathcal{E}}(\psi \circ \phi) = \mathrm{M}_{\mathcal{G},\mathcal{F}}(\psi) \cdot \mathrm{M}_{\mathcal{F},\mathcal{E}}(\phi).$$

In particular, if $V = W = U$, $\mathcal{G} = \mathcal{F}$ and $\phi = \psi = \mathrm{Id}_V$ then

$$I_n = \mathrm{M}_{\mathcal{F},\mathcal{F}}(\mathrm{Id}_V) = \mathrm{M}_{\mathcal{F},\mathcal{F}}(\mathrm{Id}_V \circ \mathrm{Id}_V) = \mathrm{M}_{\mathcal{F},\mathcal{E}}(\mathrm{Id}_V) \cdot \mathrm{M}_{\mathcal{E},\mathcal{F}}(\mathrm{Id}_V)$$

hence $\mathrm{M}_{\mathcal{F},\mathcal{E}}(\mathrm{Id}_V) = \mathrm{M}_{\mathcal{E},\mathcal{F}}(\mathrm{Id}_V)^{-1}$ and thus

$$\mathrm{M}_{\mathcal{F},\mathcal{F}}(\phi) = \mathrm{M}_{\mathcal{F},\mathcal{E}}(\mathrm{Id}_V) \cdot \mathrm{M}_{\mathcal{E},\mathcal{E}}(\phi) \cdot \mathrm{M}_{\mathcal{E},\mathcal{F}}(\mathrm{Id}_V) = \mathrm{M}_{\mathcal{E},\mathcal{F}}(\mathrm{Id}_V)^{-1} \cdot \mathrm{M}_{\mathcal{E},\mathcal{E}}(\phi) \cdot \mathrm{M}_{\mathcal{E},\mathcal{F}}(\mathrm{Id}_V).$$

So, the matrix of ϕ with respect to the basis $\mathcal{F}$ is obtained from the matrix of ϕ with respect to the basis $\mathcal{E}$ by conjugating with the matrix $\mathrm{M}_{\mathcal{E},\mathcal{F}}(\mathrm{Id}_V)$.

Definition 1.19. The matrix $\mathrm{M}_{\mathcal{E},\mathcal{F}} := \mathrm{M}_{\mathcal{E},\mathcal{F}}(\mathrm{Id}_V)$ is called the *change of basis matrix from the basis $\mathcal{F}$ to the basis $\mathcal{E}$*. It is the matrix whose columns are the coordinates of the vectors in $\mathcal{F}$ with respect to $\mathcal{E}$. If $\mathcal{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_n)$ then

$$\mathrm{M}_{\mathcal{E},\mathcal{F}} = \begin{bmatrix} \uparrow & \ldots & \uparrow \\ [\mathbf{f}_1]_{\mathcal{E}} & \ldots & [\mathbf{f}_n]_{\mathcal{E}} \\ \downarrow & \ldots & \downarrow \end{bmatrix}.$$

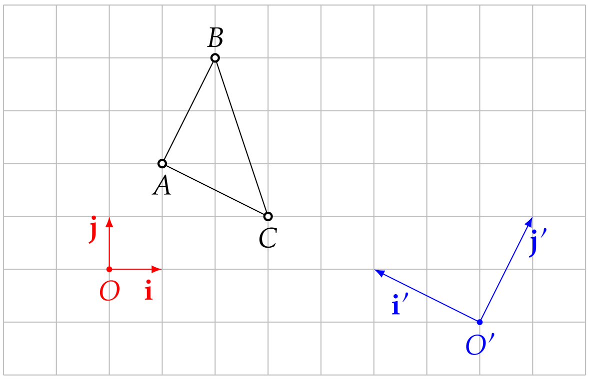
If $\mathcal{K} = (O, \mathcal{E})$ and $\mathcal{K}' = (Q, \mathcal{F})$ are two coordinate systems, we sometimes write $\mathrm{M}_{\mathcal{K},\mathcal{K}'}$ for the base change matrix $\mathrm{M}_{\mathcal{E},\mathcal{F}}$.

1.3.3 Changing Cartesian coordinate systems

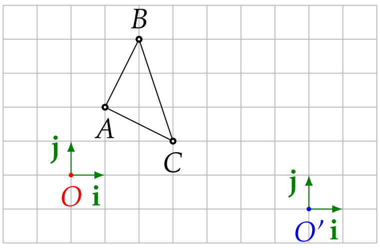
Example 1.20 (In dimension 2)**.** Let $\mathcal{K} = (O, (\mathbf{i}, \mathbf{j}))$ and $\mathcal{K}' = (O', (\mathbf{i}', \mathbf{j}'))$ be two coordinate systems (reference frames) of $\mathbb{E}^2$. Suppose that we know O', $\mathbf{i}'$ and $\mathbf{j}'$ relative to $\mathcal{K}$:

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

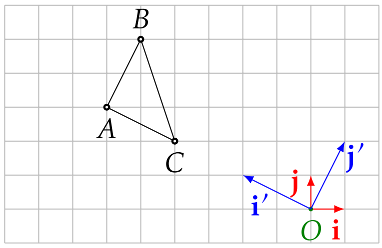
How can we translate the coordinates of points from $\mathcal{K}$ to $\mathcal{K}'$?



We can do this in two steps: (a) first we change the origin, i.e. we go from $(O,\mathbf{i},\mathbf{j})$ to $(O',\mathbf{i},\mathbf{j})$ and (b) we change the directions of the coordinate axes, i.e. we go from $(O',\mathbf{i},\mathbf{j})$ 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 of $\mathcal{K}'$ to the basis of $\mathcal{K}$. Then

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

Thus, 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, we notice the following formula for changing coordinates from the system $\mathcal{K}$ to the system $\mathcal{K}'$

$$[A]_{\mathcal{K}'} = \mathrm{M}_{\mathcal{K}',\mathcal{K}}([A]_{\mathcal{K}} - [O']_{\mathcal{K}}) = \mathrm{M}_{\mathcal{K}',\mathcal{K}}[A]_{\mathcal{K}} + [O]_{\mathcal{K}'} \tag{1.1}$$

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

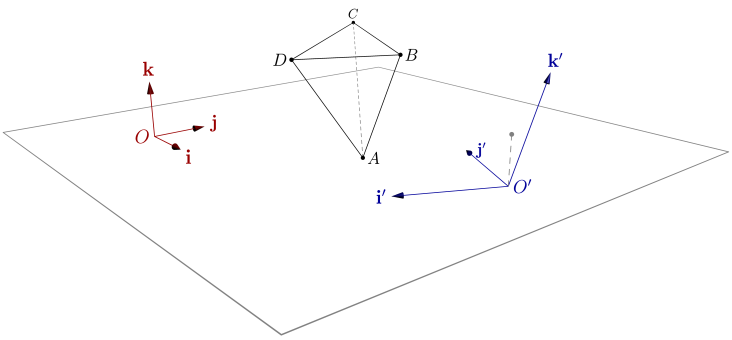
$$[A]_{\mathcal{K}'} = \mathrm{M}_{\mathcal{K}',\mathcal{K}}\begin{bmatrix}1\\2\end{bmatrix} + [O]_{\mathcal{K}'}$$

Since we know $\mathbf{i}'$ and $\mathbf{j}'$ with respect to $\mathbf{i}$ and $\mathbf{j}$, we can write down the matrix $M_{\mathcal{K},\mathcal{K}'}$ and then $M_{\mathcal{K}',\mathcal{K}} = M^{-1}_{\mathcal{K},\mathcal{K}'}$. Since we know the coordinates of O' with respect to $\mathcal{K}$, it is more convenient to use the first equality in (1.1)

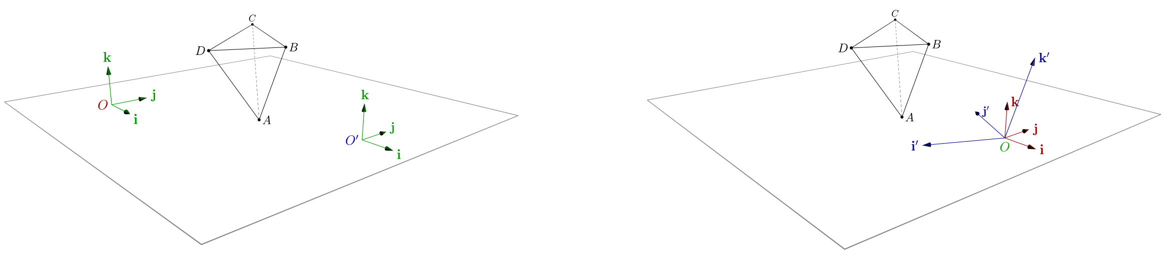
$$[A]_{\mathcal{K}'} = \begin{bmatrix} -2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \cdot \left(\begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 7 \\ -1 \end{bmatrix} \right) = \frac{1}{-5} \begin{bmatrix} 2 & -1 \\ -1 & -2 \end{bmatrix} \cdot \left(\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix}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

Example 1.21 (In dimension 3)**.** Let $\mathcal{K} = (O, (\mathbf{i}, \mathbf{j}, \mathbf{k}))$ and $\mathcal{K}' = (O', (\mathbf{i}', \mathbf{j}', \mathbf{k}'))$ be reference frames of $\mathbb{E}^3$. Suppose that we know O', $\mathbf{i}'$, $\mathbf{j}'$ and $\mathbf{k}'$ relative to $\mathcal{K}$:

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



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

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

In general, for changing reference frames in an n-dimensional affine space, we have the following result.

Theorem 1.22. Let $\mathcal{K} = (O,\mathcal{E})$ and $\mathcal{K}' = (Q,\mathcal{F})$ be two coordinate systems of $\mathbb{A}^n$. For any point $P \in \mathbb{A}^n$ we have

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

1.4 Exercises

1.1. Let $A_0,\dots,A_n$ be the vertices of a polygon. Determine $\overrightarrow{A_0A_1} + \overrightarrow{A_1A_2} + \cdots + \overrightarrow{A_{n-1}A_n} + \overrightarrow{A_nA_0}$.

1.2. In each of the following cases, decide if the indicated vectors $\mathbf{u},\mathbf{v},\mathbf{w}$ can be represented with the vertices of a triangle:

a) $\mathbf{u}(7,3)$, $\mathbf{v}(2,8)$, $\mathbf{w}(-5,5)$.

b) $\mathbf{u}(1,2,-1)$, $\mathbf{v}(2,-1,0)$, $\mathbf{w}(1,-3,1)$.

1.3. Let $ABCD$ be a quadrilateral. Let M, N, P, Q be the midpoints of $[AB]$, $[BC]$, $[CD]$ and $[DA]$ respectively. Show that

$$\overrightarrow{MN} + \overrightarrow{PQ} = 0.$$

Deduce that the midpoints of the sides of an arbitrary quadrilateral form a parallelogram.

1.4. Let $ABCD$ be a quadrilateral. Let E be the midpoint of $[AC]$ and let F be the midpoint of $[BD]$. Show that

$$\overrightarrow{EF} = \frac{1}{2}\left(\overrightarrow{AB} + \overrightarrow{CD}\right) = \frac{1}{2}\left(\overrightarrow{AD} + \overrightarrow{CB}\right).$$

1.5. Let $ABCD$ be a quadrilateral. Let E be the midpoint of $[AB]$ and let F be the midpoint of $[CD]$. Show that

$$\overrightarrow{EF} = \frac{1}{2}\left(\overrightarrow{AD} + \overrightarrow{BC}\right).$$

1.6. Let A', B' and C' be midpoints of the sides of a triangle ABC. Show that for any point O we have

$$\overrightarrow{OA'} + \overrightarrow{OB'} + \overrightarrow{OC'} = \overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC}.$$

1.7. Show that the medians in a triangle intersect in one point and deduce the ratio in which the common intersection point divides the medians.

1.8. In each of the following cases, decide if the given points are collinear:

a) $P(3,-5)$, $Q(-1,2)$, $R(-5,9)$.

b) $A(11,2)$, $B(1,-3)$, $C(31,13)$.

c) $P(1,0,-1)$, $Q(0,-1,2)$, $R(-1,-2,5)$.

d) $A(-1,-1,-4)$, $B(1,1,0)$, $C(2,2,2)$.

1.9. Let $ABCD$ be a tetrahedron. Determine the sums

a) $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD}$,

b) $\overrightarrow{AD} + \overrightarrow{BC} + \overrightarrow{DB}$,

c) $\overrightarrow{AB} + \overrightarrow{CD} + \overrightarrow{BC} + \overrightarrow{DA}$.

1.10. Let $ABCD$ be a tetrahedron. Show that

$$\overrightarrow{AD} + \overrightarrow{BC} = \overrightarrow{BD} + \overrightarrow{AC}.$$

1.11. Let $SABCD$ be a pyramid with apex S and base the parallelogram $ABCD$. Show that

$$\overrightarrow{SA} + \overrightarrow{SB} + \overrightarrow{SC} + \overrightarrow{SD} = 4\overrightarrow{SO}$$

where O is the center of the parallelogram.

1.12. Give the coordinates of the vertices of the parallelepiped whose faces lie in the coordinate planes and in the planes $x = 1$, $y = 3$ and $z = -2$.

1.13. In $\mathbb{E}^3$ consider the parallelograms $A_1A_2A_3A_4$ and $B_1B_2B_3B_4$. Show that the midpoints of the segments $[A_1B_1]$, $[A_2B_2]$, $[A_3B_3]$ and $[A_4B_4]$ are the vertices of a parallelogram.

1.14. Which of the following sets of vectors form a basis?

a) $\mathbf{v}(1,2)$, $\mathbf{w}(3,4)$;

b) $\mathbf{u}(-1,2,1)$, $\mathbf{v}(2,1,1)$, $\mathbf{w}(1,0,-1)$;

c) $\mathbf{u}(-1,2,1)$, $\mathbf{v}(2,1,1)$, $\mathbf{w}(0,5,3)$;

d) $\mathbf{v}_1(-1,2,1,0)$, $\mathbf{v}_2(2,1,1,0)$, $\mathbf{v}_3(1,0-1,1)$, $\mathbf{v}_4(1,0,0,1)$;

1.15. With respect to the basis $\mathcal{B} = (\mathbf{i},\mathbf{j},\mathbf{k})$ consider the vectors $\mathbf{u} = \mathbf{i}+\mathbf{j}$, $\mathbf{v} = \mathbf{j}+\mathbf{k}$ and $\mathbf{w} = \mathbf{i}+\mathbf{k}$. Check that $\mathcal{B}' = (\mathbf{u},\mathbf{v},\mathbf{w})$ is a basis and give the base change matrix $\mathrm{M}_{\mathcal{B}',\mathcal{B}}$.

1.16. Consider the two coordinate systems $\mathcal{K} = (O,\mathbf{i},\mathbf{j})$ and $\mathcal{K}' = (O',\mathbf{i}',\mathbf{j}')$ given in Example 1.20. Determine the base change matrix from $\mathcal{K}$ to $\mathcal{K}'$ and the coordinates of the points

$$[A]_{\mathcal{K}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad [B]_{\mathcal{K}} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad [C]_{\mathcal{K}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

in the system $\mathcal{K}'$. Further, determine the base change matrix from $\mathcal{K}'$ to $\mathcal{K}$ and use it with the previously obtained coordinates to calculate $[A]_{\mathcal{K}}$, $[B]_{\mathcal{K}}$ and $[C]_{\mathcal{K}}$.

1.17. Consider the tetrahedron $ABCD$ and the coordinate systems

$$\mathcal{K}_A = (A, \overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}), \quad \mathcal{K}'_A = (A, \overrightarrow{AB}, \overrightarrow{AD}, \overrightarrow{AC}), \quad \mathcal{K}_B = (B, \overrightarrow{BA}, \overrightarrow{BC}, \overrightarrow{BD}).$$

Determine

a) the coordinates of the vertices of the tetrahedron in the three coordinate systems,

b) the base change matrix from $\mathcal{K}_A$ to $\mathcal{K}'_A$,

c) the base change matrix from $\mathcal{K}_B$ to $\mathcal{K}_A$.

1.18. Consider the two coordinate systems $\mathcal{K} = (O,\mathbf{i},\mathbf{j},\mathbf{k})$ and $\mathcal{K}' = (O',\mathbf{i}',\mathbf{j}',\mathbf{k}')$ given in Example 1.21. Determine the base change matrix from $\mathcal{K}$ to $\mathcal{K}'$ and the coordinates of the points

$$[A]_{\mathcal{K}} = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}, \quad [B]_{\mathcal{K}} = \begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix}, \quad [C]_{\mathcal{K}} = \begin{bmatrix} -3 \\ 7 \\ 1 \end{bmatrix}, \quad [D]_{\mathcal{K}} = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$$

in the coordinate system $\mathcal{K}'$. Further, determine the base change matrix from $\mathcal{K}'$ to $\mathcal{K}$ and use it with the previously determined coordinates to calculate $[A]_{\mathcal{K}}$, $[B]_{\mathcal{K}}$, $[C]_{\mathcal{K}}$ and $[D]_{\mathcal{K}}$.

CHAPTER 2

Affine subspaces

Contents

2.1 Lines in $\mathbb{A}^2$

As discussed in Section 1.2, the Euclidean plane $\mathbb{E}^2$ is a 2-dimensional real affine space. In order to emphasize the fact that we treat $\mathbb{E}^2$ as an affine space only (i.e. without a notion of 'distance') we denote it by $\mathbb{A}^2$. With this perspective, a line in $\mathbb{A}^2$ is a set of points S such that the set of vectors which can be represented by points in S form a 1-dimensional vector subspace of $\mathbb{V}^2$ (see Theorem 1.15). In terms of the map $\phi^2_Q : \mathbb{A}^2 \to \mathbb{V}^2$ which identifies points with vectors when a point $Q \in \mathbb{A}^2$ is fixed, the subset $S \subseteq \mathbb{A}^2$ is a line if and only if for a point $Q \in S$

$$\phi^2_Q(S) = \left\{ \overrightarrow{QP} : B \in S \right\} \quad \text{is a 1-dimensional vector subspace of } \mathbb{V}^2.$$

It is not dificult to see that if the above description holds for one points $Q \in \mathbb{A}^2$, it holds for any point $Q \in \mathbb{A}^2$. If S is a line, we call the vector subspace $\phi^2_Q(S)$ of $\mathbb{V}^2$ the *direction space* of the line S and denote it $\mathrm{D}(S)$.

2.1.1 Parametric equations

If S is a line, then for any two distinct points P, Q in S the vector $\overrightarrow{QP}$ is called a *direction vector* of S. Since $\mathrm{D}(S)$ is 1-dimensional, all direction vectors are linearly dependent and $\mathbf{v}$ is a direction vector for S if and only if it is linearly dependent on $\overrightarrow{QP}$. So, for any direction vector $\mathbf{v}$ of S there is a *unique* scalar $t \in \mathbb{R}$ such that

$$\overrightarrow{QP} = t\mathbf{v}.$$

Now, if you fix Q and let P vary on the line, then t varies in $\mathbb{R}$. Since $\phi^2_Q : \mathbb{A}^2 \to \mathbb{V}^2$ is a bijection, the line S can be described as

$$S = \left\{ P \in \mathbb{A}^2 : \overrightarrow{QP} = t\mathbf{v} \text{ for some } t \in \mathbb{R} \right\}.$$

In this description, the point Q is arbitrary but fixed. If we want to emphasize that the description depends on fixing Q, we refer to this point as the *base point* of the line. Moreover, for any point $O \in \mathbb{A}^2$ we may split $\overrightarrow{QP}$ in the equation $\overrightarrow{QP} = t\mathbf{v}$ to obtain

$$\overrightarrow{OP} = \overrightarrow{OQ} + t\mathbf{v}. \tag{2.1}$$

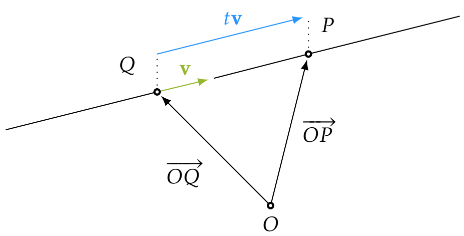


Figure 2.1: Vector equation of a line.

So, we can describe the line S as the set of points P in $\mathbb{A}^2$ which satisfy Equation (2.1) for some $t \in \mathbb{R}$. This equation is called the *vector equation of the line S relative to O, having base point Q and direction vector* $\mathbf{v}$, or simply *vector equation* of the line S.

Notice that, the vector equation depends on the choice of the base point Q and on the choice of the direction vector $\mathbf{v}$. In particular, a line does not have a unique vector equation. Notice also that, the vector equation does not depend on the coordinate system. In the above description O can be any point in $\mathbb{A}^2$.

Now fix a coordinate system $\mathcal{K} = (O, \mathcal{B})$. If we write Equation (2.1) in coordinates relative to $\mathcal{K}$, we obtain

$$S : \begin{cases} x = x_Q + t v_x \\ y = y_Q + t v_y \end{cases} \quad \text{or, in matrix form} \quad S : \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_Q \\ y_Q \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \end{bmatrix} \tag{2.2}$$

where $Q = Q(x_Q, y_Q)$, $\mathbf{v} = \mathbf{v}(v_x, v_y)$ and where t is the *parameter* - for different values t we obtain different points (x, y) on the line. The two equations in the System (2.2) are called *parametric equations* of the line S. Traditionally, they are written in the form of a system of equations as indicated on the left. Writing them as one equation, as indicated on the right, is closer to the computational perspective where we identify points with column matrices. Clearly, the two ways of writing such parametric equations are equivalent.

2.1.2 Cartesian equations

It is possible to eliminate the parameter t in (2.2). By expressing t in both equations and setting the two expressions, equal we obtain

$$\frac{x - x_Q}{v_x} = \frac{y - y_Q}{v_y}. \tag{2.3}$$

We refer to Equation (2.3) as *symmetric equation* of the line S. It could happen that v_x or v_y are zero. In that case, translate back to the parametric equations to understand what happens.

Example 2.1. The line with symmetric equation

$$\frac{x-3}{2} = \frac{y-5}{0} \quad \text{has parametric equations} \quad \begin{cases} x = 3 + 2 \cdot t \\ y = 5 + 0 \cdot t \end{cases}.$$

Thus, it is the line parallel to Ox with equation $y = 5$.

We have just described a line with a linear equation relative to the coordinate system $\mathcal{K}$. The converse is also true.

Proposition 2.2. Every line in $\mathbb{A}^2$ can be described with a linear equation in two variables

$$ax + by + c = 0 \tag{2.4}$$

relative to a fixed coordinate system and any linear equation in two variables relative to a fixed coordinate system describes a line if the constants a, b are not both zero. Moreover, if Equation (2.4) describes the line ℓ relative to a coordinate system $\mathcal{K} = (O, \mathcal{B})$, then the direction space $\mathrm{D}(\ell)$ of the line is the 1-dimensional subspace of $\mathbb{V}^2$ which, relative to the basis $\mathcal{B}$, satisfies the equation

$$\mathrm{D}(\ell) : ax + by = 0.$$

Equation (2.4) is called a *Cartesian equation* of the line which it describes. Notice that there are infinitely many Cartesian equations describing the same line, since you can multiply one equation by a non-zero constant. It is sometimes useful to rearrange the linear equation (2.4) in order to emphasize some geometric properties. For example, you can rearrange it in the form

$$\frac{x}{\alpha}+\frac{y}{\beta}=1 \quad \text{where} \quad \alpha=-\frac{c}{a} \quad \text{and} \quad \beta=-\frac{c}{b}.$$

In this form we have the *equation of the line where we can read off the intersection points with the coordinate axes* since this line intersects Ox in $(\alpha,0)$ and it intersects Oy in $(0,\beta)$.

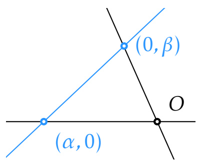


Figure 2.2: Line in $\mathbb{A}^2$ via its intersection with the coordinate axes.

2.1.3 Relative positions of two lines in $\mathbb{A}^2$

The tools of linear algebra readily apply to describe intersections of lines in $\mathbb{A}^2$. Assume that we have two lines

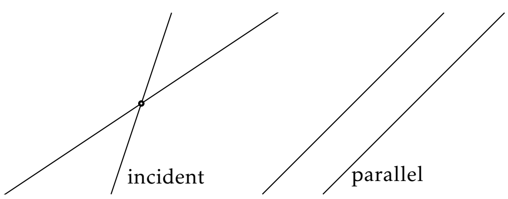
$$\ell_1 : a_1x+b_1y+c_1=0 \quad \text{and} \quad \ell_2 : a_2x+b_2y+c_2=0.$$

In order to determine if they intersect, one has to discuss the system:

$$\begin{cases} \ell_1 : a_1x+b_1y+c_1 &= 0 \\ \ell_2 : a_2x+b_2y+c_2 &= 0 \end{cases}. \tag{2.5}$$

Discussing this system is basic linear algebra (see for example [4, Section 3.6]). In the plane the situation is very simple:

- two lines intersect in a unique point, the coordinates of which are the solution to (2.5); or
- they don't intersect and (2.5) doesn't have solutions, in which case the lines are parallel; or
- System (2.5) has infinitely many solutions in which cases $\ell_1=\ell_2$.



2.1.4 Bundle of lines in $\mathbb{A}^2$

Definition 2.3. Fix a point $Q \in \mathbb{A}^2$. The set $\mathcal{L}_Q$ of all lines in $\mathbb{A}^2$ passing through Q is called a *bundle of lines* and Q is called the *center* of the bundle $\mathcal{L}_Q$.

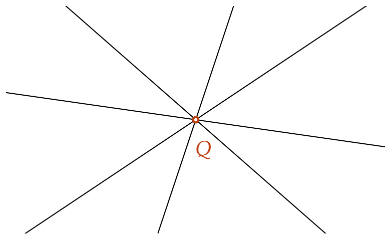


Figure 2.3: Lines in a bundle with center Q.

Proposition 2.4. If $\ell_1 : a_1x + b_1y + c_1 = 0$ and $\ell_2 : a_2x + b_2y + c_2 = 0$ are two distinct lines in the bundle $\mathcal{L}_Q$, then $\mathcal{L}_Q$ consists of lines having equations of the form

$$\ell_{\lambda,\mu} : \lambda(a_1x + b_1y + c_1) + \mu(a_2x + b_2y + c_2) = 0.$$

where $\lambda, \mu \in \mathbb{R}$ are not both zero. In particular, if $Q = Q(x_0, y_0)$, $\ell_1 : x = x_0$ and $\ell_2 : y = y_0$ then

$$\mathcal{L}_Q = \Big\{\ell_{\lambda,\mu} : \lambda(x - x_0) + \mu(y - y_0) = 0 : \lambda, \mu \in \mathbb{R} \text{ not both zero } \Big\}.$$

Bundles of lines are useful when a point Q is given as the intersection of two lines, but its coordinates are not known explicitly, and one wants to find the equation of a line passing through Q and satisfying some other conditions. For example, the condition that it contains some point P distinct from Q, or that it is parallel to a given line.

Notice that there is redundancy in the two parameters λ, μ, meaning that we don't have two independent parameters here. If $\lambda \neq 0$ then one can divide the equation of $\ell_{\lambda,\mu}$ by λ to obtain

$$\ell_{1,t} : (a_1x + b_1y + c_1) + t(a_2x + b_2y + c_2) = 0.$$

where $t = \frac{\mu}{\lambda} \in \mathbb{R}$. So $\ell_{1,\frac{\mu}{\lambda}}$ and $\ell_{\lambda,\mu}$ are in fact the same lines.

Definition 2.5. A *reduced bundle* is the set of all lines $\mathcal{L}_Q$ passing through a common point Q from which we remove one line, i.e. it is $\mathcal{L}_Q \setminus \{\ell_2\}$ for some $\ell_2 \in \mathcal{L}_Q$. With the above notation and discussion, it is the set

$$\Big\{\ell_{1,t} : (a_1x + b_1y + c_1) + t(a_2x + b_2y + c_2) = 0 : t \in \mathbb{R}\Big\}.$$

The fact that we use one parameter instead of two, to describe almost all lines passing through Q, simplifies calculations.

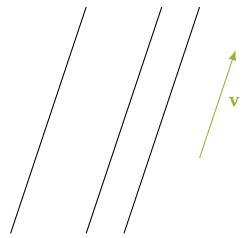


Figure 2.4: Lines in the improper bundle $\mathcal{L}_{\mathbf{v}}$.

Definition 2.6. Let $\mathbf{v} \in \mathbb{V}^2$. The set $\mathcal{L}_{\mathbf{v}}$ of all lines in $\mathbb{A}^2$ with direction vector $\mathbf{v}$ is called an *improper bundle of lines*, and $\mathbf{v}$ is called a *direction vector* of the bundle $\mathcal{L}_{\mathbf{v}}$.

The connection between bundles of lines and improper bundles of lines is best understood through projective geometry, where the improper bundle of lines is the set of all lines intersecting in the same point at infinity.

2.2 Planes in $\mathbb{A}^3$

The usual Euclidean space $\mathbb{E}^3$ is a 3-dimensional real affine space (see Section 1.2). In order to emphasize the fact that we treat $\mathbb{E}^3$ as an affine space only (i.e. without a notion of 'distance') we denote it by $\mathbb{A}^3$. A plane in $\mathbb{A}^3$ is a set of points S such that the set of vectors which can be represented by points in S form a 2-dimensional vector subspace of $\mathbb{V}^3$ (see Theorem 1.15). Considering the bijection $\phi_Q^3 : \mathbb{A}^3 \to \mathbb{V}^3$ for a point $Q \in \mathbb{A}^3$, the subset S is a plane if and only if for any $Q \in S$

$$\phi_Q^3(S) = \{\overrightarrow{QP} : P \in S\} \quad \text{is a 2-dimensional vector subspace of } \mathbb{V}^3.$$

It is not dificult to see that the above description does not depend one the point $Q \in S$. If S is a plane, we call the vector subspace $\phi_Q^3(S)$ of $\mathbb{V}^3$ the *direction space* of the plane S and denote it $\mathrm{D}(S)$.

2.2.1 Parametric equations

Since $\mathrm{D}(S)$ is 2-dimensional vector space, any basis will contain two vectors. Let $(\mathbf{v}, \mathbf{w})$ be a basis of $\mathrm{D}(S)$. Then, for any two points $P, Q \in S$, the vector $\overrightarrow{QP}$ is a linear combination of the basis vectors, i.e. there exist unique scalars $s, t \in \mathbb{R}$ such that

$$\overrightarrow{QP} = s\mathbf{v} + t\mathbf{w}.$$

Now, if we fix Q and let P vary in the plane S then s and t vary in $\mathbb{R}$. Since $\phi_Q^3 : \mathbb{A}^3 \to \mathbb{V}^3$ is a bijection, the plane S can be described as

$$S = \left\{P \in \mathbb{A}^3 : \overrightarrow{QP} = s\mathbf{v} + t\mathbf{w} \text{ for some } s, t \in \mathbb{R}\right\}.$$

In this description, the point $Q \in S$ is arbitrary but fixed. If we want to emphasize that the description depends on fixing Q, we refer to this point as the *base point*. Moreover, for any point $O \in \mathbb{A}^3$ we may split $\overrightarrow{QP}$ in the equation $\overrightarrow{QP} = s\mathbf{v} + t\mathbf{w}$ to obtain

$$\overrightarrow{OP} = \overrightarrow{OQ} + s\mathbf{v} + t\mathbf{w}. \tag{2.6}$$

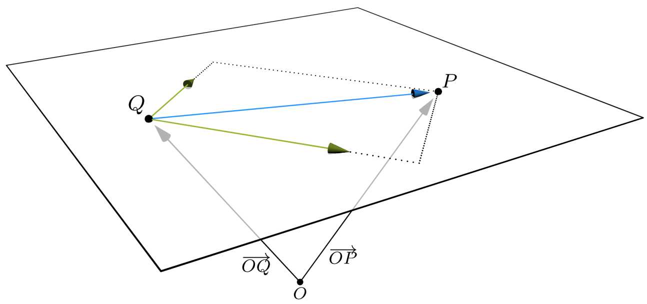


Figure 2.5: Vector equation of a plane.

So, we can describe the plane S as the set of points P in $\mathbb{A}^3$ which satisfy Equation (2.6) for some $s,t \in \mathbb{R}$. This equation is called the *vector equation of the plane S relative to O, having base point Q and direction vectors* $\mathbf{v}$ *and* $\mathbf{w}$, or simply *vector equation* of the plane S.

Notice that the vector equation depends on the choice of the base point Q and on the choice of the vectors $\mathbf{v}$ and $\mathbf{w}$. In analogy with the case of the line in $\mathbb{A}^2$ we may call such vectors *direction vectors* for the plane S. In particular, a plane does not have a unique vector equation. Notice also that the vector equation does not depend on the coordinate system. In the above description O can be any point in $\mathbb{A}^3$.

Now fix a coordinate system $\mathcal{K} = (O, \mathcal{B})$. If we write Equation (2.6) in coordinates relative to $\mathcal{K}$ then we obtain

$$S : \begin{cases} x = x_Q + sv_x + tw_x \\ y = y_Q + sv_y + tw_y \\ z = z_Q + sv_z + tw_z \end{cases} \quad \text{or, in matrix form} \quad S : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_Q \\ y_Q \\ z_Q \end{bmatrix} + s \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} + t \begin{bmatrix} w_x \\ w_y \\ w_z \end{bmatrix} \tag{2.7}$$

where $Q = A_{\mathcal{K}}(x_Q, y_Q, z_Q)$, $\mathbf{v} = \mathbf{v}_{\mathcal{K}}(v_x, v_y, v_z)$, $\mathbf{w} = \mathbf{v}_{\mathcal{K}}(w_x, w_y, w_z)$. The values s and t are called *parameters* and for different parameters we obtain different points (x, y, z) in the plane S. The three equations in the System (2.7) are called *parametric equations* for the plane S.

2.2.2 Cartesian equations

As in the case of the line in $\mathbb{A}^2$, it is possible to eliminate the parameters s,t in (2.7) to obtain

$$\left(\frac{v_x}{w_x} - \frac{v_z}{w_z}\right)\left(\frac{x - x_Q}{w_x} - \frac{y - y_Q}{w_y}\right) = \left(\frac{v_x}{w_x} - \frac{v_y}{w_y}\right)\left(\frac{x - x_Q}{w_x} - \frac{z - z_Q}{w_z}\right). \tag{2.8}$$

We will not give this equation a name, because it is a bit much to keep in mind, and one has to make sense of what happens when the denominators are zero. We simply notice that it is a linear equation in x, y and z and that it can be obtained by eliminating the parameters in (2.7).

There is an easier way of describing S with a linear equation. For this, you can interpret (2.7) as saying that the vector $\overrightarrow{QP}$ is linearly dependent on the vectors $\mathbf{v}$ and $\mathbf{w}$. With this in mind, the point $P(x,y,z)$ lies in the plane S if and only if

$$\begin{vmatrix} x-x_Q & y-y_Q & z-z_Q \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} = 0. \tag{2.9}$$

In particular, if $\mathbf{v} = \overrightarrow{QE}$, $\mathbf{w} = \overrightarrow{QF}$ and $\mathbf{w} = \overrightarrow{QG}$ for some points Q, E, F and G, then the four points are coplanar if and only if

$$\begin{vmatrix} x_E-x_Q & y_E-y_Q & z_E-z_Q \\ x_F-x_Q & y_F-y_Q & z_F-z_Q \\ x_G-x_Q & y_G-y_Q & z_G-z_Q \end{vmatrix} = 0. \tag{2.10}$$

Notice that Equation (2.10) is just a restatement of the fact that four points Q, E, F and G are coplanar if and only if the vectors $\overrightarrow{QE}$, $\overrightarrow{QF}$ and $\overrightarrow{QG}$ are linearly dependent (see Theorem 1.15). We have just described a plane with a linear equation (Equation (2.9)) relative to the coordinate system $\mathcal{K}$. The converse is also true.

Proposition 2.7. Every plane in $\mathbb{A}^3$ can be described with a linear equation in three variables

$$ax + by + cz + d = 0 \tag{2.11}$$

relative to a fixed coordinate system and any linear equation relative to a fixed coordinate system in three variables describes a plane if the constants a, b, c are not all zero. Moreover, if Equation (2.11) describes the plane π relative to a coordinate system $\mathcal{K} = (O, \mathcal{B})$, then the direction space $\mathrm{D}(\pi)$ of the plane is the 2-dimensional subspace of $\mathbb{V}^3$ which, relative to the basis $\mathcal{B}$, satisfies the equation

$$\mathrm{D}(\pi) : ax + by + cz = 0.$$

Equation (2.11) is called a *Cartesian equation* of the plane it describes. Notice that there are infinitely many Cartesian equations describing the same plane, since you can multiply one equation by a non-zero constant. In a fixed coordinate system equations of a plane are the same up to multiplication by a non-zero scalar. Here again it may be useful to rearrange the linear equation (2.11) in order to emphasize some geometric properties. For example, you can rearrange it in the form

$$\frac{x}{\alpha} + \frac{y}{\beta} + \frac{z}{\gamma} = 1 \quad \text{where} \quad \alpha = -\frac{d}{a} \quad , \beta = -\frac{d}{b}. \quad \text{and} \quad \gamma = -\frac{d}{c}.$$

In this form we have the *equation of the plane where we can read off the intersection points with the coordinate axes* since the plane intersects Ox in $(\alpha, 0, 0)$, it intersects Oy in $(0, \beta, 0)$ and it intersects Oz in $(0, 0, \gamma)$.

2.2.3 2-fold wedge product

If we expand the determinant in (2.9) on the first row, we see that the coefficients of x, y and z are respectively

$$a = \begin{vmatrix} v_y & v_z \\ w_y & w_z \end{vmatrix}, \quad b = -\begin{vmatrix} v_x & v_z \\ w_x & w_z \end{vmatrix} \quad \text{and} \quad c = \begin{vmatrix} v_x & v_y \\ w_x & w_y \end{vmatrix}. \tag{2.12}$$

We will see later that relative to an orthonormal coordinate system in the Euclidean space $\mathbb{E}^3$ the vector (a,b,c) is a normal vector for the plane S and that the wedge product (defined below) coincides with the cross product defined in Section 4.3.2.

Definition 2.8. Let $\mathbf{v}(v_x, v_y, v_z)$ and $\mathbf{w}(w_x, w_y, w_z)$ be two vectors in $\mathbb{V}^3$ with components relative to the basis $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. We define the *wedge product* of the vectors $\mathbf{v}$ and $\mathbf{w}$ to be the vector

$$\mathbf{v} \wedge_{\mathcal{B}} \mathbf{w} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix}. \tag{2.13}$$

In Section 2.4.2 we generalize this product and show that it yields the same vector for certain bases. Here we notice that the vectors $\mathbf{v}, \mathbf{w} \in \mathbb{V}^3$ are linearly dependent if and only if $\mathbf{v} \wedge_{\mathcal{B}} \mathbf{w} = 0$ for any basis $\mathcal{B}$. Moreover, Equation (2.11) of the plane π can be written in matrix form as

$$\pi : [\mathbf{v} \wedge_{\mathcal{B}} \mathbf{w}]_{\mathcal{B}} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = -d$$

where $\mathbf{v}$ and $\mathbf{w}$ form a basis of $\mathrm{D}(\pi)$ and a, b, c are as in (2.12).

2.2.4 Relative positions of two planes in $\mathbb{A}^3$

In order to describe intersections of planes we make use of linear algebra. Assume that we have two planes

$$\pi_1 : a_1x + b_1y + c_1z + d_1 = 0 \quad \text{and} \quad \pi_2 : a_2x + b_2y + c_2z + d_2 = 0.$$

We determine if they intersect or not by discussing the system:

$$\begin{cases} \pi_1 : a_1x + b_1y + c_1z + d_1 &= 0 \\ \pi_2 : a_2x + b_2y + c_2z + d_2 &= 0 \end{cases} \tag{2.14}$$

Discussing this system is basic linear algebra (see for example [4, Section 3.6]). Here again, the situation is very simple. Let M be the matrix of the system and $\widetilde{M}$ the extended matrix of the system. Then we have:

- two planes either intersect in a line, the coordinates of the points on the line will be solutions to (2.14), this happens if the rank of M and the rank of $\widetilde{M}$ equal 2; or
- they don't intersect and (2.14) doesn't have solutions, in which case the planes are parallel, this happens if the rank of M is strictly less than the rank of $\widetilde{M}$; or
- the solutions to System (2.14) depend on two parameters in which case $\pi_1 = \pi_2$, this happens if the rank of M and the rank of $\widetilde{M}$ are equal to 1.

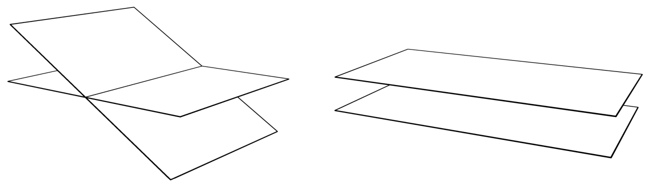


Figure 2.6: Incident and parallel planes.

2.2.5 Bundle of planes in $\mathbb{A}^3$

Definition 2.9. Let $\ell \subseteq \mathbb{A}^3$ be a line. The set Π_ℓ of all planes in $\mathbb{A}^3$ containing ℓ is called a *bundle of planes* and ℓ is called the *axis* (or *carrier line*) of the bundle Π_ℓ.

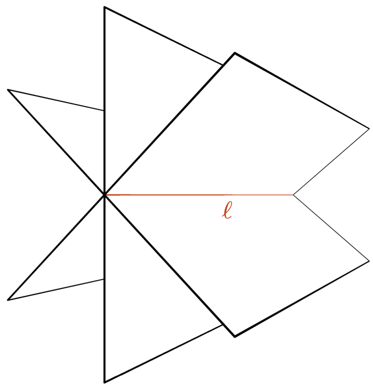


Figure 2.7: Planes in a bundle with axis (or carrier line) ℓ.

Proposition 2.10. If $\pi_1 : a_1x+b_1y+c_1z+d_1 = 0$ and $\pi_2 : a_2x+b_2y+c_2z+d_2 = 0$ are two distinct planes in the bundle Π_ℓ, then Π_ℓ consists of planes having equations of the form

$$\pi_{\lambda,\mu} : \lambda(a_1x + b_1y + c_1z + d_1) + \mu(a_2x + b_2y + c_2z + d_2) = 0.$$

where $\lambda, \mu \in \mathbb{R}$ are not both zero.

Bundles of planes are useful when a line ℓ is given as the intersection of two planes (see Subsection 2.3.2) and one wants to find the equation of a plane containing ℓ and satisfying some other conditions. For example, the condition that it contains some point P which does not belong to ℓ, or that it is parallel to a given line.

As in the case of line bundles, there is redundancy in the two parameters λ, μ. If $\lambda \neq 0$ then one can divide the equation of $\pi_{\lambda,\mu}$ by λ to obtain

$$\pi_{1,t} : (a_1x + b_1y + c_1z + d_1) + t(a_2x + b_2y + c_2z + d_2) = 0.$$

where $t = \frac{\mu}{\lambda} \in \mathbb{R}$. So $\pi_{1,\frac{\mu}{\lambda}}$ and $\pi_{\lambda,\mu}$ are in fact the same planes.

Definition 2.11. A *reduced bundle* is the set of all planes Π_ℓ with axis ℓ from which *we remove one plane*, i.e. it is $\Pi_\ell \setminus \{\pi_2\}$ for some $\pi_2 \in \Pi_\ell$. With the above notation and discussion, it is the set

$$\Big\{\pi_{1,t} : (a_1x + b_1y + c_1z + d_1) + t(a_2x + b_2y + c_2z + d_2) = 0 : t \in \mathbb{R}\Big\}.$$

The fact that we use one parameter instead of two, to describe almost all planes containing ℓ, simplifies calculations.

Definition 2.12. Let $\mathbb{W} \subseteq \mathbb{V}^3$ be a vector subspace of dimension 2. The set $\Pi_{\mathbb{W}}$ of all planes in $\mathbb{A}^3$ which admit $\mathbb{W}$ as direction space is called an *improper bundle of planes*, and $\mathbb{W}$ is called the vector subspace associated to the bundle $\Pi_{\mathbb{W}}$.

The connection between bundles of planes and improper bundles of planes is best understood through projective geometry, where we can think of the improper bundle of planes as the set of all planes intersecting in a line at infinity.

2.3 Lines in $\mathbb{A}^3$

Here again we treat the usual Euclidean space $\mathbb{E}^3$ as a 3-dimensional real affine space and denote it by $\mathbb{A}^3$. As in the case of $\mathbb{A}^2$, By Theorem 1.15, a line in $\mathbb{A}^3$ is a set of points S such that the set of vectors which can be represented by points in S form a 1-dimensional vector subspace of $\mathbb{V}^3$. Hence, the subset S is a line if for any $Q \in S$ we have that

$$\phi^3_Q(S) = \Big\{\overrightarrow{QP} : P \in S\Big\} \quad \text{is a 1-dimensional vector subspace of } \mathbb{V}^3.$$

If S is a line, we denote by $\mathrm{D}(S)$ the vector subspace $\phi^3_Q(S)$ of $\mathbb{V}^3$.

2.3.1 Parametric equations

If S is a line then for any two distinct points P, Q in S the vector $\overrightarrow{QP}$ is called a *direction vector* of S. Since $\phi^3_Q(S)$ is 1-dimensional, all direction vectors are linearly dependent and $\mathbf{v}$ is a direction vector for S if and only if it is linearly dependent on $\overrightarrow{QP}$. So, for any direction vector $\mathbf{v}$ of S there is a scalar $t \in \mathbb{R}$ such that

$$\overrightarrow{QP} = t\mathbf{v}.$$

Now, if you fix Q and let P vary on the line then t varies in $\mathbb{R}$. Since ϕ^3_Q is a bijection, the line S can be described as

$$S = \Big\{P \in \mathbb{A}^3 : \overrightarrow{QP} = t\mathbf{v} \text{ for some } t \in \mathbb{R}\Big\}.$$

In this description, the point Q is arbitrary but fixed. If we want to emphasize that this description depends on fixing Q, we refer to this point as the *base point*. Moreover, for any point $O \in \mathbb{A}^3$ we may split $\overrightarrow{QP}$ in the equation $\overrightarrow{QP} = t\mathbf{v}$ to obtain

$$\overrightarrow{OP} = \overrightarrow{OQ} + t\mathbf{v}. \tag{2.15}$$

The image that goes with this desrciption is the one in Figure 2.1. The only difference is that we interpret it in the 3-dimensional space $\mathbb{A}^3$. So, again, we can describe the line S as the set of points P in $\mathbb{A}^3$ which satisfy Equation (2.15) for some $t \in \mathbb{R}$. This equation is called the *vector equation of the line S relative to O, having base point Q and direction vector* $\mathbf{v}$, or simply *vector equation* of the line S.

So far, the description of a line in $\mathbb{A}^3$ is ad litteram the one used for $\mathbb{A}^2$. Now fix a coordinate system $\mathcal{K} = (O, \mathcal{B})$. If we write Equation (2.15) in coordinates relative to $\mathcal{K}$ then we obtain

$$\begin{cases} x = x_Q + tv_x \\ y = y_Q + tv_y \\ z = z_Q + tv_z \end{cases} \quad \text{or, in matrix form} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_Q \\ y_Q \\ z_Q \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \tag{2.16}$$

where $Q = Q(x_Q, y_Q, z_Q)$, $\mathbf{v} = \mathbf{v}(v_x, v_y, v_z)$ relative to $\mathcal{K}$ and where t is the *parameter* yielding different points (x, y, z) on the line. The three equations in the system (2.16) are called *parametric equations* for the line S.

2.3.2 Cartesian equations

It is possible to eliminate the parameter t in (2.16) in order to obtain

$$\frac{x - x_Q}{v_x} = \frac{y - y_Q}{v_y} = \frac{z - z_Q}{v_z}. \tag{2.17}$$

We refer to the Equations (2.17) as *symmetric equations* of the line S. It could happen that v_x, v_y or v_z are zero. In that case, translate back to the parametric equations to understand what happens.

We have just described a line with two linear equations (Equations (2.17)) relative to the coordinate system $\mathcal{K}$. The converse is also true.

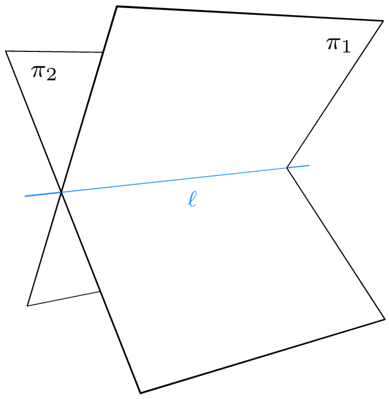
Proposition 2.13. Every line in $\mathbb{A}^3$ can be described with two linear equations in three variables

$$\begin{cases} a_1 x + b_1 y + c_1 z + d_1 &= 0 \\ a_2 x + b_2 y + c_2 z + d_2 &= 0 \end{cases} \tag{2.18}$$

relative to a fixed coordinate system and any compatible system of two linear equations of rank 2 in three variables relative to a fixed coordinate system describes a line. Moreover, if the Equations (2.19) describe the line ℓ relative to a coordinate system $\mathcal{K} = (O, \mathcal{B})$, then the direction space $\mathrm{D}(\ell)$ of the line is the 1-dimensional subspace of $\mathbb{V}^3$ which, relative to the basis $\mathcal{B}$, satisfies the equations

$$\mathrm{D}(\ell) : \begin{cases} a_1 x + b_1 y + c_1 z &= 0 \\ a_2 x + b_2 y + c_2 z &= 0 \end{cases}. \tag{2.19}$$

The Equations (2.19) are called *Cartesian equations* of the line which they describe. Notice that they describe a line as an intersection of two planes.



2.3.3 Relative positions of two lines in $\mathbb{A}^3$

Again, the intersections of lines can be determined with linear algebra. Assume we have two lines

$$\ell_1 : \begin{cases} a_1x + b_1y + c_1z + d_1 &= 0 \\ a_2x + b_2y + c_2z + d_2 &= 0 \end{cases} \quad \text{and} \quad \ell_2 : \begin{cases} a_3x + b_3y + c_3z + d_3 &= 0 \\ a_4x + b_4y + c_4z + d_4 &= 0 \end{cases}.$$

One way to determine if they intersect is to discuss the system:

$$\begin{cases} a_1x + b_1y + c_1z + d_1 &= 0 \\ a_2x + b_2y + c_2z + d_2 &= 0 \\ a_3x + b_3y + c_3z + d_3 &= 0 \\ a_4x + b_4y + c_4z + d_4 &= 0 \end{cases}. \tag{2.20}$$

Discussing this system is basic linear algebra (see for example [4, Section 3.6]). We leave this as an exercise, Excericse 2.41. It is somewhat easier to discuss the relative positions of lines in $\mathbb{A}^3$ via their parametric equations:

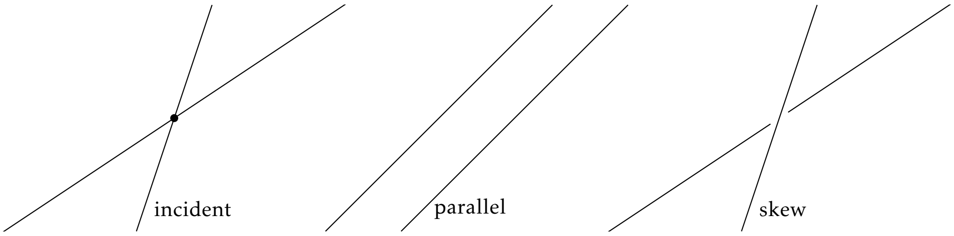
$$\ell_1 : \begin{cases} x = x_1 + tv_x \\ y = y_1 + tv_y \\ z = z_1 + tv_z \end{cases} \quad \text{und} \quad \ell_2 : \begin{cases} x = x_2 + tu_x \\ y = y_2 + tu_y \\ z = z_2 + tu_z \end{cases}.$$

We have the following cases:

- if the direction vectors $\mathbf{v}(v_1, v_2, v_3)$ and $\mathbf{u}(u_1, u_2, u_3)$ are proportional then the two lines are parallel;
- if they are parallel and have a point in common then the two lines are equal;
- if they are not parallel then they are coplanar (they lie in the same plane) if

$$\begin{vmatrix} x_1 - x_2 & y_1 - y_2 & z_1 - z_2 \\ v_x & v_y & v_z \\ u_x & u_y & u_z \end{vmatrix} = 0.$$

  in which case they intersect in exactly one point;
- if they are not parallel and they don't intersect, then we say that the two lines ℓ_1 and ℓ_2 are *skew* relative to each other.



2.3.4 Relative positions of a line and a plane in $\mathbb{A}^3$

Consider the plane

$$\pi : ax + by + cz + d = 0$$

and the line

$$\ell : \begin{cases} x = x_0 + tv_x \\ y = y_0 + tv_y \\ z = z_0 + tv_z \end{cases} .$$

In order to see if they intersect, we check to see which points in ℓ satisfy the equation of π:

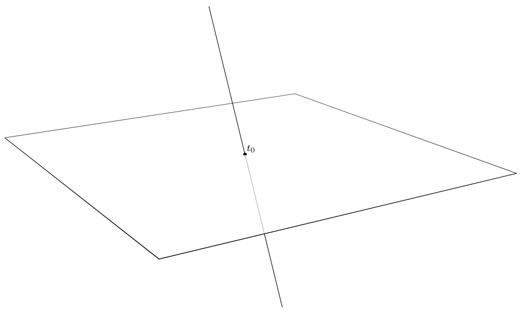
$$a(x_0 + tv_x) + b(y_0 + tv_y) + c(z_0 + tv_z) + d = 0 \quad \Leftrightarrow \quad (av_x + bv_y + cv_z)t + ax_0 + by_0 + cz_0 + d = 0. \tag{2.21}$$

The possibilities are:

- $av_x + bv_y + cv_z = 0$ and $ax_0 + by_0 + cz_0 + d \neq 0$ in which case Equation (2.21) has no solution, i.e. the plane and the line don't intersect, they are parallel; or
- $av_x + bv_y + cv_z = 0$ and $ax_0 + by_0 + cz_0 + d = 0$ in which case any $t \in \mathbb{R}$ is a solution to Equation (2.21), i.e. the line is contained in the plane, in particular they are parallel; or
- $av_x + bv_y + cv_z \neq 0$ in which case Equation (2.21) has the unique solution

$$t_0 = -\frac{ax_0 + by_0 + cz_0 + d}{av_x + bv_y + cv_z}.$$

Hence, the point corresponding to the parammeter t_0 is the intersection point $\ell \cap \pi$.



2.4 Affine subspaces of $\mathbb{A}^n$

Definition 2.14. A *d-dimensional affine subspace* of the affine space $\mathbb{A}^n$ is a subset $S \subseteq \mathbb{A}^n$ such that the set of vectors $\mathrm{D}(S)$ which can be represented by points in S form a d-dimensional vector subspace of $\mathbb{V}^n$. The vector subspace $\mathrm{D}(S)$ is then called the *direction space* of S. Moreover, given two affine subspaces S_1 and S_2 in $\mathbb{A}^n$ we say that S_1 is *parallel* to S_2, and we write $S_1 \| S_2$, if and only if $\mathrm{D}(S_1) \subseteq \mathrm{D}(S_2)$ or $\mathrm{D}(S_2) \subseteq \mathrm{D}(S_1)$. The dimension of an affine subspace S is denoted by $\dim(S)$, and $\dim(S) = \dim(\mathrm{D}(S))$.

Proposition 2.15. An affine subspace of $\mathbb{A}^n$ is an affine space with the affine structure inherited from $\mathbb{A}^n$.

Fixing a point $O \in \mathbb{A}^n$, a point $Q \in S$ and a basis $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_d)$ of $\mathrm{D}(S)$, it follows from the definition that S is a d-dimensional affine subspace if and only if

$$S = \left\{ P \in \mathbb{A}^n : \overrightarrow{OP} = \overrightarrow{OQ} + t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_d\mathbf{v}_d \text{ for some } t_1, \dots, t_d \in \mathbb{R} \right\}. \tag{2.22}$$

The equation in (2.22) is called the *vector equation of the affine subspace S relative to O, having base point Q and direction vectors* $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_d$, or simply *vector equation* of S.

Fixing a coordinate system with origin O and translating the equation in (2.22) in coordinates, one obtains *parametric equations* of the affine space S.

$$S : \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} + t_1 \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ \vdots \\ v_{1,n} \end{bmatrix} + \cdots + t_d \begin{bmatrix} v_{d,1} \\ v_{d,2} \\ \vdots \\ v_{d,n} \end{bmatrix}. \tag{2.23}$$

Another way of representing an affine subspace is by *Cartesian equations* (Equations (2.24)) as follows.

Theorem 2.16. Fix a coordinate system $\mathcal{K} = (O, \mathcal{B})$ in the affine space $\mathbb{A}^n$. Let

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{t1}x_1 + \cdots + a_{tn}x_n &= b_t \end{cases} \tag{2.24}$$

be a system of linear equations in the unknowns $x_1, \dots, x_n$. The set S of points of $\mathbb{A}^n$ whose coordinates are solutions to (2.24), if there are any, is an affine space of dimension $d = n - r$ where r is the rank of the matrix of coefficients of the system. The direction space $\mathrm{D}(S)$ is the vector subspace of $\mathbb{V}^n$ whose equations relative to $\mathcal{B}$ are given by the associated homogeneous system

$$\mathrm{D}(S) : \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{t1}x_1 + \cdots + a_{tn}x_n &= 0 \end{cases}.$$

Conversely, for every affine subspace S of $\mathbb{A}^n$ of dimension d there is a system of $n-d$ linear equations in n unknowns whose solutions correspond precisely to the coordinates of the points in S.

2.4.1 Hyperplanes

Definition 2.17. Affine subspaces in $\mathbb{A}^n$ which have dimension $n-1$ are called *hyperplanes*.

Let H be a hyperplane and let $(\mathbf{v}_1,\dots,\mathbf{v}_{n-1})$ be a basis of $\mathrm{D}(H)$ with respect to a coordinate system $\mathcal{K}=(O,\mathbf{e}_1,\dots,\mathbf{e}_n)$ of $\mathbb{A}^n$ parametric equations of H are of the from

$$H:\begin{cases} x_1 &= q_1+t_1v_{1,1}+\dots+t_{n-1}v_{n-1,1}\\ x_2 &= q_2+t_1v_{1,2}+\dots+t_{n-1}v_{n-1,2}\\ &\vdots\\ x_n &= q_n+t_1v_{1,n}+\dots+t_{n-1}v_{n-1,n}\end{cases} \quad\text{or}\quad H:\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}q_1\\q_2\\\vdots\\q_n\end{bmatrix}+t_1\begin{bmatrix}v_{1,1}\\v_{1,2}\\\vdots\\v_{1,n}\end{bmatrix}+\dots+t_{n-1}\begin{bmatrix}v_{n-1,1}\\v_{n-1,2}\\\vdots\\v_{n-1,n}\end{bmatrix} \tag{2.25}$$

where $\mathbf{v}_i=\mathbf{v}_i(v_{i,n},\dots,v_{i,n})$, $Q=Q(q_1,\dots,q_n)$ is a point in H and $t_i\in\mathbb{R}$ for each $i\in\{1,\dots,n-1\}$. These parametric equations express the fact that a point P belongs to H if and only if the vector $\overrightarrow{QP}$ is a linear combination of the basis vectors $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$, i.e. if and only if the vectors $\overrightarrow{QP},\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$ are linearly dependent. We can reformulate this as follows. A point $P(x_1,\dots,x_n)$ belongs to the hyperplane H if and only if

$$\begin{vmatrix} x_1-q_1 & x_2-q_2 & \dots & x_n-q_n\\ v_{1,1} & v_{1,2} & \dots & v_{1,n}\\ \vdots & \vdots & & \vdots\\ v_{n-1,1} & v_{n-1,2} & \dots & v_{n-1,n}\end{vmatrix}=0. \tag{2.26}$$

This is a *Cartesian equation* of the hyperplane H.

2.4.2 $(n-1)$-fold wedge product

Expanding the determinant in (2.26) we obtain

$$H: a_1x_1+a_2x_2+\dots+a_nx_n+b=0 \tag{2.27}$$

where the coefficient a_i of x_i corresponds to an $(n-1)$-minor in the determinant. We will see in Section 3.3.4 that relative to an orthonormal coordinate system in a Euclidean space $\mathbb{E}^n$ the vector $(a_1,\dots,a_n)$ is a normal vector for the hyperplane S and that the wedge product (defined below) generalizes the cross product defined for $\mathbb{E}^3$ (see Section 4.3.2).

Definition 2.18. Let $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$ be $n-1$ vectors in $\mathbb{V}^n$. Fix a basis $\mathcal{B}=(\mathbf{e}_1,\dots,\mathbf{e}_n)$ of $\mathbb{V}^n$. For each $i\in\{1,\dots,n-1\}$, let $(v_{i,n},\dots,v_{i,n})$ be the components of $\mathbf{v}_i$ with respect to $\mathcal{B}$. We define the *wedge product* of the vectors $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$ to be the vector

$$\mathbf{v}_1\wedge_{\mathcal{B}}\dots\wedge_{\mathcal{B}}\mathbf{v}_{n-1}=\begin{vmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n\\ v_{1,1} & v_{1,2} & \dots & v_{1,n}\\ \vdots & \vdots & & \vdots\\ v_{n-1,1} & v_{n-1,2} & \dots & v_{n-1,n}\end{vmatrix}. \tag{2.28}$$

Proposition 2.19. Let $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$ be $n-1$ vectors in $\mathbb{V}^n$. Consider two bases $\mathcal{B}$ and $\mathcal{B}'$ of $\mathbb{V}^n$. If $\mathrm{M}_{\mathcal{B},\mathcal{B}'}=\mathrm{M}^T_{\mathcal{B}',\mathcal{B}}$ then

$$\mathbf{v}_1\wedge_{\mathcal{B}}\dots\wedge_{\mathcal{B}}\mathbf{v}_{n-1}=\pm\mathbf{v}_1\wedge_{\mathcal{B}'}\dots\wedge_{\mathcal{B}'}\mathbf{v}_{n-1}.$$

With this notion, for a hyperplane H we can write Equation (2.27) in matrix form as

$$H : [\mathbf{v}_1 \wedge_{\mathcal{B}} \cdots \wedge_{\mathcal{B}} \mathbf{v}_{n-1}]_{\mathcal{B}}^T \cdot \left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix} \right) = 0$$

where $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ form a basis of $\mathrm{D}(H)$ and where $Q = Q_{\mathcal{K}}(q_1, \ldots, q_n)$ is a point in H.

2.4.3 Lines

A line in $\mathbb{A}^n$ is a 1-dimensional affine subspace. If ℓ is such a line, then, by definition, the vectors which can be represented by points in ℓ are linearly dependent. Any such non-zero vector $\mathbf{v}$ is called a direction vector of ℓ. Thus, ℓ can be described as

$$\ell = \left\{ P \in \mathbb{A}^n : \overrightarrow{OP} = \overrightarrow{OQ} + t\mathbf{v} \text{ for some } t \in \mathbb{R} \right\}.$$

for any point $O \in \mathbb{A}^n$ and any point $Q \in \ell$. The image that goes with this desrciption is the one in Figure 2.1, but here we interpret it in the n-dimensional space $\mathbb{A}^n$. In coordinates, relative to a given coordinate system $\mathcal{K}$ of $\mathbb{A}^n$, we obtain parametric equations for ℓ. They are of the form:

$$\ell : \begin{cases} x_1 &= q_1 + tv_1 \\ x_2 &= q_2 + tv_2 \\ &\vdots \\ x_n &= q_n + tv_n \end{cases} \quad \text{or, in matrix notation,} \quad \ell : \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} + t \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

where $Q = Q(q_1, \ldots, q_n)$ and $\mathbf{v} = \mathbf{v}(v_1, \ldots, v_n)$ relative to $\mathcal{K}$. Here again it is possible to eliminate the parameter t in order to obtain *symmetric equations* of the line ℓ:

$$\ell : \frac{x_1 - q_1}{v_1} = \frac{x_2 - q_2}{v_2} = \cdots = \frac{x_n - q_n}{v_n}.$$

These are in fact a system of $(n-1)$-linear equations which you can rearrange to look like this:

$$\ell : \begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1 \\ &\vdots \\ a_{n-1,1}x_1 + \cdots + a_{n-1,n}x_n &= b_{n-1} \end{cases}.$$

Notice that the rank of this system is $n-1$ since $\dim(\ell) = 1$. Moreover, notice that each linear equation in the above system desribes a hyperplane. So, this is saying that a line in $\mathbb{A}^n$ can be described by the intersection of $n-1$ hyperplanes.

2.4.4 Relative positions

Let S and T be two affine subspaces of $\mathbb{A}^n$. If $S \parallel T$ (see Definition 2.14) then they are disjoint or one is included in the other (see Proposition 2.20 below). Notice that if $\dim(S) = \dim(T)$, then S and T are parallel if and only if $\mathrm{D}(S) = \mathrm{D}(T)$. In particular, if S and T are lines, they are parallel if they have the same direction, i.e. any two of their direction vectors are proportional. Notice also that two hyperplanes are parallel if the coefficients of the unknowns in their equations are proportional.

Proposition 2.20. Let S and T be parallel affine subspaces of $\mathbb{A}^n$ with $\dim(S) \leq \dim(T)$.

1.) If S and T have a point in common then $S \subseteq T$.

2.) If $\dim(S) = \dim(T)$, and S and T have a point in common then $S = T$.

As a consequence of Proposition 2.20 we obtain the following corollary which implies the 'parallel postulate' of Euclidean geometry. The axioms of affine spaces therefore imply the validity of this postulate.

Corollary 2.21. If S is an affine subspace of $\mathbb{A}^n$ and $P \in \mathbb{A}^n$, there is a unique affine subspace T of $\mathbb{A}^n$ which contains P, is parallel to S and has the same dimension as S.

Definition 2.22. If two affine subspace S and T of $\mathbb{A}^n$ are not parallel, they are said to be either *skew* if they do not meet, or *incident* if they have a point in common.

In order to determine the intersection $S \cap T$, suppose that the two subspace are given by the Cartesian equations

$$S : \sum_{j=1}^{n} a_{ij} x_j = b_i \quad \text{for} \quad i = 1, \ldots, n-s \tag{2.29}$$

$$T : \sum_{j=1}^{n} c_{kj} x_j = d_k \quad \text{for} \quad k = 1, \ldots, n-t. \tag{2.30}$$

The intersection $S \cap T$ is the locus of points in $\mathbb{A}^n$ whose coordinates are simultaneously solutions to both (2.29) and (2.30), i.e. they are solutions to the system

$$S \cap T : \begin{cases} \sum_{j=1}^{n} a_{ij} x_j = b_i & \text{for} \quad i = 1, \ldots, n-s, \\ \sum_{j=1}^{n} c_{kj} x_j = d_k & \text{for} \quad k = 1, \ldots, n-t. \end{cases} \tag{2.31}$$

By Theorem 2.16, if the System (2.31) has a solution, then it describes an affine subspace. Thus, if $S \cap T$ is non-empty it is an affine subspace of $\mathbb{A}^n$.

Proposition 2.23. If the intersection $S \cap T$ of two affine subspaces of $\mathbb{A}^n$ is non-empty it is an affine subspace satisfying

$$\dim(S) + \dim(T) - \dim(\mathbb{A}^n) \leq \dim(S \cap T) \leq \min\Big\{\dim(S), \dim(T)\Big\}. \tag{2.32}$$

Notice that the second inequality is an equality if $S \subseteq T$ or $T \subseteq S$. What about the first inequality? when do we have equality there?

Proposition 2.24. Let S and T be two affine subspaces of $\mathbb{A}^n$. Then $\mathbb{V}^n = \mathrm{D}(S) + \mathrm{D}(T)$ if and only if $S \cap T \neq \emptyset$ *and*

$$\dim(S \cap T) = \dim(S) + \dim(T) - \dim(\mathbb{A}^n). \tag{2.33}$$

2.4.5 Changing the reference frame

Let S be an affine subspace of $\mathbb{A}^n$ given with respect to the reference frame $\mathcal{K} = (O, \mathcal{B})$ via the parametric equations (2.23). Then, if $\mathcal{K}' = (O', \mathcal{B}')$ is another coordinate system, by Theorem 1.22, parametric equations with respect to $\mathcal{K}'$ are

$$S : \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix} = \mathrm{M}_{\mathcal{K}',\mathcal{K}} \cdot \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} + [O]_{\mathcal{K}'} + t_1 \cdot \mathrm{M}_{\mathcal{K}',\mathcal{K}} \cdot \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ \vdots \\ v_{1,n} \end{bmatrix} + \cdots + t_d \cdot \mathrm{M}_{\mathcal{K}',\mathcal{K}} \cdot \begin{bmatrix} v_{d,1} \\ v_{d,2} \\ \vdots \\ v_{d,n} \end{bmatrix}.$$

In terms of Cartesian equations, notice that (2.24) can be written in the form

$$S : A \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = b.$$

Then, with respect to $\mathcal{K}'$, this system translates as follows

$$S : \underbrace{\left(A \cdot \mathrm{M}_{\mathcal{K},\mathcal{K}'}\right)}_{=A'} \cdot \begin{bmatrix} x_1' \\ \vdots \\ x_n' \end{bmatrix} = \underbrace{b - A \cdot [O]_{\mathcal{K}'}}_{=b'}.$$

2.5 Exercises

2.1. Determine parametric equations for the line $\ell \subseteq \mathbb{A}^2$ in the following cases:

a) ℓ contains the point $A(1,2)$ and is parallel to the vector $\mathbf{a}(3,-1)$,

b) ℓ contains the origin and is parallel to $\mathbf{b}(4,5)$,

c) ℓ contains the point $M(1,7)$ and is parallel to Oy,

d) ℓ contains the points $M(2,4)$ and $N(2,-5)$.

2.2. For the lines ℓ in the previous exercise

a) determine a Cartesian equation for ℓ,

b) describe all direction vectors for ℓ.

2.3. With the assumptions in Exercise 1.16, give parametric equations and Cartesian equations for the lines AB, AC, BC both in the coordinate system $\mathcal{K}$ and in the coordinate system $\mathcal{K}'$.

2.4. Find a Cartesian equation of the line ℓ in $\mathbb{A}^2$ containing the points $P = S \cap S'$ and $Q = T \cap T'$ where

$$S : x + 5y - 8 = 0, \quad S' : 3x + 6 = 0, \quad T : 5x - \frac{1}{2}y = 1, \quad T' : x - y = 5.$$

2.5. Deterimine an equation for the line in $\mathbb{A}^2$ parallel to $\mathbf{v}$ and passing through $S \cap T$ in each of the following cases:

1. $\mathbf{v} = (2,4)$, $S : 3x - 2y - 7 = 0$, $T : 2x + 3y = 0$,
2. $\mathbf{v} = (-5\sqrt{2}, 7)$, $S : x - y = 0$, $T : x + y = 1$.

2.6. Let ABC be a triangle in the affine space $\mathbb{A}^n$. Consider the points C' and B' on the sides AB and AC respectively, such that

$$\overrightarrow{AC'} = \lambda \overrightarrow{BC'} \quad \text{and} \quad \overrightarrow{AB'} = \mu \overrightarrow{CB'}.$$

The lines BB' and CC' meet in the point M. For a fixed but arbitrary point $O \in \mathbb{A}^n$, show that

$$\overrightarrow{OM} = \frac{\overrightarrow{OA} - \lambda \overrightarrow{OB} - \mu \overrightarrow{OC}}{1 - \lambda - \mu}.$$

Deduce a formula for $\overrightarrow{OG}$ where G is the centroid of the triangle.

2.7. In $\mathbb{A}^n$, consider the angle BOB' and the points $A \in [OB]$, $A' \in [OB']$. Show that

$$\overrightarrow{OM} = m\frac{1-n}{1-mn}\overrightarrow{OA} + n\frac{1-m}{1-mn}\overrightarrow{OA'}$$

$$\overrightarrow{ON} = m\frac{n-1}{n-m}\overrightarrow{OA} + n\frac{m-1}{m-n}\overrightarrow{OA'}$$

where $M = AB' \cap A'B$ and $N = AA' \cap BB'$ and where $\overrightarrow{OB} = m\overrightarrow{OA}$ and $\overrightarrow{OB'} = n\overrightarrow{OA'}$.

2.8. Show that the midpoints of the diagonals of a complete quadrilateral are collinear.

2.9. Determine parametric equations for the plane π in the following cases:

a) π contains the point $M(1,0,2)$ and is parallel to the vectors $\mathbf{a}_1(3,-1,1)$ and $\mathbf{a}_2(0,3,1)$,

b) π contains the points $A(-2,1,1)$, $B(0,2,3)$ and $C(1,0,-1)$,

c) π contains the point $A(1,2,1)$ and is parallel to $\mathbf{i}$ and $\mathbf{j}$,

d) π contains the point $M(1,7,1)$ and is parallel coordinate plane Oyz,

e) π contains the points $M_1(5,3,4)$ and $M_2(1,0,1)$, and is parallel to the vector $\mathbf{a}(1,3,-3)$,

f) π contains the point $A(1,5,7)$ and the coordinate axis Ox.

2.10. Determine Cartesian equations for the plane π in the following cases:

a) $\pi : x = 2 + 3u - 4v,\ y = 4 - v,\ z = 2 + 3u;$

b) $\pi : x = u + v,\ y = u - v,\ z = 5 + 6u - 4v.$

2.11. Determine parametric equations for the plane π in the following cases:

a) $3x - 6y + z = 0;$

b) $2x - y - z - 3 = 0.$

2.12. With the assumptions in Exercise 1.18, give parametric equations and Cartesian equations for the line AB and the plane ACD both in the coordinate system $\mathcal{K}$ and in the coordinate system $\mathcal{K}'$.

2.13. Show that the points $A(1,0,-1)$, $B(0,2,3)$, $C(-2,1,1)$ and $D(4,2,3)$ are coplanar.

2.14. Determine the relative positions of the planes in the following cases

a) $\pi_1 : x + 2y + 3z - 1 = 0,\ \pi_2 : x + 2y - 3z - 1 = 0.$

b) $\pi_1 : x + 2y + 3z - 1 = 0,\ \pi_2 : 2x + y + 3z - 2 = 0,\ \pi_3 : x + 2y + 3z + 2 = 0.$

2.15. Show that the planes

$$\pi_1 : 3x + y + z - 1 = 0, \quad \pi_2 : 2x + y + 3z + 2 = 0, \quad \pi_3 : -x + 2y + z + 4 = 0$$

have a point in common.

2.16. Show that the pairwise intersection of the planes

$$\pi_1 : 3x + y + z - 5 = 0, \quad \pi_2 : 2x + y + 3z + 2 = 0, \quad \pi_3 : 5x + 2y + 4z + 1 = 0$$

are parallel lines.

2.17. Determine parametric equations for the line ℓ in the following cases:

a) ℓ contains the point $M_0(2,0,3)$ and is parallel to the vector $\mathbf{a}(3,-2,-2)$,

b) ℓ contains the point $A(1,2,3)$ and is parallel to the Oz-axis,

c) ℓ contains the points $M_1(1,2,3)$ and $M_2(4,4,4)$.

2.18. Give Cartesian equations for the lines ℓ in the previous exercise.

2.19. Determine parametric equations for the line contained in the planes $x+y+2z-3=0$ and $x-y+z-1=0$.

2.20. Consider the lines $\ell_1 : x = 1+t, y = 1+2t, z = 3+t, t \in \mathbb{R}$ and $\ell_2 : x = 3+s, y = 2s, z = -2+s, s \in \mathbb{R}$. Show that ℓ_1 and ℓ_2 are parallel and find the equation of the plane determined by the two lines.

2.21. Determine parametric equations of the line passing through $P(5,0,-2)$ and parallel to the planes $\pi_1 : x-4y+2z=0$ and $\pi_2 : 2x+3y-z+1=0$.

2.22. Determine an equation of the plane containing $P(2,0,3)$ and the line $\ell : x = -1+t, y = t, z = -4+2t, t \in \mathbb{R}$.

2.23. For the points $A(2,1,-1)$ and $B(-3,0,2)$, determine an equation of the bundle of planes passing through A and B.

2.24. Determine the relative positions of the lines $x = -3t, y = 2+3t, z = 1, t \in \mathbb{R}$ and $x = 1+5s, y = 1+13s, z = 1+10s, s \in \mathbb{R}$.

2.25. Determine the parameter m for which the line $x = -1+3t, y = 2+mt, z = -3-2t$ doesn't intersect the plane $x+3y+3z-2=0$.

2.26. Determine the values a and d for which the line $\frac{x-2}{3} = \frac{y+1}{2} = \frac{z-3}{-2}$ is contained in the plane $ax+y-2z+d=0$.

2.27. In each of the following, find a Cartesian equation of the plane in $\mathbb{A}^3$ passing through Q and parallel to the lines ℓ and ℓ':

a) $Q(1,-1,-2)$, $\ell : x-y=1, x+z=5$, $\ell' : x=1, z=2$

b) $Q(0,1,3)$, $\ell : x+y=-5, x-y+2z=0$, $\ell : 2x-2y=1, x-y+2z=1$

2.28. In each of the following, find the relative positions of the line ℓ and the plane π in $\mathbb{A}^3$, and, if they are incident, determine the point of intersection.

a) $\ell : x = 1+t, y = 2-2t, z = 1-4t$, $\pi : 2x-y+z-1=0$

b) $\ell : x = 2-t, y = 1+2t, z = -1+3t$, $\pi : 2x+2y-z+1=0$

2.29. In each of the following, find a Cartesian equation for the plane in $\mathbb{A}^3$ containing the point Q and the line ℓ.

a) $Q = (3,3,1)$, $\ell : x = 2+3t, y = 5+t, z = 1+7t$

b) $Q = (2,1,0)$, $\ell : x - y + 1 = 0, 3x + 5z - 7 = 0$

2.30. In each of the following, find Cartesian equations for the line ℓ in $\mathbb{A}^3$ passing through Q, contained in the plane π and intersecting the line ℓ'

a) $Q = (1,1,0)$, $\pi : 2x - y + z - 1 = 0$, $\ell' : x = 2 - t, y = 2 + t, z = t$

b) $Q = (-1,-1,-1)$, $\pi : x + y + z + 3 = 0$, $\ell' : x - 2z + 4 = 0, 2y - z = 0$

2.31. In each of the following, find Cartesian equations for the line ℓ in $\mathbb{A}^3$ passing through Q and coplanar to the lines ℓ' and ℓ''. Furthermore, establish whether ℓ meets or is parallel to ℓ' and ℓ''

a) $Q = (1,1,2)$, $\ell' : 3x - 5y + z = -1, 2x - 3z = -9$, $\ell'' : x + 5y = 3, 2x + 2y - 7z = -7$

b) $Q = (2,0,-2)$, $\ell' : -x + 3y = 2, x + y + z = -1$, $\ell'' : x = 2 - t, y = 3 + 5t, z = -t$

2.32. In each of the following, find the value of the real parameter k for which the lines ℓ and ℓ' are coplanar. Find a Cartesian equation for the plane that contains them, and find the point of intersection whenever they meet

a) $\ell : x = k + t, y = 1 + 2t, z = -1 + kt$, $\ell' : x = 2 - 2t, y = 3 + 3t, z = 1 - t$

b) $\ell : x = 3 - t, y = 1 + 2t, z = k + t$, $\ell' : x = 1 + t, y = 1 + 2t, z = 1 + 3t$

2.33. Find a Cartesian equation for the plane π in $\mathbb{A}^3$ which contains the line of intersection of the two planes

$$x + y = 3 \quad \text{and} \quad 2y + 3z = 4$$

and is parallel to the vector $\mathbf{v} = (3,-1,2)$.

2.34. In the affine space $\mathbb{A}^4$ consider

$$\text{the plane} \quad \alpha = \left\langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\\1\end{bmatrix} \right\rangle + \begin{bmatrix}2\\4\\1\\2\end{bmatrix} \quad \text{and the line} \quad \beta = \left\langle \begin{bmatrix}1\\1\\-1\\1\end{bmatrix} \right\rangle + \begin{bmatrix}2\\3\\-1\\1\end{bmatrix}.$$

Determine $\alpha \cap \beta$.

2.35. In $\mathbb{A}^4$ consider the affine subspaces

$$\alpha = \begin{bmatrix}2\\1\\2\\1\end{bmatrix} \quad \beta = \left\langle \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \right\rangle + \begin{bmatrix}1\\3\\0\\0\end{bmatrix} \quad \gamma = \left\langle \begin{bmatrix}2\\1\\3\\-1\end{bmatrix}, \begin{bmatrix}1\\0\\2\\-2\end{bmatrix} \right\rangle + \begin{bmatrix}1\\0\\1\\0\end{bmatrix} \quad \delta = \left\langle \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \right\rangle.$$

Which of the following is true?

a) $\alpha \in \beta$

b) $\alpha \in \gamma$

c) $\alpha \in \delta$

d) $\beta \parallel \gamma$

e) $\beta \parallel \delta$

f) $\gamma \parallel \delta$

g) $\beta \subseteq \gamma$

h) $\gamma \subseteq \delta$

i) $\beta \subseteq \delta$

2.36. Consider the following affine subspaces of $\mathbb{A}^4$

$$Y : \begin{cases} x_1 + x_3 - 2 &= 0 \\ 2x_1 - x_2 + x_3 + 3x_4 - 1 &= 0 \end{cases}$$

$$Z : \begin{cases} x_1 + x_2 + 2x_3 - 3x_4 &= 1 \\ x_2 + x_3 - 3x_4 &= -1 \\ x_1 - x_2 + 3x_4 &= 3 \end{cases}$$

a) Determine the dimensions of Y and Z.

b) Find parametric equations for each of the two affine subspaces.

c) Is $Y \parallel Z$?

2.37. In Section 2.2.2 we deduce a linear equation for a plane in $\mathbb{A}^3$ via a determinant. What is the analogue of this description for lines? I.e. deduce Cartesian equations for lines starting from linear dependence of vectors (both in $\mathbb{A}^2$ and $\mathbb{A}^3$).

2.38. Consider the affine space $\mathbb{A}^3$. Show that if a line ℓ doesn't intersect a plane π then $\ell \parallel \pi$ in the sense of the Definition 2.14. Moreover, give an example in $\mathbb{A}^4$ of a line and a plane which do not intersect and which are not parallel.

2.39. Consider the affine space $\mathbb{A}^4$. Describe the relative positions of two planes.

2.40. In $\mathbb{A}^3$ discuss the relative positions of a plane and a line in terms of their Cartesian equations.

2.41. In $\mathbb{A}^3$ discuss the relative positions two lines in terms of their Cartesian equations.

CHAPTER 3

Euclidean space

Contents

3.1 Scalar product in $\mathbb{E}^2$

Let $\mathbf{a}, \mathbf{b} \in \mathbb{V}^2$ be two non-zero vectors. Given a point $O \in \mathbb{E}^2$, there are unique points $A, B \in \mathbb{E}^2$ such that $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$. Thus we obtain an angle $\measuredangle AOB$. For a different point $O' \in \mathbb{E}^2$, again there are unique points $A', B' \in \mathbb{E}^2$ such that $\mathbf{a} = \overrightarrow{O'A'}$ and $\mathbf{b} = \overrightarrow{O'B'}$ giving us a second angle $\measuredangle A'O'B'$. It is not difficult to see that the two angles are congruent. Moreover, centering a circle of radius 1 in the point O, the angle $\measuredangle AOB$ divides the circle into two arcs. Recall that the measure of the angle $\measuredangle AOB$ is the length of the shortest of these two arcs. Noticing that the measure of such angles does not depend on the choice of O we define the *angle between* $\mathbf{a}$ *and* $\mathbf{b}$ to be the measure of the angle $\measuredangle AOB$ and we denote it by $\measuredangle(\mathbf{a}, \mathbf{b})$.

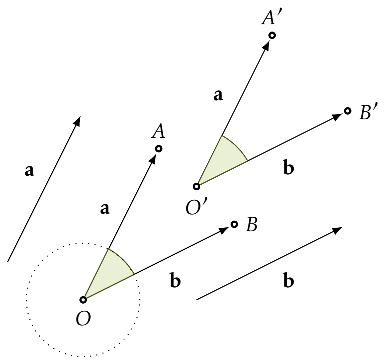


Figure 3.1: Angle of two vectors.

Definition 3.1. The *scalar product* (or, *dot product*) of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{V}^2$ is the real number

$$\langle \mathbf{a}, \mathbf{b} \rangle = \begin{cases} 0, & \text{if one of the two vector is zero;} \\ |\mathbf{a}| \cdot |\mathbf{b}| \cdot \cos\measuredangle(\mathbf{a}, \mathbf{b}) & \text{if both vectors are non-zero.} \end{cases}$$

Now, if we know how to calculate scalar products, we can calculate the distance between any two points. Indeed, the distance $d(A, B)$ from A to B is the length $|\overrightarrow{AB}|$ of the vector $\overrightarrow{AB}$, so

$$d(A, B) = |\overrightarrow{AB}| = \sqrt{\langle \overrightarrow{AB}, \overrightarrow{AB} \rangle}.$$

Moreover, if we know how to calculate scalar products, we can calculate angles since

$$\cos(\measuredangle(\mathbf{a}, \mathbf{b})) = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{|\mathbf{a}| \cdot |\mathbf{b}|}.$$

Notice that, in fact, determining $\measuredangle(\mathbf{a}, \mathbf{b})$ reduces to calculating distances, the length of a certain arc on a unit circle. So, the scalar product determines the notion of ‘length’ or ‘distance’ and we aim at giving a definition of the scalar product without relying on the notion of ‘distance’. We will do this by extracting defining properties of this product derived from Definition 3.1.

For this, fix again representatives $\overrightarrow{OA}$ and $\overrightarrow{OB}$ of $\mathbf{a}$ and $\mathbf{b}$ respectively and notice that $|\mathbf{b}| \cdot \cos\measuredangle(\mathbf{a},\mathbf{b})$ is the length b' of the orthogonal projection of the segment $[OB]$ on the line OA. Thus the length b' of this projection is $\langle \mathbf{a},\mathbf{b}\rangle/|\mathbf{a}|$. If we now construct a vector proportional to $\mathbf{a}$ and of length b' we obtain the *orthogonal projection of* $\mathbf{b}$ *on* $\mathbf{a}$:

$$\frac{\langle \mathbf{a},\mathbf{b}\rangle}{\langle \mathbf{a},\mathbf{a}\rangle}\mathbf{a}. \tag{3.1}$$

Clearly, we may interchange $\mathbf{a}$ and $\mathbf{b}$ to obtain the *orthogonal projection of* $\mathbf{a}$ *on* $\mathbf{b}$ which is $\frac{\langle \mathbf{b},\mathbf{a}\rangle}{\langle \mathbf{b},\mathbf{b}\rangle}\mathbf{b}$.

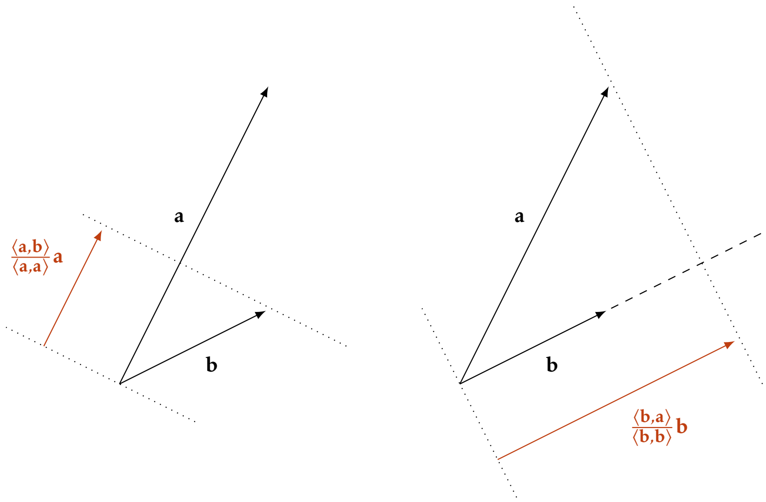


Figure 3.2: Orthogonal projection of vectors on each other.

Proposition 3.2. The scalar product $\langle \_,\_ \rangle : \mathbb{V}^2 \times \mathbb{V}^2 \to \mathbb{R}$ satisfies the following properties.

(SP1) It is *bilinear*, i.e. for all $a, b \in \mathbb{R}$ and all $\mathbf{v},\mathbf{w},\mathbf{u} \in \mathbb{V}^2$ we have

$$\langle a\mathbf{v}+b\mathbf{w},\mathbf{u}\rangle = a\langle \mathbf{v},\mathbf{u}\rangle + b\langle \mathbf{w},\mathbf{u}\rangle \quad \text{and} \quad \langle \mathbf{v}, a\mathbf{w}+b\mathbf{u}\rangle = a\langle \mathbf{v},\mathbf{w}\rangle + b\langle \mathbf{v},\mathbf{u}\rangle.$$

(SP2) It is *symmetric*, i.e. for all $\mathbf{v},\mathbf{w} \in \mathbb{V}^2$ we have

$$\langle \mathbf{v},\mathbf{w}\rangle = \langle \mathbf{w},\mathbf{v}\rangle.$$

(SP3) It is *positive definite*, i.e. for all $\mathbf{v} \in \mathbb{V}^2$

$$\text{if } \mathbf{v} \neq 0 \quad \text{then} \quad \langle \mathbf{v},\mathbf{v}\rangle > 0.$$

The bilinearity of the scalar product implies that it is enough to know its values on a basis of $\mathbb{V}^2$. Indeed, if $\mathcal{B} = (\mathbf{i},\mathbf{j})$ is a basis of $\mathbb{V}^2$, then for any vectors $\mathbf{v} = \mathbf{v}(v_1,v_2) = v_1\mathbf{i} + v_2\mathbf{j}$ and $\mathbf{w} = \mathbf{w}(w_1,w_2) = w_1\mathbf{i} + w_2\mathbf{j}$ we have

$$\langle \mathbf{v},\mathbf{w}\rangle = \langle v_1\mathbf{i}+v_2\mathbf{j}, w_1\mathbf{i}+w_2\mathbf{j}\rangle = v_1w_1\langle \mathbf{i},\mathbf{i}\rangle + v_1w_2\langle \mathbf{i},\mathbf{j}\rangle + v_2w_1\langle \mathbf{j},\mathbf{i}\rangle + v_2w_2\langle \mathbf{j},\mathbf{j}\rangle. \tag{3.2}$$

We can rearrange this expression in matrix form by considering the *Gram matrix* of the basis $\mathcal{B}$, defined as

$$\mathrm{G}_{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{i},\mathbf{i}\rangle & \langle \mathbf{i},\mathbf{j}\rangle \\ \langle \mathbf{j},\mathbf{i}\rangle & \langle \mathbf{j},\mathbf{j}\rangle \end{bmatrix}.$$

By (3.2), we then have

$$\langle \mathbf{v},\mathbf{w}\rangle = [\mathbf{v}]_{\mathcal{B}}^T \cdot \mathrm{G}_{\mathcal{B}} \cdot [\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \cdot \begin{bmatrix} \langle \mathbf{i},\mathbf{i}\rangle & \langle \mathbf{i},\mathbf{j}\rangle \\ \langle \mathbf{j},\mathbf{i}\rangle & \langle \mathbf{j},\mathbf{j}\rangle \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

Furthermore, if we choose the vectors $\mathbf{i}$ and $\mathbf{j}$ to be orthogonal and of length 1 then $\mathrm{G}_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and thus

$$\langle \mathbf{v},\mathbf{w}\rangle = v_1 w_1 + v_2 w_2. \tag{3.3}$$

3.2 The n-dimensional Euclidean space

In this section we define the n-dimensional Euclidean space to be the affine space $\mathbb{A}^n$ together with a scalar product on the set of vectors $\mathbb{V}^n$. The main task is to define the scalar product in such a way that it coincides with the one we are familiar with in dimensions 2 and 3.

3.2.1 Scalar product

In order to define a scalar product in higher dimensions, i.e. on $\mathbb{V}^n$, consider the properties extracted in Proposition 4.10 and the way in which we deduced (3.3). By (SP1), we want to have a bilinear map

$$\langle \_,\_ \rangle : \mathbb{V}^n \times \mathbb{V}^n \to \mathbb{R}.$$

With respect to a basis $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ of $\mathbb{V}^n$, such a map is of the form

$$\langle \mathbf{v}\cdot\mathbf{w}\rangle = [\mathbf{v}]_{\mathcal{B}}^T \cdot \mathrm{G}_{\mathcal{B}} \cdot [\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} \cdot \mathrm{G}_{\mathcal{B}} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix} \tag{3.4}$$

where $\mathrm{G}_{\mathcal{B}} = (g_{ij})$ is the *Gram matrix* defined by $g_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j\rangle$. Thus, one way of defining a scalar product in higher dimensions is by selecting a basis $\mathcal{B}$ and prescribing the values of the bilinear map on the basis vectors. For example, we may choose $\mathrm{G}_{\mathcal{B}}$ to be the identity matrix.

Definition 3.3. Fix a basis $\mathcal{B}$ of $\mathbb{V}^n$. Let $\mathbf{v}(v_1, \dots, v_n)$ and $\mathbf{w}(w_1, \dots, w_n)$ be two vectors with components given relative to $\mathcal{B}$. The *scalar product* of $\mathbf{v}$ and $\mathbf{w}$ is

$$\langle \mathbf{v},\mathbf{w}\rangle = v_1 w_1 + v_2 w_2 + \dots + v_n w_n. \tag{3.5}$$

Indeed, this is the prefered way of introducing the scalar product in higher dimensions since it is easy to see that it satisfies the properties in Proposition 4.10 and it coincides with (3.3) if $n = 2$. Thus Definition 3.3 implies Definition 3.1.

However, when taking (3.5) as definition, it may look like the notion of scalar product, and therefore the notion of distance, depends on the choice of the basis $\mathcal{B}$ of $\mathbb{V}^n$. Another way of defining a scalar product is to make use of all the properties in Proposition 4.10. It turns out that these properties are enough to determine the scalar product without making use of any notion of 'distance' and without fixing a basis.

Definition 3.4. A *scalar product* on $\mathbb{V}^n$ is a positive definite, symmetric, bilinear map

$$\langle \_, \_ \rangle : \mathbb{V}^n \times \mathbb{V}^n \to \mathbb{R}.$$

The following theorem shows that Definition 3.3 implies Definition 3.4. A proof of this theorem follows from Sylvester's Theorem (see Appendix A for more details).

Theorem 3.5. Given a scalar product as in Definition 3.4, there exists a basis $\mathcal{B}$ such that the Gram matrix $G_{\mathcal{B}}$ is the identity matrix. Then, with respect to the basis $\mathcal{B}$, the scalar product has the form (3.5).

By now, we have seen three definitions of scalar product: Definition 3.1, Definition 3.3 and Definition 3.4. Definition 3.1 is for the 2-dimensional setting and uses the notion of 'distance' whereas the other two definitions are free of this notion and can be *used* to define 'distance'. Moreover, Definition 3.4 is independent of any choice of a basis, it implies Definition 3.3 and it coincides with Definition 3.1 in the 2-dimensional case. In particular, from a computational perspective, there is no distinction between the last two definitions.

Definition 3.6. The n-dimensional Euclidean space $\mathbb{E}^n$ is the affine space $\mathbb{A}^n$ together with a scalar product $\langle \_, \_ \rangle$ on $\mathbb{V}^n$. The distance between two points $P, Q \in \mathbb{E}^n$ is defined by

$$d(P,Q) = |\overrightarrow{QP}| = \sqrt{\langle \overrightarrow{QP}, \overrightarrow{QP} \rangle}.$$

There is yet another advantage of the more conceptual Definition 3.4. By Proposition 2.15, a d-dimensional affine subspace S of $\mathbb{A}^n$ is itself an affine space, which can be identified with $\mathbb{A}^d$; we write $S \cong \mathbb{A}^d$. It is easy to see that a scalar product on $\mathbb{A}^n$ defined with Definition 3.4 restricts to a scalar product on $S \cong \mathbb{A}^d \subseteq \mathbb{A}^n$. Thus, an inclusion $S \cong \mathbb{A}^d \subseteq \mathbb{A}^n$ automatically translates to an inclusion $S \cong \mathbb{E}^d \subseteq \mathbb{E}^n$. Formally this can be stated as follows:

Proposition 3.7. An affine subspace of $\mathbb{E}^n$ is a Euclidean space with the scalar product inherited from $\mathbb{E}^n$.

In particular, all the results which we know to hold true for $\mathbb{E}^2$ or $\mathbb{E}^3$ will hold true when we consider 2-dimensional or 3-dimensional subspaces of the Euclidean space $\mathbb{E}^n$. This observation does not follow directly if we use Definition 3.3.

3.2.2 Orthonormal coordinate system

A vector $\mathbf{v}$ is said to be a *unit vector* if $|\mathbf{v}| = 1$. Any vector $\mathbf{w}$ is proportional to a unit vector, to $\mathbf{w}/|\mathbf{w}|$ for example. When replacing $\mathbf{w}$ by $\mathbf{w}/|\mathbf{w}|$, we say that we *normalize* the vector $\mathbf{w}$. Given two vectors $\mathbf{a}$ and $\mathbf{b}$ we define the *angle between* $\mathbf{a}$ *and* $\mathbf{b}$ to be

$$\measuredangle(\mathbf{a},\mathbf{b}) = \arccos\left(\frac{\langle \mathbf{a}, \mathbf{b} \rangle}{|\mathbf{a}| \cdot |\mathbf{b}|}\right) \in [0, \pi).$$

This definition coincides with the notion of angle discussed in dimension two (See Section 3.1). Indeed, fixing a point O, there are unique points A and B such that $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$. This defines a plane passing through O, A and B. By Proposition 3.7 we can identify this plane with $\mathbb{E}^2$. Then $\measuredangle(\mathbf{a},\mathbf{b})$ in $\mathbb{V}^n$ is the same as $\measuredangle(\mathbf{a},\mathbf{b})$ when viewing $\mathbf{a}$ and $\mathbf{b}$ as elements of $\mathbb{V}^2 \subseteq \mathbb{V}^n$. However, since the values are taken in $[0,\pi]$ this notion of angle is sometimes called *unoriented angle* - to be compared with the notion of *oriented angle* discussed in Section 4.4.2.

We say that $\mathbf{b}$ is *orthogonal to* $\mathbf{a}$ if $\langle \mathbf{b},\mathbf{a}\rangle = 0$. When this is the case, the two vectors are orthogonal to each other and we write $\mathbf{b} \perp \mathbf{a}$.

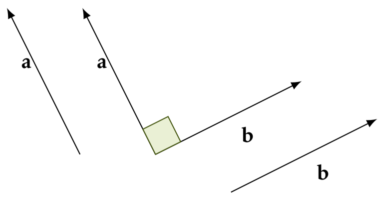


Figure 3.3: Orthogonal vectors.

Definition 3.8. A basis $\mathcal{B} = (\mathbf{e}_1,\ldots,\mathbf{e}_n)$ of $\mathbb{V}^n$ is called *orthogonal* if $\mathbf{e}_i \perp \mathbf{e}_j$ for all $i,j \in \{1,\ldots,n\}$ with $i \neq j$. The basis $\mathcal{B}$ is called *orthonormal* if it is orthogonal and all $\mathbf{e}_i$ are unit vectors. A coordinate system $\mathcal{K} = (O,\mathcal{B})$ is called *orthogonal* or *orthonormal* if the basis $\mathcal{B}$ is orthogonal or respectively orthonormal.

Theorem 3.5 says that, given a scalar product there exists an orthonormal basis in $\mathbb{V}^n$, thus, orthonormal coordinate systems exist with our definition of $\mathbb{E}^n$. Moreover, it says that with respect to such a basis the scalar product has the form (3.5). Then, one can start translating different notions into an orthonormal reference frame $\mathcal{K} = (O,\mathcal{B})$ as follows. Let $\mathbf{a}(a_1,\ldots,a_n)$ and $\mathbf{b}(b_1,\ldots,b_n)$ be vectors with components relative to $\mathcal{B}$. We have

$$|\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2},$$

$$\cos\measuredangle(\mathbf{a},\mathbf{b}) = \frac{a_1b_1 + a_2b_2 + \cdots + a_nb_n}{\sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}\sqrt{b_1^2 + b_2^2 + \cdots + b_n^2}},$$

$$\mathbf{a} \perp \mathbf{b} \Leftrightarrow a_1b_1 + a_2b_2 + \cdots + a_nb_n = 0,$$

$$d(P,Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} = |\overrightarrow{PQ}|$$

where the coordinates of $P(p_1,\ldots,p_n)$ and $Q(q_1,\ldots,q_n)$ are relative to $\mathcal{K}$.

Clearly, not all coordinate systems are orthonormal, so what do we do if we have to deal with a non-orthonormal reference frame $\mathcal{K}$? The above formulas no longer hold true. We have two options: 1. We deal with the scalar product in the given reference frame $\mathcal{K}$, or 2. We find an orthonormal reference frame $\mathcal{K}'$ starting from $\mathcal{K}$ and translate everything to $\mathcal{K}'$. We discuss these two options in the next two sections.

3.2.3 Gram matrix

We have already encountered the Gram matrix of a basis of $\mathbb{V}^n$. Given a basis $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$, it is the matrix

$$G_{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle & \langle \mathbf{e}_1, \mathbf{e}_2 \rangle & \dots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle \\ \langle \mathbf{e}_2, \mathbf{e}_1 \rangle & \langle \mathbf{e}_2, \mathbf{e}_2 \rangle & \dots & \langle \mathbf{e}_2, \mathbf{e}_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1 \rangle & \langle \mathbf{e}_n, \mathbf{e}_2 \rangle & \dots & \langle \mathbf{e}_n, \mathbf{e}_n \rangle \end{bmatrix}.$$

It is sometimes called the matrix of the bilinear map $\langle \_, \_ \rangle : \mathbb{V}^n \times \mathbb{V}^n \to \mathbb{R}$ since it consists of the values of this map on the basis vectors $\mathcal{B}$ and these uniquely determine the value of the scalar product on any other pair of vectors (see Equation (3.4)). The following proposition is an easy consequence of the definition.

Proposition 3.9. A basis $\mathcal{B}$ of $\mathbb{V}^n$ is orthonormal if and only if the Gram matrix $G_{\mathcal{B}}$ equals the identity matrix I_n.

Suppose now that the components of the basis vectors $\mathbf{e}_i$ are known relative to some orthonormal basis $\mathcal{B}'$, $\mathbf{e}_i = \mathbf{e}_i(e_{1i}, \dots, e_{ni})$ and consider the base change matrix

$$M_{\mathcal{B}'\mathcal{B}} = \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} = \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1n} \\ e_{21} & e_{22} & \dots & e_{2n} \\ \vdots & \vdots & & \vdots \\ e_{n1} & e_{n2} & \dots & e_{nn} \end{bmatrix}.$$

Notice that, since $\mathcal{B}'$ is orthonormal, we have

$$G_{\mathcal{B}} = M^T_{\mathcal{B}'\mathcal{B}} \cdot M_{\mathcal{B}'\mathcal{B}}. \tag{3.6}$$

From the above factorization of $G_{\mathcal{B}}$ we immediately deduce the following result.

Proposition 3.10. If $\mathcal{B}$ is a basis of $\mathbb{V}^n$ and $\mathcal{B}'$ is an orthonormal basis of $\mathbb{V}^n$ then

$$\det(G_{\mathcal{B}}) = \det(M_{\mathcal{B}'\mathcal{B}})^2.$$

In particular, the determinant of the Gram matrix $G_{\mathcal{B}}$ is positive. Moreover, if in addition also the basis $\mathcal{B}$ is orthonormal, then

$$M^T_{\mathcal{B}'\mathcal{B}} \cdot M_{\mathcal{B}'\mathcal{B}} = I_n.$$

Matrices satisfying the above equation are of particular importance throughout mathematics.

Definition 3.11. An $n \times n$ matrix M is called *orthogonal* if

$$M^T M = I_n.$$

In other words, the inverse of the matrix M is $M^{-1} = M^T$. The set of all such matrices form a group, the *orthogonal group* denoted by $O(n)$. We discuss $O(2)$ and $O(3)$ in Chapter 5.

Notice that, by definition, there is a bijection between orthonormal bases of $\mathbb{V}^n$ and orthogonal matrices in $O(n)$. More precisely, fixing an orthonormal basis $\mathcal{B}'$, the bijection is given by the map $\mathcal{B} \mapsto M_{\mathcal{B}'\mathcal{B}}$.

3.2.4 Gram-Schmidt process

Fix a basis $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ of $\mathbb{V}^n$. We want to construct an orthonormal basis $\mathcal{B}'$ starting from $\mathcal{B}$. Recall from (3.1) that the orthogonal projection of $\mathbf{e}_i$ on $\mathbf{e}_j$ is $\frac{\langle \mathbf{e}_j, \mathbf{e}_i \rangle}{\langle \mathbf{e}_j, \mathbf{e}_j \rangle}\mathbf{e}_j$. We construct $\mathcal{B}'$ in two steps:

1. Construct an orthogonal basis $\mathbf{e}_1', \mathbf{e}_2', \ldots, \mathbf{e}_n'$ as follows

$$\begin{aligned}
\mathbf{e}_1' &= \mathbf{e}_1 \\
\mathbf{e}_2' &= \mathbf{e}_2 - \frac{\langle \mathbf{e}_1', \mathbf{e}_2 \rangle}{\langle \mathbf{e}_1', \mathbf{e}_1' \rangle}\mathbf{e}_1' \\
\mathbf{e}_3' &= \mathbf{e}_3 - \frac{\langle \mathbf{e}_1', \mathbf{e}_3 \rangle}{\langle \mathbf{e}_1', \mathbf{e}_1' \rangle}\mathbf{e}_1' - \frac{\langle \mathbf{e}_2', \mathbf{e}_3 \rangle}{\langle \mathbf{e}_2', \mathbf{e}_2' \rangle}\mathbf{e}_2' \\
&\vdots
\end{aligned}$$

2. Normalize the vectors to obtain the basis

$$\mathcal{B}' = \left\{ \frac{\mathbf{e}_1'}{|\mathbf{e}_1'|}, \ldots, \frac{\mathbf{e}_n'}{|\mathbf{e}_n'|} \right\}.$$

This process of obtaining an orthonormal basis from a given basis is called the *Gram-Schmidt process*.

Proposition 3.12. The basis $\mathcal{B}'$ obtained from the basis $\mathcal{B}$ with the Gram-Schmidt process is an orthonormal basis.

3.3 Some applications of the scalar product

3.3.1 Area of a parallelogram

Let $\mathcal{P} = ABCD$ be a parallelogram and let $H \in AB$ be such that $DH \perp AB$. Moreover, let $\mathbf{u} = \overrightarrow{AB}$, $\mathbf{v} = \overrightarrow{AD}$ and $\mathbf{w} = \overrightarrow{HD}$.

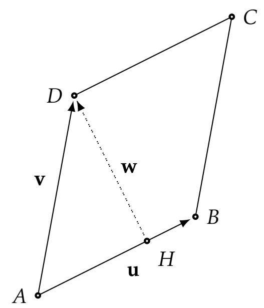


Figure 3.4: A height in a parallelogram.

With this notation $|DH|$ is the height of the parallelogram corresponding to the side $[AB]$ and the area $\text{Area}(\mathcal{P})$ of the parallelogram equals $|DH| \cdot |AB|$, i.e.

$$\text{Area}(\mathcal{P}) = |\mathbf{u}| \cdot |\mathbf{w}|.$$

From (3.1) we know that $\mathbf{w} = \mathbf{v} - \delta\mathbf{u}$ where $\delta = \frac{\langle \mathbf{u},\mathbf{v}\rangle}{|\mathbf{u}|^2}$. Thus

$$\begin{aligned} |\mathbf{w}|^2 &= \langle \mathbf{v} - \delta\mathbf{u}, \mathbf{v} - \delta\mathbf{u}\rangle \\ &= |\mathbf{v}|^2 - 2\delta\langle \mathbf{u},\mathbf{v}\rangle + \delta^2|\mathbf{u}|^2 \\ &= |\mathbf{v}|^2 - 2\frac{\langle \mathbf{u},\mathbf{v}\rangle}{|\mathbf{u}|^2}\langle \mathbf{u},\mathbf{v}\rangle + \left(\frac{\langle \mathbf{u},\mathbf{v}\rangle}{|\mathbf{u}|^2}\right)^2|\mathbf{u}|^2 \\ &= |\mathbf{v}|^2 - \frac{\langle \mathbf{u},\mathbf{v}\rangle^2}{|\mathbf{u}|^2} \end{aligned}$$

and we obtain

$$\mathrm{Area}(\mathcal{P}) = \sqrt{|\mathbf{u}|^2 \cdot |\mathbf{v}|^2 - \langle \mathbf{u},\mathbf{v}\rangle^2}.$$

Moreover, notice that if G is the Gram matrix of the basis $(\mathbf{u},\mathbf{v})$ then

$$\mathrm{Area}(\mathcal{P}) = \sqrt{\det(G)}.$$

3.3.2 Law of cosines

Let ABC be a triangle and let $\mathbf{a} = \overrightarrow{BC}$, $\mathbf{b} = \overrightarrow{AC}$ and $\mathbf{c} = \overrightarrow{AB}$ (see Figure 3.5). Denote the lengths of the sides by $a = |\mathbf{a}|$, $b = |\mathbf{b}|$ and $c = |\mathbf{c}|$ and notice that

$$\begin{aligned} a^2 &= \langle \mathbf{b} - \mathbf{c}, \mathbf{b} - \mathbf{c}\rangle \\ &= |\mathbf{b}|^2 - 2\langle \mathbf{b},\mathbf{c}\rangle + |\mathbf{c}| \\ &= b^2 - 2\langle \mathbf{b},\mathbf{c}\rangle + c^2 \end{aligned}$$

and therefore

$$\langle \mathbf{b},\mathbf{c}\rangle = \frac{1}{2}(b^2 + c^2 - a^2). \tag{3.7}$$

In particular we obtained the law of cosines: $a^2 = b^2 + c^2 - 2bc\cos(\measuredangle BAC)$.

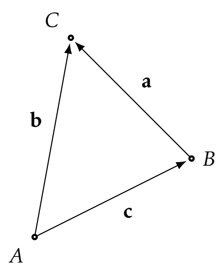


Figure 3.5: A triangle.

3.3.3 Heron's formula

Consider again the triangle ABC in Figure 3.5 with $\mathbf{a} = \overrightarrow{BC}$, $\mathbf{b} = \overrightarrow{AC}$ and $\mathbf{c} = \overrightarrow{AB}$. As before, the lengths of the sides are $a = |\mathbf{a}|$, $b = |\mathbf{b}|$ and $c = |\mathbf{c}|$. A proof of Heron's formula can be obtained with the following intermediate statement.

Lemma 3.13. Let p denote the semiperimeter $\frac{1}{2}(a+b+c)$ of the triangle ABC. Then

$$\langle \mathbf{b}, \mathbf{c}\rangle + bc = 2p(p-a) \quad \text{and} \quad \langle \mathbf{b}, \mathbf{c}\rangle - bc = -2(p-b)(p-c).$$

Theorem 3.14. Let $\mathcal{T}$ be a triangle with sides of length a, b, c. The are $\text{Area}(\mathcal{T})$ of the triangle is given by

$$\text{Area}(\mathcal{T}) = \sqrt{p(p-a)(p-b)(p-c)}$$

where p denotes the semiperimeter of the triangle $\mathcal{T}$.

3.3.4 Normal vectors of hyperplanes

Recall that, with respect to a coordinate system $\mathcal{K} = (O, \mathcal{B})$, a hyperplane $\mathcal{H}$ is given by a linear equation

$$\mathcal{H} : a_1x_1 + a_2x_2 + \cdots + a_nx_n = b. \tag{3.8}$$

Assume now that $\mathcal{K}$ is orthonormal, i.e. $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ is orthonormal. Fix a point $Q(q_1, \ldots, q_n) \in \mathcal{H}$. Since $Q \in \mathcal{H}$, we have $b = a_1q_1 + a_2q_2 + \cdots + a_nq_n$. Thus, a point $P(p_1, \ldots, p_n)$ is a solution to (3.8) if and only if

$$a_1(p_1 - q_1) + a_2(p_2 - q_2) + \cdots + a_n(p_n - q_n) = 0.$$

Therefore, if we denote by $\mathbf{n}$ the vector with components $(a_1, \ldots, a_n)$ then

$$P \in \mathcal{H} \quad \Leftrightarrow \quad \langle \mathbf{n}, \overrightarrow{QP}\rangle = 0 \quad \Leftrightarrow \quad \mathbf{n} \perp \overrightarrow{QP}.$$

In other words, the vector $\mathbf{n}$ of coefficients in (3.8) is orthogonal to any vector parallel to $\mathcal{H}$, it is orthogonal to $\mathcal{H}$.

Definition 3.15. Let $\mathcal{H}$ be a hyperplane of $\mathbb{E}^n$. A vector $\mathbf{v}$ is called a *normal vector* of $\mathcal{H}$ if it is orthogonal to $\mathcal{H}$.

Proposition 3.16. Let $\mathcal{H}$ be a hyperplane and let $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1} \in \mathrm{D}(\mathcal{H})$ be linearly independent vectors. Any normal vector of $\mathcal{H}$ is a scalar multiple of $\mathbf{v}_1 \wedge_{\mathcal{B}} \cdots \wedge_{\mathcal{B}} \mathbf{v}_{n-1}$ for any orthonormal basis $\mathcal{B}$ of $\mathbb{V}^n$.

Example 3.17. Hyperplanes in $\mathbb{E}^2$ are lines. The line with equation

$$\ell : x + 3y - 3 = 0,$$

relative to some orthonormal coordinate system, admits $\mathbf{v}(1, 3)$ as normal vector.

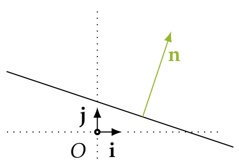


Figure 3.6: A normal vector of a line.

Example 3.18. Hyperplanes in $\mathbb{E}^3$ are planes. The plane with equation

$$\pi : 2x - y + \frac{1}{3}z + 7 = 0,$$

relative to some orthonormal coordinate system, admits $\mathbf{v}(6,-3,1)$ as normal vector.

3.3.5 Distance from a point to a hyperplane

We first consider the 2-dimensional case. Hyperplanes in $\mathbb{E}^2$ are lines. Let ℓ be a line with normal vector $\mathbf{n}$. The distance $d(P,\ell)$ from a point P to ℓ is the length of the orthogonal projection of $\overrightarrow{QP}$ on $\mathbf{n}$ for any point $Q \in \ell$ (see Section 3.1).

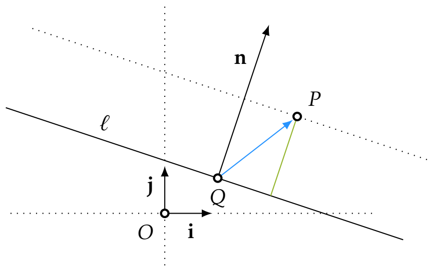


Figure 3.7: Distance from a point to a line in $\mathbb{E}^2$.

We now consider the n-dimensional case. Fix an orthonormal coordinate system $\mathcal{K}$ of $\mathbb{E}^n$ and consider a point P and the hyperplane

$$\mathcal{H} : a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

In order to describe the distance $d(P,\mathcal{H})$ from P to $\mathcal{H}$, drop a perpendicular line from P to $\mathcal{H}$, i.e. choose $H \in \mathcal{H}$ such that $PH \perp \mathcal{H}$. Then

$$d(P,\mathcal{H}) = d(P,H) = |PH|.$$

Now let Q be a point in $\mathcal{H}$ and consider the normal vector $\mathbf{n} = \mathbf{n}(a_1,\ldots,a_n)$. Then $|PH|$ is the length of the orthogonal projection of $\overrightarrow{QP}$ on $\mathbf{n}$. To see this, let N be such that $\mathbf{n} = \overrightarrow{QN}$ and look at the quadrilateral $QHPN$ (as in Figure 3.7). We have

$$d(P,\mathcal{H}) = \Big|\frac{\langle \mathbf{n}, \overrightarrow{QP}\rangle}{|\mathbf{n}|^2}\mathbf{n}\Big| = \frac{|\langle \mathbf{n}, \overrightarrow{QP}\rangle|}{|\mathbf{n}|} = \frac{|\langle \mathbf{n}, \overrightarrow{OP} - \overrightarrow{OQ}\rangle|}{|\mathbf{n}|} = \frac{|\langle \mathbf{n}, \overrightarrow{OP}\rangle - b|}{|\mathbf{n}|}.$$

If we write this explicitly in coordinates we obtain

$$d(P,\mathcal{H}) = \frac{|a_1p_1 + a_2p_2 + \cdots + a_np_n - b|}{\sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}}. \tag{3.9}$$

If $n = 2$ then you recover the distance formula from a point to a line in $\mathbb{E}^2$ and if $n = 3$ you have a formula for the distance from a point to a plane in $\mathbb{E}^3$.

Notice that a point is parallel to any affine subspace (this follows from Definition 2.14). We can generalize the distance formula (4.7) to other affine subspaces parallel to a hyperplane. Consider an affine subspace S of $\mathbb{E}^n$ which is parallel to $\mathcal{H}$. Take two points A and B and from each of them drop a perpendicular on $\mathcal{H}$ which intersects the hyperplane in M and N respectively. Since the sides MA and NB are parallel to $\mathbf{n}$, the quadrilateral $ABNM$ is planar (lies in a plane). But then AB has to be parallel to MN otherwise the two lines intersect in a point which would lie in $S \cap \mathcal{H}$ contradicting $S \| \mathcal{H}$. Since we have right angles, $ABNM$ is in fact a rectangle. This shows that $d(A,\mathcal{H}) = d(B,\mathcal{H})$. Thus, the distance from S to $\mathcal{H}$ is the distance from any point $A \in S$ to $\mathcal{H}$

$$d(S,\mathcal{H}) = d(A,\mathcal{H}) \quad \forall A \in S.$$

This, together with (4.7), gives an effective way of calculating the distance between two parallel lines in $\mathbb{E}^2$ or the distance between a plane and a parallel line in $\mathbb{E}^3$.

3.3.6 Angles between lines and hyperplanes

Let ℓ_1 and ℓ_2 be two lines in $\mathbb{E}^2$.

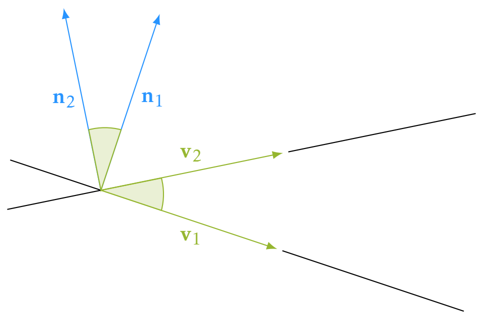


Figure 3.8: Angles between lines in $\mathbb{E}^2$.

They define two angles: if $\mathbf{v}_1$ is a direction vector for ℓ_1 and if $\mathbf{v}_2$ is a direction vector for ℓ_2 then the two angles described by ℓ_1 and ℓ_2 are $\measuredangle(\mathbf{v}_1,\mathbf{v}_2)$ and $\measuredangle(-\mathbf{v}_1,\mathbf{v}_2)$. They are supplementary angles so if you know one of them you know the other one. We may calculate this with the scalar product since

$$\cos\measuredangle(\mathbf{v}_1,\mathbf{v}_2) = \frac{\langle \mathbf{v}_1,\mathbf{v}_2\rangle}{|\mathbf{v}_1|\cdot|\mathbf{v}_2|}.$$

Notice also that the two angles can be described with normal vectors: if $\mathbf{n}_1$ and $\mathbf{n}_2$ are normal vectors for ℓ_1 and ℓ_2 respectively, then the two angles between ℓ_1 and ℓ_2 are $\measuredangle(\mathbf{n}_1,\mathbf{n}_2)$ and $\measuredangle(-\mathbf{n}_1,\mathbf{n}_2)$. So, if these vectors are known we may calculate

$$\cos\measuredangle(\mathbf{n}_1,\mathbf{n}_2) = \frac{\langle \mathbf{n}_1,\mathbf{n}_2\rangle}{|\mathbf{n}_1|\cdot|\mathbf{n}_2|}.$$

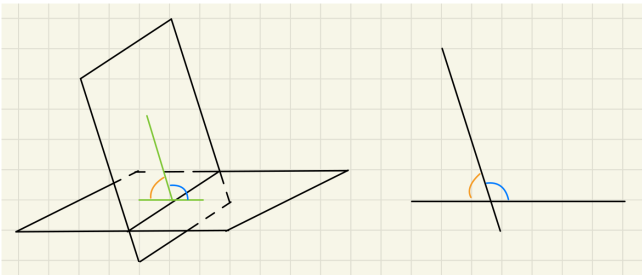
On the other hand, if we know a direction vector $\mathbf{v}_1$ for the first line and a normal vector $\mathbf{n}_2$ for the second line then the acute angle between ℓ_1 and ℓ_2 is

$$\frac{\pi}{2} - \arccos\left(\left|\frac{\langle \mathbf{v}_1, \mathbf{n}_2\rangle}{|\mathbf{v}_1| \cdot |\mathbf{n}_2|}\right|\right) \in [0, \frac{\pi}{2}).$$

This generalizes in three ways. In $\mathbb{E}^n$ consider two line ℓ_1 and ℓ_2 with direction vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ respectively as well as two hyperplanes $\mathcal{H}_1$ and $\mathcal{H}_2$ with normal vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ respectively.

1. ℓ_1 and ℓ_2 define two supplementary angles: $\measuredangle(\mathbf{v}_1, \mathbf{v}_2)$ and $\measuredangle(-\mathbf{v}_1, \mathbf{v}_2)$ which can be calculated with
$$\cos\measuredangle(\mathbf{v}_1, \mathbf{v}_2) = \frac{\langle \mathbf{v}_1, \mathbf{v}_2\rangle}{|\mathbf{v}_1| \cdot |\mathbf{v}_2|}.$$
2. $\mathcal{H}_1$ and $\mathcal{H}_2$ define two supplementary angles: $\measuredangle(\mathbf{n}_1, \mathbf{n}_2)$ and $\measuredangle(-\mathbf{n}_1, \mathbf{n}_2)$ which can be calculated with
$$\cos\measuredangle(\mathbf{n}_1, \mathbf{n}_2) = \frac{\langle \mathbf{n}_1, \mathbf{n}_2\rangle}{|\mathbf{n}_1| \cdot |\mathbf{n}_2|}.$$
3. ℓ_1 and $\mathcal{H}_1$ define two supplementary angles: if $\cos\measuredangle(\mathbf{v}_1, \mathbf{n}_1) \geq 0$ then $\measuredangle(\mathbf{v}_1, \mathbf{n}_1)$ is acute and the acute angle between ℓ_1 and $\mathcal{H}_1$ is
$$\frac{\pi}{2} - \arccos\left(\frac{\langle \mathbf{v}_1, \mathbf{n}_1\rangle}{|\mathbf{v}_1| \cdot |\mathbf{n}_1|}\right).$$
Else, if $\cos\measuredangle(\mathbf{v}_1, \mathbf{n}_1) < 0$ replace $\mathbf{n}_1$ with the normal vector $-\mathbf{n}_1$ of $\mathcal{H}$.

For (hyper)planes in $\mathbb{E}^3$ we have the notion of *dihedral angle*. Two planes, π_1 and π_2, define four dihedral angles which are the four regions in which the two planes divide $\mathbb{E}^3$. In order to measure them, let ℓ be the line $\pi_1 \cap \pi_2$. Choose a plane π orthogonal to ℓ and consider the lines $\ell_1 = \pi \cap \pi_1$ and $\ell_2 = \pi \cap \pi_2$. One can show that the measure of the angles between ℓ_1 and ℓ_2 does not depend on the choice of π. So, up to congruency we have two angles and they can be calculated using normal vectors as indicated above.



3.4 Spectral theorem

In essence, the spectral theorem says that a symmetric linear map or a symmetric bilinear map can be diagonalized in some orthonormal basis. In order to make this statement precise, we first recall some facts about linear and bilinear maps and make precise what we mean by ‘symmetric’.

Given a basis $\mathcal{B}$ of $\mathbb{V}^n$, an $n\times n$-matrix M with real entries gives rise to the linear map

$$\phi:\mathbb{V}^n\to\mathbb{V}^n,\quad \text{defined by}\quad \phi(v)=M\cdot[v]_{\mathcal{B}}$$

and to the bilinear map

$$\psi:\mathbb{V}^n\times\mathbb{V}^n\to\mathbb{R},\quad \text{defined by}\quad \psi(v,w)=[v]_{\mathcal{B}}^T\cdot M\cdot[w]_{\mathcal{B}}.$$

We say that ϕ is the *linear map associated to M in the basis $\mathcal{B}$* and ψ is the *bilinear map associated to M in the basis $\mathcal{B}$*. The other way around, given a linear map $\phi:\mathbb{V}^n\to\mathbb{V}^n$, it has an associated matrix $\mathrm{M}_{\mathcal{B},\mathcal{B}}(\phi)$ with respect to the bases $\mathcal{B}$. Similarly, given a bilinear map $\psi:\mathbb{V}^n\times\mathbb{V}^n\to\mathbb{R}$ we associate to it the Gram matrix $\mathrm{G}_{\mathcal{B}}(\psi)$ (see Section 3.2.1). Then

$$\phi(v)=\mathrm{M}_{\mathcal{B},\mathcal{B}}(\phi)\cdot[v]_{\mathcal{B}}\quad\text{and}\quad \psi(v,w)=[v]_{\mathcal{B}}^T\cdot\mathrm{G}_{\mathcal{B}}(\psi)\cdot[w]_{\mathcal{B}}.$$

If we change the basis from $\mathcal{B}$ to $\mathcal{B}'$, it is an exercise in linear algebra to show that

$$\mathrm{M}_{\mathcal{B}',\mathcal{B}'}(\phi)=\mathrm{M}_{\mathcal{B},\mathcal{B}'}^{-1}\cdot\mathrm{M}_{\mathcal{B},\mathcal{B}}(\phi)\cdot\mathrm{M}_{\mathcal{B},\mathcal{B}'}\quad\text{and}\quad \mathrm{G}_{\mathcal{B}'}(\psi)=\mathrm{M}_{\mathcal{B},\mathcal{B}'}^T\cdot\mathrm{G}_{\mathcal{B}}(\psi)\cdot\mathrm{M}_{\mathcal{B},\mathcal{B}'}\tag{3.10}$$

where $\mathrm{M}_{\mathcal{B},\mathcal{B}'}$ is the base change matrix from $\mathcal{B}'$ to $\mathcal{B}$.

Definition 3.19. A bilinear map $\psi:\mathbb{V}^n\times\mathbb{V}^n\to\mathbb{R}$ is called *symmetric* if the Gram matrix $\mathrm{G}_{\mathcal{B}}(\psi)$ is a symmetric matrix with respect to a basis $\mathcal{B}$. It follows from (3.10) that this definition doesn't depend on the basis $\mathcal{B}$.

Now consider a scalar product $\langle \_,\_\rangle$ on $\mathbb{V}^n$. Let $\mathcal{B}$ be an orthonormal basis and M and $n\times n$-symmetric matrix, i.e. $M=M^T$. Let ϕ be the linear map associated to M in the basis $\mathcal{B}$ and let ψ be the bilinear map associated to M in the basis $\mathcal{B}$. Then, since the scalar product has the form (3.5) we have

$$\psi(v,w)=[v]_{\mathcal{B}}^T\cdot M\cdot[w]_{\mathcal{B}}=\langle v,M\cdot[w]_{\mathcal{B}}\rangle=\langle v,\phi(w)\rangle.$$

Since M is symmetric, we also have

$$[v]_{\mathcal{B}}^T\cdot M\cdot[w]_{\mathcal{B}}=(M^T\cdot[v]_{\mathcal{B}})^T\cdot[w]_{\mathcal{B}}=(M\cdot[v]_{\mathcal{B}})^T\cdot[w]_{\mathcal{B}}$$

and therefore

$$\langle\phi(v),w\rangle=\psi(v,w)=\langle v,\phi(w)\rangle.\tag{3.11}$$

Definition 3.20. A linear map $\phi:\mathbb{V}^n\to\mathbb{V}^n$ is called *symmetric* if (3.11) holds for all $v,w\in\mathbb{V}^n$.

Notice that this definition depends on the scalar product. So far, we made precise what we mean by symmetric linear and bilinear maps. To simplify the proof of the spectral theorem we introduce the notion of orthogonal complement to a vector.

Definition 3.21. Let $\mathbf{v}\in\mathbb{V}^n$. The *orthogonal complement*, denoted by $\mathbf{v}^\perp$, is the set of all vectors in $\mathbb{V}^n$ which are orthogonal to $\mathbf{v}$. So

$$\mathbf{v}^\perp=\left\{\mathbf{w}\in\mathbb{V}^n:\langle\mathbf{v},\mathbf{w}\rangle=0\right\}.$$

Since the scalar product is bilinear, the map $f_{\mathbf{v}}:\mathbb{V}^n\to\mathbb{R}$ defined by $f(\mathbf{w})=\langle\mathbf{v},\mathbf{w}\rangle$ is linear and we notice that $\mathbf{v}^\perp=\ker(f_{\mathbf{v}})$. Thus, $\mathbf{v}^\perp$ is an $(n-1)$-dimensional vector subspace of $\mathbb{V}^n$.

The proof of the theorem uses the following linear algebra fact.

Lemma 3.22. The characteristic polynomial of a symmetric matrix $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ has only real roots.

Theorem 3.23 (Spectral Theorem)**.** Let $\phi : \mathbb{V}^n \to \mathbb{V}^n$ be a symmetric linear map. Then, there exists an orthonormal basis $\mathcal{B}'$ such that $\mathrm{M}_{\mathcal{B}',\mathcal{B}'}(\phi)$ is a diagonal matrix.

Corollary 3.24. Let M be a symmetric matrix. Let $\psi : \mathbb{V}^n \times \mathbb{V}^n \to \mathbb{R}$ be the bilinear map associated to M relative to an orthonormal basis $\mathcal{B}$. Then, there exists an orthonormal basis $\mathcal{B}'$ such that

$$\mathrm{G}_{\mathcal{B}'}(\psi) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of M.

3.5 Exercises

3.1. Let $\mathbf{m}$ and $\mathbf{n}$ be two unit vectors such that $\measuredangle(\mathbf{m},\mathbf{n}) = 60°$. Determine the length of the diagonals in the parallelogram spanned by the vectors $\mathbf{a} = 2\mathbf{m} + \mathbf{n}$ and $\mathbf{b} = \mathbf{m} - 2\mathbf{n}$.

3.2. Let $\mathbf{m}$ and $\mathbf{n}$ be two unit vectors such that $\measuredangle(\mathbf{m},\mathbf{n}) = 120°$. Determine the angle between the vectors $\mathbf{a} = 2\mathbf{m} + 4\mathbf{n}$ and $\mathbf{b} = \mathbf{m} - \mathbf{n}$.

3.3. You are given two vectors $\mathbf{a}(2,1,0)$ and $\mathbf{b}(0,-2,1)$ with respect to an orthonormal basis. Determine the angles between the diagonals of the parallelogram spanned by $\mathbf{a}$ and $\mathbf{b}$.

3.4. Let $(\mathbf{i},\mathbf{j},\mathbf{k})$ be an orthonormal basis. Consider the vectors $\mathbf{q} = 3\mathbf{i} + \mathbf{j}$ and $\mathbf{p} = \mathbf{i} + 2\mathbf{j} + \lambda\mathbf{k}$ with $\lambda \in \mathbb{R}$. Determine λ such that the cosine of the angle $\measuredangle(\mathbf{p},\mathbf{q})$ is $5/12$.

3.5. Using the scalar product, prove the Cauchy-Bunyakovsky-Schwarz inequality, i.e. show that for any $a_1,\ldots,a_n,b_1,\ldots,b_n \in \mathbb{R}$ we have

$$(a_1b_1 + \cdots + a_nb_n)^2 \le (a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2).$$

3.6. Let ABC be a triangle. Show that

$$\overrightarrow{AB}^2 + \overrightarrow{AC}^2 - \overrightarrow{BC}^2 = 2\overrightarrow{AB} \cdot \overrightarrow{AC}$$

and deduce the law of cosines in a triangle.

3.7. Let $ABCD$ be a tetrahedron. Show that

$$\cos(\measuredangle(\overrightarrow{AB}, \overrightarrow{CD})) = \frac{AD^2 + BC^2 - AC^2 - BD^2}{2 \cdot AB \cdot CD}.$$

This is a $3D$-version of the law of cosine.

3.8. Let $ABCD$ be a rectangle. Show that for any point O

$$\overrightarrow{OA} \cdot \overrightarrow{OC} = \overrightarrow{OB} \cdot \overrightarrow{OD} \quad \text{and} \quad \overrightarrow{OA}^2 + \overrightarrow{OC}^2 = \overrightarrow{OB}^2 + \overrightarrow{OD}^2.$$

3.9. Consider the vector $\mathbf{v}$ which is perpendicular on $\mathbf{a}(4,-2,-3)$ and on $\mathbf{b}(0,1,3)$. If $\mathbf{v}$ describes an acute angle with Ox and $|\mathbf{v}| = 26$ determine the components of $\mathbf{v}$.

3.10. In an orthonormal basis, consider the vectors $\mathbf{v}_1(0,1,0)$, $\mathbf{v}_2(2,1,0)$ and $\mathbf{v}_3(-1,0,1)$. Use the Gram-Schmidt process to find an orthonormal basis containing $\mathbf{v}_1$.

3.11. Let $\mathbf{v} \in \mathbb{V}^n$ be a vector. Show that the set $\mathbf{v}^\perp$ is an $(n-1)$-dimensional vector subspace of $\mathbb{V}^n$. Deduce that there is a basis $\mathbf{v},\mathbf{v}_2,\mathbf{v}_3,\ldots,\mathbf{v}_n$ of $\mathbb{V}^n$ with $\mathbf{v}_2,\ldots,\mathbf{v}_{n-1}$ a basis of $\mathbf{v}^\perp$. (*Hint.* use Steinitz Theorem - Algebra, Lecture 6).

3.12. Determine a Cartesian equations for the line ℓ in the following cases:

a) ℓ contains the point $A(-2,3)$ and has an angle of $60°$ with the Ox-axis,

b) ℓ contains the point $B(1,7)$ and is orthogonal to $\mathbf{n}(4,3)$.

3.13. For the lines ℓ in the previous exercise

a) give parametric equations for ℓ,

b) describe $\mathrm{D}(\ell)$.

3.14. Consider a line ℓ. Show that

c) if $\mathbf{v}(v_1,v_2)$ is a direction vector for ℓ then $\mathbf{n}(v_2,-v_1)$ is a normal vector for ℓ,

d) if $\mathbf{n}(n_1,n_2)$ is a normal vector for ℓ then $\mathbf{v}(n_2,-n_1)$ is a direction vector for ℓ.

3.15. Consider the points $A(1,2)$, $B(-2,3)$ and $C(4,7)$. Determine the medians of the triangle ABC.

3.16. Let $M_1(1,2)$, $M_2(3,4)$ and $M_3(5,-1)$ be the midpoints of the sides of a triangle. Determine Cartesian equations and parametric equations for the lines containing the sides of the triangle.

3.17. Let $A(1,3)$, $B(-4,3)$ and $C(2,9)$ be the vertices of a triangle. Determine

a) the length of the altitude from A,

b) the line containing the altitude from A.

3.18. Determine the circumcenter and the orthocenter of the triangle with vertices $A(1,2)$, $B(3,-2)$, $C(5,6)$.

3.19. Determine the angle between the lines $\ell_1 : y = 2x+1$ and $\ell_2 : y = -x+2$.

3.20. Let $A(1,-2)$, $B(5,4)$ and $C(-2,0)$ be the vertices of a triangle. Determine the equations of the angle bisectors for the angle $\measuredangle A$.

3.21. Let A' be the orthogonal reflection of $A(10,10)$ in the line $\ell : 3x+4y-20 = 0$. Determine the coordinates of A'.

3.22. Determine Cartesian equations for the lines passing through $A(-2,5)$ which intersect the coordinate axes in congruent segments.

3.23. Determine Cartesian equations for the lines situated at distance 4 from the line $12x-5y-15 = 0$.

3.24. Determine the values k for which the distance from the point $(2,3)$ to the line $8x+15y+k = 0$ equals 5.

3.25. Consider the points $A(3,-1)$, $B(9,1)$ and $C(-5,5)$. For each pair of these three points, determine the line which is equidistant from them.

3.26. The point $A(3,-2)$ is the vertex of a square and $M(1,1)$ is the intersection point of its diagonals. Determine Cartesian equations for the sides of the square.

3.27. Determine a point on the line $5x-4y-4 = 0$ which is equidistant to the points $A(1,0)$ and $B(-2,1)$.

3.28. The point $A(2,0)$ is the vertex of an equilateral triangle. The side opposite to A lies on the line $x+y-1=0$. Determine Cartesian equations for the lines containing the other two sides.

3.29. Determine an equation for each plane passing through $P(3,5,-7)$ and intersecting the coordinate axes in congruent segments.

3.30. Let $A(2,1,0)$, $B(1,3,5)$, $C(6,3,4)$, $D(0,-7,8)$ be vertices of a tetrahedron. Determine a Cartesian equation of the plane containing $[AB]$ and the midpoint of $[CD]$.

3.31. Show that a parallelepiped with faces in the planes $2x+y-2z+6=0$, $2x-2y+z-8=0$ and $x+2y+2z+1=0$ is rectangular.

3.32. Determine a Cartesian equation of the plane π if $A(1,-1,3)$ is the orthogonal projection of the origin on π.

3.33. Determine the distance between the planes $x-2y-2z+7=0$ and $2x-4y-4z+17=0$.

3.34. Solve Exercise 2.16 using normal vectors.

3.35. Let $A(1,2,-7)$, $B(2,2,-7)$ and $C(3,4,-5)$ be vertices of a triangle. Determine the equation of the internal angle bisector of $\measuredangle A$.

3.36. Determine the angles between the plane $\pi_1 : x-\sqrt{2}y+z-1=0$ and the plane $\pi_2 : x+\sqrt{2}y-z+3=0$.

3.37. Determine the values a and c for which the line $3x-2y+z+3=0 \cap 4x-3y+4z+1=0$ is perpendicular to the plane $ax+8y+cz+2=0$.

3.38. Determine the orthogonal projection of the point $A(2,11,-5)$ on the plane $x+4y-3z+7=0$.

3.39. Determine the orthogonal reflection of the point $P(6,-5,5)$ in the plane $2x-3y+z-4=0$.

3.40. Consider the point $A(1,3,5)$ and the line $\ell : 2x+y+z-1=0 \cap 3x+y+2z-3=0$.

a) Determine the orthogonal projection of A on ℓ.

b) Determine the orthogonal reflection of A in ℓ.

3.41. Determine the planes which pass through $P(0,2,0)$ and $Q(-1,0,0)$ and which form an angle of $60°$ with the z-axis.

3.42. Determine the orthogonal projection of the line $\ell : 2x-y-1=0 \cap x+y-z+1=0$ on the plane $\pi : x+2y-z=0$.

3.43. Determine the coordinates of a point A on the line $\ell : \frac{x-1}{2} = \frac{y}{3} = \frac{z+1}{1}$ which is at distance $\sqrt{3}$ from the plane $x+y+z+3=0$.

3.44. The vertices of a tetrahedron are $A(-1,-3,1)$, $B(5,3,8)$, $C(-1,-3,5)$ and $D(2,1,-4)$. Determine the height of the tetrahedron relative to the face ABC.

3.45. In $\mathbb{E}^4$, determine the angles between the hyperplanes:

$$H_1 : 3w - x - y + z + 2 = 0 \quad \text{and} \quad H_2 : -w + 2x + y - z + 2 = 0.$$

3.46. In $\mathbb{E}^4$, show that the lines

$$\ell_1 : \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 2 \\ -1 \end{bmatrix} \quad \text{and} \quad \ell_2 : \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix}$$

are skew and determine the hyperplane containing them.

3.47. In $\mathbb{E}^4$, determine a point on the first coordinate axis which is equidistant from $P_1(1,-1,0,2)$ and $P_2(0,-2,0,1)$.

3.48. In $\mathbb{E}^4$, consider the line

$$\ell : \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 2 \\ -1 \end{bmatrix} \quad \text{and the hyperplane} \quad H : \lambda w - x - y + \mu z + 2 = 0.$$

For what values λ and μ is H orthogonal to ℓ? For what values λ and μ is H parallel to ℓ?

CHAPTER 4

Orientation and volume

Contents

4.1 Orientation in $\mathbb{E}^n$

Consider the bases $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2)$ and $\mathcal{F} = (\mathbf{f}_1, \mathbf{f}_2)$ of $\mathbb{V}^2$. The vectors can be represented with points in $\mathbb{E}^2$. Represent all of them in a common point O and rotate the basis $\mathcal{F}$ such that $\mathbf{f}_1$ points in the same direction as $\mathbf{e}_1$, i.e. such that $\mathbf{f}_1 = \lambda \mathbf{e}_1$ for some positive scalar λ.

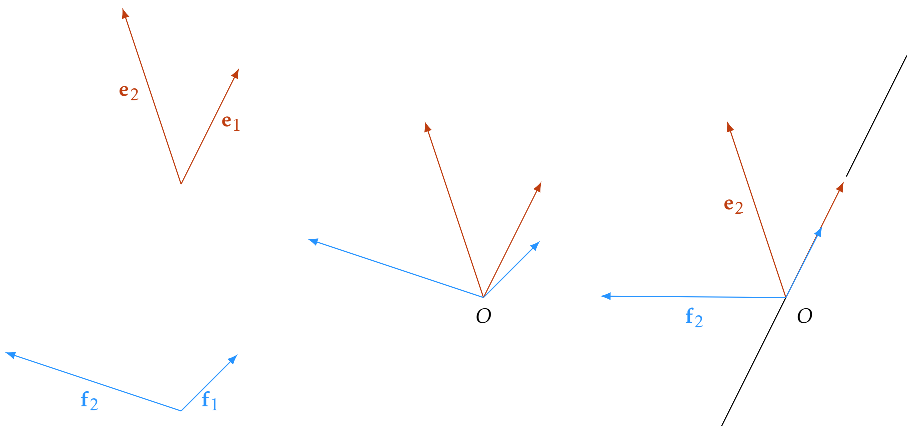


Figure 4.1: Align $\mathbf{f}_1$ with $\mathbf{e}_1$.

The line passing through O with direction vector $\mathbf{e}_1$ (and also $\mathbf{f}_1$) separates the plane $\mathbb{E}^2$ in two half-planes. Considering the positioning of the second vectors, two things can happen: $\mathbf{e}_2$ and $\mathbf{f}_2$ point towards the same half-plane or they point towards different half-planes.

In general, if we have two such bases, how can we tell the two cases apart? We look at the sign $\det(\mathrm{M}_{\mathcal{E},\mathcal{F}})$:

$$\begin{cases} \text{if } \det(\mathrm{M}_{\mathcal{E},\mathcal{F}}) > 0 & \text{then } \mathbf{e}_2 \text{ and } \mathbf{f}_2 \text{ point to the same half-plane.} \\ \text{if } \det(\mathrm{M}_{\mathcal{E},\mathcal{F}}) < 0 & \text{then } \mathbf{e}_2 \text{ and } \mathbf{f}_2 \text{ point to different half-planes.} \end{cases} \tag{4.1}$$

Why? In the above process, we changed the basis $\mathcal{F} = (\mathbf{f}_1, \mathbf{f}_2)$ to the basis $\mathcal{F}' = (\mathbf{f}_1', \mathbf{f}_2')$ with a rotation. So, the base change matrix $\mathrm{M}_{\mathcal{F}',\mathcal{F}}$ is a 2×2-rotation matrix which has determinant equal to 1 (For more on rotations see Chapter 5). Now, considering the coordinates of the vectors in $\mathcal{F}'$ with respect to $\mathcal{E}$, we have $\mathbf{f}_1' = (\lambda, 0)$ and $\mathbf{f}_2' = (a, b)$. Thus, we notice that the vectors $\mathbf{e}_1$ and $\mathbf{f}_1$ point in the same half-plane if $b > 0$. If we now calculate the above determinant, we obtain

$$\det(\mathrm{M}_{\mathcal{E},\mathcal{F}}) = \det(\mathrm{M}_{\mathcal{E},\mathcal{F}'} \cdot \mathrm{M}_{\mathcal{F}',\mathcal{F}}) = \det(\mathrm{M}_{\mathcal{E},\mathcal{F}'}) \cdot \underbrace{\det(\mathrm{M}_{\mathcal{F}',\mathcal{F}})}_{=1} = \begin{vmatrix} \lambda & a \\ 0 & b \end{vmatrix} = \lambda \cdot b$$

and since λ is positive we proved (4.1). In the first case of (4.1) we say that $\mathcal{E}$ and $\mathcal{F}$ have the *same orientation* and in the second case we say that $\mathcal{E}$ and $\mathcal{F}$ have *opposite orientation*.

Now consider the bases $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ and $\mathcal{F} = (\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3)$ of $\mathbb{V}^3$. Represent all the vectors in a common point O and rotate the basis $\mathcal{F}$ such that the plane passing through O in the direction of $\mathbf{f}_1, \mathbf{f}_2$ coincides with the plane π passing through O in the direction of $\mathbf{e}_1, \mathbf{e}_2$. If in the plane π the bases $(\mathbf{f}_1, \mathbf{f}_2)$ and $(\mathbf{e}_1, \mathbf{e}_2)$ have opposite orientation, flip the vectors $(\mathbf{f}_1, \mathbf{f}_2)$ with a rotation such that they end up having the same orientation with $(\mathbf{e}_1, \mathbf{e}_2)$. Any rotation with 180° around a line in the plane π which passes through the origin will work and such a rotation has determinant equal to 1 (See Chapter 5).

Then, the plane π separates the space $\mathbb{E}^3$ in two half-spaces. Considering the positioning of the third vectors, two things can happen: $\mathbf{e}_3$ and $\mathbf{f}_3$ point towards the same half-space or they point towards different half-spaces. How can we tell the two cases apart? A similar argument as above shows that the sign of $\det(\mathrm{M}_{\mathcal{E},\mathcal{F}})$ gives the answer.

From a different perspective, if we consider two bases $\mathcal{E}$ and $\mathcal{F}$ of $\mathbb{V}^n$, then $\det(\mathrm{M}_{\mathcal{E},\mathcal{F}})$ is not zero - it is either positive or negative. Thus, if we fix $\mathcal{E}$ the other bases fall in two classes:

$$\begin{cases} \text{if } \det(\mathrm{M}_{\mathcal{E},\mathcal{F}}) > 0 & \text{we say that } \mathcal{F} \text{ has the same orientation as } \mathcal{E}. \\ \text{if } \det(\mathrm{M}_{\mathcal{E},\mathcal{F}}) < 0 & \text{we say that } \mathcal{E} \text{ and } \mathcal{F} \text{ have opposite orientation.} \end{cases}$$

Why is this relevant? Next to the fact that it gives a geometric interpretation of the sign of $\det(\mathrm{M}_{\mathcal{E},\mathcal{F}})$, it also allows us to understand some signs which appear in calculations of areas and volumes later on. Moreover, the trigonometry that we know to hold true in $\mathbb{E}^2$ implicitly builds on the notion of oriented Euclidean plane $\mathbb{E}^2$ (See Section 4.4.2). Mathematically, the distinction in orientation is only a matter of keeping track of the signs of some determinants. However, in relation to the physical world this distinction is more concrete.

Definition 4.1. Let $(\mathbf{i}, \mathbf{j})$ be a basis of $\mathbb{V}^2$ represented in a common point $O \in \mathbb{E}^2$ such that $\mathbf{i} = \overrightarrow{OX}$ and $\mathbf{j} = \overrightarrow{OY}$. Rotate the plane such that $\mathbf{i}$ points downwards. If Y is in the right half-plane determined by the line OX, then we say that the basis $(\mathbf{i}, \mathbf{j})$ is *right oriented*. If Y lies in the left half-plane, we say that the basis $(\mathbf{i}, \mathbf{j})$ is *left oriented*. A coordinate system $(O, \mathbf{i}, \mathbf{j})$ is *left* or *right oriented* if the basis $(\mathbf{i}, \mathbf{j})$ is left respectively right oriented.

Let $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ be a basis of $\mathbb{V}^3$ represented in a common point $O \in \mathbb{E}^3$ such that $\mathbf{k} = \overrightarrow{OZ}$. We say that the basis is *right oriented* if $(\mathbf{i}, \mathbf{j})$ is a right oriented basis of the plane Oxy when observed from the point Z. We say that the basis is *left oriented* if $(\mathbf{i}, \mathbf{j})$ is a left oriented basis when observed from the point Z. A coordinate system $(O, \mathbf{i}, \mathbf{j}, \mathbf{k})$ is *left* or *right oriented* if the basis $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ is left respectively right oriented.

There are many equivalent ways of deciding if a basis is left or right oriented. The Swiss liked the three-finger rule so much, they put it on their 200-franc banknotes:

Fixing an orientation in the Euclidean plane $\mathbb{E}^2$ is equivalent to choosing a coordinate system $\mathcal{K} = (O,\mathcal{B})$ and calling it right oriented. Then, all other bases of $\mathbb{V}^2$ either have the same orientation as $\mathcal{B}$, in which case they are also called right oriented, or they have opposite orientation, in which case they are called left oriented. When it comes to a concrete configuration of points, on a sheet of paper for instance, such a choice can be made with the right-hand rule. Once we have such a choice, $\mathbb{E}^2$ is called oriented. In other words, the *oriented plane* $\mathbb{E}^2$ is the usual Euclidean plane together with a choice of which of the two opposite classes of bases contains the 'prefered' bases. In general:

Definition 4.2. We say that the Euclidean space $\mathbb{E}^n$ is *oriented* if there is a choice of a coordinate system $\mathcal{K} = (O,\mathcal{B})$ which is called *right oriented*. Then, all other bases of $\mathbb{V}^n$ with the same orientation as $\mathcal{B}$ are also *right oriented* and all other bases with opposite orientation are called *left oriented*.

Remark. Unless otherwise stated, whenever we consider a coordinate system $\mathcal{K} = (O,\mathcal{B})$ of $\mathbb{E}^n$, we will assume that it is right oriented and that $\mathbb{E}^n$ is therefore an oriented Euclidean space.

4.2 Box product in $\mathbb{E}^n$

Let $\mathbf{v}_1,\dots,\mathbf{v}_n$ be vectors in $\mathbb{V}^n$ with components $\mathbf{v}_i(v_{i,1},\dots,v_{i,n})$ relative to a basis $\mathcal{E}$. We know from linear algebra that these vectors are linearly independent if the determinant

$$\begin{vmatrix} v_{1,1} & v_{2,1} & \dots & v_{n,1} \\ v_{1,2} & v_{2,2} & \dots & v_{n,2} \\ \vdots & \vdots & & \vdots \\ v_{1,n} & v_{2,n} & \dots & v_{n,n} \end{vmatrix}$$

is non-zero. Furthermore, in the previous section we gave a geometric meaning to the sign of $\det(\mathrm{M}_{\mathcal{E},\mathcal{B}})$. Can we say more about the value of this determinant?

Consider first the Euclidean plane $\mathbb{E}^2$. Let $\mathcal{B} = (\mathbf{v},\mathbf{w})$ be a basis of $\mathbb{V}^2$ which we know relative to some other basis $\mathcal{E}$, i.e. $\mathbf{v} = \mathbf{v}(v_1,v_2)$ and $\mathbf{w} = \mathbf{w}(w_1,w_2)$. Now, if $\mathcal{E}$ is an orthonormal basis, then we have seen in Section 3.2.3 that

$$\mathrm{M}^T_{\mathcal{E},\mathcal{B}} \cdot \mathrm{M}_{\mathcal{E},\mathcal{B}} = \mathrm{G}_{\mathcal{B}}.$$

Since, by Section 3.3.1, $\sqrt{\det(\mathrm{G}_{\mathcal{B}})}$ is the area of a parallelogram spanned by $\mathbf{v}$ and $\mathbf{w}$, we conclude that

$$|\det(\mathrm{M}_{\mathcal{E},\mathcal{B}})| \quad \text{is the area of a parallelogram spanned by the basis vectors of } \mathcal{B}.$$

We have the similar statement - but with volume instead of area - in dimension 3 which we deduce in Section 4.4.5 where we also generalize to higher dimensions. Because of the above observations, we make the following definition.

Definition 4.3. Let $\mathbf{v}_1,\dots,\mathbf{v}_n$ be vectors in $\mathbb{V}^n$ with components $\mathbf{v}_i(v_{i,1},\dots,v_{i,n})$ relative to a right oriented orthonormal basis $\mathcal{B}$ of $\mathbb{V}^n$. The *(n-fold) box product* of these vectors is

$$[\mathbf{v}_1,\dots,\mathbf{v}_n] = \begin{vmatrix} v_{1,1} & v_{2,1} & \dots & v_{n,1} \\ v_{1,1} & v_{2,2} & \dots & v_{n,2} \\ \vdots & \vdots & & \vdots \\ v_{1,n} & v_{2,n} & \dots & v_{n,n} \end{vmatrix}.$$

Proposition 4.4. The definition of the box product does not depend on the choice of the right oriented orthonormal basis $\mathcal{B}$.

Definition 4.5. Let $\mathcal{B}$ be a basis of $\mathbb{V}^n$. The *oriented volume* of the basis $\mathcal{B}$, denoted by $\mathrm{Vol}_{\mathrm{or}}(\mathcal{B})$, is the value of the box product of the vectors in $\mathcal{B}$. The *volume* of the basis $\mathcal{B}$ is the absolute value $|\mathrm{Vol}_{\mathrm{or}}(\mathcal{B})|$ and we denote it by $\mathrm{Vol}(\mathcal{B})$. If we are in dimension 2, we refer to these values as the *area* and *oriented area* of $\mathcal{B}$, and denote them by $\mathrm{Area}(\mathcal{B})$ and $\mathrm{Area}_{\mathrm{or}}(\mathcal{B})$ respectively.

Since the box product is just a determinant we can immediately derive certain properties which we know from linear algebra. For example, if we restrict our attention to the 3-dimensional Euclidean space $\mathbb{E}^3$ we deduce the following properties for three vectors $\mathbf{a},\mathbf{b},\mathbf{c}$:

$$[\mathbf{a},\mathbf{b},\mathbf{c}] = [\mathbf{b},\mathbf{c},\mathbf{a}] = [\mathbf{c},\mathbf{a},\mathbf{b}] = -[\mathbf{b},\mathbf{a},\mathbf{c}] = -[\mathbf{a},\mathbf{c},\mathbf{b}] = -[\mathbf{c},\mathbf{b},\mathbf{a}];$$

the coplanarity (i.e. linear dependency) condition for $\mathbf{a},\mathbf{b},\mathbf{c}$ is

$$[\mathbf{a},\mathbf{b},\mathbf{c}] = 0;$$

the orientation of the basis $(\mathbf{a},\mathbf{b},\mathbf{c})$ is determined as follows

$$[\mathbf{a},\mathbf{b},\mathbf{c}] \begin{cases} > 0 & \text{then } (\mathbf{a},\mathbf{b},\mathbf{c}) \text{ is right oriented} \\ = 0 & \text{then } (\mathbf{a},\mathbf{b},\mathbf{c}) \text{ is not a basis} \\ < 0 & \text{then } (\mathbf{a},\mathbf{b},\mathbf{c}) \text{ is left oriented} \end{cases}.$$

4.3 Cross product (in $\mathbb{E}^3$)

Before describing the cross product, we consider the analogue notion in dimension 2, the $\mathbf{J}$-operator.

4.3.1 The J-operator in $\mathbb{E}^2$

Let us have a closer look at the wedge product in $\mathbb{E}^2$. Recall that for a vector $\mathbf{v} = (v_1, v_2)$ with respect to a basis $\mathcal{B} = (\mathbf{i},\mathbf{j})$, this is

$$\wedge_{\mathcal{B}}(\mathbf{v}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} \\ v_1 & v_2 \end{vmatrix} = v_2\mathbf{i} - v_1\mathbf{j} = \begin{bmatrix} v_2 \\ -v_1 \end{bmatrix}$$

In Section 2.4.2 we looked at this from the perspective of the affine space $\mathbb{A}^n$. If we consider a scalar product and an orientation on $\mathbb{V}^2$, we are in the oriented Euclidean plane $\mathbb{E}^2$ and we can say more.

Definition 4.6. For a vector $\mathbf{v} \in \mathbb{V}^2$ we define $\mathbf{J}(\mathbf{v})$ to be the (unique) vector in $\mathbb{V}^2$ satisfying the following properties

(a) $\mathbf{J}(\mathbf{v})$ is orthogonal to $\mathbf{v}$

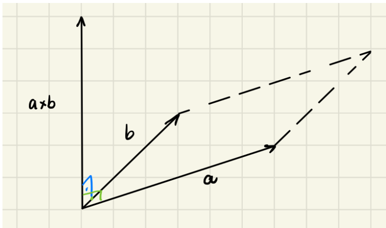
(b) $|\mathbf{J}(\mathbf{v})| = |\mathbf{v}|$

(c) $(\mathbf{v}, \mathbf{J}(\mathbf{v}))$ is a right oriented basis of $\mathbb{V}^2$.

Proposition 4.7. The $\mathbf{J}$-operator is given by $J(\mathbf{v}) = -\wedge_{\mathcal{B}}(\mathbf{v})$ where $\mathcal{B}$ is any right oriented orthonormal basis of $\mathbb{V}^2$. In particular, $\mathbf{J} : \mathbb{V}^2 \to \mathbb{V}^2$ is a linear map.

4.3.2 Definition and properties of the cross product

Here we restrict to dimension 3, i.e. we consider the oriented Euclidean space $\mathbb{E}^3$. For two vectors $\mathbf{a}$ and $\mathbf{b}$ the orthogonal complements $\mathbf{a}^\perp$ and $\mathbf{b}^\perp$ are 1-dimensional vector subspaces of $\mathbb{V}^3$. If the vectors $\mathbf{a}$ and $\mathbf{b}$ are linearly independent, then $\mathbf{a}^\perp \cap \mathbf{b}^\perp$ is a 1-dimensional vector subspace of $\mathbb{V}^3$. In other words, up to scalar multiple, there is a unique vector $\mathbf{c}$ perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. When choosing $\mathbf{c}$ to be a unit vector, we have exactly two options: $\pm\mathbf{c}$. For these two options, $(\mathbf{a},\mathbf{b},\mathbf{c})$ is either left- or right-oriented. Fixing such an orientation we have uniqueness for the choice of such $\mathbf{c}$. This brings us to the following definition.



Definition 4.8. Let $\mathbf{a},\mathbf{b} \in \mathbb{V}^3$ be two vectors. The *cross product* (or *vector product*) of $\mathbf{a}$ and $\mathbf{b}$ is the vector denoted by $\mathbf{a}\times\mathbf{b}$ and defined by the following properties:

1. if $\mathbf{a}$ and $\mathbf{b}$ are parallel then $\mathbf{a}\times\mathbf{b} = 0$.

2. if $\mathbf{a}$ and $\mathbf{b}$ are not parallel, then

   (a) $|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}|\cdot|\mathbf{b}|\cdot\sin\measuredangle(\mathbf{a},\mathbf{b})$,

   (b) $\mathbf{a}\times\mathbf{b} \perp \mathbf{a}$ and $\mathbf{a}\times\mathbf{b} \perp \mathbf{b}$,

   (c) $(\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b})$ is a right oriented basis of $\mathbb{V}^3$.

Notice that the angle $\measuredangle(\mathbf{a},\mathbf{b})$ lies between 0 and π, so $\sin\measuredangle(\mathbf{a},\mathbf{b}) \geq 0$. Notice also that the norm $|\mathbf{a}\times\mathbf{b}|$ equals the area of a parallelogram spanned by the two vectors; and, two vectors $\mathbf{a}$ and $\mathbf{b}$ are parallel if and only if $\mathbf{a}\times\mathbf{b} = 0$.

Proposition 4.9. The cross product is given by $\mathbf{a}\times\mathbf{b} = \mathbf{a}\wedge_{\mathcal{B}}\mathbf{b}$ where $\mathcal{B}$ is any right oriented orthonormal basis of $\mathbb{V}^3$.

Corollary 4.10. The cross product $\_\times\_ : \mathbb{V}^3\times\mathbb{V}^3 \to \mathbb{V}^3$ satisfies the following properties.

(CP1) It is *bilinear*, i.e. for all $a,b\in\mathbb{R}$ and all $\mathbf{v},\mathbf{w},\mathbf{u}\in\mathbb{V}^2$ we have

$$(a\mathbf{v}+b\mathbf{w})\times\mathbf{u} = a(\mathbf{v}\times\mathbf{u})+b(\mathbf{w}\times\mathbf{u}) \quad \text{and} \quad \mathbf{v}\times(a\mathbf{w}+b\mathbf{u}) = a(\mathbf{v}\times\mathbf{w})+b(\mathbf{v}\times\mathbf{u}).$$

(CP2) It is *skew-symmetric*, i.e. for all $\mathbf{v},\mathbf{w}\in\mathbb{V}^2$ we have

$$\mathbf{v}\times\mathbf{w} = -\mathbf{w}\times\mathbf{v}.$$

Since the cross product is bilinear, all its values are determined by the values on a basis. If $(\mathbf{i},\mathbf{j},\mathbf{k})$ is a right oriented orthonormal basis, the values on the basis $(\mathbf{i},\mathbf{j},\mathbf{k})^2$ of $\mathbb{V}^3\times\mathbb{V}^3$ are

$$\begin{array}{c|ccc} \times & \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \hline \mathbf{i} & 0 & \mathbf{k} & -\mathbf{j} \\ \mathbf{j} & -\mathbf{k} & 0 & \mathbf{i} \\ \mathbf{k} & \mathbf{j} & -\mathbf{i} & 0 \end{array}.$$

Moreover, since we are working in a right oriented basis, from Proposition 4.9 it also follows that

$$\langle \mathbf{a}\times\mathbf{b},\mathbf{c}\rangle = \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = [\mathbf{a},\mathbf{b},\mathbf{c}].$$

Thus, mixing the cross product with the scalar product gives the box product. For this reason $\langle\mathbf{a}\times\mathbf{b},\mathbf{c}\rangle$ is sometimes called the *mixed product* of the vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. Notice that the box product in this case is a map

$$[\_,\_,\_]:\mathbb{V}^3\times\mathbb{V}^3\times\mathbb{V}^3\to\mathbb{R} \quad (\mathbf{a},\mathbf{b},\mathbf{c})\mapsto[\mathbf{a},\mathbf{b},\mathbf{c}]=\langle\mathbf{a}\times\mathbf{b},\mathbf{c}\rangle$$

which is linear in each argument since it is the composition of linear maps. You can also deduce the linearity of this map from the properties of the determinant.

4.3.3 Some identities involving the cross product

The cross product is not associative. For example, if $(\mathbf{i},\mathbf{j},\mathbf{k})$ is a right oriented orthonormal basis then

$$(\mathbf{i}\times\mathbf{j})\times\mathbf{j}=\mathbf{k}\times\mathbf{j}=-\mathbf{i} \quad \text{whereas} \quad \mathbf{i}\times(\mathbf{j}\times\mathbf{j})=0.$$

However, there is a rule which shows what happens when we evaluate an iterated product in different ways. This is the *Jacobi identity*:

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}+(\mathbf{b}\times\mathbf{c})\times\mathbf{a}+(\mathbf{c}\times\mathbf{a})\times\mathbf{b}=0. \quad \forall\mathbf{a},\mathbf{b},\mathbf{c}\in\mathbb{V}^3. \tag{4.2}$$

One way of proving this identity is to write out the above expression in coordinates and check that the expression on the left-hand side is the zero vector. A shorter way to prove (4.2) is to make use of *Grassmann's cross product formula*:

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}=\langle\mathbf{a},\mathbf{c}\rangle\cdot\mathbf{b}-\langle\mathbf{b},\mathbf{c}\rangle\cdot\mathbf{a}. \quad \forall\mathbf{a},\mathbf{b},\mathbf{c}\in\mathbb{V}^3 \tag{4.3}$$

which can be checked separately in coordinates. Notice that (4.3) allows us to calculate two scalar products instead of calculating a determinant.

Other identities can be deduced by writing both the left- and the right-hand side in coordinates. For example, we can describe the scalar product of two cross products with *Lagrange's identity*:

$$\langle\mathbf{a}\times\mathbf{b},\mathbf{c}\times\mathbf{d}\rangle=\langle\mathbf{a},\mathbf{c}\rangle\cdot\langle\mathbf{b},\mathbf{d}\rangle-\langle\mathbf{b},\mathbf{c}\rangle\cdot\langle\mathbf{a},\mathbf{d}\rangle=\begin{vmatrix}\langle\mathbf{a},\mathbf{c}\rangle & \langle\mathbf{a},\mathbf{d}\rangle \\ \langle\mathbf{b},\mathbf{c}\rangle & \langle\mathbf{b},\mathbf{d}\rangle\end{vmatrix} \quad \forall\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}\in\mathbb{V}^3.$$

Or, we can give formulas for the 'triple cross product' in terms of the box product:

$$(\mathbf{a}\times\mathbf{b})\times(\mathbf{c}\times\mathbf{d})=\mathbf{b}\cdot[\mathbf{a},\mathbf{c},\mathbf{d}]-\mathbf{a}\cdot[\mathbf{b},\mathbf{c},\mathbf{d}]=\mathbf{c}\cdot[\mathbf{a},\mathbf{b},\mathbf{d}]-\mathbf{d}\cdot[\mathbf{a},\mathbf{b},\mathbf{c}] \quad \forall\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}\in\mathbb{V}^3.$$

4.3.4 The $(n-1)$-fold wedge product in $\mathbb{E}^n$

We have already seen in Section 2.4.2 that the $(n-1)$-fold wedge product $\mathbf{v}_1 \wedge_{\mathcal{B}} \mathbf{v}_2 \wedge_{\mathcal{B}} \cdots \wedge_{\mathcal{B}} \mathbf{v}_{n-1}$ calculated in an orthonormal basis, depends on $\mathcal{B}$ only up to sign. Moreover from the proof it follows that this product is independent of the choice of right oriented normal basis. Thus, we may define the $(n-1)$-*fold cross product* to be the $(n-1)$-fold wedge product with respect to a right oriented orthonormal basis. We observe the following analogies to the 3-dimensional case.

For each of the involved vectors $\mathbf{v}_i$, the orthogonal complement $\mathbf{v}_i^{\perp}$ is an $(n-1)$-dimensional vector subspace of $\mathbb{V}^n$. If the vectors $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ are linearly independent, then

$$W = \bigcap_{i=1}^{n-1} \mathbf{v}_i^{\perp} \quad \text{is a 1-dimensional vector subspace of } \mathbb{V}^n.$$

Thus, up to sign, there is a unique unit vector $\mathbf{w} \in W$ which generates W as a vector subspace of $\mathbb{V}^n$. Now, in order to get rid of this 'up-to-sign' ambiguity one may choose $\mathbf{w}$ or $-\mathbf{w}$ according to whether $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n-1}, \mathbf{w})$ is a right-oriented basis or not.

Furthermore, if $\mathcal{B}$ is a right oriented orthonormal basis, then

$$\langle \mathbf{v}_1 \wedge_{\mathcal{B}} \mathbf{v}_2 \wedge_{\mathcal{B}} \cdots \wedge_{\mathcal{B}} \mathbf{v}_{n-1}, \mathbf{v}_n \rangle = [\mathbf{v}_n, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{n-1}].$$

This follows from the definitions and the expressions of the scalar product with respect to an orthonormal coordinate system (See Section 2.4.2 and (3.5)).

4.4 Some applications

4.4.1 Oriented area of a parallelogram

Let $\mathcal{B} = (\mathbf{a}, \mathbf{b})$ be a basis of $\mathbb{V}^2$ and denote by $\mathcal{P}$ a parallelogram spanned by $\mathbf{a}$ and $\mathbf{b}$. We have seen in Section 4.2 that

$$\mathrm{Area}(\mathcal{P}) = \mathrm{Area}(\mathcal{B}) = \sqrt{\det(\mathrm{G}_{\mathcal{B}})} = \left|\det(\mathrm{M}_{\mathcal{E},\mathcal{B}})\right| = \left|[\mathbf{a}, \mathbf{b}]\right|$$

where $\mathcal{E}$ is any orthonormal basis of $\mathbb{V}^2$. The *oriented area* of $\mathcal{P}$ is

$$\mathrm{Area}_{\mathrm{or}}(\mathcal{P}) := \mathrm{Area}_{\mathrm{or}}(\mathcal{B}) = \det(\mathrm{M}_{\mathcal{E},\mathcal{B}}) = [\mathbf{a}, \mathbf{b}].$$

Now consider two vectors $\mathbf{a}(a_1, a_2, a_3)$ and $\mathbf{b}(b_1, b_2, b_3)$ in $\mathbb{V}^3$ and a parallelogram $\mathcal{P} \subseteq \mathbb{E}^3$ such that $\mathbf{a}$ and $\mathbf{b}$ can be represented on adjacent sides of $\mathcal{P}$. Suppose further that the components of the vectors are given with respect to an orthonormal basis $\mathcal{E}$. By the definition of the cross product we have

$$\mathrm{Area}(\mathcal{P}) = |\mathbf{a} \times \mathbf{b}| = \sqrt{\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2}.$$

Then, for a triangle ABC such that $\mathbf{a} = \overrightarrow{CA}$ and $\mathbf{b} = \overrightarrow{CB}$, its area is $\frac{1}{2}|\mathbf{a} \times \mathbf{b}|$. In particular, if the three points A, B, C lie in the Oxy plane, i.e. if $A = A(x_A, y_A, 0)$, $B = B(x_B, y_B, 0)$, $C = C(x_C, y_C, 0)$ then

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_B - x_A & y_B - y_A & 0 \\ x_C - x_A & y_C - y_A & 0 \end{vmatrix} = \begin{vmatrix} x_B - x_A & y_B - y_A \\ x_C - x_A & y_C - y_A \end{vmatrix} \mathbf{k} = \begin{vmatrix} x_A & y_A & 1 \\ x_B & y_B & 1 \\ x_C & y_C & 1 \end{vmatrix} \mathbf{k}.$$

Thus we recover a known formula for the area of the triangle:

$$\text{Area}(\triangle ABC) = \left|\frac{1}{2}\begin{vmatrix} x_A & y_A & 1 \\ x_B & y_B & 1 \\ x_C & y_C & 1 \end{vmatrix}\right| \quad \text{since} \quad |\mathbf{a}\times\mathbf{b}| = \left|\begin{vmatrix} x_A & y_A & 1 \\ x_B & y_B & 1 \\ x_C & y_C & 1 \end{vmatrix}\right|.$$

Notice that, without taking absolute value, the above determinant equals the oriented area $\text{Area}_{\text{or}}(\mathcal{P})$ of the parallelogram $\mathcal{P}$ if $\mathcal{B}$ is right oriented, and it equals minus the oriented area if $\mathcal{B}$ is left oriented.

4.4.2 Oriented angles

Let $\mathbf{a}$ and $\mathbf{b}$ be two vectors in $\mathbb{V}^2$. In Section 3.2.2, we defined the angle $\measuredangle(\mathbf{a},\mathbf{b})$ to be the measure of an angle constructed by representatives of $\mathbf{a}$ and $\mathbf{b}$. Doing so, $\measuredangle(\mathbf{a},\mathbf{b})$ takes values in $[0,\pi]$. For such values the Sine is always positive and it doesn't allow us to recover the trigonometry which we know to hold true in $\mathbb{E}^2$. The reason for this is the ambiguity in orientation. For example, if we say that $\measuredangle(\mathbf{a},\mathbf{b})$ is $2\pi/3$ and we know the position of $\mathbf{a}$, how do we draw $\mathbf{b}$? We have two choices, unless we use a convention that all angles are 'clockwise' or 'counterclockwise'. Formally, such a choice is made by choosing an orientation in $\mathbb{E}^2$ as explained in Section 4.1 (See Definition 4.2).

Definition 4.11. Let $\mathbf{a}$ and $\mathbf{b}$ be two vectors in the oriented plane $\mathbb{E}^2$. We define the *oriented angle of* $\mathbf{a}$ *and* $\mathbf{b}$ to be

$$\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b}) = \begin{cases} \measuredangle(\mathbf{a},\mathbf{b}) & \text{if} \quad [\mathbf{a},\mathbf{b}] \geq 0, \\ -\measuredangle(\mathbf{a},\mathbf{b}) & \text{if} \quad [\mathbf{a},\mathbf{b}] < 0. \end{cases}$$

Recall that $[\mathbf{a},\mathbf{b}] > 0$ if and only if $(\mathbf{a},\mathbf{b})$ is a right oriented basis of $\mathbb{V}^2$. This immediately implies the following properties

Proposition 4.12. Let $\mathbf{a}$ and $\mathbf{b}$ be two vectors in the oriented plane $\mathbb{E}^2$. Then

a) $-\pi < \measuredangle_{\text{or}}(\mathbf{a},\mathbf{b}) \leq \pi$;

b) $0 < \measuredangle_{\text{or}}(\mathbf{a},\mathbf{b}) < \pi$ if and only if $(\mathbf{a},\mathbf{b})$ is a right oriented basis;

c) $\measuredangle_{\text{or}}(\mathbf{a},\mathbf{J}(\mathbf{a})) = \frac{\pi}{2}$;

d) $\measuredangle_{\text{or}}(\mathbf{b},\mathbf{a}) = -\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})$.

Notice also that if θ is an oriented angle then the Sign is no longer positive and we have the sing rules that we are familiar with:

$$\sin(-\theta) = -\sin(\theta) \quad \text{and} \quad \cos(-\theta) = \cos(\theta).$$

Proposition 4.13. If $\mathbf{a}$ and $\mathbf{b}$ are two non-zero vectors in the oriented Euclidean plane $\mathbb{E}^2$, then

$$\cos(\theta) = \frac{\langle \mathbf{a},\mathbf{b}\rangle}{|\mathbf{a}|\cdot|\mathbf{b}|} \quad \text{and} \quad \sin(\theta) = \frac{[\mathbf{a},\mathbf{b}]}{|\mathbf{a}|\cdot|\mathbf{b}|}$$

where $\theta = \measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})$.

Proposition 4.14. If **a** is a non-zero vector in the oriented Euclidean plane $\mathbb{E}^2$, and if $c^2 + s^2 = 1$ then there is a vector **b** such that

$$\cos(\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})) = c \quad \text{and} \quad \sin(\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})) = s$$

and all such vectors are proportional.

Instead of viewing $\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})$ as a value in $(-\pi,\pi]$ it is more convenient to view it as equivalence class of $\mathbb{R}$ modulo 2π, i.e. $\theta \in \mathbb{R}$ represents the angle $\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})$ if $\theta = \measuredangle_{\text{or}}(\mathbf{a},\mathbf{b}) + 2k\pi$ for some integer k. With this convention we have.

Proposition 4.15. If $\mathbf{a},\mathbf{b},\mathbf{c}$ are three non-zero vector in the oriented Euclidean plane $\mathbb{E}^2$, then

$$\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b}) = \measuredangle_{\text{or}}(\mathbf{a},\mathbf{c}) + \measuredangle_{\text{or}}(\mathbf{c},\mathbf{b}) \mod 2\pi.$$

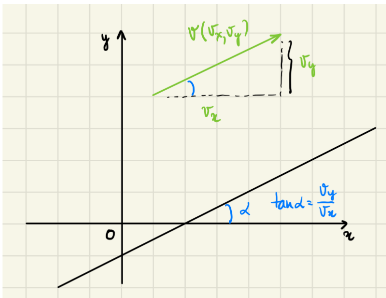
Theorem 4.16. For all $\alpha,\beta \in \mathbb{R}$, we have

$$\begin{aligned} \cos(\alpha+\beta) &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) \\ \sin(\alpha+\beta) &= \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta). \end{aligned}$$

4.4.3 Angular coefficient of a line in $\mathbb{E}^2$

Let $\mathcal{K} = (O,\mathcal{B})$ be a right oriented coordinate system of $\mathbb{E}^2$ with $\mathcal{B} = (\mathbf{i},\mathbf{j})$. Consider a line ℓ passing through a point $A(x_A,y_A)$ in the direction of a vector $\mathbf{v}(v_x,v_y)$. The symmetric equation of $\frac{x-x_A}{v_x} = \frac{y-y_A}{v_y}$ can be rearrange as

$$y = kx + m \quad \text{where} \quad k = \frac{v_y}{v_x} \quad \text{and} \quad m = -\frac{v_y}{v_x}x_A + y_A.$$



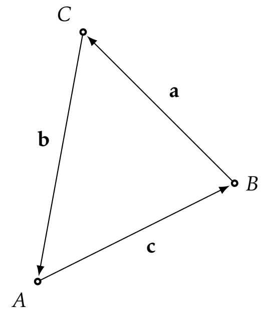
In this form we will call it *the equation where you can read-off the slope k*. Draw a picture to see that if $\alpha = \measuredangle_{\text{or}}(\mathbf{i},\mathbf{v})$ then $k = \tan(\alpha)$. In fact, you can rearrange the symmetric equation as

$$y - y_A = \tan(\alpha)(x - x_A).$$

If you like names, then you can call this *the equation of a line in* $\mathbb{E}^2$ *where you can read off the slope and a point on the line*. There is one subtle point here: the above description is unambiguous only if we consider oriented angles. Indeed, the angle is measured counterclockwise which makes sense only in an oriented Euclidean plane.

4.4.4 Law of Sines

Let ABC be a triangle in $\mathbb{E}^3$ and let $\mathbf{a} = \overrightarrow{BC}$, $\mathbf{b} = \overrightarrow{CA}$ and $\mathbf{c} = \overrightarrow{AC}$.



By the definition of the cross product we have

$$\mathbf{a} \times \mathbf{b} = \mathbf{b} \times \mathbf{c} = \mathbf{c} \times \mathbf{a}$$

in particular the length of these vectors are equal

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{b} \times \mathbf{c}| = |\mathbf{c} \times \mathbf{a}|.$$

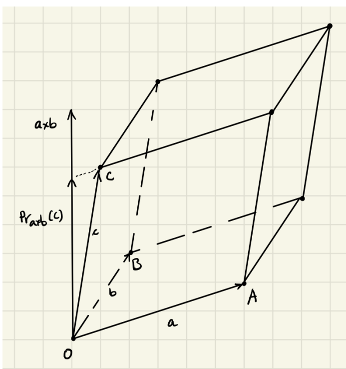
Thus, if we divide by $|\mathbf{a}| \cdot |\mathbf{b}| \cdot |\mathbf{c}|$ we obtain

$$\frac{\sin(\measuredangle(\mathbf{a}, \mathbf{b}))}{|\mathbf{c}|} = \frac{\sin(\measuredangle(\mathbf{b}, \mathbf{c}))}{|\mathbf{a}|} = \frac{\sin(\measuredangle(\mathbf{c}, \mathbf{a}))}{|\mathbf{b}|}$$

and since the interior angles of the triangle ABC are supplementary to the angles involved above we obtain the Law of Sines in a triangle.

4.4.5 Volume of an n-simplex

We first look at what happens in $\mathbb{E}^3$. In this case, a 3-simplex is a tetrahedron. Let $\mathbf{a} = \overrightarrow{OA}, \mathbf{b} = \overrightarrow{OB}, \mathbf{c} = \overrightarrow{OC} \in \mathbb{V}^3$ be non-collinear vectors and let $\mathcal{P}$ be the parallelepiped spanned by the three vectors (i.e. $[OA]$, $[OB]$ and $[OC]$ are sides of the parallelepiped $\mathcal{P}$). Then $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$ is the volume of $\mathcal{P}$ if $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is right oriented; it is minus the volume of $\mathcal{P}$ if $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is left oriented.



Indeed, since
$$[\mathbf{a},\mathbf{b},\mathbf{c}] = \langle \mathbf{a}\times\mathbf{b},\mathbf{c}\rangle$$
we have that
$$\frac{[\mathbf{a},\mathbf{b},\mathbf{c}]}{|\mathbf{a}\times\mathbf{b}|^2}\mathbf{a}\times\mathbf{b} = \frac{\langle \mathbf{a}\times\mathbf{b},\mathbf{c}\rangle}{|\mathbf{a}\times\mathbf{b}|^2}\mathbf{a}\times\mathbf{b}$$
which is the orthogonal projection of $\mathbf{c}$ on $\mathbf{a}\times\mathbf{b}$ (See Section 3.1). Therefore, the length of this vector
$$|\frac{\langle \mathbf{a}\times\mathbf{b},\mathbf{c}\rangle}{|\mathbf{a}\times\mathbf{b}|^2}\mathbf{a}\times\mathbf{b}| = |\frac{[\mathbf{a},\mathbf{b},\mathbf{c}]}{|\mathbf{a}\times\mathbf{b}|}|$$
is a height of the parallelepiped $\mathcal{P}$. It is the distance from C to the face containing the points O, A and B. It follows that
$$\mathrm{Vol}(\mathcal{P}) = \big|[\mathbf{a},\mathbf{b},\mathbf{c}]\big|.$$
Noticing that we can split the above parallelepiped $\mathcal{P}$ in 6 tetrahedra and that they share common base areas and heights. We deduce that
$$\mathrm{Vol}(\mathcal{T}) = \frac{1}{6}\mathrm{Vol}(\mathcal{P}) = \frac{1}{6}\big|[\mathbf{a},\mathbf{b},\mathbf{c}]\big|. \tag{4.4}$$

This line of reasoning can be generalized to n-dimensional space. Let $\mathcal{P}$ be a hyperparallelepiped spanned by the vectors $\mathbf{v}_1,\dots,\mathbf{v}_n$. With respect to an orthonormal basis $\mathcal{B}$, let $\mathbf{w} = \mathbf{v}_1 \wedge_{\mathcal{B}} \mathbf{v}_2 \wedge_{\mathcal{B}} \cdots \wedge_{\mathcal{B}} \mathbf{v}_{n-1}$. Consider the orthogonal projection of $\mathbf{v}_n$ on $\mathbf{w}$
$$\frac{[\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n]}{|\mathbf{w}|^2}\mathbf{w} = \frac{\langle \mathbf{w},\mathbf{v}_n\rangle}{|\mathbf{w}|^2}\mathbf{w}.$$
The length of the above vector is a height of the hyperparallelepiped $\mathcal{P}$, it is the distance between opposite $(n-1)$-dimensional faces which are parallel to the vectors $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$. Thus
$$\mathrm{Vol}(\mathcal{P}) = |\frac{[\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n]}{|\mathbf{w}|}|\cdot \mathrm{Vol}(\mathcal{F})$$
where $\mathcal{F}$ is the hyperparallelepiped spanned by $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$. To calculate $|\mathbf{w}|$ we may choose an orthonormal bases $\mathcal{E} = (\mathbf{e}_1,\dots,\mathbf{e}_n)$ such that $\mathbf{e}_n$ is orthogonal to the vectors $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$. Then
$$\mathbf{w} = \mathbf{v}_1 \wedge_{\mathcal{B}} \mathbf{v}_2 \wedge_{\mathcal{B}} \cdots \wedge_{\mathcal{B}} \mathbf{v}_{n-1} = \pm \mathbf{v}_1 \wedge_{\mathcal{E}} \mathbf{v}_2 \wedge_{\mathcal{E}} \cdots \wedge_{\mathcal{E}} \mathbf{v}_{n-1} = \pm\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \\ \leftarrow & \mathbf{v}'_1 & \rightarrow & 0 \\ & \vdots & & 0 \\ \leftarrow & \mathbf{v}'_{n-1} & \rightarrow & 0 \end{vmatrix} = \pm\begin{vmatrix} \leftarrow & \mathbf{v}'_1 & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{v}'_{n-1} & \rightarrow \end{vmatrix}\mathbf{e}_n$$
where $\mathbf{v}'_i$ is the row matrix obtained from the components of $\mathbf{v}_i$ where we remove the last component which is zero. Thus $|\mathbf{w}| = \big|[\mathbf{v}'_1,\dots,\mathbf{v}'_{n-1}]\big|$ which inductively equals the volume $\mathrm{Vol}(\mathcal{F})$ of the hyperparallelepiped $\mathcal{F}$ of dimension $n-1$.

The n-simplex spanned by the vectors $\mathbf{v}_1,\dots,\mathbf{v}_n$ at the origin is the set
$$\mathcal{S} = \{P \in \mathbb{E}^n : \overrightarrow{OP} = t_1\mathbf{v}_1 + \cdots + t_n\mathbf{v}_n : \text{ with } t_1,\dots,t_n \ge 0 \text{ and } t_1 + \cdots + t_n \le 1\}.$$

It is possible to show that the hyperparallelepiped $\mathcal{P}$ can be divided into $n!$ simplices which have the same volume, i.e. that

$$\mathrm{Vol}(\mathcal{S}) = \frac{1}{n!}\mathrm{Vol}(\mathcal{P}) = \frac{1}{n!}\big|[\mathbf{v}_1,\dots,\mathbf{v}_n]\big|.$$

Compare this to (4.4) and to Section 4.4.2.

4.4.6 Distance from a point to a hyperplane

Let $\mathcal{K} = (O,\mathcal{B})$ be an orthonormal coordinate system and consider a hyperplane $\mathcal{H}$ given by the equation

$$\mathcal{H} : a_1x_1 + a_2x_2 + \cdots + a_nx_n + a_{n+1} = 0$$

with respect to $\mathcal{K}$. If we choose $n-1$ linearly independent vectors $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$ which are parallel to $\mathcal{H}$, we may assume that

$$\mathbf{v}_1 \wedge_{\mathcal{B}} \mathbf{v}_2 \wedge_{\mathcal{B}} \cdots \wedge_{\mathcal{B}} \mathbf{v}_{n-1} = (a_1,\dots,a_n). \tag{4.5}$$

Denote the above normal vector of $\mathcal{H}$ by $\mathbf{w}$. Fix a point $Q \in \mathcal{H}$ and an arbitrary point $P \in \mathbb{E}^n$. Let $\mathcal{P}$ be a n-dimensional hyperparallelepiped spanned by $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}, \overrightarrow{QP}$ in O and let $\mathcal{F}$ be the $(n-1)$-dimensional hyperparallelepiped spanned by $\mathbf{v}_1,\dots,\mathbf{v}_{n-1}$ in O. Notice that $\mathcal{F}$ is a facet of $\mathcal{P}$ and that

$$\Big|\frac{\langle \mathbf{w}, \overrightarrow{QP}\rangle}{|\mathbf{w}|}\Big| = \frac{\mathrm{Vol}(\mathcal{P})}{\mathrm{Vol}(\mathcal{F})} \tag{4.6}$$

is the length of the orthogonal projection of $\overrightarrow{QP}$ on $\mathbf{w}$, i.e. it is the height of $\mathcal{P}$ corresponding to the facet $\mathcal{F}$ (Compare with the discussion in Section 4.4.5). Thus, the distance from the point P to the hyperplane $\mathcal{H}$ is

$$d(P,\mathcal{H}) = \Big|\frac{[\mathbf{v}_1,\dots,\mathbf{v}_{n-1},\overrightarrow{QP}]}{\mathrm{Vol}(\mathcal{F})}\Big| = \Big|\frac{[\mathbf{v}_1,\dots,\mathbf{v}_{n-1},\overrightarrow{OP}] - [\mathbf{v}_1,\dots,\mathbf{v}_{n-1},\overrightarrow{OQ}]}{\mathrm{Vol}(\mathcal{F})}\Big| = \Big|\frac{\langle \mathbf{w}, \overrightarrow{OP}\rangle - \langle \mathbf{w}, \overrightarrow{OQ}\rangle}{|\mathbf{w}|}\Big|.$$

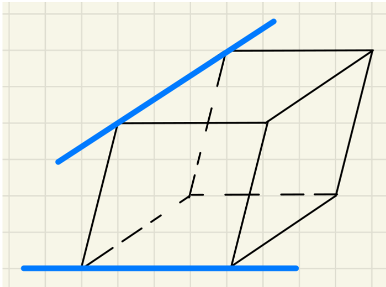
Since Q belongs to $\mathcal{H}$ and by (4.5) we obtain (again) the formula deduced in Section 3.3.5:

$$d(P,\mathcal{H}) = \frac{|a_1p_1 + a_2p_2 + \cdots + a_np_n + a_{n+1}|}{\sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}}. \tag{4.7}$$

4.4.7 Distance between lines in $\mathbb{E}^3$

Considering the distances between two lines ℓ_1 and ℓ_2 in $\mathbb{E}^3$, we have the following cases.
(Case 1) If ℓ_1 and ℓ_2 intersect then the distance between them is zero: $d(\ell_1,\ell_2) = 0$.
(Case 2) If ℓ_1 and ℓ_2 are skew relative to each other, there is a unique plane π_1 containing ℓ_1 which is parallel to ℓ_2 and there is a unique plane π_2 which contains ℓ_2 and is parallel to ℓ_1, thus

$$d(\ell_1,\ell_2) = d(\pi_1,P_2) = d(P_1,\pi_2) \quad \forall P_1 \in \ell_1 \quad \text{and} \quad \forall P_2 \in \ell_2.$$



More concretely, if the two lines are

$$\ell_1 : \begin{cases} x = x_1 + tv_x \\ y = y_1 + tv_y \\ z = z_1 + tv_z \end{cases} \quad \text{and} \quad \ell_2 : \begin{cases} x = x_2 + tu_x \\ y = y_2 + tu_y \\ z = z_2 + tu_z \end{cases}$$

with direction vectors $\mathbf{v}(v_1, v_2, v_3)$ and $\mathbf{u}(u_1, u_2, u_3)$ respectively, then

$$\pi_1 : \begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ v_x & v_y & v_z \\ u_x & u_y & u_z \end{vmatrix} = 0.$$

Hence, if we let $\mathbf{a}(a_x, a_y, a_z) = \mathbf{v} \times \mathbf{u}$, from (4.7), we obtain

$$d(\ell_1, \ell_2) = d(\pi_1, P_2) = \frac{|a_x(x_2 - x_1) + a_y(y_2 - y_1) + a_z(z_2 - z_1)|}{\sqrt{a_x^2 + a_y^2 + a_z^2}}.$$

(Case 3) If ℓ_1 and ℓ_2 do not intersect and they are parallel, then

$$d(\ell_1, \ell_2) = d(P_1, \ell_2) = d(\ell_1, P_2) \quad \forall P_1 \in \ell_1 \quad \text{and} \quad \forall P_2 \in \ell_2.$$

However, here we are dealing with lines in $\mathbb{E}^3$ which are not hyperplanes. So we cannot directly apply (4.7). Nevertheless, the idea of deducing a formula also in this last case, is similar to the one used for the formula (4.7). We formulate it as follows.

Proposition 4.17. Suppose you have a line

$$\ell_1 : \begin{cases} x = x_A + tv_x \\ y = y_A + tv_y \\ z = z_A + tv_z \end{cases} \quad \text{and a point} \quad P(x_P, y_P, z_P) \in \mathbb{E}^3.$$

The distance from P to ℓ is

$$d(P, \ell) = \frac{|\overrightarrow{PA} \times \mathbf{v}|}{|\mathbf{v}|}.$$

4.4.8 Common perpendicular line of two skew lines in $\mathbb{E}^3$

Consider two skew lines ℓ_1 and ℓ_2 in $\mathbb{E}^3$ given via parametric equations as follows:

$$\ell_1 : \begin{cases} x = x_1 + tv_x \\ y = y_1 + tv_y \\ z = z_1 + tv_z \end{cases} \quad \text{and} \quad \ell_2 : \begin{cases} x = x_2 + tu_x \\ y = y_2 + tu_y \\ z = z_2 + tu_z \end{cases}.$$

Denote the direction vectors by $\mathbf{v} = \mathbf{v}(v_1, v_2, v_3)$ and $\mathbf{u}(u_1, u_2, u_3)$ respectively. Since ℓ_1 and ℓ_2 are skew relative to each other, the vector $\mathbf{a} = \mathbf{v} \times \mathbf{u}$ is non-zero and perpendicular to both $\mathbf{v}$ and $\mathbf{u}$. The following line is called the *common perpendicular* of ℓ_1 and ℓ_2:

$$\ell : \begin{cases} \pi_1' : \begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ v_x & v_y & v_z \\ a_x & a_y & a_z \end{vmatrix} = 0, \\ \pi_2' : \begin{vmatrix} x - x_2 & y - y_2 & z - z_2 \\ u_x & u_y & u_z \\ a_x & a_y & a_z \end{vmatrix} = 0. \end{cases}$$

It is the unique line satisfying the following properties (it is a calculation to check them):

- $\ell \perp \ell_1$ and $\ell \perp \ell_2$, and
- $\ell \cap \ell_1 \neq \emptyset$ and $\ell \cap \ell_2 \neq \emptyset$.

It is possible to deduce the equations of the common perpendicular line without using the cross product, the scalar product is sufficient. However, the discussion is shorter if we use $\mathbf{a} = \mathbf{v} \times \mathbf{u}$.

4.5 Exercises

4.1. Let $(\mathbf{i},\mathbf{j},\mathbf{k})$ be a right oriented orthonormal basis of $\mathbb{V}^3$. Consider the vectors $\mathbf{a} = \mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$ and $\mathbf{b} = 7\mathbf{i} + 4\mathbf{j} + 6\mathbf{k}$. Determine $\mathbf{a} \times \mathbf{b}$ in terms of the given basis vectors.

4.2. With respect to a right oriented orthonormal basis of $\mathbb{V}^3$ consider the vectors $\mathbf{a}(3,-1,-2)$ and $\mathbf{b}(1,2,-1)$. Calculate

$$\mathbf{a} \times \mathbf{b}, \quad (2\mathbf{a} + \mathbf{b}) \times \mathbf{b}, \quad (2\mathbf{a} + \mathbf{b}) \times (2\mathbf{a} - \mathbf{b}).$$

4.3. Determine the distances between opposite sides of a parallelogram spanned by the vectors $\overrightarrow{AB}(6,0,1)$ and $\overrightarrow{AC}(1.5,2,1)$ if the coordinates of the vectors are given with respect to a right oriented orthonormal basis.

4.4. Consider the vectors $\mathbf{a}(2,3,-1)$ and $\mathbf{b}(1,-1,3)$ with respect to an orthonormal basis.

a) Determine the vector subspace $\langle \mathbf{a}, \mathbf{b} \rangle^{\perp}$.

b) Determine the vector $\mathbf{p}$ which is orthogonal to $\mathbf{a}$ and $\mathbf{b}$ and for which $\mathbf{p} \cdot (2\mathbf{i} - 3\mathbf{j} + 4\mathbf{k}) = 51$.

4.5. Consider the points $A(1,2,0)$, $B(3,0,-3)$ and $C(5,2,6)$ with respect to an orthonormal coordinate system.

a) Determine the area of the triangle ABC.

b) Determine the distance from C to AB.

4.6. Let $ABCD$ be a quadrilateral in $\mathbb{E}^3$ and let E, F be the midpoints of $[AB]$ and $[CD]$ respectively. Denote by K, L, M and N the midpoints of the segments $[AF]$, $[CE]$, $[BF]$ and $[DE]$ respectively. Prove that $KLMN$ is a parallelogram.

4.7. Let ABC be a triangle and let $\mathbf{u} = \overrightarrow{AB}$, $\mathbf{v} = \overrightarrow{BC}$, $\mathbf{w} = \overrightarrow{CA}$. Show that

$$\mathbf{u} \times \mathbf{v} = \mathbf{v} \times \mathbf{w} = \mathbf{w} \times \mathbf{u}.$$

and deduce the law of sines in a triangle.

4.8. With respect to a right oriented orthonormal coordinate system consider the vectors $\mathbf{a}(2,-3,1)$, $\mathbf{b}(-3,1,2)$ and $\mathbf{c}(1,2,3)$. Calculate $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$ and $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$.

4.9. Fix $\mathbf{v} \in \mathbb{V}^3$ and let $\psi : \mathbb{V}^3 \to \mathbb{V}^3$ be the map $\phi(\mathbf{w}) = \mathbf{v} \times \mathbf{w}$. Is the map linear? Explain why. Give the matrix of ϕ relative to a right oriented orthonormal basis. What changes if we define ϕ by $\phi(\mathbf{w}) = \mathbf{w} \times \mathbf{v}$?

4.10. Let $(\mathbf{i},\mathbf{j},\mathbf{k})$ be a right oriented orthonormal basis. Determine the matrices of the linear maps $\phi, \psi : \mathbb{V}^3 \to \mathbb{V}^3$ defined by $\phi(\mathbf{v}) = \mathbf{w} \times \mathbf{v}$ and $\psi(\mathbf{v}) = \mathbf{v} \times \mathbf{u}$ where

a) $\mathbf{w} = -\mathbf{i} + 3\mathbf{j} + \mathbf{k}$,

b) $\mathbf{w} = \mathbf{i} + \mathbf{k}$,

c) $\mathbf{u} = 2\mathbf{i} - \mathbf{j}$,

d) $\mathbf{u} = \mathbf{i} + \mathbf{j} + \mathbf{k}$.

4.11. Prove the following identities:

a) the Grassmann identity,

b) the Jacobi identity,

c) the Lagrange identity,

d) the formula for the cross product of two cross products.

4.12. Let $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ be a right oriented orthonormal basis. Consider the vectors $\mathbf{a} = \mathbf{i} + \mathbf{j}$, $\mathbf{b} = \mathbf{i} - \mathbf{k}$ and $\mathbf{c} = \mathbf{k}$. Determine if

a) $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is a basis of $\mathbb{V}^3$,

b) if it is a basis, decide if it is left or right oriented.

4.13. The points $A(1,2,-1)$, $B(0,1,5)$, $C(-1,2,1)$ and $D(2,1,3)$ are given with respect to an orthonormal coordinate system. Are the four points coplanar?

4.14. Let $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ be an orthonormal basis and consider the vectors $\mathbf{u} = -\mathbf{i} + 3\mathbf{j} + \mathbf{k}$ and $\mathbf{w} = \mathbf{k}$. Determine the matrix of the linear map $\phi : \mathbb{V}^3 \to \mathbb{R}$ defined by $\phi(\mathbf{v}) = [\mathbf{v}, \mathbf{u}, \mathbf{w}]$.

4.15. Determine the volume of the tetrahedron with vertices $A(2,-1,1)$, $B(5,5,4)$, $C(3,2,-1)$ and $D(4,1,3)$ given with respect to an orthonormal system.

4.16. The volume of a tetrahedron $ABCD$ is 5. With respect to an orthonormal system $Oxyz$ the vertices are $A(2,1,-1)$, $B(3,0,1)$, $C(2,-1,3)$ and $D \in Oy$. Determine the coordinates of D.

4.17. With respect to an orthonormal system consider the vectors $\mathbf{a}(8,4,1)$, $\mathbf{b}(2,2,1)$ and $\mathbf{c}(1,1,1)$. Determine a vector $\mathbf{d}$ satisfying the following properties

a) the angles $\measuredangle(\mathbf{d}, \mathbf{a})$ and $\measuredangle(\mathbf{d}, \mathbf{b})$ are equal,

b) $\mathbf{d}$ is orthogonal to $\mathbf{c}$,

c) $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ and $(\mathbf{a}, \mathbf{b}, \mathbf{d})$ have the same orientation.

4.18. For the tetrahedron $ABCD$ in Exercise 4.15, determine the common perpendicular of the sides AB and CD.

4.19. Consider two lines ℓ_1 and ℓ_2 in $\mathbb{E}^3$. Suppose that the common perpendicular line is

$$\ell : \begin{cases} x = 1 + t \\ y = 2 - t \\ z = t \end{cases},$$

that $P_1(1,0,1) \in \ell_1$ and that $P_2(-1,1,0) \in \ell_2$. Determine the two lines.

4.20. In $\mathbb{E}^4$ consider the affine subspaces

$$\ell : \begin{cases} w = -1 + 2t \\ x = 1 + t \\ y = 2 - t \\ z = -2 + t \end{cases} \quad \text{and} \quad \pi : \begin{cases} w + 2y - 1 = 0 \\ x - z + 2 = 0 \end{cases}.$$

Show that $\ell \parallel \pi$ and determine the distance between them.

4.21. In $\mathbb{E}^4$ consider the parallel lines

$$\ell_1 : \begin{cases} w = -1 + 2t \\ x = 1 + t \\ y = 2 - t \\ z = -2 + t \end{cases} \quad \text{and} \quad \begin{cases} w = 1 + 2t \\ x = t \\ y = 1 - t \\ z = 2 + t \end{cases}.$$

Determine the distance between them.

CHAPTER 5

Affine morphisms and isometries

Contents

5.1 Affine morphisms

A morphism is a general notion in mathematics that refers to maps which preserve certain structures. An affine morphism is a map which preserves the structure of affine spaces. In Proposition 5.4 we give a very concrete description of what an affine morphism is. If you are less inclined to abstract aspects, you can take the description in Proposition 5.4 to be the definition of an affine morphisms.

Affine spaces were introduced in Section 1.2.2. The n-dimensional affine space is a triple $(\mathbb{A}^n, \mathbb{V}^n, t)$ where $\mathbb{A}^n$ is a set, $\mathbb{V}^n$ is an n-dimensional vector space and $t: \mathbb{V}^n \times \mathbb{A}^n \to \mathbb{A}^n$ is the structure map which satisfies (AS1) and (AS2) in Definition 1.14. For another affine space $(\mathbb{A}^m, \mathbb{V}^m, \tau)$ of dimension m, the structure map is $\tau: \mathbb{V}^m \times \mathbb{A}^m \to \mathbb{A}^m$.

For the moment, consider any map $\phi: \mathbb{A}^n \to \mathbb{A}^m$. Fix a point $Q \in \mathbb{A}^n$. By (AS1), for any $P \in \mathbb{A}^n$ there is a unique vector $\mathbf{v} \in \mathbb{V}^n$ and a unique vector $\mathbf{w} \in \mathbb{V}^m$ such that

$$P = t(\mathbf{v}, Q) \quad \text{and} \quad \phi(P) = \tau(\mathbf{w}, \phi(Q)). \tag{5.1}$$

In other words $\mathbf{v} = \overrightarrow{QP}$ and $\mathbf{w} = \overrightarrow{\phi(Q)\phi(P)}$. This defines a map which we denote by $\mathrm{lin}_Q(\phi)$:

$$\mathrm{lin}_Q(\phi): \mathbb{V}^n \to \mathbb{V}^m \quad \text{defined by} \quad \mathrm{lin}_Q(\phi)(\overrightarrow{QP}) = \overrightarrow{\phi(Q)\phi(P)}.$$

Notice that, by definition we have

$$\phi(t(\mathbf{v}, Q)) = \tau(\mathrm{lin}_Q(\phi)(\mathbf{v}), \phi(Q)) \quad \forall \mathbf{v} \in \mathbb{V}^n \text{ and } \forall Q \in \mathbb{A}^n.$$

Definition 5.1. For a map $\phi: \mathbb{A}^n \to \mathbb{A}^m$ and a point $Q \in \mathbb{A}^n$, we say that ϕ *induces the map* $\mathrm{lin}_Q(\phi)$ *on vectors relative to the point* Q or that $\mathrm{lin}(\phi)$ *is the induced map on vectors relative to* Q. Moreover we say that ϕ *is compatible with the structure maps* if for any $\mathbf{v}, \mathbf{w} \in \mathbb{V}^n$ and any $Q \in \mathbb{A}^n$

$$\phi(t(\mathbf{v}, t(\mathbf{w}, Q))) = \tau(\mathrm{lin}_{Q'}(\phi)(\mathbf{v}), \tau(\mathrm{lin}_Q(\mathbf{w}), \phi(Q))) \tag{5.2}$$

where $Q' \in \mathbb{A}^n$ is any point such that $\phi(Q') = \tau(\mathrm{lin}_Q(\mathbf{w}), \phi(Q))$. Loosely speaking, this is saying that 'ϕ preserves the axiom (AS2)'.

Since an affine space is a triple $(\mathbb{A}^n, \mathbb{V}^n, t)$, for a map $\phi: \mathbb{A}^n \to \mathbb{A}^m$ to be a morphism, i.e. to preserve the structure of the affine spaces, it needs to preserve the vector space structure of the set of vectors as well as the structure maps. Preserving the vector space structure means that the induced map on vectors is linear relative to each Q. This condition implies that ϕ preserves the structure maps:

Proposition 5.2. Let $\phi: \mathbb{A}^n \to \mathbb{A}^m$ be a map. Assume that for any $Q \in \mathbb{A}^n$ the induced map on vectors $\mathrm{lin}_Q(\phi)$ is linear. Then

1. ϕ maps lines onto lines or points;
2. ϕ maps parallelograms to (possibly degenerate) parallelograms;
3. $\mathrm{lin}_Q(\phi)$ does not depend on the point Q, i.e. $\mathrm{lin}_Q(\phi) = \mathrm{lin}_{Q'}(\phi)$ for any other point Q';
4. ϕ preserves the structure maps.

Definition 5.3. A map $\phi : \mathbb{A}^n \to \mathbb{A}^m$ is called an *affine morphism* if $\mathrm{lin}_Q(\phi) : \mathbb{V}^n \to \mathbb{V}^m$ is linear for each $Q \in \mathbb{A}^n$. If this is the case then by Proposition 5.2, the map $\mathrm{lin}_Q(\phi)$ does not depend on Q, thus we denote it by $\mathrm{lin}(\phi)$ and call it *the linear map induced on vectors by the affine morphism* ϕ.

The consequence of the linearity of the induced map on vectors is the following. Fix a reference frame $\mathcal{K} = (O, \mathcal{B})$ in $\mathbb{A}^n$ and a reference frame $\mathcal{K}' = (O', \mathcal{B}')$ in $\mathbb{A}^m$. Then

$$\mathrm{lin}(\phi)(\overrightarrow{OP}) = \overrightarrow{\phi(O)\phi(P)} = \overrightarrow{O'\phi(P)} - \overrightarrow{O'\phi(O)}$$

and thus, in coordinates we have for any point P

$$[\phi(P)]_{\mathcal{K}'} = \underbrace{\mathrm{M}_{\mathcal{B}',\mathcal{B}}(\mathrm{lin}(\phi))}_{A} \cdot [P]_{\mathcal{K}} + \underbrace{[\phi(O)]_{\mathcal{K}'}}_{b}.$$

Since $\mathcal{K}$ and $\mathcal{K}'$ are fixed, in particular O and O' are fixed, we see that b and A do not depend on P. Clearly A is just the matrix of the linear map $\mathrm{lin}(\phi)$ with respect to the two indicated bases.

Proposition 5.4. A map $\phi : \mathbb{A}^n \to \mathbb{A}^m$ is an affine morphism if and only if there exists $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ and $b \in \mathrm{Mat}_{m\times 1}(\mathbb{R})$ such that

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

relative to some reference frame $\mathcal{K}$ of $\mathbb{A}^n$ and some reference frame $\mathcal{K}'$ of $\mathbb{A}^m$.

In (5.3) both A and b depend on the choice of the coordinate systems $\mathcal{K}$ and $\mathcal{K}'$ and moreover A is the matrix of $\mathrm{lin}(\phi)$ relative to the bases in $\mathcal{K}$ and $\mathcal{K}'$.

If $n = m$, then $\phi : \mathbb{A}^n \to \mathbb{A}^n$ is called an *affine endomorphisms*. The set of all such endomorphisms is denoted by $\mathrm{End}_{\mathrm{aff}}(\mathbb{A}^n)$. It is easy to see that, the morphism ϕ is invertible if and only if the map $\mathrm{lin}(\phi)$ is invertible, equivalently, ϕ is invertible if and only if the the matrix A of the linear map $\mathrm{lin}(\phi)$ is invertible. An invertible affine endomorphism ϕ is called an *affine automorphism* or *affine transformation*. The set of all affine transformations of $\mathbb{A}^n$ is denoted by $\mathrm{AGL}(\mathbb{A}^n)$

If in addition we have $O = O'$ and $b = 0$ in (5.3) then ϕ can be viewed as a linear map from $\mathbb{V}^n$ to $\mathbb{V}^n$ since it is given by multiplication with the matrix A. The set of linear maps $\mathbb{V}^n \to \mathbb{V}^n$ is denoted by $\mathrm{GL}(\mathbb{V}^n)$ and we have the following inclusion

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

Example 5.5. A homothety $\phi_{C,\lambda}$ of $\mathbb{E}^n$ with center C is the map which rescales the space with a factor λ along lines passing through the point C. With respect to a coordinate system $\mathcal{K} = (C, \mathcal{B})$ with origin in C it has the form

$$[\phi_{C,\lambda}(P)]_{\mathcal{K}} = \lambda \cdot [P]_{\mathcal{K}}.$$

Notice that if $\lambda = 1$ this is the identity map, if $\lambda < 1$ this is a contraction, if $\lambda > 1$ it is an expansion.

Example 5.6. The various parallel projections and reflections described in the next subsections are examples of affine morphisms as well as isometries discussed in Section 5.2. For other examples see [7, Chapter 5].

Proposition 5.7. Let $\phi : \mathbb{E}^n \to \mathbb{E}^m$ be an affine morphism. If a line ℓ is mapped onto a line ℓ' under ϕ, then ϕ preserves the ratio on these lines, i.e. if A and B are distinct points on ℓ, then for any other point C on the line ℓ we have

$$\frac{AC}{AB} = \frac{\phi(A)\phi(C)}{\phi(A)\phi(B)}. \tag{5.4}$$

Proposition 5.8. A map $\phi : \mathbb{E}^n \to \mathbb{E}^n$ is an affine transformation if and only if

1. ϕ is injective;
2. ϕ preserves lines;
3. ϕ preserves the ratio on lines.

5.1.1 Homogeneous coordinates and homogeneous matrix

Conceptually, homogeneous coordinates are used to describe the affine space $\mathbb{A}^n$ inside the projective space $\mathbb{P}^n$. This is outside the scope of these notes. However, there is also a computational advantage to homogeneous coordinates which is what we describe here.

Definition 5.9. Let $\mathcal{K}$ be a reference frame of $\mathbb{A}^n$. The *homogeneous coordinates* of the point $P(p_1,\dots,p_n)$ are $(p_1,\dots,p_n,1)$. Yes, they are the ordinary coordinates with an extra 1 at the end.

Definition 5.10. The *homogeneous matrix* of an affine morphism $\phi : \mathbb{A}^n \to \mathbb{A}^m$ defined with respect to some reference frames $\mathcal{K}$ and $\mathcal{K}'$ by $\phi(\mathbf{x}) = A\mathbf{x} + b$ is

$$\hat{\mathrm{M}}_{\mathcal{K},\mathcal{K}}(\phi) = \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix}.$$

The utility of introducing these notions is the following: composition of affine maps corresponds to multiplication of homogeneous matrices. Indeed let $\phi(\mathbf{x}) = A'\mathbf{x} + b'$ be another affine map defined on $\mathbb{A}^m$. Then

$$\psi \circ \phi(\mathbf{x}) = \psi(\phi(\mathbf{x})) = \psi(A\mathbf{x} + b) = A'(A\mathbf{x} + b) + b' = (A' \cdot A)\mathbf{x} + (A' \cdot b + b').$$

In terms of homogeneous matrices, we notice that

$$\hat{\mathrm{M}}_{\mathcal{K}'',\mathcal{K}}(\psi \circ \phi) = \begin{bmatrix} A' \cdot A & A' \cdot b + b' \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A' & b' \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} = \hat{\mathrm{M}}_{\mathcal{K}'',\mathcal{K}'}(\psi) \cdot \hat{\mathrm{M}}_{\mathcal{K}',\mathcal{K}}(\phi)$$

and the homogeneous coordinates of the values of the map ϕ can be obtained through a matrix multiplication as well

$$\begin{bmatrix} \phi(\mathbf{x}) \\ 1 \end{bmatrix} = \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}.$$

Thus

$$\begin{bmatrix} \psi \circ \phi(\mathbf{x}) \\ 1 \end{bmatrix} = \begin{bmatrix} A' & b' \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}.$$

5.1.2 Tensor product

Tensor products are widely used throughout Mathematics. We introduce this notion here to give a compact expressions for the projections and reflections discussed in the following subsections.

Definition 5.11. Let $\mathbf{v}(v_1,\dots,v_n)$ and $\mathbf{w}(w_1,\dots,w_n)$ be two vectors with components relative to a basis $\mathcal{B}$. The *tensor product* $\mathbf{v}\otimes_{\mathcal{B}}\mathbf{w}$ is the $n\times n$ matrix defined by $(\mathbf{v}\otimes_{\mathcal{B}}\mathbf{w})_{ij}=v_iw_j$. In other words

$$\mathbf{v}\otimes_{\mathcal{B}}\mathbf{w}=\mathbf{v}\cdot\mathbf{w}^T=\begin{bmatrix}v_1\\ \vdots\\ v_n\end{bmatrix}\cdot\begin{bmatrix}w_1\dots w_n\end{bmatrix}=\begin{bmatrix}v_1w_1 & \dots & v_1w_n\\ \vdots & & \vdots\\ v_nw_1 & \dots & v_nw_n\end{bmatrix}.$$

We write $\otimes$ instead of $\otimes_{\mathcal{B}}$ if it is clear from the context what $\mathcal{B}$ is.

Proposition 5.12. The map $\_\otimes_{\mathcal{B}}\_:\mathbb{R}^n\times\mathbb{R}^n\to\mathrm{Mat}_{n\times n}(\mathbb{R})$ given by $(\mathbf{v},\mathbf{w})\mapsto\mathbf{v}\otimes_{\mathcal{B}}\mathbf{w}$ has the following properties:

1. It is linear in both arguments,
2. $(\mathbf{v}\otimes_{\mathcal{B}}\mathbf{w})^T=\mathbf{w}\otimes_{\mathcal{B}}\mathbf{v}$.
3. If $\mathcal{B}$ is orthonormal then $(\mathbf{u}\otimes_{\mathcal{B}}\mathbf{v})\cdot\mathbf{w}=\langle\mathbf{v},\mathbf{w}\rangle\mathbf{u}$ for any $\mathbf{u},\mathbf{v},\mathbf{w}\in\mathbb{V}^n$.

5.1.3 Parallel projection on a hyperplane

Example 5.13. Fix a reference frame $\mathcal{K}$ of $\mathbb{A}^2$. We want to project on the line (hyperplane)

$$\ell: x+y-1=0$$

in the direction of the vector $\mathbf{v}(-2,-1)$. How do we do this? We do it pointwise. Take an arbitrary point $P(x_P,y_P)$ and consider the line ℓ_P passing through P in the direction of $\mathbf{v}$. It has parametric equations

$$\ell_P:\begin{cases} x=x_P-2t\\ y=y_P-t.\end{cases}$$

The projection of P on ℓ in the direction of $\mathbf{v}$ is the point $P'=\ell\cap\ell_P$.

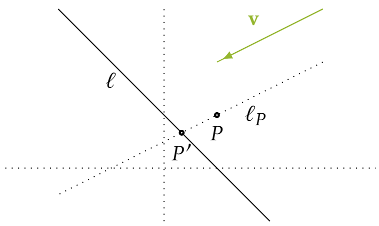


Figure 5.1: Projection of the point P on the line ℓ in the direction of the vector $\mathbf{v}$.

To determine P', we check which point on ℓ_P satisfies the equation of ℓ

$$x_P - 2t + y_P - t - 1 = 0 \quad \Rightarrow \quad t = \frac{1}{3}(x_P + y_P - 1).$$

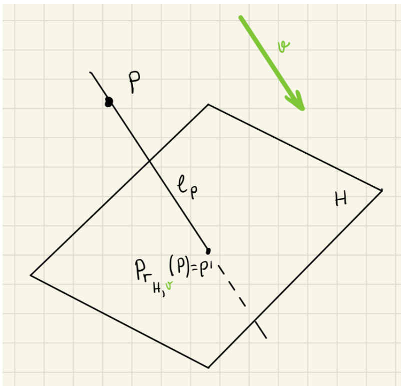
Thus, the projection of P is

$$[P']_{\mathcal{K}} = \begin{bmatrix} \frac{1}{3}x_P - \frac{2}{3}y_p + \frac{2}{3} \\ -\frac{1}{3}x_P + \frac{2}{3}y_p + \frac{1}{3} \end{bmatrix} = \underbrace{\frac{1}{3}\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}}_{A} \begin{bmatrix} x_P \\ y_P \end{bmatrix} + \underbrace{\frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{b}.$$

Definition 5.14. 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

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

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



Fix a reference frame $\mathcal{K} = (O, \mathcal{B})$ of $\mathbb{A}^n$. Consider the hyperplane

$$H : a_1x_1 + \cdots + a_nx_n + a_{n+1} = 0 \tag{5.5}$$

and a line ℓ_P passing through the point $P(p_1, \ldots, p_n)$ and having $\mathbf{v}(v_1, \ldots, v_n)$ as direction vector:

$$\ell_P = \{P + t\mathbf{v} : t \in \mathbb{R}\}. \tag{5.6}$$

The intersection $\ell_P \cap H$ can be described as follows

$$P + t\mathbf{v} \in \ell \cap H \quad \Leftrightarrow \quad a_1(p_1 + tv_1) + \cdots + a_n(p_n + tv_n) + a_{n+1} = 0.$$

So, the intersection point $\mathrm{Pr}_{H,\mathbf{v}}(P) = P'$ is

$$P' = P - \frac{a_1p_1 + \cdots + a_np_n + a_{n+1}}{a_1v_1 + \cdots + a_nv_n}\mathbf{v} = P - \frac{\mathbf{a}^T \cdot P + a_{n+1}}{\mathbf{a}^T \cdot \mathbf{v}}\mathbf{v}. \tag{5.7}$$

where $\mathbf{a} = \mathbf{a}(a_1, \dots, a_n)$ and where in the second equality we use the convention that points and vectors are identified with column matrices of their coordinates and components respectively. Hence, if we denote by $p'_1, \dots, p'_n$ the coordinates of the projected point $\mathrm{Pr}_{H,\mathbf{v}}(P)$ then

$$\begin{cases} p'_1 = p_1 + \mu v_1 \\ \quad\vdots \\ p'_n = p_n + \mu v_n \end{cases} \quad \text{where} \quad \mu = -\frac{a_1 p_1 + \cdots + a_n p_n + a_{n+1}}{a_1 v_1 + \cdots + a_n v_n}.$$

Since $\mathbf{a}^T \cdot P \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{a}^T \cdot P$, in matrix form we have

$$\mathrm{Pr}_{H,\mathbf{v}}(P) = \left(I_n - \frac{\mathbf{v} \cdot \mathbf{a}^T}{\mathbf{v}^T \cdot \mathbf{a}} \right) \cdot P - \frac{a_{n+1}}{\mathbf{v}^T \cdot \mathbf{a}} \mathbf{v}$$

where I_n is the $n \times n$-identity matrix. In particular, if $\mathcal{B}$ is orthonormal, the linear part of this map is

$$\mathrm{M}_{\mathcal{B}}(\mathrm{lin}(\mathrm{Pr}_{H,\mathbf{v}})) = \left(I_n - \frac{\mathbf{v} \otimes \mathbf{a}}{\langle \mathbf{v}, \mathbf{a} \rangle} \right).$$

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

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

Definition 5.15. The *orthogonal projection* $\mathrm{Pr}_H^{\perp}$ on the hyperplane $H \subseteq \mathbb{E}^n$ is the projection on H in the direction of a vector which is orthogonal to H, i.e.

$$\mathrm{Pr}_H^{\perp} = \mathrm{Pr}_{H,\mathbf{v}}$$

where $\mathbf{v}$ is a normal vector of H. With the above notation we see that

$$\mathrm{Pr}_H^{\perp}(P) = \left(I_n - \frac{\mathbf{a} \otimes \mathbf{a}}{|\mathbf{a}|^2} \right) \cdot P - \frac{a_{n+1}}{|\mathbf{a}|^2} \mathbf{a}$$

since we may choose $\mathbf{v} = \mathbf{a}$.

5.1.4 Parallel reflection in a hyperplane

Example 5.16. We consider again the setup in Example 5.13. But this time, we want to reflect in the line (hyperplane)

$$\ell : x + y - 1 = 0$$

in the direction of the vector $\mathbf{v}(-2, -1)$. We do it pointwise. Take an arbitrary point $P(x_P, y_P)$ and consider the line ℓ_P passing through P in the direction of $\mathbf{v}$. It has parametric equations

$$\ell_P : \begin{cases} x = x_P - 2t \\ y = y_P - t. \end{cases}$$

The reflection of P in ℓ in the direction of $\mathbf{v}$ is the point P' such that $\mathrm{Pr}_{\ell,\mathbf{v}}(P) = \ell \cap \ell_P$ is the midpoint of the segment $[PP']$. Thus, identifying points with column matrices of their coordinates relative to $\mathcal{K}$ we have

$$\frac{P + P'}{2} = \mathrm{Pr}_{\ell,\mathbf{v}}(P) \quad \Rightarrow \quad P' = 2\,\mathrm{Pr}_{\ell,\mathbf{v}}(P) - P.$$

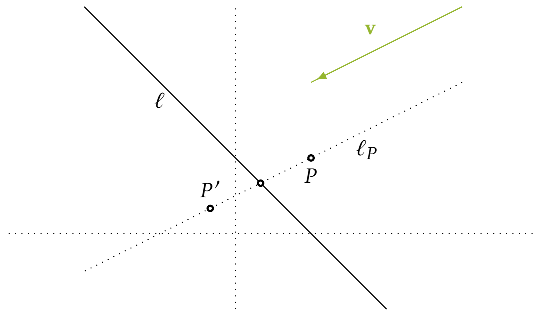


Figure 5.2: Reflection of the point P in the line ℓ parallel to the vector $\mathbf{v}$.

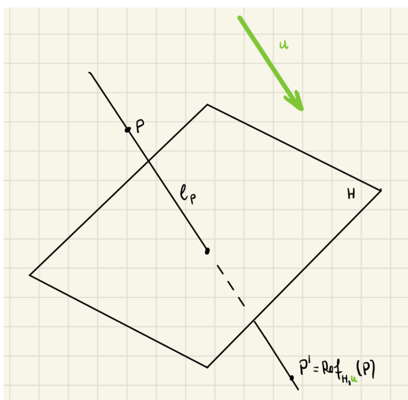
Using the calculation in Example 5.13, the reflection of P is

$$[P']_{\mathcal{K}} = \begin{bmatrix} -\frac{1}{3}x_P - \frac{4}{3}y_P + \frac{4}{3} \\ -\frac{2}{3}x_P + \frac{1}{3}y_P + \frac{2}{3} \end{bmatrix} = \underbrace{\frac{1}{3}\begin{bmatrix} -1 & -4 \\ -2 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x_P \\ y_P \end{bmatrix} + \underbrace{\frac{2}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{b}.$$

Definition 5.17. 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 $P \in \mathbb{A}^n$ 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

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

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



We keep the notation in the previous section. In particular the hyperplane H is given by the equation (5.5). The idea here is the same as in Example 5.16: $\mathrm{Pr}_{H,\mathbf{v}}(P)$ is the midpoint of the segment $[PP']$. Thus, identifying points with column matrices of their coordinates relative to $\mathcal{K}$ we have

$$\mathrm{Pr}_{H,\mathbf{v}}(P) = \frac{P+P'}{2} \quad \Rightarrow \quad \mathrm{Ref}_{H,\mathbf{v}}(P) = P - 2\frac{\mathbf{a}^T \cdot P + a_{n+1}}{\mathbf{v}^T \cdot \mathbf{a}}\mathbf{v}.$$

Again, since $\mathbf{a}^T \cdot P \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{a}^T \cdot P$, in matrix form we have

$$\mathrm{Ref}_{H,\mathbf{v}}(P) = \left(I_n - 2\frac{\mathbf{v}\cdot\mathbf{a}^T}{\mathbf{v}^T\cdot\mathbf{a}} \right) \cdot P - 2\frac{a_{n+1}}{\mathbf{v}^T\cdot\mathbf{a}}\mathbf{v}.$$

In particular, if $\mathcal{B}$ is orthonormal, the linear part of this map is

$$\mathrm{M}_{\mathcal{B}}(\mathrm{lin}(\mathrm{Ref}_{H,\mathbf{v}})) = \left(I_n - 2\frac{\mathbf{v}\otimes\mathbf{a}}{\langle \mathbf{v},\mathbf{a}\rangle} \right).$$

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

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

Definition 5.18. The *orthogonal reflection* $\mathrm{Ref}_H^{\perp}$ in the hyperplane $H \subseteq \mathbb{E}^n$ is the reflection in H parallel to a vector which is orthogonal to H, i.e.

$$\mathrm{Ref}_H^{\perp} = \mathrm{Ref}_{H,\mathbf{v}}$$

where $\mathbf{v}$ is a normal vector of H. With the above notation we see that

$$\mathrm{Ref}_H^{\perp}(P) = \left(I_n - 2\frac{\mathbf{a}\otimes\mathbf{a}}{|\mathbf{a}|^2} \right) \cdot P - 2\frac{a_{n+1}}{|\mathbf{a}|^2}\mathbf{a}$$

since we may choose $\mathbf{v} = \mathbf{a}$.

5.1.5 Parallel projection on a line

Example 5.19. Consider again Example 5.13. We have a line ℓ and a point P which we want to project on ℓ in the direction of $\mathbf{v}(-2,-1)$. We know how to do this, but let us change the role of the Cartesian equation with that of parametric equations. Parametric equations for ℓ are

$$\ell : \begin{cases} x = 1 - t \\ y = t \end{cases}.$$

For an arbitrary point $P(x_P, y_P)$ the line ℓ_P passing through P has direction space given by the equation

$$\mathrm{D}(\ell_P) : x - 2y = 0$$

with respect to the basis $\mathcal{B}$ of the current coordinate system $\mathcal{K}$. Thus, ℓ_P is described by

$$\ell_P : (x - x_P) - 2(y - y_P) = 0.$$

The projection of P on ℓ in the direction of $\mathbf{v}$ is the point $P' = \ell \cap \ell_P$ and the corresponding Figure is 5.1. The only difference is that we describe ℓ and ℓ_P with different types of equations. To determine P' we find the intersection by plugging in the points of ℓ in the equation of ℓ_P

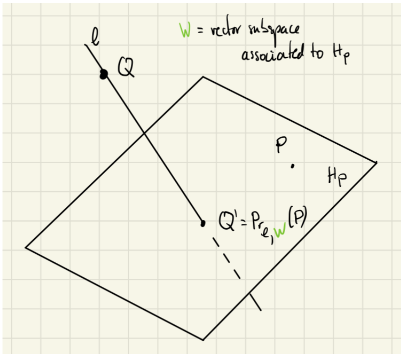
$$(1 - t - x_P) - 2(t - y_P) = 0 \quad \Rightarrow \quad t = -\frac{1}{3}(x_P + y_P - 1).$$

As expected, a short calculation shows that P' has the same expression as in Example 5.13.

Definition 5.20. Let ℓ be a line and let $\mathbb{W}$ be an $(n-1)$-dimensional vector subspace in $\mathbb{V}^n$ which is not parallel to ℓ. For any point $P \in \mathbb{A}^n$ there is a unique hyperplane H_P passing through P and having $\mathbb{W}$ as associated vector subspace. The hyperplane H_P is not parallel to ℓ, hence, it intersects ℓ in a unique point P'. We denote P' by $\mathrm{Pr}_{\ell,\mathbb{W}}(P)$ and call it the *projection of the point P on the line ℓ parallel to $\mathbb{W}$*. This gives a map

$$\mathrm{Pr}_{\ell,\mathbb{W}} : \mathbb{A}^n \to \mathbb{A}^n$$

called, the *projection on the line ℓ parallel to $\mathbb{W}$*.



With respect to the reference frame $\mathcal{K}$, the vector subspace $\mathbb{W}$ is given by a homogeneous equation

$$\mathbb{W} : a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0 \quad \Leftrightarrow \quad \mathbf{a}^T \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0 \tag{5.8}$$

where $\mathbf{a} = \mathbf{a}(a_1, \ldots, a_n)$. Thus, for a given point $P(p_1, \ldots, p_n) \in \mathbb{A}^n$, the equation of H_P is

$$H_P : a_1(x_1 - p_1) + a_2(x_2 - p_2) + \cdots + a_n(x_n - p_n) = 0 \quad \Leftrightarrow \quad \mathbf{a}^T \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{a}^T \cdot P$$

and $\mathrm{Pr}_{\ell,\mathbb{W}}(P) = \mathrm{Pr}_{H_P,\mathbf{v}}(Q)$ for a fixed but arbitrary point $Q(q_1, \ldots, q_n) \in \ell$. Hence, if we denote by $p_1', \ldots, p_n'$ the coordinates of the projected point $\mathrm{Pr}_{\ell,\mathbb{W}}(P)$ then, by (5.7),

$$\begin{cases} p_1' = q_1 + v_1\mu \\ \quad\vdots \\ p_n' = q_n + v_n\mu \end{cases} \quad \text{where} \quad \mu = -\frac{\mathbf{a}^T \cdot Q - \mathbf{a}^T \cdot P}{\mathbf{a}^T \cdot \mathbf{v}}$$

In matrix form we can rearrange this as follows

$$\mathrm{Pr}_{\ell,\mathbb{W}}(P) = \frac{\mathbf{v} \cdot \mathbf{a}^T}{\mathbf{v}^T \cdot \mathbf{a}} P + \left(I_n - \frac{\mathbf{v} \cdot \mathbf{a}^T}{\mathbf{v}^T \cdot \mathbf{a}} \right) Q.$$

In particular, if $\mathcal{B}$ is orthonormal, the linear part of this map is

$$\mathrm{M}_{\mathcal{B}}(\mathrm{lin}(\mathrm{Pr}_{\ell,\mathbb{W}})) = \frac{\mathbf{v}\otimes\mathbf{a}}{\langle\mathbf{v},\mathbf{a}\rangle}.$$

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

$$\mathrm{Pr}_{\ell,\mathbb{W}} \in \mathrm{End}_{\mathrm{aff}}(\mathbb{A}^n) \quad \text{but} \quad \mathrm{Pr}_{\ell,\mathbb{W}} \notin \mathrm{AGL}(\mathbb{A}^n).$$

Definition 5.21. The *orthogonal projection* $\mathrm{Pr}_\ell^\perp$ on the line $\ell \subseteq \mathbb{E}^n$ is the projection on ℓ parallel to vectors which are orthogonal to the line ℓ, i.e.

$$\mathrm{Pr}_\ell^\perp = \mathrm{Pr}_{\ell,\mathbf{v}^\perp}$$

where $\mathbf{v}$ is a direction vector of ℓ. With the above notation we see that

$$\mathrm{Pr}_\ell^\perp(P) = \frac{\mathbf{a}\otimes\mathbf{a}}{|\mathbf{a}|^2}P + \left(I_n - \frac{\mathbf{a}\otimes\mathbf{a}}{|\mathbf{a}|^2}\right)Q$$

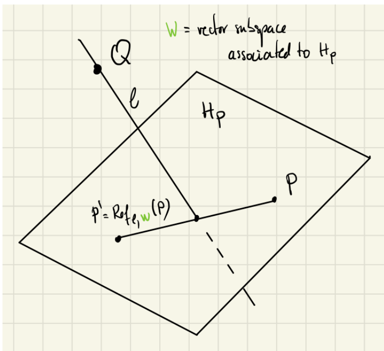
since we may choose $\mathbf{v} = \mathbf{a}$.

5.1.6 Parallel reflection in a line

Definition 5.22. Let ℓ be a line and let $\mathbb{W}$ be an $(n-1)$-dimensional vector subspace in $\mathbb{V}^n$ which is not parallel to ℓ. For any point $P \in \mathbb{A}^n$ there is a unique point P' such that $\mathrm{Pr}_{\ell,\mathbb{W}}(P)$ is the midpoint of the segment $[PP']$. We denote P' by $\mathrm{Ref}_{\ell,\mathbb{W}}(P)$ and call it the *reflection of the point P in the line ℓ parallel to* $\mathbb{W}$. This gives a map

$$\mathrm{Ref}_{\ell,\mathbb{W}} : \mathbb{A}^n \to \mathbb{A}^n$$

called, the *reflection in the line ℓ parallel to* $\mathbb{W}$.



As in Section 5.1.5, the vector subspace $\mathbb{W}$ is given by the homogeneous equation 5.8. The idea is similar to the one used in Section 5.1.4: since $\mathrm{Pr}_{H,\mathbf{v}}(P)$ is the midpoint of the segment $[PP']$, we have

$$\mathrm{Ref}_{\ell,\mathbb{W}}(P) = 2\,\mathrm{Pr}_{\ell,\mathbb{W}}(P) - P$$

Here again we use the convention that point and vectors are identified with column matrices of their coordinates and components respectively. Rearranging this in matrix form we obtain

$$\mathrm{Ref}_{\ell,\mathbb{W}}(P) = \left(2\frac{\mathbf{v}\cdot\mathbf{a}^T}{\mathbf{v}^T\cdot\mathbf{a}} - I_n\right)P + 2\left(I_n - \frac{\mathbf{v}\cdot\mathbf{a}^T}{\mathbf{v}^T\cdot\mathbf{a}}\right)Q.$$

where $Q(q_1,\dots,q_n)$ is a point on ℓ and $\mathbf{v}(v_1,\dots,v_n)$ is a direction vector for ℓ. In particular, if $\mathcal{B}$ is orthonormal, the linear part of this map is

$$\mathrm{M}_{\mathcal{B}}(\mathrm{lin}(\mathrm{Ref}_{\ell,\mathbb{W}})) = 2\frac{\mathbf{v}\otimes\mathbf{a}}{\langle\mathbf{v},\mathbf{a}\rangle} - I_n.$$

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

$$\mathrm{Ref}_{\ell,\mathbb{W}} \in \mathrm{AGL}(\mathbb{A}^n) \subseteq \mathrm{End}_{\mathrm{aff}}(\mathbb{A}^n).$$

Definition 5.23. The *orthogonal reflection* $\mathrm{Ref}_\ell^\perp$ in the line $\ell \subseteq \mathbb{E}^n$ is the reflection in ℓ parallel to vectors which are orthogonal to the line ℓ, i.e.

$$\mathrm{Ref}_\ell^\perp = \mathrm{Ref}_{\ell,\mathbf{v}^\perp}$$

where $\mathbf{v}$ is a direction vector of ℓ. With the above notation we see that for any point $Q \in \ell$

$$\mathrm{Ref}_\ell^\perp(P) = \left(2\frac{\mathbf{a}\otimes\mathbf{a}}{|\mathbf{a}|^2} - I_n\right)P + 2\left(I_n - \frac{\mathbf{a}\otimes\mathbf{a}}{|\mathbf{a}|^2}\right)Q$$

since we may choose $\mathbf{v} = \mathbf{a}$.

5.2 Isometries

Definition 5.24. 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 $P,Q \in \mathbb{E}^n$.

Proposition 5.25. Isometries are affine transformations.

Proposition 5.26. 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. $A^{-1} = A^T$.

Recall from Section 3.2.3 that a matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ such that $A^TA = I_n$ is called orthogonal and that the set of all such matrices is denoted by $O(n)$.

Definition 5.27. 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 $\mathrm{AGL}(\mathbb{A}^n)$ and $SO(n)$ is a normal subgroup of $O(n)$. With group theory notation we write this as follows:

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

Definition 5.28. Let ϕ be an isometry of $\mathbb{E}^n$ 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 $A \in SO(n)$. Else, if $\det(A) = -1$, the map ϕ is called an *indirect isometry*.

5.2.1 Rotations in dimension 2

Proposition 5.29. 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 5.30. 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

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

5.2.2 Rotations in dimension 3

Theorem 5.31 (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

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

Proposition 5.32 (Euler-Rodrigues)**.** Let $\mathbf{v}$ be a unit vector and $\theta \in \mathbb{R}$. The rotation of angle θ and axis $\mathbb{R}\mathbf{v}$ is given by

$$\mathrm{Rot}_{\mathbf{v},\theta}(P) = \cos(\theta)P + \sin(\theta)(\mathbf{v} \times P) + (1 - \cos(\theta))\langle \mathbf{v}, P\rangle \mathbf{v}. \tag{5.9}$$

Where we convene that P is identified with its position vector $\overrightarrow{OP}$.

5.2.3 Classification of isometries

Theorem 5.33 (Chasles)**.** A direct isometry of the plane $\mathbb{E}^2$ is either

a) the identity, or

b) a translation, or

c) a rotation.

Theorem 5.34. An indirect isometry of the plane $\mathbb{E}^2$ fixes a line ℓ and is either

a) a reflection in ℓ, or

b) the composition of a reflection in ℓ with a translation parallel to ℓ, in which case it is called a *glide-reflection*.

Definition 5.35. A *glide-rotation* (or *helical displacement*) is the composition of a rotation in $\mathbb{E}^3$ with a translation parallel to the rotation axis.

Theorem 5.36 (Chasles)**.** A direct isometry of the Euclidean space $\mathbb{E}^3$ is either

a) the identity, or

b) a translation, or

c) a rotation, or

d) a glide-rotation.

Theorem 5.37. An indirect isometry of the Euclidean space $\mathbb{E}^2$ fixes a plane π and is either

a) a reflection in π, or

b) the composition of a reflection in π with a translation parallel to π, in which case it is called a *glide-reflection*, or

c) the composition of a reflection in π with a rotation of axis orthogonal to π, in which case it is called a *rotation-reflection*.

5.3 Some applications

5.3.1 Cycloids

Here we restrict to the Euclidean plane $\mathbb{E}^2$. Rotation matrices give an effective way of describing circular motions which can be used to construct curves as trajectories of a particle. Here we look at some cycloids. For this, let us first deduce the homogeneous matrix of a rotation around a point $C(c_1,c_2)\in\mathbb{E}^2$.

$$\begin{bmatrix}1 & 0 & c_1\\ 0 & 1 & c_2\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix}\cos(\theta) & -\sin(\theta) & 0\\ \sin(\theta) & \cos(\theta) & 0\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix}1 & 0 & -c_1\\ 0 & 1 & -c_2\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}\cos(\theta) & -\sin(\theta) & -c_1\cos(\theta)+c_2\sin(\theta)+c_1\\ \sin(\theta) & \cos(\theta) & -c_1\sin(\theta)-c_2\cos(\theta)+c_2\\ 0 & 0 & 1\end{bmatrix}$$

Now, if you choose the center C to be the point $(0,1)$, you have the following homogenous rotation matrix

$$\begin{bmatrix}\cos(\theta) & -\sin(\theta) & \sin(\theta)\\ \sin(\theta) & \cos(\theta) & -\cos(\theta)+1\\ 0 & 0 & 1\end{bmatrix}$$

If you move the origin with this rotation you obtain

$$\begin{bmatrix}\cos(\theta) & -\sin(\theta) & \sin(\theta)\\ \sin(\theta) & \cos(\theta) & -\cos(\theta)+1\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}\sin(\theta)\\ -\cos(\theta)+1\\ 1\end{bmatrix}$$

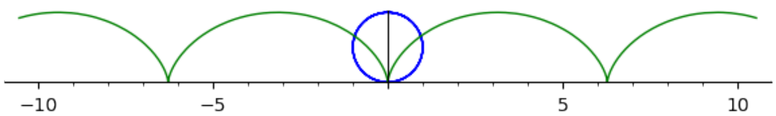
since the homogeneous coordinates of the origin are $(0,0,1)$. If you 'vary θ with time t' you are rotating the origin around C counterclockwise. If you want to have a clockwise motion you just change the sign of the angle to obtain

$$\begin{bmatrix}0\\0\\1\end{bmatrix}\mapsto\begin{bmatrix}-\sin(t)\\ -\cos(t)+1\\ 1\end{bmatrix}.$$

Now, if at the same time t you translate the point along the x-axis in the direction of $\mathbf{i}$, you get

$$\begin{bmatrix}0\\0\\1\end{bmatrix}\mapsto\begin{bmatrix}1 & 0 & t\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}-\sin(t)\\ -\cos(t)+1\\ 1\end{bmatrix}=\begin{bmatrix}-\sin(t)+t\\ -\cos(t)+1\\ 1\end{bmatrix}.$$

What you obtain is the trajectory of the point O as it rotates on the blue circle while the circle is moving like a wheel on the x-axis. The corresponding curve is called a cycloid:



Let's do something else. Instead of rotating the circle on the x-axis let us rotate it on a bigger circle centered at the origin. The small circle we can think of as the trajectory of the point $P(3,0)$ rotated around the center $C(4,0)$. The corresponding rotation matrix is

$$\begin{bmatrix}\cos(\theta) & -\sin(\theta) & -x_C\cos(\theta)+y_C\sin(\theta)+x_C\\ \sin(\theta) & \cos(\theta) & -x_C\sin(\theta)-y_C\cos(\theta)+y_C\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}\cos(\theta) & -\sin(\theta) & -4\cos(\theta)+4\\ \sin(\theta) & \cos(\theta) & -4\sin(\theta)\\ 0 & 0 & 1\end{bmatrix}$$

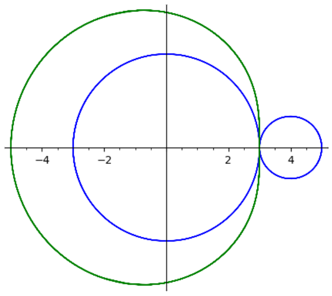
and thus, rotating $P(2,0)$ with time t we obtain

$$\begin{bmatrix}\cos(t) & -\sin(t) & -4\cos(t)+4\\ \sin(t) & \cos(t) & -4\sin(t)\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}3\\0\\1\end{bmatrix}=\begin{bmatrix}3\cos(t)-4\cos(t)+4\\ 3\sin(t)-4\sin(t)\\ 1\end{bmatrix}$$

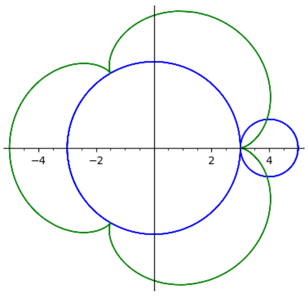
If at the same time we rotate around the origin, P will move on the small circle which rotates around a big circle of radius 3 centered at the origin:

$$\begin{bmatrix}\cos(t') & -\sin(t') & 0\\ \sin(t') & \cos(t') & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}3\cos(t)-4\cos(t)+4\\ 3\sin(t)-4\sin(t)\\ 1\end{bmatrix}=\dots$$

If we do this simultaneously, i.e. if we choose $t'=t$ then we obtain the following trajectory for P:



However, if we want the small circle to rotate like a wheel on the big circle, then, after an entire revolution of the small circle we need to have traversed the length of this circle on the big circle, i.e. 2π. But the big circle is 3 times longer, so we need to choose $t = 3t'$, i.e. the rotation on the small circle is 3-times faster:



This is an example of an epicycloid.

5.3.2 Euler angles

Let us first discuss the effect of rotations on coordinates and on basis vectors in dimension 2. Let $\mathcal{K} = (O, \mathcal{B})$ be a right-oriented orthonormal reference frame of $\mathbb{E}^2$ with $\mathcal{B} = (\mathbf{i}, \mathbf{j})$. A rotation around the origin with angle θ is given by the map

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{=\mathrm{Rot}_\theta} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

Thus, Rot_θ is a base change matrix $\mathrm{M}_{\mathcal{B}',\mathcal{B}}$ where $\mathcal{B}' = (\mathbf{i}', \mathbf{j}')$ is another orthonormal basis. The components of $\mathbf{i}'$ and $\mathbf{j}'$ with respect to $\mathcal{B}$ are the columns of the matrix $\mathrm{M}^{-1}_{\mathcal{B}',\mathcal{B}} = \mathrm{M}^{T}_{\mathcal{B}',\mathcal{B}}$:

$$\mathbf{i}' = \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix}, \quad \mathbf{j}' = \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix}.$$

We notice that the versors in $\mathcal{B}'$ are obtained by rotating the versors in $\mathcal{B}$ by $-\theta$. This is what we expect since rotating points counterclockwise with respect to $\mathcal{K}$ is equivalent to rotating $\mathcal{K}$ clockwise.

Let us now consider rotations in $\mathbb{E}^3$. Fix a right oriented orthonormal reference frame $\mathcal{K} = (O, \mathcal{B})$. Rotations around the coordinate axes are given by the following matrices:

$$\mathrm{Rot}_{x,\theta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \quad \mathrm{Rot}_{y,\theta} = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}, \quad \mathrm{Rot}_{z,\theta} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Consider the composition of the following three rotations $\mathrm{Rot}_{z,\gamma}\mathrm{Rot}_{x,\beta}\mathrm{Rot}_{z,\alpha} =$

$$\begin{bmatrix} \cos\gamma\cos\alpha - \sin\gamma\cos\beta\sin\alpha & -\cos\gamma\sin\alpha - \sin\gamma\cos\beta\cos\alpha & \sin\gamma\sin\beta \\ \sin\gamma\cos\alpha + \cos\gamma\cos\beta\sin\alpha & -\sin\gamma\sin\alpha + \sin\gamma\cos\beta\cos\alpha & -\cos\gamma\sin\beta \\ \sin\beta\sin\alpha & \sin\beta\cos\alpha & \cos\beta \end{bmatrix}.$$

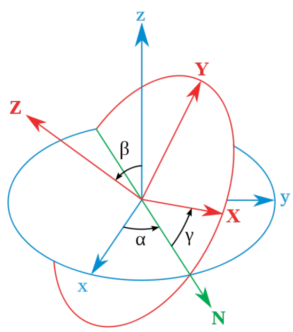


Figure 5.3: Euler angles[1]

You may think of the composition of these three rotations as follows: Each of the three rotations is a base change matrix. The first rotation, $\mathrm{Rot}_{z,\alpha} = \mathrm{M}_{\mathcal{K}',\mathcal{K}}$, rotates the versor of the x-axis and that of the y-axis by $-\alpha$ (and therefore rotates points with respect to $\mathcal{K}$ by α). The next rotation $\mathrm{Rot}_{x,\beta} = \mathrm{M}_{\mathcal{K}'',\mathcal{K}'}$, rotates the versors of the y'-axis and that of the z'-axis by $-\beta$ (and therefore rotates points with respect to $\mathcal{K}'$ by β). Similarly with the last rotation. The observation here is that $\mathrm{Rot}_{x,\beta}$ is a rotation around the *current x-axis*, i.e. a rotation around the first axis of the coordinate system that you are in. If we want to point out that at each step the coordinate system is changing we may write $\mathrm{Rot}_{z'',\gamma}\mathrm{Rot}_{x',\beta}\mathrm{Rot}_{z,\alpha}$ for the overall rotation.

Proposition 5.38. All rotations in dimension 3 are of this form, i.e. any matrix in SO(3) can be written in this form for some $\alpha, \beta, \gamma \in \mathbb{R}$.

Definition 5.39. Moreover, you may restrict the range of the values α, β, γ by $\alpha \in [0, 2\pi[$, $\beta \in [0, \pi[$ and $\gamma \in [0, 2\pi[$. Then, each triple (α, β, γ) corresponds to a unique rotation $\mathrm{Rot}_{z,\gamma}\mathrm{Rot}_{x,\beta}\mathrm{Rot}_{z,\alpha} \in \mathrm{SO}(3)$. The angles α, β and γ are called *Euler angles*. Another way of describing rotations in $\mathbb{E}^3$ is via quaternions (see Appendix B).

[1] Image source: Wikipedia

Another way of looking at $\mathrm{Rot}_{z,\gamma}\mathrm{Rot}_{x,\beta}\mathrm{Rot}_{z,\alpha}$ is by consider the trajectory that you obtain on a given point when you vary α, β and γ. For instance

$$\mathrm{Rot}_{z,\gamma}\mathrm{Rot}_{x,\beta}\mathrm{Rot}_{z,\alpha}\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}\cos\gamma\cos\alpha-\sin\gamma\cos\beta\sin\alpha\\ \sin\gamma\cos\alpha+\cos\gamma\cos\beta\sin\alpha\\ \sin\beta\sin\alpha\end{bmatrix}.$$

If in this expression you fix $\gamma=0$ and let $\alpha\in[0,2\pi[$, $\beta\in[0,\pi[$ vary, you obtain

$$\begin{bmatrix}1\\0\\0\end{bmatrix}\mapsto\begin{bmatrix}\cos\alpha\\ \cos\beta\sin\alpha\\ \sin\beta\sin\alpha\end{bmatrix},$$

i.e. the trajectory of the point $(1,0,0)$ is a sphere and varying α and β corresponds to the map

$$[0,2\pi[\times[0,\pi[\ni(\alpha,\beta)\mapsto\begin{bmatrix}\cos\alpha\\ \cos\beta\sin\alpha\\ \sin\beta\sin\alpha\end{bmatrix}, \tag{5.10}$$

which is a parametrization of the sphere. You can get other parametrizations of the sphere if you fix α and let β and γ vary.

5.3.3 Surfaces of revolution

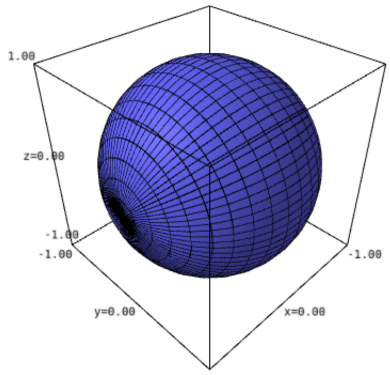
First, notice that a plane in $\mathbb{E}^3$ can be described as the set of points which you touch if you translate a line in a given direction. This can be seen with the parametric equations:

$$\pi:\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x_Q\\y_Q\\z_Q\end{bmatrix}+s\begin{bmatrix}v_x\\v_y\\v_z\end{bmatrix}+t\begin{bmatrix}w_x\\w_y\\w_z\end{bmatrix}=\underbrace{\left(\begin{bmatrix}x_Q\\y_Q\\z_Q\end{bmatrix}+s\begin{bmatrix}v_x\\v_y\\v_z\end{bmatrix}\right)}_{\text{line }\ell}\quad\underbrace{+t\begin{bmatrix}w_x\\w_y\\w_z\end{bmatrix}}_{\text{translation with }t\mathbf{w}} \tag{5.11}$$

This is a general method of constructing surfaces starting from curves: you start with a curve in $\mathbb{E}^3$ and apply a motion to it. What you obtain, if non-degenerate, is a parametrization of a surface. This can be exemplified with the parametrization of the sphere in (5.10) which you can rewrite as follows

$$(\alpha,\beta)\mapsto\underbrace{\begin{bmatrix}1&0&0\\0&\cos\beta&-\sin\beta\\0&\sin\beta&\cos\beta\end{bmatrix}}_{\text{rotation }\mathrm{Rot}_{x,\beta}}\quad\underbrace{\begin{bmatrix}\cos\alpha\\ \sin\alpha\\ 0\end{bmatrix}}_{\text{circle in }Oxy\text{-plane}}\quad=\begin{bmatrix}\cos\alpha\\ \cos\beta\sin\alpha\\ \sin\beta\sin\alpha\end{bmatrix}.$$

This describes the sphere as the points which you touch with the unit circle centered at the origin in the Oxy-plane if you rotate the circle around the x-axis.



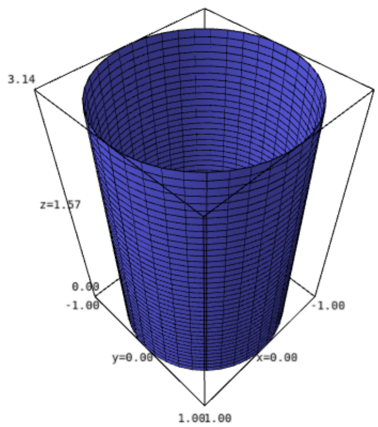
Definition 5.40. A *surface of revolution* in $\mathbb{E}^3$ is a surface obtained by rotating a curve around a line ℓ. The line ℓ is called the *axis* of the surface.

Example 5.41. In (5.11), instead of translating the line ℓ we can rotate it around a line which is parallel to ℓ. In this way we obtain a cylinder. For example, if

$$\ell : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

and if we rotate around the z-axis we obtain a parametrization of a cylinder of radius 1 and axis the z-axis:

$$(\theta, s) \mapsto \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} \cos\theta \\ \sin\theta \\ s \end{bmatrix}$$



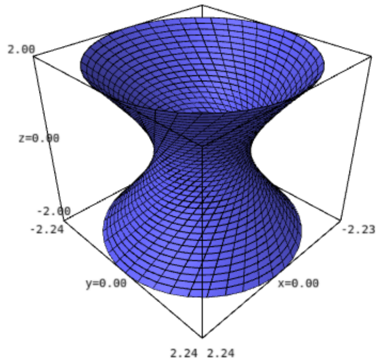
Example 5.42. If instead we consider two skew lines and rotate one around the other, we obtain

hyperboloids of revolution. For example, if

$$\ell : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

and if we rotate around the z-axis we obtain a parametrization of a hyperboloid

$$(\theta, s) \mapsto \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} \cos\theta - s\sin\theta \\ \sin\theta + s\cos\theta \\ s \end{bmatrix}.$$



5.4 Exercises

5.1. Consider an orthonormal coordinate system $\mathcal{K}$ of $\mathbb{E}^n$ where $n = 2$ or 3. Starting from the matrix form of the projections and reflections described in this Chapter, deduce the matrices of

a) the orthogonal projections on the coordinate axes and on the coordinate hyperplanes of $\mathcal{K}$.

b) the orthogonal reflections in coordinate axes and in coordinate hyperplanes of $\mathcal{K}$.

5.2. Consider the vector $\mathbf{v}(2,1,1) \in \mathbb{V}^3$.

a) Give the matrix form for the parallel projection on the plane $\pi : z = 0$ parallel to $\mathbf{v}$.

b) Give the matrix form for the parallel reflection in the plane $\pi : z = 0$ parallel to $\mathbf{v}$.

5.3. Determine the orthogonal projection of the point $A(2,11,-5)$ on the plane $x + 4y - 3z + 7 = 0$ by determining the matrix form of the projection.

5.4. Determine the orthogonal reflection of the point $P(6,-5,5)$ in the plane $2x - 3y + z - 4 = 0$ by determining the matrix form of the reflection.

5.5. Consider an orthonormal coordinate system $\mathcal{K}$ of $\mathbb{E}^2$. Starting from the matrix form of the projections and reflections described in this Section 5.1, show that

$$\mathrm{Pr}^{\perp}_{\mathbf{a}}(\mathbf{b}) = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}\mathbf{a}$$

Compare this to the projections described in Section 3.1.

5.6. Determine the orthogonal projection of the line $\ell : 2x - y - 1 = 0 \cap x + y - z + 1 = 0$ on the plane $\pi : x + 2y - z = 0$ by determining the matrix form of the projection. (Compare your result with the previous seminar.)

5.7. Give Cartesian equations for the line passing through the point $M(1,0,7)$, parallel to the plane $\pi : 3x - y + 2z - 15 = 0$ and intersecting the line

$$\ell : \frac{x-1}{4} = \frac{y-3}{2} = \frac{z}{1}.$$

5.8. In $\mathbb{E}^3$, show that the orthogonal reflection $\mathrm{Ref}^{\perp}_{\pi}(x)$ in the plane $\pi : \langle n, x \rangle = p$ is given by

$$\mathrm{Ref}_{\pi}(x) = Ax + b$$

where $A = \left(I - 2\frac{n \otimes n}{|n|^2}\right)$ and $b = \frac{2p}{|n|^2}n$.

5.9. Give the matrix form for the orthogonal reflections in the planes

$$\pi_1 : 3x - 4z = -1 \quad \text{and} \quad \pi_2 : 10x - 2y + 3z = 4 \quad \text{respectively.}$$

5.10. Write down the vector forms and matrix forms for parallel projections and reflections in $\mathbb{E}^3$.

5.11. In $\mathbb{E}^2$, for the lines/hyperplanes

$$\pi : ax + by + c = 0, \quad \ell : \frac{x - x_0}{v_1} = \frac{y - y_0}{v_2}$$

with $\pi \nparallel \ell$, deduce the matrix forms of $\mathrm{Pr}_{\pi,\ell}$ and $\mathrm{Ref}_{\pi,\ell}$.

5.12. Let H be a hyperplane and let $\mathbf{v}$ be a vector which is not parallel to H. Use the deduced matrix forms to show that

a) $\mathrm{Pr}_{H,\mathbf{v}} \circ \mathrm{Pr}_{H,\mathbf{v}} = \mathrm{Pr}_{H,\mathbf{v}}$ and

b) $\mathrm{Ref}_{H,\mathbf{v}} \circ \mathrm{Ref}_{H,\mathbf{v}} = \mathrm{Id}$.

5.13. Let $\phi(\mathbf{x}) = A\mathbf{x} + b$ be an affine transformation. Give the homogenous matrix of the inverse transformation ϕ^{-1}.

5.14. The vertices of a triangle are $A(1,1)$, $B(4,1)$ and $C(2,3)$. Determine the image of the triangle ABC under a rotation by 90° around C followed by an orthogonal reflection relative to the line AB.

5.15. Determine the sum-of-angles formulas for sine and cosine using rotation matrices.

5.16. Let T be the isometry obtained by applying a rotation of angle $-\frac{\pi}{3}$ around the origin after a transation with vector $(-2,5)$. Determine the inverse transformation, T^{-1}.

5.17. Determine the matrix form of a rotation with angle 45° having the same center of rotation as the rotation

$$f(\mathbf{x}) = \frac{1}{\sqrt{13}}\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

5.18. Determine the cosine of the angle of the rotation f given in the previous exercise and find the inverse rotation, f^{-1}.

5.19. Verify that the matrices

$$A = \frac{1}{3}\begin{bmatrix} -1 & 2 & -2 \\ -2 & -2 & -1 \\ -2 & 1 & 2 \end{bmatrix} \quad \text{and} \quad B = \frac{1}{11}\begin{bmatrix} -9 & -2 & 6 \\ 6 & -6 & 7 \\ 2 & 9 & 6 \end{bmatrix}$$

belong to $SO(3)$. Moreover, determine the axis of rotation and the rotation angle.

5.20. Check

1. the calculations in Section 5.3.1.
2. the calculations in Section 5.3.2.
3. the calculations in Section 5.3.3.

5.21. Using Euler-Rodrigues formula, deduce the known matrix forms of rotations around the coordinate axes.

5.22. Using Euler-Rodrigues formula, write down the matrix form of a rotation around the axis $\mathbb{R}\mathbf{v}$ where $\mathbf{v} = (1,1,0)$. Use this matrix form to give a parametrization of a cylinder with axis $\mathbb{R}\mathbf{v}$ and diameter $\sqrt{2}$.

5.23. Using rotations around the coordinate axes, give a parametrization of a cylinder with axis $\mathbb{R}\mathbf{v}$ and diameter $\sqrt{2}$.

5.24. Using rotations around the coordinate axes, give a parametrization of a cone containing the line $\ell = \{(0,t,t) : t \in \mathbb{R}\}$ and with axis the z-axis.

CHAPTER 6

Quadratic curves (conics)

Contents

Definition 6.1. A *quadratic curve* (or *conic*) in $\mathbb{E}^2$ is a curve described by a quadratic equation

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

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

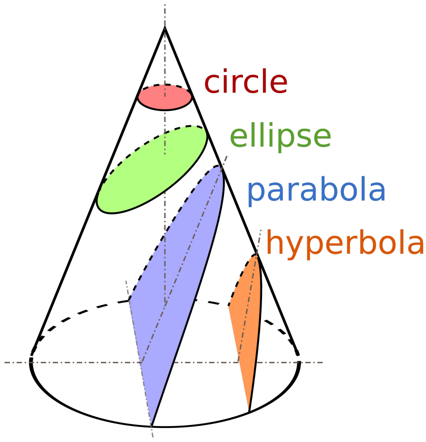
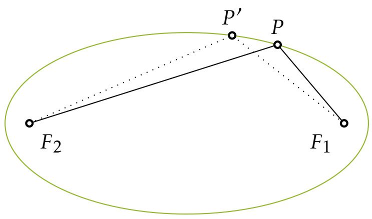


Figure 6.1: Conic sections[1]

6.1 Ellipse

6.1.1 Geometric description



Definition 6.2. 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.

6.1.2 Canonical equation

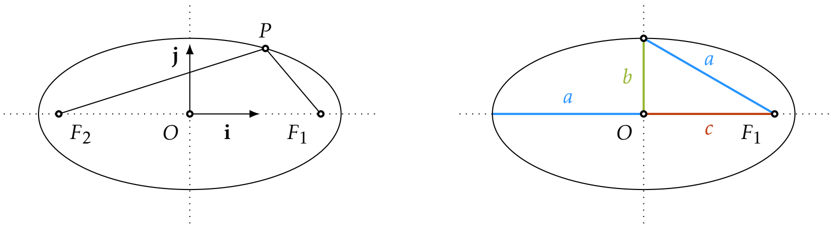
In general, we can describe a conic with a so-called canonical equation. Such an equation is with respect to a well chosen coordinate system.

[1] Image source: Wikipedia

Proposition 6.3. 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

$$\mathcal{E}_{a,b} : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{6.1}$$

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



- Equation (6.1) 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 $(\pm x_M, \pm y_M) \in \mathcal{E}_{a,b}$.

6.1.3 Parametric equations

Parametric equations are never unique. Depending on what your intentions are you may prefer one over the other. Equation (6.1) 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

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

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

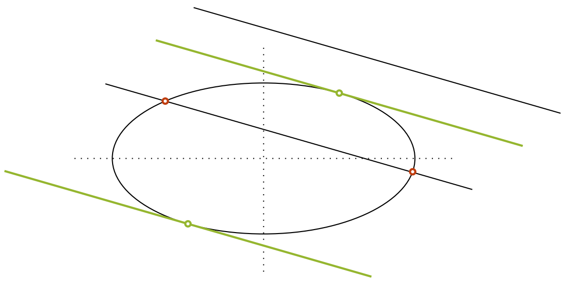
$$\phi : \mathbb{R} \to \mathbb{E}^2 \quad \text{defined by} \quad \phi(t) = (a\cos(t), b\sin(t)).$$

It is easy to check using equation (6.1) that this is a parametrization. Moreover, with this parametrization, we may view the ellipse as the orbit of a rotation followed by a dilation along the coordinate axes:

$$t \mapsto \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

We call such a transformation *elliptical motion*.

6.1.4 Relative position of a line



Consider the canonical equation (6.1) of the ellipse $\mathcal{E}_{a,b}$. Let ℓ be a line with equation $y = kx + m$. 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:

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

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 with this slope:
$$y = kx \pm \sqrt{a^2k^2+b^2}.$$
- $m < \sqrt{a^2k^2+b^2}$ or $m > \sqrt{a^2k^2+b^2}$ in which case there is no intersection point between ℓ and $\mathcal{E}_{a,b}$.

6.1.5 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

$$\frac{x_0v_x}{a^2} + \frac{y_0v_y}{b^2} = 0 \quad \Leftrightarrow \quad \langle \mathbf{n}, \mathbf{v} \rangle = 0$$

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:

$$T_{(x_0,y_0)}\mathcal{E}_{a,b} : \frac{x_0x}{a^2} + \frac{y_0y}{b^2} = 1. \tag{6.2}$$

6.1.6 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

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

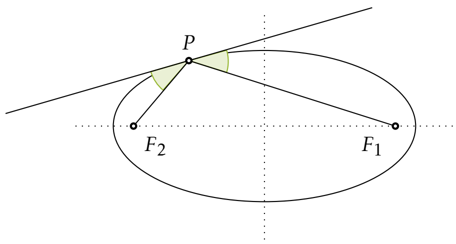
and notice that $\mathcal{E}_{a,b}=\psi^{-1}(1)$. The gradient in a point $(x_0,y_0)\in\mathcal{E}_{a,b}$ is

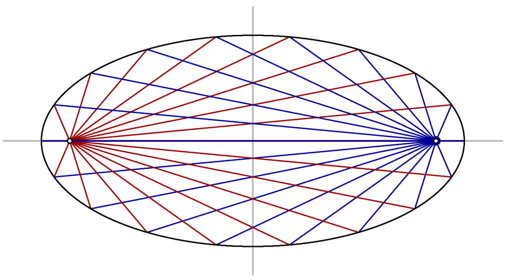
$$\nabla_{(x_0,y_0)}(\psi)=\left(2\frac{x}{a^2},2\frac{y}{b^2}\right)_{(x_0,y_0)}=2\left(\frac{x_0}{a^2},\frac{y_0}{b^2}\right).$$

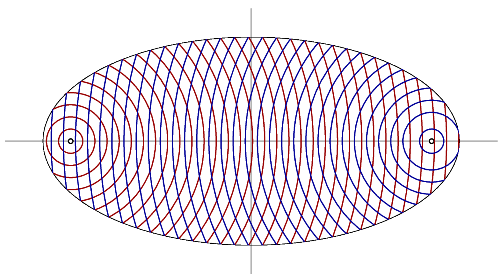
Using a parametrization $\phi:I\to\mathbb{E}^2$ of $\mathcal{E}_{a,b}$ and calculating $\partial_t\psi(\phi(t))$ with the chain rule, one shows that $\nabla_{(x_0,y_0)}(\psi)$ is orthogonal to the tangent vectors at the point (x_0,y_0). In other words, $\nabla_{(x_0,y_0)}(\psi)$ is a normal vector at the point $(x_0,y_0)\in\mathcal{E}_{a,b}$. This gives a different way of obtaining the equation (6.2).

6.1.7 Reflective properties

An ellipse has the following reflective properties: a ray starting in one focal point is being reflected by the ellipse to the other focal point, i.e. if P is a point on the ellipse then the tangent line in P is an exterior angle bisector of the triangle F_1PF_2.

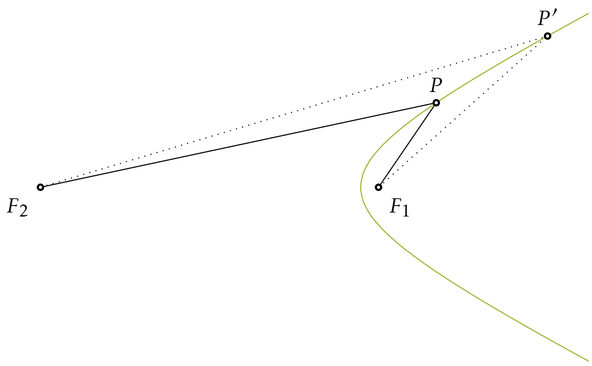






6.2 Hyperbola

6.2.1 Geometric description



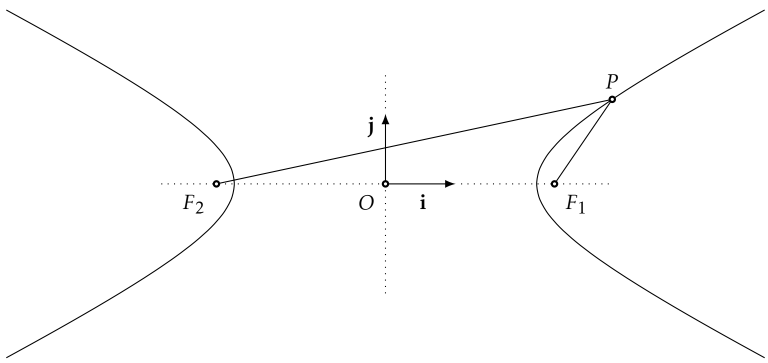
Definition 6.4. A *hyperbola* is the geometric locus of points in $\mathbb{E}^2$ for which the difference of the distances from two given points, the *focal points*, is constant.

6.2.2 Canonical equation

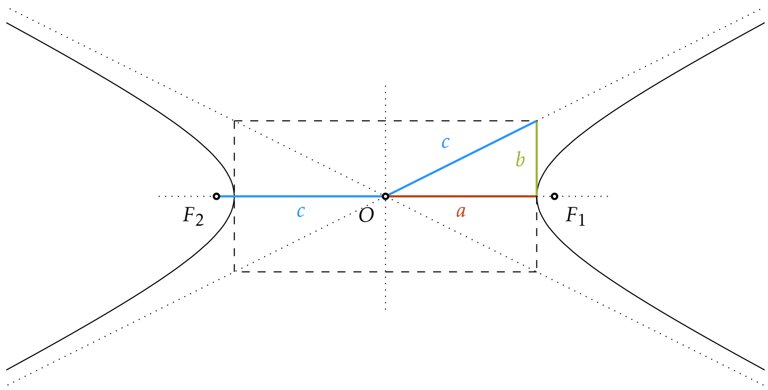
Proposition 6.5. Let F_1 and F_2 be two points in $\mathbb{E}^2$ and let a be a positive real scalar. Choose the coordinate system $Oxy = (O, \mathbf{i}, \mathbf{j})$ such that F_1 and F_2 are on the Ox axis, such that $\overrightarrow{F_2F_1}$ has the same direction as $\mathbf{i}$ and such that O is 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

$$\mathcal{H}_{a,b} : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{6.3}$$

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



- Equation (6.3) 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 $(\pm x_M, \pm y_M) \in \mathcal{H}_{a,b}$.



6.2.3 Parametric equations

Parametric equations are never unique. Depending on what your intentions are you may prefer one over the other. Equation (6.3) 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

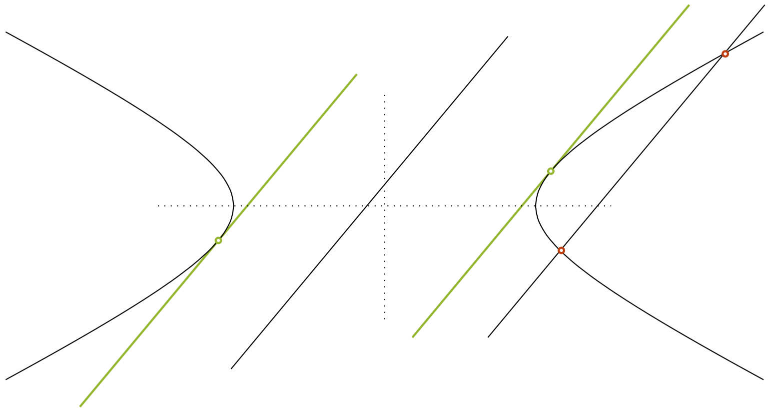
$$\phi : (-\infty, -a] \cup [a, \infty) \to \mathbb{E}^2 \quad \text{given by} \quad \phi(x) = \big(x, y(x)\big) = \Big(x, \frac{b}{a}\sqrt{x^2 - a^2}\Big).$$

This is the graph of the function

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

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

6.2.4 Relative position of a line



Consider the canonical equation (6.3) of the hyperbola $\mathcal{H}_{a,b}$. Let ℓ be a line with equation $y = kx + m$. 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:

$$(b^2 - a^2k^2)x^2 - 2kma^2x - a^2(m^2 + b^2) = 0 \tag{6.4}$$

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 (6.4) 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 such tangent lines to the hyperbola:

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

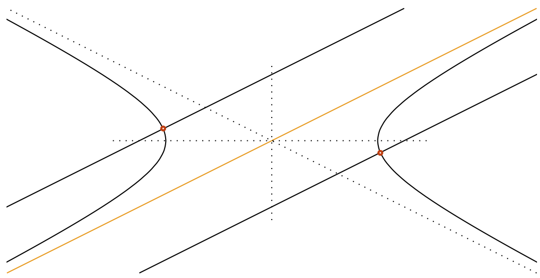
- $-\sqrt{a^2k^2 - b^2} < m < \sqrt{a^2k^2 - b^2}$ in which case there is no intersection point between ℓ and $\mathcal{H}_{a,b}$.

Suppose equation (6.4) 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.



6.2.5 Tangent line in a given point

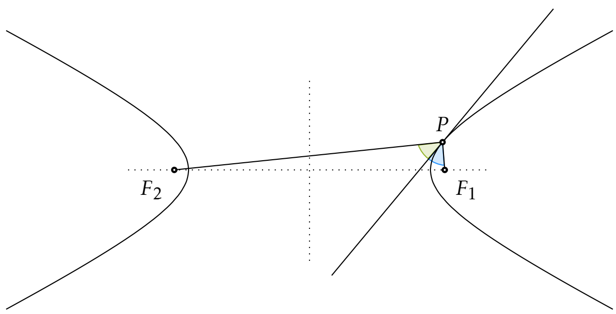
The *tangent line to $\mathcal{H}_{a,b}$ at the point* $(x_0, y_0) \in \mathcal{H}_{a,b}$ has an equation of the form

$$T_{(x_0,y_0)}\mathcal{H}_{a,b} : \frac{x_0 x}{a^2} - \frac{y_0 y}{b^2} = 1. \tag{6.5}$$

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

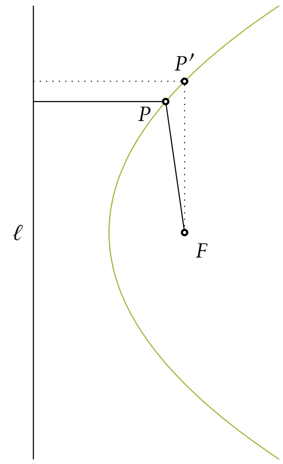
6.2.6 Reflective properties

A hyperbola has the following reflective properties: for a point P on the hyperbola, the tangent line at P is the angle bisector of the angle $\measuredangle F_1 P F_2$.



6.3 Parabola

6.3.1 Geometric description



Definition 6.6. 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*.

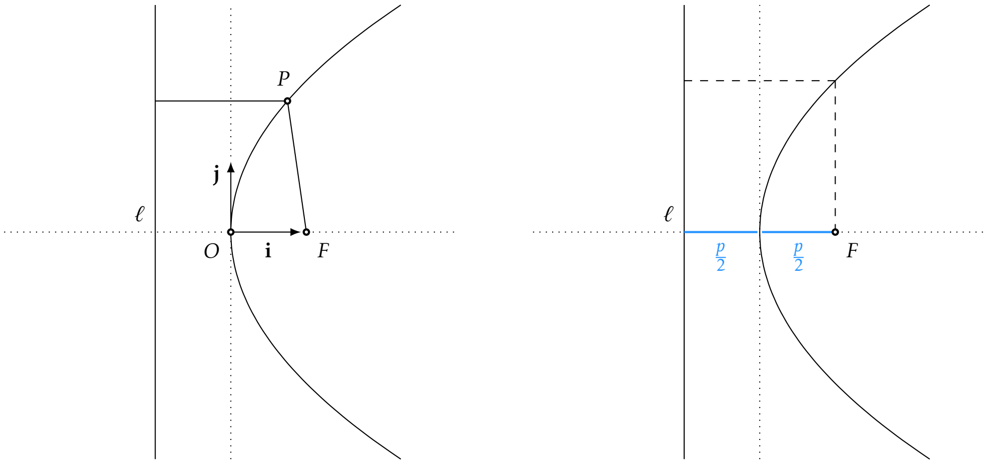
6.3.2 Canonical equation

Proposition 6.7. Let F be a point, let d be a line in $\mathbb{E}^2$ and let p be a positive 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 $d(F, d) = p$ has an equation of the form

$$\mathcal{P}_p : y^2 = 2px \tag{6.6}$$

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

- Equation (6.6) 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 $(x_M, \pm y_M) \in \mathcal{P}_p$.



6.3.3 Parametric equations

Parametric equations are never unique. Depending on what your intentions are you may prefer one over the other. Equation (6.6) 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

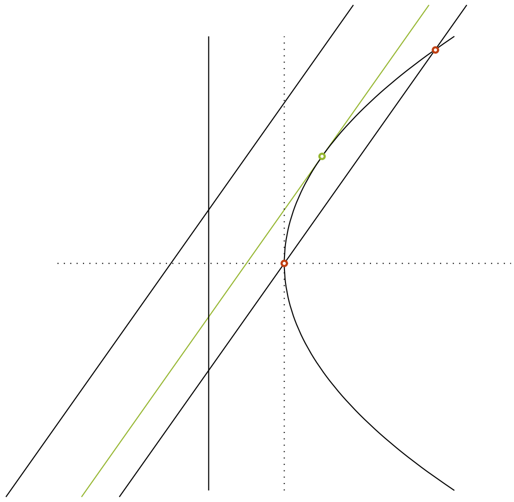
$$y(x) = \sqrt{2px} \quad \text{for which} \quad y'(x) = \frac{\sqrt{2p}}{2\sqrt{x}} \quad \text{and} \quad y''(x) = -\frac{\sqrt{2p}}{4x^{3/2}}$$

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 (6.6). We then have the parametrization

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

6.3.4 Relative position of a line



Consider the canonical equation (6.6) of the parabola $\mathcal{P}_p$. Let ℓ be a line with equation $y = kx + m$. 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:

$$k^2x^2 + 2(km - p)x + m^2 = 0 \tag{6.7}$$

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 having the given slope:
$$y = kx + \frac{p}{2k}.$$
- $km > p/2$ in which case there is no intersection point between ℓ and $\mathcal{P}_p$.

6.3.5 Tangent line in a given point

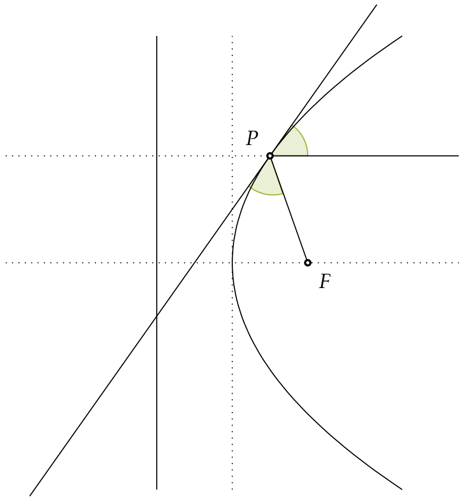
The *tangent line to $\mathcal{P}_p$ at the point* $(x_0, y_0) \in \mathcal{P}_p$ has an equation of the form

$$T_{(x_0,y_0)}\mathcal{P}_p : yy_0 = p(x + x_0) \tag{6.8}$$

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

6.3.6 Reflective properties

A parabola has the following reflective properties: rays starting in the focal point are reflected by the parabola in rays parallel to the axis of the parabola. Equivalently, rays inside the parabola which are parallel to the axis are reflected in the focal point.



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

6.4 Exercises

6.1. Find the equation of the circle:

a) of diameter $[A,B]$, with $A(1,2)$ and $B(-3,-1)$,

b) with center $I(2,-3)$ and radius $R=7$,

c) with center $I(-1,2)$ and passing through $A(2,6)$,

d) centered at the origin and tangent to $\ell: 3x-4y+20=0$,

e) passing through $A(3,1)$ and $B(-1,3)$ and having the center on the line $\ell: 3x-y-2=0$,

f) passing through $A(1,1)$, $B(1,-1)$ and $C(2,0)$,

g) tangent to both $\ell_1: 2x+y-5=0$ and $\ell_2: 2x+y+15=0$ if one tangency point is $M(3,-1)$.

6.2. For a circle $\mathcal{C}$ of radius R:

a) Use the parametrization $x \mapsto (x, \pm\sqrt{R^2-x^2})$ to deduce a parametrization of tangent lines to $\mathcal{C}$.

b) Use the parametrization $\theta \mapsto (R\cos(\theta), R\sin(\theta))$ to deduce a parametrization of tangent lines to $\mathcal{C}$.

c) Compare these to the equation of the tangent line $xx_0+yy_0=R^2$ where $(x_0,y_0)\in\mathcal{C}$.

6.3. Determine the foci (focal points) of the Ellipse $9x^2+25y^2-225=0$

6.4. Determine the intersection of the line $\ell: x+2y-7=0$ and the ellipse $\mathcal{E}: x^2+3y^2-25=0$.

6.5. Determine the position of the line $\ell: 2x+y-10=0$ relative to the ellipse $\mathcal{E}: \frac{x^2}{9}+\frac{y^2}{4}-1=0$.

6.6. Determine an equation of a line which is orthogonal to $\ell: 2x-2y-13=0$ and tangent to the ellipse $\mathcal{E}: x^2+4y^2-20=0$.

6.7. A *diameter* of an ellipse is the line segment determined by the intersection points of the ellipse with a line passing through the center of the ellipse. Show that the tangent lines to an ellipse at the endpoints of a diameter are parallel.

6.8. Consider the family of ellipses $\mathcal{E}_a: \frac{x^2}{a^2}+\frac{y^2}{16}=1$. For what value $a\in\mathbb{R}$ is $\mathcal{E}_a$ tangent to the line $\ell: x-y+5=0$?

6.9. Consider the family of lines $\ell_c: \sqrt{5}x-y+c=0$. For what values $c\in\mathbb{R}$ is ℓ_c tangent to the ellipse $\mathcal{E}: x^2+\frac{y^2}{4}=1$?

6.10. Determine the common tangents to the ellipses

$$\frac{x^2}{45}+\frac{y^2}{9}=1 \quad\text{and}\quad \frac{x^2}{9}+\frac{y^2}{18}=1.$$

6.11. Consider the ellipse $\mathcal{E} : \frac{x^2}{4} + y^2 - 1 = 0$ with focal points F_1 and F_2. Determine the points M, situated on the ellipse, for which

a) the angle $\measuredangle F_1MF_2$ is right;

b) the angle $\measuredangle F_1MF_2$ is θ;

c) the angle $\measuredangle F_1MF_2$ is maximal.

6.12. Using a rotation of the coordinate system, find the equation of an ellipse centered at the origin, with focal points on the line $x = y$ and having the large diameter equal to 4 and the distance between the focal points equal to $2\sqrt{3}$.

6.13. Consider the ellipse $\mathcal{E} : x^2 + 4y^2 = 25$. Find the chords on the ellipse which have the point $A(7/2, 7/4)$ as midpoint.

6.14. Consider the ellipse $\mathcal{E} : \frac{x^2}{25} + \frac{y^2}{9} = 1$. Determine the geometric locus of the midpoints of the chords on the ellipse which are parallel to the line $\ell : x + 2y = 1$.

6.15. Using the gradient, prove the reflective properties of an ellipse.

6.16. Determine the intersection points between the line $\ell : 2x - y - 10 = 0$ and the hyperbola $\mathcal{H} : \frac{x^2}{20} - \frac{y^2}{5} - 1 = 0$.

6.17. Determine the tangents to the hyperbola $\mathcal{H} : \frac{x^2}{16} - \frac{y^2}{8} - 1 = 0$ which are parallel to the line $\ell : 4x + 2y - 5 = 0$.

6.18. Determine the tangents to the hyperbola $\mathcal{H} : x^2 - y^2 = 16$ which contain the point $M(-1, 7)$.

6.19. Determine the relations between the coordinates (x_P, y_P) of the point P such that P does not belong to any tangent line to the hyperbola

$$\frac{x^2}{4} - \frac{y^2}{9} = 1.$$

6.20. Find the area of the triangle determined by the asymptotes of the hyperbola $\mathcal{H} : \frac{x^2}{4} - \frac{y^2}{9} - 1 = 0$ and the line $\ell : 9x + 2y - 24 = 0$.

6.21. Find an equation for the tangent lines to:

a) the hyperbola $\mathcal{H} : \frac{x^2}{20} - \frac{y^2}{5} - 1 = 0$, orthogonal to the line $\ell : 4x + 3y - 7 = 0$;

b) the parabola $\mathcal{P} : y^2 - 8x = 0$, parallel to $\ell : 2x + 2y - 3 = 0$.

6.22. Find an equation for the tangent lines to:

a) the hyperbola $\mathcal{H} : \frac{x^2}{3} - \frac{y^2}{5} - 1 = 0$, passing through $P(1, -5)$;

b) the paraola $\mathcal{P} : y^2 - 36x = 0$, passing through $P(2, 9)$.

6.23. Consider the hyperbola $\mathcal{H} : x^2 - \frac{y^2}{4} - 1 = 0$ with focal points F_1 and F_2. Find the points M situated on the hyperbola such that

a) The angle $\measuredangle F_1MF_2$ is right;

b) The angle $\measuredangle F_1MF_2$ is $60°$;

c) The angle $\measuredangle F_1MF_2$ is θ.

6.24. Consider the tangents to the parabola $\mathcal{P} : y^2 - 10x = 0$ passing through the point $P(-3, 12)$. Calculate the distance from the point P to the chord of the parabola which is formed by the two contact points.

6.25. Consider the hyperbola $\mathcal{H} : x^2 - 2y^2 = 1$. Determine the geometric locus described by the midpoints of the chords of $\mathcal{H}$ which are parallel to the line $2x - y = 0$.

6.26. For which value k is the line $y = kx + 2$ tangent to the parabola $\mathcal{P} : y^2 = 4x$?

6.27. Consider the parabola $\mathcal{P} : y^2 = 16x$. Determine the tangents to $\mathcal{P}$ which are

a) parallel to the line $\ell : 3x - 2y + 30 = 0$;

b) perpendicular to the line $\ell : 4x + 2y + 7 = 0$.

6.28. Determine the tangents to the parabola $\mathcal{P} : y^2 = 16x$ which contain the point $P(-2, 2)$.

6.29. Using the gradient, prove the reflective properties of the hyperbola and of the parabola.

CHAPTER 7

Hyperquadrics

Contents

7.1 Hyperquadrics

Definition 7.1. A *hyperquadric* $\mathcal{Q}$ in $\mathbb{E}^n$ is a the set of points whose coordinates satisfy a quadratic equation, i.e.

$$\mathcal{Q} : \sum_{i,j=1}^{n} a_{ij}x_i x_j + \sum_{i}^{n} b_i x_i + c = 0, \tag{7.1}$$

with respect to some coordinate system.

- Hyperquadrics in $\mathbb{E}^2$ are called *conic sections* (see previous chapter).
- In Chapter 8 we consider hyperquadrics in $\mathbb{E}^3$. In dimension 3 they are called *quadrics*.
- Notice that Equation (7.1) is equivalent to

$$\mathcal{Q} : \sum_{i,j=1}^{n} q_{ij}x_i x_j + \sum_{i}^{n} b_i x_i + c = 0, \tag{7.2}$$

  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* (7.2) *of the quadric* $\mathcal{Q}$.
- The matrix Q defines a homogeneous polynomial of degree 2 in the above equation.
- Notice that Equation (7.2) can be rearranged in matrix form as follows

$$\mathcal{Q} : \mathbf{x}^T \cdot Q \cdot \mathbf{x} + \mathbf{b}^T \cdot \mathbf{x} + c = 0 \tag{7.3}$$

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

7.2 Reducing to the canonical form

Let $\mathcal{Q}$ be the hyperquadric described by Equation (7.2) with respect to the right oriented orthonormal coordinate system $\mathcal{K} = (O, \mathcal{B})$ where $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$. Let Q be the matrix associated to Equation (7.2) of $\mathcal{Q}$.

[Step 1 - Rotation] By the Spectral Theorem (see Corollary 3.24), there is an orthonormal basis $\mathcal{B}'$ of eigenvectors for Q which diagonalizes Q. Changing the coordinate system from $\mathcal{K} = (O, \mathcal{B})$ to $\mathcal{K}(O, \mathcal{B}')$, the equation of the hyperquadric becomes

$$\mathcal{Q} : \mathbf{y}^T \cdot D \cdot \mathbf{y} + \mathbf{v}^T \cdot \mathbf{y} + c = 0 \quad \text{where} \quad D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \tag{7.4}$$

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 $\mathrm{M}_{\mathcal{B},\mathcal{B}'}\mathbf{y}$ since $\mathbf{x} = \mathrm{M}_{\mathcal{B},\mathcal{B}'}\mathbf{y}$. Notice that, since the two bases $\mathcal{B}$ and $\mathcal{B}'$ are orthonormal, the matrix $\mathrm{M}_{\mathcal{B},\mathcal{B}'}$ is an orthogonal matrix, i.e. $\mathrm{M}_{\mathcal{B},\mathcal{B}'} \in \mathrm{O}(n)$.

For the proof of the Spectral theorem, we used the fact that, since Q is symmetric, the eigenvalues $\lambda_1,\dots,\lambda_n$ are real. Hence, eventually after permuting the basis vectors in $\mathcal{B}'$ 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 $\mathcal{B}' = (\mathbf{e}'_1,\dots,\mathbf{e}'_n)$ to $\mathcal{B}'' = (\mathbf{e}_{\sigma(1)},\dots,\mathbf{e}_{\sigma(n)})$ for some permutation σ of $\{1,\dots,n\}$. The base change matrix $\mathrm{M}_{\mathcal{B}',\mathcal{B}''}$ is again orthogonal. Thus, so far, we have a change of coordinates from $\mathcal{K} = (O,\mathcal{B})$ to $\mathcal{K} = (O,\mathcal{B}'')$ given by the base change matrix $\mathrm{M}_{\mathcal{B},\mathcal{B}''} = \mathrm{M}_{\mathcal{B},\mathcal{B}'}\mathrm{M}_{\mathcal{B}',\mathcal{B}''} \in \mathrm{O}(n)$.

Recall that $\det(\mathrm{M}_{\mathcal{B},\mathcal{B}''}) = 1$ if $\mathrm{M}_{\mathcal{B},\mathcal{B}''} \in \mathrm{SO}(n)$ and $\det(\mathrm{M}_{\mathcal{B},\mathcal{B}''}) = -1$ if $\mathrm{M}_{\mathcal{B},\mathcal{B}''}$ is not special orthogonal. In the latter case, the matrix $\mathrm{M}_{\mathcal{B},\mathcal{B}''}$ changes the orientation of the basis $\mathcal{B}''$, i.e. the change of coordinates is an indirect isometry. So, if we replace one vector in $\mathcal{B}''$ by minus that vector, for example $\mathbf{e}''_1 \leftrightarrow -\mathbf{e}''_1$, then $\mathcal{B}''$ has the same orientation as $\mathcal{B}$ and $\mathrm{M}_{\mathcal{B},\mathcal{B}'} \in \mathrm{SO}(n)$, i.e. the change of coordinates is a displacement. In conclusion, we may assume that $\mathcal{B}''$ is such that $\mathrm{M}_{\mathcal{B},\mathcal{B}''} \in \mathrm{SO}(n)$.

Writing out the equation of $\mathcal{Q}$ in $\mathcal{K}''$ we have:

$$\mathcal{Q} : \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_r y_r^2 + v_1 y_1 + v_2 y_2 + \cdots + v_n y_n + c = 0. \tag{7.5}$$

[Step 2 - translation] Notice that $r > 0$, i.e. there is at least one eigenvalue distinct from 0 otherwise $Q = 0_n$ (the $n \times n$ zero matrix) and (7.2) is not a quadratic polynomial. Now, 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 $i \in \{1,\dots,r\}$ we let $z_i = y_i + \frac{v_i}{2\lambda_i}$ and $z_i = y_i$ for $i > r$ then Equation (7.5) becomes

$$\mathcal{Q} : \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_r z_r^2 + v_{r+1} z_{r+1} + \cdots + v_n z_n = k \tag{7.6}$$

for some $k \in \mathbb{R}$. This change of variables corresponds to a change of coordinates from $\mathcal{K}'' = (O,\mathcal{B}'')$ to $\mathcal{K}''' = (O',\mathcal{B}'')$ given by the translation

$$\begin{bmatrix} z_1 \\ \vdots \\ z_r \\ z_{r+1} \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} \frac{v_1}{2\lambda_1} \\ \vdots \\ \frac{v_r}{2\lambda_r} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus, up until now, we changed the coordinates from $\mathcal{K}$ to $\mathcal{K}''$ with a direct isometry which fixes the origin (corresponding to a matrix in $\mathrm{SO}(n)$) and from $\mathcal{K}''$ to $\mathcal{K}'''$ with a translation. These are isometries, so the composition of these two transformations is an isometry.

[Step 3 - non-isometric deformation] It is possible to further simplify Equation (7.6) of $\mathcal{Q}$, by making other changes of coordinates which scale the axes such that all the coefficients are ± 1. However these will in general not correspond to isometries (See the discussion in Section 7.3.3).

We *reduced* Equation (7.2) to Equation (7.6). Equation (7.6) is not yet the canonical form but it is an important step towards the canonical form. We may refer to Equation (7.6) as an *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 $\mathbb{E}^3$.

7.3 Classification of conics

7.3.1 Isometric classification

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

$$\mathcal{C}: q_{11}x^2 + 2q_{12}xy + q_{22}y^2 + b_1x + b_2y + c = 0. \tag{7.7}$$

From the discussion in Section 7.2, we may apply a rotation and a translation to change the coordinate system such that Equation (7.7) becomes

$$\mathcal{C}: \lambda_1 x^2 + \lambda_2 y^2 = k \quad \text{or} \quad \mathcal{C}: \lambda_1 x^2 + v_2 y = k. \tag{7.8}$$

We may also assume that $\lambda_1 > 0$ since if $\lambda_1, \lambda_2 < 0$ we can multiply the whole equation by -1.

Case 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.
Case 2: If $\lambda_2 > 0$ and $k < 0$, then there are no real solutions to the equation.
Case 3: 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 rotating the reference frame with 90° which is a direct isometry.
Case 4: 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.
Case 5: 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 rotating the reference frame with 90° which is a direct isometry. Then we end up in the following case.
Case 6: 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 a hyperbola.
Case 7: If $\lambda_2 = 0$ we have the equation

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

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, $x^2 = 0$.

Case 8: If $v_2 = 0$ and $k < 0$ then we have no solutions.
Case 9: If $v_2 \neq 0$ then, dividing by $|v_2|$, the equation becomes

$$\frac{\lambda_1}{|v_2|}x^2 = \frac{k}{|v_2|} - \frac{v_2}{|v_2|}y$$

and we may change the coordinates with the translation/glide reflection $(x, \frac{k}{|v_1|} - \frac{v_2}{|v_2|}y) \to (x, y)$ such that the equation becomes

$$x^2 = \frac{|v_2|}{\lambda_1}y.$$

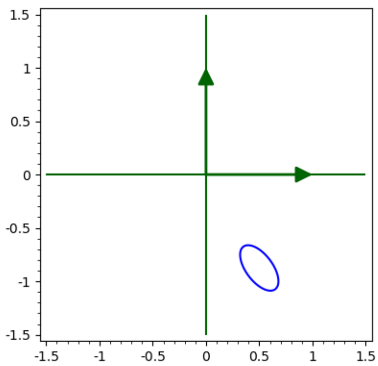
This change of coordinates corresponds to a reflection and a translation (again an isometric change of coordinates) and we recognize the equation of a parabola with parameter $p = \frac{|v_1|}{2\lambda_1}$. However, here again, we don't yet have the canonical form of a parabola. In order to obtain $\mathcal{P}_p$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

[Conclusion] Starting with an equation of the form (7.7), we may use rotations, translations and reflections in order to recognize that we obtain either

- degenerate cases: two lines, double lines, points, or
- non-degenerate cases:

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

Moreover, we can either carefully select direct isometries at each step, or, we can ensure at the end that the composition of all isometries used is direct. So, the curves described by an equation of the form (7.7) are conic sections and can be reduced to such curves via direct isomtries (displacements).



Example 7.2 (Ellipse)**.** Consider the curve with equation

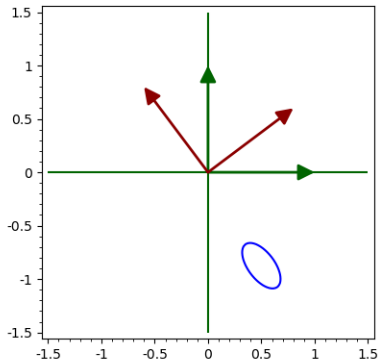
$$\mathcal{C} : 73x^2 + 72xy + 52y^2 - 10x + 55y + 25 = 0. \tag{7.9}$$

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 (7.9) becomes

$$\mathcal{C} : \begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} 73 & 36 \\ 36 & 52 \end{bmatrix}}_{Q} \begin{bmatrix} x \\ y \end{bmatrix} + \underbrace{\begin{bmatrix} -10 & 55 \end{bmatrix}}_{b} \begin{bmatrix} x \\ y \end{bmatrix} + 25 = 0. \tag{7.10}$$



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 $\mathcal{B}' = (\mathbf{e}_1'(4/5, 3/5), \mathbf{e}_2'(3/5, -4/5))$. The base change matrix from $\mathcal{B}'$ to $\mathcal{B}$ is

$$M_{\mathcal{B},\mathcal{B}'} = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}.$$

We know that this matrix is orthogonal, which in this example is easy to check directly. In particular $M_{\mathcal{B},\mathcal{B}'}^{-1} = M_{\mathcal{B},\mathcal{B}'}^{T}$. Moreover, the determinant is -1 which means that $M_{\mathcal{B},\mathcal{B}'}$ is not a direct isometrie, i.e. $M_{\mathcal{B},\mathcal{B}'} \in O(2) \setminus SO(2)$. 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 $\mathcal{B}' = (\mathbf{e}_1'(4/5, 3/5), \mathbf{e}_2'(-3/5, 4/5))$ we have

$$M_{\mathcal{B},\mathcal{B}'} = \frac{1}{5}\begin{bmatrix} 4 & -3 \\ 3 & 4 \end{bmatrix} \in SO(2).$$

Changing the coordinate system from $\mathcal{K}$ to $\mathcal{K}' = (O, \mathcal{B}')$, Equation (7.9) becomes

$$\mathcal{C} : \begin{bmatrix} x' & y' \end{bmatrix} M_{\mathcal{B},\mathcal{B}'}^{T}\, Q\, M_{\mathcal{B},\mathcal{B}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + b\, M_{\mathcal{B},\mathcal{B}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + 25 = 0$$

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

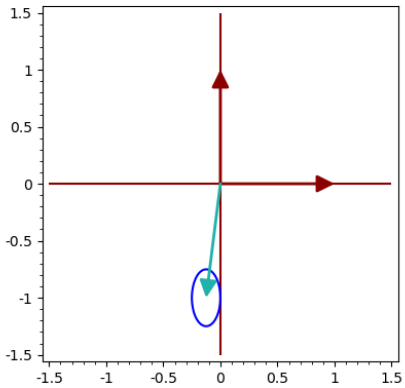
$$\mathcal{C} : 4x'^2 + y'^2 + x' + 2y' + 1 = 0$$

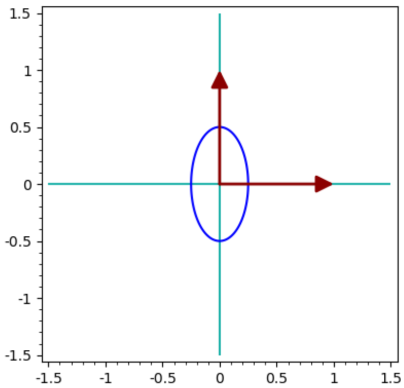
equivalently

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

equivalently

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





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

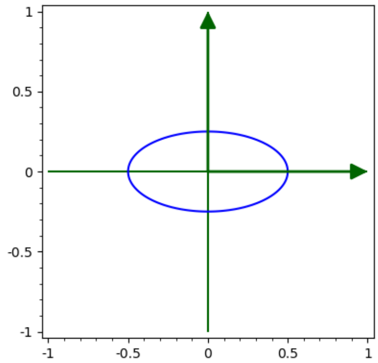
$$\mathcal{C} : 4x''^2 + y''^2 - \frac{1}{16} = 0 \quad \Leftrightarrow \quad \frac{x''^2}{\frac{1}{64}} + \frac{y''^2}{1} - 1 = 0.$$

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

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

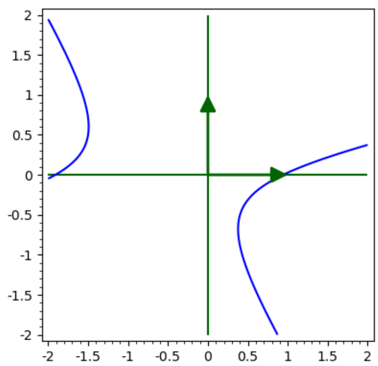
With any of the two transformations we obtain

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



To recap:

- We changed the coordinates from $\mathcal{K}$ to $\mathcal{K}'$ with the rotation $M_{\mathcal{B},\mathcal{B}'}$ 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.
- We obtained $\mathcal{E}_{1,\frac{1}{8}}$.



Example 7.3 (Hyperbola)**.** Consider the curve with equation

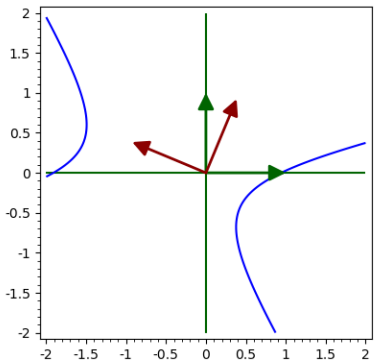
$$\mathcal{C} : -94x^2 + 360xy + 263y^2 - 91x + 221y + 169 = 0. \tag{7.11}$$

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 (7.11) becomes

$$\mathcal{C}:\begin{bmatrix}x & y\end{bmatrix}\underbrace{\begin{bmatrix}-94 & 180\\ 180 & 263\end{bmatrix}}_{Q}\begin{bmatrix}x\\ y\end{bmatrix}+\underbrace{\begin{bmatrix}-91 & 221\end{bmatrix}}_{b}\begin{bmatrix}x\\ y\end{bmatrix}+25=0. \tag{7.12}$$



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 $\mathcal{B}' = (\mathbf{e}_1'(5/13,12/13),\mathbf{e}_2'(12/13,-5/13))$. The base change matrix from $\mathcal{B}'$ to $\mathcal{B}$ is

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

We know that this matrix is orthogonal, which in this example is easy to check directly. In particular $\mathrm{M}_{\mathcal{B},\mathcal{B}'}^{-1} = \mathrm{M}_{\mathcal{B},\mathcal{B}'}^{T}$. Moreover, the determinant is -1 which means that $\mathrm{M}_{\mathcal{B},\mathcal{B}'}$ is not a direct isometrie, i.e. $\mathrm{M}_{\mathcal{B},\mathcal{B}'} \in \mathrm{O}(2)\backslash\mathrm{SO}(2)$. 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 $\mathcal{B}' = (\mathbf{e}_1'(5/13,12/13),\mathbf{e}_2'(-12/13,5/13))$ we have

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

Changing the coordinate system from $\mathcal{K}$ to $\mathcal{K}' = (O,\mathcal{B}')$, Equation (7.11) becomes

$$\mathcal{C}:\begin{bmatrix}x' & y'\end{bmatrix}\mathrm{M}_{\mathcal{B},\mathcal{B}'}^{T}\,Q\,\mathrm{M}_{\mathcal{B},\mathcal{B}'}\begin{bmatrix}x'\\ y'\end{bmatrix}+b\,\mathrm{M}_{\mathcal{B},\mathcal{B}'}\begin{bmatrix}x'\\ y'\end{bmatrix}+169=0$$

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

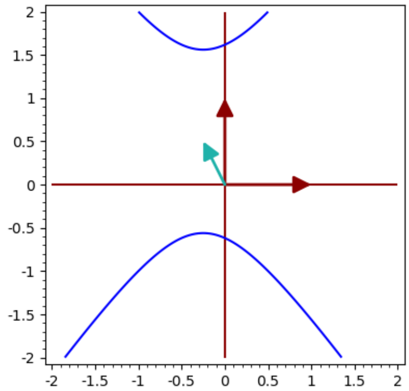
$$\mathcal{C}: 2x'^2 - y'^2 + x' + y' + 1 = 0$$

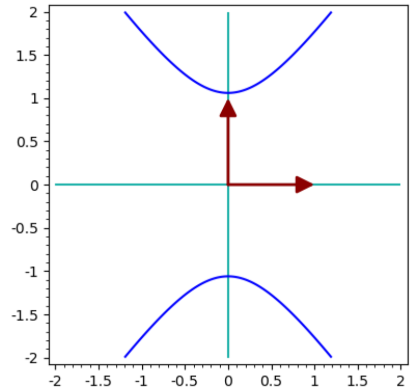
equivalently

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

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 $\mathcal{K}'' = (O'', \mathcal{B}'')$ where $O'' = (-\frac{1}{4}, \frac{1}{2})$ and the basis $\mathcal{B}'' = \mathcal{B}'$ doesn't change. In $\mathcal{K}''$ the equation of $\mathcal{C}$ becomes

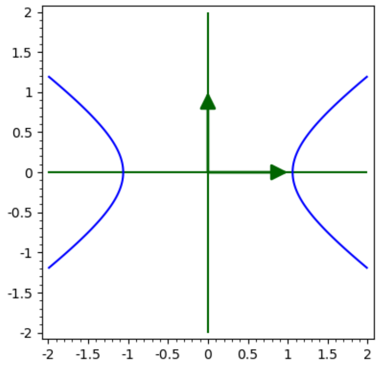
$$\mathcal{C}: 2x''^2 - y''^2 + \frac{9}{8} = 0 \quad \Leftrightarrow \quad -\frac{x''^2}{\frac{9}{16}} + \frac{y''^2}{\frac{9}{8}} - 1 = 0.$$

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

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

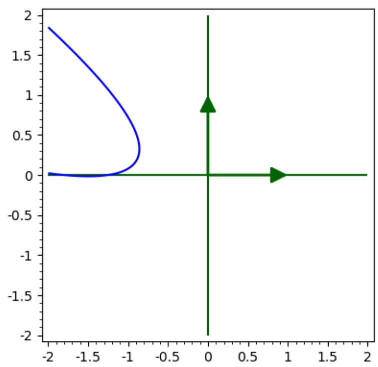
With any of the two transformations we obtain

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



To recap:

- We changed the coordinates from $\mathcal{K}$ to $\mathcal{K}'$ with the rotation $M_{\mathcal{B},\mathcal{B}'}$ 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.
- We obtained $\mathcal{H}_{\frac{3}{4},\frac{3}{\sqrt{2}}}$.



Example 7.4 (Parabola)**.** Consider the curve with equation

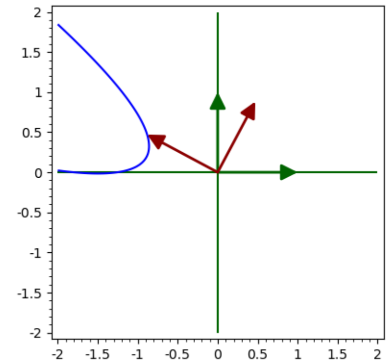
$$\mathcal{C} : 128x^2 + 480xy + 450y^2 + 391x + 119y + 289 = 0. \tag{7.13}$$

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 (7.13) becomes

$$\mathcal{C} : \begin{bmatrix} x & y \end{bmatrix} \underbrace{\begin{bmatrix} 128 & 240 \\ 240 & 225 \end{bmatrix}}_{Q} \begin{bmatrix} x \\ y \end{bmatrix} + \underbrace{\begin{bmatrix} 391 & 119 \end{bmatrix}}_{b} \begin{bmatrix} x \\ y \end{bmatrix} + 25 = 0. \tag{7.14}$$



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 $\mathcal{B}' = (\mathbf{e}_1'(8/17,15/17), \mathbf{e}_2'(15/17,-8/17))$. The base change matrix from $\mathcal{B}'$ to $\mathcal{B}$ is

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

We know that this matrix is orthogonal, which in this example is easy to check directly. In particular $\mathrm{M}_{\mathcal{B},\mathcal{B}'}^{-1} = \mathrm{M}_{\mathcal{B},\mathcal{B}'}^{T}$. Moreover, the determinant is -1 which means that $\mathrm{M}_{\mathcal{B},\mathcal{B}'}$ is not a direct isometrie, i.e. $\mathrm{M}_{\mathcal{B},\mathcal{B}'} \in \mathrm{O}(2)\setminus\mathrm{SO}(2)$. 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 $\mathcal{B}' = (\mathbf{e}_1'(8/17,15/17), \mathbf{e}_2'(-15/17,8/17))$ we have

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

Changing the coordinate system from $\mathcal{K}$ to $\mathcal{K}' = (O,\mathcal{B}')$, Equation (7.13) becomes

$$\mathcal{C} : \begin{bmatrix} x' & y' \end{bmatrix} \mathrm{M}_{\mathcal{B},\mathcal{B}'}^{T}\, Q\, \mathrm{M}_{\mathcal{B},\mathcal{B}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + b\,\mathrm{M}_{\mathcal{B},\mathcal{B}'} \begin{bmatrix} x' \\ y' \end{bmatrix} + 289 = 0$$

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

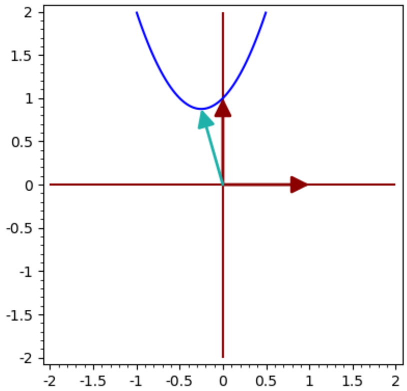
$$\mathcal{C} : 2x'^2 + x' - y' + 1 = 0$$

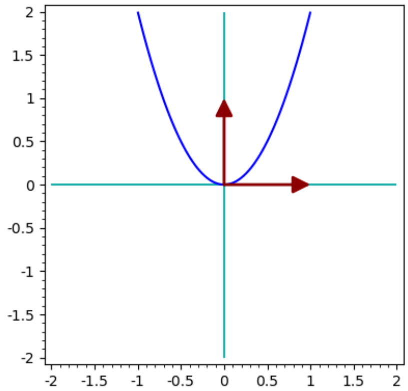
equivalently

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

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 $\mathcal{K}'' = (O'', \mathcal{B}'')$ where $O'' = (-\frac{1}{4}, \frac{7}{8})$ and the basis $\mathcal{B}'' = \mathcal{B}'$ doesn't change. In $\mathcal{K}''$ the equation of $\mathcal{C}$ becomes

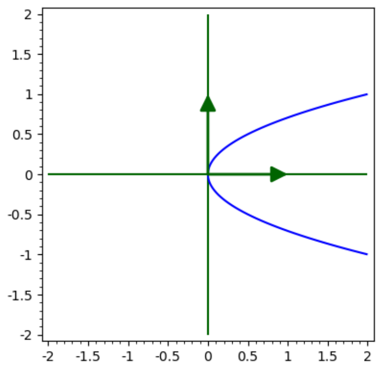
$$\mathcal{C} : 2x''^2 - y'' = 0 \quad \Leftrightarrow \quad x''^2 = \frac{1}{2}y''.$$

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

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

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 $\mathrm{M}_{\mathcal{B},\mathcal{B}'}$ 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.
- We obtained $\mathcal{P}_{\frac{1}{4}}$.

7.3.2 Algorithm 1: Isometric invariants

The discussion in Subsections 7.3.1 is an algebraic case-by-case analysis which may appear lengthy. It is the honest way of doing mathematics. However, it may be difficult to perform such steps.

If we only want to know what type of conic we are dealing with, it is not difficult to extract a recipt which allows us to decide what type of curve a given quadratic equation describes. For this, notice that we may write the Equation (7.1) of $\mathcal{Q}$ in a slightly different matrix form. We only do this in dimension 2, so consider the quadratic curve:

$$\mathcal{C} : q_{11}x^2 + 2q_{12}xy + q_{22}y^2 + 2b_1x + 2b_2y + c = 0. \tag{7.15}$$

The equation can be rewritten as

$$\begin{bmatrix} x & y & 1 \end{bmatrix} \underbrace{\begin{bmatrix} q_{11} & q_{12} & b_1 \\ q_{21} & q_{22} & b_2 \\ b_1 & b_2 & c \end{bmatrix}}_{=\widehat{Q}} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0$$

The matrix $\widehat{Q}$ is symmetric and we call it the *extended symmetric matrix associated to Equation* (7.15) *of the conic* $\mathcal{C}$. Notice that changing coordinates with a rotation R [Step 1] followed by a translation with a vector $\mathbf{v}$ [Step 2] amounts to a single matrix multiplication:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} R & \mathbf{v} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \Leftrightarrow \quad \underbrace{\begin{bmatrix} R^{-1} & -R^{-1}\mathbf{v} \\ 0 & 1 \end{bmatrix}}_{=M} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

after which the equation becomes:

$$\begin{bmatrix} x' & y' & 1 \end{bmatrix} M^T \widehat{Q} M \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = 0.$$

Observe that $\det(\widehat{Q})$ *does not change* when reducing to the canonical form with displacements. We say that $\det(\widehat{Q})$ is *invariant*. Indeed $\det(M) = \det(R^{-1}) = 1$ so it doesn't affect the determinant of $\widehat{Q}$:

$$\det(M^T\widehat{Q}M) = \det(M^T)\det(\widehat{Q})\det(M) = \det(\widehat{Q}).$$

In order to distinguish between the different conics, we use three invariants. As before, Q denotes the symmetric matrix associated to Equation (7.15) of the conic $\mathcal{C}$. The invariants are:

$$\widehat{D} = \det(\widehat{Q}), \quad D = \det(Q) \quad \text{and} \quad T = \operatorname{tr}(Q).$$

By checking Cases 1-9 in Section 7.3.1 and because $\widehat{D}$, D and T do not change under displacements, we have the following result:

Proposition 7.5. The type of curve described by Equation (7.15) is given with the following table.

$\widehat{D}$	D	T	curve C
$\widehat{D}=0$	$D>0$	-	A point
	$D=0$	-	Two lines or the empty set
	$D<0$	-	Two lines
$\widehat{D}\neq 0$	$D>0$	$\widehat{D}T<0$	An ellipse
	$D>0$	$\widehat{D}T>0$	The empty set
	$D=0$	-	A parabola
	$D<0$	-	A hyperbola

Table 7.1: Classification in dimension 2 via invariants.

7.3.3 Affine classification

In the isometric classification, we allowed only changes of coordinates which are isometries. In fact we noticed that it suffices to use direct isometries in order to reduce to the canonical form. 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.

Example 7.6. 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 (7.6) in general for hypersurfaces. However, 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. Here, the second column indicates the signature of Q.

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

Table 7.2: Affine classification in dimension 2.

7.3.4 Algorithm 2: Lagrange's method

One reason the discussion in Subsections 7.3.1 is lengthy is the interpretation that we added to each step (rotation, translation, eigenvectors give the directions of the new axes,...etc). We did this because we aimed at changing the coordinate system via displacements, without metrically altering the objects.

If we are only interested in the shape of the object, we are allowed to use affine transformations, which are changes of coordinates. This allows us to focus just on the algebra that is behind the calculations and 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:

$$\mathcal{C} : q_{11}x^2 + 2q_{12}xy + q_{22}y^2 + b_1x + b_2y + c = 0. \tag{7.16}$$

(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

$$ax^2 + by^2 + c = 0 \quad \text{or} \quad ax^2 + by + c = 0.$$

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 7.2).

The reason this works is because Step 1 and Step 2 correspond to affine changes of coordinates.

7.4 Classification of quadrics

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

$$\mathcal{C} : q_{11}x^2 + q_{22}y^2 + q_{33}z^2 + 2q_{12}xy + 2q_{12}yz + 2q_{12}xz + b_1x + b_2y + b_3z + c = 0. \tag{7.17}$$

From the discussion in Section 7.2, we may apply an orthogonal change of coordinates and a translation to change the coordinate system such that Equation (7.17) becomes

$$\mathcal{Q}: \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_r x_r^2 + v_{r+1} x_{r+1} + \cdots + v_n x_n = k. \tag{7.18}$$

The isometric 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 $M_{\mathcal{B},\mathcal{B}''}$ used to obtain Equation (7.18), is an element of the group SO(3). So, by Euler's theorem (Theorem 5.31), this is indeed a rotation around an axis.

However, as in Section 7.3.3 and Section 7.3.4, 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 (7.17) is one of the possibilities listed in the following table.

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

7.5 Exercises

7.1. For each of the equations in Table 7.2, discuss the geometric locus of points satisfying them.

7.2. For each of the following matrices A, write down a quadratic equation with associated matrix A and find the matrix $M \in \mathrm{SO}(2)$ which diagonalizes A.

a) $\begin{bmatrix} 6 & 2 \\ 2 & 9 \end{bmatrix}$

b) $\begin{bmatrix} 5 & -13 \\ -15 & 5 \end{bmatrix}$

c) $\begin{bmatrix} 7 & -2 \\ -2 & 5/3 \end{bmatrix}$

7.3. Check the calculations in Examples 7.2, 7.3 and 7.4.

7.4. For each of the following equations write down the associated matrix and bring the equation in canonical form.

a) $-x^2 + xy - y^2 = 0$,

b) $6xy + x - y = 0$.

7.5. In each of the following cases, decide the type of the quadratic curve based on the parameter $a \in \mathbb{R}$.

a) $x^2 - 4xy + y^2 = a$,

b) $x^2 + 4xy + y^2 = a$.

7.6. Consider the rotation R_{90° of $\mathbb{E}^2$ around the origin and the translation $T_{\mathbf{v}}$ of $\mathbb{E}^2$ with vector $\mathbf{v}(1,0)$.

a) Give the algebraic form of the isometries R_{90°, $T_{\mathbf{v}}$ and $T_{\mathbf{v}} \circ R_{90^\circ}$.

b) Determine the equations of the hyperbola $\mathcal{H} : \frac{x^2}{4} - \frac{y^2}{9} - 1 = 0$ and the parabola $\mathcal{P} : y^2 - 8x = 0$ after transforming them with R_{90° and with $T_{\mathbf{v}} \circ R_{90^\circ}$ respectively.

7.7. Find the canonical equation for each of the following cases

a) $5x^2 + 4xy + 8y^2 - 32x - 56y + 80 = 0$,

b) $8y^2 + 6xy - 12x - 26y + 11 = 0$,

c) $x^2 - 4xy + y^2 - 6x + 2y + 1 = 0$.

7.8. For each of the conics in the previous exercise, indicate the affine change of coordinates which brings the equation in canonical form.

7.9. Discuss the type of the curve

$$x^2 + \lambda xy + y^2 - 6x - 16 = 0$$

in terms of $\lambda \in \mathbb{R}$.

7.10. Using the classification of quadrics, decide what surfaces are described by the following equations.

a) $x^2 + 2y^2 + z^2 + xy + yz + zx = 1$,

b) $xy + yz + zx = 1$,

c) $x^2 + xy + yz + zx = 1$,

d) $xy + yz + zx = 0$.

CHAPTER 8

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

$$\mathcal{S} : q_{11}x^2 + q_{22}y^2 + q_{33}z^2 + 2q_{12}xy + 2q_{13}xz + 2q_{23}yz + a_1x + a_2y + a_3z + c = 0. \tag{8.1}$$

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.

8.1 Ellipsoid

8.1.1 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 $a, b, c \in \mathbb{R}$.

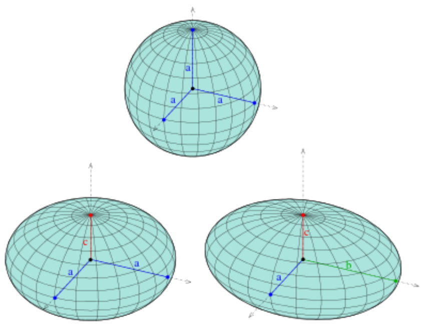


Figure 8.1: Ellipsoid[1]

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.

8.1.2 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

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

[1] Image source: Wikipedia

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.

8.1.3 Parametrizations - local description

A parametrization of this surface is

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

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

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

8.2 Elliptic Cone

8.2.1 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 $a,b,c \in \mathbb{R}$.

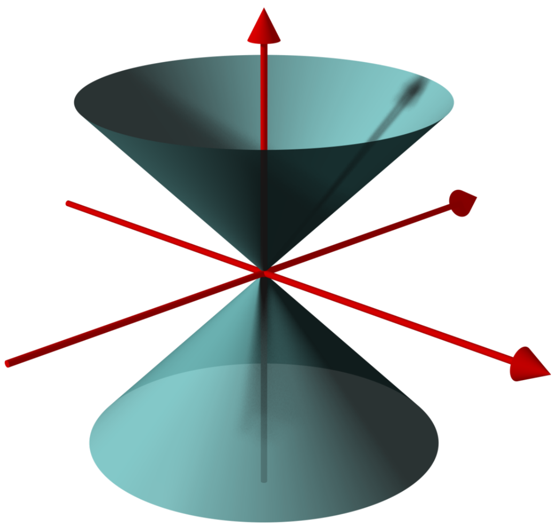


Figure 8.2: Elliptic cone[2]

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

8.2.2 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 8.1. The intersection of an elliptic cone with an arbitrary plane is a (possibly degenerate) quadratic curve.

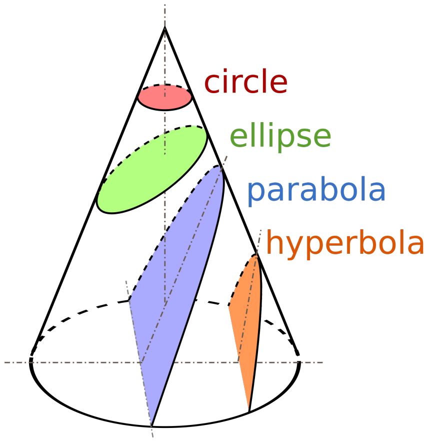


Figure 8.3: Conic sections[3]

8.2.3 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

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

8.2.4 $\mathcal{C}_{a,b,c}$ 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.

8.2.5 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:

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

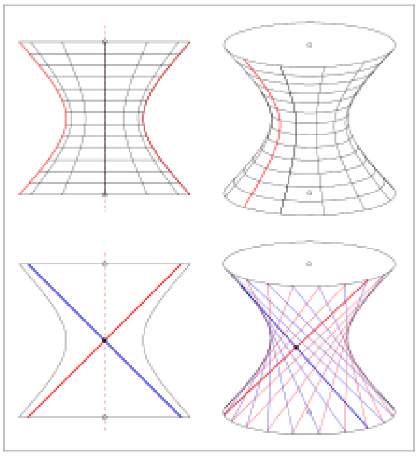
8.3 Hyperboloid of one sheet

8.3.1 Canonical equation - global description

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

$$\mathcal{H}^1_{a,b,c} : \underbrace{\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}}_{\varphi(x,y,z)} = 1 \quad \text{so} \quad \mathcal{H}^1_{a,b,c} = \varphi^{-1}(1) \tag{8.2}$$

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.

8.3.2 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

$$\varphi\left(\begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}\right) = 1 \quad \Leftrightarrow \quad \frac{(x_p + tv_x)^2}{a^2} + \frac{(y_p + tv_y)^2}{b^2} - \frac{(z_p + tv_z)^2}{c^2} - 1 = 0. \tag{8.3}$$

[4] Image source: Wikipedia

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

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

$$\Leftrightarrow\quad \underbrace{\Big(\frac{v_x^2}{a^2}+\frac{v_y^2}{b^2}-\frac{v_y^2}{c^2}\Big)}_{=0?}t^2+2\Big(\frac{x_pv_x}{a^2}+\frac{y_pv_y}{b^2}-\frac{z_pv_z}{c^2}\Big)t=0. \tag{8.4}$$

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

$$\frac{v_x^2}{a^2}+\frac{v_y^2}{b^2}-\frac{v_y^2}{c^2}=0 \tag{8.5}$$

then equation (8.4) 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 (8.5), 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 8.5).

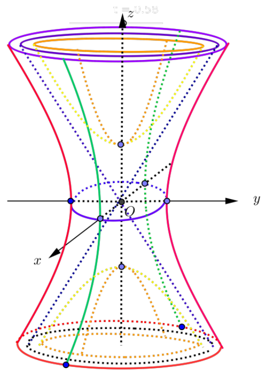


Figure 8.5: Hyperboloid and asymptotic cone[5]

This should help to see that, when the vector $v = (v_x, v_y, v_z)$ satisfies equation (8.5), 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 (8.3) without the assumption that p lies on the surface $\mathcal{H}^1_{a,b,c}$. How would you do this?

8.3.3 $\mathcal{H}^1_{a,b,c}$ as ruled surface

Here is a fact: the hyperboloid with one sheet is a *doubly ruled surface* (this is visible in Figure 8.4a). 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 (8.2):

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad \Leftrightarrow \quad \frac{x^2}{a^2} - \frac{z^2}{c^2} = 1 - \frac{y^2}{b^2} \quad \Leftrightarrow \quad \left(\frac{x}{a} - \frac{z}{c}\right)\left(\frac{x}{a} + \frac{z}{c}\right) = \left(1 - \frac{y}{b}\right)\left(1 + \frac{y}{b}\right) \tag{8.6}$$

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 (8.6) 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 $\mathcal{L}_2$.

8.3.4 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. $p = p(x(\theta_{1,p}, \theta_{2,p}), y(\theta_{1,p}, \theta_{2,p}), z(\theta_{1,p}, \theta_{2,p}))$.

8.4 Hyperboloid of two sheets

8.4.1 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 $a, b, c \in \mathbb{R}$.

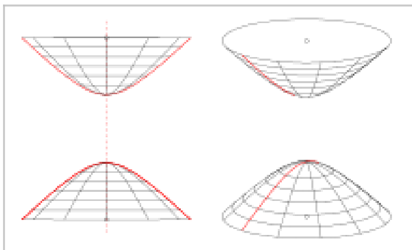


Figure 8.6: Hyperboloid of two sheets[6]

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.

8.4.2 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

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

$$\Leftrightarrow\quad \underbrace{\Big(\frac{v_x^2}{a^2}+\frac{v_y^2}{b^2}-\frac{v_y^2}{c^2}\Big)}_{=0?}t^2+2\Big(\frac{x_pv_x}{a^2}+\frac{y_pv_y}{b^2}-\frac{z_pv_z}{c^2}\Big)t=0. \tag{8.7}$$

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

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

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

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

Deduce this equation also with the gradient.

How do we interpret the first condition? If

$$\frac{v_x^2}{a^2}+\frac{v_y^2}{b^2}-\frac{v_y^2}{c^2}=0 \tag{8.8}$$

then equation (8.7) 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 (8.8), 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 8.5). 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.

8.4.3 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 $\varepsilon\in\{\pm1\}$. 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.

8.5 Elliptic paraboloid

8.5.1 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 $a,b\in\mathbb{R}$.

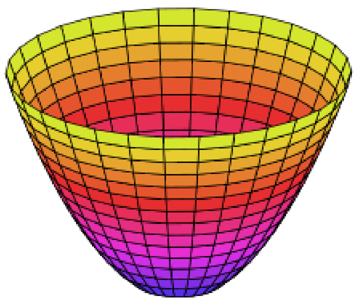


Figure 8.7: Elliptic paraboloid[7]

[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.

8.5.2 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

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

$$\Leftrightarrow \quad \underbrace{\Big(\frac{v_x^2}{a} + \frac{v_y^2}{b}\Big)}_{\neq 0}t^2 + 2\Big(\frac{x_p v_x}{a} + \frac{y_p v_y}{b} - v_z\Big)t = 0. \tag{8.9}$$

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

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

Deduce this equation also with the gradient.

8.5.3 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. $p = p(x(\theta_{1,p},\theta_{2,p}), y(\theta_{1,p},\theta_{2,p}), z(\theta_{1,p},\theta_{2,p}))$.

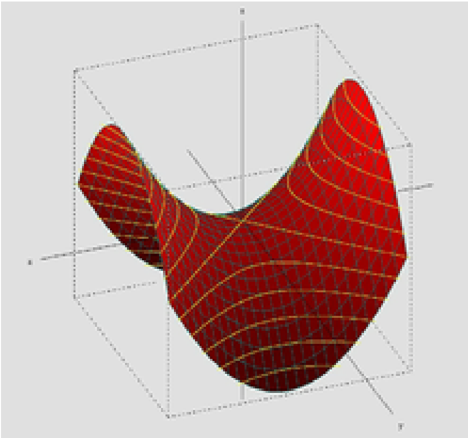
8.6 Hyperbolic paraboloid

8.6.1 Canonical equation - global description

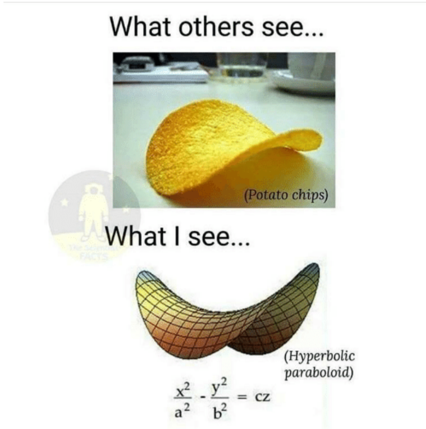
A *hyperbolic paraboloid* is a surface which (in some coordinate system) satisfies an equation of the form

$$\mathcal{P}^h_{a,b} : \underbrace{\frac{x^2}{a} - \frac{y^2}{b} - 2z}_{\varphi(x,y,z)} = 0 \quad \text{so} \quad \mathcal{P}^h_{a,b} = \varphi^{-1}(0) \tag{8.10}$$

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.

8.6.2 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: the internet

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

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

$$\Leftrightarrow\quad \underbrace{\Big(\frac{v_x^2}{a}-\frac{v_y^2}{b}\Big)}_{=0?}t^2+2\Big(\frac{x_pv_x}{a}-\frac{y_pv_y}{b}-v_z\Big)t=0. \tag{8.11}$$

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

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

Deduce this equation also with the gradient.

How do we interpret the first condition? If

$$\frac{v_x^2}{a}-\frac{v_y^2}{b}=0\quad\Leftrightarrow\quad\Big(\frac{v_x}{\sqrt{a}}-\frac{v_y}{\sqrt{b}}\Big)\Big(\frac{v_x}{\sqrt{a}}+\frac{v_y}{\sqrt{b}}\Big)=0 \tag{8.12}$$

then equation (8.11) 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 (8.12), 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 (8.12). 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 8.8a). 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 (8.11), 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).

8.6.3 $\mathcal{P}^h_{a,b}$ 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 8.8a.

Again, one way to see where the two families of lines come from is to rearrange (8.10)

$$\frac{x^2}{a} - \frac{y^2}{b} - 2z = 0 \quad \Leftrightarrow \quad \frac{x^2}{a} - \frac{y^2}{b} = 2z \quad \Leftrightarrow \quad \left(\frac{x}{\sqrt{a}} - \frac{y}{\sqrt{b}}\right)\left(\frac{x}{\sqrt{a}} + \frac{y}{\sqrt{b}}\right) = 2z \tag{8.13}$$

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 8.8a, since it lies in the plane $z = 0$. The other one, $\tilde{l}_\infty$, belongs to the family $\mathcal{L}_2$.

8.6.4 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. $p=p(x(\theta_{1,p},\theta_{2,p}),y(\theta_{1,p},\theta_{2,p}),z(\theta_{1,p},\theta_{2,p}))$.

8.7 Exercises

8.1. Determine the intersection of the ellipsoid

$$\mathcal{E}_{4,2\sqrt{3},2} : \frac{x^2}{16} + \frac{y^2}{12} + \frac{z^2}{4} - 1 = 0 \quad \text{with the line} \quad \ell = \begin{bmatrix} 4 \\ -6 \\ -2 \end{bmatrix} + \langle \begin{bmatrix} 2 \\ -3 \\ -2 \end{bmatrix} \rangle.$$

Write down the equations of the tangent planes in the intersection points.

8.2. Determine the intersection of the ellipsoid

$$\mathcal{E}_{2,3,4} : \frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{16} = 1$$

with planes parallel to the coordinate planes. Treat the various cases separately.

8.3. Determine the intersection of the ellipsoid

$$\mathcal{E}_{2,\sqrt{3},3} : \frac{x^2}{4} + \frac{y^2}{3} + \frac{z^2}{9} = 1 \quad \text{with the line} \quad \ell : x = y = z.$$

Write down the equations of the tangent planes in the intersection points.

8.4. Determine the tangent planes to the ellipsoid

$$\mathcal{E}_{2,3,2\sqrt{2}} : \frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{8} = 1$$

which are parallel to the plane $\pi : 3x - 2y + 5z + 1 = 0$.

8.5. Determine the points P of the ellipsoid

$$\mathcal{E} : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

for which the tangent space $T_P\mathcal{E}$ intersects the coordinate axis in congruent segments.

8.6. Show that the line

$$\begin{bmatrix} 2 \\ -3 \\ 6 \end{bmatrix} + \langle \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} \rangle \quad \text{is tangent to the quadric} \quad \frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{16} - 1 = 0$$

and determine the tangency point.

8.7. Prove that the intersection of a quadric in $\mathbb{E}^3$ with a plane is either the empty set or a point or a line or two lines or an ellipse or a hyperbola or a parabola.

8.8. Prove that the intersection of an ellipsoid with a plane is either the empty set or a point or an ellipse.

8.9. Show that the ellipsoid $\mathcal{E}_{a,b,b}$ is the locus of points for which the sum of the distances to two given points is constant. Such a surface is called *ellipsoid of revolution*.

8.10. Use a parametrization of an ellipse and a rotation matrix to deduce a parametrization of an ellipsoid of revolution.

8.11. For the surface $\mathcal{S}$ with parametrization

$$\mathcal{S}: \begin{cases} x = 4\cos(s)\cos(t) \\ y = 4\sin(s)\cos(t) \\ z = 2\sin(t) \end{cases} \qquad s \in [0, 2\pi[\quad t \in [-\frac{\pi}{2}, \frac{\pi}{2}[$$

a) Give an equation of $\mathcal{S}$.

b) Find the parameters of the point $P(3, \sqrt{3}, 1)$.

c) Calculate a parametrization of the tangent plane $T_P\mathcal{S}$ using partial derivatives.

d) Give an equation of $T_P\mathcal{S}$.

8.12. Prove that the intersection of an elliptic cone with a plane is either a point or a line or an ellipse or a hyperbola or a parabola.

8.13. Determine the intersection of the hyperboloid

$$\mathcal{H}^1_{4,3,1} : \frac{x^2}{16} + \frac{y^2}{9} - \frac{z^2}{1} = 1 \quad \text{with the line} \quad \ell = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix} + \langle \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} \rangle.$$

Write down the equations of the tangent planes in the intersection points.

8.14. Determine the tangent plane of the hyperboloid

$$\mathcal{H}^1_{2,3,1} : \frac{x^2}{4} + \frac{y^2}{9} - \frac{z^2}{1} = 1$$

in the point $M(2, 3, 1)$. Show that the tangent plane intersects the surface in two lines.

8.15. Determine the generators of the hyperboloid

$$\frac{x^2}{36} + \frac{y^2}{9} - \frac{x^2}{4} = 1$$

which are parallel to the plane $x + y + z = 0$.

8.16. Determine the intersection of the hyperboloid

$$\mathcal{H}^2_{2,1,3} : \frac{x^2}{4} + \frac{y^2}{1} - \frac{z^2}{9} = -1 \quad \text{with the line} \quad \ell = \begin{bmatrix} 3 \\ 1 \\ 6 \end{bmatrix} + \langle \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} \rangle.$$

Write down the equations of the tangent planes in the intersection points.

8.17. Determine the intersection of the paraboloid

$$\mathcal{P}^h_{2,\frac{1}{2}} : x^2 - 4y^2 = 4z \quad \text{with the line} \quad \ell = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix} + \langle \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} \rangle.$$

Write down the equations of the tangent planes in the intersection points.

8.18. Determine the tangent plane of

a) the elliptic paraboloid $\frac{x^2}{5} + \frac{y^2}{3} = z$ and of

b) the hyperbolic paraboloid $x^2 - \frac{y^2}{4} = z$

which are parallel to the plane $x - 3y + 2z - 1 = 0$.

8.19. Determine the plane which contains the line

$$\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + \langle \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \rangle \quad \text{and is tangent to the quadric} \quad x^2 + 2y^2 - z^2 + 1 = 0.$$

8.20. Show that the parabolid $\mathcal{P}^e_{p,p}$ is the locus of points for which the distance from a point equals the distance to a plane. Such a surface is called *elliptic paraboloid of revolution*.

8.21. Use a parametrization of a parabola and a rotation matrix to deduce a parametrization of an elliptic paraboloid of revolution.

8.22. For the surface $\mathcal{S}$ with parametrization

$$\mathcal{S} : \begin{cases} x = \sqrt{1+t^2}\cos(s) \\ y = \sqrt{1+t^2}\sin(s) \\ z = 2t \end{cases}$$

a) Give the equation of $\mathcal{S}$.

b) Find the parameters of the point $P(1,1,2)$.

c) Calculate a parametrization of the tangent plane $T_P\mathcal{S}$ using partial derivatives.

d) Give the equation of $T_P\mathcal{S}$.

8.23. Determine the generators of the paraboloid

$$\frac{x^2}{16} - \frac{y^2}{4} = z$$

which are parallel to the plane $3x + 2y - 4z = 0$.

8.24. Which of the following is a hyperboliod?

a) $\mathcal{S} : 2xz + 2xy + 2yz = 1$

b) $\mathcal{S} : 5x^2 + 3y^2 + xz = 1$

c) $\mathcal{S} : 2xy + 2yz + y + z = 2$

APPENDIX A

Bilinear forms

Throughout we let $\mathbb{V}$ denote a real vector space. The proofs for the statements in this appendix can be found in [6, Chapter 15 and 16].

Definition A.1. A map $\phi : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ is called a *bilinear map* (or *bilinear form*) if it is linear in both arguments, i.e.

$$\phi(a\mathbf{u} + b\mathbf{v}, \mathbf{w}) = a\phi(\mathbf{u}, \mathbf{w}) + b\phi(\mathbf{v}, \mathbf{w}) \quad \text{and} \quad \phi(\mathbf{u}, a\mathbf{v} + b\mathbf{w}) = a\phi(\mathbf{u}, \mathbf{v}) + b\phi(\mathbf{u}, \mathbf{w}).$$

for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{V}$ and all $a, b \in \mathbb{R}$.

With respect to a basis $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ of $\mathbb{V}$, the values of a bilinear form ϕ are determined by its values on the basis vectors. Thus, ϕ is determined by its Gram matrix $\mathrm{G}_{\mathcal{B}}(\phi) := \phi(\mathbf{e}_i, \mathbf{e}_j)$ since

$$\phi(\mathbf{v}, \mathbf{w}) = [\mathbf{v}]_{\mathcal{B}}^T \cdot \mathrm{G}_{\mathcal{B}}(\phi) \cdot [\mathbf{w}]_{\mathcal{B}}.$$

Moreover if $\mathcal{B}'$ is another basis, then

$$\mathrm{G}_{\mathcal{B}'}(\phi) = \mathrm{M}_{\mathcal{B},\mathcal{B}'}^T \cdot \mathrm{G}_{\mathcal{B}}(\phi) \cdot \mathrm{M}_{\mathcal{B},\mathcal{B}'}$$

in particular we obtain the following result.

Proposition A.2. Let ϕ be a bilinear form on the vector space $\mathbb{V}$. The rank of $\mathrm{G}_{\mathcal{B}}(\phi)$ does not depend on $\mathcal{B}$. It depends only on ϕ and we denote it by $\mathrm{rank}(\phi)$.

Definition A.3. Let ϕ be a bilinear form on the vector space $\mathbb{V}$. The *quadratic form associated to* ϕ is the map

$$q_\phi : \mathbb{V} \to \mathbb{R} \quad \text{defined by} \quad q_\phi(\mathbf{v}) = \phi(\mathbf{v}, \mathbf{v}).$$

Proposition A.4. Let ϕ be a symmetric bilinear form on the vector space $\mathbb{V}$. The quadratic form q_ϕ associated to ϕ satisfies

$$\begin{array}{rcl} q_\phi(\lambda \mathbf{v}) & = & \lambda^2 q_\phi(\mathbf{v}) \\ 2\phi(\mathbf{v},\mathbf{w}) & = & q(\mathbf{v}+\mathbf{w}) - q(\mathbf{v}) - q(\mathbf{w}) \end{array}$$

for every $\lambda \in \mathbb{R}$ and every $\mathbf{v}, \mathbf{w} \in \mathbb{V}$.

In particular, Proposition A.4 implies that the quadratic form q_ϕ determines uniquely the symmetric bilinear form ϕ, i.e. the correspondence $\phi \leftrightarrow q_\phi$ is a bijection between symmetric bilinear forms and quadratic forms.

Definition A.5. Let ϕ be a symmetric bilinear form on the vector space $\mathbb{V}$. A *diagonalizing basis* for ϕ is a basis $\mathcal{B}$ of $\mathbb{V}$ such that $\mathrm{G}_\mathcal{B}(\phi)$ is a diagonal matrix.

Theorem A.6. Let ϕ be a symmetric bilinear form on the vector space $\mathbb{V}$. Then there exists a diagonalizing basis for ϕ.

Theorem A.7. (Sylvester) Let ϕ be a symmetric bilinear form of rank r on the vector space $\mathbb{V}$. Then there is an integer p depending only on ϕ and a basis $\mathcal{B}$ of $\mathbb{V}$ such that

$$\mathrm{G}_\mathcal{B}(\phi) = \begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where 0 stands for zero matrices of appropriate sizes.

APPENDIX B

Quaternions and rotations

Contents

B.1 Algebraic considerations

Some of the aspects considered here are also covered in [5, Section 4.4].

Definition B.1. Denote the standard basis of $\mathbb{R}^4$ by $1, \mathbf{i}, \mathbf{j}, \mathbf{k}$ and consider the bilinear map

$$\_\cdot\_ : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}^4$$

given on the basis vectors by

$$\begin{array}{c|cccc} & 1 & \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \hline 1 & 1 & \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \mathbf{i} & \mathbf{i} & -1 & \mathbf{k} & -\mathbf{j} \\ \mathbf{j} & \mathbf{j} & -\mathbf{k} & -1 & \mathbf{i} \\ \mathbf{k} & \mathbf{k} & \mathbf{j} & -\mathbf{i} & -1 \end{array}\;.$$

We denote $\mathbb{R}^4$ with the above multiplication by $\mathbb{H}$. The elements of $\mathbb{H}$ are called *quaternions*. The product is the *Hamilton product*.

Remark. From the definition we observe that

1. The multiplication map on arbitrary quaternions $p = a_1 + b_1\mathbf{i} + c_1\mathbf{j} + d_1\mathbf{k}$ and $q = a_2 + b_2\mathbf{i} + c_2\mathbf{j} + d_2\mathbf{k}$ is
$$\begin{aligned} pq &= (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) + (a_1b_2 + a_2b_1 + c_1d_2 - c_2d_1)\mathbf{i} \\ &+ (a_1c_2 + a_2c_1 - b_1d_2 + b_2d_1)\mathbf{j} + (a_1d_2 + a_2d_1 + b_1c_2 - b_2c_1)\mathbf{k} \end{aligned} \tag{B.1}$$

2. Direct calculations show that $\mathbb{H}$ is an algebra, usually called *quaternion algebra*.

3. $\mathbb{H}$ is not commutative, $\mathbf{i} \cdot \mathbf{j} = \mathbf{k} = -\mathbf{j} \cdot \mathbf{i}$.

4. $\mathbb{R} \cdot 1$ is a subfield of $\mathbb{H}$ so we just write $\mathbb{R}$ for it.

5. $\mathbb{C} = \mathbb{R} \cdot 1 + \mathbb{R} \cdot \mathbf{i}$ is a subfield of $\mathbb{H}$.

Definition B.2. For a quaternion $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$, a is the *real part* $\mathfrak{Re}(q)$ *of* q and $b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ the *imaginary part* $\mathfrak{Im}(q)$ *of* q. We say that q is *real* if it equals its real part. We say that q is *purely imaginary* if it equals its imaginary part.

Proposition B.3. A quaternion is real if and only if it commutes with all quaternions, i.e. the center of $\mathbb{H}$ is $\mathbb{R}$.

Proposition B.4. A quaternion is purely imaginary if and only if its square is real and non-pozitive.

Definition B.5. For a quaternion $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$, the *conjugate of* q is

$$\overline{q} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k} = \mathfrak{Re}(q) - \mathfrak{Im}(q) \in \mathbb{H}.$$

Proposition B.6. For $p, q \in \mathbb{H}$ and $a \in \mathbb{R}$ we have

1. $\overline{p+q} = \overline{p} + \overline{q}$

2. $\overline{ap} = a\overline{p}$

3. $\overline{\overline{p}} = p$

4. $\overline{p \cdot q} = \overline{q} \cdot \overline{p}$

5. $p \in \mathbb{R} \Leftrightarrow \overline{p} = p$

6. p is purely imaginary $\Leftrightarrow \overline{p} = -p$

7. $\mathfrak{Re}(p) = \frac{1}{2}(p + \overline{p})$

8. $\mathfrak{Im}(p) = \frac{1}{2}(p - \overline{p})$

B.2 $\mathbb{H}$, $\mathbb{E}^4$ and $\mathbb{E}^3$

By construction $\mathbb{H}$ is $\mathbb{R}^4$ as real vector space, so we may view it as a 4-dimensional real affine space. If in addition we consier the 4-dimensional Euclidean structure we may identify $\mathbb{H}$ with $\mathbb{E}^4$. In particular, we may consider the standard scalar product $\langle \_, \_ \rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{H} \cong \mathbb{R}^4$.

Proposition B.7 (Compare this with the similar statements for $\mathbb{C} \cong \mathbb{E}^2$)**.** For $p, q \in \mathbb{H}$ we have

1. $\langle p, q \rangle = \frac{1}{2}(\overline{p}q + \overline{q}p)$
2. $\langle p, p \rangle = \overline{p}p$
3. $|p| = \sqrt{\overline{p}p}$

If in addition p and q are purely imaginary, we have

4. $\langle p, q \rangle = -\frac{1}{2}(pq + qp) = -\mathfrak{Re}(pq)$
5. $\langle p, p \rangle = -p^2$
6. $|p| = \sqrt{-p^2}$
7. $\langle p, q \rangle = 0 \Leftrightarrow pq = -qp$.

Definition B.8. With our identification, quaternions are vectors in $\mathbb{V}^4 \cong D(\mathbb{H}) \cong \mathbb{H}$ and the *norm* $|q|$ of a quaternion q equals $(\overline{q}q)^{\frac{1}{2}}$. If $|q| = 1$ we say that q is a *unit quaternion*.

Proposition B.9. For any $p, q \in \mathbb{H}$ we have

$$|pq| = |p| \cdot |q|.$$

In particular, left and right multiplication by unit quaternions are isometries.

Proposition B.10. $\mathbb{H}$ is a skew field. The inverse of $q \in \mathbb{H} \setminus \{0\}$ is

$$q^{-1} = \frac{\overline{q}}{|q|^2}.$$

We identified $\mathbb{H}$ with $\mathbb{E}^4$. Next, we view $\mathbb{E}^3$ as a subspace of $\mathbb{H}$ identifying it with purely imaginary quaterions $\mathfrak{Im}(\mathbb{H}) = \mathbb{R}\mathbf{i} + \mathbb{R}\mathbf{j} + \mathbb{R}\mathbf{k}$.

Proposition B.11. Let q_1, q_2 be two quaternions with $a_i = \mathfrak{Re}\, q_i$, $v_i = \mathfrak{Im}\, q_i$. Making use of the scalar product and the cross product in $\mathbb{E}^3$, we have

$$q_1 q_2 = (a_1 + v_1)(a_2 + v_2) = a_1 a_2 - \langle v_1, v_2 \rangle + a_2 v_1 + a_1 v_2 + v_1 \times v_2. \tag{B.2}$$

Proposition B.12. Let $v = v_i\mathbf{i} + v_j\mathbf{j} + v_k\mathbf{k} \in D(\mathbb{E}^3) \cong \mathfrak{Im}(\mathbb{H})$ be a unit quaternion and $p \in \mathbb{E}^3 \cong \mathfrak{Im}(\mathbb{H})$ a point. The rotation of p around the axis $\mathbb{R}v$ by an angle θ is given by

$$p' = qpq^{-1}$$

where

$$q = \cos\left(\frac{\theta}{2}\right) + \sin\left(\frac{\theta}{2}\right)v.$$

Bibliography

[1] D. Andrica, L. Ţopan, Analytical Geometry, Cluj University Press, 2005.

[2] M. Berger, Geometry I, Translated from the French by M. Cole, S. Levy, Springer, 2009.

[3] P.A. Blaga, Geometrie Liniară - Cu un ochi către grafica pe calculator, Presa Universitară Clujeană, 2022.

[4] S. Crivei, Basic Linear Algebra, Presa Universitară Clujeană, 2022.

[5] A.D. Mărcuș, Introducere în algebră pentru fizicieni, Casa Cărții de Știință, Cluj-Napoca, 2022.

[6] E. Sernesi, Linear Algebra; A geometric approach, (Translated by J. Montaldi), Chapman & Hall/CRC, 1993.

[7] M. Troyanov, Cours de géométrie, Press polytechniques universitaires romandes, 2011.