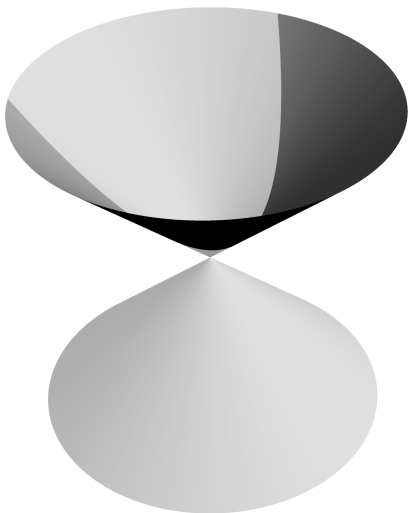
Geometry I

Draft 2025



Contents

1 Affine space **4**
1.1 Geometric Vectors . . . 5
1.2 Vector space structure of geometric vectors . . . 12
1.3 Affine space structure of the Euclidean space . . . 16

2 Cartesian coordinates **18**
2.1 Frames in dimension 2 . . . 19
2.2 Frames in dimension 3 . . . 24
2.3 Frames in dimension n . . . 27

3 Affine subspaces **30**
3.1 Lines in $\mathbb{A}^2$ . . . 31
3.2 Planes in $\mathbb{A}^3$ . . . 34
3.3 Lines in $\mathbb{A}^3$ . . . 37
3.4 Affine subspaces of $\mathbb{A}^n$ . . . 41

4 Euclidean space **48**
4.1 Angles . . . 49
4.2 Scalar product . . . 55
4.3 Distance . . . 61

5 Area and volume **66**
5.1 Area . . . 67
5.2 Cross product . . . 70
5.3 Volume . . . 78

6 Affine maps **86**
6.1 Properties of affine maps . . . 87

6.2 Projections and reflections . . . 90

7 Isometries **98**
7.1 Affine form of isometries . . . 99
7.2 Isometries in dimension 2 . . . 101
7.3 Isometries in dimension 3 . . . 104
7.4 Moving points with isometries . . . 108

Appendices **113**

A Axioms **114**

B Lines and the real numbers **117**

C Changing the basis in a vector space **120**

D Coordinate systems **122**
D.1 Polar coordinates . . . 122
D.2 Cylindrical coordinates . . . 123
D.3 Spherical coordinates . . . 123
D.4 Barycentric coordinates . . . 124

E Bundles of lines and planes **125**
E.1 Bundle of lines in $\mathbb{A}^2$ . . . 125
E.2 Bundle of planes in $\mathbb{A}^3$ . . . 126

F Some classical theorems in affine geometry **128**

G Eigenvalues and Eigenvectors **140**
G.1 Characteristic polynomial . . . 141

H Bilinear forms and symmetric matrices **143**
H.1 Affine diagonalization . . . 143
H.2 Isometric diagonalization . . . 147

I Trigonometric functions **151**

J Some classical theorems in Euclidean geometry **154**

K Quaternions and rotations **156**
K.1 Algebraic considerations . . . 156
K.2 Geometric considerations . . . 157

Bibliography **159**

CHAPTER 1

Affine space

Contents

1.1 Geometric Vectors 5
1.2 Vector space structure of geometric vectors 12
1.3 Affine space structure of the Euclidean space 16

1.1 Geometric Vectors

In this section, we introduce the concept of a *vector* based on the axiomatic framework provided by Hilbert in his *Foundations of Geometry* [14]. Vectors are at the heart of our understanding of Euclidean geometry. Axioms are the starting point of any process of logical reasoning. Hilbert's axioms can be found in Appendix A.

We denote by $\mathbb{E}$ the Euclidean space. It consists of a set of elements called *points*, which are governed by the axioms. Points are primitives and the axioms describe how primitives interact, not what they are. The interaction is expressed through relations, such as *collinearity*, *betweenness*, *coplanarity*, *incidence* and *congruence*.

One approach to understanding a theory from its axioms is to systematically explore the information derived from a given set of primitives. If a single point is given, no additional information can be inferred. If two points are given, we arrive at the concept of a vector. A vector encapsulates all the information carried by two points through the axioms. In Euclidean geometry, the information carried by two points is *length* and *direction*. To formally define vectors, we treat two (not necessarily distinct) points A and B as an *ordered pair* (A,B).

We begin with the concept of *length*. Given two points A and B, Axiom I.1 ensures the existence of a *straight line* passing through them, commonly denoted as AB. For brevity, we use the shorter term *line* for a straight line. Using the *betweenness* relation, the Axioms of Order allow us to define segments. The *segment* $[AB]$ is the set of points on the line AB which lie between the points A and B together with the points A and B. Segments $[AA]$, with equal endpoints, are called *trivial segments*. If $[AB]$ is congruent to $[CD]$ we write $[AB]\equiv[CD]$.

Definition 1.1. We say that two ordered pairs of points (A,B) and (C,D) are *equidistant*, and we write $(A,B)\equiv(C,D)$ if and only if the segments $[AB]$ and $[CD]$ are congruent[1].

Proposition 1.2. The equidistance relation is an equivalence relation.

Proof. The equidistance relation is equivalent to the congruence relation on segments. We show that the latter relation is an equivalence relation. We need to show reflexivity, symmetry and transitivity. In order to show that the relation is reflexive, fix $[AB]$ and construct congruent a segment $[A'B']$ using Axiom III.1. Then, applying Axiom III.2 to the congruences $[AB]\equiv[A'B']$ and $[AB]\equiv[A'B']$ we obtain $[AB]\equiv[AB]$. For symmetry, assume that $[AB]\equiv[CD]$. Applying Axiom III.2 to the congruences $[CD]\equiv[CD]$ and $[AB]\equiv[CD]$ we obtain $[CD]\equiv[AB]$ as desired. For transitivity, assume that $[AB]\equiv[CD]$ and that $[CD]\equiv[EF]$. By symmetry we have $[EF]\equiv[CD]$. Then, applying Axiom III.2 to the congruences $[AB]\equiv[CD]$ and $[EF]\equiv[CD]$ we obtain $[AB]\equiv[EF]$ as desired. □

Definition 1.3. The equivalence classes of the equidistance relation are called *distances* or *lengths*. The distance defined by the pair (A,B) is denoted by $|AB|$. It is also called the length of the segment $[AB]$. Explicitly, we have

$$|AB| = \Big\{\text{ordered pairs } (X,Y) \text{ such that } (X,Y)\equiv(A,B)\Big\}.$$

[1] The axioms sometimes refer to segments as being 'congruent or equal' which may suggest that congruence is reserved for unequal segments. For us, if two segments are equal they are congruent.

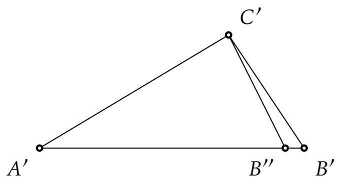
We say that (A,B) *represents the distance* $|AB|$ and that $[AB]$ *represents the length* $|AB|$. If $|AB| = |CD|$ we say that $[AB]$ *and* $[CD]$ *have the same length*. The set of distances/lengths defined by segments in $\mathbb{E}$ is the set of equivalence classes:

$$\mathbb{L} = \Big\{ |AB| : (A,B) \in \mathbb{E} \times \mathbb{E} \Big\} = \mathbb{E} \times \mathbb{E} \Big/ \equiv .$$

Next, we introduce the concept of *direction*. The Axioms of Order also allow us to define the notion of sides of the line AB with respect to the point A. The points $B, C \in AB$ are *on the same side of AB relative to A* if A is not between B and C (see [14, p.8]). The set of points on the same side of a line relative to A is also called a *ray emanating from A*. If it contains the point B, we denote it by $(AB$.

Lemma 1.4. For any segment $[AB]$, any line ℓ and any point $A' \in \ell$ there is a unique segment $[A'B']$, congruent to $[AB]$, lying on ℓ on a given side relative to A'.

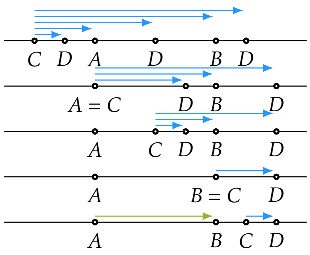
Proof. The existence of the point B', and therefore of the segment $[A'B']$, is guaranteed by Axiom III.1. In order to prove uniqueness, suppose for a contradiction that there is another point B'' on the ray $(A'B'$ such that $[A'B'] \equiv [A'B'']$. Choose a point C' not on the line $A'B'$. The existence of such a point is guaranteed by Axioms I.2 and I.8. We have $[A'C'] \equiv [A'C']$, $[A'B'] \equiv [A'B'']$ and $\measuredangle C'A'B' \equiv \measuredangle C'A'B''$. Applying Axiom III.6 we obtain $\measuredangle A'C'B' \equiv \measuredangle A'C'B''$ which contradicts Axiom III.4. □

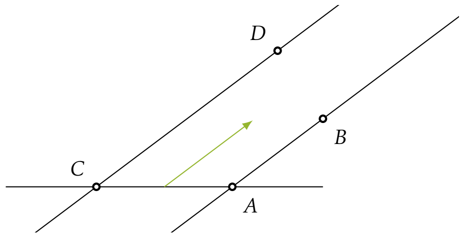


When doing geometry it is customary to avoid degenerate cases and overlaps. However, for our purpose of building on the axioms, we take degenerate cases into account when defining directions. This will open the way to linear algebra tools. Before giving the definition, let us recall that the Axioms of Order allow us to define the notion of sides of a plane α with respect to a line contained in α (see [14, p.8]).

Definition 1.5. We say that two ordered pairs of points (A,B) and (C,D) are *equidirectional*, and we write $(A,B) \triangleq (C,D)$, in the following three cases:

1. if $A = B$ and $C = D$;
2. if $A \neq B$, the points A, B, C are collinear and
   - 2.1. $A = C$ and B, D are on the same side relative to A, or
   - 2.2. B, C are on the same side of A and A, D are on opposite sides of C, or
   - 2.3. B, C are on opposite sides of A and A, D are on the same side of C;





3. if $A \neq B$, the points A, B, C are not collinear, AB and CD are parallel and B and D lie on the same side of AC in the plane that the four points determine.

Proposition 1.6. Consider two pairs of distinct points (A,B) and (C,D).

1. If $AB = CD$ then $(A,B) \overset{\triangle}{=} (C,D)$ if and only if $(AB \cap (CD = (AB$ or $(AB \cap (CD = (CD$.

2. If AB, CD are distinct but parallel then $(A,B) \overset{\triangle}{=} (C,D)$ if and only if $[AD]$ intersects $[BC]$.

Proof. 1. If $A = C$ then, by definition, we have $(A,B) \overset{\triangle}{=} (C,D)$ if and only if B, D are on the same side of A. Thus, by definition, we have $(AB = (CD$, hence $(AB \cap (CD = (AB = (CD$. For the remaining cases we make use of the following fact which intuitively is clear: $(*)$ four points on a line can always be relabeled P_1, P_2, P_3, P_4 such that P_2 and P_3 lie between P_1 and P_4, and furthermore, that P_2 lies between P_1 and P_3 and P_3 lies between P_2 and P_4. This is Theorem 5 in [14].

Consider the case where $A \neq C$. For the implication from left to right, assume first that B, C are on the same side of A. Since C is between A and D, by $(*)$, for any point X on $(CD$ we have C between A and X, thus $X \in (AC$. Since $(AC = (AB$ we showed that $(AB \cap (CD = (CD$. Next assume that B, C are on opposite sides of A. By $(*)$, for any $X \in (AB$ we have A between C and X, thus $X \in (CA$. Since A, D are on the same side of C we also have $(CA = CD$, hence $(AB \cap (CD = (AB$. The implication from right to left is easier.

2. Now assume that the lines AB and CD are distinct and parallel. We show the implication from left to right since the other direction is easier. Assume for a contradiction that B and C lie on the same side of the line AD. Then $(DC$ lies on the same side of AD as B. By assumption, since $(A,B) \overset{\triangle}{=} (C,D)$, we also have $(CD$ on the same side of AC as B. By Axiom II.2, we may choose a point P between C and D. Since P lies on both $(CD$ and $(DC$ it lies on the same side of AC and AD as B. Thus B lies in the interior of the angle described by $(AC$ and $(AD$. Hence AB intersects $[CD]$, contradicting $AB \| CD$. Thus B and C lie on opposite sides of AD, i.e. $[BC]$ intersects AD. Interchanging the role of B, C with A, D we find that $[AD]$ intersects BC. Therefore $[BC]$ intersects $[AD]$. □

Proposition 1.7. The equidirectional relation is an equivalence relation.

Proof. We need to show that the equidirectional relation is reflexive, symmetric and transitive. Reflexivity and symmetry follow directly from the definition. For transitivity, let (A,B), (C,D) and (E,F) be three ordered pairs of points. Assuming that $(A,B) \overset{\triangle}{=} (C,D)$ and that $(C,D) \overset{\triangle}{=} (E,F)$, we need to show that $(A,B) \overset{\triangle}{=} (E,F)$. If $A = B$ then $C = D$, thus $E = F$ and the claim follows in this case. For the rest of the proof assume that the considered pairs consist of distinct points.

Next, we prove transitivity along lines by treating the case when the four points are collinear. By Proposition 1.6, we have $(AB \cap (CD = (AB$ or $(CD$ and similarly, $(EF \cap (CD = (EF$ or $(CD$. In all four cases, by Axiom V.1 we find a point P such that $(AB = (AP$, $(CD = (CP$ and $(EF = (EP$. Then, if $(AP \cap (CP = (AP$ and $(EP \cap (CP = (EP$, we have A, E between C and P. It follows from Theorem 5 in [14] that E lies between C and A or that A lies between C and E. In the first case $(AP \cap (EP = (EP$ and in the second case $(AP \cap (EP = (AP$. The other three cases are treated similarly. Thus, for the rest of the proof we may assume that we consider pairs of distinct points and that the six points are not collinear.

Consider the cases where two of the lines overlap. The cases where AB or EF equal CD follow directly from transitivity along lines proved in the previous paragraph. It remains to consider the case where $AC = EF = \ell$. Since the six points are not collinear, the line CD is distinct from ℓ and parallel to ℓ. We need to show that $(A,B) \triangleq (E,F)$. By transitivity along lines, we may replace (A,B) and (E,F) by other equidirectional pairs of points on ℓ. Thus, we may assume that $A = E$. Then B and F are on the same side of the line $AC = EC$ since $(A,B) \triangleq (C,D)$ and since $(E,F) \triangleq (C,D)$. Hence B,F are on the same side of A.

Finally, consider the case where the three lines AB, CD and EF are distinct. Again, by transitivity along lines, we may shift the pairs of points along these three lines. Thus we may assume that C is the intersection of AE with CD. Then B, D are on the same side of the line $AC = AE$ since $(A,B) \triangleq (C,D)$ and D, F are on the same side of the line $CE = AE$ since $(A,B) \triangleq (C,D)$. Thus B,F are on the same side relative to the line AE. □

Definition 1.8. The equivalence classes of the equidirectional relation are called *directions*. The direction containing the ordered pair (A,B) is denoted by $|AB\rangle$. Explicitly, we have

$$|AB\rangle = \Big\{\text{ordered pairs } (X,Y) \text{ such that } (X,Y) \triangleq (A,B)\Big\}.$$

We say that (A,B) *is a representative of the direction* $|AB\rangle$. If $|AB\rangle = |CD\rangle$ we say that (A,B) *and* (C,D) *define the same direction*. The set of directions defined with points in $\mathbb{E}$ is the set of equivalence classes:

$$\mathbb{D} = \Big\{|AB\rangle : (A,B) \in \mathbb{E} \times \mathbb{E}\Big\} = \mathbb{E} \times \mathbb{E}\Big/ \triangleq .$$

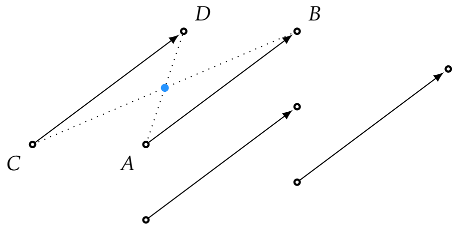
The direction $|BA\rangle$ is called *the opposite direction of* $|AB\rangle$ and is denoted by $-|AB\rangle$. This defines an involution $-\square : \mathbb{D} \to \mathbb{D}$.

At the intersection of the equidistance relation and the equidirectional relation lies the equipollence relation which is used to define vectors. Before defining this relation, we define degenerate parallelograms first.

Definition 1.9. A *parallelogram* $ABCD$ is an ordered quadruple of points (A,B,C,D) with the usual property of having parallel opposite sides if the four points are not collinear. In this case, the labeling is such that C, D are on the same side of AB and the points D, A are on the same side of BC.

If the four points are collinear we require that the segments $[AC]$ and $[BD]$ have the same midpoint. In this case, we say that the parallelogram is *degenerate*.

Definition 1.10. Two ordered pairs of points (A,B) and (C,D) are called *equipollent*, and we write $(A,B) \sim (C,D)$, if the segments $[AD]$ and $[BC]$ have the same midpoints.



Proposition 1.11. For two ordered pairs of points (A,B) and (C,D) the following statements are equivalent:

1. $(B,A) \sim (D,C)$.
2. $(A,B) \sim (C,D)$.
3. $ABDC$ is a parallelogram, possibly degenerate.
4. $|AB\rangle = |CD\rangle$ and $[AB] \equiv [CD]$.
5. $|AB\rangle = |CD\rangle$ and $|AB| = |CD|$.

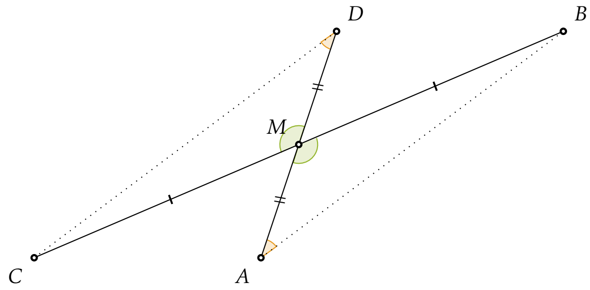
Proof. First notice that the equivalence (1. $\Leftrightarrow$ 2.) follows directly from Definition 1.10 and that (4. $\Leftrightarrow$ 5.) is by Definition 1.1. Thus, it suffices to show (2. $\Rightarrow$ 3.), (3. $\Rightarrow$ 4.) and (4. $\Rightarrow$ 2.).

(2. $\Rightarrow$ 3.) Let $(A,B) \sim (C,D)$. Denote by M the common midpoint of the segments $[AD]$ and $[BC]$. If $A = B$ then $[AD]$ and $[AC]$ have M as common midpoint. Then, if $A = C$ or $A = D$ it follows that all four points coincide, hence $ABCD$ is a parallelogram by definition. If on the other hand $A \neq C$ and $A \neq D$ then M lies between A and C and between A and D. From Lemma 1.4 it follows that $C = D$. Thus, for $A = B$ we showed that $ABDC$ is a degenerate parallelogram. A similar argument shows that $ABDC$ is a parallelogram if $C = D$. For the rest of the proof we may assume that $A \neq B$ and that $C \neq D$.

Assume further that A, B, C are collinear. Since M is the common midpoint of $[AD]$ and $[BC]$ we have $DM = AM = BC$. Thus, the points are collinear and $ABDC$ is a degenerate parallelogram by definition.

Assume next that $A \neq B$, $C \neq D$ and that A, B, C are not collinear. In this case the lines AB and CD are distinct. Therefore, the common midpoint M of $[AD]$ and $[BC]$ cannot lie on any of these two lines. Thus, the angles $\measuredangle CMD$ and $\measuredangle AMB$ are defined (see [14, p.11]). They are so-called vertical angles (see [14, §6]) and it follows from [14, Theorem 14] that vertical angles are congruent. Thus, by the first congruence theorem for triangles [14, Theorem 12], the triangles MCD and MBA are congruent. In particular $\measuredangle MDC$ is congruent to $\measuredangle MAB$ and, again by [14, Theorem 14], they have congruent supplementary angles. It then follows from [14, Theorem 30] that the lines AB and CD are parallel. Similarly, one shows that AC and BD are parallel. Thus $ABDC$ is a parallelogram.

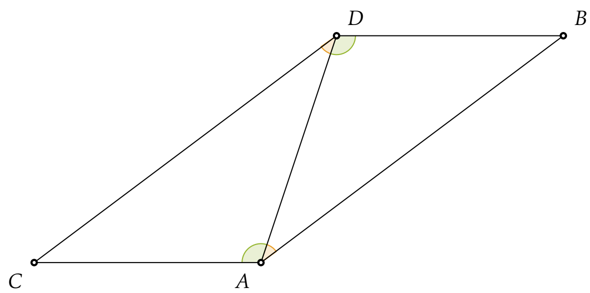
(3. $\Rightarrow$ 4.) Let $ABDC$ be a parallelogram. Assume first that it is degenerate, i.e. the four points are collinear. Let M be the common midpoint of $[AD]$ and $[CB]$. First consider the case where $A = B$. If $C = D$ we are done since $[AB] \equiv [CD]$ and $|AB\rangle = |CD\rangle$. Assume for a contradiction that $C \neq D$. Then



D, M are on the same side of A and C, M are on the same side of $B = A$. From Lemma 1.4 it follows that $C = D$. A similar argument shows that if $C = D$ then $A = B$ and our claim follows. Thus, we may assume that $A \neq B$ and $C \neq D$. Assume now that $A = C$. Then D, M are on the same side of A and B, M are on the same side of $C = A$. From Lemma 1.4 it follows that $B = D$ and our claim follows. Next, assume that $A = D$. Then $M = A = D$. Since M is the midpoint of $[CB]$, we have that $A = D$ lies between C and B, hence $|CD\rangle = |AB\rangle$. Moreover, since $A = D$ is the midpoint of $[CB]$ we have $[CD] \equiv [AB]$ by Lemma 1.4. We may assume for the rest of the proof that the four points are distinct.

Since M is the common midpoint of the segments $[AD]$ and $[BC]$, from Axiom III.3 we deduce that $[AB] \equiv [CD]$. It remains to show that $|AB\rangle = |CD\rangle$. Since M is the midpoint of $[AD]$, the points A, D lie on distinct sides of AD relative to M. Similarly for B and C. Thus, we have two possibilities: either A and B are on the same side relative to M or A and C are on the same side relative to M. In the first case, there are two possibilities: either B is between A and M or A is between B and M. If B is between A and M, since M is the common midpoint of $[AD]$ and $[BC]$, it follows from Axiom III.3 that C lies between M and D. Therefore A and D are on opposite sides relative to C, i.e. (A, B) and (C, D) define the same direction. The remaining three cases are treated similarly.

Now assume that $ABDC$ is not degenerate. By definition, the opposite sides of $ABDC$ are parallel. By [14, Theorem 30], we have the following congruences $\measuredangle CDA \equiv \measuredangle DAB$ and $\measuredangle DAC \equiv \measuredangle ADB$. By the second congruence theorem for triangles [14, Theorem 13], the triangles ACD and DBA are congruent, in particular $[AB] \equiv [CD]$. Moreover, by definition, the labeling of the vertices are such that B and D lie on the same side of the line determined by A and C. Thus, since the sides are parallel, the pairs (A, B) and (C, D) define the same direction.



(4. $\Rightarrow$ 2.) Assume that $[A, B] \equiv [C, D]$ and $|AB\rangle = |CD\rangle$. Then AB is parallel to CD. If the four

points are on the same line ℓ, the configuration is degenerate. By definition, since (A,B) and (C,D) define the same direction, we have two cases. In the first case B, C are on the same side of ℓ relative to A and A, D are on distinct sides of ℓ relative to C. Assume that C is between B and D and let M be the midpoint of $[AD]$. Then, by Axiom III.3 applied to M,A,B and M,D,C we have $[MB] \equiv [MC]$, i.e. M is also the midpoint of $[BC]$. The other cases are treated similarly. The only thing to pay attention to is that Axiom III.3 requires segments which do not overlap.

Finally, assume that the four points are not collinear. Denote by M the intersection of the diagonals. It exists by Point 2. of Proposition 1.6. By [14, Theorem 30] and the second congruence theorem for triangles [14, Theorem 13], the triangles MAB and MDC are congruent. Thus $[AM] \equiv [MD]$ and $[BM] \equiv [MC]$, i.e. M is the common midpoint of $[AD]$ and $[BC]$. □

Vectors are defined as equivalence classes of the equipollence relation (see Definition 1.13 below). To this end, the following proposition shows that this relation is indeed an equivalence relation.

Proposition 1.12. The equipollence relation is an equivalence relation.

Proof. By Proposition 1.11 we have that $(A,B) \sim (C,D)$ if and only if $|AB\rangle = |CD\rangle$ and $[AB] \equiv [CD]$. By Proposition 1.2 and Proposition 1.7 both the equidistance relation on the equidirectional relation are equivalence relations. Thus, reflexivity, symmetry and transitivity for the equipollence relation follow from the corresponding properties of these two relations. □

Definition 1.13. The equivalence classes of the equipollence relation are called *vectors*. The vector containing the ordered pair (A,B) is denoted by $\overrightarrow{AB}$. Explicitly, we have

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

We say that (A,B) *is a representative of the vector* $\overrightarrow{AB}$. If $\overrightarrow{AB} = \overrightarrow{CD}$ we say that (A,B) *and* (C,D) *define the same vector*. By Proposition 1.11 we have that $\overrightarrow{AB} = \overrightarrow{CD}$ if and only if $|AB\rangle = |CD\rangle$ and $|AB| = |CD|$. The set of vectors defined with points in $\mathbb{E}$ is the set of equivalence classes:

$$\mathbb{V} = \Big\{\overrightarrow{AB} : (A,B) \in \mathbb{E}\times\mathbb{E}\Big\} = \mathbb{E}\times\mathbb{E}\Big/\sim.$$

The vector $\overrightarrow{AA}$ is called *the zero vector* and we denote it by $\overrightarrow{0}$ or simply by 0 when there is no risk of confusion. Since all representatives of a vector $\overrightarrow{AB}$ define the same length $|AB|$ this will also be the *length* of the vector $\overrightarrow{AB}$ and we denote it by $|\overrightarrow{AB}|$. The vector $\overrightarrow{BA}$ is called *the opposite of the vector* $\overrightarrow{AB}$ and is denoted by $-\overrightarrow{AB}$. This defines an involution $-\square : \mathbb{V} \to \mathbb{V}$.

The first observation about vectors (which we prove below) is that for any fixed but arbitrary point O there is a 1-to-1 correspondence between points A and vectors $\overrightarrow{OA}$. In this correspondence $\overrightarrow{OA}$ is called *the position vector of A relative to O* or, if it is clear from the context what O is, we simply say *the position vector of A*.

Proposition 1.14. For any ordered pair of points (A,B) and any point O, there exists a unique point X such that $(A,B) \sim (O,X)$.

Proof. Assume first that A, B and O are not collinear. By Axiom III.4 there is a line ℓ passing through O and having the same angles with AO as AB. By [14, Theorem 22] the lines ℓ and AB cannot have a point in common, i.e. they are parallel. By Axiom IV, the line ℓ is the unique line passing through O which is parallel to AB. Similarly, there is a unique line ℓ' passing through B and which is parallel to OA. Let X be intersection point of ℓ and ℓ'. Then $ABXO$ is a parallelogram. Hence, by Proposition 1.11, we have $(A,B) \sim (O,X)$.

Now assume that A, B and O lie on a line ℓ. If $O = A$ then $[AB] = [OB]$, thus $(A,B) \sim (O,X)$ if and only if $X = B$. If $O \neq A$, consider the two sides in which A divides ℓ. If O and B are on the same side, there exists a unique segment $[OX]$ on ℓ congruent to $[AB]$ and such that A and X are not on the same side of ℓ relative to O (by Lemma 1.4). Thus (A,B) and (O,X) define the same direction, hence $(A,B) \sim (O,X)$. If O and B are on different sides of ℓ relative to A, there exists a unique segment $[OX]$ on ℓ which is congruent to $[AB]$ and such that A and X lie on the same side of ℓ relative to O (by Lemma 1.4). As in the previous case, X is the unique point such that $(A,B) \sim (O,X)$. $\square$

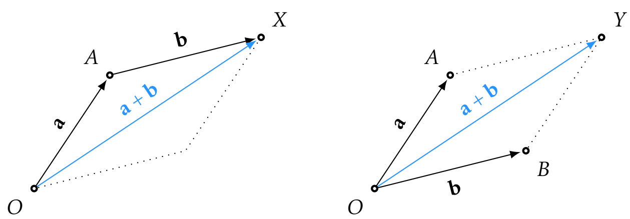
Corollary 1.15. For any point O, the map $\phi_O : \mathbb{E} \to \mathbb{V}$ defined by $\phi_O(A) = \overrightarrow{OA}$ is a bijection.

Proof. Fix a vector $\overrightarrow{CD}$. We show that there is a point A such that $\phi_O(A) = \overrightarrow{CD}$. By Proposition 1.14, there is a unique point A such that $(C,D) \sim (O,A)$, i.e. there exists a unique point A such that $\overrightarrow{CD} = \overrightarrow{OA} = \phi_O$. The existence of A implies the surjectivity of ϕ_O and the uniqueness implies that ϕ_O is injective. Thus, ϕ_O is bijective. $\square$

Remark. It is clear that if in the set of ordered pairs of points $\mathbb{E} \times \mathbb{E}$ we fix the first entry then we obtain a bijection with $\mathbb{E}$. What Corollary 1.15 is saying is that, starting from the Axioms, vectors do not carry *more* information than two points do. So, why not simply work with pairs of points instead? We *are* simply working with pairs of points to which we formally attached the concept of *length* and *direction*.

1.2 Vector space structure of geometric vectors

For a fixed point O, Corollary 1.15 allows us to identify points with geometric vectors. However, the set $\mathbb{V}$ of vectors has more structure than the set of points. In this section we show that $\mathbb{V}$ is a real vector space.



Definition 1.16. Consider two vectors $\mathbf{a}$ and $\mathbf{b}$. If we fix a point O then, by Proposition 1.14, 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 1.17. The addition of vectors is well defined.

Proof. Let $\mathbf{a},\mathbf{b} \in \mathbb{V}$ and O,A,X be as in Definition 1.16. The sum maps $(\mathbf{a},\mathbf{b})$ to $\mathbf{a}+\mathbf{b}$ as in the definition. It is a map $\square+\square : \mathbb{V}\times\mathbb{V} \to \mathbb{V}$ on equivalence classes which is defined using representatives. The proposition claims that the definition *does not depend on the choice of representatives*, i.e. if we replace O with a point distinct from O, then the above construction yields the same equivalence class.

Fix a point $O' \neq O$. Let A' be the unique point such that $\mathbf{a} = \overrightarrow{O'A'}$ and let X' be the unique point such that $\mathbf{b} = \overrightarrow{A'X'}$. We need to show that $\overrightarrow{AX} = \overrightarrow{A'X'}$. Since $\overrightarrow{O'A'} = \mathbf{a} = \overrightarrow{OA}$, we have that $OAA'O'$ is a parallelogram. Similarly $AXX'A'$ is a parallelogram and it suffices to show that $OXX'O'$ is a parallelogram.

If $O = A$, then $O' = A'$ and $OXX'O'$ is a parallelogram since $AXX'A'$ is a parallelogram. The cases where $A = X$ or $X = O$ are similar. Thus, we may assume that the points O,A,X are pairwise distinct. Now, if O,A,X are collinear then $[OX]$ is congruent to $[O'X']$ by Axiom III.3. Moreover, by the construction in Definition 1.16, we have $|OX\rangle = |OX'\rangle$ and therefore $\overrightarrow{OX} = \overrightarrow{O'X'}$.

Finally, assume that $A \notin OX$. By the third congruence theorem for triangles [14, Theorem 18], we have that the triangles OAX and $OA'X'$ are congruent. In particular $\measuredangle AOX \equiv \measuredangle A'O'X'$. Moreover, since the lines OA and OA' are parallel, they meet OO' in congruent angles (by [14, Theorem 30]). From this we deduce that OX and OX' meet OO' in congruent angles. Thus OX and $O'X'$ are parallel. Since $OO' \| AA'$ and $AA' \| XX'$ we also have $OO' \| XX'$. Thus $OXX'O'$ is a parallelogram. □

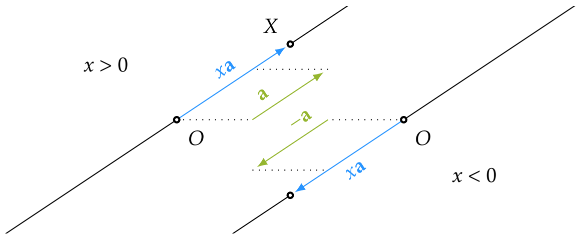
Proposition 1.18. The set of vectors $\mathbb{V}$ with addition is an abelian group.

Proof. Proposition 1.17 show that addition is indeed an operation on vectors. Thus we may choose convenient representatives if needed. Let $\mathbf{a},\mathbf{b},\mathbf{c}$ be three vectors. Addition is associative if $(\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c})$. Fix a point A. Then, by Proposition 1.14, there exist unique points B, C and D such that $\overrightarrow{AB} = \mathbf{a}$, $\overrightarrow{BC} = \mathbf{b}$, $\overrightarrow{CD} = \mathbf{c}$. But then $(\overrightarrow{AB} + \overrightarrow{BC}) + \overrightarrow{CD} = \overrightarrow{AC} + \overrightarrow{CD} = \overrightarrow{AD} = \overrightarrow{AB} + \overrightarrow{BD} = \overrightarrow{AB} + (\overrightarrow{BC} + \overrightarrow{CD})$ which proves associativity. The neutral element is $0 = \overrightarrow{AA}$ since $\overrightarrow{AA} + \overrightarrow{AB} = \overrightarrow{AB} = \overrightarrow{AB} + \overrightarrow{BB} = \overrightarrow{AB} + \overrightarrow{AA}$. The inverse element of $\overrightarrow{AB}$ is the opposite vector $-\overrightarrow{AB} = \overrightarrow{BA}$ since $\overrightarrow{AB} + (-\overrightarrow{AB}) = \overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = 0$. Finally, to show commutativity we notice the following. Since $\mathbf{a}+\mathbf{b}$ is constructed on the diagonal of a parallelogram with $\mathbf{a}$ and $\mathbf{b}$ represented on the sides, the construction is symmetric and does not depend on the order of $\mathbf{a}$ and $\mathbf{b}$ in the sum. □

Up to this point we have used the concept of length solely as equivalence classes given by the congruence relation on segments (Definition 1.8). While this was enough for our discussion so far, in what follows we need more structure on the set $\mathbb{L}$ of lengths. It is natural to ask about *all the consequences* that the Axioms have on $\mathbb{L}$. *If we choose* a segment $[AB]$, a *unit segment*, one can deduce from the axioms that the line AB can be identified with the set of real numbers $\mathbb{R}$ and that $\mathbb{L}$ can be identified with the set of non-negative real numbers $\mathbb{R}_{\geq 0}$ such that $|AB| = 1$. These identifications require care. They are available in Appendix B for the interested reader (see in particular Theorem B.9). The identification is necessary for example in the existence and uniqueness claim of the following definition which uses Proposition B.11.

Definition 1.19. Assume that a unit segment was chosen. Consider a non-zero vector $\mathbf{a} = \overrightarrow{OA}$ and a scalar $x \in \mathbb{R}$. If $x > 0$, there is a unique point X on the ray $(OA$ such that $|OX| = x \cdot |OA|$. The *multiplication of the vector* **a** *with the scalar* x, denoted $x \cdot \mathbf{a}$ (or simply $x\mathbf{a}$), is given by

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



Proposition 1.20. Assume that a unit segment was chosen. The multiplication of vectors with scalars is well defined.

Proof. Let $\overrightarrow{OA} = \overrightarrow{O'A'}$. For $r > 0$, construct X from O and A as in Definition 1.19 and similarly, construct X' from O' and A'. We need to show that $\overrightarrow{OX} = \overrightarrow{O'X'}$. By construction $\overrightarrow{OX}$ and $\overrightarrow{O'X'}$ define the same direction, namely both define the same direction as $\overrightarrow{OA} = \overrightarrow{O'A'}$. Thus, by Proposition 1.11, it suffices to show that $|OX| = |O'X'|$. Since $r > 0$, this is clear since $|OX| = r \cdot |OA| = r \cdot |O'A'| = |O'X'|$. If $r < 0$ the claim also follows since $|XO| = |OX|$. The case where $r = 0$ is trivially true. □

Proposition 1.21. Assume that a unit segment was chosen. For $\mathbf{a}, \mathbf{b} \in \mathbb{V}$ and $x, y \in \mathbb{R}$ we have

1. $0 \cdot \mathbf{a} = \overrightarrow{0}$.
2. $1 \cdot \mathbf{a} = \mathbf{a}$
3. $-1 \cdot \mathbf{a} = -\mathbf{a}$
4. $(x + y) \cdot \mathbf{a} = x \cdot \mathbf{a} + y \cdot \mathbf{a}$
5. $x \cdot (y \cdot \mathbf{a}) = (xy) \cdot \mathbf{a}$
6. $x \cdot (\mathbf{a} + \mathbf{b}) = x \cdot \mathbf{a} + x \cdot \mathbf{b}$

Proof. The first five claims follow from Proposition B.12 and Definition 1.19. We prove the last claim. If $x=0$ the statement follows from the first assertion. Moreover, we may assume that $x>0$. Indeed, if $x<0$ we have $(-x)\cdot(-(\mathbf{a}+\mathbf{b}))=(-x)\cdot(-\mathbf{a})+(-x)\cdot(-\mathbf{b})$ which follows from the case $x>0$. Now, by Proposition 1.17 and Proposition 1.20 the operations do not depend on the choice of representatives. Let $OAQB$ be a parallelogram such that $\mathbf{a}=\overrightarrow{OA}$ and $\mathbf{b}=\overrightarrow{OB}$. By Proposition B.11, there is a unique point $A'\in(OA$ such that $x\mathbf{a}=\overrightarrow{OA'}$ and there is a unique point $B'\in O$ such that $x\mathbf{b}=\overrightarrow{OB'}$. Moreover, we have $x|\mathbf{a}|=|\overrightarrow{OA'}|$ and $x|\mathbf{b}|=|\overrightarrow{OB'}|$. Let Q be the unique point such that $OA'QB'$ is a parallelogram. We have $\overrightarrow{OQ'}=x\mathbf{a}+x\mathbf{b}$ and we need to show that $x\cdot\overrightarrow{OQ}=\overrightarrow{OQ'}$. Then the points O,Q,Q' are collinear and the angles in the triangles OAQ and $OA'Q'$ are pairwise congruent by [14, Theorem 30]. It then follows from the proportionality of the sides of similar triangles (see Theorem 41 in [14]) that $|OQ'|:|OQ|=|OA'|:|OA|=r$. □

Theorem 1.22. The set of vectors $\mathbb{V}$ with vector addition and scalar multiplication is a vector space.

Proof. For the axioms of a vector space see for example [13, Definition 1.1]. They follow directly from Proposition 1.18 and Proposition 1.21. □

Lemma 1.23. Let A,B,C be three non-collinear points and let π be the unique plane containing them. Then, a point Q lies in π if and only if there exists a parallelogram $AXYQ$ with $X\in AB$ and $Y\in AC$. Moreover, if such a parallelogram exists, it is unique.

Theorem 1.24. Assume that a unit segment was chosen. Let S be a subset of $\mathbb{E}$ 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.
2. The vectors $\overrightarrow{OA},\overrightarrow{OB}$ are linearly dependent if and only if the points O,A,B are collinear.
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.

Proof. Assume that S is a line. By definition of the vector space operations (Definitions 1.16 and 1.19), $\phi_O(S)$ is stable under multiplication with scalars and vector addition. Moreover, by Proposition B.11, any two vectors represented on the line S are linearly dependent. Thus $\phi_O(S)$ is a vector subspace of dimension 1.

For the other implication, assume that $\phi_O(S)$ is a 1-dimensional vector subspace and let $\mathbf{e}$ be a basis vector. Let A be the unique point such that $\phi_O(A)=\mathbf{e}$. By the argument in the first paragraph, $\phi_O(OA)$ is a 1-dimensional vector subspace of $\mathbb{V}$. Since $\mathbf{e}$ is a basis vector for both $\phi_O(OA)$ and $\phi_O(S)$ we have $\phi_O(OA)=\phi_O(S)$ and since ϕ_O is bijective we have $OA=S$. For 2., notice that O,A,B are collinear if an only if $\overrightarrow{OB}$ belongs to the 1-dimensional vector subspace $\phi_O(OA)$ which in turn is equivalent to the two vectors being linearly dependent.

Assume now that S is a plane. Then there exist other two points A and B such that O,A,B are non-collinear. By the above paragraph $\phi_O(OA)$ and $\phi_O(OB)$ are two distinct 1-dimensional vector subspaces contained in $\phi_O(S)$. Now, for any point Q in the plane S there is a unique parallelogram $OXQY$ with $X \in OA$ and $Y \in OB$ (by Lemma 1.23). Thus $\overrightarrow{OQ} = \overrightarrow{OX} + \overrightarrow{OY}$. In other words, any vector in $\phi_O(S)$ is a linear combination of vectors in $\phi_O(OA)$ and $\phi_O(OB)$, i.e. $\phi_O(S)$ is the vector space generated by these two one dimensional subspaces. For the other implication assume that $\phi_O(S)$ is a 2-dimensional vector subspace. Let $\mathbf{e}_1, \mathbf{e}_2$ be a basis of $\phi_O(S)$ and let A, B be the unique points such that $\phi_O(A) = \mathbf{e}_1$ and $\phi_O(B) = \mathbf{e}_2$. Then O, A, B are non-collinear. Let π be the unique plane containing these three points. By the first part of the argument $\phi_O(\phi)$ is a 2-dimensional vector subspace. Since it is included in $\phi_O(S)$ we must have $\phi_O(S) = \phi_O(\pi)$. Since ϕ_O is bijective, we have $S = \pi$.

The proof for 4. and 5. is similar. □

Definition 1.25. Because of the above theorem, there is an overlap in terminology which we accept. We say that two vectors are *collinear* if they are linearly dependent. Moreover, since two vectors are linearly dependent if and only if they have the same or opposite directions, we say that such vectors are *parallel*. Similarly, we say that three vectors are *coplanar* if they are linearly dependent.

Proposition 1.26. The dimension of $\mathbb{V}$ is at least 3.

Proof. Assume for a contradiction that $\dim \mathbb{V} \leq 2$. Then $\dim \mathbb{V}$ cannot be 0 since by Axiom I.8 there are at least 3 vectors in $\mathbb{V}$. If $\dim \mathbb{V} = 1$ then by Theorem 1.24 there are no planes - this contradicts Axiom I.3 which guarantees that there are at least 3 non-collinear points, i.e. there is at least on plane by Axiom I.4. If $\dim \mathbb{V} = 2$, by Theorem 1.24, all points lie in one plane which is a contradiction with Axiom I.8. □

1.3 Affine space structure of the Euclidean space

The vector space structure deduced in Theorem 1.22 consists in particular of two maps

$$\square + \square : \mathbb{V} \times \mathbb{V} \to \mathbb{V} \quad \text{and} \quad \square \cdot \square : \mathbb{R} \times \mathbb{V} \to \mathbb{V}.$$

Moreover, with Corollary 1.15 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 so-called *translation map*

$$\square + \square : \mathbb{V} \times \mathbb{E} \to \mathbb{E} \quad \text{given by} \quad \mathbf{a} + O = X. \tag{1.1}$$

We say that *the vectors in* $\mathbb{V}$ *act on the set of points* $\mathbb{E}$ *by translations*. This is one key observation which allows us to use real vector spaces as an underlying model for the Euclidean space $\mathbb{E}$. This is made precise by the following definition.

Definition 1.27. A *real affine space* $\mathbb{A}$ is a triple $(\mathbb{P}, \mathbb{V}, t)$ where $\mathbb{P}$ is a non-empty set whose elements are called *points*, where $\mathbb{V}$ is a real vector space called *direction space of* $\mathbb{A}$, and where t is a map $\mathbb{V} \times \mathbb{P} \to \mathbb{P}$ called *translation map* which satisfies the following two axioms:

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

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

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

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

The *dimension* of the affine space $\mathbb{A}$ is by definition the dimension of $\mathbb{V}$ and is denoted by $\dim \mathbb{A}$. The set of points is rarely mentioned separately. When we refer to 'points in $\mathbb{A}$' we mean elements of $\mathbb{P}$. When dealing with an affine space it is common for the vector space $\mathbb{V}$ not to have a label. In that case, we may invoke it as $\mathrm{D}(\mathbb{A})$.

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

Theorem 1.28. The Euclidean space $\mathbb{E}$ has the structure of a real affine space of dimension $\dim \mathbb{E} \geq 3$.

Proof. Considering the set $\mathbb{P}$ of points in $\mathbb{E}$ and the set of geometric vectors $\mathbb{V}$, we observed in (1.1) the existence of a translation map $\mathbb{V} \times \mathbb{E} \to \mathbb{E}$ using Corollary 1.15. This gives $\mathbb{E}$ the structure of a real affine space by definition. The claim on the dimension follows from Proposition 1.26. $\square$

Example 1.29. The main, and in fact the only examples of finite dimensional real affine spaces are 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

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

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. Consequently, if $\dim \mathbb{V} = n$, we denote the corresponding affine space by $\mathbb{A}^n$.

Remark. The concept of a real affine space unlocks all linear algebra tools for Euclidean geometry. In this set-up we identify the set of points with $\mathbb{V}$ and view the elements of $\mathbb{V}$ in two distinct ways: as points and as vectors which act on points. A vector space has an origin, the zero vector, but on a line or 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' [4, Chapter 2].

Remark. The mathematical library *mathlib* [22] written in Lean [21] also uses the concept of affine space as the underlying model for Euclidean geometry (see [23]).

CHAPTER 2

Cartesian coordinates

Contents

2.1 Frames in dimension 2 . . . . **19**

2.1.1 Coordinates as projections 20

2.1.2 Changing frames (an example) 21

2.1.3 Orientation 22

2.2 Frames in dimension 3 . . . . **24**

2.2.1 Coordinates as projections 25

2.2.2 Changing frames (an example) 26

2.2.3 Orientation 27

2.3 Frames in dimension n **27**

2.3.1 Algorithm for changing frames 28

2.3.2 Orientation 29

Coordinates are tuples of numbers associated to points in a given object or space. For an object or space S, we seek a subset $C \subseteq \mathbb{R}^n$ and a bijective map $C \to S$ that establishes a correspondence $(x_1, \ldots, x_n) \leftrightarrow P$ between coordinates and points. Typically, there are many choices of such maps, and the specific choice determines the level of control you have over the object or space S. This choice often depends on the intended purpose or application involving S. In this chapter, we focus exclusively on Cartesian frames, which are named after René Descartes[1]. For other types of coordinate systems we refer to Appendix D.

Here is a passage from [15, Section 6.8]: "in the eighteenth century, historians of mathematics (French ones, in particular) considered Descartes *the* revolutionary who had freed them from bondage to the tedious methods of the ancient Greeks, by reducing hard geometric problems to

[1] 1596–1650

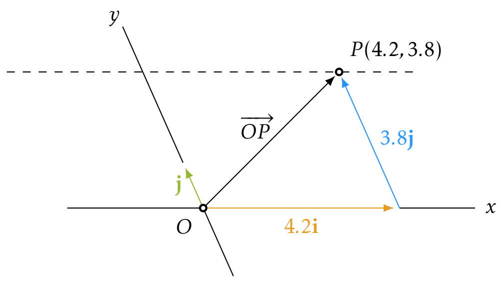
simple algebraic ones. This is a view which is often now regarded with some suspicion, although Descartes himself promoted it [...] In any case, his project was [...] specific: the relation of geometry and algebra. A standard modern textbook criticizes Descartes for not being more practical [...] This criticism is interesting, but, I think, misplaced. Coordinate geometry even today is not 'intrinsically' practical – even the statistician who studies whether points in a scatter graph lie near a straight line $y = ax + b$, let alone the geometer who wishes to picture the curve $y^2 = x^3 + x^2$ (Fig. 2) are not thinking as surveyors or geographers. On the other hand, for *some* practical tasks, the new ideas were very well adapted, as Newton and Leibniz were to understand."

It should be clear that Descartes did not consider the concept of affine space in the form that we introduced it in the previous chapter. However, the main idea is the same. Examples of classical theorems which can be proved using Cartesian frames and algebraic manipulations are given in Appendix F.

2.1 Frames in dimension 2

We denote by $\mathbb{E}^2$ an arbitrary plane. 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 1.24, $\phi_O(\ell_1)$ and $\phi_O(\ell_2)$ are 1-dimensional, linearly independent vector subspaces of the 2-dimensional vector space $\mathbb{V}^2 = \mathrm{D}(\mathbb{E}^2)$. If we select non-zero vectors $\mathbf{i} \in \phi_O(\ell_1)$ and $\mathbf{j} \in \phi_O(\ell_2)$, then $(\mathbf{i}, \mathbf{j})$ is a basis of $\mathbb{V}^2$. Consequently, there exist *unique* scalars $x_P, y_P \in \mathbb{R}$ such that

$$\phi_O(P) = \overrightarrow{OP} = x_P\mathbf{i} + y_P\mathbf{j}. \tag{2.1}$$



This establishes a bijection $\mathbb{E}^2 \leftrightarrow \mathbb{R}^2$ between points in the plane and ordered pairs of real numbers. This bijection arises as the composition of two bijections:

1. The bijection $\phi_O : \mathbb{E}^2 \to \mathbb{V}^2$, which assigns to each point P its position vector $\overrightarrow{OP}$.
2. The decomposition of vectors with respect to a basis, establishing the identification $\mathbb{V}^2 \cong \mathbb{R}^2$.

Thus, the bijection $\mathbb{E}^2 \leftrightarrow \mathbb{R}^2$ depends on two choices:

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

Definition 2.1. A *frame in* $\mathbb{E}^2$ is a pair $\mathcal{K} = (O, \mathcal{B})$, where O is a point in $\mathbb{E}^2$ and $\mathcal{B} = (\mathbf{i}, \mathbf{j})$ is a basis of $\mathbb{V}^2$. A frame is also referred to as a *Cartesian coordinate system*, or a *Cartesian frame*. We use the shorter term for convenience. Given a frame $\mathcal{K}$, the unique pair (x_P, y_P) in (2.1) is called *coordinates of* P *with respect to the frame* $\mathcal{K}$. In other words, the coordinates of P are the components of the position vector of P relative to O in the basis $\mathcal{B}$. 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 omit the subscript $\mathcal{K}$ and simply write $P(x_P, y_P)$. By convention, the first coordinate is typically denoted x, and the second y. The line ℓ_1 is called the x-axis, denoted Ox, while the line ℓ_2 is the y-axis, denoted Oy. The point O is referred to as the *origin of the frame* $\mathcal{K}$.

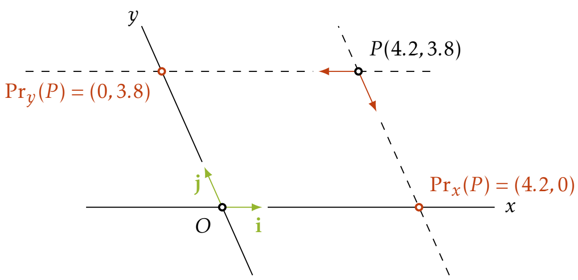
For computational purposes, points $P_{\mathcal{K}}(x_P, y_P)$ and vectors $\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j}$ are identified with column matrices, using their coordinates and components, respectively:

$$[P]_{\mathcal{K}} = \begin{bmatrix} x_P \\ y_P \end{bmatrix} \in \mathbb{R}^2 \quad \text{and} \quad [\mathbf{a}]_{\mathcal{B}} = \begin{bmatrix} a_x \\ a_y \end{bmatrix} \in \mathbb{R}^2.$$

Here, the subscript $\mathcal{K}$ indicates the frame relative to which P has the indicated coordinate, while the subscript $\mathcal{B}$ indicates the basis in which the components of the vector $\mathbf{a}$ are expressed.

2.1.1 Coordinates as projections

We explore projections extensively in Chapter 6. Here, we simply observe that the coordinates can be interpreted as projections onto coordinate axes. Let $\mathcal{K} = (O, \mathcal{B})$ be a frame in $\mathbb{E}^2$, where $\mathcal{B} = (\mathbf{i}, \mathbf{j})$ is a basis of $\mathbb{V}^2$. By definition, a point P has coordinates (x_P, y_P) if and only if $\overrightarrow{OP} = x_P\mathbf{i} + y_P\mathbf{j}$. Since $O, \mathbf{i}, \mathbf{j}$ are fixed, this is equivalent to the existence of unique points $X \in Ox$ and $Y \in Oy$ such that $OXPY$ is a parallelogram.



Definition 2.2. This defines the following maps

$$\mathrm{Pr}_x : \mathbb{E}^2 \to Ox, \quad \mathrm{Pr}_x(P) = X(x_P, 0) \quad \text{and} \quad \mathrm{Pr}_y : \mathbb{E}^2 \to Oy, \quad \mathrm{Pr}_y(P) = Y(0, y_P).$$

The map Pr_x is called *the projection on* Ox *along* Oy, and Pr_y is called *the projection on* Oy *along* Ox. For vectors $\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j}$, we have similar maps defined as follows

$$\mathrm{pr}_x : \mathbb{V}^2 \to \mathbb{R}, \quad \mathrm{pr}_x(\mathbf{a}) = a_x \quad \text{and} \quad \mathrm{pr}_y : \mathbb{V}^2 \to \mathbb{R}, \quad \mathrm{pr}_y(\mathbf{a}) = a_y.$$

The map pr_x is called *the projection on the first component*, and pr_y is called *the projection on the second component*. From the definitions we immediately deduce the following identities

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

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

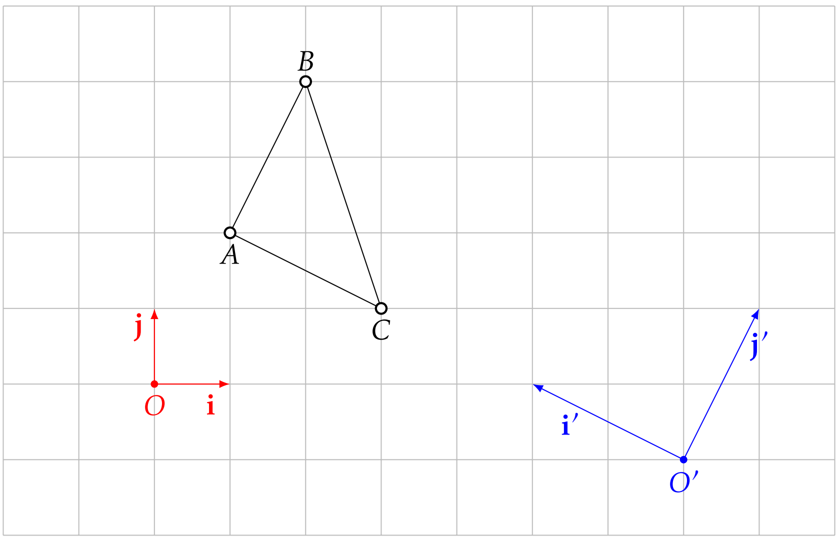
Remark. In the process of showing that $\mathbb{E}$ has the structure of an affine space, we made a choice, we fixed a unit segment. If $\mathbf{i}$ and $\mathbf{j}$ are unit vectors, then for a point P as above, we have $|OX| = x_P, |OY| = y_P$. In other words, the coordinates of P are the lengths of the sides of the parallelogram $OXPY$.

2.1.2 Changing frames (an example)

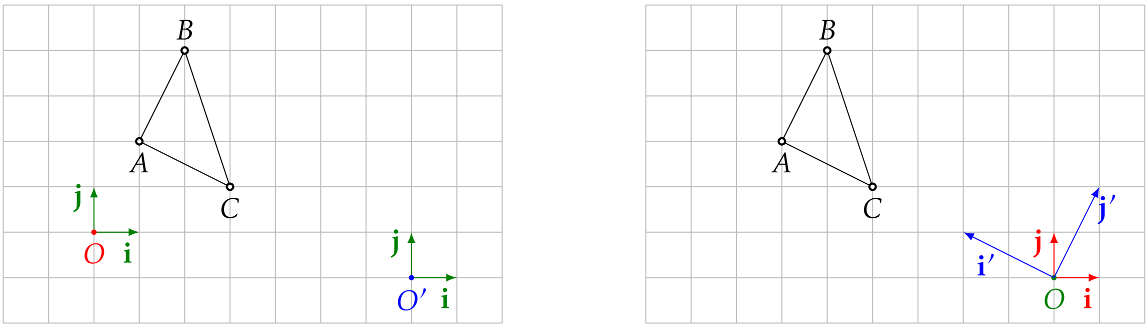
With a fixed frame $\mathcal{K}$, we may identify $\mathbb{E}^2$ with $\mathbb{R}^2$. However, a pair of numbers (x_P, y_P) has no intrinsic geometric meaning in the absence of a frame. Moreover, the coordinates of a point vary depending on the frame used. The process of translating from one Cartesian frame to another is made precise in Section 2.3. Here, we illustrate this process with an example. Let $\mathcal{K} = (O, \mathcal{B})$ and $\mathcal{K}' = (O', \mathcal{B}')$ be two frames in $\mathbb{E}^2$, with $\mathcal{B} = (\mathbf{i}, \mathbf{j})$ and $\mathcal{B}' = (\mathbf{i}', \mathbf{j}')$. Assume that O', $\mathbf{i}'$ and $\mathbf{j}'$ are known relative to $\mathcal{K}$:

$$[O']_{\mathcal{K}} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}, \quad [\mathbf{i}']_{\mathcal{B}} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad [\mathbf{j}']_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \tag{2.2}$$

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



We can achieve this in two steps: (a) first, we change the origin, i.e., we go from $(O, \mathcal{B})$ to $(O', \mathcal{B})$, and (b) we change the direction of the coordinate axes, i.e. we go from $(O', \mathcal{B})$ to $(O', \mathcal{B}')$. The first step is simply a translation, while the second corresponds to the usual base change from linear algebra.



(a) Change the origin. (b) Change the direction of the axes.

In the first step, when going from $\mathcal{K}=(O,\mathcal{B})$ to $\tilde{\mathcal{K}}=(O',\mathcal{B})$, we are looking for the components of the position vector of A relative to O' with respect to $\mathcal{B}$. We find these components by noticing that

$$\overrightarrow{O'A}=\overrightarrow{OA}-\overrightarrow{OO'}\quad \text{and therefore}\quad [\overrightarrow{O'A}]_{\mathcal{B}}=[\overrightarrow{OA}]_{\mathcal{B}}-[\overrightarrow{OO'}]_{\mathcal{B}}.$$

In the second step, when going from $\tilde{\mathcal{K}}=(O',\mathcal{B})$ to $\mathcal{K}'=(O',\mathcal{B}')$, we are looking for the components of the position vector of A relative to O' with respect to $\mathcal{B}'$. From linear algebra, we know that this is done with the base change matrix (see Appendix C). Let $\mathrm{M}_{\mathcal{B}',\mathcal{B}}$ be the base change matrix from the basis $\mathcal{B}$ to the basis $\mathcal{B}'$. Then,

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

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

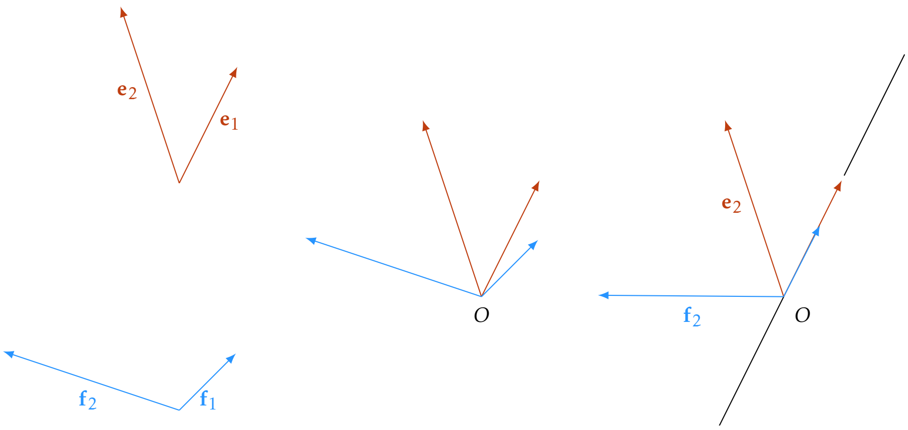
$$[A]_{\mathcal{K}'}=[\overrightarrow{O'A}]_{\mathcal{B}'}=\mathrm{M}_{\mathcal{B}',\mathcal{B}}\cdot[\overrightarrow{O'A}]_{\mathcal{B}}=\mathrm{M}_{\mathcal{B}',\mathcal{B}}\cdot\Big([\overrightarrow{OA}]_{\mathcal{B}}-[\overrightarrow{OO'}]_{\mathcal{B}}\Big) \tag{2.3}$$

Now, suppose that A has coordinates $(1,2)$ with respect to the frame $\mathcal{K}$. Since $[\overrightarrow{OO'}]_{\mathcal{B}}=[O']_{\mathcal{K}}$, by (2.2), we already know the terms in the parentheses on the right-hand side of (2.3). Furthermore, from our assumptions (2.2), it is easy to write down the matrix $\mathrm{M}_{\mathcal{B},\mathcal{B}'}$ and then $\mathrm{M}_{\mathcal{B}',\mathcal{B}}=\mathrm{M}_{\mathcal{B},\mathcal{B}'}^{-1}$. Thus

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

2.1.3 Orientation

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



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.

Proposition 2.3. Let $\mathcal{E}$ and $\mathcal{F}$ be as above.

1. If $\det(M_{\mathcal{E},\mathcal{F}}) > 0$ then $\mathbf{e}_2$ and $\mathbf{f}_2$ point to the same half-plane.
2. If $\det(M_{\mathcal{E},\mathcal{F}}) < 0$ then $\mathbf{e}_2$ and $\mathbf{f}_2$ point to different half-planes.

Proof. 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 $M_{\mathcal{F}',\mathcal{F}}$ is a 2×2-rotation matrix which has determinant equal to 1 (For more on rotations see Chapter 7). 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

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

and, since λ is positive, the proof is finished. □

Definition 2.4. In the first case of Proposition (2.3) 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*.

Why is this relevant? Next to the fact that it gives a geometric interpretation of the sign of the determinant $\det(M_{\mathcal{E},\mathcal{F}})$, it also allows us to understand some signs which appear in calculations of areas (see Chapter 5). Moreover, the trigonometry that we know to hold true in $\mathbb{E}^2$ implicitly builds on the notion of oriented Euclidean plane $\mathbb{E}^2$. 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 2.5. 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.

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.

2.2 Frames in dimension 3

The situation in dimension 3 is similar to the one discussed for dimension 2 (Section 2.1) and is a particular case of the n-dimensional case presented in Section 2.3. 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 1.24, the sets $\phi_O(\ell_1)$, $\phi_O(\ell_2)$ and $\phi_O(\ell_3)$ are 1-dimensional, linearly independent vector subspaces of the 3-dimensional vector space $\mathbb{V}^3 = \mathrm{D}(\mathbb{E}^3)$. Thus, if we choose non-zero vectors $\mathbf{i} \in \phi_O(\ell_1)$, $\mathbf{j} \in \phi_O(\ell_2)$ and $\mathbf{k} \in \phi_O(\ell_3)$, we obtain a basis $(\mathbf{i},\mathbf{j},\mathbf{k})$ of $\mathbb{V}^3$ and there exist *unique* scalars $x_P, y_P, z_P \in \mathbb{R}$ such that

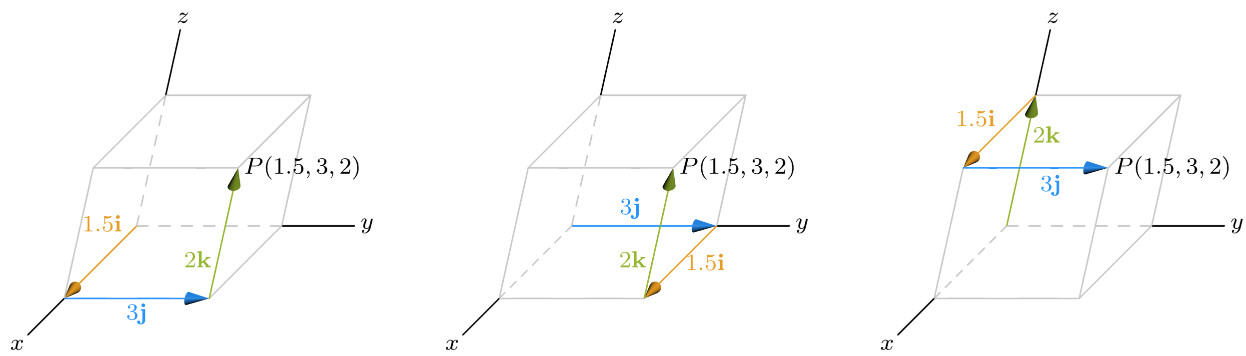
$$\phi_O(P) = \overrightarrow{OP} = x_P\mathbf{i} + y_P\mathbf{j} + z_P\mathbf{k}. \tag{2.4}$$

This defines a bijection between points $P \in \mathbb{E}^3$ and triples of real numbers (x_P, y_P, z_P) 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 2.6. A *frame in* $\mathbb{E}^3$ is a pair $\mathcal{K} = (O,\mathcal{B})$ 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 a frame $\mathcal{K}$, the unique triple (x_P, y_P, z_P) in (2.4) is the *coordinates of* P *with respect to the frame* $\mathcal{K}$. Again, the coordinates of P are just the components of the position vector of P relative to O. 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. By convention, we use x for the first coordinate, y for the second, and z for the third. The line ℓ_1 is the x-axis, denoted Ox, the line ℓ_2 is the y-axis, denoted Oy, and ℓ_3 is the z-axis, denoted Oz. The point O is the *origin of the frame* $\mathcal{K}$. The planes containing two coordinate axes are the *coordinate planes*. We let Oxy denote the plane containing Ox and Oy, and similarly for the other two.

Here again, for computational purposes, points $P(x_P, y_P, z_p)$ and vectors $\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}$ are represented as column matrices, using their coordinates and components, respectively:

$$[P]_{\mathcal{K}} = \begin{bmatrix} x_P \\ y_P \\ z_P \end{bmatrix} \in \mathbb{R}^3 \quad \text{and} \quad [\mathbf{a}]_{\mathcal{B}} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \in \mathbb{R}^3.$$



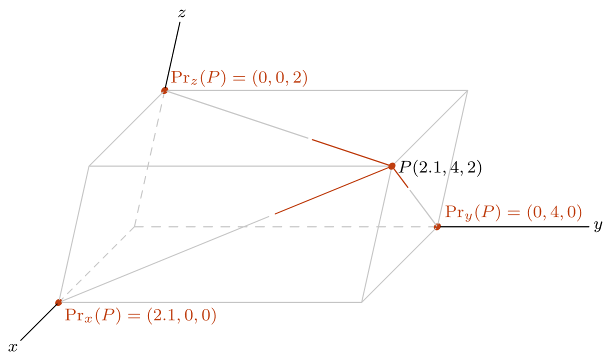
2.2.1 Coordinates as projections

Projections are explored in detail in Chapter 6. Here we invoke the 3-dimensional image for viewing coordinates as projections on coordinate axes. Let $\mathcal{K}=(O,\mathcal{B})$ be frame in $\mathbb{E}^3$, where $\mathcal{B}=(\mathbf{i},\mathbf{j},\mathbf{k})$. By definition, a point P has coordinates (x_P,y_P,z_P) if and only if $\overrightarrow{OP}=x_P\mathbf{i}+y_P\mathbf{j}+z_P\mathbf{k}$. Since $O,\mathbf{i},\mathbf{j},\mathbf{k}$ are fixed, this is equivalent to the existence of unique points $X\in Ox, Y\in Oy, Z\in Oy$ such that O,X,Y,Z,P are vertices of a parallelepiped.

In dimension 3, we have projection 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 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 $\mathbf{a}=a_x\mathbf{i}+a_y\mathbf{j}+a_z\mathbf{k}$ in $\mathbb{V}^3$, we have maps $\mathrm{pr}_x,\mathrm{pr}_y,\mathrm{pr}_z:\mathbb{V}^3\to\mathbb{R}$ defined by $\mathrm{pr}_x(\mathbf{a})=a_x$, $\mathrm{pr}_y(\mathbf{a})=a_y$, $\mathrm{pr}_z(\mathbf{a})=a_z$. From the definitions we immediately deduce the following identities

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

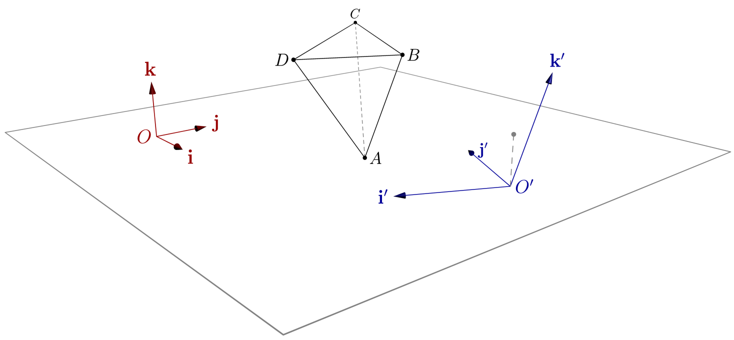
$$\overrightarrow{OP}=\mathrm{pr}_x(\overrightarrow{OP})\mathbf{i}+\mathrm{pr}_y(\overrightarrow{OP})\mathbf{j}+\mathrm{pr}_z(\overrightarrow{OP})\mathbf{k}\quad\text{and}\quad [P]_{\mathcal{K}}=\begin{bmatrix}x_P\\y_P\\z_P\end{bmatrix}=\begin{bmatrix}\mathrm{pr}_x(\overrightarrow{OP})\\\mathrm{pr}_y(\overrightarrow{OP})\\\mathrm{pr}_z(\overrightarrow{OP})\end{bmatrix}=[\overrightarrow{OP}]_{\mathcal{B}}.$$



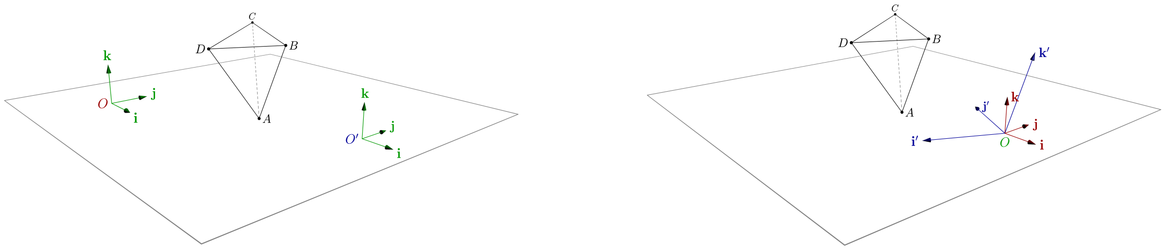
2.2.2 Changing frames (an example)

The process of translating from one Cartesian frame to another is made precise in Section 2.3. Here, we illustrate this process with an example. Let $\mathcal{K} = (O, \mathcal{B})$ and $\mathcal{K}' = (O', \mathcal{B}')$ be two frames in $\mathbb{E}^3$ with $\mathcal{B} = (\mathbf{i}, \mathbf{j}, \mathbf{k})$ and $\mathcal{B}' = (\mathbf{i}', \mathbf{j}', \mathbf{k}')$. Assume that O', $\mathbf{i}', \mathbf{j}'$ and $\mathbf{k}'$ are known relative to $\mathcal{K}$:

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



In any dimension, the coordinates with respect to one frame can be obtained from the coordinates with respect to another frame in two steps.



(a) Change the origin.

(b) Change the direction of the axes.

Let B be the point with coordinates $(1, 5, 1)$ relative to $\mathcal{K}$. The argument used for dimension 2 (Section 2.2.2) literally translates to our 3-dimensional setting and we have

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

2.2.3 Orientation

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 six 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 7).

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 in dimension 2 shows that the sign of $\det(M_{\mathcal{E},\mathcal{F}})$ gives the answer. This fact is true in any dimension (see Section 2.3.2). In dimension 3, the orientation of a basis explains some signs which appear in calculations of volumes (see Chapter 5). In relation to the physical world this distinction is more concrete.

Definition 2.7. 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 of $\mathbb{V}^3$ is left or right oriented. The Swiss liked the three-finger rule so much, they put it on their 200-franc banknotes:

2.3 Frames in dimension n

Definition 2.8. A *frame in* $\mathbb{A}^n$ is a pair $\mathcal{K} = (O, \mathcal{B})$, where O is a point in the affine space $\mathbb{A}^n$ and $\mathcal{B} = (\mathbf{i}_1, \ldots, \mathbf{i}_n)$ is a basis of the direction space $\mathrm{D}(\mathbb{A}^n)$. A frame is also called *Cartesian coordinate system*, or *Cartesian frame*. We use the shorter term for convenience. Given a frame $\mathcal{K}$, any point $P \in \mathbb{A}^n$ has a unique corresponding position vector $\overrightarrow{OP}$ (see Section 1.27) which relative to $\mathcal{B}$ has (unique) components $(x_1, \ldots, x_n)$, i.e.

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

The n-tuple $(x_1, \ldots, x_n)$ is called *coordinates of P with respect to the frame $\mathcal{K}$* and we write $P_{\mathcal{K}}(x_1, \ldots, x_n)$ when we want to indicate the coordinates, or simply $P(x_1, \ldots, x_n)$ if it is clear from the context what $\mathcal{K}$

is. The point O is the *origin of the frame* $\mathcal{K}$. For computational purposes, points $P(x_1,\dots,x_n)$ and vectors $\mathbf{a}=a_1\mathbf{i}_1+\cdots+a_n\mathbf{i}_n$ are identified with column matrices, using their coordinates and components, respectively:

$$[P]_{\mathcal{K}}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\in\mathbb{R}^n \quad\text{and}\quad [\mathbf{a}]_{\mathcal{B}}=\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix}\in\mathbb{R}^n.$$

Definition 2.9. For vectors $\mathbf{a}=a_1\mathbf{i}_1+\cdots+a_n\mathbf{i}_n$ in $\mathrm{D}(\mathbb{A}^n)$ we have *projection maps* $\mathrm{pr}_1,\dots,\mathrm{pr}_n:\mathrm{D}(\mathbb{A}^n)\to\mathbb{R}$ defined by $\mathrm{pr}_1(\mathbf{a})=a_1,\dots,\mathrm{pr}_n(\mathbf{a})=a_n$. From the definition we see that for a point $P(x_1,\dots,x_n)$ we have

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

A frame $\mathcal{K}$ allows us to identify the set of points in $\mathbb{A}^n$ with $\mathbb{R}^n$. However, an n-tuple of numbers has no geometric meaning in the absence of a frame. Moreover, with respect to different frames, points have different coordinates. The process of translating from one Cartesian frame to another is made precise with the following theorem.

Theorem 2.10. Let $\mathcal{K}=(O,\mathcal{B})$ and $\mathcal{K}'=(O',\mathcal{B}')$ be two frames in $\mathbb{A}^n$. For any point $P\in\mathbb{A}^n$ we have

$$[P]_{\mathcal{K}'}=\mathrm{M}_{\mathcal{B}',\mathcal{B}}\cdot([P]_{\mathcal{K}}-[O']_{\mathcal{K}})=\mathrm{M}^{-1}_{\mathcal{B},\mathcal{B}'}\cdot([P]_{\mathcal{K}}-[O']_{\mathcal{K}})=\mathrm{M}_{\mathcal{B}',\mathcal{B}}\cdot[P]_{\mathcal{K}}+[O]_{\mathcal{K}'}. \tag{2.5}$$

Proof. Since the coordinates of a point are the components of its position vector, Equation (2.5) is equivalent to:

$$[\overrightarrow{O'P}]_{\mathcal{B}'}=\mathrm{M}_{\mathcal{B}',\mathcal{B}}\cdot([\overrightarrow{OP}]_{\mathcal{B}}-[\overrightarrow{OO'}]_{\mathcal{B}})=\mathrm{M}^{-1}_{\mathcal{B},\mathcal{B}'}\cdot([\overrightarrow{OP}]_{\mathcal{B}}-[\overrightarrow{OO'}]_{\mathcal{B}})=\mathrm{M}_{\mathcal{B}',\mathcal{B}}\cdot[\overrightarrow{OP}]_{\mathcal{B}}+[\overrightarrow{O'O}]_{\mathcal{B}'}. \tag{2.6}$$

Since $[\overrightarrow{O'P}]_{\mathcal{B}}=[\overrightarrow{OP}]_{\mathcal{B}}-[\overrightarrow{OO'}]_{\mathcal{B}}$ and since $\mathrm{M}_{\mathcal{B}',\mathcal{B}}$ is the base change matrix from $\mathcal{B}$ to $\mathcal{B}'$, the first equality follows. The second equality follows from the property that $\mathrm{M}_{\mathcal{B}',\mathcal{B}}=\mathrm{M}^{-1}_{\mathcal{B},\mathcal{B}'}$ (see Appendix C). The last equality is obtained by opening the parentheses and noticing that

$$\mathrm{M}^{-1}_{\mathcal{B},\mathcal{B}'}[\overrightarrow{OO'}]_{\mathcal{B}}=[\overrightarrow{OO'}]_{\mathcal{B}'}=-[\overrightarrow{O'O}]_{\mathcal{B}'}.$$

□

2.3.1 Algorithm for changing frames

Let us flesh out the steps needed to translate coordinates from one Cartesian frame to another, as exemplified in Sections 2.1.2 and 2.2.2. We expand Theorem 2.5 using Appendix C.

Let $\mathcal{K}=(O,\mathcal{B})$ and $\mathcal{K}'=(O',\mathcal{B}')$ be two frames in $\mathbb{A}^n$ with $\mathcal{B}=(\mathbf{i}_1,\dots,\mathbf{i}_n)$ and $\mathcal{B}'=(\mathbf{i}'_1,\dots,\mathbf{i}'_n)$. Suppose we know $\mathcal{K}'$ in terms of $\mathcal{K}$, i.e., the coordinates of O' relative to $\mathcal{K}$ and the components of $\mathbf{i}'_1,\dots,\mathbf{i}'_n$ with respect to $\mathcal{B}$ are given. If the coordinates of a point P are given relative to $\mathcal{K}$, we can find the coordinates of P relative to $\mathcal{K}'$ with the following steps:

1. Construct the base change matrix $\mathrm{M}_{\mathcal{B},\mathcal{B}'}$ by placing $[\mathbf{i}'_1]_{\mathcal{B}},\ldots,[\mathbf{i}'_n]_{\mathcal{B}}$ in the columns of the matrix.
2. Calculate $\mathrm{M}_{\mathcal{B}',\mathcal{B}}$ by inverting $\mathrm{M}_{\mathcal{B},\mathcal{B}'}$.
3. Calculate $[P]_{\mathcal{K}'} = \mathrm{M}_{\mathcal{B}',\mathcal{B}} \cdot ([P]_{\mathcal{K}} - [O']_{\mathcal{K}})$ since $[P]_{\mathcal{K}}$ and $[O']_{\mathcal{K}}$ are given.

2.3.2 Orientation

Definition 2.11. Two bases $\mathcal{E}$ and $\mathcal{F}$ of the space $\mathbb{V}^n$ of geometric vectors are said to have the *same orientation* if $\det(\mathrm{M}_{\mathcal{E},\mathcal{F}}) > 0$. They have *opposite orientation* if $\det(\mathrm{M}_{\mathcal{E},\mathcal{F}}) < 0$. Two frames $\mathcal{K}$ and $\mathcal{K}'$ have the *same orientation* if their bases have the same orientation and they have *opposite orientation* otherwise.

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

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

CHAPTER 3

Affine subspaces

Contents

3.1 Lines in $\mathbb{A}^2$ **31**
3.1.1 Parametric equations 31
3.1.2 Cartesian equations 32
3.1.3 Relative positions of two lines in $\mathbb{A}^2$ 33
3.2 Planes in $\mathbb{A}^3$ **34**
3.2.1 Parametric equations 34
3.2.2 Cartesian equations 35
3.2.3 Relative positions of two planes in $\mathbb{A}^3$ 37
3.3 Lines in $\mathbb{A}^3$ **37**
3.3.1 Parametric equations 38
3.3.2 Cartesian equations 38
3.3.3 Relative positions of two lines in $\mathbb{A}^3$ 39
3.3.4 Relative positions of a line and a plane in $\mathbb{A}^3$ 40
3.4 Affine subspaces of $\mathbb{A}^n$ **41**
3.4.1 Hyperplanes 43
3.4.2 Lines 44
3.4.3 Relative positions 45
3.4.4 Changing the reference frame 47

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

Let $\mathbb{E}^2$ denote an arbitrary plane. It is a 2-dimensional real affine space. In order to emphasize the fact that we treat $\mathbb{E}^2$ as an affine space only we denote it by $\mathbb{A}^2$. In this setting, 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 1.24). 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 $\mathrm{D}(S)$.

3.1.1 Parametric equations

For a line S and 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, i.e. $\mathbf{v}$ 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

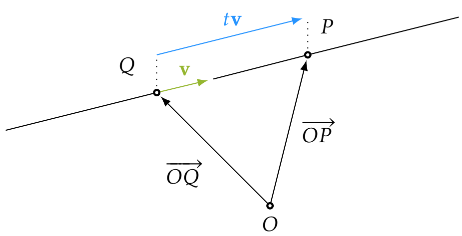
$$\overrightarrow{QP} = t\mathbf{v}.$$

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 $Q \in S$ 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* of the line. Moreover, for any point $O \in \mathbb{A}^2$ we may split the vector $\overrightarrow{QP}$ in the equation $\overrightarrow{QP} = t\mathbf{v}$ to obtain

$$\overrightarrow{OP} = \overrightarrow{OQ} + t\mathbf{v}. \tag{3.1}$$



So, we can describe the line S as the set of points P in $\mathbb{A}^2$ which satisfy Equation (3.1) 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 *a vector equation* of the line S. If $\mathbf{v} = \overrightarrow{QA}$, this equation also describes the segment $[QA]$ if we restrict the parameter to $t \in [0,1]$. Moreover, for $t > 0$ and $t < 0$ we obtain two rays amanating from Q.

Notice that, a 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 frame $\mathcal{K} = (O,\mathcal{B})$. If we write Equation (3.1) in coordinates relative to $\mathcal{K}$, we obtain

$$S : \begin{cases} x = x_Q + tv_x \\ y = y_Q + tv_y \end{cases} \quad \text{or, in matrix form} \quad S : \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_Q \\ y_Q \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \end{bmatrix} \tag{3.2}$$

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 (3.2) 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.

3.1.2 Cartesian equations

It is possible to eliminate the parameter t in (3.2). By expressing t in both equations and setting the two expressions equal, we obtain

$$\frac{x - x_Q}{v_x} = \frac{y - y_Q}{v_y}. \tag{3.3}$$

We refer to Equation (3.3) 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 3.1. 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 described by the equation $y = 5$.

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

Proposition 3.2. Every line in $\mathbb{A}^2$ can be described with a linear equation in two variables

$$ax + by + c = 0 \tag{3.4}$$

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 (3.4) 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

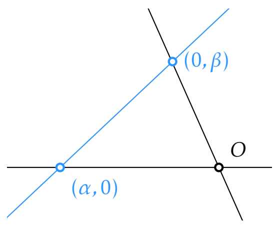
$$\mathrm{D}(\ell) : ax + by = 0.$$

Proof. This is a particular case of Theorem 3.7. □

Equation (3.4) 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 (3.4) 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 $(0,\beta)$.



Or, we may express the linear equation (3.4) with a determinant

$$\begin{vmatrix} x-x_Q & y-y_Q \\ v_x & v_y \end{vmatrix}=0$$

which says that a point P belongs to the line if $\overrightarrow{QP}$ is linearly dependent on $\mathbf{v}$. In this form we may describe the two half-planes which are separated by the line with the inequalities

$$\begin{vmatrix} x-x_Q & y-y_Q \\ v_x & v_y \end{vmatrix}<0 \quad \text{and} \quad \begin{vmatrix} x-x_Q & y-y_Q \\ v_x & v_y \end{vmatrix}>0$$

Indeed, any point P in the plane either lies on the line or $(\overrightarrow{OP},\mathbf{v})$ is a left oriented basis or $(\overrightarrow{OP},\mathbf{v})$ is a right oriented basis.

3.1.3 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

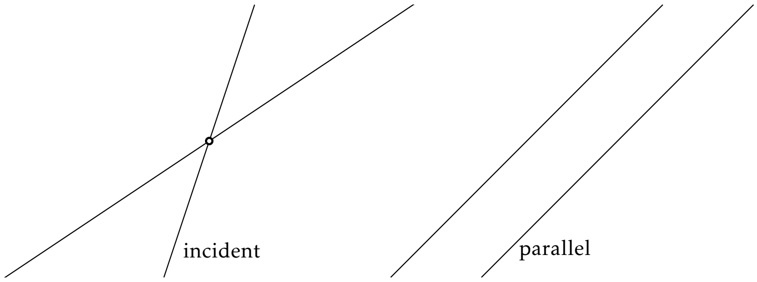
$$\ell_1 : a_1x+b_1y+c_1=0 \quad \text{and} \quad \ell_2 : a_2x+b_2y+c_2=0.$$

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

$$\begin{cases} \ell_1 : a_1x+b_1y+c_1 & = & 0 \\ \ell_2 : a_2x+b_2y+c_2 & = & 0 \end{cases}. \tag{3.5}$$

Discussing this system is basic linear algebra (see for example [9, 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 (3.5); or
- they don't intersect and (3.5) doesn't have solutions, in which case the lines are parallel; or
- System (3.5) has infinitely many solutions in which cases $\ell_1 = \ell_2$.



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

The usual Euclidean space $\mathbb{E}^3$ is a 3-dimensional real affine space In order to emphasize the fact that we treat $\mathbb{E}^3$ as an affine space only 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 1.24). Considering the bijection $\phi^3_Q : \mathbb{A}^3 \to \mathbb{V}^3$ for a point $Q \in \mathbb{A}^3$, the subset S is a plane if and only if for any $Q \in S$

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

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 $\phi^3_Q(S)$ of $\mathbb{V}^3$ the *direction space* of the plane S and denote it $\mathrm{D}(S)$.

3.2.1 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

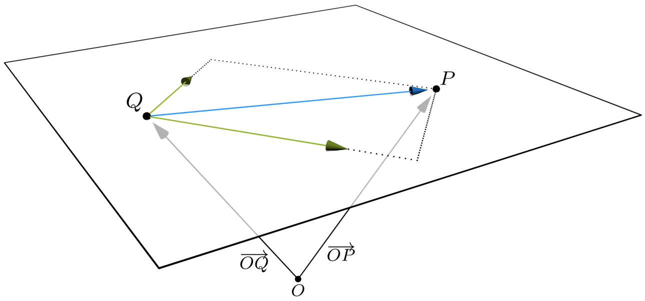
$$\overrightarrow{QP} = s\mathbf{v} + t\mathbf{w}.$$

Now, if we fix Q and let P vary in the plane S then s and t vary in $\mathbb{R}$. Since $\phi^3_Q : \mathbb{A}^3 \to \mathbb{V}^3$ 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 the vector $\overrightarrow{QP}$ in the equation $\overrightarrow{QP} = s\mathbf{v} + t\mathbf{w}$ to obtain

$$\overrightarrow{OP} = \overrightarrow{OQ} + s\mathbf{v} + t\mathbf{w}. \tag{3.6}$$



So, we can describe the plane S as the set of points P in $\mathbb{A}^3$ which satisfy Equation (3.6) 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 *a vector equation* of the plane S. If $\mathbf{v} = \overrightarrow{QA}$, $\mathbf{w} = \overrightarrow{QC}$ and $\mathbf{v} + \mathbf{w} = \overrightarrow{QB}$, this equation also described the interior of the parallelogram $QABC$ if we restrict the parameters s,t to $(0,1)$. The interior of the triangle QAC is obtained if we further impose the condition that $s+t<1$. Moreover, for $t>0$ and $t<0$ we obtain two half-planes separated by the line QA.

Notice that a 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 frame $\mathcal{K} = (O, \mathcal{B})$. If we write Equation (3.6) in coordinates relative to $\mathcal{K}$ then we obtain

$$S : \begin{cases} x = x_Q + sv_x + tw_x \\ y = y_Q + sv_y + tw_y \\ z = z_Q + sv_z + tw_z \end{cases} \quad \text{or, in matrix form} \quad S : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_Q \\ y_Q \\ z_Q \end{bmatrix} + s \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} + t \begin{bmatrix} w_x \\ w_y \\ w_z \end{bmatrix} \tag{3.7}$$

where $Q = Q_{\mathcal{K}}(x_Q, y_Q, z_Q)$, $\mathbf{v} = \mathbf{v}_{\mathcal{K}}(v_x, v_y, v_z)$, $\mathbf{w} = \mathbf{v}_{\mathcal{K}}(w_x, w_y, w_z)$. 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 (3.7) are called *parametric equations* for the plane S.

3.2.2 Cartesian equations

As in the case of the line in $\mathbb{A}^2$, it is possible to eliminate the parameters s,t in (3.7) to obtain

$$\left(\frac{v_x}{w_x} - \frac{v_z}{w_z}\right)\left(\frac{x-x_Q}{w_x} - \frac{y-y_Q}{w_y}\right) = \left(\frac{v_x}{w_x} - \frac{v_y}{w_y}\right)\left(\frac{x-x_Q}{w_x} - \frac{z-z_Q}{w_z}\right). \tag{3.8}$$

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 (3.7).

There is an easier way of describing S with a linear equation. For this, you can interpret (3.7) 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

$$\begin{vmatrix} x-x_Q & y-y_Q & z-z_Q \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{vmatrix} = 0. \tag{3.9}$$

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

$$\begin{vmatrix} x_E-x_Q & y_E-y_Q & z_E-z_Q \\ x_F-x_Q & y_F-y_Q & z_F-z_Q \\ x_G-x_Q & y_G-y_Q & z_G-z_Q \end{vmatrix} = 0. \tag{3.10}$$

Notice that Equation (3.10) 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 1.24). Notice also that if we replace the equals sign in (3.9) with inequalities, we describe the two half spaces separated by this plane since, for a point P not in the given plane, $(\overrightarrow{QP}, \mathbf{v}, \mathbf{w})$ is a basis which is either left or right oriented.

We have just described a plane with a linear equation (Equation (3.9)) relative to the coordinate system $\mathcal{K}$. The converse is also true.

Proposition 3.3. Every plane in $\mathbb{A}^3$ can be described with a linear equation in three variables

$$ax + by + cz + d = 0 \tag{3.11}$$

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 (3.11) 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

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

Proof. This is a particular case of Theorem 3.7. □

Equation (3.11) 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. Here again it may be useful to rearrange the linear equation (3.11) 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 $(0, 0, \gamma)$.

3.2.3 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

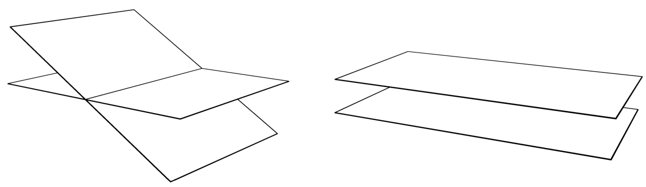
$$\pi_1 : a_1x + b_1y + c_1z + d_1 = 0 \quad \text{and} \quad \pi_2 : a_2x + b_2y + c_2z + d_2 = 0.$$

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

$$\begin{cases} \pi_1 : a_1x + b_1y + c_1z + d_1 &= 0 \\ \pi_2 : a_2x + b_2y + c_2z + d_2 &= 0 \end{cases} \tag{3.12}$$

Discussing this system is basic linear algebra (see for example [9, 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 (3.12), this happens if the rank of M and the rank of $\widetilde{M}$ equal 2; or
- they don't intersect and (3.12) 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 (3.12) 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.



3.3 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 1.24, 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

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

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

3.3.1 Parametric equations

For a line S and any two (distinct) points P,Q in S the vector $\overrightarrow{QP}$ is called a *direction vector* of S. Since $\phi_Q^3(S)$ 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_Q^3 is a bijection, the line S can be described as

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

In this description, the point Q is arbitrary but fixed. If we want to emphasize that 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 the vector $\overrightarrow{QP}$ in the equation $\overrightarrow{QP} = t\mathbf{v}$ to obtain

$$\overrightarrow{OP} = \overrightarrow{OQ} + t\mathbf{v}. \tag{3.13}$$

The image that goes with this description is the same as the one in dimension 2. 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 (3.13) 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 *a 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 frame $\mathcal{K} = (O, \mathcal{B})$. If we write Equation (3.13) in coordinates relative to $\mathcal{K}$ then we obtain

$$\begin{cases} x = x_Q + tv_x \\ y = y_Q + tv_y \\ z = z_Q + tv_z \end{cases} \quad \text{or, in matrix form} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_Q \\ y_Q \\ z_Q \end{bmatrix} + t \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} \tag{3.14}$$

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 (3.14) are called *parametric equations* for the line S.

3.3.2 Cartesian equations

It is possible to eliminate the parameter t in (3.14) in order to obtain

$$\frac{x - x_Q}{v_x} = \frac{y - y_Q}{v_y} = \frac{z - z_Q}{v_z}. \tag{3.15}$$

We refer to the Equations (3.15) 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 (3.15)) relative to the coordinate system $\mathcal{K}$. The converse is also true.

Proposition 3.4. Every line in $\mathbb{A}^3$ can be described with two linear equations in three variables

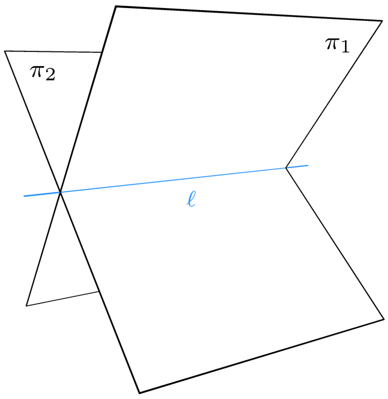
$$\begin{cases} a_1x + b_1y + c_1z + d_1 &= 0 \\ a_2x + b_2y + c_2z + d_2 &= 0 \end{cases} \tag{3.16}$$

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 (3.17) 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

$$\mathrm{D}(\ell) : \begin{cases} a_1x + b_1y + c_1z &= 0 \\ a_2x + b_2y + c_2z &= 0 \end{cases}. \tag{3.17}$$

Proof. This is a particular case of Theorem 3.7. □

The Equations (3.17) are called *Cartesian equations* of the line which they describe. Notice that they describe a line as an intersection of two planes.



3.3.3 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:

$$\begin{cases} a_1x + b_1y + c_1z + d_1 &= 0 \\ a_2x + b_2y + c_2z + d_2 &= 0 \\ a_3x + b_3y + c_3z + d_3 &= 0 \\ a_4x + b_4y + c_4z + d_4 &= 0 \end{cases}. \tag{3.18}$$

Discussing this system is basic linear algebra (see for example [9, Section 3.6]). It is somewhat easier to discuss the relative positions of lines in $\mathbb{A}^3$ via their parametric equations:

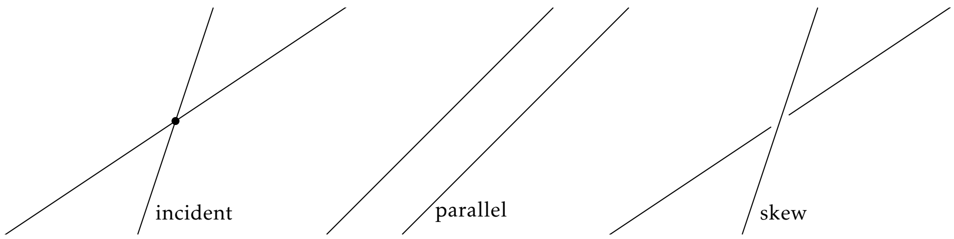
$$\ell_1 : \begin{cases} x = x_1 + tv_x \\ y = y_1 + tv_y \\ z = z_1 + tv_z \end{cases} \quad \text{und} \quad \ell_2 : \begin{cases} x = x_2 + tu_x \\ y = y_2 + tu_y \\ z = z_2 + tu_z \end{cases}.$$

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.



3.3.4 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:

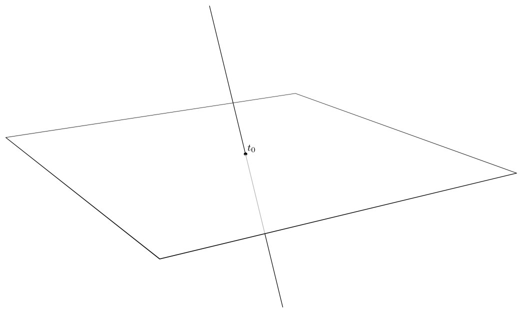
$$a(x_0 + tv_x) + b(y_0 + tv_y) + c(z_0 + tv_z) + d = 0 \quad \Leftrightarrow \quad (av_x + bv_y + cv_z)t + ax_0 + by_0 + cz_0 + d = 0. \tag{3.19}$$

The possibilities are:

- $av_x + bv_y + cv_z = 0$ and $ax_0 + by_0 + cz_0 + d \neq 0$ in which case Equation (3.19) 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 (3.19), 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 (3.19) has the unique solution

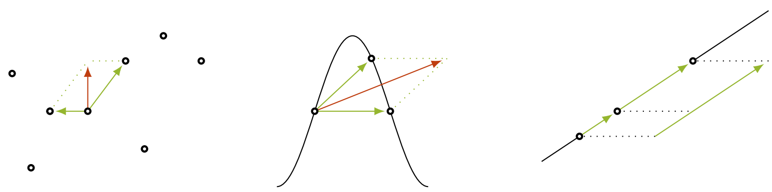
$$t_0 = -\frac{ax_0 + by_0 + cz_0 + d}{av_x + bv_y + cv_z}.$$

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



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

Definition 3.5. 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 is defined by $\dim(S) = \dim(\mathrm{D}(S))$.



Proposition 3.6. An affine subspace of $\mathbb{A}^n$ is an affine space with the affine structure inherited from $\mathbb{A}^n$.

Proof. The proof is a simple matter of unpacking the definition of affine spaces (Definition 1.27). The space $\mathbb{A}^n$ is a triple $(\mathbb{P}, \mathbb{V}, t)$ where $\mathbb{P}$ is the set of points, $\mathbb{V}$ is an n-dimensional vector space and

$t : \mathbb{V} \times \mathbb{P} \to \mathbb{P}$ is the translation map. If S is an affine subspace, it is in particular a subset of $\mathbb{P}$. The direction space $\mathrm{D}(S)$ consists of vectors in $\mathbb{V}$ which can be represented by points in S, thus $\mathrm{D}(S)$ is a vector subspace of $\mathbb{V}$. The inherited affine structure is obtained by restricting the translation map to obtain $t' : D(S) \times S \to S$. Since t satisfies the axioms of an affine space, so does t', hence the triple $(S, \mathrm{D}(S), t')$ is an affine space. □

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

$$S = \left\{ P \in \mathbb{A}^n : \overrightarrow{OP} = \overrightarrow{OQ} + t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_d\mathbf{v}_d \text{ for some } t_1, \ldots, t_d \in \mathbb{R} \right\}. \tag{3.20}$$

The equation in (3.20) is called the *vector equation of the affine subspace S relative to O, having base point Q and direction vectors* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_d$, or simply *a vector equation* of S.

Fixing a coordinate system with origin O and translating the equation in (3.20) in coordinates, one obtains *parametric equations* of the affine space S.

$$S : \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} + t_1 \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ \vdots \\ v_{1,n} \end{bmatrix} + \cdots + t_d \begin{bmatrix} v_{d,1} \\ v_{d,2} \\ \vdots \\ v_{d,n} \end{bmatrix} \quad t_1, \ldots, t_d \in \mathbb{R}. \tag{3.21}$$

Another way of representing an affine subspace is by *Cartesian equations* (Equations (3.22)) as follows.

Theorem 3.7. Fix a coordinate system $\mathcal{K} = (O, \mathcal{B})$ in the affine space $\mathbb{A}^n$. Let

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ & \vdots \\ a_{t1}x_1 + \cdots + a_{tn}x_n &= b_t \end{cases} \tag{3.22}$$

be a system of linear equations in the unknowns $x_1, \ldots, x_n$. The set S of points of $\mathbb{A}^n$ whose coordinates are solutions to (3.22), 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

$$\mathrm{D}(S) : \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ & \vdots \\ a_{t1}x_1 + \cdots + a_{tn}x_n &= 0 \end{cases} . \tag{3.23}$$

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.

Proof. We follow the proof in [19, Section 8]. Denote by $\mathbb{W}$ the vector subspace defined by the homogeneous system (3.23). First we show that the set of solutions to (3.22) is an affine subspace of $\mathbb{A}^n$

with the indicated properties. By assumption we only consider the cases where $S \neq \emptyset$. Thus, we may fix a point $Q(q_1,\dots,q_n) \in S$. Then, for any point $P(p_1,\dots,p_n)$ belonging to S we have

$$a_{j1}(p_1-q_1)+\cdots+a_{jn}(p_n-q_n) = \underbrace{(a_{j1}p_1+\cdots+a_{jn}p_n)}_{b_j} - \underbrace{(a_{j1}q_1+\cdots+a_{jn}q_n)}_{b_j} = 0$$

for each $j = 1,\dots,t$. Thus, $\overrightarrow{QP} \in \mathbb{W}$. This shows that S is contained in the affine subspace T passing through Q and parallel to $\mathbb{W}$. Conversely, if $R(r_1,\dots,r_n) \in T$, then $\overrightarrow{QR} \in \mathbb{W}$ an so the components $(x_1-q_1,\dots,r_n-q_n)$ of $\overrightarrow{QR}$ are solutions to (3.23).Thus,

$$0 = a_{j1}(r_1-q_1)+\cdots+a_{jn}(r_n-q_n) = a_{j1}r_1+\cdots+a_{jn}r_n - \underbrace{(a_{j1}q_1+\cdots+a_{jn}q_n)}_{b_j}$$

for each $j = 1,\dots,t$. That is, $R \in S$. Thus $S = T$, hence S is an affine subspace. Moreover, $\dim(S) = \dim(\mathbb{W}) = n-$ the rank fo the matrix of coefficients of (3.23).

Next we show that an affine subspace $S \subseteq \mathbb{A}^n$ has a description by a linear system of the form (3.22). Let S be any affine subspace of $\mathbb{A}^n$ with direction space $\mathbb{W}$ of dimension s. Being an s-dimensional subspace of $\mathbb{V}$, $\mathbb{W}$ can be described by a homogeneous system with $n-s$ equations

$$\mathbb{W}: \left\{ \begin{array}{rcl} a_{11}x_1+\cdots+a_{1n}x_n & = & 0 \\ & \vdots & \\ a_{n-s,1}x_1+\cdots+a_{n-s,n}x_n & = & 0 \end{array} \right. .$$

Fix a point $Q \in S$. The points $P(p_1,\dots,p_n)$ of S are characterized by the condition that $\overrightarrow{QP} \in \mathbb{W}$, i.e.

$$a_{j1}(p_1-q_1)+\cdots+a_{jn}(p_n-q_n) = 0$$

for each $j = 1,\dots,n-s$, equivalently,

$$a_{j1}p_1+\cdots+a_{jn}p_n = b_j$$

where we have put $b_j = a_{j1}q_1+\cdots+a_{jn}q_n$. Thus, the points $P(p_1,\dots,p_n) \in S$ are precisely those points in $\mathbb{A}^n$ whose coordinates satisfy the equations

$$\left\{ \begin{array}{rcl} a_{11}x_1+\cdots+a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{n-s,1}x_1+\cdots+a_{n-s,n}x_n & = & b_{n-s} \end{array} \right. .$$

□

3.4.1 Hyperplanes

Definition 3.8. 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 frame $\mathcal{K} = (O,\mathbf{e}_1,\dots,\mathbf{e}_n)$ of $\mathbb{A}^n$, parametric equations of H are of the from

$$H:\begin{cases} x_1 &= q_1+t_1v_{1,1}+\dots+t_{n-1}v_{n-1,1}\\ x_2 &= q_2+t_1v_{1,2}+\dots+t_{n-1}v_{n-1,2}\\ &\vdots\\ x_n &= q_n+t_1v_{1,n}+\dots+t_{n-1}v_{n-1,n}\end{cases} \quad\text{or}\quad H:\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}q_1\\q_2\\\vdots\\q_n\end{bmatrix}+t_1\begin{bmatrix}v_{1,1}\\v_{1,2}\\\vdots\\v_{1,n}\end{bmatrix}+\dots+t_{n-1}\begin{bmatrix}v_{n-1,1}\\v_{n-1,2}\\\vdots\\v_{n-1,n}\end{bmatrix} \tag{3.24}$$

where $\mathbf{v}_i=\mathbf{v}_i(v_{i,n},\dots,v_{i,n})$, where $Q=Q(q_1,\dots,q_n)$ is a point in H and where $t_i\in\mathbb{R}$ 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

$$\begin{vmatrix} x_1-q_1 & x_2-q_2 & \dots & x_n-q_n\\ v_{1,1} & v_{1,2} & \dots & v_{1,n}\\ \vdots & \vdots & & \vdots\\ v_{n-1,1} & v_{n-1,2} & \dots & v_{n-1,n}\end{vmatrix}=0. \tag{3.25}$$

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

3.4.2 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 description is the same as the one in dimension 2, but here we interpret it in the n-dimensional space $\mathbb{A}^n$. In coordinates, relative to a given frame $\mathcal{K}$ 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 $Q=Q(q_1,\dots,q_n)$ and $\mathbf{v}=\mathbf{v}(v_1,\dots,v_n)$ relative to $\mathcal{K}$. Here again it is possible to eliminate the parameter t in order to obtain *symmetric equations* of the line ℓ:

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

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 describes a hyperplane. So, this is saying that a line in $\mathbb{A}^n$ can be described by the intersection of $n-1$ hyperplanes.

3.4.3 Relative positions

Let S and T be two affine subspaces of $\mathbb{A}^n$. If $S \parallel T$ (see Definition 3.5) then they are disjoint or one is included in the other (see Proposition 3.9 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 3.9. 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 $S = T$.

Proof. Fix $Q \in S \cap T$. Since S and T are parallel, without loss of generality we may assume that $\mathrm{D}(S) \subseteq \mathrm{D}(T)$. Then, we obtain 1.) from

$$S = \{P \in \mathbb{A}^n : \overrightarrow{QP} \in \mathrm{D}(S)\} \subseteq \{P \in \mathbb{A}^n : \overrightarrow{QP} \in \mathrm{D}(T)\} = T.$$

We obtain 2.) by noticing that $\dim(S) = \dim(T)$ implies equality in the above equation. Indeed, by definition $\dim(S) = \dim(T)$ means that $\dim(\mathrm{D}(S)) = \dim(\mathrm{D}(T))$, but then $\mathrm{D}(S)$ is a vector subspace of $\mathrm{D}(T)$ of maximal dimension, hence $\mathrm{D}(S) = \mathrm{D}(T)$. □

As a consequence of Proposition 3.9 we obtain the following corollary which implies the 'parallel postulate' of Euclidean geometry (Axiom IV in Appendix A). The axioms of affine spaces therefore imply the validity of this postulate.

Corollary 3.10. If S is an affine subspace of $\mathbb{A}^n$ and Q a point in $\mathbb{A}^n$, there is a unique affine subspace T of $\mathbb{A}^n$ which contains Q, is parallel to S and has the same dimension as S.

Proof. Fix a point $Q \in \mathbb{A}^n$ and a point $P \in S$. To see that T exists, translate all points of S with $\overrightarrow{PQ}$, i.e. consider

$$T = \overrightarrow{PQ} + S = \overrightarrow{PQ} + \{P' \in \mathbb{A}^n : \overrightarrow{PP'} \in \mathrm{D}(S)\} = \{P' \in \mathbb{A}^n : \overrightarrow{QP'} \in \mathrm{D}(S)\}$$

where for the last equality we use $\overrightarrow{PQ} + \overrightarrow{PP'} = \overrightarrow{QP}$. Then T is an affine subspace with $\mathrm{D}(T) = \mathrm{D}(S)$, in particular it is parallel to S and $\dim(T) = \dim(S)$. To see that T is unique, assume that T' is another affine subspace passing through Q which is parallel to S and of the same dimension. By point 2.) of Proposition 3.9 we see that T' has to equal T. □

Definition 3.11. 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

$$S : \sum_{j=1}^{n} a_{ij} x_j = b_i \quad \text{for} \quad i = 1, \ldots, n-s \tag{3.26}$$

$$T : \sum_{j=1}^{n} c_{kj} x_j = d_k \quad \text{for} \quad k = 1, \ldots, n-t. \tag{3.27}$$

The intersection $S \cap T$ is the locus of points in $\mathbb{A}^n$ whose coordinates are simultaneously solutions to both (3.26) and (3.27), i.e. they are solutions to the system

$$S \cap T : \begin{cases} \sum_{j=1}^{n} a_{ij} x_j = b_i & \text{for} \quad i = 1, \ldots, n-s, \\ \sum_{j=1}^{n} c_{kj} x_j = d_k & \text{for} \quad k = 1, \ldots, n-t. \end{cases} \tag{3.28}$$

By Theorem 3.7, if the System (3.28) 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 3.12. If the intersection $S \cap T$ of two affine subspaces of $\mathbb{A}^n$ is non-empty it is an affine subspace satisfying

$$\dim(S) + \dim(T) - \dim(\mathbb{A}^n) \leq \dim(S \cap T) \leq \min\Big\{\dim(S), \dim(T)\Big\}. \tag{3.29}$$

Proof. Let $s = \dim(S)$ and let $t = \dim(T)$. By Theorem 3.7 we have

$$\begin{cases} S : \sum_{j=1}^{n} a_{ij} x_j = 0 & \text{for } i = 1, \ldots, n-s \\ T : \sum_{j=1}^{n} b_{ij} x_j = 0 & \text{for } i = 1, \ldots, n-t \end{cases}$$

Since $S \cap T$ is non-empty, the above system is compatible and the dimension of $S \cap T$ is $n - r$ where r is the rank of the matrix of coefficients of this system. Notice that

$$r \leq n - s + n - t = 2n - (s + t)$$

Thus,

$$\dim(S \cap T) = n - r \geq n - [2n - (s + t)] = s + t - n = \dim(S) + \dim(T) - \dim(\mathbb{A}^n).$$

The last inequality is clear since $S \cap T \subseteq S, T$ implies $\mathrm{D}(S \cap T) \subseteq \mathrm{D}(S), \mathrm{D}(T)$ and therefore $\dim(S \cap T) \leq \dim(S), \dim(T)$. □

Notice that in the previous proposition 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 3.13. Let S and T be two affine subspaces of $\mathbb{A}^n$ and denote the direction space of $\mathbb{A}^n$ by $\mathbb{V}^n$. Then $\mathbb{V}^n = \mathrm{D}(S) + \mathrm{D}(T)$ if and only if $S \cap T \neq \emptyset$ *and*

$$\dim(S \cap T) = \dim(S) + \dim(T) - \dim(\mathbb{A}^n). \tag{3.30}$$

Proof. We use Grassmann's identity which states that for any vector subspaces $\mathbb{W}$ and $\mathbb{U}$ we have

$$\dim(\mathbb{W}) + \dim(\mathbb{U}) = \dim(\mathbb{W} \cap \mathbb{U}) + \dim(\mathbb{W} + \mathbb{U}).$$

Let $s = \dim(S)$ and let $t = \dim(T)$. By Theorem 3.7 we have

$$\begin{cases} S : \sum_{j=1}^{n} a_{ij} x_j = 0 & \text{for } i = 1, \dots, n-s \\ T : \sum_{j=1}^{n} b_{ij} x_j = 0 & \text{for } i = 1, \dots, n-t \end{cases}$$

By convention $\dim(\emptyset) = -\infty$, thus (3.30) holds only if $S \cap T \neq \emptyset$. Assume that $S \cap T$ is non-empty. Then, the above system is compatible and the dimension of $S \cap T$ is $n-r$ where r denotes the rank of the matrix of coefficients of this system. Moreover we notice that (3.30) is equivalent to

$$n - r = s + t - n \quad \Leftrightarrow \quad \dim(\mathrm{D}(S \cap T)) = \dim(\mathrm{D}(S)) + \dim(\mathrm{D}(T)) - \dim(\mathbb{V}^n)$$

and since $S \cap T \neq \emptyset$ this is in turn equivalent to

$$\dim(\mathrm{D}(S) \cap \mathrm{D}(T)) = \dim(\mathrm{D}(S)) + \dim(\mathrm{D}(T)) - \dim(\mathbb{V}^n)$$

rearranging the equation and using Grassmann's identity the equation is equivalent to

$$\dim(\mathbb{V}^n) = \dim(\mathrm{D}(S)) + \dim(\mathrm{D}(T)) - \dim(\mathrm{D}(S) \cap \mathrm{D}(T)) = \dim(\mathrm{D}(S) + \mathrm{D}(T)).$$

At this point we use the fact that the vector subspace $\mathrm{D}(S) + \mathrm{D}(T)$ of $\mathbb{V}^n$ has maximal dimension if and only if it equals the ambient space, i.e. if and only if $\mathrm{D}(S) + \mathrm{D}(T) = \mathbb{V}^n$. $\square$

3.4.4 Changing the reference frame

Let S be an affine subspace of $\mathbb{A}^n$ given with respect to the frame $\mathcal{K} = (O, \mathcal{B})$ via the parametric equations (3.21). Then, if $\mathcal{K}' = (O', \mathcal{B}')$ is another frame, by Theorem 2.10, parametric equations with respect to $\mathcal{K}'$ are

$$S : \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{bmatrix} = \mathrm{M}_{\mathcal{B}',\mathcal{B}} \cdot \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} + [O]_{\mathcal{K}'} + t_1 \cdot \mathrm{M}_{\mathcal{B}',\mathcal{B}} \cdot \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ \vdots \\ v_{1,n} \end{bmatrix} + \dots + t_d \cdot \mathrm{M}_{\mathcal{B}',\mathcal{B}} \cdot \begin{bmatrix} v_{d,1} \\ v_{d,2} \\ \vdots \\ v_{d,n} \end{bmatrix}.$$

In terms of Cartesian equations, notice that (3.22) 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

$$S : \underbrace{\left(A \cdot \mathrm{M}_{\mathcal{B},\mathcal{B}'}\right)}_{=A'} \cdot \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix} = \underbrace{b - A \cdot [O']_{\mathcal{K}}}_{=b'}.$$

CHAPTER 4

Euclidean space

Contents

4.1 Angles **49**
4.1.1 Orthonormal frames 52
4.1.2 Oriented angles 52
4.2 Scalar product **55**
4.2.1 The Euclidean space $\mathbb{R}^n$ (first revision) 56
4.2.2 Gram-Schmidt algorithm 57
4.2.3 Normal vectors 58
4.2.4 Angles between lines and hyperplanes 59
4.3 Distance **61**
4.3.1 Distance from a point to a hyperplane 61
4.3.2 Loci of points equidistant from affine subspaces 62
4.3.3 Convergence 65

4.1 Angles

In Chapter 1 we extracted the information that the Axioms encode in two points and arrived at the notion of affine space. One natural way to proceed is to consider the information that the Axiom encodes in three points. Three non-collinear points define a plane, 12 angles, a triangle and its area, etc. Here, we focus on angles which is a second key concept in the formulation of the Axioms.

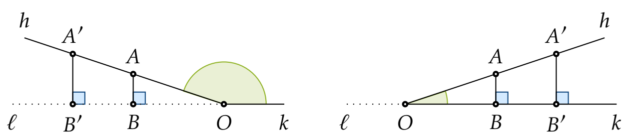
There was much debate among philosophers as to the particular category (according to the Aristotelian scheme) in which an angle should be placed; is it, namely, a *quantity*, a *quality* or a *relation* (see [10, p.177]). Such questions are for philosophers. Based on Hilbert's Axioms we will take an angle $\measuredangle(h,k)$ to be the information carried by two rays h and k, emanating from a common point O. One can deduce from the Axioms that an angle defines a unique plane which the two rays separate in a convex and a concave subset. Visually, it is common to identify an angle with the convex subset they define. As a matter of notation, for points $A \in h$ and $B \in k$ we may write $\measuredangle AOB$ for the angle $\measuredangle(h,k)$. Notice that if three points O, A, B are given, the symbols $\measuredangle AOB$ require that $A \neq O$ and $B \neq O$.

In this section we derive properties of angles needed to introduce the scalar product. Standard results concerning the trigonometry of the oriented Euclidean plane are deduced in Appendix I. Throughout, we are interested in properties of angles up to congruence, i.e. in those properties which all congruent angles share. From Axioms III.4 and III.5 one can deduce that congruence of angles is indeed an equivalence relation and one may consider equivalence classes of angles under this relation. This treatment is implicit in the notion of angles between two vectors. Instead of adding more notation, we will simply say *angle up to congruence* to mean that the angle may be replaced with a congruent angle.

Two lines which intersect in exactly one point form four angles. The opposite angles are congruent (see [14]) and the adjacent angles are called *supplementary*. A *right angle* is an angle which is congruent to its supplementary angle. If one of the angles of two intersecting lines is a right angle then all of them are right angles and we say that the lines are *orthogonal*, or that the lines are *perpendicular* to eachother. An *acute angle* is an angle less than a right angle and an *obtuse angle* is an angle greater than a right angle.

Definition 4.1. Let $\measuredangle(h,k)$ be an angle with the two rays emanating from the point O. Let ℓ be the line containing k and choose $A \in h$ and $B \in \ell$ such that AB is orthogonal to ℓ. The *sine* and *cosine of the angle* $\measuredangle(h,k)$ are defined by

$$\sin\measuredangle(h,k) = \frac{|AB|}{|OA|} \quad \text{and} \quad \cos\measuredangle(h,k) = \begin{cases} 0 & \text{if the angle is a right angle;} \\ \frac{|OB|}{|OA|} & \text{if the angle is acute;} \\ -\frac{|OB|}{|OA|} & \text{if the angle is obtuse.} \end{cases}$$

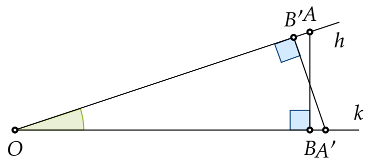


Proposition 4.2. The sine and cosine of an angle are well defined. Moreover, the following hold:

1. For an angle θ we have $\sin(\theta) \in [0,1]$, $\cos(\theta) \in [-1,1]$ and $\cos(\theta)^2 + \sin(\theta)^2 = 1$.
2. Two angles are congruent if and only if their cosines are equal.
3. Two angles have the same sine if and only if they are congruent or supplementary up to congruence.

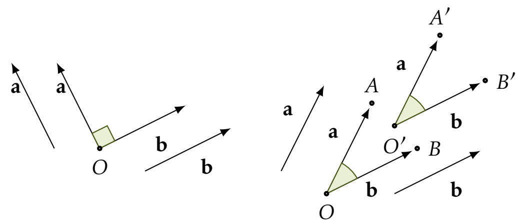
Proof. The sine function and the cosine function each attribute a real value to an angle by means of certain choices. We need to show that these values are independent of the choices made. We show this for acute angles, the other cases are similar. Let $\measuredangle(h,k)$ be an acute angle and let O,A,B be as in Definition 4.1. Consider other two points $A' \in h$ and $B' \in k$ such that $A'B'$ is orthogonal to k. By Thales' Intercept Theorem (see Theorem F.1) the ratio $|A'B'|/|OA'|$ equals $|AB|/|OA|$ and the ratio $|OB'|/|OA'|$ equals $|OB|/|OA|$.

It remains to show that the definition does not depend on the order of the two rays. By Lemma 1.4 there is a unique point $A' \in k$ such that $[OA]$ is congruent to $[OA']$. Let $B' \in h$ be such that $A'B'$ is orthogonal to h. By the Second Congruence Theorem [14, Theorem 13], the triangles OAB and $OB'A$ are congruent, hence the ratio $|OB'|/|OA'|$ equals $|A'B|/|B'A|$ and the ratio $|OB'|/|OA'|$ equals $|OB|/|OA|$.



It remains to consider the last three claims. Claim 1. follows from the fact that the length of the catheti in a right angle triangle are always less than the length of the hypothenuse and, $\cos(\theta)^2 + \sin(\theta) = 1$ follows from Pythagoras' Theorem. Since congruence of triangles is an equivalence relation, Claim 2. and 3. can be deduced with Axiom III.4. □

Definition 4.3. For two non-zero vectors $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$, the *unoriented angle between* $\mathbf{a}$ *and* $\mathbf{b}$, denoted $\measuredangle(\mathbf{a},\mathbf{b})$, is the angle $\measuredangle AOB$ up to congruence. By Proposition 4.2, the values of sine and cosine do not change under congruence, thus, the *sine* $\sin\measuredangle(\mathbf{a},\mathbf{b})$ and *cosine* $\cos\measuredangle(\mathbf{a},\mathbf{b})$ *of the unoriented angle* $\measuredangle(\mathbf{a},\mathbf{b})$ are well defined. If $\cos\measuredangle(\mathbf{a},\mathbf{b}) = 0$ we say that $\mathbf{a}$ *and* $\mathbf{b}$ *are orthogonal* and we write $\mathbf{a} \perp \mathbf{b}$. We denote the set of all unoriented angles by $\mathbb{W}$ (from the German word 'Winkel').

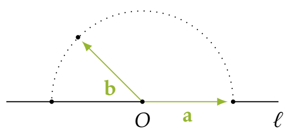


Proposition 4.4. The unoriented angle of two (non-zero) vectors is well defined. Moreover, for any two non-zero vectors $\mathbf{a},\mathbf{b}$ and a real number $x>0$, the following hold:

1. $\measuredangle(\mathbf{a},\mathbf{b}) = \measuredangle(x\mathbf{a},\mathbf{b}) = \measuredangle(\mathbf{a},x\mathbf{b})$,
2. $\measuredangle(-x\mathbf{a},\mathbf{b}) = \measuredangle(\mathbf{a},-x\mathbf{b})$,
3. $\cos\measuredangle(\mathbf{a},\mathbf{b}) = -\cos\measuredangle(-\mathbf{a},\mathbf{b})$,
4. $\sin\measuredangle(\mathbf{a},\mathbf{b}) = \cos\measuredangle(-\mathbf{a},\mathbf{b})$.

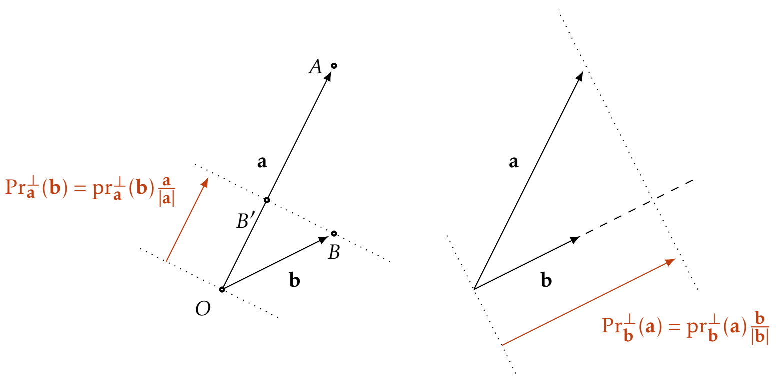
Proof. Let $\mathbf{a},\mathbf{b}\in\mathbb{V}^2$ be two non-zero vectors. Given a point $O\in\mathbb{E}^2$, there are unique points $A,B\in\mathbb{E}^2$ such that $\mathbf{a}=\overrightarrow{OA}$ and $\mathbf{b}=\overrightarrow{OB}$. Thus we obtain an angle $\measuredangle AOB$. For a different point $O'\in\mathbb{E}^2$, 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. □

Remark. Let ℓ be a line in a plane π and let O be a point on ℓ. Choose a side of π relative to ℓ and consider the semicircle $\mathbb{S}^{\frac{1}{2}} = \left\{ A : \text{ with } A \text{ on the given side of } \ell \text{ or on } \ell \text{ and } |OA|=1 \right\}$. It follows from Proposition 4.4 that the set of unoriented angles $\mathbb{W}$ is in bijection with $\mathbb{S}^{\frac{1}{2}}$.



Definition 4.5. For a vector $\mathbf{a}$ we have *orthogonal projection maps* $\mathrm{pr}^{\perp}_{\mathbf{a}}:\mathbb{V}\to\mathbb{R}$ and $\mathrm{Pr}^{\perp}_{\mathbf{a}}:\mathbb{V}\to\mathbb{V}$ defined as follows. For a vector $\mathbf{b}$ let O,A,B be such that $\mathbf{a}=\overrightarrow{OA}$, $\mathbf{b}=\overrightarrow{OB}$. Drop a perpendicular BB' on OA with $B'\in OA$. Then

$$\mathrm{Pr}^{\perp}_{\mathbf{a}}(\mathbf{b}) = \overrightarrow{OB'} \quad\text{and}\quad \mathrm{Pr}^{\perp}_{\mathbf{a}}(\mathbf{b}) = \mathrm{pr}^{\perp}_{\mathbf{a}}(\mathbf{b})\frac{\mathbf{a}}{|\mathbf{a}|}.$$



Proposition 4.6. The orthogonal projection map on vectors is a well defined linear map. Moreover, for two vectors $\mathbf{a}$ and $\mathbf{b}$ we have

$$\cos\measuredangle(\mathbf{a},\mathbf{b}) = \frac{\mathrm{pr}^{\perp}_{\mathbf{a}}(\mathbf{b})}{|\mathbf{b}|} = \frac{\mathrm{pr}^{\perp}_{\mathbf{b}}(\mathbf{a})}{|\mathbf{a}|}.$$

Proof. Showing that the map is well defined is left as an exercise. It is enough to show that $\mathrm{pr}^{\perp}_{\mathbf{a}}$ is linear, since then

$$\mathrm{Pr}^{\perp}_{\mathbf{a}}(x\mathbf{b}+y\mathbf{c}) = \mathrm{pr}^{\perp}_{\mathbf{a}}(x\mathbf{b}+y\mathbf{c})\frac{\mathbf{a}}{|\mathbf{a}|} = x\,\mathrm{pr}^{\perp}_{\mathbf{a}}(\mathbf{b})\frac{\mathbf{a}}{|\mathbf{a}|} + y\,\mathrm{pr}^{\perp}_{\mathbf{a}}(\mathbf{c})\frac{\mathbf{a}}{|\mathbf{a}|} = x\,\mathrm{Pr}^{\perp}_{\mathbf{a}}(\mathbf{b}) + y\,\mathrm{Pr}^{\perp}_{\mathbf{a}}(\mathbf{c}).$$

Let $\mathbf{i}$ be the unit vector $\frac{\mathbf{a}}{|\mathbf{a}|}$ and let $\mathbf{j}$ be a vector orthogonal to $\mathbf{i}$. The linearity of $\mathrm{pr}^{\perp}_{\mathbf{a}}$ follows from the fact that it is a projection on the coordinate x-axis of the frame having O as origin and $(\mathbf{i},\mathbf{j})$ as basis. The last claim follows from the definition of cosine. □

4.1.1 Orthonormal frames

Definition 4.7. A basis $\mathcal{B} = (\mathbf{e}_1,\dots,\mathbf{e}_n)$ of $\mathbb{V}^n$ is called *orthogonal* if the vectors $\mathbf{e}_i$ are mutually orthogonal, i.e. if $\mathbf{e}_i \perp \mathbf{e}_j$ for all $i,j\in\{1,\dots,n\}$ 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.

In dimension 2, the existence of orthonormal frames is a consequence of the existence of right angles. In general the existence of these frames follows from the defining properties of the scalar product (Proposition 4.15) in the context of bilinear forms (see Corollary H.9) or constructively using the Gram-Schmidt algorithm (see Section 4.2.2).

Notice that by Proposition 4.6, for any vector $\mathbf{a}$ we have $\mathrm{pr}^{\perp}_{\mathbf{e}_i}(\mathbf{a}) = |\mathbf{a}|\cos\measuredangle(\mathbf{a},\mathbf{e}_i)$. Thus, with respect to an orthonormal frame the coordinates of a point $P(x_1,\dots,x_n)$ and its position vector $\mathbf{a} = \overrightarrow{OP}$ are

$$[P]_{\mathcal{K}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} |\mathbf{a}|\cos\measuredangle(\mathbf{a},\mathbf{e}_1) \\ \vdots \\ |\mathbf{a}|\cos\measuredangle(\mathbf{a},\mathbf{e}_n) \end{bmatrix} = [\overrightarrow{OP}]_{\mathcal{B}}.$$

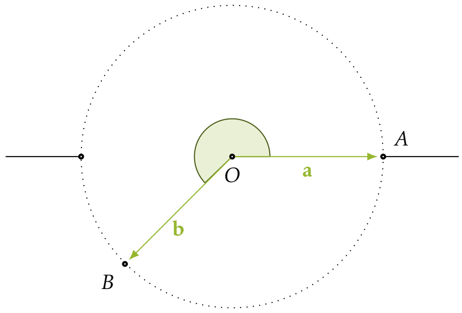
4.1.2 Oriented angles

Proposition 4.4 shows that the set of (unoriented) angles can be represented with a semicircle. If we are in dimension 2, i.e., when considering the Euclidean plane $\mathbb{E}^2$, we may consider the exterior of an angle to be an angle as well. This only works in dimension 2 because lines are hyperplanes here.

Definition 4.8. For an angle $\measuredangle(h,k)$ denote by $-\measuredangle(h,k)$ the *exterior* of the angle, i.e. the region in the plane which is *not between* the two rays h and k. Assume that a right oriented frame in $\mathbb{E}^2$ has been choosen. Let h and k be two rays emanating from the same point O and let $A\in h$ and $B\in k$. The oriented angle defined by h and k is

$$\measuredangle_{\mathrm{or}}(h,k) = \begin{cases} \measuredangle(h,k) & \text{if } (\overrightarrow{OA},\overrightarrow{OB}) \text{ is right-oriented,} \\ -\measuredangle(h,k) & \text{otherwise.} \end{cases}$$

For two vectors $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$ the oriented angle $\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})$ is the oriented angle defined by the rays $(OA$ and $(OB$. The orientation of $\measuredangle_{\text{or}}(h,k)$ is the orientation of the basis $(\mathbf{a},\mathbf{b})$. We denote the set of all oriented angles up to congruence by $\mathbb{W}_{\text{or}}$.

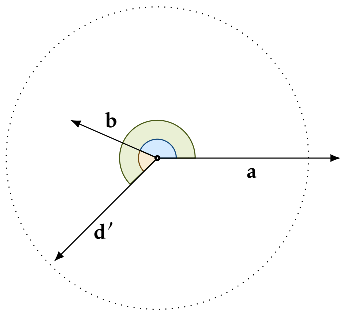


Proposition 4.9. There is a bijection between the set $\mathbb{W}_{\text{or}}$ of oriented angles and the unit circle $\mathbb{S}^1$.

Lemma 4.10. Fix an orientation in $\mathbb{E}^2$, an oriented angle $\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})$ and a length x. For any non-zero vector $\mathbf{c}$ there is a unique vector $\mathbf{d}$ of length x such that $\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b}) = \measuredangle_{\text{or}}(\mathbf{c},\mathbf{d})$.

Definition 4.11 (Counterclockwise sum of angles)**.** We define the sum of two oriented angles $\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})$ and $\measuredangle_{\text{or}}(\mathbf{c},\mathbf{d})$ as follows. By Lemma 4.10, there is a unique unit vector $\mathbf{d}'$ such that $\measuredangle_{\text{or}}(\mathbf{c},\mathbf{d}) = \measuredangle_{\text{or}}(\mathbf{b},\mathbf{d}')$ and we define

$$\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b}) + \measuredangle_{\text{or}}(\mathbf{c},\mathbf{d}) = \measuredangle_{\text{or}}(\mathbf{a},\mathbf{d}')$$



Proposition 4.12. The set $\mathbb{W}_{\text{or}}$ of oriented angles with addition is an abelian group.

Definition 4.13. For a vector $\mathbf{v} \in \mathbb{V}^2$ let $\mathbf{J}(\mathbf{v})$ to be the unique vector in $\mathbb{V}^2$ satisfying the following properties

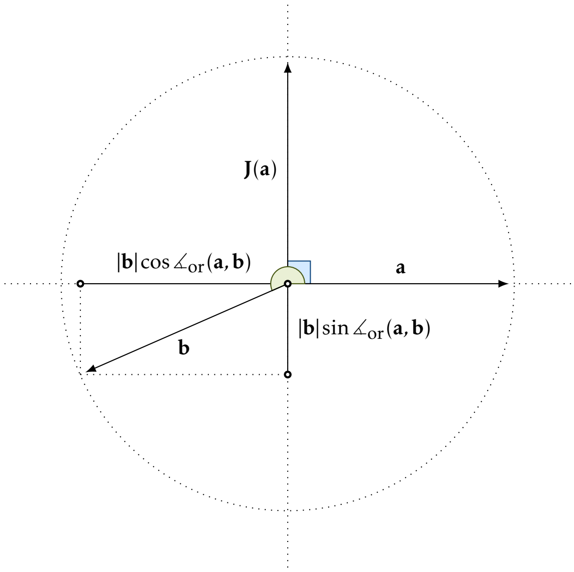
(a) $\mathbf{J}(\mathbf{v}) \perp \mathbf{v}$,

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

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

The *sine of the oriented angle* $\measuredangle_{\text{or}}(\mathbf{a}, \mathbf{b})$ is defined to be

$$\sin \measuredangle_{\text{or}}(\mathbf{a}, \mathbf{b}) = \frac{\mathrm{pr}^{\perp}_{\mathbf{J}(\mathbf{a})}(\mathbf{b})}{|\mathbf{b}|} = \frac{\mathrm{pr}^{\perp}_{\mathbf{J}(\mathbf{b})}(\mathbf{a})}{|\mathbf{a}|}. \tag{4.1}$$



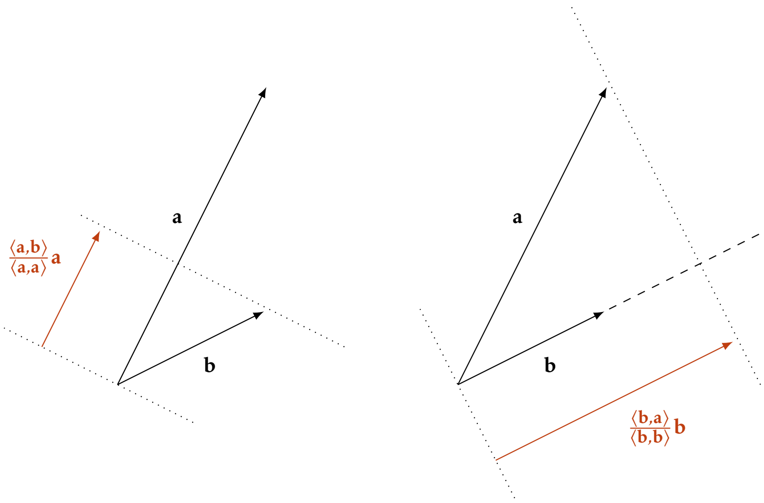
4.2 Scalar product

Definition 4.14. The *scalar product* (or, *dot product*) of two vectors $\mathbf{a},\mathbf{b}\in\mathbb{V}$ is the real number

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

It is a map $\langle\_,\_\rangle:\mathbb{V}\times\mathbb{V}\to\mathbb{R}$. Notice that from the definitions we have

$$\mathrm{Pr}^{\perp}_{\mathbf{a}}(\mathbf{b}) = \frac{\langle \mathbf{a},\mathbf{b}\rangle}{\langle \mathbf{a},\mathbf{a}\rangle}\mathbf{a} \quad\text{and}\quad \mathrm{Pr}^{\perp}_{\mathbf{b}}(\mathbf{a}) = \frac{\langle \mathbf{b},\mathbf{a}\rangle}{\langle \mathbf{b},\mathbf{b}\rangle}\mathbf{b}. \tag{4.2}$$



Proposition 4.15. The scalar product 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 $\mathbf{v}\in\mathbb{V}^2$

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

(SP4) It recognizes right angles and unit lengths, i.e. for all non-zero vectors $\mathbf{v},\mathbf{w}\in\mathbb{V}^2$

$$\mathbf{v}\perp\mathbf{w} \quad\Leftrightarrow\quad \langle \mathbf{v},\mathbf{w}\rangle = 0 \qquad\text{and}\qquad |\mathbf{v}|=1 \quad\Leftrightarrow\quad \langle \mathbf{v},\mathbf{v}\rangle = 1.$$

Proof. Since $\cos\measuredangle(\mathbf{a},\mathbf{b}) = \cos\measuredangle(\mathbf{b},\mathbf{a})$ scalar product is symmetric. Since $\cos\measuredangle(\mathbf{a},\mathbf{a}) = 0$ the scalar product is positive definite. In order to show bilinearity, notice that by symmetry it is enough to show that the scalar product is linear in the first argument. For any non-zero vectors $\mathbf{a},\mathbf{b}$, by Proposition 4.6 we have

$$\langle \mathbf{a},\mathbf{b}\rangle = \frac{1}{|\mathbf{a}|\cdot|\mathbf{b}|}\cos\measuredangle(\mathbf{a},\mathbf{b}) = \frac{1}{|\mathbf{a}|\cdot|\mathbf{b}|}\mathrm{pr}^{\perp}_{\mathbf{b}}(\mathbf{a})|\mathbf{a}| = \frac{1}{|\mathbf{b}|}\mathrm{pr}^{\perp}_{\mathbf{b}}(\mathbf{a})$$

which is linear in $\mathbf{a}$ since $\mathbf{b}$ is fixed and since $\mathrm{pr}^{\perp}_{\mathbf{b}}$ is a linear map. The last property follows directly from the definition. □

Proposition 4.16. Let $\mathcal{K} = (O,\mathcal{B})$ be an orthonormal frame. For two vectors $\mathbf{a}(a_1,\dots,a_n)$ and $\mathbf{b}(b_1,\dots,b_n)$ we have

$$\langle \mathbf{a},\mathbf{b}\rangle = a_1b_1 + \cdots + a_nb_n \tag{4.3}$$

and therefore

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

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

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

4.2.1 The Euclidean space $\mathbb{R}^n$ (first revision)

In Chapter 1 we denoted by $\mathbb{E}$ the Euclidean space, a set of points $\mathbb{P}$ governed by the Hilbert's Axioms listed in Appendix A. Fixing a unit segment, one shows that any line can be identified with the real numbers $\mathbb{R}$ (see Appendix B). Moreover, from the Axioms we extracted the concept of geometric vectors and, having fixed a unit segment, we showed that the set of vectors $\mathbb{V}$ is a vector space. Then, we concluded that the set of points $\mathbb{P}$ is a real affine space over $\mathbb{V}$. This structure encapsulates in particular the Axioms of Continuity.

At this point we made the extra assumtion that the dimension of $\mathbb{V}$ is n which we indicate with the notation $\mathbb{E}^n$. Then, the Cartesian frames introduced in Chapter 2 allow us to identify $\mathbb{P}$ with the affine space $\mathbb{R}^n$ which opend the way to linear algebra. This identification amounts to a choice of a frame and in particular contains the choice of an orientation.

In Chapter 3 we saw that lines, planes and higher dimensional analogues correspond to systems of linear equations. This allows one to revise the relations of incidence, betweenness and parallelism. In particular, by considering $\mathbb{E}^n$ as a real affine space we encapsulate the Axioms of Incidence, the Axioms of Order and the Axiom of Parallels.

Thus, it remains to shed more light on the Axioms of Congruence. In other words, we wish for a more precise description of the congruence relation for segments and angles. Congruence of segments was already extracted in the notion of length. By definition, two segments are congruent if and only if they define the same length and the set of lengths $\mathbb{L}$ can be identified with $\mathbb{R}_{\geq 0}$ (Appendix B). By Proposition 4.2, two angles are congruent if and only if they have the same cosine. Therefore, the scalar product allows us not only to efficiently calculate lengths and measure angles, but also to give an explicit description of the congruence relation – we finish this in Section 7.1.1.

By Corollary H.10, the scalar product is the unique positive definite symmetric bilinear form which recognizes right angles and unit lengths. By Corollary H.9 for any positive definite bilinear form (i.e. satisfying (SP1), (SP3) and (SP3) in Proposition 4.15) there is a basis in which it looks like the scalar product (4.3). Thus, computationally such bilinear forms are indistinguishable.

Definition 4.17. The *n-dimensional Euclidean space* $\mathbb{E}^n$ is the pair $(\mathbb{A}^n, \langle \_, \_ \rangle)$ where $\mathbb{A}^n$ is the n-dimensional real affine space and where $\langle \_, \_ \rangle$ is the unique positive definite symmetric bilinear form on $\mathrm{D}(\mathbb{A}^n)$ which recognizes right angles and unit lengths, i.e. with respect to an orthonormal basis $\mathcal{B}$ it has the expression (4.3). Then, the *distance between two points* $P, Q \in \mathbb{E}^n$ is

$$d(P,Q) = |\overrightarrow{QP}| = \sqrt{\langle \overrightarrow{QP}, \overrightarrow{QP} \rangle} = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2}. \tag{4.4}$$

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

Remark (The Euclidean space $\mathbb{R}^n$)**.** Choosing an orthonormal frame we may identify both $\mathbb{A}^n$ and $\mathbb{V}^n$ with $\mathbb{R}^n$. Since computationally there is no difference, it is more economical to say that the Euclidean space is $\mathbb{R}^n$ with the standard basis an orthonormal basis. This is the starting point of any Analysis course, and it is the advantage that Newton and Leibniz in particular saw in Descartes work.

Remark. The advantages of Definition 4.17 are conceptual. For example, by Proposition 3.6, 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 the properties in Proposition 4.15 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: 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 $\mathbb{E}^n$.

4.2.2 Gram-Schmidt algorithm

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 familiar formulas in Proposition 4.16 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 $\mathcal{K}'$. For the first option one makes use of the Gram matrix (see Appendix H). We discuss the second option in this section.

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 (4.2) 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

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

This process of obtaining an orthonormal basis from a given basis is called the *Gram-Schmidt process*. It can be used in an infinite dimensional vector space, hence the name *process*. If the vector space is finite, the process terminates and we may call it *algorithm*.

Proposition 4.18. The basis $\mathcal{B}'$ obtained from the basis $\mathcal{B}$ with the Gram-Schmidt algorithm is an orthonormal basis.

Proof. A more general statement and proof is given in [19, Theorem 17.4]. The vectors $\mathbf{e}_1', \mathbf{e}_2', \dots, \mathbf{e}_k', \dots$ are constructed by induction on k. For each k we consider the vector subspace $\mathbb{V}_k$ generated by $\mathcal{B}_k = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k)$ and show that $\mathcal{B}_k' = (\mathbf{e}_1', \mathbf{e}_2', \dots, \mathbf{e}_k')$ is an orthogonal basis for $\mathbb{V}_k$.

If $k = 1$ then $\mathbf{e}_1' = \mathbf{e}_1$ and the claim is obvious. Assume that the claim holds for k. By definition we have

$$\mathbf{e}_{k+1}' = \mathbf{e}_{k+1} - \underbrace{\sum_{i=1}^{k} \frac{\langle \mathbf{e}_i', \mathbf{e}_{k+1} \rangle}{\langle \mathbf{e}_i', \mathbf{e}_i' \rangle} \mathbf{e}_i'}_{\mathbf{v}}.$$

Since the vectors $\mathbf{e}_i'$ form a basis, $\mathbf{e}_{k+1} \neq 0 \Leftrightarrow \mathbf{v} \neq \mathbf{e}_{k+1}$ else $\mathbf{e}_{k+1}$ would be a linear combination $\mathcal{B}_k'$ and therefore of $\mathcal{B}_k$ (since both are bases of $\mathbb{V}_k$). It follows that $\mathcal{B}_{k+1}'$ is a basis of $\mathbb{V}_{k+1}$. Moreover, for each $i = 1, \dots, k$ we have

$$\begin{aligned} \langle \mathbf{e}_{k+1}', \mathbf{e}_i' \rangle &= \langle \mathbf{e}_{k+1} - \textstyle\sum_{j=1}^{k} \frac{\langle \mathbf{e}_j', \mathbf{e}_{k+1} \rangle}{\langle \mathbf{e}_j', \mathbf{e}_j' \rangle} \mathbf{e}_j', \mathbf{e}_i' \rangle \\ &= \langle \mathbf{e}_{k+1}, \mathbf{e}_i \rangle - \frac{\langle \mathbf{e}_i', \mathbf{e}_{k+1} \rangle}{\langle \mathbf{e}_i', \mathbf{e}_i' \rangle} \langle \mathbf{e}_i', \mathbf{e}_i' \rangle \\ &= \langle \mathbf{e}_{k+1}, \mathbf{e}_i \rangle - \langle \mathbf{e}_{k+1}, \mathbf{e}_i' \rangle = 0 \end{aligned}$$

since, by the inductive hypothesis, $\langle \mathbf{e}_i', \mathbf{e}_j' \rangle = 0$ for all $1 \leq i, j \leq k$. Thus, $\mathcal{B}_{k+1}'$ is an orthogonal bases and renormalizing it we obtain an orthonormal bases for $\mathbb{V}_{k+1}$. □

4.2.3 Normal vectors

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

$$\mathcal{H} : a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b. \tag{4.5}$$

Assume now that $\mathcal{K}$ is orthonormal, i.e. that $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ is orthonormal. Fix a point $Q(q_1, \dots, q_n)$ in $\mathcal{H}$. Since Q lies in $\mathcal{H}$, it satisfies the equation (4.5), i.e we have $b = a_1 q_1 + a_2 q_2 + \cdots + a_n q_n$. Having expressed the constant b in terms of Q, any other point $P(p_1, \dots, p_n)$ is a solution to (4.5) 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 $(a_1, \dots, a_n)$ then

$$P \in \mathcal{H} \quad \text{if and only if} \quad \langle \mathbf{n}, \overrightarrow{QP} \rangle = 0, \quad \text{equivalently, if and only if} \quad \mathbf{n} \perp \overrightarrow{QP}.$$

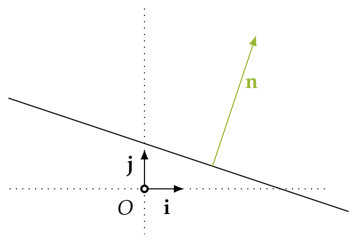
In other words, the vector $\mathbf{n}$ of coefficients in (4.5) is orthogonal to any vector parallel to $\mathcal{H}$, and we say that it is *orthogonal to $\mathcal{H}$*.

Definition 4.19. 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 $\mathcal{H}$.

Example 4.20. 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.



Example 4.21. 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.

Proposition 4.22 (Hesse normal form)**.** Let $\mathcal{K} = (O,\mathcal{B})$ be an orthonormal frame with $\mathcal{B} = (\mathbf{e}_1,\dots,\mathbf{e}_n)$. Any hyperplane $\mathcal{H}$ has, up to sign, a unique normal vector $\mathbf{n}$ of length 1. The components of $\mathbf{n}$ are $(\cos(\theta_1),\cos(\theta_2),\dots,\ \cos(\theta_n))$ where θ_i is the angle $\measuredangle(\mathbf{n},\mathbf{e}_i)$. Consequently, relative to $\mathcal{K}$ any hyperplane $\mathcal{H}$ has a unique equation of the form

$$\mathcal{H} : \cos(\theta_1)x_1 + \cos(\theta_2)x_2 + \cdots + \cos(\theta_n)x_n = c. \tag{4.6}$$

for some $c > 0$. Moreover, the distance $d(O,\mathcal{H})$ from the origin to the hyperplane equals c.

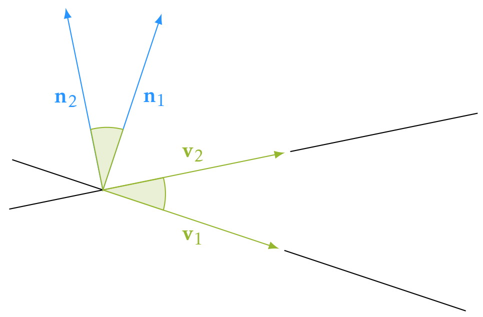
Proof. Let $\mathcal{H}$ be a hyperplane with equation (4.5). A normal vector of $\mathcal{H}$ is $\mathbf{n}(a_1,\dots,a_n)$. By the discussion in Section 4.1.1, the normalized vector $\tilde{\mathbf{n}} = \mathbf{n}/|\mathbf{n}|$ has components $(\cos(\theta_1),\dots,\cos(\theta_n))$ for some angles $\theta_1,\dots,\theta_n \in [0,\pi)$. Thus, with $c = b/|\mathbf{n}|$ the equation of $\mathcal{H}$ is has the form (4.6). If $c < 0$ replace each θ_i by $\pi - \theta_i$ to change the sign of c (this is equivalent to changing the sign of $\tilde{\mathbf{n}}$). The last claim follows from the distance formula (4.8) deduced below in Proposition 4.25. □

4.2.4 Angles between lines and hyperplanes

The scalar product offers an efficient way of calculating angles with respect to an orthonormal bases due to (4.3). In Section 4.1 we defined the sine and cosine of angles. The familiar properties of these functions are derived in Appendix I. Notice however that due to Proposition 4.2 the cosine is enough to distinguish between (unoriented) angles.

Let ℓ_1 and ℓ_2 be two 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

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

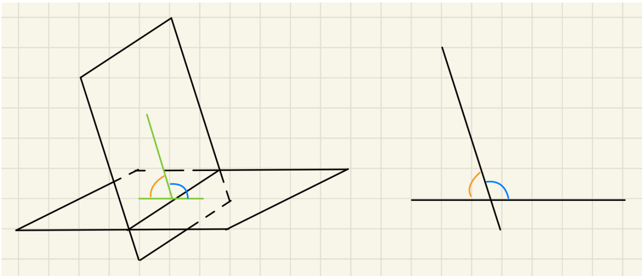
On the other hand, if we know a direction vector $\mathbf{v}_1$ for the first line and a normal vector $\mathbf{n}_2$ for the second line then the acute angle between ℓ_1 and ℓ_2 is

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

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 $\mathcal{H}$.

The angles between (hyper)planes in $\mathbb{E}^3$ are referred to as *dihedral angles*. Two planes, π_1 and π_2, define four dihedral angles which are the four regions in which the two planes divide $\mathbb{E}^3$. More precisely, 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 $\ell_2 = \pi \cap \pi_2$. The angles between ℓ_1 and ℓ_2 do not depend on the choice of π, i.e if we choose a different plane orthogonal to ℓ we obtain congruent angles. So, up to congruence we have two angles and they can be calculated using normal vectors as indicated above.



4.3 Distance

Definition 4.23. The *distance between two points* P,Q in $\mathbb{E}^n$, denoted $d(P,Q)$, is the length of the segment $[PQ]$ and we calculate it with (4.4). The *distance between two sets of points* S_1 *and* S_2 is

$$d(S_1,S_2) := \inf\Big\{d(P,Q) : P \in S_1 \text{ and } Q \in S_2\Big\}. \tag{4.7}$$

4.3.1 Distance from a point to a hyperplane

Proposition 4.24. Consider a hyperplane $\mathcal{H}$ and a point P not in $\mathcal{H}$. Drop a perpendicular line from P to $\mathcal{H}$ and let P' be the point in which the line intersects the hyperplane. Then

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

Proof. For any other point Q in $\mathcal{H}$, distinct from P', we have a right-angled triangle $PP'Q$. Since the hypothenuse is larger than the catheti we have

$$d(P,\mathcal{H}) \leq |PP'| < |PQ|$$

for any point $Q \in \mathcal{H}$. Thus $d(P,\mathcal{H}) \leq |PP'|$. □

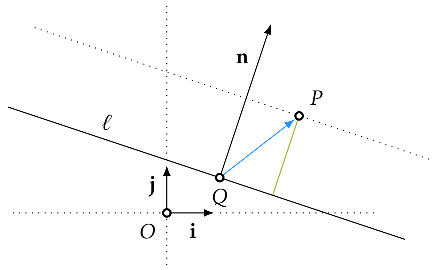
Proposition 4.25. Let $\mathcal{K}$ be a frame of $\mathbb{E}^n$. Consider the point $P(p_1,\dots,p_n)$ and a hyperplane $\mathcal{H}$: $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$. Then

$$d(P,\mathcal{H}) = \frac{|a_1p_1 + a_2p_2 + \cdots + a_np_n - b|}{\sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}}. \tag{4.8}$$

Proof. Drop a perpendicular line from P to $\mathcal{H}$, intersecting $\mathcal{H}$ in P'. By Proposition 4.24, we have $d(P,\mathcal{H}) = |PP'|$. Now let Q be a point in $\mathcal{H}$ and consider the normal vector $\mathbf{n} = \mathbf{n}(a_1,\dots,a_n)$. Then $|PP'|$ 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 $QHPN$. We have

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

If we write this explicitly in coordinates we obtain the claim. □



Proposition 4.26. Let S be an affine subspace of $\mathbb{E}^n$ parallel to a hyperplane $\mathcal{H}$. Then

$$d(S,\mathcal{H}) = d(P,\mathcal{H})$$

for any P in $\mathcal{S}$.

Proof. Notice that a point is parallel to any affine subspace (this follows from Definition 3.5). We can generalize the distance formula (4.8) 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 $\mathcal{H}$

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

□

4.3.2 Loci of points equidistant from affine subspaces

Definition 4.27. Let S be a set of points in $\mathbb{E}^n$. For $c \in \mathbb{R}$, the set of points at distance c from S is

$$\mathcal{L}(S,c) = \left\{P \in \mathbb{E}^n : d(P,S) = c\right\}.$$

Proposition 4.28. Let S be an affine subspace of $\mathbb{E}^n$ and let $c > 0$ be a constant. Table 4.1 classifies the possible shapes of $\mathcal{L}(S,c)$.

n	a point	a line	a plane
1	two points	-	-
2	circle	two lines	-
3	sphere	circular cylinder	two planes

Table 4.1: Loci of points at constant distance from an affine subspace.

Proof. We notice first that if $c = 0$ then $\mathcal{L}(S,c)$ is the set S itself by definition, (4.7). Moreover, if $c < 0$ then $\mathcal{L}(S,c)$ is empty since distances are positive. The non-trivial cases appear when $c > 0$. If S consists of a single point, in dimension $n = 1$, then $\mathcal{L}(S,c)$ consists of two points, namely the endpoints of a segment with midpoint S. We can discuss the diagonal entries in Table 4.1 together by considering hyperplanes since, in dimension 1, these are points, in dimension 2, these are lines and, in dimension 3, these are planes. Let the hyperplane S be given by the equation $a_1x_1+a_2x_2+\cdots+a_nx_n = b$. By (4.8), a point $Q(q_1,\ldots,q_n)$ belongs to $\mathcal{L}(S,c)$ if and only if for all $P \in S$ we have

$$d(P,Q) = c \quad \Leftrightarrow \quad |a_1q_1 + a_2q_2 + \cdots + a_nq_n| = c'$$

where $c' = c \cdot \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$. Thus, $\mathcal{L}(S,c)$ is the union of two hyperplanes with equations

$$a_1q_1 + a_2q_2 + \cdots + a_nq_n = c' \quad \text{and} \quad a_1q_1 + a_2q_2 + \cdots + a_nq_n = -c'.$$

Consider the case where S is a line in $\mathbb{E}^3$. We notice that in dimension 3 we don't yet have a formula for the distance between a point and a line. Such a formula will be deduced in Section 5.2.2 and can be used to deduce equations for $\mathcal{L}(S,c)$ if S is an arbitrary line. However, we have other options as well, for instance we can 'scan' the three dimensional space with planes passing through S and notice that in each such plane we obtain two lines at equal distance from S. Yet another option, since we are only interested in the possible shapes of $\mathcal{L}(S,c)$, is to make a good choise of a frame. Choose the frame $\mathcal{K}$ such that S is the z-axis. An arbitrary point in S is $P(0,0,t)$ with $t \in \mathbb{R}$. Then, for a point $Q(x_Q,y_Q,z_Q)$ we have

$$d(P,Q) = c \quad \Leftrightarrow \quad x_Q^2 + y_Q^2 = c^2$$

Thus, this locus of points has equation $\mathcal{L}(S,c) : x^2 + y^2 = c^2$ which describes a cylinder with axis S.

Lastly, in any dimension the set of points at distance c from a point $Q(x_Q,y_Q,z_Q)$ is the (hyper)sphere of radius c centered in Q, since

$$d(P,Q) = c \quad \Leftrightarrow \quad (p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2 = c^2.$$

□

Definition 4.29. If S' is another set, the *locus of points equidistant from S and S'* is

$$\mathcal{L}(S,S') = \left\{P \in \mathbb{E}^n : d(P,S) = d(P,S')\right\}.$$

Proposition 4.30. Let S and S' be two affine subspace of $\mathbb{E}^n$. Table 4.2 classifies the possible shapes of $\mathcal{L}(S,S')$.

n	2 points	point + line	point + plane	2 lines	line + plane	2 planes
1	point	-	-	-	-	-
2	line	parabola	-	line/lines	-	-
3	plane	parabolic cylinder	paraboloid	planes or hyperboloid	cone or parabolic cylinder	plane or two planes

Table 4.2: Loci of points equidistant from two distinct affine subspaces.

Proof. Consider the first column in Table 4.2. Let $A(a_1,\dots,a_n)$ and $B(b_1,\dots,b_n)$ be t wo fixed points. A point $P(p_1,\dots,p_n)$ is equidistant from A and B if and only if $d(A,P) = d(P,B)$. In coordinates this gives the equation

$$(a_1-p_1)^2+\cdots+(a_n-p_n)^2=(b_1-p_1)^2+\cdots+(b_n-p_n)^2.$$

which is equivalent to P satisfying the equation

$$(a_1-b_1)x_1+\cdots+(a_n-b_n)x_n+\frac{1}{2}(a_1^2-b_1^2+\cdots+a_n^2-b_n^2)=0$$

which is the equation of a hyperplane with normal vector $\overrightarrow{AB}$. Moreover, it is easy to check that the hyperplane contains the midpoint of the segment $[AB]$. It is called the *perpendicular bisecting hyperplane* of the segment $[AB]$. In dimension 2 it is the perpendicular bisector of the segment $[AB]$.

For the first diagonal in Table 4.2, consider a hyperplane $\mathcal{H}$ and a point Q outside $\mathcal{H}$. We choose the frame $\mathcal{K}$ such that Q is on the positive part of the last coordinate axes, such that the last coordinate axes is orthogonal to $\mathcal{H}$ and such that the origin O is at the same distance from $\mathcal{H}$ and Q. Moreover, eventually changing the unit segment, we may assume that Q has coordinates $(0,\dots,0,1)$ and $\mathcal{H}$ has equation $x_n+1=0$. Then, a point P is equidistant from $\mathcal{H}$ and Q if and only if

$$d(P,\mathcal{H})=d(P,Q) \quad\Leftrightarrow\quad |p_n+1|=\sqrt{p_1^2+\cdots+p_{n-1}^2+(p_n-1)^2}=0$$

which is equivalent to P satisfying the equation

$$x_1^2+\cdots+x_{n-1}^2=4x_n$$

which is the equation of a hyperparaboloid. In dimension 1 this is a point (the origin, 0). In dimension 2 it is a parabola and in dimension 3 a paraboloid (see Chapter **??**).

Next, consider the case of two hyperplanes

$$\mathcal{H}: a_1x_1+\cdots+a_nx_n+a_{n+1}=0 \quad\text{and}\quad \mathcal{H}': b_1x_1+\cdots+b_nx_n+b_{n+1}=0.$$

We may assume that the normal vectors which can be read of from the equations are unit vectors. A point $P(p_1,\dots,p_n)$ is at the same distance from $\mathcal{H}$ and $\mathcal{H}'$ if

$$d(P,\mathcal{H})=d(P,\mathcal{H}') \quad\Leftrightarrow\quad |a_1x_1+\cdots+a_nx_n+a_{n+1}|=|b_1x_1+\cdots+b_nx_n+b_{n+1}|.$$

This translates to two equations

$$(a_1-b_1)x_1+\cdots+(a_n-b_n)x_n+(a_{n+1}-b_{n+1})=0 \quad\text{and}\quad (a_1+b_1)x_1+\cdots+(a_n+b_n)x_n+(a_{n+1}+b_{n+1})=0$$

and we notice that if the hyperplanes are parallel then only one of the equation has solutions (the hyperplane lying at half-distance between $\mathcal{H}$ and $\mathcal{H}'$). In dimension 1 this is the midpoint of the segment $[\mathcal{H}\mathcal{H}']$. Assume that $\mathcal{H}$ and $\mathcal{H}'$ are not parallel. In dimension 2 these are the angle bisectors of the angles described by the two lines $\mathcal{H}$ and $\mathcal{H}'$. In dimension 3 these are bisecting planes of the dihedral angles between $\mathcal{H}$ and $\mathcal{H}'$ and in general these are angle bisecting hyperplanes.

Next consider the case of a point $S = \{Q\}$ and a line S' in $\mathbb{E}^3$. We choose a frame $\mathcal{K}$ such that the point Q has coordinates $(1,0,0)$ and such that the line S' contains the point $(-1,0,0)$ and has direction vector $\mathbf{j}(0,1,0)$. We anticipate and use the distance formula from a point to a line in dimension 3 (Section 5.2.2). Notice however that because of the choice of the frame, the distance formula can be easily deduced in this case. For a point P, we have

$$d(P,S) = d(P,S') = |\overrightarrow{QP} \times \mathbf{j}| \quad \Leftrightarrow \quad (x_P - 1)^2 + y_P^2 + z_P^2 = z_P^2 + (x_P + 1)^2 \quad \Leftrightarrow \quad y_P^2 = 4x_P$$

which is the equation of a parabolic cylinder.

Next consider the case of two lines in $\mathbb{E}^3$. Assume first that the line intersect in one point. We assume that the two lines S and S' are the x-axis and the y-axis respectively. Then, for a point P, we have

$$d(P,S) = d(P,S') \quad \Leftrightarrow \quad |\overrightarrow{OP} \times \mathbf{i}| = |\overrightarrow{OP} \times \mathbf{j}| \quad \Leftrightarrow \quad y_P^2 + z_P^2 = z_P^2 + x_P^2 \quad \Leftrightarrow \quad (x_P - y_P)(x_P + y_P) = 0.$$

These are two planes orthogonal to the plane containing the two lines and which bisect the angles formed by the two lines.

Assume now that the lines are skew. We choose a frame such that S is the x-axis and S' contains the point $Q(0,1,0)$ and has $\mathbf{v}(\lambda,0,1)$ as direction vector. Then, for a point P, we have

$$d(P,S) = d(P,S') \quad \Leftrightarrow \quad |\overrightarrow{OP} \times \mathbf{i}| = |\overrightarrow{QP} \times \mathbf{v}| \quad \Leftrightarrow \quad y_P^2 + z_P^2 = (x_P - \lambda z_P)^2 + (\lambda^2 + 1)(y_P - 1)^2$$

which corresponds to a one-sheeted hyperboloid (see Chapter **??**). In the last subcase, where the two lines S and S' are parallel on checks that $\mathcal{L}(S,S')$ is a plane.

Finally, the last case that we need to consider is that of a line S and a plane S' in $\mathbb{E}^3$. With a similar argument we see that in this case we obtain an elliptic cone if the line punctures the plane and, if the line and the plane are parallel we obtain a parabolic cylinder. □

4.3.3 Convergence

Once we have a clear notion of distance, the notion of proximity to a point also becomes clear. Being close to a point P means lying in a ball of radius ε centered at the point P and you may adjust ε at will. The set of open balls centered at all points defines a topology on $\mathbb{E}^n$, called the *standard topology*. Recall Chapter 7 of you Analysis course [16]. All the results from your analysis course hold true for $\mathbb{E}^n$ by fixing an orthonormal frame which identifies $\mathbb{E}^n$ with $\mathbb{R}^n$.

CHAPTER 5

Area and volume

Contents

5.1 Area **67**
- 5.1.1 Area of polygons 67
- 5.1.2 Oriented area 70

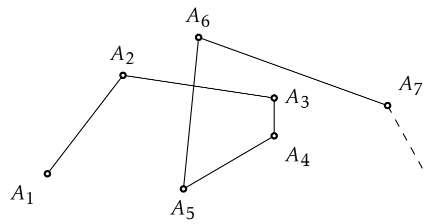
5.2 Cross product **70**
- 5.2.1 Algebraic identities 73
- 5.2.2 Distance between a point and a line 75
- 5.2.3 Common perpendicular line of two skew lines 76
- 5.2.4 Distance between two skew lines 77

5.3 Volume **78**
- 5.3.1 Volume of polyhedra 78
- 5.3.2 Oriented volume 82
- 5.3.3 Hypervolume 83

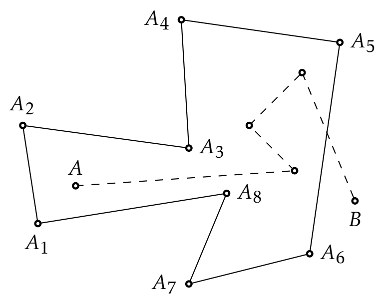
5.1 Area

5.1.1 Area of polygons

Definition 5.1. A *polygonal segment* $A_1A_2\dots A_n$ is a union of segments $[A_0A_1],\dots,[A_{n-1}A_n]$, such that consecutive segments $[A_iA_{i+1}]$ and $[A_{i+1}A_{i+2}]$ intersect only in their shared *vertex* A_{i+1}. If all the vertices lie in the same plane, the polygonal segment is said to be *planar*. The polygonal segment $A_1A_2\dots A_n$ is said to *intersect itself* in a point P if there exist indices i and j with $i+1<j$, such that $[A_iA_{i+1}]$ intersects $[A_jA_{j+1}]$ at P.

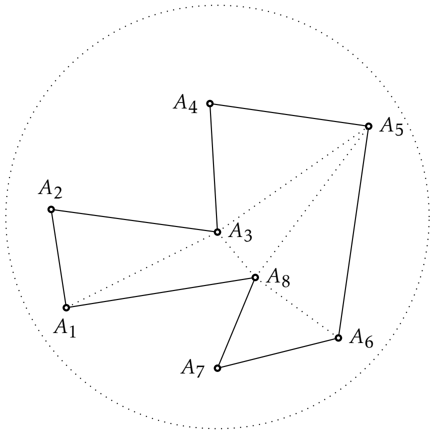


A *polygon* is an n-tuple of points $(A_1,A_2,\dots,A_n)$, denoted $A_1A_2\dots A_n$, such that the polygonal segment $A_1A_2\dots A_nA_1$ intersects itself only at the vertex A_1. The segments $[A_iA_{i+1}]$ are called the *sides of the polygon*. Two polygons are said to be *congruent* if their corresponding sides and the angles between them are pairwise congruent. Two polygons $P_1 = A_1A_2\dots A_n$ and $P_2 = B_1B_2\dots B_m$ are said to be *similar* if their corresponding angles are congruent. In this case, the ratio of the lengths of corresponding sides is constant, and we denote this *ratio* by $P_1 : P_2$.



We will only consider planar polygons. A planar polygon separates the points of the plane not lying on its sides into two regions (see [14, Theorem 9]): the *interior* and the *exterior*. These regions have the following property: if A is a point of the interior (an *inner point*) and B is a point of the exterior (an *exterior point*), then every polygonal segment lying in the plane of the polygon and connecting A with B must intersect the sides of the polygon at least once.

The interior S of a (planar) polygon is *bounded* in the sense that, for any point P in the plane, there is an $r>0$ such that all points in S are at distance at most r from P. Indeed, take r to be the maximum of the distances from P to all vertices. Our goal is to measure such bounded regions of the plane - namely, interiors of polygons. We do this by means of triangles, which are the simplest possible polygons. Any polygon can be subdivided into triangles (see for example [3, Theorem 3.1]), for instance with the so-called ear clipping method. Such a subdivision partitions the interior of a polygon into interiors of triangles, if we are willing to disregard segments. Given all this, we may measure the interior of polygons by comparing them with the the interior of a unit square.



Definition 5.2. For brevity, we use the term *area of a polygon* to mean the *area of the interior of a polygon*. To define area, we consider two polygons to be *disjoint* if their interiors are disjoint. Denote by $\mathcal{P}$ the set of all finite unions of polygons. We define an *area* function $\text{Area} : \mathcal{P} \to \mathbb{R}_{\geq 0}$ through the following properties:

(A1) If S is a square of side length 1 then $\text{Area}(S) = 1$.

(A2) If S_1 and S_2 are similar polygons with ratio x then $\text{Area}(S_1) = x^2 \text{Area}(S_2)$.

(A3) If S_1 and S_2 are disjoint polygons then $\text{Area}(S_1 \cup S_2) = \text{Area}(S_1) + \text{Area}(S_2)$.

Remark. The above definition captures basic principles of the intuitive notion of area. We note that in property (A2), the term 'similar polygons' can be replaced with 'congruent polygons', resulting in a weaker assumption. Furthermore, in property (A3) the finite union can be extended to a countable union, in which case the sum is replaced by a countable sum (see Section 5.3.3). Our definition of area allows us to deduce the following well-known facts.

Proposition 5.3. We have

1. The area of a rectangle of sides a and b is ab.
2. The area of a triangle ABC with side length a and corresponding height h is $ah/2$.
3. The area of a parallelogram $ABCD$ is twice the area of the triangle ABC.

Proof. Denote by S_a a square of side length a and by $R_{a,b}$ a rectangle of sides length a and b. First, notice that since the ratio $S_a : S_1$ equals a, we have $\text{Area}(S_a) = a^2 \text{Area}(S_1) = a^2$, by (A2) and (A1).

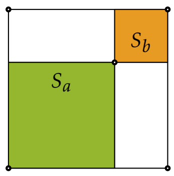
1. Place S_a and S_b in opposite corners of S_{a+b}. The complement of the two small squares in the big square are two disjoint rectangles which are congruent to $R_{a,b}$. Thus, by (A3) we have

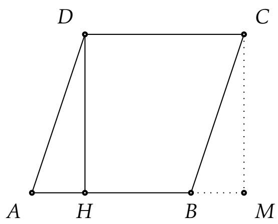
$$\mathrm{Area}(S_{a+b}) = \mathrm{Area}(S_a \cup R_{a,b} \cup R_{a,b} \cup S_b) = \mathrm{Area}(S_a) + \mathrm{Area}(S_b) + 2\,\mathrm{Area}(R_{a,b})$$

hence

$$(a+b)^2 = a^2 + b^2 + 2\,\mathrm{Area}(R_{a,b})$$

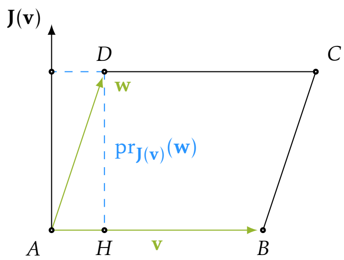
and therefore $\mathrm{Area}(R_{a,b}) = ab$.





For claim 3, it suffices to notice that a diagonal divides a parallelogram in two congruent triangles. For claim 2, we use claim 3 and check the known area formula for a parallelogram $ABCD$, i.e. $\mathrm{Area}(ABCD) = ah$ where h is a height and a the length of the corresponding side. Let H be a point on AB such that $h = |DH|$. The triangle ADH can be moved on the opposite side of the parallelogram to form a rectangle with side lengths a and h. Thus, by claim 1, $\mathrm{Area}(ABCD) = \mathrm{Area}(R_{a,h}) = ah$. $\square$

Orthonormal frames offer an efficient method for calculating the area of a parallelogram, and therefore, the areas of triangles and polygons.



Proposition 5.4. Suppose the vertices $A(x_A, y_A)$, $B(x_B, y_B)$ and $D(x_D, y_D)$ of the parallelogram $ABCD$ are given with respect to an orthonormal frame. Then,

$$\mathrm{Area}(ABCD) = \left|\begin{vmatrix} x_A & y_A & 1 \\ x_B & y_B & 1 \\ x_D & y_D & 1 \end{vmatrix}\right| = \left|\begin{vmatrix} v_x & v_y \\ w_x & w_y \end{vmatrix}\right|$$

where $\mathbf{v}(v_x, v_y) = \overrightarrow{AB}$ and $\mathbf{w}(w_x, w_y) = \overrightarrow{AD}$.

Proof. Drop a perpendicular line from D on AB and denote by H its intersection with AB. By Proposition 5.3, we have

$$\mathrm{Area}(ABCD) = |\overrightarrow{AB}| \cdot |\overrightarrow{HA}| = |\mathbf{v}| \cdot |\mathrm{pr}_{\mathbf{J}(\mathbf{v})}(\mathbf{w})| = |\mathbf{v}| \cdot |\frac{\langle \mathbf{J}(\mathbf{v}), \mathbf{w} \rangle}{|\mathbf{J}(\mathbf{v})|}| = |\langle \mathbf{J}(\mathbf{v}), \mathbf{w} \rangle| = |\begin{vmatrix} v_x & v_y \\ w_x & w_y \end{vmatrix}|.$$

Moreover, using properties of the determinants we obtain

$$\begin{vmatrix} v_x & v_y \\ w_x & w_y \end{vmatrix} = \begin{vmatrix} x_B - x_A & y_B - y_A \\ x_D - x_A & y_D - y_A \end{vmatrix} = \begin{vmatrix} x_A & y_A & 1 \\ x_B - x_A & y_B - y_A & 0 \\ x_D - x_A & y_D - y_A & 0 \end{vmatrix} = \begin{vmatrix} x_A & y_A & 1 \\ x_B & y_B & 1 \\ x_D & y_D & 1 \end{vmatrix}.$$

□

5.1.2 Oriented area

Fix an orientation in $\mathbb{E}^2$, i.e. fix a right-oriented basis $\mathcal{B}$. Proposition 5.4 shows that, up to sign, the area of a parallelogram $ABDC$ is given by the determinant of the base change matrix $M_{\mathcal{B},\mathcal{B}'}$, where $\mathcal{B}' = (\overrightarrow{AB}, \overrightarrow{AC})$. The value $\det(M_{\mathcal{B},\mathcal{B}'})$ does not depend on the orthonormal basis $\mathcal{B}$ in which the vectors are expressed. The proof of this claim is the same in any dimension (see Proposition 5.21).

Definition 5.5. With the above notation, the *box product of* $\mathbf{v}$ *and* $\mathbf{w}$ is $\det(M_{\mathcal{B},\mathcal{B}'})$ and we denote it by $[\mathbf{v}, \mathbf{w}]$. The *oriented area of the parallelogram* $ABDC$ spanned by $\mathbf{v} = \overrightarrow{AB}$ and $\mathbf{w} = \overrightarrow{AC}$ is

$$\mathrm{Area}_{\mathrm{or}}(ABDC) = [\mathbf{v}, \mathbf{w}].$$

Similarly, the oriented area of the triangle ABC is $\mathrm{Area}_{\mathrm{or}}(ABC) = \mathrm{Area}_{\mathrm{or}}(ABDC)/2$.

Proposition 5.6. Let $\mathbf{v} = \overrightarrow{AB}$ and $\mathbf{w} = \overrightarrow{AC}$ be two vectors with ABC a triangle.

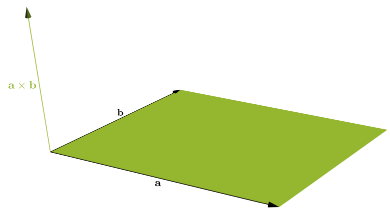
$$\sin \measuredangle_{\mathrm{or}}(\mathbf{v}, \mathbf{w}) = \frac{[\mathbf{v}, \mathbf{w}]}{|\mathbf{v}| \cdot |\mathbf{w}|} = \frac{\mathrm{Area}_{\mathrm{or}}(ABC)}{|AB| \cdot |AC|}.$$

where $\mathcal{B}' = (\mathbf{v}, \mathbf{w})$.

Proof. This is a direct consequence of the definition of the sine function for oriented angles and of the definition of oriented area of a triangle. □

5.2 Cross product

The cross product is the 3-dimensional analogue of the operator J introduced in Section 4.1.2. Throughout this section, we consider $\mathbb{E}^3$. For two non-zero vectors $\mathbf{a}$ and $\mathbf{b}$, the orthogonal complements $\mathbb{V}^{\perp \mathbf{a}}$ and $\mathbb{V}^{\perp \mathbf{b}}$ are 2-dimensional vector subspaces of $\mathbb{V}^3$. If $\mathbf{a}$ and $\mathbf{b}$ are linearly independent, then $\mathbb{V}^{\perp \mathbf{a}} \cap \mathbb{V}^{\perp \mathbf{b}}$ 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 prescribing the length of $\mathbf{c}$, we have exactly two options: $\pm\mathbf{c}$. For these options, $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is either left- or right-oriented. By fixing such an orientation, the choice of $\mathbf{c}$ becomes unique. This brings us to the following definition.



Definition 5.7. Let $\mathbf{a}, \mathbf{b} \in \mathbb{V}^3$ be two vectors. The *cross product* (or *vector product*) of $\mathbf{a}$ and $\mathbf{b}$, denoted $\mathbf{a} \times \mathbf{b}$, is the vector defined by the following properties:

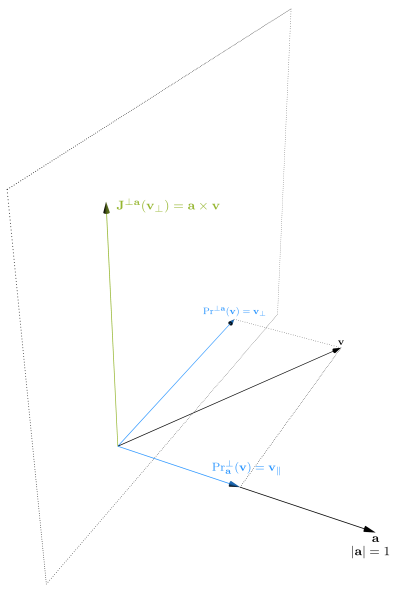
1. if $\mathbf{a}$ and $\mathbf{b}$ are parallel then $\mathbf{a} \times \mathbf{b} = 0$.
2. if $\mathbf{a}$ and $\mathbf{b}$ are not parallel, then
   (a) $|\mathbf{a} \times \mathbf{b}|$ equals the area of a parallelogram spanned by $\mathbf{a}$ and $\mathbf{b}$,
   (b) $\mathbf{a} \times \mathbf{b} \perp \mathbf{a}$ 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 $\mathbb{V}^3$.

In particular, the vectors $\mathbf{a}$ and $\mathbf{b}$ are parallel if and only if $\mathbf{a} \times \mathbf{b} = 0$.

The above geometric definition can be translated into a more accurate algebraic description as follows. Fix a non-zero vector $\mathbf{a}$. Consider the orthogonal complement of $\mathbf{a}$, i.e. $\mathbb{V}^{\perp \mathbf{a}} = \{\mathbf{v} \in \mathbb{V}^3 : \langle \mathbf{v}, \mathbf{a} \rangle = 0\}$. It is a 2-dimensional vector subspace of $\mathbb{V}^3$. For any vector $\mathbf{v} \in \mathbb{V}^3$, we have a unique decomposition

$$\mathbf{v} = \mathbf{v}_{||} + \mathbf{v}_{\perp} \quad \text{with } \mathbf{v}_{||} \text{ parallel to } \mathbf{a} \text{ and } \mathbf{v}_{\perp} \text{ orthogonal to } \mathbf{a}.$$

This gives a projection map $\mathrm{Pr}^{\perp \mathbf{a}} : \mathbb{V}^3 \to \mathbb{V}^{\perp \mathbf{a}}$ defined by $\mathrm{Pr}^{\perp \mathbf{a}}(\mathbf{v}) = \mathbf{v}_{\perp}$. Let $\mathbf{J}^{\perp \mathbf{a}}(\mathbf{v}_{\perp})$ denote the unique vector in $\mathbb{V}^{\perp \mathbf{a}}$ of length $|\mathbf{v}_{\perp}|$ and such that $(\mathbf{v}_{||}, \mathbf{v}_{\perp}, \mathbf{J}^{\perp \mathbf{a}}(\mathbf{v}_{\perp}))$ is right-oriented.



Proposition 5.8. With the above notation, for any two non-zero vectors $\mathbf{a}, \mathbf{b} \in \mathbb{V}^3$ we have

$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}| \cdot \mathbf{J}^{\perp \mathbf{a}}(\mathrm{Pr}^{\perp \mathbf{a}}(\mathbf{b})).$$

Proof. By construction, we see that $\mathbf{J}^{\perp \mathbf{a}}(\mathrm{Pr}^{\perp \mathbf{a}}(\mathbf{b}))$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. Furthermore, by the definition of the operator $\mathbf{J}^{\perp \mathbf{a}}$, the vector $\mathbf{J}^{\perp \mathbf{a}}(\mathrm{Pr}^{\perp \mathbf{a}}(\mathbf{b}))$ is a positive scalar multiple of $\mathbf{a} \times \mathbf{b}$. Therefore, all we need to show is that the vector on the right-hand side has length $|\mathbf{a} \times \mathbf{b}|$. We have

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| \cdot |\mathbf{b}| \cdot \sin \measuredangle(\mathbf{a}, \mathbf{b}) = |\mathbf{a}| \cdot |\mathrm{Pr}^{\perp \mathbf{a}}(\mathbf{b})| = |\mathbf{a}| \cdot |\mathbf{J}^{\perp \mathbf{a}}(\mathrm{Pr}^{\perp \mathbf{a}}(\mathbf{b}))|$$

where the last equality follows from the fact that $\mathbf{J}^{\perp \mathbf{a}}$ is a rotation with a right angle - in particular it doesn't change the length of vectors. □

Proposition 5.9. The cross product $\square \times \square : \mathbb{V}^3 \times \mathbb{V}^3 \to \mathbb{V}^3$ 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

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

Proof. Skew-symmetry follows from the orientation requirement in the definition of the cross product. Indeed, if $(\mathbf{a}, \mathbf{b}, \mathbf{v})$ is right-oriented then $(\mathbf{b}, \mathbf{a}, \mathbf{v})$ is left-oriented, hence $(\mathbf{b}, \mathbf{a}, -\mathbf{v})$ is right-oriented. To check bilinearity, we need to verify that the cross product is linear in each argument. If the first argument is fixed, the map $\mathbf{a} \times \square : \mathbb{V}^3 \to \mathbb{V}^3$ is a composition of linear maps (by Proposition 5.8), so it is linear. For the second argument we may use the skew-symmetry of the cross product. □

Since the cross product is bilinear, its values are determined by the values on a basis. If $\mathcal{B} = (\mathbf{i}, \mathbf{j}, \mathbf{k})$ is a right-oriented orthonormal basis, one can check with the definition of the cross product that the values on the basis vectors are:

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

This table allows us to calculate the cross product for arbitrary vectors $\mathbf{a}(a_1, a_2, a_3)$ and $\mathbf{b}(b_1, b_2, b_3)$, with components relative to $\mathcal{B}$:

$$(a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \times (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) = (a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

We can therefore derive, for example, a formula for the area of a parallelogram P spanned by $\mathbf{a}$ and $\mathbf{b}$:

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

5.2.1 Algebraic identities

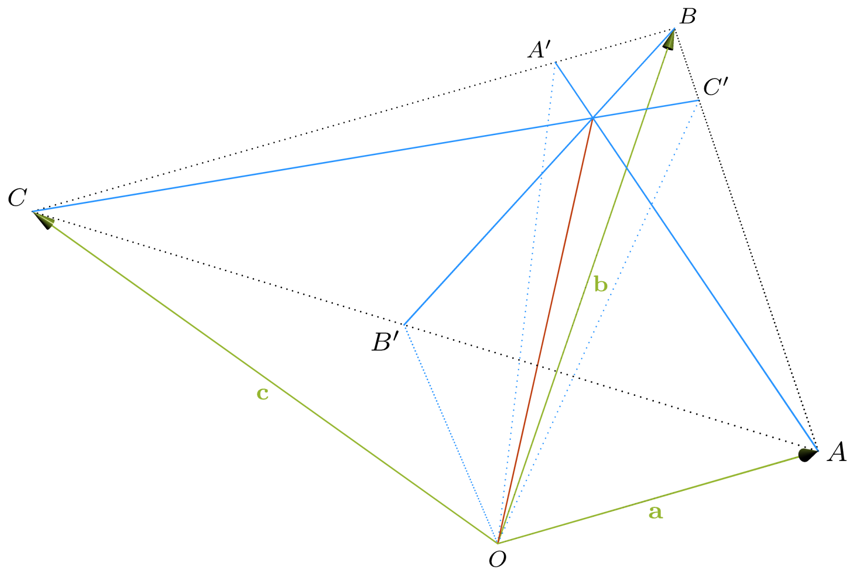
When calculating consecutive cross products, we notice that the cross product is not associative. For example, if $(\mathbf{i},\mathbf{j},\mathbf{k})$ is a right oriented orthonormal basis, then:

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

However, there is a rule which explains how iterated cross products behave when evaluated in different ways. This rule is the *Jacobi identity*:

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}+(\mathbf{b}\times\mathbf{c})\times\mathbf{a}+(\mathbf{c}\times\mathbf{a})\times\mathbf{b}=0.\quad\forall\mathbf{a},\mathbf{b},\mathbf{c}\in\mathbb{V}^3. \tag{5.1}$$

One way to prove this identity is to write out the above expression in coordinates and check that the left-hand side simplifies to the zero vector. The geometric interpretation of the identity is as follows: given a tetrahedron $OABC$, the three planes passing through the edges adjacent to O and orthogonal to the respective opposite faces intersect in one line.



A short way to prove (5.1) is by using *the double cross formula*. This formula is of independent interest, as it provides an efficient way of calculating iterated cross products. Specifically, it replaces the calculation of a determinant with the calculation of two scalar products. Other beautiful identities, which can be derived with linear algebra only, can be found in the exercises and in [11, Chapter 4].

Theorem 5.10 (Double Cross Formula)**.** For any vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ we have

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}=\langle\mathbf{a},\mathbf{c}\rangle\cdot\mathbf{b}-\langle\mathbf{b},\mathbf{c}\rangle\cdot\mathbf{a}. \tag{5.2}$$

Proof. Notice that (5.2) can be checked directly in coordinates. A different proof is provided in [11, p.74] or in [5, p.70]. Our proof has three steps. First, note that we may assume all three vectors to be non-zero; otherwise, it is straightforward to check that both sides of the equality are zero.

(Step 1) It suffices to prove (5.2) in the special cases where all three vectors are unit vectors. Indeed, under this assumption, the linearity of the cross product (Proposition 5.9) and the linearity of the scalar product we obtain:

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c} = |\mathbf{a}|\cdot|\mathbf{b}|\cdot|\mathbf{c}|\cdot\left[\left(\frac{\mathbf{a}}{|\mathbf{a}|}\times\frac{\mathbf{b}}{|\mathbf{b}|}\right)\times\frac{\mathbf{c}}{|\mathbf{c}|}\right] = |\mathbf{a}|\cdot|\mathbf{b}|\cdot|\mathbf{c}|\cdot\left[\langle\frac{\mathbf{a}}{|\mathbf{a}|},\frac{\mathbf{c}}{|\mathbf{c}|}\rangle\frac{\mathbf{b}}{|\mathbf{b}|} - \langle\frac{\mathbf{b}}{|\mathbf{b}|},\frac{\mathbf{c}}{|\mathbf{c}|}\rangle\frac{\mathbf{a}}{|\mathbf{a}|}\right] = \langle\mathbf{a},\mathbf{c}\rangle\cdot\mathbf{b} - \langle\mathbf{b},\mathbf{c}\rangle\cdot\mathbf{a}.$$

(Step 2) It suffices to prove (5.2) in the special cases where $\mathbf{c}=\mathbf{a}$ and $\mathbf{c}=\mathbf{b}$. Since $(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$ is orthogonal to $\mathbf{a}\times\mathbf{b}$, it lies in the linear span of $\mathbf{a}$ and $\mathbf{b}$, i.e. $\mathbf{c}=\alpha\mathbf{a}+\beta\mathbf{b}$ for some real numbers α and β. Then, using the linearity of the cross product (Proposition 5.9) and the properties of the scalar product, we have:

$$\begin{aligned}
(\mathbf{a}\times\mathbf{b})\times\mathbf{c} &= (\mathbf{a}\times\mathbf{b})\times(\alpha\mathbf{a}+\beta\mathbf{b})\\
&= \alpha(\mathbf{a}\times\mathbf{b})\times\mathbf{a}+\beta(\mathbf{a}\times\mathbf{b})\times\mathbf{b}\\
&= \alpha\big(\langle\mathbf{a},\mathbf{a}\rangle\mathbf{b}-\langle\mathbf{b},\mathbf{a}\rangle\mathbf{a}\big)+\beta\big(\langle\mathbf{a},\mathbf{b}\rangle\mathbf{b}-\langle\mathbf{b},\mathbf{b}\rangle\mathbf{a}\big)\\
&= \big(-\alpha\langle\mathbf{b},\mathbf{a}\rangle-\beta\langle\mathbf{b},\mathbf{b}\rangle\big)\mathbf{a}+\big(\alpha\langle\mathbf{a},\mathbf{a}\rangle+\beta\langle\mathbf{a},\mathbf{b}\rangle\big)\mathbf{b}\\
&= \langle-\alpha\mathbf{a}-\beta\mathbf{b},\mathbf{b}\rangle\mathbf{a}+\langle\alpha\mathbf{a}+\beta\mathbf{b},\mathbf{a}\rangle\mathbf{b}\\
&= \langle\mathbf{a},\mathbf{c}\rangle\mathbf{b}-\langle\mathbf{b},\mathbf{c}\rangle\mathbf{a}.
\end{aligned}$$

(Step 3) It remains to show that (5.2) holds in the special cases where $\mathbf{c}=\mathbf{a}$ and $\mathbf{c}=\mathbf{b}$, assuming $\mathbf{a}$ and $\mathbf{b}$ are unit vectors. Using the skew-symmetry of the cross product, it is easy to see that the two cases are equivalent. By Step 1, we may also assume that all three vectors are unit vectors. In particular $\langle\mathbf{a},\mathbf{a}\rangle=1$ and $\langle\mathbf{b},\mathbf{a}\rangle=\cos\theta$, where $\theta=\measuredangle(\mathbf{a},\mathbf{b})$. Thus, it suffices to prove the identity:

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{a}=\mathbf{b}-\cos\theta\cdot\mathbf{a}.$$

Let $\mathbf{v}$ denote the left-hand side and $\mathbf{w}$ the right-hand side. Notice that $\mathcal{B}_{\mathbf{v}}=(\mathbf{a}\times\mathbf{b},\mathbf{a},\mathbf{v})$ is a right-oriented orthogonal frame. Since there is a unique vector of length $|\mathbf{v}|$ with this property, it suffices to show that $\mathcal{B}_{\mathbf{w}}=(\mathbf{a}\times\mathbf{b},\mathbf{a},\mathbf{w})$ is a right oriented orthogonal frame and that $|\mathbf{v}|=|\mathbf{w}|$. The first property can be verified by calculating the determinant of the base-change matrix to the basis $\mathcal{B}=(\mathbf{a},\mathbf{b},\mathbf{a}\times\mathbf{b})$, which is also right-oriented:

$$\det(M_{\mathcal{B},\mathcal{B}_{\mathbf{w}}})=\begin{vmatrix}0&1&-\cos\theta\\0&0&1\\1&0&0\end{vmatrix}=1+\cos\measuredangle(\mathbf{b},\mathbf{a}).$$

Since $\mathbf{a}$ and $\mathbf{b}$ are linearly independent, the angle θ cannot equal two right angles. Consequently, $1+\cos\measuredangle(\mathbf{b},\mathbf{a})>0$ and therefore $\mathcal{B}_{\mathbf{w}}$ is right-oriented. Furthermore, because $\langle\mathbf{a},\mathbf{w}\rangle=\langle\mathbf{b},\mathbf{a}\rangle-\langle\mathbf{b},\mathbf{a}\rangle=0$, it follows that $(\mathbf{a}\times\mathbf{b},\mathbf{a},\mathbf{w})$ is an orthogonal frame (the remaining orthogonality requirements are straightforward to verify). It remains to check the length of $\mathbf{w}$. Since the vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are unit vectors, and because $\measuredangle(\mathbf{a}\times\mathbf{b},\mathbf{a})$ is a right angle, we have

$$|\mathbf{v}|^2=|\mathbf{a}\times\mathbf{b}|^2=(\sin\theta)^2=1-2(\cos\theta)^2+(\cos\theta)^2=\langle\mathbf{b}-\langle\mathbf{b},\mathbf{a}\rangle\cdot\mathbf{a},\mathbf{b}-\langle\mathbf{b},\mathbf{a}\rangle\cdot\mathbf{a}\rangle=|\mathbf{w}|^2.$$

Thus, $\mathcal{B}_{\mathbf{w}}$ equals $\mathcal{B}_{\mathbf{v}}$, hence $\mathbf{v}=\mathbf{w}$. This completes the proof of Step 3 and of the theorem. □

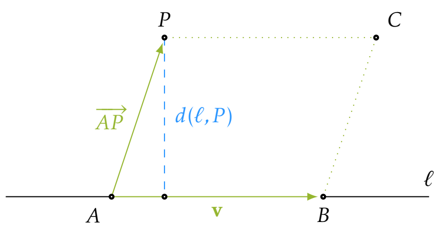
5.2.2 Distance between a point and a line

A line ℓ in dimension 2 is a hyperplane and for any point P of $\mathbb{E}^2$ we have an efficient way of calculating $d(\ell,P)$ with respect to an orthonormal frame of $\mathbb{E}^2$.

Now consider the 3-dimensional case. Let ℓ be a line in $\mathbb{E}^3$ and let P be a point which does not lie on ℓ. Suppose that P has coordinates (x_P, y_P, z_P) with respect to an orthonormal frame $\mathcal{K}$ of $\mathbb{E}^3$, and that ℓ is given by a point $A(x_A, y_A, z_A) \in \ell$ and a direction vector $\mathbf{v}(v_x, v_y, v_z)$ with respect to $\mathcal{K}$. There is a plane π containg both ℓ and P. It is possible to calculate a frame $\mathcal{K}_\pi$ of π, extend it to a frame $\mathcal{K}'$ of $\mathbb{E}^3$, translate the given coordinates and components from $\mathcal{K}$ to $\mathcal{K}'$ and then to the calculations in π (with respect to $\mathcal{K}_\pi$) where ℓ is a hyperplane.

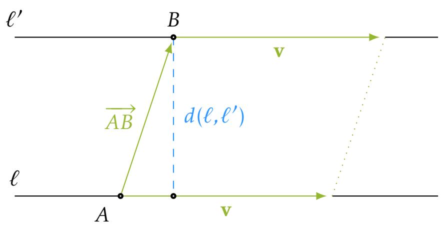
A much more efficient way to calculate $d(\ell,P)$ can be deduced with the cross product. Let B be another point on ℓ such that $\overrightarrow{AB} = \mathbf{v}$ and consider the parallelogram $ABCP$. Then $d(\ell,P)$ is the height of the parallelogram corresponding to the side $[AB]$. Thus

$$d(\ell,P) = \frac{\text{Area}\,(ABCP)}{|AB|} = \frac{|\overrightarrow{AP} \times \mathbf{v}|}{|\mathbf{v}|}.$$



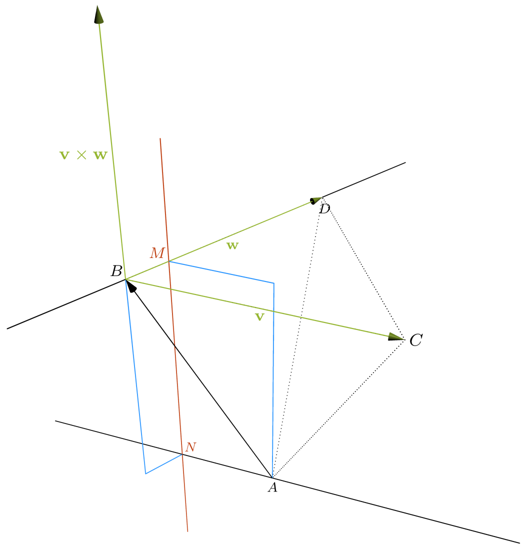
From this we immediately deduce a formula for the distance between two parallel lines in $\mathbb{E}^3$. Let ℓ' be a line parallel to ℓ which passes through a point $B(x_B, y_B, z_B)$. We may apply Porposition 4.26 to conclude that $d(\ell,\ell') = d(\ell,Q)$ for any point Q in ℓ'. Thus,

$$d(\ell,\ell') = d(\ell,B) = \frac{|\overrightarrow{AB} \times \mathbf{v}|}{|\mathbf{v}|}.$$



5.2.3 Common perpendicular line of two skew lines

Consider two vectors $\mathbf{v}$ and $\mathbf{w}$. When proving Theorem 5.10 we reduced the claim to $(\mathbf{v}\times\mathbf{w})\times\mathbf{v}$. Let us give more insight into this expression by deducing equations for the common perpendicular line of two skew lines. Let ℓ and ℓ' be two skew lines in $\mathbb{E}^3$ passing respectively through A and B. Let $\mathbf{v}$ be the direction vector for ℓ and let $\mathbf{w}$ be the direction vector for ℓ'.



The *common perpendicular line* of ℓ and ℓ' is is the unique line d which intersects the two lines and is orthogonal to both of them, i.e. the line satisfying the following properties

- $d\perp\ell$ and $d\perp\ell'$, and
- $d\cap\ell\neq\emptyset$ and $d\cap\ell'\neq\emptyset$.

Since the two lines are skew relative to each other, the vectors $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, hence the vector $\mathbf{v}\times\mathbf{w}$ is non-zero and perpendicular to both $\mathbf{v}$ and $\mathbf{w}$. Let π be the plane passing through A and parallel to $\mathbf{v}$ and $\mathbf{v}\times\mathbf{w}$. Notice that $\mathbf{n}=(\mathbf{v}\times\mathbf{w})\times\mathbf{v}$ is a normal vector for π and that ℓ is included in π. Similarly, there is a plane π' containing ℓ' and having normal vector $\mathbf{n}'=(\mathbf{v}\times\mathbf{w})\times\mathbf{w}$. The intersection of these two planes is a line d and we claim that it is the common perpendicular of ℓ and ℓ'. Since $d=\pi\cap\pi'$, a direction vector for d is

$$\mathbf{n}\times\mathbf{n}'=\big((\mathbf{v}\times\mathbf{w})\times\mathbf{v}\big)\times\big((\mathbf{v}\times\mathbf{w})\times\mathbf{w}\big).$$

Denote $\mathbf{v}\times\mathbf{w}$ by $\mathbf{a}$ and notice that, by Theorem 5.10, we have

$$\mathbf{n}\times\mathbf{n}'=(\mathbf{a}\times\mathbf{v})\times(\mathbf{a}\times\mathbf{w})=\underbrace{\langle\mathbf{a},\mathbf{a}\times\mathbf{w}\rangle}_{=0}\mathbf{v}-\langle\mathbf{v},\mathbf{a}\times\mathbf{w}\rangle\mathbf{a}\quad(\text{proportional to } \mathbf{a}=\mathbf{v}\times\mathbf{w})$$

where $\langle\mathbf{a},\mathbf{a}\times\mathbf{w}\rangle=0$ since the two vectors are orthogonal. It follows that $\mathbf{a}$ is a direction vector for d. But $\mathbf{a}=\mathbf{v}\times\mathbf{w}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$, hence it is orthogonal to both ℓ and ℓ'. Moreover, since d

and ℓ lie in the plane π and have non-parallel direction vectors they necessarely intersect. Similarly, we see that $d \cap \ell' \neq \emptyset$. This shows that d is indeed the common perpendicular of ℓ and ℓ'.

Writing the equations of π and π' in coordinates we obtain

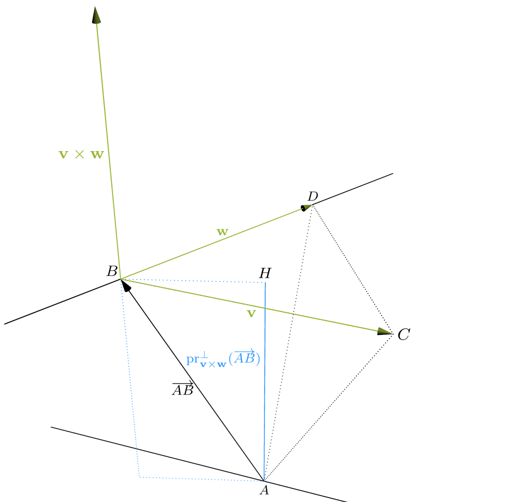
$$d = \pi \cap \pi' : \begin{cases} \pi : \begin{vmatrix} x - x_A & y - y_A & z - z_A \\ v_x & v_y & v_z \\ a_x & a_y & a_z \end{vmatrix} = 0 \\ \pi' : \begin{vmatrix} x - x_B & y - y_B & z - z_B \\ w_x & w_y & w_z \\ a_x & a_y & a_z \end{vmatrix} = 0 \end{cases}$$

where we used the coordinates of A and B with respect to an orthonormal frame $\mathcal{K} = (O, \mathcal{B})$ and the components of $\mathbf{v}$, $\mathbf{w}$ and $\mathbf{a}$ with respect to $\mathcal{B}$.

5.2.4 Distance between two skew lines

As in the previous section, let ℓ and ℓ' be two skew lines in $\mathbb{E}^3$ passing respectively through $A(x_A, y_A, z_A)$ and $B(x_B, y_B, z_B)$. Let $\mathbf{v}(v_x, v_y, v_z)$ be the direction vector for ℓ and let $\mathbf{w}(w_x, w_y, w_z)$ be the direction vector for ℓ'. We assume that the coordinates of the points and the components of the vectors are known.

Let MN be the common perpendicular line of ℓ and ℓ' with $M \in \ell$ and $N \in \ell'$. Consider the cuboid spanned by $\overrightarrow{MA}, \overrightarrow{NM}, \overrightarrow{NB}$ it is easy to show that $d(A, B) \geq d(MN) = |MN|$. Hence $d(\ell, \ell') = |MN|$. Moreover, we could calculate this distance by writing down equations for MN (as in Section 5.2.3), intersecting with ℓ and ℓ' to determine the coordinates of M and N, and then, calculate $|MN|$. However, there is a shorter way of obtaining this distance.



Since the lines are skew, there is a unique plane π containing ℓ which is parallel to ℓ'. Let P be a point in π. Considering the cuboid with vertices M, N and P we see that $d(\ell, \pi) = |MN|$. Hence

$d(\ell,\ell') = d(\ell,\pi)$. Now, dropping a perpendicular from A on π which intersects π in H we see that $d(\ell,\pi) = |AH|$. Therefore

$$d(\ell,\ell') = d(\ell,\pi) = |AH| = |\mathrm{Pr}_{\mathbf{v}\times\mathbf{w}}(\overrightarrow{AB})| = |\frac{\langle \mathbf{v}\times\mathbf{w}, \overrightarrow{AB}\rangle}{|\mathbf{v}\times\mathbf{w}|}|$$

Hence, if we let $\mathbf{a}(a_x,a_y,a_z) = \mathbf{v}\times\mathbf{w}$, we obtain

$$d(\ell_1,\ell_2) = |\frac{\langle \mathbf{v}\times\mathbf{w}, \overrightarrow{AB}\rangle}{|\mathbf{v}\times\mathbf{w}|}| = \frac{\left|a_x(x_B-x_A)+a_y(y_B-y_A)+a_z(z_B-z_A)\right|}{\sqrt{a_x^2+a_y^2+a_z^2}}.$$

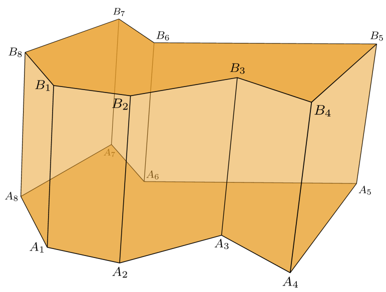
A further interpretation of this formula is the following. Let P be a parallelepiped spanend by the three vectors $\mathbf{v},\mathbf{w},\overrightarrow{AB}$ and denote by F a face spanned by $\mathbf{v}$ and $\mathbf{w}$. The distance between the two lines is the height of P corresponding to the face F. Moreover, anticipating the discussion in Section 5.3, we have

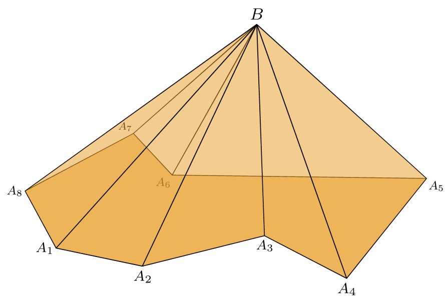
$$d(\ell_1,\ell_2) = \frac{\mathrm{Vol}(P)}{\mathrm{Area}(F)} = |\frac{[\mathbf{v},\mathbf{w},\overrightarrow{AB}]}{|\mathbf{v}\times\mathbf{w}|}|.$$

5.3 Volume

5.3.1 Volume of polyhedra

Definition 5.11. 'A polyhedron may be defined as a finite, connected set of plane polygons, such that every side of each polygon belongs also to just one other polygon, with the proviso that the polygons surrounding each vertex form a single circuit' [7]. Well known examples are *prisms* having two congruent parallel faces, in particular triangular prisms and parallelepipeds, or *pyramids* having just one point, the *apex*, outside the plane of the polygonal *base*. The *height of a prism* is the distance between the two main faces, and the *height of a pyramid* is the distance from the apex to the base.





Triangulation of a polyhedra, or *tetrahedralization*, refers to the subdivision of a polyhedra into tetrahedral meshes. Similar to polygons in dimension 2, it is always possible to break a polyhedron into tetrahedra. Unlike the two dimensional case, it is not always possible to do this without adding vertices. The smallest example of a polyhadron where this is not possible is the Schönhardt polyhedron. For a discussion of the various challanges and algorithms see [20, Chapter 25].

For our purposes, the following non-standard definition suffices. The interior of a tetrahedron is a polyhedral set. If $ABCD$ and $ABCD'$ are two tetrahedra with disjoint interior then the union of the interior of the two tetrahedra together with the interior of the triangle ABC is a polyhedral set obtained by *gluing the two tetrahedra*. Any polyhedral set is obtained after a finite number of gluings of tetrahedra.

Definition 5.12. Let $\mathcal{P}$ denote the set of all polyhedral sets in $\mathbb{E}^3$. We define a *volume* function $\mathrm{Vol} : \mathcal{P} \to \mathbb{R}_{\geq 0}$ through the following properties:

(V1) If S is a cuboid of side lengths 1, 1 and x then $\mathrm{Vol}(S) = x$.

(V2) If S_1 and S_2 are similar polyhedral sets with ratio $x = S_1 : S_2$ then $\mathrm{Vol}(S_1) = x^3 \cdot \mathrm{Vol}(S_2)$.

(V3) If S_1 and S_2 are disjoint polyhedral sets then $\mathrm{Vol}(S_1 \cup S_2) = \mathrm{Vol}(S_1) + \mathrm{Vol}(S_2)$.

Proposition 5.13. We have

1. The volume of a prism of height h and base area a is ah. In particular, the volume of a parallelogram with a face of area a and corresponding height h is ah.

2. The volume of a pyramid of height h and base area a is $ah/3$. In particular, the volume of a tetrahedron $ABCD$ with $\mathrm{Area}(BCD) = a$ and $d(A, BCD) = h$ is $ah/3$.

3. The volume of a rectangular parallelepiped of side lengths a, b and c is abc.

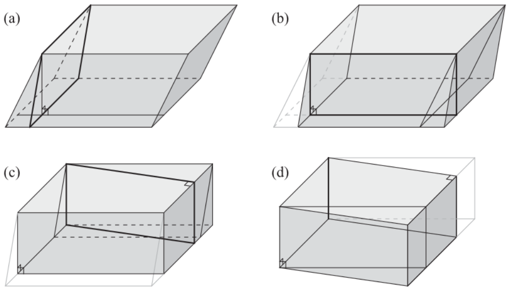
Proof. It is clear that 3. follows from 1. However, the proof needs to start from the definitions. We do this by first proving 3. Denote by $R_{a,b,c}$ a rectangular parallelepiped with side lengths a, b and c. By (V1) and (V2), we have

$$\mathrm{Vol}(R_{a,a,1}) = a^3 \cdot \mathrm{Vol}(R_{1,1,\frac{1}{a}}) = a^3 \cdot \frac{1}{a} = a^2$$

for any $a > 0$. Now, as in the proof of Proposition 5.3, consider $R_{a,a,1}$ and $R_{b,b,1}$ in opposite corners of $R_{a+b,a+b,1}$. By (V3), we have $\mathrm{Vol}(R_{a+b,a+b,1}) = \mathrm{Vol}(R_{a,a,1}) + \mathrm{Vol}(R_{b,b,1}) + 2 \cdot \mathrm{Vol}(R_{a,b,1})$ hence $(a+b)^2 = a^2 + b^2 + 2 \cdot \mathrm{Vol}(R_{a,b,1})$ and therefore $\mathrm{Vol}(R_{a,b,1}) = ab$ for any $a, b > 0$. Then, for an arbitrary rectangular parallelepiped we have

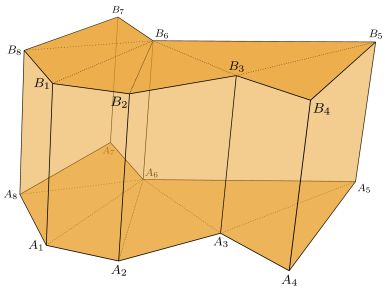
$$\mathrm{Vol}(R_{a,b,c}) = c^3 \cdot \mathrm{Vol}(R_{\frac{a}{c},\frac{b}{c},1}) = c^3 \cdot \frac{ab}{c^2} = abc.$$

Next, we use 3. to deduce the volume of an arbitrary parallelepiped. For this we section and rearrange the parallelepiped as in the figure (this is Figure 4.1.2 from [1]) below.

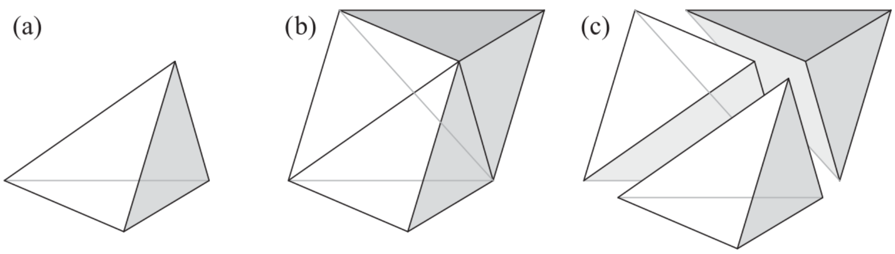


In (a), (b) and (c) the parallelepiped is transformed into a prism with base a parallelogram by cutting and pasting congruent triangular prisms. In particular, the height h and the base area a are unchanged. In the last step (d), the base is rearranged into a rectangle without changing its area - as in the proof of Proposition 5.3. The final result is a rectangular parallelepiped $R_{x,y,h}$ with volume ah equal to the volume of the initial parallelepiped.

Now, slicing a parallelepiped along the diagonals of two opposite faces we obtain a triangular prism with volume ha where a is half the area of the face which was cut. Thus, a triangular prism has the claimed area. For an arbitrary prism, use a triangulation of tha base polygon to deduce the claim.



For 2. we first consider tetrahedra and follow the argument in [1, §4.1]. Slice a rectangular prism as in the following figure[1]. The proof of Proposition 5 of Book XII of Euclid's *Elements* (see for example [10]), is correct for our settings and assumptions. It implies that two triangular pyramids (tetrahedra) with the same height and with congruent bases have the same volume [10, Volume 3, p.390].



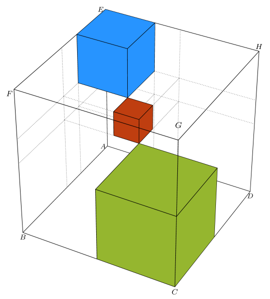
[1] Figure 4.1.4 from [1]

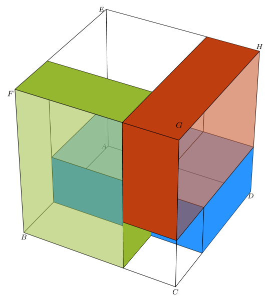
In the above slicing the two tetrahedra with white base have the same heights and the bottom and top tetrahedra are clearly congruent. Thus, if T is any of the three tetrahedra of the triangular prism P, then

$$\mathrm{Vol}(T) = \frac{1}{3}\mathrm{Vol}(P) = \frac{ah}{3}$$

where a is the area of a base triangle of P and h is the height of P. Then, for an arbitrary pyramid, the claim follows by considering a tetrahedralization. □

Remark. Similar to our definition of area, the above definition of volume tries to capture basic principles of the intuitive notion of volume. It is an over-simplified version of a so-called Lebesgue measure (see Section 5.3.3). In property (V1) we may replace the rectangular parallelepiped with a cube of side-length 1. However, after weakening the assumption, it is not possible to cover an arbitrary rectangle allowing only the uniform scaling given in (V2). To see this difficulty, keep the notation in the proof of Proposition 5.13 and place three cubes C_a, C_b and C_c on the diagonal of a cube C_{a+b+c}.





The planes of the small cubes divide $C_{a,b,c}$ in cuboids the sides of which can be checked to correspond to the algebraic formula for the cube of $a+b+c$. We have

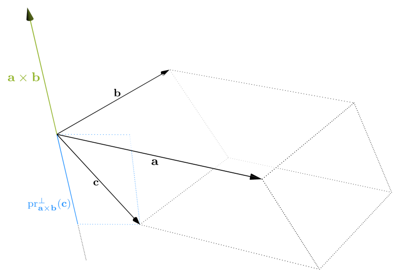
$$\mathrm{Vol}(C_{a,b,c}) = (a+b+c)^3 = a^3+b^3+c^3+3a^2b+3a^2c+3ab^2+3b^2c+3ac^2+3bc^2+6abc. \tag{5.3}$$

Assuming that the volume $\mathrm{Vol}(C_{x,y,y})$ of a cuboid with two equal adjacent sides is xy^2, it follows from (V2), (V3) and (5.3) that $\mathrm{Vol}(R_{a,b,c}) = abc$. Thus, it is necessary and sufficient to show that $\mathrm{Vol}(C_{x,y,y}) = xy^2$. For this, place two cubes C_a and C_b on the diagonal of C_{a+b} and notice that the subdivision of the big cube corresponds to

$$\mathrm{Vol}(C_{a+b}) = (a+b)^3 = a^3+b^3+3ab(a+b).$$

From this equation and (V2), (V3) it follows that $\mathrm{Vol}(R_{a,b,(a+b)}) = ab(a+b)$ for any $a,b>0$. However, this does not yield a decomposition of a cube into smaller cubes and rectangles congruent to $R_{x,y,y}$ for arbitrary x and y. Compare this to the two dimensional case.

Orthonormal frames offer an efficient way of calculating the volume of a parallelogram, and therefore volumes of tetrahedra and polyhedra provided that the coordinates of vertices are known.



Proposition 5.14. The volume of a parallelepiped P spanned by the vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ is

$$\mathrm{Vol}(P) = |\langle \mathbf{a}\times\mathbf{b},\mathbf{c}\rangle| = |\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}|$$

where the components of $\mathbf{a}(a_1,a_2,a_3)$, $\mathbf{b}(b_1,b_2,b_3)$ and $\mathbf{c}(c_1,c_2,c_3)$ are with respect to an orthonormal basis.

Proof. The area of the face spanned by $\mathbf{a}$ and $\mathbf{b}$ is $|\mathbf{a}\times\mathbf{b}|$ and the height corresponding to this face is $|\mathrm{pr}^{\perp}_{\mathbf{a}\times\mathbf{b}}(\mathbf{c})|$. Thus

$$\mathrm{Vol}(P) = |\mathbf{a}\times\mathbf{b}|\cdot|\mathrm{pr}^{\perp}_{\mathbf{a}\times\mathbf{b}}(\mathbf{c})| = |\mathbf{a}\times\mathbf{b}|\cdot|\frac{\langle \mathbf{a}\times\mathbf{b},\mathbf{c}\rangle}{\langle \mathbf{a}\times\mathbf{b},\mathbf{a}\times\mathbf{b}\rangle}\mathbf{a}\times\mathbf{b}| = |\langle \mathbf{a}\times\mathbf{b},\mathbf{c}\rangle|.$$

Moreover, with respect to an orthonormal basis we have

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

□

5.3.2 Oriented volume

Definition 5.15. The above proposition shows that, up to sign, the volume of a parallelepiped spanned by $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ is the determinant of the base change matrix $\mathrm{M}_{\mathcal{B},\mathcal{B}'}$ there $\mathcal{B}$ is a right-oriented basis and $\mathcal{B}' = (\mathbf{a},\mathbf{b},\mathbf{c})$. The value $\det(\mathrm{M}_{\mathcal{B},\mathcal{B}'})$ does not depend on the orthonormal basis $\mathcal{B}$ in which the vectors are expressed. The proof of this claim is the same in any dimension (see Proposition 5.21). We call this determinant the *box product of* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$ and denote it by $[\mathbf{a},\mathbf{b},\mathbf{c}]$. We define the *oriented volume of a parallelepiped P spanned by the vectors* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$ to be

$$\mathrm{Vol}_{\mathrm{or}}(P) = [\mathbf{a},\mathbf{b},\mathbf{c}].$$

Similar, the oriented area of the tetrahedron $ABCD$ is $\mathrm{Vol}_{\mathrm{or}}(ABCD) = [\overrightarrow{AB},\overrightarrow{AC},\overrightarrow{AD}]/6$. Proposition 5.14 shows that mixing the cross product with the scalar product gives the box product. For this reason $\langle \mathbf{a}\times\mathbf{b},\mathbf{c}\rangle = [\mathbf{a},\mathbf{b},\mathbf{c}]$ is sometimes called the *mixed product* of the vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$.

Proposition 5.16. Box product defines a map $[\_,\_,\_]:\mathbb{V}^3\times\mathbb{V}^3\times\mathbb{V}^3\to\mathbb{R}\quad(\mathbf{a},\mathbf{b},\mathbf{c})\mapsto[\mathbf{a},\mathbf{b},\mathbf{c}]$ which is linear in each argument. Moreover, for three vectors $\mathbf{a},\mathbf{b},\mathbf{c}$ we have

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

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

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

The sign of the box product determines the orientation of the basis $(\mathbf{a},\mathbf{b},\mathbf{c})$

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

Proof. In dimension 3, linearity of the box product follows from the fact that it is the composition of linear maps, cross product and scalar product. You can also deduce the linearity of this map from the fact that with respect to an orthonormal basis it is the determinant of a base change matrix. The other properties follow from this latter description and from known properties of the determinant. □

Proposition 5.17 (Triple cross product formula)**.** For vectors $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}\in\mathbb{V}^3$ we have

$$(\mathbf{a}\times\mathbf{b})\times(\mathbf{c}\times\mathbf{d})=\mathbf{b}\cdot[\mathbf{a},\mathbf{c},\mathbf{d}]-\mathbf{a}\cdot[\mathbf{b},\mathbf{c},\mathbf{d}]=\mathbf{c}\cdot[\mathbf{a},\mathbf{b},\mathbf{d}]-\mathbf{d}\cdot[\mathbf{a},\mathbf{b},\mathbf{c}]$$

5.3.3 Hypervolume

The higher-dimensional analogues of polygons and polyhedra are *polytopes*. The simplest polytope in dimension n is an n-simplex. If $n=2$ these are triangles and if $n=3$ these are tetrahedra. In general, in dimension n, consider $n+1$ points $P_0,P_1,\dots,P_n$ which do not lie in a hyperplane. This is equivalent to asking for the vectors $\overrightarrow{P_0P_1},\dots,\overrightarrow{P_0P_n}$ to be linearly independent, i.e. to form a basis of $\mathbb{V}^n$. The convex hull of $P_0,P_1,\dots,P_n$ is an n-simplex. The $(n-1)$-faces of an n-simplex are $n-1$-simplices, in particular, the faces of a 4-simplex are tetrahedra. We may define polytopes as sets obtained by gluing n-simplices along their faces.

A hyperparallelepiped spanned by $\mathbf{v}_1,\dots,\mathbf{v}_n$ is the set of all points obtained from a given point P with linear combinations of the given vectors considering only coefficients in the interval $[0,1]$, concretely

$$\mathcal{H}=P+\{\alpha_1\mathbf{v}_1+\cdots+\alpha_n\mathbf{v}_n:\alpha_1,\dots,\alpha_n\in[0,1]\}.$$

A *hypercube* is a hyperparallelepiped with all 1-faces of equal length and all angles right angles.

Definition 5.18. Let $\mathcal{P}$ denote the set of all polytopes in $\mathbb{E}^3$. We define a *rational hypervolume* function $\mathrm{Vol}:\mathcal{P}\to\mathbb{R}_{\geq 0}$ through the following properties:

(V1) If S is hypercube of side length 1 then $\mathrm{Vol}(S)=1$.

(V2) If S_1 and S_2 are congruent polytopes then $\mathrm{Vol}(S_1)=\mathrm{Vol}(S_2)$.

(V3) If S_1 and S_2 are disjoint polytopes then $\mathrm{Vol}(S_1\cup S_2)=\mathrm{Vol}(S_1)+\mathrm{Vol}(S_2)$.

This volume function is called rational because it does not allow us to measure all polytopes. For instance it is not possible to use it to deduce the volume of a cube with sides an arbitrary irrational numbers.

Definition 5.19. We say that a set S is *measurable*, if it contains a sequence of strictly growing internal polytopes I_i and if it is contained in a sequance of strictly shrinking exterior polytopes E_i such that the volumes of these sets converge to a common value. Formaly, the following properties need to hold

$$I_i \subseteq I_j \subseteq S \subseteq E_j \subseteq E_i \quad \forall i < j \quad \text{and} \quad \lim_{i\to\infty} \text{Area}(I_i) = \lim_{i\to\infty} \text{Area}(E_i)$$

If the limit exists, it is unique and we denote it by $\text{Vol}(S)$. Denoting by $\mathcal{M}$ the set of all measurable sets of $\mathbb{E}^n$, we extended the (hyper)volume function to $\text{Vol} : \mathcal{M} \to \mathbb{R}_{\geq 0}$. This is a version of the Lebesgue measure (see for example [18, Chapter 11]).

From this definition one may deduce that a cuboid with side lengths $a_1, \ldots, a_n$ has volume $\prod a_i$. Or that if P is a hyperparallelepiped in dimension n, and if S is any n-simplex constructed from vertices of P then

$$\text{Vol}(S) = \frac{1}{n!}\text{Vol}(P)$$

which is the higher-dimensional analogue of the volume formula for a tetrahedron.

As in the case of dimension 2 and 3, the efficient way of calculating hypervolume is given by the box product which allows us to calculate the hypervolume of hyperparallelepipeds and therefore of n-simplices and of polytopes in general.

Definition 5.20. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors in $\mathbb{V}^n$ with components $\mathbf{v}_i(v_{i,1}, \ldots, v_{i,n})$ relative to a right oriented orthonormal basis $\mathcal{B}$ of $\mathbb{V}^n$. The *(n-fold) box product* of these vectors is

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

Proposition 5.21. The definition of the box product does not depend on the choice of the right oriented orthonormal basis $\mathcal{B}$.

Proof. Let $\mathcal{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ be an n-tuple of n-dimensional vectors and let $\text{M}_{\mathcal{B},\mathcal{V}}$ be the matrix whose column are the components of the $\mathbf{v}_i's$. Notice that

$$[\mathbf{v}_1, \ldots, \mathbf{v}_n] = \det(\text{M}_{\mathcal{B},\mathcal{V}})$$

If the vectors are linearly dependent, then $\det(\text{M}_{\mathcal{B},\mathcal{V}}) = 0$ for any basis $\mathcal{B}$. If the vectors are linearly independent, then $\mathcal{V}$ is a basis and if $\mathcal{B}'$ is another left oriented orthonormal bases then

$$\det(\text{M}_{\mathcal{B},\mathcal{V}}) = \det(\text{M}_{\mathcal{B}',\mathcal{B}}\text{M}_{\mathcal{B},\mathcal{V}}) = \det(\text{M}_{\mathcal{B}',\mathcal{B}})\det(\text{M}_{\mathcal{B},\mathcal{V}})$$

and it suffices to show that $\det(M_{\mathcal{B}',\mathcal{B}}) = 1$. Let $\mathbf{e}_i$ denote the vectors in $\mathcal{B}$. Then, since $\mathcal{B}'$ is orthonormal, we have

$$M^T_{\mathcal{B}',\mathcal{B}} M_{\mathcal{B}',\mathcal{B}} = \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle & \langle \mathbf{e}_1, \mathbf{e}_2 \rangle & \dots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle \\ \langle \mathbf{e}_2, \mathbf{e}_1 \rangle & \langle \mathbf{e}_2, \mathbf{e}_2 \rangle & \dots & \langle \mathbf{e}_2, \mathbf{e}_n \rangle \\ \vdots & \vdots & \dots & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1 \rangle & \langle \mathbf{e}_n, \mathbf{e}_2 \rangle & \dots & \langle \mathbf{e}_n, \mathbf{e}_n \rangle \end{bmatrix} = I_n$$

hence

$$1 = \det(I_n) = \det(M^T_{\mathcal{B}',\mathcal{B}} M_{\mathcal{B}',\mathcal{B}}) = \det(M^T_{\mathcal{B}',\mathcal{B}}) \det(M_{\mathcal{B}',\mathcal{B}}) = \det(M_{\mathcal{B}',\mathcal{B}})^2.$$

It follows that $\det(M_{\mathcal{B}',\mathcal{B}}) = \pm 1$ and since the two bases are both right-oriented, we deduce that $\det(M_{\mathcal{B}',\mathcal{B}}) = 1$ and the claim follows. □

Definition 5.22. 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.

One can also generalize the **J**-operator and the cross product to higher dimensions. For $n-1$ vectors $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$, the *wedge product* of these vectors, denoted by $\mathbf{v}_1 \wedge \dots \wedge \mathbf{v}_{n-1}$ is the unique vector $\mathbf{w}$ such that $|\mathbf{w}|$ is the hypervolume of the hyperparallelepiped spanned by $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$, the vector $\mathbf{w}$ is orthogonall to each $\mathbf{v}_i$ and $(\mathbf{v}_1, \dots, \mathbf{v}_{n-1}, \mathbf{w})$ is right oriented.

CHAPTER 6

Affine maps

Contents

6.1 Properties of affine maps **87**
6.1.1 Homogeneous coordinates and homogeneous matrix 89
6.2 Projections and reflections **90**
6.2.1 Tensor product 90
6.2.2 Parallel projection on a hyperplane 90
6.2.3 Parallel reflection in a hyperplane 93
6.2.4 Parallel projection on a line 95
6.2.5 Parallel reflection in a line 97

6.1 Properties of affine maps

Definition 6.1. A map $\phi : \mathbb{A}^n \to \mathbb{A}^m$ is an affine map 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

$$[\phi(P)]_{\mathcal{K}'} = A \cdot [P]_{\mathcal{K}} + b \tag{6.1}$$

relative to some frame $\mathcal{K} = (O, \mathcal{B})$ of $\mathbb{A}^n$ and some frame $\mathcal{K}' = (O', \mathcal{B}')$ of $\mathbb{A}^m$. Such a map defines a linear map $\mathrm{lin}(\phi) : \mathrm{D}(\mathbb{A}^n) \to \mathrm{D}(\mathbb{A}^m)$ where

$$[\mathrm{lin}(\phi)(\overrightarrow{PQ})]_{\mathcal{B}'} = A \cdot [\overrightarrow{PQ}]_{\mathcal{B}} \tag{6.2}$$

since

$$[\mathrm{lin}(\phi)(\overrightarrow{PQ})]_{\mathcal{B}'} = [\overrightarrow{\phi(P)\phi(Q)}]_{\mathcal{B}'} = (A \cdot [Q]_{\mathcal{K}} + b) - (A \cdot [P]_{\mathcal{K}} + b) = A \cdot ([Q]_{\mathcal{K}} - [P]_{\mathcal{K}}) = A \cdot [\overrightarrow{PQ}]_{\mathcal{B}}.$$

We notice that both A and b depend on the choice of the frames $\mathcal{K}$ 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 endomorphism*. The set of all such endomorphisms is denoted by $\mathrm{End}_{\mathrm{aff}}(\mathbb{A}^n)$. It is easy to see that, the affine map 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 (see proof of Proposition 6.5). 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 $\mathrm{AGL}(\mathbb{A}^n)$

Moreover, if in addition to $n = m$, $O = O'$ and $b = 0$ in (6.1) 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 invertible 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 6.2. 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 6.3. The various parallel projections and reflections described in the next sections are examples of affine maps as well as isometries discussed in Chapter 7.

Proposition 6.4. Let $\phi : \mathbb{E}^n \to \mathbb{E}^m$ be an affine map. If a line ℓ is mapped onto a line ℓ' under ϕ, then ϕ preserves the oriented ratio on ℓ, i.e. if A, B, C, D are points on ℓ, then

$$\frac{\overrightarrow{AC}}{\overrightarrow{AB}} = \frac{\overrightarrow{\phi(A)\phi(C)}}{\overrightarrow{\phi(A)\phi(B)}}. \tag{6.3}$$

In particular the ratios on ℓ are preserved.

Proof. The orented ratio $\frac{\overrightarrow{AC}}{\overrightarrow{AB}}$ denotes a scalar λ such that $\overrightarrow{AC} = \lambda\overrightarrow{AB}$. Since $\overrightarrow{\phi(A)\phi(C)} = \mathrm{lin}(\phi)(\overrightarrow{AC})$ and since $\mathrm{lin}(\phi)$ is a linear map we have

$$\overrightarrow{\phi(A)\phi(C)} = \mathrm{lin}(\phi)(\overrightarrow{AC}) = \mathrm{lin}(\phi)(\lambda\overrightarrow{AB}) = \lambda\,\mathrm{lin}(\phi)(\overrightarrow{AB}) = \lambda\overrightarrow{\phi(A)\phi(B)}$$

and the claim follows. □

Proposition 6.5. 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 oriented ratio on lines.

Proof. Throughout we let ϕ be an affine map as in (6.1). First we notice that ϕ is an affine transformation, i.e. invertible, if the matrix A is invertible. Indeed, for two points P,Q we have

$$[\phi(P)]_{\mathcal{K}'} = [\phi(Q)]_{\mathcal{K}'} \quad\Leftrightarrow\quad A\cdot[P]_{\mathcal{K}} + b = A\cdot[Q]_{\mathcal{K}} + b \quad\Leftrightarrow\quad A\cdot[P]_{\mathcal{K}} = A\cdot[Q]_{\mathcal{K}}.$$

Therefore, injectivity and surjectivity of ϕ are equivalent to injectivity respectively surjectivity of multiplication with A, i.e. ϕ is bijective if and only if A is invertible.

Now we start with the proof of our proposition. Assume that ϕ is an affine transformation. It is bijective, so, in particular injective. By the above paragraph, the matrix A is invertible. Then, a point on a line $\ell = \{P + t\mathbf{v} : t \in \mathbb{R}\}$ is mapped to

$$A\cdot[P + t\mathbf{v}]_{\mathcal{K}} + b = (A\cdot[P] + b) + t\cdot A\cdot[\mathbf{v}]_{\mathcal{B}}$$

which shows that $\phi(\ell)$ is a line passing through $(A\cdot[P]+b) \in \mathbb{A}^n$ and having direction vector $A\cdot[\mathbf{v}]_{\mathcal{B}} \in \mathrm{D}(\mathbb{A}^n)$ (which is a non-zero vector since A is invertible). The third claim follows from Proposition 6.3.

For the converse, assume that $\phi : \mathbb{A}^n \to \mathbb{A}^n$ is a map with the three indicated properties. Since ϕ preserves oriented ratios, it preserves midpoints of segments, i.e. if M is the midpoint of the segment $[AB]$ then $\phi(M)$ is the midpoint of the segment $[\phi(A)\phi(B)]$. Therefore, ϕ preserves parallelograms. Hence ϕ preserves sums of any two vectors, i.e. if $\overrightarrow{OA} + \overrightarrow{OC} = \overrightarrow{OB}$ then $\overrightarrow{\phi(O)\phi(A)} + \overrightarrow{\phi(O)\phi(C)} = \overrightarrow{\phi(O)\phi(B)}$. Moreover, since ϕ preserves oriented ratios, if $\overrightarrow{OA} = \lambda\overrightarrow{OB}$ then $\overrightarrow{\phi(O)\phi(A)} = \lambda\overrightarrow{\phi(O)\phi(B)}$. This two observations show that ϕ induces a linear map on vectors, i.e. that the map

$$\mathrm{lin}(\phi) : \overrightarrow{OA} \mapsto \overrightarrow{\phi(O)\phi(A)}$$

is linear. Now, Since ϕ is injective, it is easy to see that $\mathrm{lin}(\phi) : \mathrm{D}(\mathbb{A}^n) \to \mathrm{D}(\mathbb{A}^n)$ is injective. Moreover, it is a linear algebra fact that an injective linear map from a vectorspace to itself is bijective. Therefore, $\mathrm{lin}(\phi)$ is bijective.

Now let $\mathcal{K} = (O, \mathcal{B})$ be a frame in $\mathbb{A}^n$ where $\mathcal{B} = (\overrightarrow{OX_1}, \dots, \overrightarrow{OX_n})$. Let $O' = \phi(O)$, $Y_i = \phi(X_i)$ and $\mathcal{B}' = (\overrightarrow{O'Y_1}, \dots, \overrightarrow{O'Y_n})$. Since $\mathrm{lin}(\phi)$ is bijective, $\mathcal{B}'$ is a basis of $\mathrm{D}(\mathbb{A}^n)$ and for any point P we have

$$[\phi(P)]_{\mathcal{K}'} = [\,\overrightarrow{O'\phi(P)}\,]_{\mathcal{B}'} = [\,\overrightarrow{\phi(O)\phi(P)}\,]_{\mathcal{B}'} + [\,\overrightarrow{O'\phi(O)}\,]_{\mathcal{B}'} = \underbrace{\mathrm{M}_{\mathcal{B},\mathcal{B}'}(\mathrm{lin}(\phi))}_{A}[P]_{\mathcal{B}'} + \underbrace{[\,\overrightarrow{O'\phi(O)}\,]_{\mathcal{B}'}}_{b}.$$

This shows that ϕ has the expression (6.1), i.e. it is an affine map, and since A is invertible, ϕ is a transformation by the first paragraph of the proof. □

6.1.1 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 6.6. 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 6.7. The *homogeneous matrix* of an affine map $\phi : \mathbb{A}^n \to \mathbb{A}^m$ 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

$$\begin{bmatrix} \psi \circ \phi(\mathbf{x}) \\ 1 \end{bmatrix} = \begin{bmatrix} A' & b' \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}.$$

6.2 Projections and reflections

6.2.1 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 6.8. 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 6.9. 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 $\mathbf{u},\mathbf{v},\mathbf{w}\in\mathbb{V}^n$.

Proof. 1. Linearity of from the linearity of matrix multiplication. To see this for the first argument, take two scalars α,β, three vectors $\mathbf{u},\mathbf{v},\mathbf{w}$ and notice that

$$(\alpha\mathbf{u}+\beta\mathbf{v})\otimes\mathbf{w}=(\alpha\mathbf{u}+\beta\mathbf{v})\cdot\mathbf{w}^T=\alpha(\mathbf{u}\cdot\mathbf{w}^T)+\beta(\mathbf{v}\cdot\mathbf{w}^T)=\alpha\mathbf{u}\otimes\mathbf{w}+\beta\mathbf{v}\otimes\mathbf{w}.$$

2. For the second claim we use properties of transposition of matrices

$$(\mathbf{v}\otimes_{\mathcal{B}}\mathbf{w})^T=(\mathbf{v}\cdot\mathbf{w}^T)^T=(\mathbf{w}^T)^T\cdot\mathbf{v}^T=\mathbf{w}\cdot\mathbf{v}^T=\mathbf{w}\otimes\mathbf{v}.$$

3. For the last point we use the fact that $\langle\mathbf{v},\mathbf{w}\rangle=\mathbf{v}^T\cdot\mathbf{w}$ since we are in an orthonormal basis. Then

$$(\mathbf{u}\otimes_{\mathcal{B}}\mathbf{v})\cdot\mathbf{w}=(\mathbf{u}\cdot\mathbf{v}^T)\cdot\mathbf{w}=\mathbf{u}\cdot(\mathbf{v}^T\cdot\mathbf{w})=\mathbf{u}\cdot\langle\mathbf{v},\mathbf{w}\rangle=\langle\mathbf{v},\mathbf{w}\rangle\cdot\mathbf{u}.$$

□

6.2.2 Parallel projection on a hyperplane

Example 6.10. 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 $P'=\ell\cap\ell_P$.

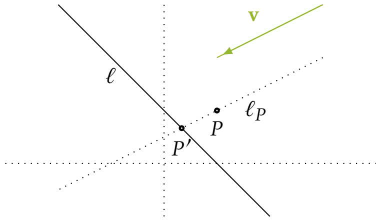


Figure 6.1: Projection of the point P on the line ℓ in the direction of the vector $\mathbf{v}$.

To determine P', we check which point on ℓ_P satisfies the equation of ℓ

$$x_P - 2t + y_P - t - 1 = 0 \quad \Rightarrow \quad t = \frac{1}{3}(x_P + y_P - 1).$$

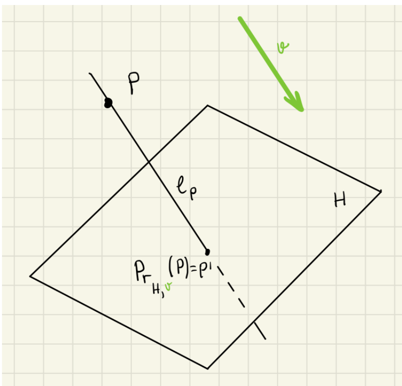
Thus, the projection of P is

$$[P']_{\mathcal{K}} = \begin{bmatrix} \frac{1}{3}x_P - \frac{2}{3}y_p + \frac{2}{3} \\ -\frac{1}{3}x_P + \frac{2}{3}y_p + \frac{1}{3} \end{bmatrix} = \underbrace{\frac{1}{3}\begin{bmatrix} 1 & -2 \\ -1 & 2 \end{bmatrix}}_{A}\begin{bmatrix} x_P \\ y_P \end{bmatrix} + \underbrace{\frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{b}.$$

Definition 6.11. 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

$$H : a_1x_1 + \cdots + a_nx_n + a_{n+1} = 0 \tag{6.4}$$

and a line ℓ_P passing through the point $P(p_1,\dots,p_n)$ and having $\mathbf{v}(v_1,\dots,v_n)$ as direction vector:

$$\ell_P = \{P + t\mathbf{v} : t \in \mathbb{R}\}. \tag{6.5}$$

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

$$P' = P - \frac{a_1p_1 + \cdots + a_np_n + a_{n+1}}{a_1v_1 + \cdots + a_nv_n}\mathbf{v} = P - \frac{\mathbf{a}^T \cdot P + a_{n+1}}{\mathbf{a}^T \cdot \mathbf{v}}\mathbf{v}. \tag{6.6}$$

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

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

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 maps. 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 6.12. 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 $\mathbf{v} = \mathbf{a}$.

6.2.3 Parallel reflection in a hyperplane

Example 6.13. We consider again the setup in Example 6.10. 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

$$\frac{P + P'}{2} = \mathrm{Pr}_{\ell,\mathbf{v}}(P) \quad \Rightarrow \quad P' = 2\,\mathrm{Pr}_{\ell,\mathbf{v}}(P) - P.$$

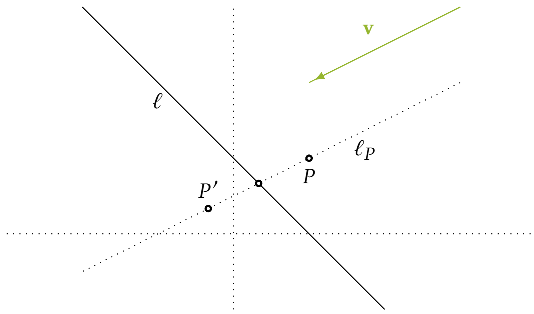


Figure 6.2: Reflection of the point P in the line ℓ parallel to the vector $\mathbf{v}$.

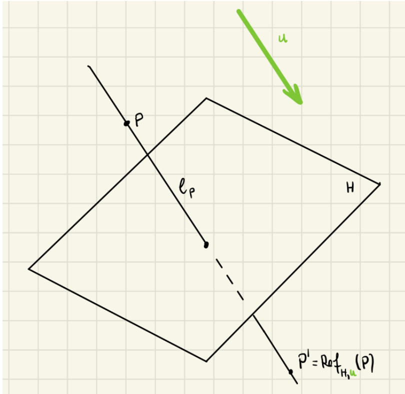
Using the calculation in Example 6.10, the reflection of P is

$$[P']_{\mathcal{K}} = \begin{bmatrix} -\frac{1}{3}x_P - \frac{4}{3}y_p + \frac{4}{3} \\ -\frac{2}{3}x_P + \frac{1}{3}y_p + \frac{2}{3} \end{bmatrix} = \underbrace{\frac{1}{3}\begin{bmatrix} -1 & -4 \\ -2 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x_P \\ y_P \end{bmatrix} + \underbrace{\frac{2}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix}}_{b}.$$

Definition 6.14. 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 $P \in \mathbb{A}^n$ 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 (6.4). The idea here is the same as in Example 6.13: $\mathrm{Pr}_{H,\mathbf{v}}(P)$ 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 maps. 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 6.15. 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

$$\mathrm{Ref}_H^{\perp}(P) = \left(I_n - 2\frac{\mathbf{a} \otimes \mathbf{a}}{|\mathbf{a}|^2}\right) \cdot P - 2\frac{a_{n+1}}{|\mathbf{a}|^2}\mathbf{a} \tag{6.7}$$

since we may choose $\mathbf{v} = \mathbf{a}$.

6.2.4 Parallel projection on a line

Example 6.16. Consider again Example 6.10. 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 6.1. 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

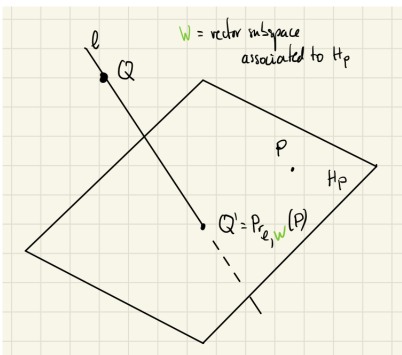
$$(1 - t - x_P) - 2(t - y_P) = 0 \quad \Rightarrow \quad t = -\frac{1}{3}(x_P - 2y_P - 1).$$

As expected, a short calculation shows that P' has the same expression as in Example 6.10.

Definition 6.17. 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

$$\mathbb{W}: a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0 \quad \Leftrightarrow \quad \mathbf{a}^T \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 0 \tag{6.8}$$

where $\mathbf{a} = \mathbf{a}(a_1, \dots, a_n)$. Thus, for a given point $P(p_1, \dots, p_n) \in \mathbb{A}^n$, 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 $Q(q_1, \dots, q_n) \in \ell$. Hence, if we denote by $p_1', \dots, p_n'$ the coordinates of the projected point $\mathrm{Pr}_{\ell,\mathbb{W}}(P)$ then, by (6.6),

$$\begin{cases} p_1' = q_1 + v_1\mu \\ \quad \vdots \\ p_n' = q_n + v_n\mu \end{cases} \quad \text{where} \quad \mu = -\frac{\mathbf{a}^T \cdot Q - \mathbf{a}^T \cdot P}{\mathbf{a}^T \cdot \mathbf{v}}$$

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

$$M_{\mathcal{B}}(\mathrm{lin}(\mathrm{Pr}_{\ell,\mathbb{W}})) = \frac{\mathbf{v} \otimes \mathbf{a}}{\langle \mathbf{v}, \mathbf{a} \rangle}.$$

Parallel projections on lines are affine maps. 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 6.18. 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

$$\mathrm{Pr}_{\ell}^{\perp}(P) = \frac{\mathbf{a} \otimes \mathbf{a}}{|\mathbf{a}|^2} P + \left(I_n - \frac{\mathbf{a} \otimes \mathbf{a}}{|\mathbf{a}|^2} \right) Q$$

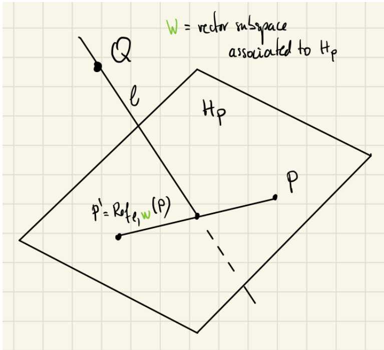
since we may choose $\mathbf{v} = \mathbf{a}$.

6.2.5 Parallel reflection in a line

Definition 6.19. 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 6.2.4, the vector subspace $\mathbb{W}$ is given by the homogeneous equation 6.8. The idea is similar to the one used in Section 6.2.3: 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 $Q(q_1,\ldots,q_n)$ is a point on ℓ and $\mathbf{v}(v_1,\ldots,v_n)$ 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 maps. 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 6.20. 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 $\mathbf{v} = \mathbf{a}$.

CHAPTER 7

Isometries

Contents

7.1 Affine form of isometries . . . 99
7.1.1 The Euclidean space $\mathbb{E}^n$ (second revision) . . . 101
7.2 Isometries in dimension 2 . . . 101
7.2.1 Rotations . . . 101
7.2.2 Classification . . . 103
7.3 Isometries in dimension 3 . . . 104
7.3.1 Rotations . . . 104
7.3.2 Euler angles . . . 106
7.3.3 Classification . . . 107
7.4 Moving points with isometries . . . 108
7.4.1 Cycloids . . . 108
7.4.2 Surfaces of revolution . . . 111

7.1 Affine form of isometries

Definition 7.1. 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 7.2. Isometries are affine transformations.

Proof. Let $\phi : \mathbb{E}^n \to \mathbb{E}^n$ be an isometry. We prove our claim using the characterization of affine transformations in Proposition 6.5.

The map ϕ is *injective*: if $\phi(P) = \phi(Q)$ then

$$0 = d(\phi(P), \phi(Q)) = d(P, Q)$$

which implies that $P = Q$.

The map ϕ *preserves lines*: consider three collinear points $A, B, C \in \mathbb{E}^n$. One of the points will lie between the other two, assume that $C \in [AB]$. Then $d(A,B) = d(A,C) + d(C,B)$. Let $A' = \phi(A)$, $B' = \phi(B')$ and $C' = \phi(C)$. Then $d(A',B') = d(A,B)$, $d(A'C') = d(A,C)$ and $d(B',C') = d(B,C)$. Therefore

$$d(A',B') = d(A,B) = d(A,C) + d(C,B) = d(A',C') + d(C',B')$$

which implies $C' \in [A'B']$.

The map ϕ *preserves oriented ratios*: notice that

$$\frac{d(A',C')}{d(A',B')} = \frac{d(A,C)}{d(A,B)}$$

hence

$$\frac{\overrightarrow{A'C'}}{\overrightarrow{A'B'}} = \pm\frac{d(A',C')}{d(A',B')} = \pm\frac{d(A,C)}{d(A,B)} = \pm\frac{\overrightarrow{AC}}{\overrightarrow{AB}}.$$

But, the two ratios have the same sign since, as in the previous paragraph, C is between A and B if and only if C' is between A' and B'. □

Remark (Notation)**.** At this point some simplifications in notation are useful. When it is clear from the context that we are working in a frame $\mathcal{K} = (O, \mathcal{B})$, we will simply write $\mathbf{v}$ instead of $[\mathbf{v}]_{\mathcal{B}}$. We identify points with column matrices. For example, we write $\mathbf{x}$ to mean the column matrix with entries $(x_1, \ldots, x_n)$. This will represent both the coordinates of a point and the components of the position vector.

Moreover, when there is no risk of confusion, we denote the linear map induced by an afine transformation with the same letter, i.e. we write ψ instead of $\mathrm{lin}(\psi)$. The arguments are either points or vectors, and the context will make it clear which one it is.

Example 7.3. Translation maps are isometries. Let $\mathbf{v}$ be a vector then, the *translation with the vector* $\mathbf{v}$ is the map $T_{\mathbf{v}} : \mathbb{E}^n \to \mathbb{E}^n$ given by $T_{\mathbf{v}}(\mathbf{x}) = \mathbf{x} + \mathbf{v}$. It is just the translation map of the affine structure of $\mathbb{E}^n$ where we fixed the vector argument to be $\mathbf{v}$.

Proposition 7.4. 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 frame. The following are equivalent:

1. ϕ is an isometry
2. ϕ preserves lengths of vectors
3. ϕ preserves the scalar product, i.e. for any vectors $\mathbf{v}, \mathbf{w}$ we have
$$\langle \phi(\mathbf{v}), \phi(\mathbf{w}) \rangle = \langle \mathbf{v}, \mathbf{w} \rangle.$$

Proof. The length of a vector $\overrightarrow{AB}$ is by definition the length of the segment $[AB]$ which is the distance between A and B, so 1. is equivalent to 2. If ϕ is an isometry, it is an affine map, by Proposition 7.2. In particular it preserves oriented ratios by Proposition 6.5. From this it follows that it preserves the cosine of angles. Then
$$\langle \phi(\mathbf{v}), \phi(\mathbf{w}) \rangle = |\phi(\mathbf{v})| \cdot |\phi(\mathbf{w})| \cdot \cos \measuredangle(\phi(\mathbf{v}), \phi(\mathbf{w})) = |\mathbf{v}| \cdot |\mathbf{w}| \cdot \cos \measuredangle(\mathbf{v}, \mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle.$$
For the converse, let $\mathbf{v} = \overrightarrow{AB}$ and notice that if ϕ preserves the scalar product then
$$d(\phi(A), \phi(B)) = |\phi(\overrightarrow{AB})| = \sqrt{\langle \phi(\overrightarrow{AB}), \phi(\overrightarrow{AB}) \rangle} = \sqrt{\langle \overrightarrow{AB}, \overrightarrow{AB} \rangle} = |\overrightarrow{AB}| = d(A, B).$$
This shows that 3. implies 1. and 2. $\square$

Proposition 7.5. 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 frame. The following are equivalent:

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

Proof. Consider $\psi(\mathbf{x}) = \phi(\mathbf{x}) - b = A\mathbf{x}$. Since translations are isometries, ϕ is an isometry if and only if ψ is an isometry. Let $\mathcal{K} = (O, \mathcal{B})$ with $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be the orthonormal frame with respect to which ϕ is given in the form $\phi(\mathbf{x}) = A\mathbf{x} + b$. Let $\mathbf{f}_1 = \psi(\mathbf{e}_1) = A\mathbf{e}_1$, $\mathbf{f}_2 = \psi(\mathbf{e}_2) = A\mathbf{e}_2, \ldots, \mathbf{f}_n = \psi(\mathbf{e}_n) = A\mathbf{e}_n$.

To see that 1. implies 2., notice that if ψ is an isometry, by Proposition 7.4 we have
$$\langle \mathbf{f}_i, \mathbf{f}_j \rangle = \langle \psi(\mathbf{e}_i), \psi(\mathbf{e}_j) \rangle = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$
Since the components of $\mathbf{f}_i = A\mathbf{e}_i$ are the entries in the i-th column of the matrix A, it follows that $A^T A$ is the identity matrix I_n.

To see that 2. implies 1. consider two arbitrary points $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{E}^n$. We have $|\phi(\mathbf{x}) - \phi(\mathbf{y})| = |(A\mathbf{x} + b) - (A\mathbf{y} + b)| = |A(\mathbf{x} - \mathbf{y})|$ and therefore
$$\begin{aligned} |\phi(\mathbf{x}) - \phi(\mathbf{y})|^2 &= |A(\mathbf{x} - \mathbf{y})|^2 \\ &= \langle A(\mathbf{x} - \mathbf{y}), A(\mathbf{x} - \mathbf{y}) \rangle \\ &= (A(\mathbf{x} - \mathbf{y}))^T \cdot A(\mathbf{x} - \mathbf{y}) \\ &= (\mathbf{x} - \mathbf{y})^T A^T A(\mathbf{x} - \mathbf{y}) \\ &= (\mathbf{x} - \mathbf{y})^T (\mathbf{x} - \mathbf{y}) \\ &= |\mathbf{x} - \mathbf{y}|^2 \end{aligned}$$
which is equivalent to $d(\phi(\mathbf{x}), \phi(\mathbf{y})) = d(\mathbf{x}, \mathbf{y})$. $\square$

Definition 7.6. A matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ such that $A^T A = I_n$ is called *orthogonal* and the set of all such matrices is denoted by $\mathrm{O}(n)$.

Proposition 7.7. If A is an orthogonal matrix then $\det(A) = \pm 1$.

Proof. Since $A^T A = I_n$ we have $1 = \det(I_n) = \det(A^T A) = \det(A^T)\det(A) = \det(A)^2$. □

Definition 7.8. The set of matrices in $\mathrm{O}(n)$ with determinant 1 is denoted by $\mathrm{SO}(n)$. Such matrices are called *special orthogonal*. The set $\mathrm{O}(n)$ is a subgroup of $\mathrm{AGL}(\mathbb{A}^n)$ and $\mathrm{SO}(n)$ is a normal subgroup of $\mathrm{O}(n)$. With group theory notation we write this as follows:

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

Definition 7.9. Let ϕ be an isometry of $\mathbb{E}^n$ given by $\phi(\mathbf{x}) = A\mathbf{x} + b$ with respect to some right-oriented orthonormal frame. Then ϕ is called a *displacement*, or a *direct isometry*, if $A \in \mathrm{SO}(n)$. Else, if $\det(A) = -1$, the map ϕ is called an *indirect isometry*.

7.1.1 The Euclidean space $\mathbb{E}^n$ (second revision)

In Section 4.2.1 we have revised $\mathbb{E}^n$ to be $\mathbb{R}^n$ with the standard scalar product. This point of view encapsulates all aspects of the Hilbert's axioms (Appendix A) but doesn't immediately give a transparent description of the congruence relation. This description can be done now: a segment $[AB]$ is congruent to a segment $[CD]$ if and only if there is an isometry ϕ such that $\phi(A) = C$ and $\phi(B) = D$. This shows that the notion of congruence (in dimension n) is completely described by translations and by the orthogonal group $\mathrm{O}(n)$.

It is tautological to say that congruence is described by isometries. The point made here is that isometries can be classified and described precisely by focusing the analysis on $\mathrm{O}(n)$. We do this for dimension 2 and 3 in the following two sections.

7.2 Isometries in dimension 2

7.2.1 Rotations

Proposition 7.10. A matrix A is in $\mathrm{SO}(2)$ if and only if A equals

$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \tag{7.1}$$

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

Proof. It is easy to check that R_θ is special orthogonal. For the converse consider the matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SO}(2).$$

We have $\det(A) = ad - bc = 1$ and

$$A^T A = \begin{bmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus, the entries of the matrix A satisfy the system

$$\begin{cases} a^2 + c^2 = 1 \\ b^2 + d^2 = 1 \\ ab + cd = 0 \\ ad - bc = 1 \end{cases}.$$

From the first two equations it follows that there exist $\theta, \tilde{\theta}$ such that $a = \cos(\theta)$, $c = \sin(\theta)$, $d = \cos(\tilde{\theta})$ and $b = \sin(\tilde{\theta})$. Then, the last two equations become

$$\begin{cases} 0 = ab + cd = \cos(\theta)\sin(\tilde{\theta}) + \sin(\theta)\cos(\tilde{\theta}) = \sin(\theta + \tilde{\theta}) \\ 0 = ad - bc = \cos(\theta)\cos(\tilde{\theta}) - \sin(\theta)\sin(\tilde{\theta}) = \cos(\theta + \tilde{\theta}) \end{cases}$$

and therefore $\tilde{\theta} = -\theta + 2k\pi$ for some integer k. □

Corollary 7.11. 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

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

Proof. Let $\phi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ with respect to a right oriented orthonormal frame. By Proposition 7.10, it is a direct isometry if $A \in \mathrm{SO}(2)$ which is equivalent to $A = R_\theta$ where R_θ is the rotation matrix (7.1). If $\theta = 0$ then ϕ is a (possibly trivial) translation by b thus, it has a fixed point only if $\mathbf{b} = 0$ in which case ϕ equals the identity map, i.e. $\phi = \mathrm{Id}$. For the rest of the proof assume that $\theta \neq 0$.

Let $\mathbf{p}$ be a fixed point for ϕ. Then

$$\phi(\mathbf{p}) = \mathbf{p} \quad \Leftrightarrow \quad R_\theta \mathbf{p} + \mathbf{b} = \mathbf{p} \quad \Leftrightarrow \quad \mathbf{b} = (I_2 - R_\theta)\mathbf{p}. \tag{7.2}$$

Now notice that

$$\det(I_2 - R_\theta) = \det \begin{bmatrix} 1 - \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & 1 - \cos(\theta) \end{bmatrix} = (1 - \cos(\theta))^2 - \sin(\theta) = 2 - 2\cos(\theta).$$

So, if $\theta \neq 0$ then equation (7.2) has a unique solution $\mathbf{p}$, i.e. the fixed point of ϕ is unique. To finish the proof we show that ϕ is a rotation around the point $\mathbf{p}$. We notice that

$$\phi(\mathbf{x}) - \phi(\mathbf{p}) = \phi(\mathbf{x}) - \mathbf{p} = R_\theta \mathbf{x} + \mathbf{b} - \mathbf{p} = R_\theta \mathbf{x} + (I_2 - R_\theta)\mathbf{p} - \mathbf{p} = R_\theta(\mathbf{x} - \mathbf{p})$$

which means that ϕ rotates $\overrightarrow{\mathbf{px}}$ with θ around $\mathbf{p}$. Finally, the trace of the liner map $\operatorname{lin}(\phi)$ is the trace of its matrix with respect to any orthonormal basis. Thus, we may use $A = R_\theta$ to conclude that

$$\operatorname{tr}(\operatorname{lin}(\phi)) = \operatorname{tr}(R_\theta) = 2\cos(\theta).$$

□

Remark. Let us notice the effect of rotations on coordinates and on basis vectors in dimension 2. Let $\mathcal{K} = (O,\mathcal{B})$ be a right-oriented orthonormal 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 vectors in $\mathcal{B}'$ are obtained by rotating the vectors in $\mathcal{B}$ by $-\theta$. Indeed rotating points counterclockwise with respect to $\mathcal{K}$ is equivalent to rotating $\mathcal{K}$ clockwise.

7.2.2 Classification

Theorem 7.12 (Chasles)**.** A direct isometry of the plane $\mathbb{E}^2$ is either

a) the identity, or

b) a translation, or

c) a rotation.

Proof. Consider a direct isometry $\phi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ given with respect to a right-oriented orthonormal frame. By Proposition 7.10 $A = R_\theta$ where R_θ is the rotation matrix (7.1) for some θ. If $\theta = 0$ then ϕ the identity map if $\mathbf{b} = 0$ and it is a translation if $\mathbf{b} \neq 0$. Suppose therefore that $\theta \neq 0$. Then, by the proof of Corollary 7.11 the map ϕ has a (unique) fixed point $\mathbf{p}$. Thus, ϕ is a direct isometry which fixes a point and the proof is finished by applying Corollary 7.11. □

Definition 7.13. A *glide reflection in the line* ℓ is the composition of an orthogonal reflection in ℓ and a translation in the direction of ℓ.

Example 7.14. A reflection in the x-axis followed by a translation with $\lambda\mathbf{i}$ is a glide rotation. In matrix form the map is given by

$$\mathbf{x} \mapsto \mathrm{Ref}^{\perp}_{Ox}(\mathbf{x}) + \lambda\mathbf{i} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \lambda \\ 0 \end{bmatrix}.$$

Lemma 7.15. A rotation with angle θ around the origin after a reflection in the x-axis is a reflection in the line $y = \tan(\theta/2)x$.

Proof. The mentioned map, the reflection in the x-axis followed by a rotatation in the origin, has the form $\phi : \mathbf{x} \mapsto A\mathbf{x}$ where

$$A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Notice that $\det(A - I_2) = 0$ thus ϕ fixes at least a line. Consider the vector $\mathbf{v}(\cos(\frac{\theta}{2}), \sin(\frac{\theta}{2}))$. A calculation shows that $(A - I_2)\mathbf{v} = 0$ which means that $\mathbf{v}$ is fixed by A. Hence the line passing through the origin in the direction of $\mathbf{v}$ is fixed by ϕ. To see that it is a reflection, check that $\mathbf{J}(\mathbf{v})$ is mapped to $-\mathbf{J}(\mathbf{v})$. □

Theorem 7.16. An indirect isometry of the plane $\mathbb{E}^2$ fixes a line ℓ and is either

a) a reflection in ℓ, or

b) a glide-reflection in ℓ.

Proof. Let $\phi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ be an indirect isometry given with respect to a right-oriented orthonormal basis. Consider the reflection $\psi(\mathbf{x}) = \mathrm{Ref}^{\perp}_{Ox}(\mathbf{x}) = C\mathbf{x}$, which is also an indirect isometry. Since $(\phi \circ \psi)(\mathbf{x}) = AC\mathbf{x} + \mathbf{b}$ and since $\det(A) = \det(C) = -1$ we see that AC is an orthogonal matrix of determinant 1, i.e. $\phi \circ \psi$ is a direct isometry. Then by Theorem 7.12, we have 3 cases.

If $\phi \circ \psi$ is the idenity or a translation, then $AC = I_2$ which implies $A = C^{-1}$. Since C is the matrix of a reflection we have $C^2 = I_n$, hence $A = C$. Therefore ϕ is the composition of $\mathrm{Ref}^{\perp}_{Ox}$ followed by a translation with $\mathbf{b}(b_1, b_2)$. Then

$$\phi(\mathbf{x}) = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{x} + b_2\mathbf{j}}_{\psi'(\mathbf{x})} + b_2\mathbf{i}$$

and one checks with (for example with (6.7)) that ψ' is a reflection in the line $y = b_1/2$. Hence ϕ is the composition of a reflection and a (possibly trivial) translation along the reflection axis. Thus, if $b_1 = 0$ it is a reflection and if $b_1 \neq 0$ it is a glied reflection.

If $\phi \circ \psi$ is a rotation then $AC = R_\theta$ for some θ. In this case we have $A = R_\theta C$. By Lemma 7.15, ϕ is a reflection in a line ℓ followed by a translation and the claim follows as in the previous paragraph. □

7.3 Isometries in dimension 3

7.3.1 Rotations

Theorem 7.17 (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

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

Proof. By choosing the fixed point of ϕ to be the origin, we may assume that ϕ has the form $\phi(\mathbf{x}) = A\mathbf{x}$ with respect to a right-oriented orthonormal frame. Since ϕ is a direct isometry, we have $A \in \mathrm{SO}(3)$. A rotation around an axis fixes the rotation axis. To see that this is the case for ϕ, it suffices to show that A has an eigenvector $\mathbf{v}$ for the eigenvalue 1 since then

$$\phi(t\mathbf{v}) = A(t\mathbf{v}) = t(A\mathbf{v}) = t\mathbf{v} \quad \forall t \in \mathbb{R}$$

which means that ϕ fixes the line passing through the origin in the direction of $\mathbf{v}$. Notice that

$$\begin{aligned}
\det(A - I_3) &= \det(A^T)\det(A - I_3) && \text{since } \det(A) = 1\\
&= \det(A^T(A - I_3))\\
&= \det(A^T A - A^T)\\
&= \det(I_3 - A^T) && \text{since } A \text{ is orthogonal}\\
&= \det((I_3 - A)^T)\\
&= \det(I_3 - A)
\end{aligned}$$

and since $A - I_3$ is a 3×3 matrix we have $\det(I_3 - A) = -\det(A - I_3)$. Thus

$$\det(A - I_3) = -\det(A - I_3) \quad \Rightarrow \quad \det(A - I_3) = 0.$$

Thus, A admits 1 as eigenvalue. Let $\mathbf{v}$ be a corresponding eigenvector of length 1.

Next we show that ϕ is a rotation around the axis $\mathbb{R}\mathbf{v}$. Choose a unit vector $\mathbf{u}_1 \perp \mathbf{v}$ and let $\mathbf{u}_2 = \mathbf{v} \times \mathbf{u}_1$. Then $A\mathbf{u}_1$ and $A\mathbf{u}_2$ are also orthogonal to $\mathbf{v}$ since

$$\langle A\mathbf{u}_i, \mathbf{v}\rangle = \langle A\mathbf{u}_i, A\mathbf{v}\rangle = \langle \mathbf{u}_i, \mathbf{v}\rangle = 0.$$

In fact $(\mathbf{u}_1, \mathbf{u}_2)$ is a basis for $\mathbf{v}^\perp$, the orthogonal complement to $\mathbf{v}$. Since ϕ is an isometry it maps $\mathbf{v}^\perp$ to itself. In particular, restricting ϕ to $\mathbf{v}^\perp$ we have an isometry in dimension 2. Thus, with respect to the basis $\mathcal{B} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{v})$ we have

$$A = \begin{bmatrix} B & 0 \\ 0 & 1 \end{bmatrix}$$

where B is a 2×2 matrix. Since $\mathcal{B}$ is right-oriented and ϕ is a direct isometry, we have $\det(A) = 1$. Therefore $\det(B) = 1$. Hence ϕ restricts to a direct isometry on $\mathbf{v}^\perp$. Since it fixes the origin, it cannot be a translation. Therefore, by Theorem 7.12, it must be a (possibly trivial) rotation. Then, the matrix of ϕ with respect to $\mathcal{B}$ is

$$[\mathrm{lin}(\phi)]_{\mathcal{B}} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

for some $\theta \in \mathbb{R}$. This is a rotation around the axis $\mathbb{R}\mathbf{v}$. Finally, the trace of the linear map $\mathrm{lin}(\phi)$ is the trace of the matrix $[\mathrm{lin}(\phi)]_{\mathcal{B}}$. Thus

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

□

Proposition 7.18 (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

$$\mathrm{Rot}_{\mathbf{v},\theta}(\mathbf{x}) = \cos(\theta)\mathbf{x} + \sin(\theta)(\mathbf{v}\times\mathbf{x}) + (1-\cos(\theta))\langle \mathbf{v},\mathbf{x}\rangle\mathbf{v}. \tag{7.3}$$

Proof. Let $\mathbf{x} = \mathbf{x}_{\parallel} + \mathbf{x}_{\perp}$ where $x_{\parallel}$ is parallel to $\mathbf{v}$ and $\mathbf{x}_{\perp}$ is orthogonal to $\mathbf{v}$. Since $\mathbf{v}$ is a unit vector, we know that

$$\mathbf{x}_{\parallel} = \langle \mathbf{v},\mathbf{x}\rangle\mathbf{v} \quad \text{and} \quad \mathbf{x}_{\perp} = \mathbf{x} - \mathbf{x}_{\parallel}.$$

Let $\mathbf{x}'_{\perp} = \mathbf{v}\times\mathbf{x}$ and notice that

$$\mathrm{Rot}_{\mathbf{v},\theta}(\mathbf{x}) = \mathbf{x}_{\parallel} + \cos(\theta)\mathbf{x}_{\perp} + \sin(\theta)\mathbf{x}'_{\perp}. \tag{7.4}$$

Moreover, the vector $\mathbf{x}'_{\perp}$ is orthogonal to $\mathbf{x}_{\parallel}$ and $\mathbf{x}_{\perp}$ and has the same length as $\mathbf{x}_{\perp}$ since

$$\begin{aligned} |\mathbf{x}'_{\perp}|^2 &= \langle \mathbf{v}\times\mathbf{x}, \mathbf{v}\times\mathbf{x}\rangle \\ &= \langle \mathbf{v},\mathbf{v}\rangle\langle \mathbf{x},\mathbf{x}\rangle - \langle \mathbf{v},\mathbf{x}\rangle\langle \mathbf{v},\mathbf{x}\rangle \\ &= |\mathbf{x}|^2 - |\mathbf{x}_{\parallel}|^2 \\ &= |\mathbf{x}_{\perp}| \end{aligned}$$

where for the second equality we used Lagrange's identity. Therefore, formula (7.4) shows that $\mathrm{Rot}_{\mathbf{v},\theta}$ rotates $\mathbf{x}_{\perp}$ by θ around the axis $\mathbb{R}\mathbf{v}$. □

7.3.2 Euler angles

Fix a right oriented orthonormal 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 $\mathrm{Rot}_{z,\gamma}\mathrm{Rot}_{x,\beta}\mathrm{Rot}_{z,\alpha} =$

$$\begin{bmatrix} \cos\gamma\cos\alpha - \sin\gamma\cos\beta\sin\alpha & -\cos\gamma\sin\alpha - \sin\gamma\cos\beta\cos\alpha & \sin\gamma\sin\beta \\ \sin\gamma\cos\alpha + \cos\gamma\cos\beta\sin\alpha & -\sin\gamma\sin\alpha + \sin\gamma\cos\beta\cos\alpha & -\cos\gamma\sin\beta \\ \sin\beta\sin\alpha & \sin\beta\cos\alpha & \cos\beta \end{bmatrix}.$$

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.

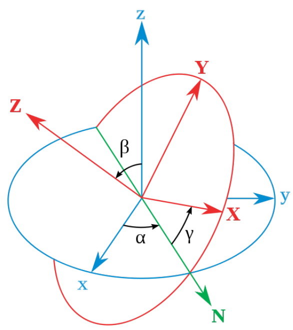


Figure 7.1: Euler angles[1]

Proposition 7.19. 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 7.20. 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 K).

Remark. Euler angles give coordinates on SO(3). Not to be confused with spherical coordinates (See Appendix D).

7.3.3 Classification

Definition 7.21. 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 7.22 (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.

Proof. Let ϕ be the direct isometry given by $\phi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ with respect to a right-oriented frame. We know that $A \in \mathrm{SO}(3)$. If $A = I_3$, then ϕ is a translation or the identity. Suppose that $A \neq I_3$. Applying Theorem 7.17 to $\mathbf{x} \mapsto A\mathbf{x}$ we see that A is a rotation matrix. Let $\mathbf{v}$ be a direction vector of length 1 for the rotation axis of A and decompose the vector $\mathbf{b}$ into its components parallel to $\mathbf{v}$ and orthogonal to $\mathbf{v}$:

$$\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2 \quad \text{where} \quad \mathbf{b}_1 \| \mathbf{v} \quad \text{and} \quad \mathbf{b}_2 \perp \mathbf{v}.$$

[1] Image source: Wikipedia

We have $\mathbf{b}_1 = \langle \mathbf{v}, \mathbf{b} \rangle \mathbf{v}$ and $\mathbf{b}_2 = (\mathbf{v} \times \mathbf{b}) \times \mathbf{v}$. Now consider the isometries

$$\phi_1(\mathbf{x}) = \mathbf{x} + \mathbf{b}_1 \quad \text{and} \quad \phi_2(\mathbf{x}) = A\mathbf{x} + \mathbf{b}_2.$$

Let π be the plane passing through the origin and orthogonal to the rotation axis. Then ϕ_2 is an isomtry which leaves π invariant ($\phi_2(\pi) = \pi$). By Chasles' Theorem in dimension 2 (Theorem 7.12), the restriction of ϕ_2 to π is a rotation around a fixed point $\mathbf{p} \in \pi$. Thus, ϕ_2 is a rotation around the axis $\mathbf{p} + \mathbb{R}\mathbf{v}$. On the other hand, ϕ_1 is a translation by $\mathbf{b}_1$ parallel to the rotation axis. This finishes the proof since $\phi = \phi_1 \circ \phi_2$. □

Theorem 7.23. 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*.

Proof. Let ϕ be the direct isometry given by $\phi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ with respect to a right-oriented frame. One can show that a matrix $A \in O(3)$ of determinant -1 admits -1 as an eigenvalue. Let $\mathbf{v}$ be an eigenvector for the eigenvalue -1, i.e. $A\mathbf{v} = -\mathbf{v}$. Notice that we also have $A^T\mathbf{v} = A^{-1}\mathbf{v} = -\mathbf{v}$. Calculating we obtain

$$\begin{aligned} \langle \mathbf{v}, \phi(\mathbf{x}) \rangle &= \langle \mathbf{v}, A\mathbf{x} + \mathbf{b} \rangle \\ &= \langle \mathbf{v}, A\mathbf{x} \rangle + \langle \mathbf{v}, \mathbf{b} \rangle \\ &= \langle A^T\mathbf{v}, \mathbf{x} \rangle + \langle \mathbf{v}, \mathbf{b} \rangle \\ &= \langle -\mathbf{v}, \mathbf{x} \rangle + \langle \mathbf{v}, \mathbf{b} \rangle \end{aligned}$$

Thus, the plane

$$\pi : \langle \mathbf{v}, \mathbf{x} \rangle = \frac{1}{2} \langle \mathbf{v}, \mathbf{b} \rangle$$

is invariant under the isometry ϕ. Moreover, if we choose the frame with the first basis vectors parallel to π then

$$A = \begin{bmatrix} B & 0 \\ 0 & -1 \end{bmatrix}$$

and $\det(A) = \det(B)$. Thus, the restriction of ϕ to the plane π is a direct isometry. Therefore, by Theorem 7.12, it is either the identity, or a translation or a rotation, which correspond to the three cases stated in the theorem. □

7.4 Moving points with isometries

7.4.1 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

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) & -\cos(\theta)+1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sin(\theta) \\ -\cos(\theta)+1 \\ 1 \end{bmatrix}$$

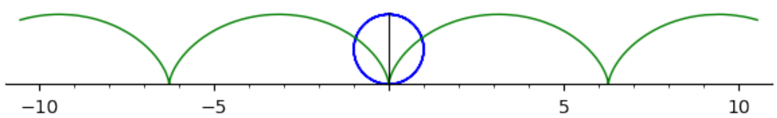
since the homogeneous coordinates of the origin are $(0,0,1)$. If you 'vary θ with time t' you are rotating the origin around C counterclockwise. If you want to have a clockwise motion you just change the sign of the angle to obtain

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} -\sin(t) \\ -\cos(t)+1 \\ 1 \end{bmatrix}.$$

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

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) & -x_C\cos(\theta)+y_C\sin(\theta)+x_C \\ \sin(\theta) & \cos(\theta) & -x_C\sin(\theta)-y_C\cos(\theta)+y_C \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & -4\cos(\theta)+4 \\ \sin(\theta) & \cos(\theta) & -4\sin(\theta) \\ 0 & 0 & 1 \end{bmatrix}$$

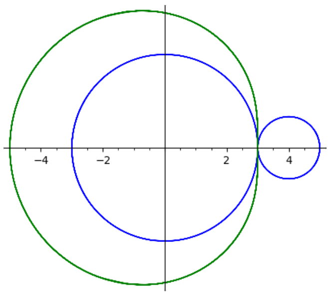
and thus, rotating $P(2,0)$ with time t we obtain

$$\begin{bmatrix} \cos(t) & -\sin(t) & -4\cos(t)+4 \\ \sin(t) & \cos(t) & -4\sin(t) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3\cos(t)-4\cos(t)+4 \\ 3\sin(t)-4\sin(t) \\ 1 \end{bmatrix}$$

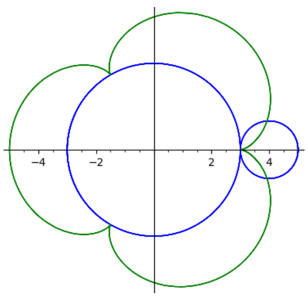
If at the same time we rotate around the origin, P will move on the small circle which rotates around a big circle of radius 3 centered at the origin:

$$\begin{bmatrix} \cos(t') & -\sin(t') & 0 \\ \sin(t') & \cos(t') & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3\cos(t)-4\cos(t)+4 \\ 3\sin(t)-4\sin(t) \\ 1 \end{bmatrix} = \dots$$

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.

7.4.2 Surfaces of revolution

Another way of looking at Euler angles $\mathrm{Rot}_{z,\gamma}\,\mathrm{Rot}_{x,\beta}\,\mathrm{Rot}_{z,\alpha}$ (see Section 7.3.2) 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

$$[0,2\pi[\times[0,\pi[\ni(\alpha,\beta)\mapsto\begin{bmatrix}\cos\alpha\\ \cos\beta\sin\alpha\\ \sin\beta\sin\alpha\end{bmatrix},\tag{7.5}$$

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.

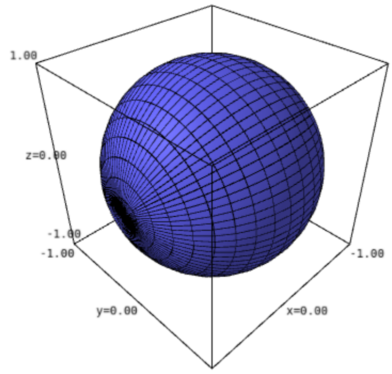
From a different perspective, notice that a plane in $\mathbb{E}^3$ can be described as the set of points which you touch if you translate a line in a given direction. This can be seen with the parametric equations:

$$\pi:\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x_Q\\y_Q\\z_Q\end{bmatrix}+s\begin{bmatrix}v_x\\v_y\\v_z\end{bmatrix}+t\begin{bmatrix}w_x\\w_y\\w_z\end{bmatrix}=\underbrace{\left(\begin{bmatrix}x_Q\\y_Q\\z_Q\end{bmatrix}+s\begin{bmatrix}v_x\\v_y\\v_z\end{bmatrix}\right)}_{\text{line }\ell}\quad\underbrace{+t\begin{bmatrix}w_x\\w_y\\w_z\end{bmatrix}}_{\text{translation with }t\mathbf{w}}\tag{7.6}$$

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 (7.5) which you can rewrite as follows

$$(\alpha,\beta)\mapsto\underbrace{\begin{bmatrix}1&0&0\\0&\cos\beta&-\sin\beta\\0&\sin\beta&\cos\beta\end{bmatrix}}_{\text{rotation }\mathrm{Rot}_{x,\beta}}\underbrace{\begin{bmatrix}\cos\alpha\\ \sin\alpha\\ 0\end{bmatrix}}_{\text{circle in }Oxy\text{-plane}}=\begin{bmatrix}\cos\alpha\\ \cos\beta\sin\alpha\\ \sin\beta\sin\alpha\end{bmatrix}.$$

This describes the unit sphere centered at the origin as the set of 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.



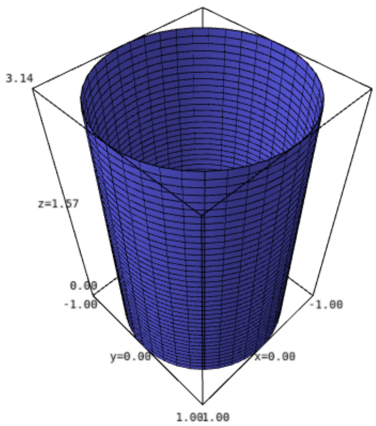
Definition 7.24. A *surface of revolution* in $\mathbb{E}^3$ is a surface obtained by rotating a curve around a line ℓ. The line ℓ is called the *axis* of the surface.

Example 7.25. In (7.6), 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:

$$(\theta, s) \mapsto \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} \cos\theta \\ \sin\theta \\ s \end{bmatrix}$$



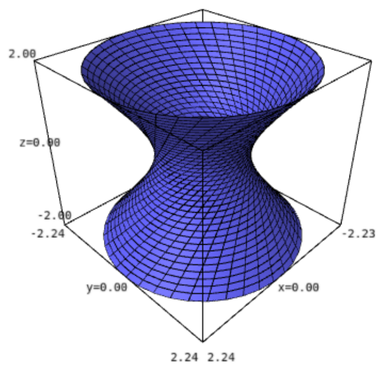
Example 7.26. If instead we consider two skew lines and rotate one around the other, we obtain

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

$$(\theta, s) \mapsto \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right) = \begin{bmatrix} \cos\theta - s\sin\theta \\ \sin\theta + s\cos\theta \\ s \end{bmatrix}.$$



APPENDIX A

Axioms

'That it was the Greeks who added the element of logical structure to geometry is virtually universally admitted today' [6, p.47]. According to Aristotle[1], for Thales[2] 'the primary question was not *what do we know*, but *how we know it*' [6, Chapter 4]. The ancient Greeks revolutionized mathematics through their systematic use of ordered sequences of logical deductions. In Euclid's[3] *Elements* these deductions are grounded in what are considered self-evident truths – definitions and postulates.

A major revision of the fundamental assumptions underlying Euclidean geometry was carried out by Hilbert[4] in his *Foundations of Geometry* [14]. The axioms Hilbert proposed as the basis for Euclidean geometry are presented below. Hilbert showed that these axioms are independent. They are necessary and sufficient assumptions for describing the interactions between primitives (*points*, *lines*, and *planes*). The purpose of any axiomatic treatment is to put all statements on a solid ground, the axioms from which they can be deduced. As Blumenthal recounts, Hilbert remarked that 'one must be able to say "tables, chairs, beer-mugs" each time in place of "points, lines, planes" ' [12, p.208].

Remark. We use the notation $[AB]$ for 'the segment AB'. Moreover, Axiom III.5 does not appear in [14] but it is needed for the transitivity of the congruence relation on angles.

Axiom Group I: Axioms of Incidence

I.1 For every two points A, B there exists a line a that contains each of the points A, B.

I.2 For every two points A, B there exists no more than one line that contains each of the points A, B.

[1] 384–322 BC
[2] c.626/623 – c.548/545 BC
[3] ∼ 300 BC
[4] 1862 – 1943

I.3 There exist at least two points on a line. There exist at least three points that do not lie on a line.

I.4 For any three points A, B, C that do not lie on the same line there exists a plane α that contains each of the points A, B, C. For every plane there exists a point which it contains.

I.5 For any three points A, B, C that do not lie on one and the same line there exists no more than one plane that contains each of the three points A, B, C.

I.6 If two points A, B of a line a lie in a plane α then every point of a lies in the plane α.

I.7 If two planes α, β have a point A in common then they have at least one more point B in common.

I.8 There exist at least four points which do not lie in a plane.

Axiom Group II: Axioms of Order

II.1 If a point B lies between a point A and a point C then the points A, B, C are three distinct points of a line, and B then also lies between C and A.

II.2 For two points A and C, there always exists at least one point B on the line AC such that C lies between A and B.

II.3 Of any three points on a line there exists no more than one that lies between the other two.

II.4 (Pasch's Axiom) Let A, B, C be three points that do not lie on a line and let a be a line in the plane ABC which does not meet any of the points A, B, C. If the line a passes through a point of the segment $[AB]$, it also passes through a point of the segment $[AC]$, or through a point of the segment $[BC]$.

Axiom Group III: Axioms of Congruence

III.1 If A, B are two points on a line a, and A' is a point on the same or on another line a' then it is always possible to find a point B' on a given side of the line a' through A' such that the segment $[AB]$ is congruent or equal to the segment $[A'B']$. In symbols $[AB] \equiv [A'B']$.

III.2 If a segment $[A'B']$ and a segment $[A''B'']$ are congruent to the same segment $[AB]$, then the segment $[A'B']$ is also congruent to the segment $[A''B'']$, or briefly, if two segments are congruent to a third one they are congruent to each other.

III.3 On the line a let $[AB]$ and $[BC]$ be two segments which except for B have no point in common. Furthermore, on the same or on another line a' let $[A'B']$ and $[B'C']$ be two segments which except for B' also have no point in common. In that case, if $[AB] \equiv [A'B']$ and $[BC] \equiv [B'C']$ then $[AC] \equiv [A'C']$.

III.4 Let $\measuredangle(h,k)$ be an angle in a plane α and a' a line in a plane α' and let a definite side of a' of α' be given. Let h' be a ray on the line a' that emanates from the point O'. Then there exists in the plane α' one and only one ray k' such that the angle $\measuredangle(h,k)$ is congruent or equal to the angle $\measuredangle(h',k')$ and at the same time all interior points of the angle $\measuredangle(h',k')$ lie on the given side of a'. Symbolically $\measuredangle(h,k) \equiv \measuredangle(h',k')$. Every angle is congruent to itself, i.e., $\measuredangle(h,k) \equiv \measuredangle(h,k)$ is always true.

III.5 If an angle $\measuredangle(h',k')$ and an angle $\measuredangle(h'',k'')$ are congruent to the same angle $\measuredangle(h,k)$, then the angle $\measuredangle(h',k')$ is also congruent to the angle $\measuredangle(h'',k'')$, or briefly, if two angles are congruent to a third one they are congruent to each other.

III.6 If for two triangles ABC and $A'B'C'$ the congruences $AB \equiv A'B'$, $AC \equiv A'C'$, $\measuredangle BAC \equiv \measuredangle B'A'C'$ hold, then the congruence $\measuredangle ABC \equiv \measuredangle A'B'C'$ is also satisfied.

Axiom Group IV: Axiom of Parallels

IV (Euclid's Axiom). Let a be any line and A a point not on it. Then there is at most one line in the plane, determined by a and A, that passes through A and does not intersect a.

Axiom Group V: Axioms of Continuity

V.1 (Axiom of measure or Archimedes' Axiom). If $[AB]$ and $[CD]$ are any segments then there exists a number n such that n segments $[CD]$ constructed contiguously from A, along the ray from A through B, will pass beyond the point B.

V.2 (Axiom of line completeness). An extension of a set of points on a line with its order and congruence relations that would preserve the relations existing among the original elements as well as the fundamental properties of line order and congruence that follows from Axioms I-III, and from V.1 is impossible.

APPENDIX B

Lines and the real numbers

We are used to thinking about lengths of segments as positive real numbers $\mathbb{R}_{\geq 0}$. We are also used to drawing the field of real numbers $\mathbb{R}$ as a line with 0 represented by a point which separates positive and negative numbers. In this section we describe a path which connects Hilbert's Axioms with real numbers using vectors. The zero vector corresponds to $0 \in \mathbb{R}$, but $\mathbb{R}$ also has an identity element 1 which needs a correspondent. For this, we need to make a choice.

Definition B.1. A *unit segment* is a non-trivial segment in $\mathbb{E}$ which we choose. Once a unit segment $[AB]$ is chosen the length $|AB|$ is called the *unit length* and, equivalently the distance from A to B is called *unit distance* (see Definition 1.3). Then, vectors having unit length are called *unit vectors* or *versors*.

Definition B.2. For two points A and B, we denote by $\mathbb{F}_{AB}$ the set of vectors represented with points on the line AB. Observe that we have a partition

$$\mathbb{F}_{AB} = \underbrace{\left\{\overrightarrow{XY} : X, Y \in AB \text{ and } |XY\rangle = |AB\rangle\right\}}_{\mathbb{F}^+_{AB}} \cup \left\{\overrightarrow{0}\right\} \cup \underbrace{\left\{\overrightarrow{XY} : X, Y \in AB \text{ and } |XY\rangle = -|AB\rangle\right\}}_{\mathbb{F}^-_{AB}}.$$

Proposition B.3. Let O be a point on a line AB. The following maps are bijections:

1. The map $\phi_O : AB \to \mathbb{F}_{AB}$ defined by $\phi_O(P) = \overrightarrow{OP}$,
2. The map $\psi_+ : \mathbb{F}^+_{AB} \cup \{\overrightarrow{0}\} \to \mathbb{L}$ defined by $\psi_+(\overrightarrow{XY}) = |\overrightarrow{XY}|$,
3. The map $\psi_- : \mathbb{F}^-_{AB} \cup \{\overrightarrow{0}\} \to \mathbb{L}$ defined by $\psi_-(\overrightarrow{XY}) = |\overrightarrow{XY}|$.

Sketch of proof. For 1., if $\overrightarrow{OP} = \overrightarrow{OQ}$ then $P = Q$ by Lemma 1.4. For 2. and 3. let X, Y be two distinct points on the line AB. By Lemma 1.4, there are exactly two points Z, Z' with O between them such that the segment $[XY]$ is congruent to $[OZ]$ and to $[Z'O]$. Then $\overrightarrow{OZ} = -\overrightarrow{OZ'}$ and exactly one of them will have direction $|AB\rangle$. $\square$

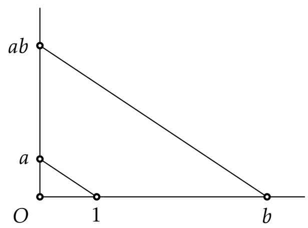
Definition B.4. Under the bijections in Proposition B.3 we identify $\mathbb{F}^+_{AB} \cup \{0\}$ with $\mathbb{L}$ and we call these elements *positive*. The elements in $\mathbb{F}^+_{AB}$ are called *strictly positive*. We call the elements in $\mathbb{F}^-_{AB} \cup \{0\}$ *negative* and the ones in $\mathbb{F}^-_{AB}$ are called *strictly negative*. Notice that the involution in Definition 1.13 restricts to an involution $-\square : \mathbb{F}_{AB} \to \mathbb{F}_{AB}$ which interchanges positive and negative elements.

Definition B.5. We define an ordering on $\mathbb{F}_{AB}$. If $\overrightarrow{AC}, \overrightarrow{AD}$ are positive, we let $\overrightarrow{AC} \leq \overrightarrow{AD}$ if and only if $[AC] \subseteq [AD]$. If $\overrightarrow{AC}, \overrightarrow{AD}$ are negative, we let $\overrightarrow{AC} \leq \overrightarrow{AD}$ if and only if $[AD] \subseteq [AC]$.

Proposition B.6. The ordering in Definition B.5 is a total order. With this ordering and with addition of vectors, $\mathbb{F}_{AB}$ is an abelian totally ordered group and $\mathbb{L}$ is an ordered submonoid of $\mathbb{F}_{AB}$.

Sketch of proof. Since $\mathbb{F}_{AB}$ is stable under addition, it follows from Proposition 1.18 that it is an abelian group. The other claims follow by inspecting the definitions. $\square$

Definition B.7 (Multiplication of segments)**.** Assume that a unit segment 1 was chosen. Let a and b be two segments which represent the lengths x and y respectively. If $x = 0$ or $y = 0$ by definition we have $x \cdot y = 0$. If they are not both zero, let us cite the construction from [14, p.52]: "Lay off the segments 1 and b from the vertex O on a side of a right angle. Then lay off the segment a on the other side. Join the end points of the segment 1 and a with a line and through the end point of the segment b draw a parallel to this line. It will delineate a segment c on the other side. This segment is then called the *product* of the segments a by the segment b" and we write $c = ab$.



This defines a multiplication on $\mathbb{L}$ which we identified with the positive elements $\mathbb{F}^+_{AB} \cup \{0\}$. The product can be extended to the whole of $\mathbb{F}_{AB}$ by requiring that $x \cdot y = (-x) \cdot (-y)$ whenever x, y are negative and that $(-x) \cdot y = -(x \cdot y) = x \cdot (-y)$ whenever x is negative and y is positive.

Proposition B.8. Assume that a unit segment was chosen. With addition of vector and the above multiplication and ordering, $\mathbb{F}_{AB}$ is a totally ordered field and $\mathbb{L}$ is an ordered submonoid of $\mathbb{F}_{AB}$ with respect to both addition and multiplication.

Sketch of proof. The sum and product defined above are shown to be associative, commutative and distributive [14, §5]. Moreover it is easy to see that 1 is the neutral element for the product and it is easy to construct inverses. The other claims follow by inspecting the definitions. $\square$

The upshot of this section is the next theorem stating that the Axioms imply that $\mathbb{F}_{AB}$ is isomorphic to $\mathbb{R}$. Let us point out first that there are several ways of describing $\mathbb{R}$. It can be described as the 'smallest complete ordered field'. Concretely, this description requires that $\mathbb{R}$ satisfies a certain set of axioms (see for example [18, Chapter 1]). In order to show that it exists one can rely on several constructions. The construction closest to our setting is the construction of $\mathbb{R}$ via Dedekind cuts (see for example [18, Appendix to Chapter 1]).

Theorem B.9. For any unit segment $[AB]$ there is a unique isomorphism $\phi_{AB} : \mathbb{F}_{AB} \to \mathbb{R}$ of ordered fields mapping $\overrightarrow{0}$ to 0 and $\overrightarrow{AB}$ to 1. This isomorphism maps $\mathbb{L}$ to $\mathbb{R}_{\geq 0}$.

Sketch of proof. By Proposition B.8, we have that $\mathbb{F}_{AB}$ is an ordered field. Then, by Archimedes' Axiom (Axiom V.1) the field $\mathbb{F}_{AB}$ contains $\mathbb{Z}$ as a subring. Thus, it contains $\mathbb{Q}$ as an ordered subfield. By the Axiom of line completeness (Axiom V.2), the field $\mathbb{F}_{AB}$ is a complete field. However, any complete ordered field which containing $\mathbb{Q}$ as an ordered subfield is isomorphic to $\mathbb{R}$ through an order preserving isomorphism [17, p.17]. $\square$

Definition B.10. Assume that a unit segment $[AB]$ was chosen. By Theorem B.9, we may identify $\mathbb{F}_{AB}$ with $\mathbb{R}$ which gives a multiplication of real numbers with vectors. More precisely we have a map $\square \cdot \square : \mathbb{R} \times \mathbb{F}_{AB} \to \mathbb{F}_{AB}$ given by $(r, \mathbf{a}) \mapsto r \cdot \mathbf{a} := \phi_{AB}^{-1}(r)\mathbf{a}$.

Proposition B.11. Assume that a unit segment and a point O have been chosen. For any point X and any $r \in \mathbb{R}$ there is a unique point Y such that $\overrightarrow{OY} = r \cdot \overrightarrow{OX}$. Moreover we have $|\overrightarrow{OY}| = |r| \cdot |\overrightarrow{OX}|$.

Proposition B.12. In the setting of Definition B.10, for any vector a and any scalars $x, y \in \mathbb{R}$, we have $(x + y) \cdot \mathbf{a} = x \cdot \mathbf{a} + y \cdot \mathbf{a}$.

APPENDIX C

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 $\mathcal{F}$ [9, 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 components of the $\phi(\mathbf{e}_i)$'s with respect to the basis $\mathcal{F}$:

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

You have also learned [9, 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 C.1. 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 components of the vectors in $\mathcal{F}$ with respect to $\mathcal{E}$. If $\mathcal{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_n)$ then

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

APPENDIX D

Coordinate systems

D.1 Polar coordinates

Polar coordinates use oriented angles with a line. They establish a bijection between points distinct from the origin and pairs of numbers (r,φ) in $\mathbb{R}_{>0}\times[0,2\pi[$.

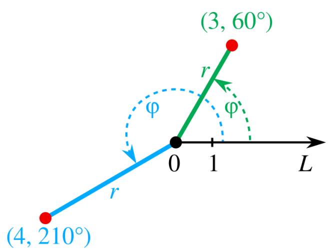


Figure D.1: Polar coordinates[1].

The correspondence between Cartesian coordinates (x,y) and polar coordinates (r,φ) is as follows:

$$\begin{cases} x = r\cos\varphi \\ y = r\sin\varphi \end{cases}, \quad \begin{cases} r = r(x,y) = \sqrt{x^2+y^2} \\ \varphi = f(x,y) \end{cases}$$

where, with $r \neq 0$,

$$f(x,y) = \begin{cases} \arccos(\frac{x}{r}) & \text{if } y \geq 0 \\ -\arccos(\frac{x}{r}) & \text{if } y < 0 \end{cases}.$$

[1] Image source: Wikipedia

D.2 Cylindrical coordinates

Cylindrical coordinates extend polar coordinates to dimension 3. They establish a bijection between points distinct from the origin and triples of numbers (r, φ, z) in $\mathbb{R}_{>0} \times [0, 2\pi[\times\mathbb{R}$.

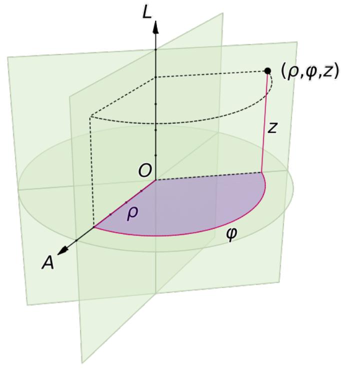


Figure D.2: Cylindrical coordinates[2].

The correspondence between Cartesian coordinates (x, y) and polar coordinates (r, φ) is as follows:

$$\begin{cases} x = r\cos\varphi \\ y = r\sin\varphi \\ z = z \end{cases}, \quad \begin{cases} r = r(x,y) = \sqrt{x^2+y^2} \\ \varphi = f(x,y) \\ z = z \end{cases}$$

where f is as in the previous Section D.1.

D.3 Spherical coordinates

Spherical coordinates also extend polar coordinates to dimension 3 but use angles for the third coordinate. They establish a bijection between points distinct from origin and triples of numbers (r, φ, θ) in $\mathbb{R}_{>0} \times [0, 2\pi[\times[0, \pi[$.

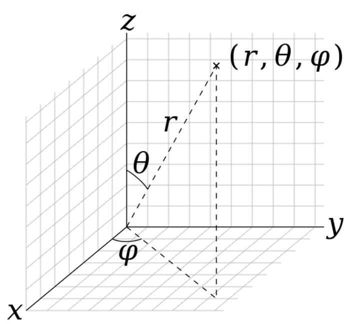


Figure D.3: Spherical coordinates[3].

[2]Image source: Wikipedia

The correspondence between Cartesian coordinates (x,y,z) and polar coordinates (r,φ,θ) is as follows:

$$\begin{cases} x = r\cos\varphi\sin\theta \\ y = r\sin\varphi\sin\theta \\ z = r\cos\theta \end{cases}, \quad \begin{cases} r = r(x,y,z) = \sqrt{x^2+y^2+z^2} \\ \varphi = f(x,y) \\ \theta = \arccos(\frac{z}{r}) \end{cases}$$

where f is as in the previous Section D.1.

D.4 Barycentric coordinates

In the n-dimensional affine space $\mathbb{A}^n$, let $P_0,P_1,\ldots,P_n$ be points such that $\mathcal{B} = (\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_1})$ is a basis of the direction space $\mathrm{D}(\mathbb{A}^n)$. This is the case if and only if the given points are the vertices of an n-simplex. Then $\mathcal{K} = (P_0,\mathcal{B})$ is a Cartesian frame. Every point A has a unique expression of the form

$$\begin{aligned} A &= P_0 + a_1\overrightarrow{P_0P_1} + \cdots + a_n\overrightarrow{P_0P_n} \\ &= (1-a_1-\cdots-a_n)P_0 + a_1P_1 + \cdots + a_nP_n \\ &= \lambda_0P_0 + \lambda_1P_1 + \cdots + \lambda_nP_n. \end{aligned}$$

The $n+1$ values $(\lambda_0,\lambda_1,\ldots,\lambda_n)$ are the barycentric coordinates of the point A. Notice that the sum of these coordinates equals 1. The above equation also shows how to translate from Cartesian coordinates to barycentric coordinates. This is the algebraic point of view.

The geometric interpretation in dimension 2 is the following. Let ABC be a triangle of area 1. A point P in the plane of the triangle has barycentric coordinates

$$\lambda_1 = \mathrm{Area}_{\mathrm{or}}(\triangle PAB), \quad \lambda_2 = \mathrm{Area}_{\mathrm{or}}(\triangle PBC), \quad \lambda_3 = \mathrm{Area}_{\mathrm{or}}(\triangle PCA).$$

Where $\mathrm{Area}_{\mathrm{or}}$ stands for the oriented area of a triangle. Similar, in higher dimensions, the barycentric coordinates are give by oriented volume relative to an n-simplex.

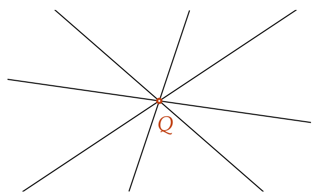
[3]Image source: Wikipedia

APPENDIX E

Bundles of lines and planes

E.1 Bundle of lines in $\mathbb{A}^2$

Definition E.1. 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 $\mathcal{L}_Q$.



Proposition E.2. 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

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

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

$$\ell_{1,t} : (a_1 x + b_1 y + c_1) + t(a_2 x + b_2 y + c_2) = 0.$$

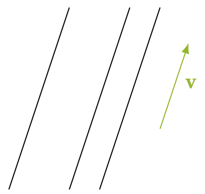
where $t = \frac{\mu}{\lambda} \in \mathbb{R}$. So $\ell_{1,\frac{\mu}{\lambda}}$ and $\ell_{\lambda,\mu}$ are in fact the same lines.

Definition E.3. 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 $\mathcal{L}_Q \backslash \{\ell_2\}$ for some $\ell_2 \in \mathcal{L}_Q$. With the above notation and discussion, it is the set

$$\left\{\ell_{1,t} : (a_1 x + b_1 y + c_1) + t(a_2 x + b_2 y + c_2) = 0 : t \in \mathbb{R}\right\}.$$

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

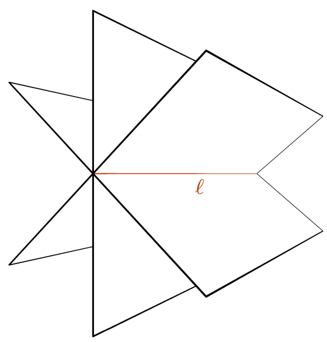
Definition E.4. Let $\mathbf{v} \in \mathbb{V}^2$. The set $\mathcal{L}_{\mathbf{v}}$ of all lines in $\mathbb{A}^2$ with direction vector $\mathbf{v}$ is called an *improper bundle of lines*, and $\mathbf{v}$ is called a *direction vector* of the bundle $\mathcal{L}_{\mathbf{v}}$.



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.

E.2 Bundle of planes in $\mathbb{A}^3$

Definition E.5. 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.



Proposition E.6. 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 3.3.2) 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 E.7. 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 E.8. 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.

APPENDIX F

Some classical theorems in affine geometry

In this section we collect classical theorems in affine geometry. The proofs of the first three theorems are vectorially (the theorem of Thales F.1, the affine version of Pappus' theorem F.2 and the theorem of Desargue F.3). We then introduce Lemma F.4 and Lemma F.4 which allow us to prove the next four theorem by means of frames (Pappus' hexagon theorem F.6, Newton-Gauss theorem F.7, and the theorems of Menelaus F.8 and Ceva F.9).

It is important to point out that in the process of deducing the affine space structure from Hilbert's Axioms we made use of some of these results. Indeed, in [14], Theorem 40 is Pappus' affine theorem F.2 and is a particular case of Pascal's theorem F.11. This result was implicitly used to define and derive properties of the multiplication of vectors with scalars. Therefore, proving the following theorem after using results from [14] may introduce circular reasoning.

The value of the following proofs lies in the fact that we prove them starting from the notion of an affine space only. Moreover, the proofs are such that they work not only for real affine spaces but for affine spaces over other commutative fields. These latter affine spaces are outside the scope of these notes. However, the proofs which allow this type of generalization are certainly of interest since they capture the essence of the statements.

The following theorem asserts that the ratio in which parallel lines cut a transversal line ℓ does not depend on ℓ but only on the parallel lines.

Theorem F.1 (Thales'[1] intercept theorem)**.** Let H, H' and H'' be three distinct parallel lines in the affine plane $\mathbb{A}^2$. Let ℓ_1 and ℓ_2 be two lines not parallel to H, H' and H''. For $i = 1, 2$ let

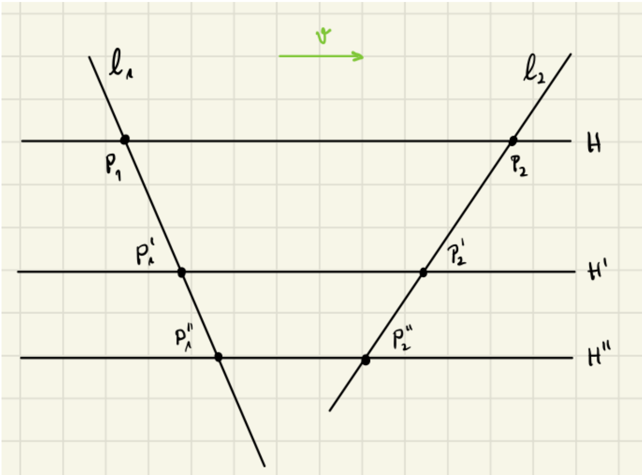
$$\begin{aligned} P_i &= \ell_i \cap H \\ P_i' &= \ell_i \cap H' \\ P_i'' &= \ell_i \cap H'' \end{aligned}$$

[1] c.626/623 – c.548/545 BC

and let k_1, k_2 be scalars such that

$$\overrightarrow{P_iP_i''} = k_i\,\overrightarrow{P_iP_i'}.$$

Then $k_1 = k_2$.



Proof. We follow the proof in [19]. If $\ell_1 = \ell_2$ the theorem is trivial. Suppose that $\ell_1 \neq \ell_2$. Then, if $P_1 = P_2$ the points P_1' and P_2' must be distinct. Thus, interchanging H with H' we may assume that $P_1 \neq P_2$. Now let $\mathbf{v} = \overrightarrow{P_1P_2}$ and consider the vectors

$$\begin{aligned} \overrightarrow{P_2P_2'} - \overrightarrow{P_1P_1'} = \overrightarrow{P_1'P_2'} - \overrightarrow{P_1P_2} &= \alpha\mathbf{v} \\ \overrightarrow{P_2P_2''} - \overrightarrow{P_1P_1''} = \overrightarrow{P_1''P_2''} - \overrightarrow{P_1P_2} &= \beta\mathbf{v} \end{aligned}$$

where α and β are scalars.

If $\alpha = 0$, from the first equation we have $\overrightarrow{P_2P_2'} = \overrightarrow{P_1P_1'}$. Since these are direction vectors of ℓ_1 and ℓ_2 it follows that the two lines are parallel. Since $\overrightarrow{P_2P_2''}$ and $\overrightarrow{P_1P_1''}$ are also direction vectors for the two lines, we see from the second equation that $\beta\mathbf{v}$ is a direction vector for ℓ_1 and ℓ_2. But $\mathbf{v}$ is a direction vector for H which is not parallel to ℓ_1. Thus, $\beta = 0$. Moreover, in this case we have

$$\begin{aligned} &\overrightarrow{P_2P_2''} = k_2\,\overrightarrow{P_2P_2'} = k_2\,\overrightarrow{P_1P_1'} \\ &\overrightarrow{P_1P_1''} = k_1\,\overrightarrow{P_1P_1'} \end{aligned}$$

and since $\overrightarrow{P_2P_2''} = \overrightarrow{P_1P_1''}$ it follows that $k_1 = k_2$.

If $\alpha \neq 0$, then

$$\overrightarrow{P_2P_2''} - \overrightarrow{P_1P_1''} = \beta\mathbf{v} = \alpha^{-1}\beta(\alpha\mathbf{v}) = \alpha^{-1}\beta\,\overrightarrow{P_2P_2'} - \alpha^{-1}\beta\,\overrightarrow{P_1P_1'}.$$

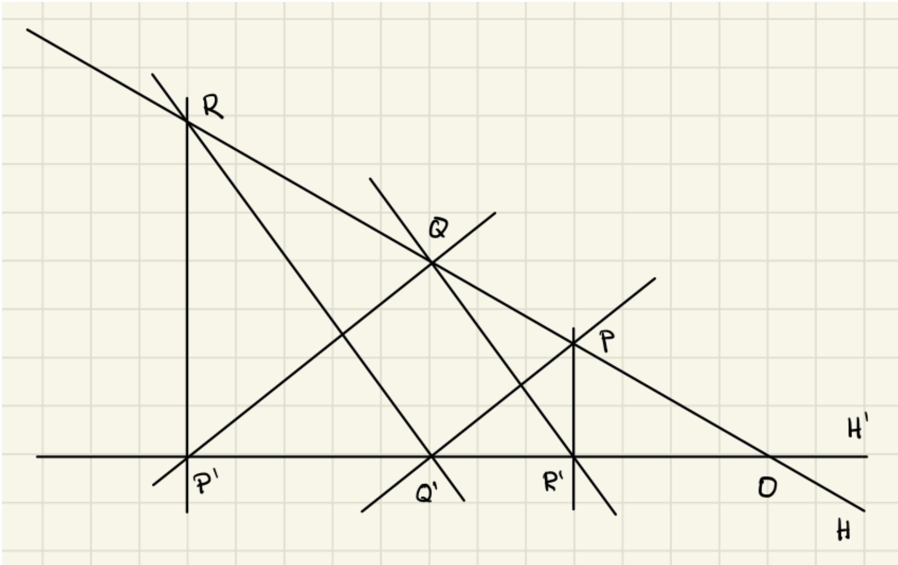
On the other hand,

$$\overrightarrow{P_2P_2''} - \overrightarrow{P_1P_1''} = k_2\,\overrightarrow{P_2P_2'} - k_1\,\overrightarrow{P_1P_1'}.$$

Now, since $\alpha \neq 0$, the vectors $\overrightarrow{P_1P_1'}$ and $\overrightarrow{P_2P_2'}$ are not parallel, i.e. they are linearly independent. Therefore, the last two equations imply $k_2 = \alpha^{-1}\beta = k_1$. $\square$

Remark. Theorem F.1 can be generalize to the case where the three parallel lines are replaced by parallel hyperpanes. The proof does not change if H, H' and H'' are replaced by hyperplanes.

Theorem F.2 (Pappus'[2] affine theorem)**.** Let H, H' be two distinct lines in the affine plane $\mathbb{A}^2$. Let $P, Q, R \in H$ and $P', Q', R' \in H'$ be distinct points, none of which lies at the intersection $H \cap H'$. If $PQ' \parallel P'Q$ and $QR' \parallel Q'R$ then $PR' \parallel P'R$.



Proof. We follow the proof in [19]. Suppose H and H' are not parallel, and let $H \cap H' = \{O\}$. By Theorem F.1, for some scalars h, k we have

$$\overrightarrow{OP'} = k\,\overrightarrow{OQ'}, \quad \overrightarrow{OQ} = k\,\overrightarrow{OP}, \quad \text{since} \quad PQ' \parallel P'Q,$$

$$\overrightarrow{OQ'} = h\,\overrightarrow{OR'}, \quad \overrightarrow{OR} = h\,\overrightarrow{OQ}, \quad \text{since} \quad QR' \parallel Q'R.$$

But then,

$$\overrightarrow{PR'} = \overrightarrow{OR'} - \overrightarrow{OP} = h^{-1}\,\overrightarrow{OQ'} - k^{-1}\,\overrightarrow{OQ}$$
$$\overrightarrow{RP'} = \overrightarrow{OP'} - \overrightarrow{OR} = k^{-1}\,\overrightarrow{OQ'} - h^{-1}\,\overrightarrow{OQ},$$

and so $\overrightarrow{RP'} = hk\,\overrightarrow{PR'}$, that is, $RP' \parallel PR'$. If $H \parallel H'$, then

$$\overrightarrow{PQ} = \overrightarrow{Q'P'}, \quad \text{since} \quad PQ \parallel Q'P',$$

$$\overrightarrow{QR} = \overrightarrow{R'Q'}, \quad \text{since} \quad QR \parallel R'Q',$$

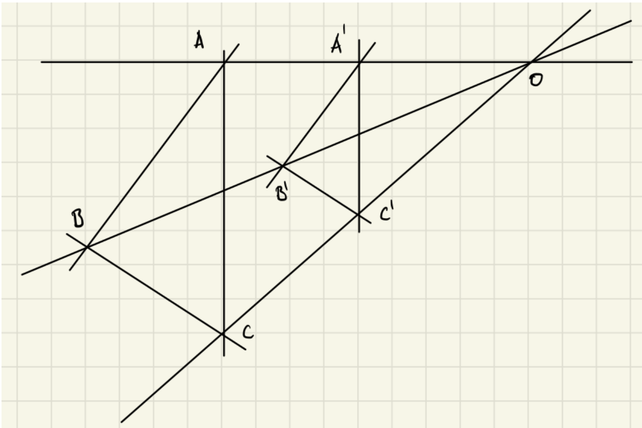
and so

$$\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{Q'P'} + \overrightarrow{R'Q'} = \overrightarrow{R'P'}.$$

Thus $PR' \parallel P'R$. □

[2] c.290 – c.350

Theorem F.3 (Desargue's[3] theorem)**.** Let $A,B,C,A',B',C' \in \mathbb{A}^n$ be points such that no three are collinear, and such that $AB \parallel A'B'$, $BC \parallel B'C'$ and $AC \parallel A'C'$. Then the three lines AA', BB' and CC' are either parallel or have a point in common.



Proof. We follow the proof in [19]. Suppose first that AA', BB' and CC' are not parallel. Then two of them meet, and we may assume that $AA' \cap BB' = \{O\}$. By Theorem F.1 applied to AB and $A'B$ we have

$$\overrightarrow{OA'} = k\,\overrightarrow{OA}, \quad \text{and} \quad \overrightarrow{OB'} = k\,\overrightarrow{OB}$$

for some scalar k. Let $\{C''\} = OC \cap A'C'$. By Theorem F.1, now applied to AC and $A'C'$ we have

$$\overrightarrow{OC''} = k\,\overrightarrow{OC}$$

since $\overrightarrow{OA'} = k\,\overrightarrow{OA}$. On the other hand, putting $\{C'''\} = OC \cap B'C'$, Theorem F.1 applied to the lines BC and $B'C'$ implies

$$\overrightarrow{OC'''} = k\,\overrightarrow{OC}$$

since $\overrightarrow{OB'} = k\,\overrightarrow{OB}$. Then, the last two equations imply that $C'' = C''' = C'$, and so O, C and C' are collinear. □

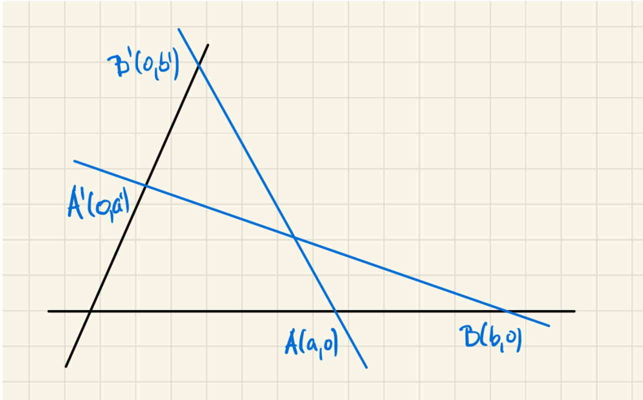
Lemma F.4. Let $\mathcal{K}$ be a frame of $\mathbb{A}^2$. Consider the lines

$$\ell_1 : \frac{x}{a} + \frac{y}{b'} = 1 \quad \text{and} \quad \ell_2 : \frac{x}{a'} + \frac{y}{b} = 1$$

for some non-zero scalars a, a', b, b'. Then the lines are parallel if and only if $aa' - bb' = 0$. If they are not parallel they meet in the point with coordinates

$$\left(\frac{ab(a'-b')}{aa'-bb'}, \frac{a'b'(a-b)}{aa'-bb'} \right).$$

[3] 1591 – 1661



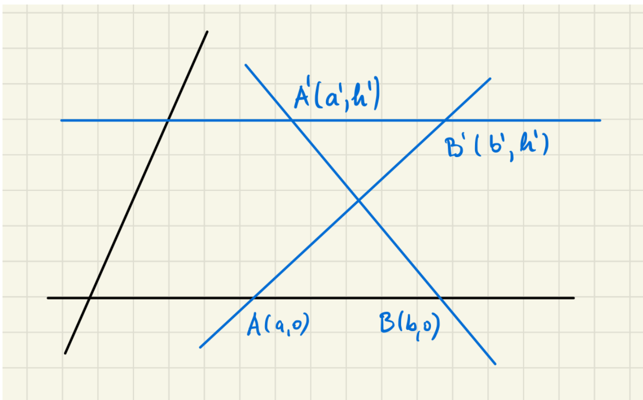
Proof. A direction vector for ℓ_1 is $(-a, b')$ and a direction vector for ℓ_2 is $(-b, a')$. The two vectors are linearly dependent if and only if

$$\frac{a}{b} = \frac{b'}{a'} \quad \Leftrightarrow \quad aa' = bb' \quad \Leftrightarrow \quad aa' - bb' = 0.$$

If ℓ_1 and ℓ_2 are not parallel, the intersection point is obtained by solving the system given by the equations of the two lines. □

Lemma F.5. Let $\mathcal{K}$ be a frame of $\mathbb{A}^2$. Consider the points $A(a,0)$, $B(b,0)$, $A'(a',h')$ and $B'(b',h')$ with $a \neq b$, $a' \neq b'$ and $h \neq 0$. If the lines AB' and $A'B$ are not parallel, they meet in the point with coordinates

$$\left(\frac{bb' - aa'}{b + b' - a - a'}, \frac{h'(b-a)}{b + b' - a - a'}\right).$$

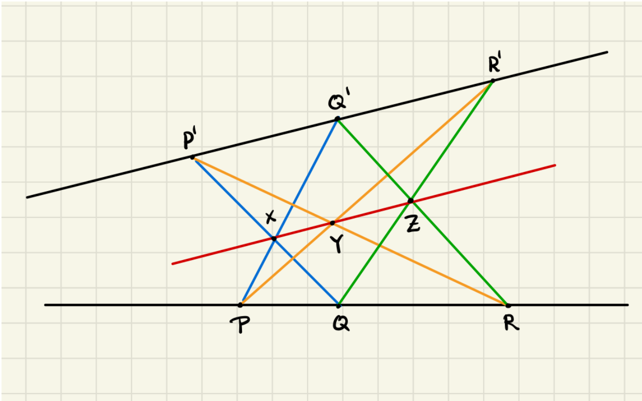


Proof. A direction vector for AB' is $(b' - a, h')$ and a direction vector for $A'B$ is $(a' - b, h')$. The equations of the two lines are

$$AB' : \frac{x-a}{b'-a} = \frac{y}{h'} \quad \text{and} \quad A'B : \frac{x-b}{a'-b} = \frac{y}{h'}$$

If ℓ_1 and ℓ_2 are not parallel, the intersection point is obtained by solving the system given by the equations of the two lines. □

Theorem F.6 (Pappus'[4] hexagon theorem)**.** Let H, H' be two distinct lines in the affine plane $\mathbb{A}^2$. Let $P,Q,R \in H$ and $P',Q',R' \in H'$ be distinct points, none of which lies at the intersection $H \cap H'$. Assume that $PQ' \cap P'Q = \{X\}$, $PR' \cap P'R = \{Y\}$ and $QR' \cap Q'R = \{Z\}$. Then the points X,Y,Z are collinear.



Proof. We first consider the case where H and H' are not parallel. Let $H \cap H' = \{O\}$ and choose a frame $\mathcal{K}$ with the origin in O, with the first coordinate axis H and the second coordinate axis H'. With respect to $\mathcal{K}$ the coordinates of the given points are as follows

$$P(p,0), \quad Q(q,0), \quad R(r,0), \quad P'(0,p'), \quad Q'(0,q'), \quad R(0,r').$$

Applying Lemma F.4 three times with the frame $\mathcal{K}$, we obtain the coordinates of the intersection points:

$$X\left(\frac{pq(p'-q')}{pp'-qq'}, \frac{p'q'(p-q)}{pp'-qq'}\right), \quad Y\left(\frac{pr(p'-r')}{pp'-rr'}, \frac{p'r'(p-r)}{pp'-rr'}\right), \quad Z\left(\frac{qr(q'-r')}{qq'-rr'}, \frac{q'r'(q-r)}{qq'-rr'}\right).$$

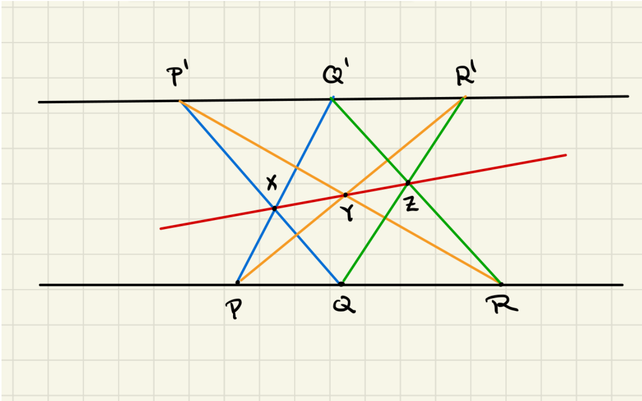
We may check the collinearity of these points by calculating a determinant. This amounts to an algebraic calculation. We may simplify this calculation by noticing that the frame $\mathcal{K}$ can be choosen such that $p = p' = 1$. A calculation shows that

$$\begin{vmatrix} \frac{q(1-q')}{1-qq'} & \frac{q'(1-q)}{1-qq'} & 1 \\ \frac{r(1-r')}{1-rr'} & \frac{r'(1-r)}{1-rr'} & 1 \\ \frac{qr(q'-r')}{qq'-rr'} & \frac{q'r'(q-r)}{qq'-rr'} & 1 \end{vmatrix} = \frac{1}{(1-qq')(1-rr')(qq'-rr')} \begin{vmatrix} q(1-q') & q'(1-q) & 1-qq' \\ r(1-r') & r'(1-r) & 1-rr' \\ qr(q'-r') & q'r'(q-r) & qq'-rr' \end{vmatrix} = 0.$$

Now consider the case where H and H' are parallel. Choose a frame $\mathcal{K}$ with origin P and with the first coordinate axis H and the second coordinate axis the line PP'. With respect to $\mathcal{K}$ the coordinates of the points are as follows

$$P(0,0), \quad Q(q,0), \quad R(r,0), \quad P'(0,h'), \quad Q'(q',h'), \quad R(r',h').$$

[4] c.290 – c.350



Applying Lemma F.5 three times with the frame $\mathcal{K}$, we obtain the coordinates of the intersection points:

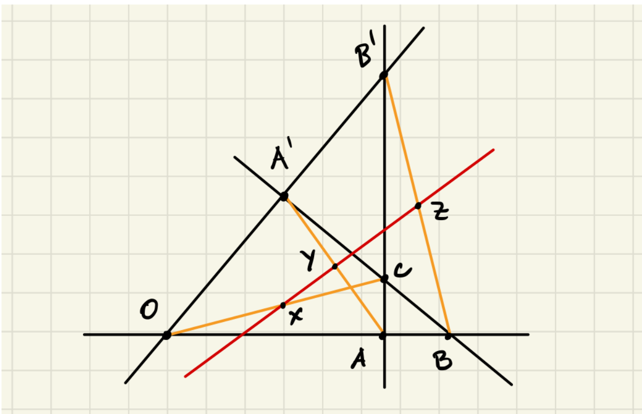
$$X\left(\frac{qq'}{q+q'}, \frac{h'q}{q+q'}\right) \quad Y\left(\frac{rr'}{r+r'}, \frac{h'r}{r+r'}\right) \quad Z\left(\frac{qq'-rr'}{q+q'-r-r'}, \frac{h'(q-r)}{q+q'-r-r'}\right)$$

Similar to the previous case, the collinearity of these points can be checked by calculating a determinant. It is easy to see that this determinant is zero:

$$\frac{h'}{(q+q')(r+r')(q+q'-r-r')}\begin{vmatrix} qq' & q & q+q' \\ rr' & r & r+r' \\ qq'-rr' & q-r & q+q'-r-r' \end{vmatrix} = 0.$$

□

Remark. The two theorems of Pappus can be unified in the context of projective geometry. If in the previous theorem (Theorem F.6) we let the points X, Y, Z go to infinity, we obtain Theorem F.2. Moreover, the union of the two lines H and H' in these theorems is a degenerate conic. Both these theorems are particular cases of Pascal's Theorem F.11.



Theorem F.7 (Newton-Gauss line)**.** A complete quadrilateral is the configuration of four lines, no three of which pass through the same point. Let O,A,B,C,A',B' be the intersection points of such lines, i.e. the vertices of the complete quadrilateral, such that A lies between O and B and such that A' lies between O and B'. The diagonals of this quadrilateral are the segments $[OC]$, $[AA']$ and $[BB']$. The midpoints of a complete quadrilateral are collinear.

Proof. Choose the coordinate frame with origin O and basis $(\overrightarrow{OA},\overrightarrow{OA'})$. Then the coordinates of the points are $O(0,0)$, $A(1,0)$, $A'(0,1)$, $B(b,0)$, $B'(0,b')$ and by Lemma F.4

$$C\left(\frac{b(1-b')}{1-bb'},\frac{b'(1-b)}{1-bb'}\right).$$

Thus, if X,Y,Z are the midpoints of $[OC]$, $[AA']$, $[BB']$ respectively, then

$$X\left(\frac{b(1-b')}{2(1-bb')},\frac{b'(1-b)}{2(1-bb')}\right),\quad Y\left(\frac{1}{2},\frac{1}{2}\right),\quad Z\left(\frac{b}{2},\frac{b'}{2}\right).$$

The collinearity of the three points is equivalent to the vanishing of the following determinant:

$$\frac{1}{4(1-bb')}\begin{vmatrix} b(1-b') & b'(1-b) & 1-bb' \\ 1 & 1 & 1 \\ b & b' & 1 \end{vmatrix}=0.$$

□

The following two theorems give necessary and sufficient conditions for three points to be collinear and, respectively, for three lines to be concurrent in terms of oriented ratios of the sides of a triangle. In the context of projective geometry, Menelaus' theorem and Ceva's theorem may be seen as projective duals (see for example [2]). In order to state them, we need to introduce *oriented ratios*. Let P pe a point on the line AB. The oriented ratio $AP:PA$, in which P divides the segment $[AB]$ is the ordinary ratio $|AP|:|PB|$ if P is between A and B and it is $-|AP|:|PB|$ otherwise. Notice that the signed ratio is defined by the equation

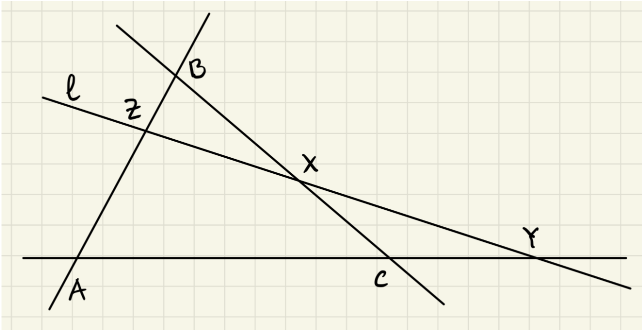
$$\overrightarrow{AP}=\big(AP:PB\big)\overrightarrow{PB} \tag{F.1}$$

which is why it is sometimes written as $\overrightarrow{AP}/\overrightarrow{PB}$.

Theorem F.8 (Menelaus'[5] Theorem)**.** Let ABC be a triangle and ℓ a line which does not pass through the vertices of the triangle. Let X,Y,Z be points on the lines BC, CA and AB respectively and consider the signed ratios $\alpha=BX:XC$, $\beta=CY:YA$ and $\gamma=AZ:ZB$. Then X,Y,Z are collinear if and only if

$$\alpha\cdot\beta\cdot\gamma=-1.$$

[5] c.70–c.130



Proof. By Proposition 6.4, affine transformations preserve oriented ratios, i.e. oriented ratios are independent of the choice of the Cartesian frame. We choose a frame with origin A and basis $(\overrightarrow{AC}, \overrightarrow{AZ})$. Consider the coordinates of the points in this frame: $A(0,0)$, $B(0,b)$, $C(1,0)$, $Y(y,0)$, $Z(0,1)$. By Lemma F.4, the point $X \in BC$ lies on the line YZ if and only if it has the following coordinates

$$X\left(\frac{y(1-b)}{1-by}, \frac{b(1-y)}{1-by}\right). \tag{F.2}$$

Considering the components of the relevant vectors

$$\overrightarrow{BX} = \frac{y(1-b)}{1-by}(1,-b) \quad \text{and} \quad \overrightarrow{XC} = \frac{y-1}{1-by}(1,-b)$$

we see that (F.2) holds if and only if

$$\beta = BX : XC = \frac{y(1-b)}{y-1}. \tag{F.3}$$

Considering the other vectors we have

$$\overrightarrow{AZ} = (0,1), \quad \overrightarrow{ZB} = (b-1)(0,1) \quad \Rightarrow \quad \gamma = AZ : ZB = \frac{1}{b-1} \quad \text{and}$$

$$\overrightarrow{CY} = (y-1)(1,0), \quad \overrightarrow{YA} = -y(1,0) \quad \Rightarrow \quad \alpha = CY : YA = \frac{y-1}{-y}.$$

It follows that (F.3) is equivlanet to

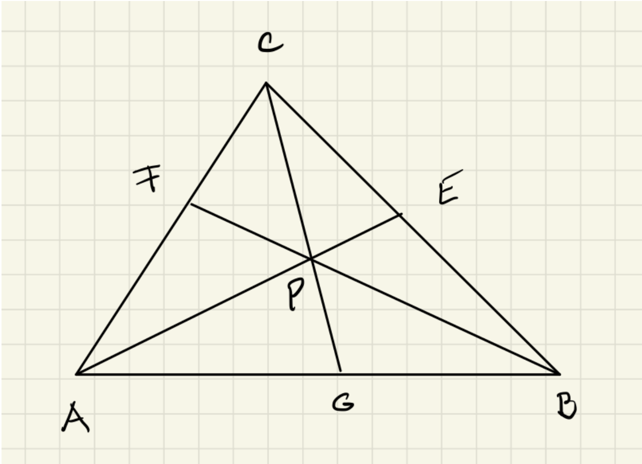
$$\beta = -\frac{1}{\alpha \cdot \beta} \quad \Leftrightarrow \quad \alpha \cdot \beta \cdot \gamma = -1.$$

□

Theorem F.9 (Ceva's[6] Theorem)**.** Let ABC be a triangle. Let $E \in BC$, $F \in CA$ and $G \in AB$ and consider the signed ratios $\alpha = BE : EC$, $\beta = CF : FA$ and $\gamma = AG : GB$. The three lines AE, BF and CG are concurrent if and only if

$$\alpha \cdot \beta \cdot \gamma = 1.$$

[6] 1647-1734



Proof. By Proposition 6.4, affine transformations preserve oriented ratios, i.e. oriented ratios are independent of the choice of the Cartesian frame. We choose a frame with origin A and basis $(\overrightarrow{AG}, \overrightarrow{AF})$. Consider the coordinates of the points in this frame: $A(0,0)$, $B(0,b)$, $C(c,0)$, $F(0,1)$, $G(1,0)$. By Lemma F.4, the intersection point P of the lines BF and CG has coordinates

$$P\left(\frac{b(1-c)}{1-bc}, \frac{c(1-b)}{1-bc}\right).$$

Then, the lines AE, BF, CG are concurrent if and only if P lies on AE, i.e. if and only if E is the intersection of AP with BC. These two lines are described by the equations

$$BC: \frac{x}{b} + \frac{y}{c} = 1 \quad \text{and} \quad AP: \begin{cases} x = \frac{b(1-c)}{1-bc}t \\ y = \frac{c(1-b)}{1-bc}t \end{cases}.$$

Their intersection point is

$$\left(\frac{b(1-c)}{2-b-c}, \frac{c(1-b)}{2-b-c}\right) \tag{F.4}$$

Considering the components of the relevant vectors

$$\overrightarrow{BE} = \frac{b-1}{2-b-c}(b,-c) \quad \text{and} \quad \overrightarrow{EC} = \frac{c-1}{2-b-c}(b,-c)$$

we see that E has coordinates (F.4) if and only if

$$\alpha = \frac{b-1}{c-1}. \tag{F.5}$$

Considering the other vectors we have

$$\overrightarrow{AG} = (1,0), \quad \overrightarrow{GB} = (b-1)(1,0) \quad \Rightarrow \quad \gamma = AG : GB = \frac{1}{b-1}$$

$$\text{and} \overrightarrow{CF} = (1-c)(0,1), \quad \overrightarrow{FA} = -(1,0) \quad \Rightarrow \quad \beta = CF : FA = \frac{1-c}{-1}$$

and we see that (F.5) is equivlanet to

$$\alpha = \frac{1}{\beta \cdot \gamma} \quad \Leftrightarrow \quad \alpha \cdot \beta \cdot \gamma = 1.$$

□

Remark. The lines AE, BF and CG are known as *cevians*. As is often the case with mathematical discoveries, attributing a theorem to a single mathematician can be challenging. Today, we know that a proof of Ceva's theorem was discovered much earlier by Al-Mu'taman[7].

The intersection point of two cevians can be described vectorially as follows.

Proposition F.10. Let ABC be a triangle. Let $G \in AB$ and $F \in AC$ be two points distinct from the vertices of the triangle and let $P = BF \cap CG$. Consider the signed ratios $\lambda = AG : GB$ and $\mu = AF : FC$. For any point O we have

$$\overrightarrow{OP} = \frac{\overrightarrow{OA} + \lambda \overrightarrow{OB} + \mu \overrightarrow{OC}}{1 + \lambda + \mu}. \tag{F.6}$$

Proof. By (F.1) we have $\overrightarrow{AG} = \lambda \overrightarrow{GB}$ and $\overrightarrow{AF} = \mu \overrightarrow{FC}$. Therefore

$$\overrightarrow{AB} = \overrightarrow{AG} + \overrightarrow{GB} = (1+\lambda)\overrightarrow{GB} = (1+\lambda)\frac{\overrightarrow{GB}}{\overrightarrow{AG}}\overrightarrow{AG} = \frac{1+\lambda}{\lambda}\overrightarrow{AG}.$$

and

$$\overrightarrow{AC} = \overrightarrow{AF} + \overrightarrow{FC} = (1+\mu)\overrightarrow{FC} = (1+\mu)\frac{\overrightarrow{FC}}{\overrightarrow{AF}}\overrightarrow{AF} = \frac{1+\mu}{\mu}\overrightarrow{AG}.$$

In other words, with respect to the frame with origin A and basis $(\overrightarrow{AG}, \overrightarrow{AF})$, we have the following coordinates $G(1,0)$, $F(0,1)$, $B(\lambda/(1+\lambda), 0)$ and $C(0, \mu/(1+\mu))$. Thus, by Lemma F.4, the intersection point P has coordinates

$$P\left(\frac{\lambda}{1+\lambda+\mu}, \frac{\mu}{1+\lambda+\mu}\right).$$

Hence, for any point O we have

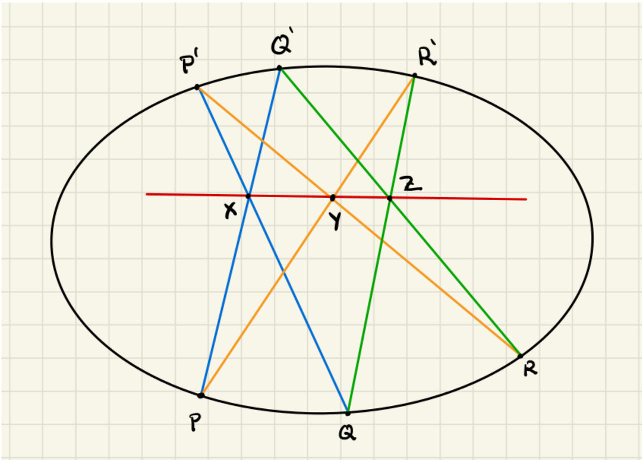
$$\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{AP} = \overrightarrow{OA} + \frac{\lambda\overrightarrow{AB} + \mu\overrightarrow{AC}}{1+\lambda+\mu} = \overrightarrow{OA} + \frac{(\lambda+\mu)\overrightarrow{AO} + \lambda\overrightarrow{OB} + \mu\overrightarrow{OC}}{1+\lambda+\mu} = \frac{\overrightarrow{OA} + \lambda\overrightarrow{OB} + \mu\overrightarrow{OC}}{1+\lambda+\mu}$$

□

Remark. Equation (F.6) in the above proposition does not depend on the point O. The coefficients of the three vectors are the barycentric coordinates of the point P relative to the triangle ABC (See Section D.4). The above proposition can be applied to derive the barycentric coordinates of the intersection point of different cevians, such as the medians, the altitudes, etc.

[7] 11th century

Theorem F.11 (Pascal's hexagrammum mysticum theorem)**.** If all six vertices of a hexagon lie on a conic section and the three pairs of opposite sides intersect, then the three points of intersection are collinear.



Proof. In the case where the conic section is a circle, a proof of this theorem based on Menelaus' theorem can be found in [8, §.3.8]. □

APPENDIX G

Eigenvalues and Eigenvectors

Definition G.1. Let $\mathbb{V}$ be a $\mathbb{R}$-vector space and let $\phi : \mathbb{V} \to \mathbb{V}$ be a linear map. A non-zero vector $\mathbf{v} \in \mathbb{V}$ is called an *eigenvector of* ϕ if there is a scalar $\lambda \in \mathbb{R}$ such that $\phi(\mathbf{v}) = \lambda\mathbf{v}$. The scalar λ is then called *the eigenvalue associated to the eigenvector* $\mathbf{v}$. The set of eigenvalues of ϕ is called the *spectrum of* ϕ.

For $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ an *eigenvector of* A is an eigenvector $\mathbf{v} \in \mathbb{R}^n$ for the map $\phi_A : \mathbb{R}^n \to \mathbb{R}^n$ defined by A through $\phi(\mathbf{x}) = A\mathbf{x}$, and an *eigenvalue* of A is an eigenvalue for ϕ_A.

- If $\phi = \mathrm{Id}_{\mathbb{V}}$, then every non-zero vector $\mathbf{v}$ is an eigenvector of ϕ with eigenvalue $\lambda = 1$.
- Every non-zero vector in $\ker(\phi)$ is an eigenvector for ϕ with eigenvalue $\lambda = 0$.

Proposition G.2. The eigenvalue associated to an eigenvector is unique.

Proposition G.3. If $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{V}$ are eigenvectors with the same eigenvalue λ, then for every $c_1, c_2 \in \mathbb{R}$ the vector $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, if it is non-zero, is also an eigenvector with eigenvalue λ.

Definition G.4. From the above proposition it follows that for each $\lambda \in \mathbb{R}$ the set

$$\mathbb{V}_\lambda(\phi) = \{\mathbf{v} \in \mathbb{V} : \mathbf{v} \text{ is an eigenvector of } \phi \text{ with eigenvalue } \lambda\} \cup \{0\}$$

is a vector subspace of $\mathbb{V}$, called the *eigenspace for the eigenvalue* λ. For a matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ the *eigenspace for the eigenvalue* λ is defined to be the subspace $\mathbb{V}_\lambda(A) := \mathbb{V}_\lambda(\phi_A)$ in $\mathbb{R}^n$.

Proposition G.5. If $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{V}$ are eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_k$ respectively, and these λ_i are pairwise distinct, then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent.

Proposition G.6. If every $\mathbf{v} \in \mathbb{V} \setminus \{0\}$ is an eigenvector of ϕ then there exists $\lambda \in \mathbb{R}$ such that $\phi = \lambda\,\mathrm{Id}_{\mathbb{V}}$.

G.1 Characteristic polynomial

In order to find the eigenvalues of a linear map $\mathbb{V} \to \mathbb{V}$ one uses the characteristic polynomial.

Proposition G.7. Let $\mathbb{V}$ be a finite dimensional vector space and let $\phi : \mathbb{V} \to \mathbb{V}$ be a linear map. A scalar $\lambda \in \mathbb{R}$ is an eigenvalue of ϕ if and only if the map

$$\phi - \lambda \operatorname{Id}_{\mathbb{V}} : \mathbb{V} \to \mathbb{V} \quad \text{defined by} \quad (\phi - \lambda \operatorname{Id}_{\mathbb{V}})(\mathbf{v}) = \phi(\mathbf{v}) - \lambda \mathbf{v}$$

is *not* bijective, that is, if and only if $\det(\phi - \lambda \operatorname{Id}_{\mathbb{V}}) = 0$.

- Let $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ be a basis of $\mathbb{V}$. The matrix associated to the map $\lambda \operatorname{Id}_{\mathbb{V}}$ is

$$[\lambda \operatorname{Id}_{\mathbb{V}}]_{\mathcal{B}} = \begin{bmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda \end{bmatrix}$$

  and if $A = (a_{ij}) = [\phi]_{\mathcal{B}}$ then

$$[\phi - \lambda \operatorname{Id}_{\mathbb{V}}]_{\mathcal{B}} = \begin{bmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{bmatrix}.$$

Definition G.8. Let $A \in \operatorname{Mat}_{n \times n}(\mathbb{R})$. The determinant

$$P_A(T) : \det(A - T \operatorname{Id}_n) = \begin{vmatrix} a_{11} - T & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - T & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - T \end{vmatrix}$$

is a polynomial of degree n in T, called *the characteristic polynomial of* A. If $\phi : \mathbb{V} \to \mathbb{V}$ is linear and $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ is a basis of $\mathbb{V}$ then the *characteristic polynomial of* ϕ is $P_\phi := P_{[\phi]_{\mathcal{B}}}$.

Proposition G.9. The definition of P_ϕ is independent of the basis.

Corollary G.10. Let $\mathbb{V}$ be a vector space of dimension n, and let $\phi : \mathbb{V} \to \mathbb{V}$ be linear. Then $\lambda \in \mathbb{R}$ is an eigenvalue of ϕ if and only if λ is a root of the polynomial P_ϕ. In particular, ϕ has at most n eigenvalues.

Proposition G.11. Let $\mathbb{V}$ be a finite dimensional vector space. A linear map $\phi : \mathbb{V} \to \mathbb{V}$ is diagonalizable if and only if there is a basis of $\mathbb{V}$ consisting entirely of eigenvectors of ϕ.

Theorem G.12. Let $\mathbb{V}$ be a real vector space of dimension n, and let $\phi : \mathbb{V} \to \mathbb{V}$ be linear. If $\{\lambda_1, \ldots, \lambda_k\} \subseteq \mathbb{R}$ is the spectrum of ϕ, then

$$\dim(V_{\lambda_1}(\phi)) + \cdots + \dim(V_{\lambda_k}(\phi)) \leq n$$

with equality if and only if ϕ is diagonalizable.

Corollary G.13. If $\dim(\mathbb{V}) = n$ and ϕ has n distinct eigenvalues then it is diagonalizable.

- We have a practical method for finding eigenvalues and eigenvectors.
- Suppose we are given $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$.
  1. Calculate the characteristic polynomial P_A.
  2. Find the eigenvalues of A by calculating the roots of P_A which lie in $\mathbb{R}$.
  3. For each eigenvalue $\lambda \in \mathbb{R}$, the homogenous system of n equations in n unknowns:

     $$(A - \lambda\,\mathrm{Id}_n)\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

     has rank $r < n$ and therefore has nontrivial solutions. The space of solutions is the eigenspace $\mathbb{V}_\lambda(A)$.
- If the sum of the dimensions of the eigenspaces found by letting λ vary over the roots of P_A is equal to n then A is diagonalizable (by Theorem G.12).
- A linear map $\mathbb{V} \to \mathbb{V}$ need not have any eigenvalues nor eigenvectors.
- If $\dim(\mathbb{V})$ is odd, then the characteristic polynomial has odd degree and so has at least one real root. Thus, every linear map $\mathbb{V} \to \mathbb{V}$ on an odd dimensional real vector space has at least one eigenvalue, and so at least one eigenvector.
- If we replace $\mathbb{R}$ by $\mathbb{C}$, by the fundamental theorem of algebra, P_A has roots in $\mathbb{C}$. Thus, every linear map $\mathbb{V} \to \mathbb{V}$ on a finite dimensional complex vector space has at least one eigenvalue, and so at least one eigenvector. Of course, this does not mean that the linear map is diagonalizable.

Definition G.14. Let $\mathbb{V}$ be a finite dimensional $\mathbb{R}$-vector space and let $\phi : \mathbb{V} \to \mathbb{V}$ be linear admitting λ as eigenvalue. The number $\dim(\mathbb{V}_\lambda(\phi))$ is called the *geometric multiplicity of λ for ϕ*. The *algebraic multiplicity of λ for ϕ* is instead the multiplicity of λ as a root of the characteristic polynomial P_ϕ; this is denoted by $h_\phi(\lambda)$.

Proposition G.15. For any linear map $\phi : \mathbb{V} \to \mathbb{V}$ and $\lambda \in \mathbb{R}$ one has

$$\dim(\mathbb{V}_\lambda(\phi)) \le h_\phi(\lambda),$$

that is, the geometric multiplicity is not larger than the algebraic multiplicity.

APPENDIX H

Bilinear forms and symmetric matrices

The purpose of this appendix is threefold:

1. We prove Sylvester's theorem (Theorem H.7) which shows that the properties in Proposition 4.15 define scalar products (Corollary H.10).

2. From the proof of Sylvester's theorem one obtains a simple algorithm for the affine classification of quadrics (see Chapter **??**).

3. We prove the Spectral theorem (Theorem H.14) which is used in the isometric classification of quadratic surfaces (see Chapter **??**).

Bilinear forms are the natural context for the proofs of both theorems.

H.1 Affine diagonalization

Throughout we let $\mathbb{V}$ denote a finite dimensional real vector space. The proofs for the statements in this appendix can be found in [19, Chapter 15 and 16].

Definition H.1. A map $\phi : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ is called a *bilinear form* if it is linear in both arguments, i.e.

$$\phi(a\mathbf{u} + b\mathbf{v}, \mathbf{w}) = a\phi(\mathbf{u}, \mathbf{w}) + b\phi(\mathbf{v}, \mathbf{w}) \quad \text{and} \quad \phi(\mathbf{u}, a\mathbf{v} + b\mathbf{w}) = a\phi(\mathbf{u}, \mathbf{v}) + b\phi(\mathbf{u}, \mathbf{w}). \tag{H.1}$$

for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{V}$ and all $a, b \in \mathbb{R}$.

With respect to a basis $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ of $\mathbb{V}$, the values of a bilinear form ϕ are determined by its values on the basis vectors. These values are the entries in the *Gram matrix* $\mathrm{G}_{\mathcal{B}}(\phi)$ *of* ϕ *relative to* $\mathcal{B}$, which is defined by $\mathrm{G}_{\mathcal{B}}(\phi) = \phi(\mathbf{e}_i, \mathbf{e}_j)$. This matrix determines ϕ 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

$$G_{\mathcal{B}'}(\phi) = M^T_{\mathcal{B},\mathcal{B}'} \cdot G_{\mathcal{B}}(\phi) \cdot M_{\mathcal{B},\mathcal{B}'}. \tag{H.2}$$

In particular, the rank of $G_{\mathcal{B}}(\phi)$ does not depend on $\mathcal{B}$. It depends only on ϕ. We denote it by $\operatorname{rank}(\phi)$ and call it *the rank of* ϕ.

Definition H.2. A bilinear form $\phi : \mathbb{V}^n \times \mathbb{V}^n \to \mathbb{R}$ is called *symmetric* if the Gram matrix $G_{\mathcal{B}}(\phi)$ is a symmetric matrix with respect to a basis $\mathcal{B}$. It follows from (H.2) that this definition doesn't depend on the basis $\mathcal{B}$.

Definition H.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 H.4. Let ϕ be a symmetric bilinear form on the vector space $\mathbb{V}$. The quadratic form q_ϕ associated to ϕ satisfies

$$\begin{aligned} q_\phi(\lambda \mathbf{v}) &= \lambda^2 q_\phi(\mathbf{v}) \\ 2\phi(\mathbf{v},\mathbf{w}) &= q(\mathbf{v}+\mathbf{w}) - q(\mathbf{v}) - q(\mathbf{w}) \end{aligned}$$

for every $\lambda \in \mathbb{R}$ and every $\mathbf{v},\mathbf{w} \in \mathbb{V}$.

Proof. The first property is an immediate consequence of (H.1). Further,

$$\begin{aligned} q_\phi(\mathbf{v}+\mathbf{w}) - q_\phi(\mathbf{v}) - q_\phi(\mathbf{w}) &= \phi(\mathbf{v}+\mathbf{w},\mathbf{v}+\mathbf{w}) - \phi(\mathbf{v},\mathbf{v}) - \phi(\mathbf{w},\mathbf{w}) \\ &= \phi(\mathbf{v},\mathbf{w}) + \phi(\mathbf{w},\mathbf{v}) \\ &= 2\phi(\mathbf{v},\mathbf{w}). \end{aligned}$$

□

Remark. In particular, Proposition H.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 H.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 $G_{\mathcal{B}}(\phi)$ is a diagonal matrix.

Theorem H.6. Let ϕ be a symmetric bilinear form on the vector space $\mathbb{V}$. Then there exists a diagonalizing basis for ϕ.

Proof. We take the proof from [19, p.232]. The proof is 'by completing the squares' and induction on n. If $n = 1$ there is nothing to prove. Suppose therefore that $n \geq 2$, and that every symmetric bilinear form on a space of dimension less than n has a diagonalizing basis. Choose a basis $\mathcal{B} = (\mathbf{b}_1,\dots,\mathbf{b}_n)$ of $\mathbb{V}$. If ϕ is the zero form, then $\mathcal{B}$ is diagonalizing and there is nothing to prove. Otherwise we can obtain from $\mathcal{B}$ a second basis $\mathcal{B}' = (\mathbf{c}_1,\dots,\mathbf{c}_n)$ for which $\phi(\mathbf{c}_1,\mathbf{c}_n) \neq 0$. Indeed, if there is some i for which $\phi(\mathbf{b}_i,\mathbf{b}_i) \neq 0$ then it suffices to exchange $\mathbf{b}_1$ and $\mathbf{b}_i$. If, on the other hand, $\phi(\mathbf{b}_i,\mathbf{b}_i) = 0$ for all i, then there are $i \neq j$ for which $\phi(\mathbf{b}_i,\mathbf{b}_j) \neq 0$ (otherwise ϕ is the zero form), and again we can exchange these with $\mathbf{b}_1$ and $\mathbf{b}_2$ so that $\phi(\mathbf{b}_1,\mathbf{b}_2) \neq 0$. The new basis

$$\mathcal{B}' = (\mathbf{b}_1 + \mathbf{b}_2, \mathbf{b}_2, \dots, \mathbf{b}_n)$$

has the required property. In the basis $\mathcal{B}'$, the quadratic form q_ϕ associated to ϕ has the form

$$q(\mathbf{v}(y_1,\dots,y_n)) = h_{11}y_1^2 + 2\sum_{i=2}^{n} h_{1i}y_1y_i + \sum_{i,j=2}^{n} h_{ij}y_iy_j,$$

where $h_{ij} = \phi(\mathbf{c}_i,\mathbf{c}_j)$. Since $h_{11} = \phi(\mathbf{c}_1,\mathbf{c}_1) \neq 0$ we can rewrite the above equation as

$$q(\mathbf{v}(y_1,\dots,y_n)) = h_{11}\left(y_1^2 + 2\sum_{i=2}^{n} h_{11}^{-1}h_{1i}y_1y_i\right) + (\text{terms not involving } y_1).$$

we now change coordinates as follows

$$z_1 = y_1 + \sum_{i=2}^{n} h_{11}^{-1}h_{1i}y_i, \quad z_2 = y_2,\dots,z_n = y_n,$$

which corresponds to a change of basis from $\mathcal{B}'$ to $\mathcal{B}'' = (\mathbf{d}_1,\dots,\mathbf{d}_n)$ given by $\mathbf{d}_1 = \mathbf{c}_1$, and for $i > 1$, $\mathbf{d}_1 = \mathbf{c}_1 - h_{11}^{-1}h_{1i}\mathbf{c}_1$. In these coordinates, q_ϕ has the form

$$q(\mathbf{v}(y_1,\dots,y_n)) = h_{11}z_1^2 + q'(z_2,\dots,z_n),$$

where q' is a homogeneous polynomial of degree 2 in $z_1,\dots,z_n$, and so defines a quadratic form on the space $\langle \mathbf{d}_2,\dots,\mathbf{d}_n\rangle$. By the inductive hypothesis, $\langle \mathbf{d}_2,\dots,\mathbf{d}_n\rangle$ has a basis $(\mathbf{e}_2,\dots,\mathbf{e}_n)$ which diagonalizes q'. Thus, the basis $(\mathbf{d}_1,\mathbf{e}_2,\dots,\mathbf{e}_n)$ is diagonalizing for q. □

Theorem H.7. (Sylvester's[1] Theorem) 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 the Gram matrix $\mathrm{G}_{\mathcal{B}}(\phi)$ has the form

$$\begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{H.3}$$

where 0 stands for zero matrices of appropriate sizes.

Proof. We take the proof from [19, p.232]. By Theorem H.6 there is a basis $\mathcal{B} = (\mathbf{b}_1,\dots,\mathbf{b}_n)$ for which

$$q_\phi(\mathbf{v}) = b_{11}y_1^2 + \cdots + b_{nn}y_n^2.$$

for each $\mathbf{v} = y_1\mathbf{b}_1 + \cdots + y_n\mathbf{b}_n \in \mathbb{V}$. The number of non-zero coefficients b_{ii} is equal to the rank of the quadratic from q_ϕ, and so, depends only on ϕ. After possibly reordering the basis, we can suppose that the first p coefficients are positive, the next $r-p$ are negative, and the remaining $n-r$ are zero. Then one has

$$b_{11} = \beta_1^2,\dots,b_{1p} = \beta_p^2, \quad b_{p+1,p+1} = -\beta_{p+1}^2,\dots,b_{r,r} = -\beta_r^2, \quad b_{r+1,r+1} = \cdots = a_{n,n} = 0.$$

[1] 1814 – 1897

for appropriate $\beta_1,\ldots,\beta_r \in \mathbb{R}$, which can be taken to be positive. Then, with respect to the basis

$$\mathbf{c}_1 = \mathbf{b}_1/\beta_1,\ldots,\mathbf{c}_r = \mathbf{b}_r/\beta_r, \quad \mathbf{c}_{r+1} = \mathbf{b}_{r+1},\ldots,\mathbf{c}_n = \mathbf{b}_n,$$

the matrix of ϕ is that given in (H.3), an so the quadratic form q_ϕ is

$$q_\phi(\mathbf{v}(x_1,\ldots,x_n)) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{r^2}.$$

for each $\mathbf{v} = x_1\mathbf{c}_1 + \cdots + x_n\mathbf{e}_n \in \mathbb{V}$.

There remains only to prove that p depends only on ϕ and not on the particular basis $\mathcal{B}$. Suppose that with respect to another basis $\mathcal{B}' = (\mathbf{c}_1,\ldots,\mathbf{c}_n)$, we have

$$q_\phi = z_1^2 + \cdots + z_t^2 - z_{t+1}^2 - \cdots - z_n^2,$$

for each $\mathbf{v} = z_1\mathbf{c}_1,\ldots,z_2\mathbf{c}_n$, for some integer $t \leq r$. We need to show that $t = p$. If $p \neq t$ we can suppose that $t < p$. Consider the subspaces of $\mathbb{V}$:

$$S = \langle \mathbf{b}_1,\ldots,\mathbf{b}_p\rangle \quad \text{and} \quad T = \langle \mathbf{c}_{t+1},\ldots,\mathbf{c}_n\rangle.$$

Since $\dim(S) + \dim(T) > n$ it follows from Grassman's formula that $S \cap T \neq \{0\}$, and so there is a $\mathbf{v} \in S \cap T$ with $\mathbf{v} \neq 0$. Then,

$$\mathbf{v} = x_1\mathbf{b}_1 + \cdots + x_p\mathbf{b}_p = z_{t+1}\mathbf{c}_{t+1} + \cdots + z_n\mathbf{c}_n.$$

Since $\mathbf{v} \neq 0$ it follows, using the expression of q_ϕ relative to $\mathcal{B}$, that

$$q_\phi(\mathbf{v}) = x_1^2 + \cdots + x_p^2 > 0$$

and, using the expression of q_ϕ relative to $\mathcal{B}'$, we have

$$q_\phi(\mathbf{v}) = -z_{t+1}^2 - \cdots - z_r^2 < 0.$$

This is clearly a contradiction, so $t = p$. □

Theorem H.8. (Sylvester's Theorem - matrix formulation) Let A be a symmetric matrix of rank r. Then there is an integer p depending only on A, and an invertible matrix P such that

$$P^T A P = \begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where 0 stands for zero matrices of appropriate sizes.

Proof. Let ϕ be the symmetric bilinear form associated to the matrix A with respect to some basis $\mathcal{B}'$ of $\mathbb{V}$. Then $\mathrm{G}_{\mathcal{B}'}(\phi) = A$. By Theorem H.7 there is a basis $\mathcal{B}$ and an integer p depending only on ϕ such that

$$P^T A P = \mathrm{M}_{BB'}^T \mathrm{G}_{\mathcal{B}'}(\phi)\mathrm{M}_{B'B} = \mathrm{G}_{\mathcal{B}}(\phi) = \begin{bmatrix} I_p & 0 & 0 \\ 0 & -I_{r-p} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

where $P = \mathrm{M}_{B'B}$. Since p depends only on ϕ, it does not depend on $\mathcal{B}$ and $\mathcal{B}'$, i.e. it depends only on A and the claim follows. □

Corollary H.9. Let ϕ be a positive definite symmetric bilinear form on an n-dimensional vector space $\mathbb{V}$. There is a basis $\mathcal{B}$ of $\mathbb{V}$ such that $\mathrm{G}_{\mathcal{B}}(\phi) = I_n$.

Proof. By Theorem H.7 there is a basis $\mathcal{B} = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ such that the Gram matrix of ϕ has the form (H.3). Since ϕ is positive definite, i.e. $\phi(\mathbf{v}, \mathbf{v}) > 0$ for all non-zero vectors $\mathbf{v}$, its Gram matrix must equal the identity matrix I_n (otherwise $\langle \mathbf{b}_n, \mathbf{b}_n \rangle$ equals 0 or -1, a contradiction). $\square$

Corollary H.10. Assume that the dimension of the space $\mathbb{V}$ of geometric vectors is n and assume that a unit segment has been chosen. Then, the properties in Proposition 4.15 define the scalar product. More concretely, there is a unique positive definite symmetric bilinear form ϕ which recognizes right angles and unit lengths.

Proof. The first three properties in Proposition 4.15, assert that the scalar product is a positive definite symmetric bilinear form on the space $\mathbb{V}^n$ of geometric vectors. By Corollary H.9 there is a basis such that the Gram matrix of the scalar product equals I_n. Therefore, by property (SP4), $\mathcal{B}$ is an orthonormal basis.

Let ϕ be another positive definite symmetric bilinear form. Since ϕ is symmetric, the Gram matrix $\mathrm{G}_{\mathcal{B}}(\phi)$ is symmetric. If ϕ satisfies property (SP4), i.e. if it recognizes right angles and unit lengths, then we must have $\mathrm{G}_{\mathcal{B}}(\phi) = I_n$ since $\mathcal{B}$ is orthonormal. Hence ϕ is the scalar product. $\square$

H.2 Isometric diagonalization

In essence, the spectral theorem says that a symmetric linear map or a symmetric bilinear form can be diagonalized with orthogonal transformations. In order to make this statement precise, we first put side by side some facts about linear maps and bilinear forms, and make precise what we mean by a 'symmetric' linear map. Symmetric bilinear forms were discussed in the previous section H.2.

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

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

and to the bilinear form

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

We say that ϕ is the *linear map associated to M in the basis $\mathcal{B}$* and ψ is the *bilinear form 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 form $\psi : \mathbb{V}^n \times \mathbb{V}^n \to \mathbb{R}$ we associate to it the Gram matrix $\mathrm{G}_{\mathcal{B}}(\psi)$ (see Section H.1). Then

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

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

$$\mathrm{M}_{\mathcal{B}',\mathcal{B}'}(\phi) = \mathrm{M}_{\mathcal{B},\mathcal{B}'}^{-1} \cdot \mathrm{M}_{\mathcal{B},\mathcal{B}}(\phi) \cdot \mathrm{M}_{\mathcal{B},\mathcal{B}'} \quad \text{and} \quad \mathrm{G}_{\mathcal{B}'}(\psi) = \mathrm{M}_{\mathcal{B},\mathcal{B}'}^T \cdot \mathrm{G}_{\mathcal{B}}(\psi) \cdot \mathrm{M}_{\mathcal{B},\mathcal{B}'} \tag{H.4}$$

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

In this setting we add the following assumptions: we consider the scalar product $\langle \_,\_\rangle$ on $\mathbb{V}^n$, we let $\mathcal{B}$ be an orthonormal basis and we assume that M is a symmetric matrix, i.e. $M = M^T$. Then, we let ϕ be the linear map associated to M in the basis $\mathcal{B}$ and we let ψ be the bilinear form associated to M in the basis $\mathcal{B}$. Then, since the gram matrix of the scalar product is I_n we have

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

Since M is symmetric, we also have

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

and therefore

$$\langle \phi(\mathbf{v}), \mathbf{w}\rangle = \psi(\mathbf{v},\mathbf{w}) = \langle \mathbf{v}, \phi(\mathbf{w})\rangle. \tag{H.5}$$

Definition H.11. A linear map $\phi : \mathbb{V}^n \to \mathbb{V}^n$ is called *symmetric (relative to the scalar product)* if (H.5) holds for all $\mathbf{v},\mathbf{w} \in \mathbb{V}^n$.

The proof of the Spectral Theorem uses the concept of orthogonal complement to a vector.

Definition H.12. 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

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

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, if $\mathbf{v}$ is non-zero, $\mathbf{v}^\perp$ is an $(n-1)$-dimensional vector subspace of $\mathbb{V}^n$.

The proof of the Spectral Theorem also uses the following linear algebra fact.

Lemma H.13. The characteristic polynomial of a symmetric matrix $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ has only real roots. Equivalently, the eigenvalues of a symmetric linear map are all real.

Proof. The equivalence of the two statements is obvious, by considering the linear map associated to M. We take the proof of the first claim from [19, p.232].

We can consider M as a matrix over $\mathbb{C}$ (that is, with complex entries), and so we may view ϕ as a linear map $\mathbb{C}^n \to \mathbb{C}^n$. Let $\lambda \in \mathbb{C}$ be a root of the characteristic polynomial of M, and let $\mathbf{x}(x_1,\ldots,x_n) \in \mathbb{C}^n$ be a corresponding eigenvector. Then

$$M\mathbf{x} = \lambda\mathbf{x}.$$

Taking the complex conjugates of both sides gives

$$M\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}.$$

Consider the scalar $\overline{\mathbf{x}}^T M\mathbf{x}$. Writing it in two different ways using the above equations gives

$$\overline{\mathbf{x}}^T M\mathbf{x} = \overline{\mathbf{x}}^T(M\mathbf{x}) = \overline{\mathbf{x}}^T(\lambda\mathbf{x}) = \lambda\overline{\mathbf{x}}^T\mathbf{x} \tag{H.6}$$

$$\overline{\mathbf{x}}^T M\mathbf{x} = (\overline{\mathbf{x}}^T M)\mathbf{x} = (M\overline{\mathbf{x}})^T\mathbf{x} = (\overline{\lambda}\overline{\mathbf{x}})^T\mathbf{x} = \overline{\lambda}\overline{\mathbf{x}}^T\mathbf{x} \tag{H.7}$$

Note that

$$\overline{\mathbf{x}}^T\mathbf{x} = \overline{x_1}x_1 + \ldots \overline{x_n}x_n$$

is a strictly positive real number, since $\mathbf{x} \neq 0$. We can therefore deduce from (H.6) and (H.7) that $\overline{\lambda} = \lambda$, that is that λ is real. □

Theorem H.14 (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.

Equivalently, let $\psi : \mathbb{V}^n \times \mathbb{V}^n \to \mathbb{R}$ be a symmetric bilinear form. Then, there exists an orthonormal basis $\mathcal{B}$ such that $\mathrm{G}_{\mathcal{B}}(\psi)$ is a diagonal matrix.

Proof. We take the proof of the first claim from [19, p.232]. The proof is by induction on $n = \dim(\mathbb{V}^n)$. If $n = 1$ there is nothing to prove. Suppose therefore that $n \geq 2$, and that the theorem holds for spaces of dimension $n-1$. Since ϕ is symmetric, it has only real eigenvalues, by Lemma H.13. Thus ϕ has an eigenvalue λ; let $\mathbf{e}_1$ be a corresponding eigenvector, which we can take to be of length 1. Let $\mathbb{U} = \mathbf{e}_1^{\perp}$, the orthogonal complement to $\mathbf{e}_1$. For each $\mathbf{u} \in \mathbb{U}$,

$$\langle \phi(\mathbf{u}), \mathbf{e}_1 \rangle = \langle \mathbf{u}, \phi(\mathbf{e}_1) \rangle = \langle \mathbf{u}, \lambda \mathbf{e}_1 \rangle = \lambda \langle \mathbf{u}, \mathbf{e}_1 \rangle = \lambda \cdot 0 = 0,$$

and so $\phi(\mathbf{u}) \in \mathbb{U}$, that is ϕ induces a map $\phi_{\mathbb{U}} : \mathbb{U} \to \mathbb{U}$. Since $\phi_{\mathbb{U}}(\mathbf{u}) = \phi(\mathbf{u})$ for every $\mathbf{u} \in \mathbb{U}$, the map ϕ is a symmetric linear map. By the inductive hypothesis, $\mathbb{U}$ has an orthonormal basis $(\mathbf{e}_2, \ldots, \mathbf{e}_n)$ which diagonalizes $\phi_{\mathbb{U}}$. Thus $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ is an orthonormal basis of $\mathbb{V}$ which diagonalizes ϕ.

For the second claim we use (H.5). Let ϕ be the symmetric linear map associated to the Gram matrix $\mathrm{G}_{\mathcal{B}'}(\psi)$ of ψ for some basis $\mathcal{B}'$. Let $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ be the diagonalizing orthonormal basis for ϕ obtained in the first part of the proof. By construction, it consists of eigenvectors of ϕ. Then, by (H.5), we have

$$\psi(\mathbf{e}_i, \mathbf{e}_j) = \langle \phi(\mathbf{e}_i), \mathbf{e}_j \rangle = \langle \lambda_i \mathbf{e}_i, \mathbf{e}_j \rangle = \lambda_i \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \begin{cases} \lambda_i & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

hence $\mathcal{B}$ is a diagonalizing basis for ψ. □

Theorem H.15 (Spectral Theorem - matrix formulation)**.** Let $M \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ be a symmetric matrix. There exists an orthogonal matrix P such that

$$P^{-1}MP = P^T MP = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

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

Proof. Let $\mathcal{B}'$ be an orthonormal basis of $\mathbb{V}^n$. Let ϕ be the linear map associated to M in the basis $\mathcal{B}'$. By the proof of Theorem H.14, there is an orthonormal basis $\mathcal{B}$ such that

$$\mathrm{M}_{\mathcal{B},\mathcal{B}}(\phi) = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of ϕ, and therefore of M. Then we may choose $P = \mathrm{M}_{\mathcal{B}',\mathcal{B}}$ since

$$P^{-1}MP = \mathrm{M}_{\mathcal{B}',\mathcal{B}}^{-1} \mathrm{M}_{\mathcal{B}',\mathcal{B}'}(\phi) \mathrm{M}_{\mathcal{B}',\mathcal{B}} = \mathrm{M}_{\mathcal{B},\mathcal{B}}(\phi)$$

by (H.4). Moreover, since $\mathcal{B}$ and $\mathcal{B}'$ are both orthogonal, it follows that P is an orthogonal matrix, i.e. $P^{-1} = P^T$ and the proof is finished. Indeed, since $\mathcal{B}'$ is orthonormal and since $P = \mathrm{M}_{\mathcal{B}',\mathcal{B}}$ is the matrix with columns consisting of the components of the basis vectors in $\mathcal{B}$ (which is an orthonormal basis), we have $P^T \cdot P = 1$, hence $P^{-1} = P^T$. $\square$

APPENDIX I

Trigonometric functions

Let $\mathbf{a}$ and $\mathbf{b}$ be two vectors in $\mathbb{V}^2$. In Section 4.1 we noticed that there is a bijection between the set $\mathbb{W}$ of unoriented angles $\measuredangle(\mathbf{a},\mathbf{b})$ and the semicircle $\mathbb{S}^{\frac{1}{2}}$ as well as a bijection between the set $\mathbb{W}_{\text{or}}$ of oriented angles $\measuredangle_{\text{or}}(\mathbf{a},\mathbf{b})$ and the unit circle $\mathbb{S}^1$. It is clearly uncomfortable to do calculations in this setting. For a more comfortable manipulation, one attaches numbers to angles (we parametrize angles). The standard way to do this is to define a measure of an angle, a numerical value which we may use instead of describing the angle as two rays emanating from a point. The standard measure is that of radians, which can be introduced as follows

Definition I.1. Let $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$ be two non-zero vectors in $\mathbb{V}^2$. The *interior of the angle* $\measuredangle AOB$ is the set of points P with the property that there are scalars $\alpha,\beta \geq 0$ such that $\overrightarrow{OP} = \alpha\mathbf{a}+\beta\mathbf{b}$. The *sector* $\mathcal{S}(AOB)$ is the set of points P in the interior of the angle and in the interior of the circle of radius 1 centered in O, i.e. $d(O,P)\leq 1$.

The *measure of the angle* $\measuredangle(\mathbf{a},\mathbf{b})$, denoted by $m(\measuredangle(\mathbf{a},\mathbf{b}))$, is

$$m(\measuredangle(\mathbf{a},\mathbf{b})) = \begin{cases} 2\cdot\text{Area}(\mathcal{S}(AOB)) & \text{if } \mathbf{a} \text{ and } \mathbf{b} \text{ are linearly independent} \\ 0 & \text{if } \mathbf{a} \text{ and } \mathbf{b} \text{ have the same direction} \\ \pi & \text{if } \mathbf{a} \text{ and } \mathbf{b} \text{ have opposite direction} \end{cases}$$

where π is by definition the area of a unit circle. The definition does not depend on the representatives.

Theorem I.2. The properties of the area function translate into properties of angles. For two non-zero vectors $\mathbf{a}$ and $\mathbf{b}$ we have

1. $0 \leq m(\measuredangle(\mathbf{a},\mathbf{b})) \leq \pi$
2. If $\mathbf{v}$ is between $\mathbf{a}$ and $\mathbf{b}$, then $m(\measuredangle(\mathbf{a},\mathbf{v}))+m(\measuredangle(\mathbf{v},\mathbf{b})) = m(\measuredangle(\mathbf{a},\mathbf{b}))$.
3. For any $\theta\in[0,\pi]$ there is a vector $\mathbf{c}\in\mathbb{V}$ such that $m(\measuredangle(\mathbf{a},\mathbf{c})) = \theta$

4. $m(\measuredangle(\mathbf{a},\mathbf{b})) = m(\measuredangle(\mathbf{c},\mathbf{d}))$ if and only if $\frac{\langle \mathbf{a},\mathbf{b}\rangle}{|\mathbf{a}|\cdot|\mathbf{b}|} = \frac{\langle \mathbf{c},\mathbf{d}\rangle}{|\mathbf{c}|\cdot|\mathbf{d}|}$.
5. $m(\measuredangle(\mathbf{a},\mathbf{b})) = m(\measuredangle(\mathbf{b},\mathbf{a}))$.
6. $m(\measuredangle(\mathbf{a},\mathbf{b})) + m(\measuredangle(\mathbf{b},-\mathbf{a})) = \pi$.
7. $m(\measuredangle(\mathbf{a},\mathbf{b})) = 0$ if and only if $\mathbf{b}$ and $\mathbf{a}$ have the same direction.
8. If $\mathbf{a} \perp \mathbf{b}$, then $m(\measuredangle(\mathbf{a},\mathbf{b})) = \frac{\pi}{2}$.
9. If $\langle \mathbf{a},\mathbf{b}\rangle > 0$, then $0 \le m(\measuredangle(\mathbf{a},\mathbf{b})) < \frac{\pi}{2}$ and we say that the angle is *obtuse*.
10. If $\langle \mathbf{a},\mathbf{b}\rangle < 0$, then $\frac{\pi}{2} < m(\measuredangle(\mathbf{a},\mathbf{b})) \le \pi$ and we say that the angle is *acute*.
11. In particular, $\mathbf{a} \perp \mathbf{b}$ if and only if $m(\measuredangle(\mathbf{a},\mathbf{b})) = \frac{\pi}{2}$.

Definition I.3. The *measure of the angle* $\measuredangle_{\mathrm{or}}(\mathbf{a},\mathbf{b})$, denoted by $m(\measuredangle_{\mathrm{or}}(\mathbf{a},\mathbf{b}))$, is

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

Proposition I.4. The measure of an oriented angle satisfies the following properties

1. $-\pi < m(\measuredangle_{\mathrm{or}}(\mathbf{a},\mathbf{b})) \le \pi$
2. $0 < m(\measuredangle_{\mathrm{or}}(\mathbf{a},\mathbf{b})) < \pi$ if and only if $(\mathbf{a},\mathbf{b})$ is right oriented.
3. $m(\measuredangle_{\mathrm{or}}(\mathbf{a},\mathbf{J}(\mathbf{a}))) = \frac{\pi}{2}$
4. $m(\measuredangle_{\mathrm{or}}(\mathbf{a},\mathbf{b})) = -m(\measuredangle_{\mathrm{or}}(\mathbf{b},\mathbf{a}))$

These definitions of measures of angles give the standard bijections

$$\mathbb{W} \leftrightarrow [0,\pi] \quad \text{and} \quad \mathbb{W}_{\mathrm{or}} \leftrightarrow (-\pi,\pi]. \tag{I.1}$$

This allows us to view $\sin, \cos, \tan$ as functions $(-\pi,\pi] \to \mathbb{R}$. Notice also that if θ is an oriented angle then the sign is no longer positive and the sine function changes the sign while the cosine function doesn't, i.e. we have the sign rules that we are familiar with:

$$\sin(-\theta) = -\sin(\theta) \quad \text{and} \quad \cos(-\theta) = \cos(\theta).$$

In what follows we work out main properties of the trigonometric functions. The bijections (I.1) allow us to simplify notation: we write $\measuredangle_{\mathrm{or}}(\mathbf{a},\mathbf{b})$ instead of $m(\measuredangle_{\mathrm{or}}(\mathbf{a},\mathbf{b}))$.

Proposition I.5. 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 $\theta = \measuredangle_{\mathrm{or}}(\mathbf{a},\mathbf{b})$.

Proposition I.6. If $\mathbf{a}$ 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 $\mathbf{b}$ 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 an 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 I.7. If $\mathbf{a},\mathbf{b},\mathbf{c}$ are three non-zero vectors 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 I.8. For all $\alpha,\beta \in \mathbb{R}$, we have

$$\begin{aligned}\cos(\alpha+\beta) &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)\\ \sin(\alpha+\beta) &= \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta).\end{aligned}$$

Proposition I.9. The following half angle formulas hold true

$$\begin{aligned}\tan\left(\tfrac{\alpha}{2}\right) &= \tfrac{\sin(\alpha)}{1+\cos(\alpha)}\\ 2\cos^2\left(\tfrac{\alpha}{2}\right) &= 1+\cos\alpha\\ 2\sin^2\left(\tfrac{\alpha}{2}\right) &= 1-\cos\alpha\end{aligned}$$

Proposition I.10. For all $\theta \geq 0$, we have

$$\sin(\theta) \leq \theta \leq \tan(\theta).$$

Corollary I.11. For all θ, we have

$$0 \leq 1 - \cos(\theta) \leq \theta^2.$$

Proposition I.12. We have the following limits

$$\lim_{\theta\to 0}\left(\frac{1-\cos(\theta)}{\theta}\right) = 0 \quad \text{and} \quad \lim_{\theta\to 0}\left(\frac{\sin(\theta)}{\theta}\right) = 1.$$

Corollary I.13. The sine and cosine functions are differentiable and the derivatives are

$$\cos'(\theta) = -\sin(\theta) \quad \text{and} \quad \sin'(\theta) = \cos(\theta).$$

APPENDIX J

Some classical theorems in Euclidean geometry

In this section we collect classical theorems in Euclidean geometry. The theorems in Appendix F are purely affine but they also belong to this list.

Theorem J.1 (Euclid's first theorem)**.** Let ABC be a rectangular triangle with $\overrightarrow{BA} \perp \overrightarrow{BC}$, and let $H \in [AC]$ be such that $\overrightarrow{BH} \perp \overrightarrow{AC}$. Then

$$d(A,B)^2 = d(A,C) \cdot d(A,H).$$

Theorem J.2 (Pythagoras' theorem)**.** Let the triangle ABC be rectangular in B. Then

$$d(A,C)^2 = d(A,B)^2 + d(A,C)^2.$$

Theorem J.3 (Theorem of heights)**.** Let the triangle ABC be rectangular in B and let H be the foot of the altitude on the leg $[AC]$. Then

$$d(B,H)^2 = d(A,H)^2 \cdot d(H,C)^2.$$

Theorem J.4 (Thales' circle theorem)**.** If A, B, C are points on a circle, then $[AC]$ is a diameter if and only if $\measuredangle ABC$ is a right angle.

Theorem J.5 (Central angle theorem)**.** Let ABC be a triangle and consider the circumcircle with center O. We say that the angle $\varphi = \measuredangle(ACB)$ is subtended by the chord $[AB]$ and that the angle $\psi = \measuredangle AOB$ is the corresponding central angle. We have $\psi = 2\varphi$.

Theorem J.6 (Ptolemy's theorem)**.** A quadrilateral is called *inscribed* if its vertices lie on a circle. A quadrilateral is inscribed if and only if the sum of the products of the lengths of its two pairs of opposite sides is equal to the product of the lengths of its diagonals.

Theorem J.7 (Euler's line)**.** The orthocenter H, the centroid G and the circumcenter U of a triangle are collinear. Moreover show that

$$\overrightarrow{HG} = 2\overrightarrow{GU}.$$

Theorem J.8 (Feuerbach Circle)**.** In every triangle the three midpoints of the sides, the three base points of the altitudes, and the midpoints of the three altitude sections touching the vertexes lie on a circle. Moreover, the center of this circle lies on Euler's line half way between the orthocenter and the circumcenter.

APPENDIX K

Quaternions and rotations

K.1 Algebraic considerations

Some of the aspects considered here are also covered in [**?**, Section 4.4].

Definition K.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
$$\begin{aligned} pq &= (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) + (a_1b_2 + a_2b_1 + c_1d_2 - c_2d_1)\mathbf{i} \\ &+ (a_1c_2 + a_2c_1 - b_1d_2 + b_2d_1)\mathbf{j} + (a_1d_2 + a_2d_1 + b_1c_2 - b_2c_1)\mathbf{k} \end{aligned} \tag{K.1}$$

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 K.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 K.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 K.4. A quaternion is purely imaginary if and only if its square is real and non-positive.

Definition K.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 K.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. $\mathfrak{Im}(p) = \frac{1}{2}(p - \overline{p})$

K.2 Geometric considerations

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 consider 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 K.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 K.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 K.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 K.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 quaternions $\mathfrak{Im}(\mathbb{H}) = \mathbb{R}\mathbf{i}+\mathbb{R}\mathbf{j}+\mathbb{R}\mathbf{k}$.

Proposition K.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

$$q_1q_2 = (a_1+v_1)(a_2+v_2) = a_1a_2 - \langle v_1,v_2\rangle + a_2v_1 + a_1v_2 + v_1\times v_2. \tag{K.2}$$

Proposition K.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] C. Alsina, B. Nelsen, A mathematical space odyssey. Solid geometry in the 21st century. The Dolciani Mathematical Expositions, 50. Mathematical Association of America, Washington, DC, 2015.

[2] J. Benítez, A unified proof of Ceva and Menelaus' theorems using projective geometry. J. Geom. Graph. 11 (2007), no. 1, 39–44.

[3] M. de Berg, O. Cheong, M. van Kreveld, M. Overmars, Computational geometry. Algorithms and applications. Third edition. Springer-Verlag, Berlin, 2008.

[4] M. Berger, Geometry I. Translated from the 1977 French original by M. Cole and S. Levy. Fourth printing of the 1987 English translation Universitext. Springer-Verlag, Berlin, 2009.

[5] P.A. Blaga, Geometrie liniară. Cu un ochi către grafica pe calculator, Vol. I, Presa Universitară Clujeană, 2022.

[6] C.B. Boyer, A history of mathematics. Second edition. Edited and with a preface by Uta C. Merzbach. John Wiley & Sons, Inc., New York, 1989.

[7] H.S.M. Coxeter, Regular polytopes. Methuen, London, 1948.

[8] H.S.M. Coxeter, S.L. Greitzer, Geometry revisited. New Mathematical Library, 19. Random House, Inc., New York, 1967.

[9] S. Crivei, Basic Linear Algebra. Presa Universitară Clujeană, 2022.

[10] Euclid. The thirteen books of Euclid's Elements translated from the text of Heiberg. Vol. I: Introduction and Books I, II. Vol. II: Books III–IX. Vol. III: Books X–XIII and Appendix. Translated with introduction and commentary by Thomas L. Heath. 2nd ed. Dover Publications, Inc., New York, 1956.

[11] A.E. Fekete, Real linear algebra. Monographs and Textbooks in Pure and Applied Mathematics, 91. Marcel Dekker, Inc., New York, 1985.

[12] I. Grattan-Guinness, The search for mathematical roots, 1870–1940. Logics, set theories and the foundations of mathematics from Cantor through Russell to Gödel. Princeton Paperbacks. Princeton University Press, Princeton, NJ, 2000.

[13] J. Hefferon, Linear Algebra, 2020.

[14] D. Hilbert, Foundations of Geometry. Second english edition. Translated by Unger L. from 10th ed. Revised and Enlarged by Bernays P. Open Court, Illinois, 1971.

[15] L. Hodgkin, A history of mathematics. From Mesopotamia to modernity. Oxford University Press, Oxford, 2005.

[16] M. Nechita, Lecture notes for mathematical analysis, 2025.

[17] C.C. Pugh, Real mathematical analysis. Second edition. Undergraduate Texts in Mathematics. Springer, Cham, 2015.

[18] W. Rudin, Principles of mathematical analysis. Third edition. International Series in Pure and Applied Mathematics. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976.

[19] E. Sernesi, Linear Algebra. A geometric approach. Translated by Montaldi J., Chapman & Hall/CRC, 1993.

[20] Handbook of discrete and computational geometry. Second edition. Edited by Jacob E. Goodman and Joseph O'Rourke. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2004.

[21] Lean 4, Programming Language and Theorem Prover, `https://lean-lang.org/`

[22] The mathlib Community: The Lean mathematical library. In: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs. p. 367–381. CPP 2020, New York, NY, USA (2020).

[23] A mathlib overview, `https://leanprover-community.github.io/mathlib-overview.html`