CHAPTER 1

Vectors and coordinate systems

Contents

We denote by $\mathbb{E}^2$ the Euclidean plane and by $\mathbb{E}^3$ the Euclidean space.

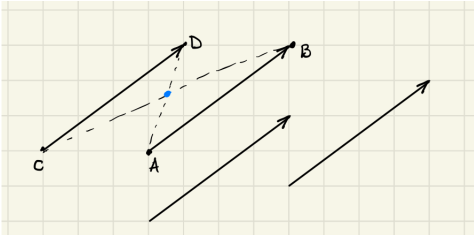
1.1 Geometric Vectors

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. 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.



Proposition 1.1. 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.

3. (A,B) and (C,D) have the same distance and direction.

Proposition 1.2. 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 1.3. The equipollence relation is an equivalence relation.

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

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

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 we have

- $\overrightarrow{AB} = \overrightarrow{CD}$ if and only if $(A,B) \sim (C,D)$, in particular all representatives of the vector $\overrightarrow{AB}$ define the *same* distance and the *same* direction, therefore

- if $\overrightarrow{AB} = \overrightarrow{CD}$ then $|AB| = |CD|$ and we define the *length* $|\overrightarrow{AB}|$ *of the vector* $\overrightarrow{AB}$ to be the distance $|AB|$ of the line segment $[A,B]$.

The set of all vectors for $\mathbb{E}^2$ and respectively for $\mathbb{E}^3$ are

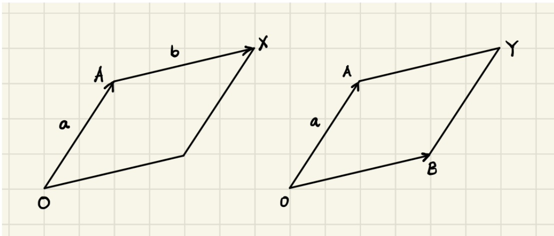
$$\mathbb{V}^2 = \{\overrightarrow{AB} : (A,B) \in \mathbb{E}^2 \times \mathbb{E}^2\} = \mathbb{E}^2 \times \mathbb{E}^2/\sim \quad \text{and} \quad \mathbb{V}^3 = \{\overrightarrow{AB} : (A,B) \in \mathbb{E}^3 \times \mathbb{E}^3\} = \mathbb{E}^3 \times \mathbb{E}^3/\sim.$$

Proposition 1.4. For any point O, the map ϕ_O defined by $\phi_O(A) = \overrightarrow{OA}$ is a bijection between points and vectors.

1.2 Vector space structure

Definition. Consider two vectors $\mathbf{a}$ and $\mathbf{b}$. If we fix a point O then, by Proposition 1.2, 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 $\mathbf{a}+\mathbf{b} = \overrightarrow{OY} = \overrightarrow{OX}$.

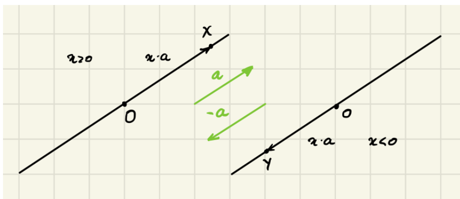


Proposition 1.5. The addition of vectors is well defined.

Proposition 1.6. The set of vectors $\mathbb{V}^2$ with addition is a commutative group. Similarly, $(\mathbb{V}^3,+)$ is a commutative group.

Definition. 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 by $x \cdot \mathbf{a}$ (or simply by $x\mathbf{a}$), is by definition

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



Proposition 1.7. The multiplication of scalars with vectors is well defined.

Proposition 1.8. 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 1.9. 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.

Remark. The vector space structure 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 $\mathbb{V}^3$). With Proposition 1.4 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

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

This is an example of a group action. We say that the group $\mathbb{V}^2$ acts on the set of point $\mathbb{E}^2$ by *translations*.

Proposition 1.4 gives a way of identifying vectors with points. If O is a fixed point then we have a bijective map

$$\phi_O : \mathbb{E}^2 \to \mathbb{V}^2 \quad \text{given by} \quad \phi_O(A) = \overrightarrow{OA}.$$

(similarly for $\mathbb{V}^3$). With this map we can compare objects in $\mathbb{E}^2$ with objects in $\mathbb{V}^2$ and objects in $\mathbb{E}^3$ with objects in $\mathbb{V}^3$ as follows.

Theorem 1.10. Let S be a subset of $\mathbb{E}^2$ and let O be a point in S.

1. The set S is a line if and only if $\phi_O(S)$ is a 1-dimensional vector subspace of $\mathbb{V}^2$.

2. If S is a line then the vector subspace $\phi_O(S)$ is independent of the choice of O in S.

3. Two vectors $\overrightarrow{OA}, \overrightarrow{OB}$ are linearly dependent $\Leftrightarrow O, A, B$ are collinear.

4. $\dim \mathbb{V}^2 = 2$, i.e. $\mathbb{V}^2 \cong \mathbb{R}^2$.

Similarly, let S be a subset of $\mathbb{E}^3$ and let O be a point in M.

5. The set S is a line if and only if $\phi_O(S)$ is a 1-dimensional vector subspace of $\mathbb{V}^3$.

6. If S is a line then the vector subspace $\phi_O(S)$ is independent of the choice of O in S.

7. Two vectors $\overrightarrow{OA}, \overrightarrow{OB}$ are linearly dependent $\Leftrightarrow O, A, B$ are collinear.

8. S is a plane if and only if $\phi_O(S)$ is a 2-dimensional vector subspace of $\mathbb{V}^3$.

9. If S is a plane then the vector subspace $\phi_O(S)$ is independent of the choice of O in S.

10. Three vectors $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ are linearly dependent $\Leftrightarrow O, A, B, C$ are coplanar.

11. $\dim \mathbb{V}^3 = 3$, i.e. $\mathbb{V}^3 \cong \mathbb{R}^3$.

Remark. A different perspective on the above theorem is the following. Given a subset $S \subseteq \mathbb{E}^3$ we can consider the *set of vectors represented by pairs of points in S*, i.e. the set

$$V_S = \{\overrightarrow{AB} : A, B \in S\}.$$

The theorem says that S is a line if and only if V_S is a 1-dimensional vector subspace of $\mathbb{V}^3$ and S is a plane if and only if V_S is a 2-dimensional vector subspace of $\mathbb{V}^3$. A similar interpretation holds for $\mathbb{E}^2$ and $\mathbb{V}^2$.

1.3 Coordinate systems

Coordinates are a way of associating tuples of numbers to points in a space. Given a 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 $(x_1, \ldots, x_n) \leftrightarrow p$ 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.

Fix two lines ℓ_1 and ℓ_2 in $\mathbb{E}^2$ which intersect in a (unique) point O. We can describe any point P in $\mathbb{E}^2$ as follows. By Theorem 1.10, $\phi_O(\ell_1)$ 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 and y (in $\mathbb{R}$) such that

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

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

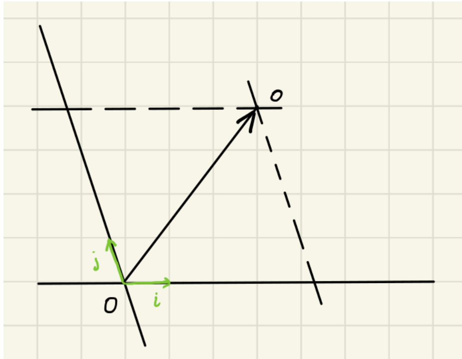
For this bijection we made two choices:

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

Definition. A *coordinate system* for $\mathbb{E}^2$ is a triple $\mathcal{K} = (O, \mathbf{i}, \mathbf{j})$ where O is a point in $\mathbb{E}^2$ and $(\mathbf{i}, \mathbf{j})$ is a basis of $\mathbb{V}^2$. Given a coordinate system $\mathcal{K} = (O, \mathbf{i}, \mathbf{j})$, 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 $P(x_P, y_P)$ when we want to indicate the coordinates. 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 $\mathcal{K}$.

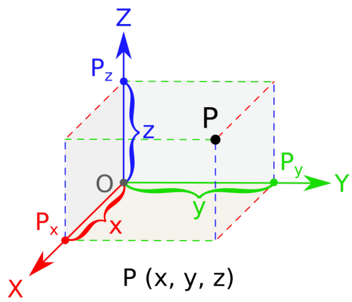


Remark. When we fix a coordinate system $\mathcal{K} = (O, \mathbf{i}, \mathbf{j})$ with $\mathbf{i}$ and $\mathbf{j}$ of length 1 then we can interpret the coordinates of a point P as follows. For each such point P there is a unique parallelogram $OXPY$ with $X \in Ox$ and $Y \in Oy$ and the coordinates (x_P, y_P) of the point P with respect to $\mathcal{K}$ are the lengths of the sides of this parallelogram.

Similarly, if we fix three non-coplanar lines ℓ_1, ℓ_2 and ℓ_3 in $\mathbb{E}^3$ which intersect in a (unique) point O, we can describe any point P in $\mathbb{E}^3$ as follows. By Theorem 1.10, $\phi_O(\ell_1)$, $\phi_O(\ell_2)$ and $\phi_O(\ell_3)$ are 1-dimensional linearly independent vector subspaces of the two dimensional vector space $\mathbb{V}^3$. Thus, if we choose $\mathbf{i} \in \phi_O(\ell_1)$, $\mathbf{j} \in \phi_O(\ell_2)$ and $\mathbf{k} \in \phi_O(\ell_3)$ 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 and z (in $\mathbb{R}$) such that

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

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. A *coordinate system* for $\mathbb{E}^3$ is a quadruple $\mathcal{K} = (O, \mathbf{i}, \mathbf{j}, \mathbf{k})$ where O is a point in $\mathbb{E}^3$ and $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ is a basis of $\mathbb{V}^3$. Given a coordinate system $\mathcal{K} = (O, \mathbf{i}, \mathbf{j}, \mathbf{k})$, for any point $P \in \mathbb{E}^2$ 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 $P(x_P, y_P, z_P)$ when we want to indicate the coordinates. 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.

Remark. 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 then 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.

Notation. We write points and vectors as column matrices when we work with their coordinates respectively with their components. For example, in dimension 2

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

With the subscript $\mathcal{K}$ we indicate the coordinate system with respect to which P has the indicated coordinates. For vectors, the subscript $\mathcal{K}$ indicates the basis $(\mathbf{i}, \mathbf{j})$ with respect to which the vector $\mathbf{a}$ has components a_x and a_y. Similarly, in dimension 3

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

1.4 Coordinates as projections

Let $\mathcal{K}$ be a coordinate system of $\mathbb{E}^2$. The correspondence between points and coordinates in the system $\mathcal{K}$

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

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*.

If $\mathcal{B} = (\mathbf{i}, \mathbf{j})$ is a basis for $\mathbb{V}^2$ then for vectors, 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

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

It follows that

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

Similarly, 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 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

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

and by

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

It follows that

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

1.5 Changing coordinate systems

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. We will look at the details of how this is done later during the course.