EDWIN MONTUFAR

A TREATISE ON AFFINE GEOMETRY (DRAFT)

CONTENTS

1 — Affine Space

1.1. Characterization of an Affine Space

Definition 1.1.1. An ***affine space*** is either the degenerate space reduced to the empty set or a triple $(\mathcal{A}, V, +)$ consisting of a nonempty set of points $\mathcal{A}$, a vector space V of free vectors , and an action $+ : \mathcal{A} \times V \to \mathcal{A}$, satisfying the following conditions:

1. $A + 0_V = A$ for every $A \in \mathcal{A}$ and 0_V is the zero vector in V. **(A1)**
2. $(A + u) + v = A + (u \hat{+} v)$, for every $a \in \mathcal{A}$ and $u, v \in V$. **(A2)**
3. For any two points $A, B \in \mathcal{A}$ there exists a unique free vector $u \in V$ such that $A + u = B$. **(A3)**

Remark 1.1.2. We will see later on that we can denote this vector u by $\overrightarrow{AB}$ or $(B - A)$, thereby also writing $B = A + \overrightarrow{AB}$. The, notation $\hat{+}$, serves to distinguish itself as vector addition, whereas $+$, serves to indicate addition between a point and a vector. This is done to prevent us from overloading the addition symbol; however, when the context is clear, we shall avoid making any such distinction. Also, note that it doesn't make sense to add or subtract points in an affine space. This is why it's preferred to write u as $\overrightarrow{AB}$ instead of $(B - A)$, though it's acceptable as long as we remember that $(B - A)$ is a free vector belonging to V and not a point in $\mathcal{A}$, plus using the latter form will facilitate algebraic manipulations if we use following associativity conditions:

1. $P + (Q - P) = Q$
2. $(P - Q) \hat{+} u = (P + u) - Q$.

Where $P, Q \in \mathcal{A}$ and $u \in V$.

1.2. The Identification of a Point Space with a Vector Space

Theorem 1.2.1. *Let $(\mathcal{A}, V, +)$ be an affine space, then the sets $\mathcal{A}$ and V are isomorphic, that is, $\mathcal{A} \cong V$.*

Proof. Let $A, B \in \mathcal{A}$ be arbitrary points. For every $A \in \mathcal{A}$ consider the mapping $\varphi_A : V \to \mathcal{A}$ defined by $\varphi_A(v) = A + v$ and the mapping $\psi_A : \mathcal{A} \to V$ defined by $\psi_A(B) = \overrightarrow{AB}$. Then

$$(\varphi_A \circ \psi_A)(B) = \varphi_A(\psi_A(B)) = \varphi_A\left(\overrightarrow{AB}\right) \tag{1.1}$$

$$= A + \overrightarrow{AB} = B. \tag{1.2}$$

and

$$(\psi_A \circ \varphi_A)(v) = \psi_A(\varphi_A(v)) = \psi_A(A + v) \tag{1.3}$$

$$= \overrightarrow{A(A + v)} = v \tag{1.4}$$

Thus $\varphi \circ \psi : \mathcal{A} \to \mathcal{A}$ is the identity from $\mathcal{A} \to \mathcal{A}$, whereas $\psi \circ \varphi : V \to V$ is the identity from V to V , that is, $\varphi \circ \psi = \mathrm{id}_{\mathcal{A}}$ and $\psi \circ \varphi = \mathrm{id}_V$. Consequently, φ and ψ are invertible, thereby establishings bijections between $\mathcal{A}$ and V. Hence $\mathcal{A}$ and V are isomorphic in the category of sets. ■

Remark 1.2.2. We see that $\mathcal{A}$ is identified with V via the mapping ψ_A by designating A as the origin in $\mathcal{A}$. Hence, ψ_A identifies a point B with its position vector $\overrightarrow{AB}$. Conversely, we identify V with $\mathcal{A}$ via the mapping φ_A by translation of A via the free vectors in V. Essentially every point in $\mathcal{A}$ can be resolved via translation of a chosen point by all free vectors in V. As a consequence of these results, we have an isomorphism (in the category of sets) between $\mathcal{A}$ and V, hence we write $\mathcal{A} \cong V$. Moreover, since V is isomorphic to $\mathbb{F}^n$(that is, $V \cong \mathbb{F}^n$) via the coordinate map and a choice of basis, it follows via transitivity of isomorphisms that $\mathcal{A} \cong \mathbb{F}^n$, as well. Hence points in $\mathcal{A}$ can be identified with n-tuples taking entries from an algebraic field $\mathbb{F}$; for example $\mathbb{F} = \mathbb{R}$. This enables us to coordinatize points in $\mathcal{A}$, more on that later.

Definition 1.2.3. The ***dimension*** of an affine space $(\mathcal{A}, V, +)$ is the dimension of its associated vector space V.

Some Useful Lemmas

Lemma 1.2.4 (*Cancellation Law*). *Let $(\mathcal{A}, V, +)$ be an affine space. If for arbitrary $A \in \mathcal{A}$ and $u, v \in V$, we have $A + u = A + v$, then $u = v$.*

Proof. Let $B \in \mathcal{A}$ be such that $A + u = B$. Then, by **A3**, u must be unique. Now, since $A + u = A + v$, it must follow by uniqueness of u that $u = v$. ■

Lemma 1.2.5 (Chasle's Relation). *Let A, B, C be points in an affine space $(\mathcal{A}, V, +)\mathcal{A}$ such that $C = A + \overrightarrow{AC}$, and $B = A + \overrightarrow{AB}$, and $C = B + \overrightarrow{BC}$. Then, $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$.*

Proof. Indeed, if $C = B + \overrightarrow{BC}$, then

$$\begin{aligned} C &= \left(A + \overrightarrow{AB}\right) + \overrightarrow{BC} \\ &= A + \left(\overrightarrow{AB} + \overrightarrow{BC}\right) \qquad \text{(By A2)} \end{aligned}$$

Now, by A3, the unique vector $u \in V$ such that

$$A + u = C = A + \overrightarrow{AC}$$

is $u = \left(\overrightarrow{AB} + \overrightarrow{BC}\right)$, whereby uniqueness, it follows that $\left(\overrightarrow{AB} + \overrightarrow{BC}\right) = \overrightarrow{AC}$, as required. ■

Corollary 1.2.6. *Consider the affine space $(\mathcal{A}, V, +)$. Then, for $A, B, C, D \in \mathcal{A}$, it follows that:*

1. $\overrightarrow{AA} = 0_V$.
2. $\overrightarrow{BA} = -\overrightarrow{AB}$
3. $\overrightarrow{AB} = \overrightarrow{DC}$ if and only if $\overrightarrow{BC} = \overrightarrow{AD}$.

Proof. (1) In view of Theorem 1.2.1, $A = A + \overrightarrow{AA}$, then by **A1**, it follows that $A = A + 0_V$, whereby the cancellation law $\overrightarrow{AA} = 0_V$. (2) Let $C = A$ in Chasle's relation, then

$$\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = 0_V.$$

Hence, it follows that $\overrightarrow{BA} = -\overrightarrow{AB}$. (3) By Chasle's relation, it follows that

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC} \quad \text{and} \quad \overrightarrow{AD} + \overrightarrow{DC} = \overrightarrow{AC}.$$

Hence, $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AD} + \overrightarrow{DC}$. Now, $\left(\overrightarrow{AB} - \overrightarrow{DC}\right) = 0_V = \left(\overrightarrow{BC} - \overrightarrow{AD}\right)$ implies that $\overrightarrow{AB} = \overrightarrow{DC}$ and $\overrightarrow{BC} = \overrightarrow{AD}$. ■

1.3. Vectorization of an Affine Space

In view of Theorem 1.2.1, we know that we can identify $\mathcal{A}$ with a vector space V by making a choice of origin. Now, let us take this a step further to define the notion of an affine frame or affine basis. Just as vector spaces have bases, we can extend this notion of a basis to affine space provided that we have a choice of origin. Recall that bases vectors must be linearly independent to span the vector space they belong to. In a similar manner we must establish point-wise independence or affine independence amongst points in $\mathcal{A}$. We proceed to define such a notion.

Definition 1.3.1. Given an affine space $(\mathcal{A}, V, +)$, a family $(A_i)_{i\in I}$ of points in $\mathcal{A}$ is ***affinely independent*** or ***point-wise independent*** if the family of vectors $\left(\overrightarrow{A_iA_j}\right)_{j\in I\setminus\{I\}}$ is linearly independent for some $i \in I$.

Remark 1.3.2. By making a fixed choice of $i \in I$, we are implicitly designating the point corresponding to this index as a base point or origin per se'.

Example 1.3.3. Suppose we have three non-collinear points A, B, C in the plane $\mathcal{A}$ making a triangle $\triangle ABC$, then these points are affinely independent because the family $\left(\overrightarrow{AA}, \overrightarrow{AB}, \overrightarrow{AC}\right)$ is linearly independent.

Points are Independent of the Choice of Origin

Lemma 1.3.4. *Let $(\mathcal{A}, V, +)$ be an affine space, and let $(A_i)_{i\in I}$ be a family of points in $\mathcal{A}$ and let $(\lambda_i)_{i\in I}$be a family of scalars in $\mathbb{F}$. For any two points $\Omega_1, \Omega_2 \in \mathcal{A}$, the following properties are true:*

1. If $\sum_{i\in I} \lambda_i = 1$, then
$$\Omega_1 + \sum_{i\in I} \lambda_i \overrightarrow{\Omega_1 A_i} = \Omega_2 + \sum_{i\in I} \lambda_i \overrightarrow{\Omega_2 A_i}. \tag{1.5}$$
2. If $\sum_{i\in I} \lambda_i = 0$, then
$$\sum_{i\in I} \lambda_i \overrightarrow{\Omega_1 A_i} = \sum_{i\in I} \lambda_i \overrightarrow{\Omega_2 A_i}. \tag{1.6}$$

Proof. (1) By Chasle's relation, it follows that $\overrightarrow{\Omega_1\Omega_2} + \overrightarrow{\Omega_2 A_i} = \overrightarrow{\Omega_1 A_i}$, thus

$$\begin{aligned}
\Omega_1 + \sum_{i\in I} \lambda_i \overrightarrow{\Omega_1 A_i} = \Omega_1 + \sum_{i\in I} \lambda_i \left(\overrightarrow{\Omega_1\Omega_2} + \overrightarrow{\Omega_2 A_i}\right) \\
&= \Omega_1 + \left[\sum_{i\in I} \lambda_i\right] \overrightarrow{\Omega_1\Omega_2} + \sum_{i\in I} \lambda_i \overrightarrow{\Omega_2 A_i} \\
&= \Omega_1 + \overrightarrow{\Omega_1\Omega_2} + \sum_{i\in I} \lambda_i \overrightarrow{\Omega_2 A_i} \qquad \because \sum_{i\in I} \lambda_i = 1 \\
&= \Omega_2 + \sum_{i\in I} \lambda_i \overrightarrow{\Omega_2 A_i} \qquad \because \Omega_1 + \overrightarrow{\Omega_1\Omega_2} = \Omega_2.
\end{aligned}$$

Now, for (2), it follows that $\overrightarrow{\Omega_1\Omega_2} + \overrightarrow{\Omega_2 A_i} = \overrightarrow{\Omega_1 A_i}$, now

$$\begin{aligned}
\sum_{i\in I} \lambda_i \overrightarrow{\Omega_1 A_i} &= \sum_{i\in I} \lambda_i \left(\overrightarrow{\Omega_1\Omega_2} + \overrightarrow{\Omega_2 A_i}\right) \\
&= \left[\sum_{i\in I} \lambda_i\right] \overrightarrow{\Omega_1\Omega_2} + \sum_{i\in I} \lambda_i \overrightarrow{\Omega_2 A_i} = \sum_{i\in I} \lambda_i \overrightarrow{\Omega_2 A_i} \qquad \because \sum_{i\in I} \lambda_i = 0
\end{aligned}$$

In conlusion, we see that for a point $X \in \mathcal{A}$, $X = \Omega + \sum_{i\in I} \lambda_i \overrightarrow{\Omega A_i}$ is independent of the choice of origin Ω in $\mathcal{A}$ whenever $\sum \lambda_i = 1$. ■

Affine Combinations and Barycentres

Definition 1.3.5. Let $(A_i)_{i\in I}$ be a family of points in an affine space $\mathcal{A}$, then the point

$$X = \sum_{i\in I} \lambda_i A_i \tag{1.7}$$

where $(\lambda_i)_{i\in I}$ is a family of scalars such that $\sum_{i\in I} \lambda_i = 1$ is an ***affine combination*** or ***barycentre*** of the family of weighted points $((A_i, \lambda_i))_{i\in I}$. Moreover, given an arbitrary choice of origin Ω for $\mathcal{A}$, then X is the unique point such that

$$X = \Omega + \sum_{i\in I} \lambda_i \overrightarrow{\Omega A_i}, \tag{1.8}$$

and

$$\overrightarrow{\Omega X} = \sum_{i\in I} \lambda_i \overrightarrow{\Omega A_i} \tag{1.9}$$

for all $\Omega \in \mathcal{A}$. We see that this sum is a ***linear combination*** of the vectors $\overrightarrow{\Omega A_i}$.

Remark 1.3.6. In view of the previous lemma, the point X is independent of the choice of origin Ω in $\mathcal{A}$. To see how we can resolve, 1.7 from 1.8, we observe that

$$X = \Omega + \sum_{i\in I} \lambda_i (A_i - \Omega) = \Omega - \sum_{i\in I} \lambda_i \Omega + \sum_{i\in I} \lambda_i A_i \tag{1.10}$$

$$= \underbrace{\left(1 - \sum_{i\in I} \lambda_i\right)}_{=0} \Omega + \sum_{i\in I} \lambda_i A_i = \sum_{i\in I} \lambda_i A_i \tag{1.11}$$

Since $(1 - \sum_{i\in I} \lambda_i) + \sum_{i\in I} \lambda_i = 1$.

Example 1.3.7. Let A and B be points in the Euclidean plane $\mathcal{E}^2$, then a ***line in the Euclidean plane,*** ℓ_{AB} is the set of points

$$\left\{P \in \mathcal{E}^2 : P = aA + bB,\ a + b = 1\right\}, \tag{1.12}$$

hence such points are expressed as affine combinations of A and B. Another way to characterize such a line is as follows: the set of points $P \in \mathcal{E}^2$, where

$$P = (1-t)A + tB, \quad t \in \mathbb{R}. \tag{1.13}$$

Notation 1.3.8. We shall often denote ℓ_{AB} by (AB), and to say that $P \in \ell_{AB}$, we shall often express this as (PAB). Also, the intersection of two lines ℓ and m will oft be expressed as (ℓm).

Definition 1.3.9. A ***convex combination*** of the points A_i in an affine space $\mathcal{A}$, where $i = 1, 2, \ldots, n$ is essentially an affine combination with the added condition that the coefficients be non-negative, that is,

$$\sum_{i=1}^{n} \alpha_i A_i, \quad \text{where } (\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1) \tag{1.14}$$

and $a_i \geq 0$ for all $i = 1, \ldots, n$.

Example 1.3.10. Let the ℓ_{AB} be a line in $\mathcal{E}^2$, then the segment joining A and B is comprised of all points $P \in \ell_{AB}$ that are convex combinations of A and B in $\mathcal{E}^2$, where

$$P = (1-t)A + tB, \quad t \in [0,1]. \tag{1.15}$$

We shall denote the ***segment of the line*** ℓ_{AB}, comprised of points $P \in \ell_{AB}$, lying between A and B, with A and B inclusive, as $[AB]$.

Definition 1.3.11. The ***affine span or hull*** of an affine space $\mathcal{A}$ is the set of all affine combinations of points from $\mathcal{A}$,that is

$$\operatorname{aff}(\mathcal{A}) := \left\{ \sum_{i \in I} \lambda_i A_i \mid A_i \in \mathcal{A}, \ \sum_{i \in I} \lambda_i = 1 \right\} \tag{1.16}$$

Affine Subspaces

Definition 1.3.12. Let $(\mathcal{A}, V, +)$ be an affine space, and $\mathcal{B}$ a subset of $\mathcal{A}$, with associated vector space W as a subspace of V. We define an ***affine subspace*** *or* ***linear variety*** of $(\mathcal{A}, V, +)$ to be a triplet $(\mathcal{B}, W, +)$ which satisfies the following equivalent statements:

1. $\underline{\mathcal{B}}$ is closed under affine combinations, that is, for every family of weighted points $((A_i, \lambda_i))_{i \in I}$ in $\mathcal{B}$ where $\sum_{i \in I} \lambda_i = 1$, the affine combination $\sum_{i \in I} \lambda_i A_i$ belongs to $\mathcal{B}$.
2. For all $A \in \mathcal{B}$ and for all vectors $w \in W$:
   (a) We have $A + w \in \mathcal{B}$, that is closure under addition of points in $\mathcal{B}$ with vectors in W.
   (b) The map $\Phi_w : \mathcal{B} \to \mathcal{B}$ given by $\Phi_w(A) = A + w$, satisfies:
      i. $A + 0_V = A$, for all $A \in \mathcal{B}$, where $0_V \in W$;
      ii. $A + (v + w) = (A + v) + w$, for all $A \in \mathcal{B}$ and $v, w \in W$;
      iii. For each pair $A, B \in \mathcal{B}$, we have $\overrightarrow{AB} \in W$.
3. For any two points $A, B \in \mathcal{B}$, the line

$$\ell := \{A + \lambda(B - A) = (1-\lambda)A + \lambda B \mid \lambda \in \mathbb{F}\} \tag{1.17}$$

   is contained in $\mathcal{B}$.

Remark 1.3.13. Notice that this is similar to how subspaces of vector spaces must satisfy closure under linear combinations of vectors. *For 2. we need to only verify that (a) and (iii) are satisfied, since (a) essentially implies conditions (i) and (ii); reason being, the translation action (+) of the vector space V over $\mathcal{A}$ restricts to an action of the subspace W over $\mathcal{B}$.*

Theorem 1.3.14. *The statements above are equivalent.*

Proof. $1 \Rightarrow 3$ We must verify an arbitrary element of ℓ is in $\mathcal{B}$. Indeed, let $X \in \ell$ be chosen such that

$$X = (1-t)\,A + tB$$

where $A, B \in \mathcal{B}$. Since $\mathcal{B}$ is closed under affine combinations, there exist families of weighted points $((A_i,\lambda_i))_{i=1}^{k}$ and $\left((B_j,\mu_j)\right)_{j=1}^{l}$in $\mathcal{B}$, such that the barycenters

$$A = \sum_{i=1}^{k} \lambda_i A_i, \quad B = \sum_{j=1}^{l} \mu_j B_j \tag{1.18}$$

are in $\mathcal{B}$, where $\sum \lambda_i = 1$ and $\sum \mu_j = 1$. Then

$$X = (1-t)\,A + tB = (1-t)\sum_{i=1}^{k} \lambda_i A_i + t\sum_{j=1}^{l} \mu_j B_j \tag{1.19}$$

$$= \sum_{i=1}^{k} (1-t)\,\lambda_i A_i + \sum_{j=1}^{l} t\mu_j B_j \tag{1.20}$$

Now, observe that

$$\sum_{i=1}^{k} (1-t)\,\lambda_i + \sum_{j=1}^{l} t\mu_j = (1-t)\sum_{i=1}^{k} \lambda_i + t\sum_{j=1}^{l} \mu_j \tag{1.21}$$

$$= (1-t) + t = 1, \tag{1.22}$$

because $\sum \lambda_i = 1$, and $\sum \mu_j = 1$. Thus, X is an affine combination of points is $\mathcal{B}$, and so $\ell \subseteq \mathcal{B}$, as desired.$3 \Rightarrow 1$ Suppose that for any two points $A, B \in \mathcal{B}$, the line

$$\ell := \{A + \lambda\,(B-A) = (1-\lambda)\,A + \lambda B \mid \lambda \in \mathbb{F}\} \tag{1.23}$$

is contained in $\mathcal{B}$, then for arbitrary $X \in \ell$, it is of the form

$$X = (1-\lambda)\,A + \lambda B \Leftrightarrow A + \lambda\,(B-A) \Leftrightarrow X = A + \lambda\overrightarrow{AB} \tag{1.24}$$

whereby definition, $X = A + \lambda\overrightarrow{AB} \in \mathcal{B}$ is the barycenter or affine combination of the weighted points $((A, 1-\lambda), (B,\lambda))$, and $(1-\lambda) + \lambda = 1$. Hence, $(\mathcal{B}, W, +)$ is closed under affine combinations. $3 \Rightarrow 2$. Suppose the line ℓ as described above is in $\mathcal{B}$, then for $X \in \ell$, we have $X = A + \lambda\overrightarrow{AB} \in \mathcal{B}$ since $\ell \subseteq \mathcal{B}$. Therefore closure under addition of points with vectors is satisfied. For $w \in W \subseteq V$, it follows that $X = A + w = A + \lambda\overrightarrow{AB}$, whereby the cancellation law, $w = \lambda\overrightarrow{AB}$. Since W is a subspace of V and therefore closed under scalar multiplication of vectors, it follows that $\overrightarrow{AB} \in W$. $2 \Rightarrow 3$. Suppose that for arbitrary $A \in \mathcal{B}$ and $w \in W$, that $A + w \in \mathcal{B}$ and that $\overrightarrow{AB} \in W$ for the pair $A, B \in \mathcal{B}$. Then, $A + \overrightarrow{AB} = B$ is in $\mathcal{B}$.Now, it remains to be shown that $(1-\lambda)\,A + \lambda B$ is in $\mathcal{B}$. Indeed, we have

$$(1-\lambda)\,A + \lambda B = A + \lambda\overrightarrow{AB} \tag{1.25}$$

Now, since W is a subspace and $\overrightarrow{AB} \in W$, it follows that $\lambda\overrightarrow{AB} \in W$. Thus, by our hypothesis, it follows that $A + \lambda\overrightarrow{AB} \in \mathcal{B}$. Since A, B were arbitrary, it follows that the line ℓ as defined above is contained in $\mathcal{B}$, that is, $\ell \subseteq \mathcal{B}$. ∎

Theorem 1.3.15. *If $(\mathcal{B}, W, +)$ is an affine subspace of $(\mathcal{A}, V, +)$, then*

$$\mathcal{B} = A + W$$

where $W \subseteq V$, and $A \in \mathcal{B}$.

Proof. We see that ■

A very useful result

Lemma 1.3.16. *For $n \geq 3$, the barycenter of n weighted points can be obtained by repeated computations of barycenters of two weighted points.*

Proof. We prove this by induction on n and make the assumption that the weights are taken from a scalar field of characteristic zero. **Base case:** for $n = 3$, suppose we have a family of two weighted points $((A_i, \lambda_i))_{i=1}^{3}$ in an affine space $\mathcal{A}$, then we have the barycenter

$$X = \lambda_1 A_1 + \lambda_2 A_2 + \lambda_3 A_3 \tag{1.26}$$

which we want to express in the form:

$$X = \mu \left(\frac{\lambda_1 A_1 + \lambda_2 A_2}{\mu} \right) + (1 - \mu) A_3 \tag{1.27}$$

Indeed, since $\lambda_1 + \lambda_2 + \lambda_3 = 1$, we set $\mu = \lambda_1 + \lambda_2 = 1 - \lambda_3$, thus,

$$X = (1 - \lambda_3) \left(\frac{\lambda_1 A_1 + \lambda_2 A_2}{1 - \lambda_3} \right) + \lambda_3 A_3 \tag{1.28}$$

$$= (1 - \lambda_3) \underbrace{\left(\frac{\lambda_1 A_1 + \lambda_2 A_2}{\lambda_1 + \lambda_2} \right)}_{\overline{A}} + \lambda_3 A_3. \tag{1.29}$$

Hence, X is the barycentre of two weighted points, namely $\left(\left(\overline{A}, 1 - \lambda_3 \right), (A_3, \lambda_3) \right)$. **Induction hypothesis:** suppose the lemma holds true for $i < n$. Then, for the family of $n \geq 4$ weighted points $((A_i, \lambda_i))_{i=1}^{n} \in \mathcal{A}$ where $\lambda_i \neq 0$ for all $i = 1, \ldots, n$, we have

$$X = \sum_{i=1}^{n} \lambda_i A_i \tag{1.30}$$

Then, as in 1.27, we want to express this as a barycenter of two weighted points:

$$X = \mu \underbrace{\left(\frac{\sum_{i=1}^{n-1} \lambda_i A_i}{\mu} \right)}_{\overline{A}} + (1 - \mu) A_n \tag{1.31}$$

So, set $\mu = \sum_{i=1}^{n-1} \lambda_i = 1 - \lambda_n$ and observe that $\overline{A} = \sum_{i=1}^{n-1} (\lambda_1/\mu) A_i$ is the barycenter of $n - 1$ weighted points, whereby the induction hypothesis, such a barycentre may be obtained by repeated computations of barycentres of two weighted points. Thus, we deduce that the barycenter in 1.30 may be ultimately obtained by barycenter of two weighted points, namely $\left(\left(\overline{A}, 1 - \lambda_n \right), (A_n, \lambda_n) \right)$. So we conclude that the barycenter of n weighted points can be obtained by repeated computations of barycentres of two weighted points. ■

Theorem 1.3.17. *Let $(\mathcal{A}, V, +)$ be an affine space, then:*

1. A nonempty subset $\mathcal{W}$ of $\mathcal{A}$ is an affine subspace if and only if for every point $\Omega \in \mathcal{W}$, the set
$$W_\Omega = \left\{\overrightarrow{\Omega X} \mid X \in \mathcal{W}\right\}$$
is a subspace of V. Consequently, $\mathcal{W} = \Omega + W_\Omega$.

2. The set $W = \left\{\overrightarrow{AB} \mid A, B \in \mathcal{W}\right\}$ is a subspace of V and $W_\Omega = W$ for all $\Omega \in \mathcal{A}$. Hence, $\mathcal{W} = \Omega + W$.

3. For any subspace W of V and any $\Omega \in \mathcal{A}$, the set $\mathcal{W} = \Omega + W$ is an affine subspace.

Proof. (1) Let $\Omega \in \mathcal{W}$ be chosen arbitrarily as the origin of $\mathcal{W}$. Then, by Corrolary[], we know that $\overrightarrow{\Omega\Omega} = 0_V$ is in W. Moreover, if $\mathcal{W}$ is an affine subspace of $\mathcal{A}$, then it must satisfy closure under affine combinations. Indeed, if $X \in \mathcal{B}$, then by eqn 1.9, the barycentre X is the unique point such that

$$\overrightarrow{\Omega X} = \sum_{i=1}^{n} \lambda_i \overrightarrow{\Omega A_i} \Leftrightarrow X = \Omega + \sum_{i=1}^{n} \lambda_i \overrightarrow{\Omega A_i}.$$

In which case, it follows by definition that X is the barycentre or affine combination of the weighted points $\Omega, A_1, \ldots A_n \in \mathcal{W}$, expressed as

$$\sum_{i=1}^{n} \lambda_i A_i + \underbrace{\left(1 - \sum_{i=1}^{n} \lambda_i\right)}_{=0} \Omega = \sum_{i=1}^{n} \lambda_i A_i,$$

where $\sum_{i=1}^{n} \lambda_i = 1$. Hence, W_Ω satisfies closure under linear combinations, implying that it's a subspace of V. Consequently, $\mathcal{W} = \Omega + W_\Omega$. (2) Now, since $W_\Omega = \left\{\overrightarrow{\Omega X} \mid X \in \mathcal{W}\right\}$ and $W = \left\{\overrightarrow{AB} \mid A, B \in \mathcal{W}\right\}$it follows that $W_A \subseteq W$ because $\Omega \in \mathcal{W}$. And so $W_\Omega = W$ for every choice of origin Ω in $\mathcal{W}$. Consequently, $\mathcal{W} = \Omega + W$. (3) Now, suppose $\mathcal{W} = \Omega + W$, where W is a subspace of V. Let $\Omega \in \mathcal{A}$, be chosen abitrarily, then for every family of weighted points $((\Omega + w_i, \lambda_i))_{1 \le i \le n}$ where $w_i \in W$ and $\sum_i \lambda_i = 1$, the affine combination

$$X = \Omega +$$

■

Definition 1.3.18. The point Ω, serving as a choice of origin for $\mathcal{A}$, and a basis $\xi = (e_i)_{i \in I}$ for its associated vector space V, constitute a ***frame of reference*** or ***affine frame*** in the space $\mathcal{A}$, denoted by $(O; e_1, e_2, \ldots, e_n)$. The n-tuple of the form $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ associated with a point $A \in \mathcal{A}$ is called the ***coordinates*** of the point A with respect to the associated frame of reference in $\mathcal{A}$. Moreover, the basis vectors e_i can be expressed as

$$e_i = \overrightarrow{OA_i}$$

thus, a frame of reference in $\mathcal{A}$, may also be expressed as the the collection of $n + 1$ points,

$$(O, A_1, \ldots, A_n) \tag{1.32}$$

for which the vectors $\overrightarrow{OA_1}, \ldots, \overrightarrow{OA_n}$ form a basis of V. Furthermore, a point $P \in \mathcal{A}$, may be expressed as

$$P = O + (P - O) = O + \overrightarrow{OP} = O + \sum_{i=1}^{n} x_i e_i,$$

where $x = \overrightarrow{OP} = \sum_{i=1}^{n} x_i e_i$ is a vector in V expressed as a unique linear combination of the basis vectors in ξ.

Affine Ratios: a Precursor to the Main Collinearity Lemmas

Corollary 1.3.19. *There is a unique line ℓ passing through a pair of distinct points A and B of an affine space $(\mathcal{A}, V, +)$.*

Proof. Since points A and B are distinct, then they are point-wise independent, that is, if A and B are distinct, then $\lambda\overrightarrow{AB} = \lambda (B - A) = \mathbf{0}$ if and only $\lambda = 0$, since $B \neq A$. Therefore the line containing A and B must coincide with a set of points $P \in \mathcal{A}$ for which

$$\overrightarrow{AP} \in \operatorname{span}\left(\overrightarrow{AB}\right)$$

or $\overrightarrow{BP} \in \operatorname{span}\left(\overrightarrow{AB}\right)$. Hence, if $\overrightarrow{AP} = \lambda\overrightarrow{AB}$ and $\overrightarrow{AP'} = \beta\overrightarrow{AB}$ for $P' \in \mathcal{A}$, then $\overrightarrow{PP'} = (\beta - \lambda)\overrightarrow{AB}$. Hence we see that $\overrightarrow{PP'} \in \operatorname{span}\left(\overrightarrow{AB}\right)$ implying that ℓ is a line. ■

Lemma 1.3.20. *Let A and B be two distinct points in an affine space. Then the points P on the unique line ℓ_{AB} determined by A and B are those for which*

$$P = aA + bB$$

and $a + b = 1$. Such a representation is unique for each individual point P on the line ℓ_{AB}.

Proof. Suppose a point P lies on the line determined by two distinct points A and B in an affine space $\mathcal{A}$, then

$$\overrightarrow{AP} \in \operatorname{span}\left(\overrightarrow{AB}\right),$$

In view of our previous lemma or the distinctness of A and B, we know that $\overrightarrow{AB}$ is linearly independent, and so for some unique λ, we have

$$\overrightarrow{AP} = \lambda\overrightarrow{AB}$$

Thus, $P = A + \lambda (B - A) \Leftrightarrow P = (1 - \lambda) A + \lambda B$. So, set $a = 1 - \lambda$ and $b = \lambda$, and we see that

$$P = aA + bB$$

where $a + b = 1$. Moreover, this representation is unique, for if $\overrightarrow{AP} = \lambda\overrightarrow{AB}$ and $\overrightarrow{AP} = \beta\overrightarrow{AB}$, then

$$(\lambda - \beta)\overrightarrow{AB} = \mathbf{0}$$

if and only if $(\lambda - \beta) = 0 \Leftrightarrow \lambda = \beta$, by virtue of $\overrightarrow{AB}$ being linearly independent. ■

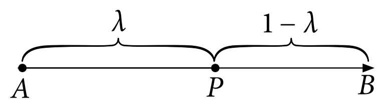


Figure 1.3.1. Three collinear points A, P, B: the number λ may viewed as the ratio between the vectors $\overrightarrow{AP}$ and $\overrightarrow{AB}$.

In view of these results, we see that if P belongs to a line determined by two distinct points A and B in $\mathcal{A}$, then P is the unique point such that

$$P = A + \lambda\overrightarrow{AB} \Leftrightarrow P = (1-\lambda)\,A + \lambda B \Leftrightarrow P = aA + bB \tag{1.33}$$

where $a + b = 1$. We must bear in mind that P could just as easily extend past B, in which case $\lambda > 1$. Also, if P is behind A, then $\lambda < 1$. Now, if P is within the segment $[A, B]$, then $0 < \lambda < 1$, therefore we can express P as a convex combination of A and B. Also, the number λ is really just the ratio of the distance between A and P to the distance between A and B. However, notice that the notion of distance or length hasn't been defined anywhere as of yet. So, instead we may view λ as the ratio between the vectors $\overrightarrow{AP}$ and $\overrightarrow{AB}$, denoted by

$$\overrightarrow{AP} : \overrightarrow{AB} \quad \text{or} \quad \frac{P-A}{B-A}$$

as used by Tondeur [?]. So, let us proceed to define the notion of an affine ratio.

Vector Ratios	Ratio Quantity
$(A,P,B) = \overrightarrow{AP} : \overrightarrow{AB}$	λ
$(A,B,P) = \overrightarrow{AB} : \overrightarrow{AP}$	$\frac{1}{\lambda}$
$(B,A,P) = \overrightarrow{BA} : \overrightarrow{BP}$	$\frac{1}{1-\lambda}$
$(B,P,A) = \overrightarrow{BP} : \overrightarrow{BA}$	$1-\lambda$
$(P,A,B) = \overrightarrow{PA} : \overrightarrow{PB}$	$\frac{\lambda}{\lambda-1}$
$(P,B,A) = \overrightarrow{PB} : \overrightarrow{PA}$	$\frac{\lambda-1}{\lambda}$

Table 1.3.1: Affine Ratios for Three Collinear Points A, P, B.

Definition 1.3.21. Let $A, P, B \in \mathcal{A}$ be three distinct collinear points. Then the ***affine ratio*** of these points is the unique scalar $\lambda = (A, P, B) \in \mathbb{R}$ such that

$$\overrightarrow{AP} = \lambda\overrightarrow{PB}.$$

Hence, $\overrightarrow{AP} : \overrightarrow{AB} = \lambda : 1$ or $\frac{P-A}{B-A} = \lambda$.

Remark 1.3.22. Now, permuting the last two points in the triple (A, P, B), thereby obtaining (A, B, P), it follows that

$$(A, B, P) = \frac{1}{\lambda},$$

which is the reciprocal of λ. Now, there are a total of 6 permutations of the letters A, P, and B. So going through the remaining four triples, we have

$$\begin{aligned} (B,A,P) &= \frac{-1}{-(1-\lambda)} = \frac{1}{1-\lambda} \\ (B,P,A) &= 1-\lambda \end{aligned}$$

and

$$\begin{aligned} (P,A,B) &= \frac{-\lambda}{1-\lambda} = \frac{\lambda}{\lambda-1} \\ (P,B,A) &= \frac{\lambda-1}{\lambda}. \end{aligned}$$

In each case, we see that keeping the first letter the same and permuting the last points, we obtain a reciprocal quantity. All these affine ratios can be using Figure 1.3.1. We tabulate these quantities for future reference in Table [].

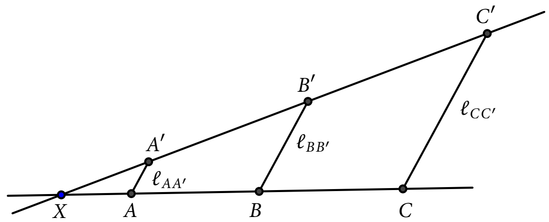


Figure 1.3.2. Thale's Theorem

An Application of Affine Ratios: Thale's Theorem

Collinearity Conditions: the Main Collinearity Lemma

1.4. Coordinatization of Affine Structures and Vector Spaces

The purpose of this section is to make the distinction between vectors and points very clear. It was mentioned in the remark to Theorem 1.2.1 that points and vectors can be identified with n-tuples whose entries are derived from a scalar field $\mathbb{F}$. For example, let $\mathbb{F} = \mathbb{R}$, and consider the affine space $(\mathcal{A}, V, +)$ whose associated vector space has dimension n. Then, let $\xi = (e_1, e_2, \ldots, e_n)$ be a basis for V over $\mathbb{F} = \mathbb{R}$. Then every vector $u, v \in V$ can be expressed as a unique linear combination of the vectors in basis ξ, that is,

$$\begin{aligned} u &= u_1 e_1 + u_2 e_2 + \cdots + u_n e_n \\ v &= v_1 e_1 + v_2 e_2 + \cdots + v_n e_n, \quad u_i, v_i \in \mathbb{R}. \end{aligned}$$

Then, the coordinate representations of vectors u and v with respect to basis ξ are given by

$$[u]_\xi = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_2 \end{bmatrix}, \quad [v]_\xi = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

respectively. Now, taking any linear combination $\lambda u + \mu v$ of these two vectors using scalars $\mu, \lambda \in \mathbb{R}$, the coordinate respresentation of the vector resulting from this combination is

$$[\lambda u + \mu v]_\xi = \begin{bmatrix} \lambda u_1 + \mu v_1 \\ \lambda u_2 + \mu u_2 \\ \vdots \\ \lambda u_n + \mu u_n \end{bmatrix}.$$

So, suppose we decide on using another basis for V , that is, $\xi' = (e'_i)_{i=1}^n$. Then, the change of basis matrix, changing ξ' coordinates to ξ coordinates is given by

$$P_{\xi \leftarrow \xi'} = \big[[e'_1]_\xi \mid \quad [e'_2]_\xi \mid \quad \cdots \quad \mid [e'_n]_\xi \big]$$

whose columns are the coordinate representations of the basis vectors in ξ' with respect to basis ξ. Furthermore, we have u, v as unique linear combinations of the vectors in ξ and ξ', respectively.

$$\sum_{i=1}^n u_i e_i = u = \sum_{i=1}^n u'_i e'_i \quad \text{and} \quad \sum_{i=1}^n v_i e_i = v = \sum_{i=1}^n v'_i e'_i.$$

This indicates that **linear combinations of vectors are basis independent**. Furthermore, it follows that

$$[u]_{\xi} = P_{\xi \leftarrow \xi'} [u]_{\xi'} \quad \text{and} \quad [v]_{\xi} = P_{\xi \leftarrow \xi'} [v]_{\xi}.$$

Also,

$$[u]_{\xi'} = P^{-1}_{\xi \leftarrow \xi'} [u]_{\xi} \quad \text{and} \quad [v]_{\xi'} = P^{-1}_{\xi \leftarrow \xi'} [v]_{\xi}.$$

Now, by linearity of the coordinate map with respect to a basis, it follows that

$$\begin{aligned} [\lambda u + \mu v]_{\xi'} &= \lambda [u]_{\xi'} + \mu [v]_{\xi'} \\ &= \lambda P^{-1}_{\xi \leftarrow \xi'} [u]_{\xi} + \mu P^{-1}_{\xi \leftarrow \xi'} [v]_{\xi} \\ &= P^{-1}_{\xi \leftarrow \xi'} [\lambda u + \mu v]_{\xi}. \end{aligned}$$

Now, let us examine the set of points $\mathcal{A}$, and suppose we identify its origin O with the n-tuple $O = (0, 0, \ldots, 0)$. Now, ξ be the standard basis in $\mathbb{R}^n$

Theorem 1.4.1. *.All the vectors of the form $\overrightarrow{AB}$, where A and B are arbitrarily chosen in an affine space $\mathcal{A}$ make a subspace V' of V.*

Proof. Indeed, $\overrightarrow{AA} = 0_V$ is in V'. Now, let C and D be arbitrarily chosen in $\mathcal{A}$, by (2) it follows that there exists a unique free vectors $u, v \in V'$ such that

$$A + u = B$$

and

$$C + v = D.$$

Moreover, $u = \overrightarrow{BC}$ and $v = \overrightarrow{CD}$. Then, by (2), it follows again that there exits a unique free vector in V' such that

$$B + w = D.$$

Indeed, let $w = \overrightarrow{BD}$, which is equivalent to $u + v = \overrightarrow{BC} + \overrightarrow{CD}$ by Chasle's Identity. Now, let $c \in \mathbb{F}$, then

Then $u + v \in V'$. and let $x = \overrightarrow{AB}$ and $y = \overrightarrow{CD}$ be vectors in V', then by (2) of Definition 1.3.1, there exists a unique point $P \in \mathcal{A}$, such that $\overrightarrow{BP} = \overrightarrow{CD}$, in which case, it follows by condition (1), that

$$\overrightarrow{AP} = \overrightarrow{AB} + \overrightarrow{BP} = \overrightarrow{AB} + \overrightarrow{CD} = x + y$$

hence, closure under addition is satisfied. Now, for any vector $x = \overrightarrow{AB}$, we know know by condition (3) that there exists a unique point $C \in \mathcal{A}$ for which

$$\overrightarrow{AC} = \alpha \overrightarrow{AB} = \alpha x.$$

Thus, closure under scalar multiplication is satisfied, and the proof is complete. ∎

Definition 1.4.2. A translation of an affine space $(\mathcal{A}, V)$ by a vector v is a mapping $\tau_v : \mathcal{A} \to \mathcal{A}$ which assigns to the point A the point B such that $v = \overrightarrow{AB}$. Hence, $\tau_v(A) = B$, where $\overrightarrow{AB} = v$.

$$\begin{tikzcd} \mathcal{A} \arrow[rr, "\tau_v"] & & \mathcal{A} \\ & V \arrow[ul, "\varphi^{-1}"] \arrow[ur, "\varphi'"'] & \end{tikzcd}$$

The diagram commutes since $\tau_v = \varphi' \circ \varphi^{-1}$.

Vectorization of an Affine Space

Definition 1.4.3. The point O, serving as a choice of origin for $\mathcal{A}$, and a basis $\xi = \{e_1, e_2, \ldots, e_n\}$ for its associated vector space V, constitute a ***frame of reference*** or ***affine frame*** in the space $\mathcal{A}$, denoted by $(O; e_1, e_2, \ldots, e_n)$. The n-tuple of the form $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ associated with a point $A \in \mathcal{A}$ is called the ***coordinates*** of the point A with respect to the associated frame of reference in $\mathcal{A}$. Moreover, the basis vectors e_i can be expressed as

$$e_i = \overrightarrow{OA_i}$$

thus, a frame of reference in $\mathcal{A}$, may also be expressed as the the collection of $n+1$ points,

$$(O, A_1, \ldots, A_n) \tag{1.34}$$

for which the vectors $\overrightarrow{OA_1}, \ldots, \overrightarrow{OA_n}$ form a basis of V. Furthermore, a point $P \in \mathcal{A}$, may be expressed as

$$P = O + (P - O) = O + \overrightarrow{OP} = O + \sum_{i=1}^{n} x_i e_i,$$

where $x = \overrightarrow{OP} = \sum_{i=1}^{n} x_i e_i$ is a vector in V expressed as a unique linear combination of the basis vectors in ξ.

Example 1.4.4. The non-collinear points A, B, C or the vertices of $\triangle ABC$ in $\mathcal{E}^2$ form an affine frame.

Remark 1.4.5. Suppose we have points $A, B \in \mathcal{A}$, with coordinates $(\alpha_1, \ldots, \alpha_n)$, and $(\beta_1, \beta_2, \ldots, \beta_n)$, then the vector $\overrightarrow{AB} \in V$ has coordinates $(\beta_1 - \alpha_1, \beta_2 - \alpha_2, \ldots, \beta_n - \alpha_n)$ with respect a basis ξ in V.

Definition 1.4.6. Let $\mathfrak{b} = (\Omega, \beta)$, be an affine basis for the affine space $(\mathcal{A},)$ where $\Omega \in \mathcal{A}$ is a choice of origin and $\beta = (e_1, e_2, \ldots, e_n)$ is the standard basis for the vector space $\mathbb{R}^n$.

1.5. Projective Completion of an Affine Space

TODO...

2 — Recurrent Themes

We are referring to points within the context of the Euclidean plane; however, bear in mind that no mention of its inner product structure has been made. This is intentional and will be reserved for later chapters.

2.1. The Theorems of Ceva and Menelaus

The motivation behind the Theorem of Menelaus is to determine a condition on the points A', B', C' on each side of triangle $\triangle ABC$ so that they may be collinear. For Ceva's theorem we're trying to determine a condition for which the lines $\ell_{AA'}, \ell_{BB'}, \ell_{CC'}$ are concurrent.

The Main Collinearity Lemma

The results of the main collinearity lemma will be used to shorten our proofs for the theorems of Menelaus and Ceva.

Lemma 2.1.1. *Let A, B, X be different points on the Euclidean plane. Then the following statements are equivalent:*

(a) *A,B,X are collinear.*

(b) *There exist $a, b \in \mathbb{R}$ such that*

$$X = \frac{aA + bB}{a + b}$$

where $a + b \neq 0$.

(c) *There exist $a, b \in \mathbb{R}$ for which the ratios:*

$$\frac{X - A}{X - B} = -\frac{b}{a} \quad \text{and} \quad \frac{X - B}{X - A} = -\frac{a}{b}.$$

(d) *There exist $a, b, x \in \mathbb{R}$ such that the equation*

$$aA + bB + xX = O$$

where $a + b + x = 0$, but $abx \neq 0$.

Proof. First, assume that $A \neq B$ and let

$$\frac{X - A}{X - B} = -\frac{b}{a} \tag{$\clubsuit$}$$

Then multiplying both sides by a and $(X - B)$, we then obtain

$$a(X - A) = -b(X - B)$$

which is equivalent to

$$(a+b)X = aA + bB. \tag{♠}$$

We should observe that $a+b \neq 0$ for it would contradict our hypothesis otherwise. Indeed, if $a+b=0$, then with $a=-b$, we have for ($\star$), the following

$$\frac{X-A}{X-B} = 1 \Leftrightarrow X-A = X-B \Leftrightarrow A = B,$$

therefore violating our hypothesis that A and B are not equal. Finally, dividing both sides of ♠ by $(a+b)$ we obtain

$$X = \frac{aA+bB}{a+b}. \tag{♡}$$

Using the same reasoning, we can deduce why the ratio $X-B/X-A = -a/b$ implies ♡ as well. Now, to see why the converse is also true, we simply work backwards with ♡ using algebra to obtain the ratios in (c). We can therefore see that statements **(b)** and **(c)** are equivalent. Now, let us assume A, B, X are distinct and that there exist $a, b, x \in \mathbb{R}$ such that

Consider the following cases where $X = A$, or $X = B$, or $X = A = B$ occurs in the ratio $X - A/X - B$. Then for each respective case we obtain $0, \infty$, and 1. These are going to be accounted for in the projective plane.

The product $abx \neq 0$ simply suggests that none of quantities involved are zero.

$$aA + bB + xX = 0 \quad \text{where } a+b+x = 0 \text{ and } abx \neq 0.$$

Set $x = -(a+b)$, then we have for the equation above $aA + bB - (a+b)X = 0$, which is equivalent to

$$aA + bB = (a+b)X.$$

Then dividing both sides by $(a+b)$, we obtain

$$X = \frac{aA+bB}{a+b}.$$

We've thus shown that **(c)** implies **(d)**. To prove the converse, simply assume **(d)** and reverse the algebra in the steps above to obtain **(c)**. Consequently, we see that **(b)**⇔**(c)** and **(c)**⇔**(d)**, and so **(b)**⇔**(d)**. The proof is complete. ■

Definition 2.1.2. A line ℓ that passes through a vertex of a triangle $\triangle ABC$ and the side opposite to it is called a **cevian**.

Theorem of Ceva. *Let A', B', C' be points on the sides of triangle $\triangle ABC$ that are opposite to vertices A, B, C respectively. The lines (cevians) $\ell_{AA'}, \ell_{BB'}, \ell_{CC'}$ are concurrent if and only if*

$$\frac{A'-B}{A'-C} \cdot \frac{B'-C}{B'-A} \cdot \frac{C'-A}{C'-B} = -1.$$

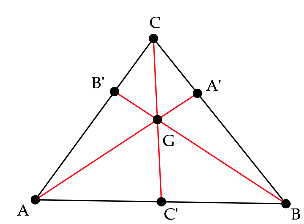


Figure 2.1.1: Ceva's Theorem

Proof. Let us suppose that the cevians $\ell_{AA'}, \ell_{BB'}, \ell_{CC'}$ concur at a point $G = aA + bB + cC$ where $a+b+c = 1$. Now, take note that $b + c \neq 0$, becuase otherwise we'd have $a = 1$, thereby implying that

$$G = A + bB + (-b)C \Leftrightarrow G - A = b(B-C)$$

meaning that $\ell_{AG} \parallel \ell_{BC}$. Furthermore, this would suggest that $G \notin \ell_{AA'}$ and so it would contradict our hypothesis. Therefore, proceeding with the notion that $b + c \neq 0$, we can write

$$G = aA + (b+c)\frac{bB+cC}{b+c} = aA + (b+c)A'.$$

Applying the main collinearity lemma to

$$A' = \frac{bB + cC}{b + c}$$

we deduce that

$$\frac{A' - B}{A' - C} = -\frac{c}{b}. \tag{$\spadesuit$}$$

With the similar reasoning as above, we deduce that

$$G = (a + b)\,\frac{aA + bB}{a + b} + cC \quad \text{and} \quad G = (a + c)\,\frac{aA + cC}{a + c} + bB$$

where $(a + b) \neq 0$ and $(a + c) \neq 0$. Thus, by virtue of the main collinearity lemma we have

$$B' = \frac{aA + cC}{a + c} \quad \text{and} \quad C' = \frac{aA + bB}{a + b}.$$

Furthermore,

$$\frac{B' - C}{B' - A} = -\frac{a}{c} \quad and \quad \frac{C' - A}{C' - B} = -\frac{b}{c}. \tag{$\clubsuit$}$$

Finallly, we can take the product of the ratios in $\clubsuit$ and $\spadesuit$, thereby obtaining

$$\frac{A' - B}{A' - C} \cdot \frac{B' - C}{B' - A} \cdot \frac{C' - A}{C' - B} = -\left(\frac{c}{b} \cdot \frac{a}{c} \cdot \frac{b}{a}\right) = -1, \tag{$\heartsuit$}$$

as desired. Conversely, let us assume that $\heartsuit$ is given and that $A' \in \ell_{BC}, B' \in \ell_{AC}$, and $C' \in \ell_{AB}$. Now, we know by hypothesis that A' is distinct from B or C, and similarly for B' and C' otherwise the left hand side of $\heartsuit$ would be undefined or 0, thereby giving rise to a contradiction. Then it follows that for some $b, c \neq 0$, we have

$$\frac{A' - B}{A' - C} = -\frac{c}{b} \quad \text{and} \quad \frac{B' - C}{B' - A} = -\frac{a}{c} \quad \text{for some } a \neq 0. \tag{$\diamondsuit$}$$

Therefore substituting these values into the left hand side of $\heartsuit$, we have

$$\left(-\frac{c}{b}\right)\left(-\frac{a}{c}\right)\frac{C' - A}{C' - B} = \frac{a}{b} \cdot \frac{C' - A}{C' - B} = -1$$

which is further equivalent to

$$\frac{C' - A}{C' - B} = -\frac{b}{a}.$$

Then, by virtue of the main collinearity lemma we see that this implies that

$$C' = \frac{aA + bB}{a + b}.$$

Applying the MCL to the ratios in $\diamondsuit$, we also obtain

$$A' = \frac{bB + cC}{b + c} \quad \text{and} \quad B' = \frac{aA + cC}{a + c}.$$

Now, let us define G as the barycenter of mass points $(a, A), (b, B), (c, C)$, then

$$G = \frac{aA + bB + cC}{a + b + c}$$

whence it follows from previous results that

$$G = \frac{aA + (b + c)\,\frac{bB + cC}{b + c}}{a + b + c} = \frac{aA + (b + c)\,A'}{a + b + c}$$

and
$$G = \frac{bB + (a+c)\frac{aA+cC}{a+c}}{a+b+c} = \frac{bB + (a+c)B'}{a+b+c}$$
as well as
$$G = \frac{cC + (a+b)\frac{aA+bB}{a+b}}{a+b+c} = \frac{cC + (a+b)C'}{a+b+c}.$$
So, by the MCL it follows that the triples (A, A', G), (B, B', G), (C, C', G) are comprised of collinear points. Furthermore, we that see that the barycenter G is common to lines $\ell_{AA'}, \ell_{BB'}, \ell_{CC'}$, thereby indicating that they concur at G. ∎

Lemma 2.1.3. *Let A, B and X be three distinct points on the Euclidean plane. Then the following statements are equivalent:*

(a) *The points A, B, X are collinear.*

(b) *There exist $a, b, x \in \mathbb{R}$ such that*
$$X = \frac{aA - bB}{a - b}$$
where $a \neq b$ and a, b are nonzero.

(c) *There exist $a, b, x \in \mathbb{R}$ such that*
$$\frac{X - A}{X - B} = \frac{b}{a} \quad \text{and} \quad \frac{X - B}{X - A} = \frac{a}{b}$$
where $a \neq b$ and $a, b \neq 0$.

Remark 2.1.4. This is simply a modification of the main collinearity lemma, where we've taken away the last result, and also replaced $+b$ with $-(-b)$, which is essentially the same thing. No proof is necessary, but this modification will be of some utility when we attempt to prove the theorem of Menelaus.

Definition 2.1.5. A point lying on the line determined by the side of a triangle is a **Menelaus point** for that given side.

Theorem of Menelaus. *Let points A', B', C' be Menelaus points corresponding to sides BC, CA, and AB of triangle $\triangle ABC$, respectively. These points are collinear if and only if the product*
$$\frac{A' - B}{A' - C} \cdot \frac{B' - C}{B' - A} \cdot \frac{C' - A}{C' - B} = 1.$$

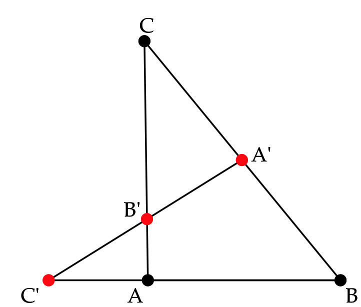


Figure 2.1.2: Theorem of Menelaus

Proof. Going forward, let us suppose that the points A', B', C' are collinear. Then by virtue of the modified main collinearity lemma, it follows that there exist $c, b \in \mathbb{R}$ such that
$$\frac{A' - B}{A' - C} = \frac{c}{b}. \quad \text{where } c \neq b \text{ and } bc \neq 0. \tag{1}$$
Similarly, there exists a unique $a \in \mathbb{R}$ such that
$$\frac{B' - C}{B' - A} = \frac{a}{c} \quad \text{where } a \neq c \neq 0. \tag{2}$$

Now, with a little algebra both the equations above can be rewritten as such

$$(b-c)\,A' = bB - cC \tag{$\clubsuit$}$$

and

$$(c-a)\,B' = cC - aA. \tag{$\spadesuit$}$$

Now, adding $\clubsuit$ and $\spadesuit$, we obtain

$$(b-c)\,A' + (c-a)\,B' = bB - aA. \tag{$\diamondsuit$}$$

Bear in mind that we've yet to see how a and b and relate. If $a = b = t$, then

$$(t-c)\,A' + (c-t)\,B' = tB - tA \Leftrightarrow (t-c)\,(A'-B') = t\,(B-A)$$

where $t - c \neq 0 \neq t$. This however would mean that $\ell_{A'B'} \parallel \ell_{AB}$, thereby giving rise to a contradiction with our hypothesis. To see why, let us recall that we've assumed that points A', B', C' are collinear, but keep in mind these are also Menelaus points of their respective sides. So, if $\ell_{A'B'} \parallel \ell_{AB}$, then there's no way can have a point $C' \in \ell_{AB}$ that is collinear with A' and B'. Now, if $a \neq b$ it is clear that $b - a \neq 0$. Furthermore, dividing $\diamondsuit$ by $b - a$, we get

$$\frac{b-c}{b-a}A' + \frac{c-a}{b-a}B' = \frac{bB-aA}{b-a}$$

which essentially means that we've obtained a point in the intersection $\ell_{A'B'} \cap \ell_{AB}$, namely C'. Thus

$$C' = \frac{bB-aA}{b-a}$$

and so by virtue of the modified mcl, it follows that

$$\frac{C'-A}{C'-B} = \frac{b}{a}. \tag{3}$$

Finally, taking the product of the ratios obtained in (1),(2), and (3), we get

$$\frac{A'-B}{A'-C} \cdot \frac{B'-C}{B'-A} \cdot \frac{C'-A}{C'-B} = \left(\frac{c}{b} \cdot \frac{a}{c} \cdot \frac{b}{a}\right) = 1, \tag{$\heartsuit$}$$

as desired. Conversely, let us assume $\heartsuit$, then substituting the right hand side of (1) and (2) into $\heartsuit$, we get

$$\frac{c}{b} \cdot \frac{a}{c} \cdot \frac{C'-A}{C'-B} = \frac{a}{b} \cdot \frac{C'-A}{C'-B} = 1.$$

Whence it follows that

$$\frac{C'-A}{C'-B} = \frac{b}{a}.$$

Furthermore, clearing denominators and distributing real coefficients we get $aC' - aA = bC' - bB$, which is equivalent to

$$(a-b)\,C' = aA - bB. \tag{✠}$$

Then, adding ✠ to $\diamondsuit$, we get

$$(b-c)\,A' + (c-a)\,B' + (a-b)\,C' = O.$$

We've established before that the coefficients in this equation are non-zero, but their sum is. So, by the last condition of the MCL it follows that points A', B', C' are collinear. ■

The Equivalence of the Theorems of Menelaus and Ceva

S

Proposition 2.1.6. *The theorems of Menelaus and Ceva are equivalent.*

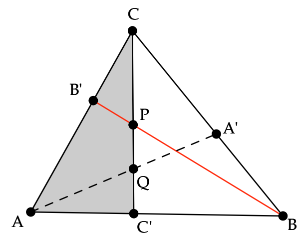


Figure 2.1.3.

In the figure above, you'll notice that the Menelaus points P and Q do not coincide. They simply don't have to since they are each chosen with respect to their corresponding triangles, namely $\triangle CAC'$ and $\triangle CBC'$.

Proof. **(Menelaus implies Ceva)** Consider[1] triangle $\triangle CAC'$ and line $\ell_{BB'}$ as demonstrated in the figure above. We denote P as the intersection of lines $\ell_{CC'}$ and $\ell_{BB'}$. Since the points B', P, B are collinear, it follows by Menelaus' Theorem that

$$\frac{B'-A}{B'-C}\cdot\frac{P-C}{P-C'}\cdot\frac{B-C'}{B-A}=1. \tag{$\spadesuit$}$$

[1] The trick to proving this implication is to apply Menelaus' Theorem twice on triangles $\triangle CAC'$ and $\triangle CBC'$, considering lines $\ell_{AA'}$ and $\ell_{BB'}$.

Now, let us consider triangle $\triangle CBC'$ and line $\ell_{AA'}$. We denote Q as the point of intersection of lines $\ell_{AA'}$ and $\ell_{CC'}$. Since the points A, Q, A' are collinear, it follows by Menelaus' Theorem that

$$\frac{A-B}{A-C'}\cdot\frac{Q-C'}{Q-C}\cdot\frac{A'-C}{A'-B}=1. \tag{$\clubsuit$}$$

Multiplying $\clubsuit$ and $\spadesuit$, we get

$$\left(\frac{B'-A}{B'-C}\cdot\frac{P-C}{P-C'}\cdot\frac{B-C'}{B-A}\right)\left(\frac{A-B}{A-C'}\cdot\frac{Q-C'}{Q-C}\cdot\frac{A'-C}{A'-B}\right)=1$$

which is equivalent to

$$\left(\frac{P-C}{P-C'}\cdot\frac{Q-C'}{Q-C}\right)\left(\frac{B'-A}{B'-C}\cdot\frac{A'-C}{A'-B}\right)\left(\frac{B-C'}{B-A}\cdot\frac{A-B}{A-C'}\right)=1.$$

Observe that

$$\left(\frac{B-C'}{B-A}\cdot\frac{A-B}{A-C'}\right)=\frac{A-B}{B-A}\cdot\frac{B-C'}{A-C'}=(-1)\left(\frac{-(C'-B)}{-(C'-A)}\right)=-\frac{C'-B}{C'-A}.$$

Consequently,

$$\left(\frac{P-C}{P-C'}\cdot\frac{Q-C'}{Q-C}\right)\left(\frac{B'-A}{B'-C}\cdot\frac{A'-C}{A'-B}\right)\left(-\frac{C'-B}{C'-A}\right)=1. \tag{$\diamond$}$$

Now, let us assume that $P=Q=(\ell_{AA'}\cap\ell_{BB'}\cap\ell_{CC'})$. This would then mean that the first factor on RHS of $\diamond$ is equal to 1, thereby giving us

$$\left(\frac{B'-A}{B'-C}\cdot\frac{A'-C}{A'-B}\right)\left(-\frac{C'-B}{C'-A}\right)=1 \tag{$\star$}$$

which is equivalent to

$$\frac{A'-C}{A'-B}\cdot\frac{B'-A}{B'-C}\cdot\frac{C'-B}{C'-A}=-1. \tag{$\heartsuit$}$$

Conversely, let us assume $\heartsuit$, then by substituting rhs of $\star$, into lhs of $\diamond$, we get

$$\left(\frac{P-C}{P-C'}\cdot\frac{Q-C'}{Q-C}\right)=1 \Leftrightarrow \frac{P-C}{P-C'}=\frac{Q-C}{Q-C'}.$$

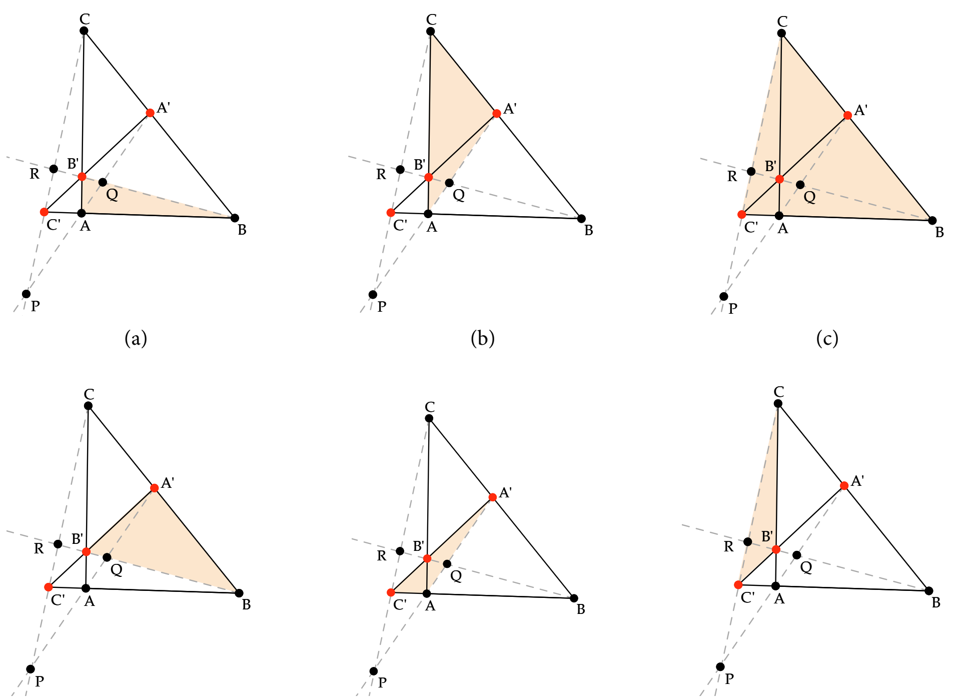
Thus, by the the modified mcl, we know there exist non-zero $x, y \in \mathbb{R}$ such that

$$P=\frac{xC'-yC}{x-y}=Q \quad \text{where } x \neq y.$$

And so, we conclude that $P=Q$ and the lines $\ell_{AA'}, \ell_{BB'}, \ell_{CC'}$ concur at the aforementioned points. ■

Table 2.1.1. Things to Consider

Triangles	Lines and Points of Concurrence	Cevian Product
$\triangle ABB'$	$A' = \ell_{BC}\cap\ell_{AQ}\cap\ell_{B'C'}$	$\frac{C-A}{C-B'}\cdot\frac{Q-B'}{Q-B}\cdot\frac{C'-B}{C'-A}=-1$
$\triangle AA'C$	$C' = \ell_{AB}\cap\ell_{CP}\cap\ell_{A'B'}$	$\frac{B-C}{B-A'}\cdot\frac{P-A'}{P-A}\cdot\frac{B'-A}{B'-C}=-1$
$\triangle C'BC$	$B' = \ell_{CA}\cap\ell_{BR}\cap\ell_{C'A'}$	$\frac{A-B}{A-C'}\cdot\frac{R-C'}{R-C}\cdot\frac{A'-C}{A'-B}=-1$
$\triangle B'BA'$	$A = \ell_{BC'}\cap\ell_{B'C}\cap\ell_{A'Q}$	$\frac{C'-B'}{C'-A'}\cdot\frac{C-A'}{C-B}\cdot\frac{Q-B}{Q-B'}=-1$
$\triangle C'AA'$	$C = \ell_{AB'}\cap\ell_{A'B}\cap\ell_{C'P}$	$\frac{B'-A'}{B'-C'}\cdot\frac{B-C'}{B-A}\cdot\frac{P-A}{P-A'}=-1$
$\triangle C'B'C$	$B = \ell_{CA'}\cap\ell_{C'A}\cap\ell_{B'R}$	$\frac{A'-C'}{A'-B'}\cdot\frac{A-B'}{A-C}\cdot\frac{R-C}{R-C'}=-1$



(a) (b) (c) (d) (e) (f)

Figure 2.1.4. Triangles to consider

Proof. **(Ceva implies Menelaus)** Consider the triangles, lines, and points of concurrence in Table 2.1.1. By application of Ceva's theorem we obtain the products in the third column of the table above. Multiplying these products, we ultimately obtain

$$\left(\underbrace{\frac{A'-C}{A'-B}}_{-}\cdot\underbrace{\frac{B'-A}{B'-C}}_{-}\cdot\underbrace{\frac{C'-B}{C'-A}}_{+}\right)^2=1,$$

from which it follows that

$$\frac{A'-C}{A'-B}\cdot\frac{B'-A}{B'-C}\cdot\frac{C'-B}{C'-A}=1,$$

as desired. ∎

Applications of Ceva's and Menelaus' Theorems

Example 2.1.7. The Medians of a triangle $\triangle ABC$ are concurrent. Let (AA'), (BB'), and (CC') denote the medians of $\triangle ABC$. Since A', B', and C' are midpoint of their respective sides for $\triangle ABC$, it follows that (draw a picture!)

$$C'-A=B-C',\quad A'-B=C-A',\quad B'-C=A-B'. \tag{2.1}$$

Now, multiplying the inequalities in (2.1), we obtain

$$(C'-A)(A'-B)(B'-C)=(B-C')(C-A')(A-B') \tag{2.2}$$

whereby division of both sides by the LHS, we obtain

$$\frac{C'-A}{B-C'}\cdot\frac{A'-B}{C-A'}\cdot\frac{B'-C}{A-B'}=1 \Leftrightarrow \frac{C'-A}{C'-B}\cdot\frac{A'-B}{A'-C}\cdot\frac{B'-C}{B'-A}=1.$$

Hence, by Ceva's Theorem, it follows that the medians of a triangle are concurrent.

Example 2.1.8. Suppose we have a triangle $\triangle ABC$. Given that a point C' (AB) has $\frac{C'-A}{C'-B}=-3$. What is $\frac{A-C'}{A-B}$? **Answer:** we know by the Main Collinearity Lemma that

$$\frac{C'-A}{C'-B}=-\frac{b}{a}=-\left(\frac{3}{1}\right).$$

So,

$$C'=\frac{aA+bB}{a+b}=\frac{1A+3B}{4}=\frac{1}{4}A+\frac{3}{4}B.$$

We may visualize this as C'as being $\frac{3}{4}$ the distance from A to B. So, $1A+3B+(-4)C'=0$, where $a+b+c'=0$.

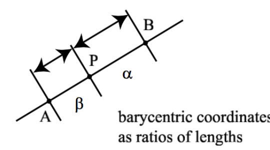


Figure 2.1.5: Distance.

2.2. Barycentric Coordinates

Definition 2.2.1. Consider a triangle $\triangle ABC$, a ***median*** is a line a joining a vertex to the midpoint of the opposite side. Thus, the intersection of all medians of a triangle is the ***centroid*** of the triangle. Now, consider $A_1,\dots,A_n$ points in the plane and let $a_1,\dots,a_n$ be arbitrary scalars, which we may think of masses or weights associated with each A_i. The pairs (a_i, A_i) are referrred to as ***weighted points*** or ***mass-points.*** The centroid of the mass-points is defined by

$$G=\frac{1}{\sum_{i=1}^n a_i}\left(\sum_{i=1}^n a_iA_i\right) \tag{2.3}$$

Definition 2.2.2. Let A_1, A_2, A_3 be points in $\mathbb{R}^2$, which do not lie on a line, that is, they form a triangle $\triangle A_1A_2A_3$. Furthermore the components of A_i comprise the pair (a_i, b_i) for $i = 1, 2, 3$. Now, if $P = (x, y) \in \mathbb{R}^2$ is a point in the plane, then taking the moments about x and y axis, we may express the components of P as such:

$$\begin{aligned} (\alpha + \beta + \gamma)\,x &= \alpha a_1 + \beta a_2 + \gamma a_3 \\ (\alpha + \beta + \gamma)\,y &= \alpha b_1 + \beta b_2 + \gamma b_3. \end{aligned}$$

Thus, P can be uniquely expressed as (to-show):

$$P = \frac{\alpha}{\alpha + \beta + \gamma}(a_1, b_1) + \frac{\beta}{\alpha + \beta + \gamma}(a_2, b_2) + \frac{\gamma}{\alpha + \beta + \gamma}(a_3, b_3)\,. \tag{2.4}$$

where $\alpha + \beta + \gamma = 1$, thus

$$x = \alpha a_1 + \beta a_2 + \gamma a_3 \quad \text{and} \quad y = \alpha b_1 + \beta b_2 + \gamma b_3. \tag{2.5}$$

Hence, (α, β, γ) are the ***barycentric coordinates*** of $P = (x, y)$ in the Euclidean plane. If, on the other hand we're initially given barycentric coordinates of P, then to obtain the cartesian coordinates, we solve for α, β, and γ using

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \tag{2.6}$$

which boils down to solving a system of equations in α, β, and γ.

Remark 2.2.3. This is useful because it helps us express a point in the plane with cartesian coordinates to one with barycentric coordinates provided that we have a triangle of reference to work with.

Theorem 2.2.4. *Every point P of the plane has a unique representation*

$$P = aA + aB + cC, \quad \text{where } a + b + c = 1. \tag{2.7}$$

That is, every point P in the plane can be written as a unique combination of three non-collinear points A, B, and C. The triple (a, b, c) represents the barycentric coordinates of P with respect to the points A, B, and C.

Proof. Let A, B, C be points on the plane which do not lie on a line, that is, they comprise a reference triangle. Now, consider X and Y as points that are projections of P onto the lines ℓ_{CA} and ℓ_{CB} parallel to the line ℓ_{CB} and ℓ_{CA}, respectively...TODO ∎

Remark 2.2.5. Tondeur's book has a proof of the unique representation above.

Definition 2.2.6. The equation $pa + qb + rc = 0$ is the ***barycentric equation of a line ℓ.***

Remark 2.2.7. If a line ℓ is given by the equation $ax + by + c = 0$ where $a \neq 0$ or $b \neq 0$ wrt to nonbarycentric coordinates (x, y), then in barycentric coordinates (z, x, y), where $x + y + z = 1$, the line D is given by

$$(a + c)\,x + (b + c)\,y + cz = 0.$$

Moreover, if $a + c = b + c = c$, then $a = b = 0$, a contradiction.

Theorem 2.2.8. *Three points with barycentric coordinates* $(a_1, b_1, c_1), (a_2, b_2, c_2), (a_3, b_3, c_3)$ *are collinear iff*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0.$$

Corollary 2.2.9. *The line* ℓ *joining* $(a_1, b_1, c_1), (a_2, b_2, c_2)$ *consists of all* (a, b, c) *such that*

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a & b & c \end{vmatrix} = 0.$$

Theorem 2.2.10. *The lines with barycentric equations* $p_1a + q_1b + r_1c = 0$, $p_2a + q_2b + r_2c = 0$, *and* $p_3a + q_3b + r_3c = 0$ *are either concurrent or all parallel iff*

$$\begin{vmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{vmatrix} = 0.$$

Corollary 2.2.11. *The lines with barycentric equations* $p_1a+q_1b+r_1c = 0$ *and* $p_2a+q_2b+r_2c = 0$ *are parallel iff*

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

Question 2.2.12. *How do we find the barycentric coordinates of a point* P *in a plane with respect to an affine frame of reference* (A_0, A_1, A_2). ***Answer:*** first observe that A_0, A_1, A_2 are affinely independent in the sense that they are non-collinear. Thus, they form the vertices of a triangle, call it, $\triangle A_0A_1A_2$. Now, (draw this out: figure TODO), we proceed as follows:

1. First we determine the intersection of the lines (PA_0) and (A_1A_2), and denote it as Q. Then, the point Q has barycentric coordinates

   $$Q = \beta_1A_1 + \beta_2A_2.$$

   Also, the point P, with respect to A_0 and Q has coordinates $P = \delta_0A_0 + \delta_1Q$, thus

   $$\begin{aligned} P &= \delta_0A_0 + \delta_1Q = \delta_0A_0 + \delta_1(\beta_1A_1 + \beta_2A_2) \\ &= \delta_0A_0 + \delta_1\beta_1A_1 + \delta_1\beta_1A_2. \end{aligned}$$

Examples

Example 2.2.13. Let (a, b, c) be barycentric coordinates of P relative to the vertices of $\triangle ABC$, that is, the affine frame (A, B, C), for which we shall recall that

$$P = aA + bB + cC, \quad a + b + c = 1. \tag{2.8}$$

Now, through what points does the line $2a + 3b + 5c = 0$ cross the sides of the triangle? **Answer:** let $a = 0$, then we have

$$3b + 5c = 0. \tag{2.9}$$

In which case points P lying on(BC) satisfy, $P = bB + cC$, where $b + c = 1$. So, we must find numbers b, c that satisfy eqn. (17), and for which $b + c = 1$. That is, we must solve the system of equations

$$\begin{aligned} 3b + 5c &= 0 \\ b + c &= 1. \end{aligned}$$

Thus, let $b = 1 - c$, in which case $3b + 5c = 3(1 - c) + 5c = 0$, implies $c = -3/2$. From this, we further deduce that $b = \frac{5}{2}$. So, $A' = \frac{5}{2}B - \frac{3}{2}C$ lies on (BC) and on the line $2a + 3b + 5c = 0$. Now, suppose that $c = 0$, then we must solve for

$$\begin{aligned} 2a + 3b &= 0 \\ a + b &= 1 \end{aligned}$$

and proceed as in the previous parts.

Example 2.2.14. (Paint mixing.) Let A be the color red, B the color blue, C the color yellow. If one mixes a quarts of red paint with b quarts of blue paint and c quarts of yellow paint, one gets $a + b + c$ quarts of paint of color defined by the formula

$$\frac{1}{a+b+c}(aA + bB + cC). \tag{2.10}$$

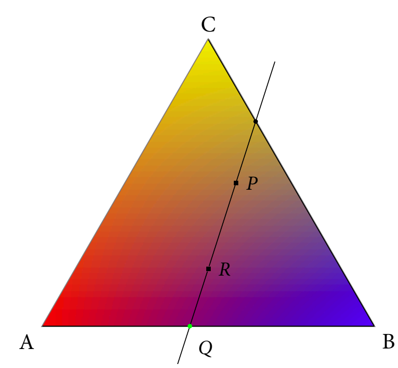


Figure 2.2.1: Q is the unknown color we suspect to lie the side AB.

Mix one quart of paint of color $\frac{1}{6}(A + 2B + 3C)$ with an unknown quantity of mixed paint using only two colors, so as to obtain an unknown quantity of mixed paint of color $\frac{1}{5}(2A + 2B + C)$. Determine the unknown quantities and color. **Answer:** Let P be the given color that was mixed with the unknown color Q to obtain R. Thus,

$$P = \frac{1A}{6} + \frac{2B}{6} + \frac{3C}{6} = \frac{3}{6}\underbrace{\left(\frac{1}{3}A + \frac{2}{3}B\right)}_{P_1} + \frac{3}{6}C \tag{2.11}$$

$$= \frac{1}{6}A + \frac{5}{6}\underbrace{\left(\frac{2}{5}B + \frac{3}{5}C\right)}_{P_2} = \frac{4}{6}\underbrace{\left(\frac{1}{4}A + \frac{3}{4}C\right)}_{P_3} + \frac{2}{6}B. \tag{2.12}$$

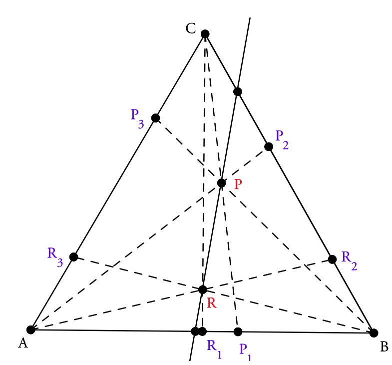


Figure 2.2.2: Location of P and R

And for R, we have the following

$$R = \frac{2}{5}A + \frac{2}{5}B + \frac{1}{5}C = \frac{4}{5}\underbrace{\left(\frac{2}{4}A + \frac{2}{4}B\right)}_{R_1} + \frac{1}{5}C \tag{2.13}$$

$$= \frac{2}{5}A + \frac{3}{5}\underbrace{\left(\frac{2}{3}B + \frac{1}{3}C\right)}_{R_2} = \frac{2}{5}B + \frac{3}{5}\underbrace{\left(\frac{2}{3}A + \frac{1}{3}C\right)}_{R_3}. \tag{2.14}$$

We can then locate the points P_1, P_2, P_3 and R_1, R_2, R_3 on the sides of an equilateral triangle $\triangle ABC$ and draw cevians connecting these points to the vertex opposite to them. The intersection of each cevian provides the exact location of P and R as indicated in figure 2. From the illustration we can see that Q is most likely situated on ℓ_{AB}. This is because $R = (1 - t)Q + tP$, being the combination of paints P and Q has to have positive coordinates in relation to P and Q . So, R must lie between P and Q. With this in mind, we express Q as an affine combination of A and B, that is, we let $Q = \alpha A + \beta B$, and solve for t in the equation

$$\begin{aligned} R &= (1-t)\,Q + tP = (1-t)\,(\alpha A + \beta B) + tP \\ &= \alpha A + \beta B - \alpha t A - \beta t B + \frac{1}{6}tA + \frac{2}{6}tB + \frac{3}{6}tC \\ &= \left(\alpha - \alpha t + \frac{1}{6}t\right)A + \left(\beta - \beta t + \frac{2}{6}t\right)B + \frac{3}{6}tC \end{aligned}$$

By equating the coeffcients of the equation above with those of $R = \frac{2}{5}A + \frac{2}{5}B + \frac{1}{5}C$, we can solve for t and obtain the values of α and β. So, for the coefficients of C we have $\frac{3}{6}t = \frac{1}{5}$, which means that $t = \frac{2}{5}$, and $1 - t = \frac{3}{5}$. So, we plug these values into

$$\alpha(1-t) + \frac{1}{6}t = \frac{2}{5}$$

and

$$\beta\,(1-t) + \frac{2}{6}t = \frac{2}{5}$$

thereby obtaining $\alpha = \frac{5}{9}$ and $\beta = \frac{4}{9}$. Therefore $Q = \frac{5}{9}A + \frac{4}{9}B$. Now, to discern the amount of Q mixed with one quart of P, and the amount of R obtained from this mixture, we simply solve for x in the equation $P + xQ = R\,(1+x)$. Therefore

$$\begin{aligned} x = \frac{P-R}{R-Q} &= \frac{\left(\frac{1}{6}A + \frac{2}{6}B + \frac{3}{6}C\right) - \left(\frac{2}{5}A + \frac{2}{5}B + \frac{1}{5}C\right)}{\left(\frac{2}{5}A + \frac{2}{5}B + \frac{1}{5}C\right) - \left(\frac{5}{9}A + \frac{4}{9}B\right)} \\ &= \frac{-\frac{7}{30}A - \frac{1}{15}B + \frac{3}{10}C}{-\frac{7}{45}A - \frac{2}{45}B + \frac{1}{5}C} = \frac{\frac{-7A-2B+9C}{30}}{\frac{-7A-2B+9C}{45}} = \frac{45}{30} = \frac{3}{2}. \end{aligned}$$

So, $x = 1.5$ quarts. Therefore, 1.5 quarts of Q were mixed with 1 quart of P to obtain 2.5 quarts of R.

2.3. Desargues' Theorem

Theorem of Desargues. *Let $\triangle ABC$ and $\triangle A'B'C'$ be two triangles in the Euclidean plane . If lines $\ell_{AA'}, \ell_{BB'}, \ell_{CC'}$ are concurrent with D as the point of intersection and*

$$A^* = \ell_{BC} \cap \ell_{B'C'}, \quad B^* = \ell_{AC} \cap \ell_{A'C'} \quad C^* = \ell_{AB} \cap \ell_{A'B'}$$

then A^*, B^*, C^* are collinear.

Proof. Consider triangle $\triangle ABD$, the line $\ell_{A'B'}$, and menelaus points $C^\star \in \ell_{AB}$, $A' \in \ell_{DA}$, and $B' \in \ell_{DB}$. Applying Menelaus' Theorem, we thereby obtain

$$\frac{C^\star - A}{C^\star - B} \cdot \frac{A' - D}{A' - A} \cdot \frac{B' - D}{B' - B} = 1. \tag{1}$$

Similarly, for $\triangle CBD$, the line $\ell_{B'C'}$, and menelaus points $A^\star \in \ell_{BC}$, $C' \in \ell_{CD}$, and $B' \in \ell_{BD}$, we apply Menelaus' Theorem and get

$$\frac{A^\star - B}{A^\star - C} \cdot \frac{C' - C}{C' - D} \cdot \frac{B' - B}{B' - D} = 1. \tag{2}$$

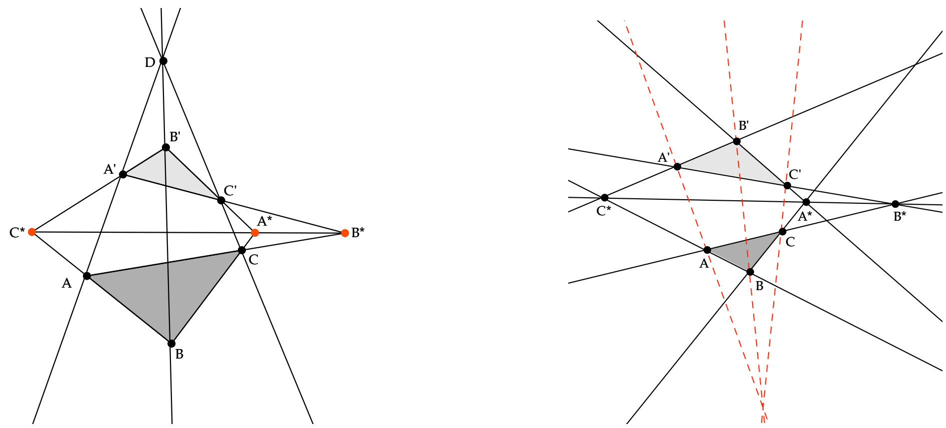


Figure 2.3.1. The Duality of Desargue's Theorem

Lastly, for $\triangle ACD$, the line $\ell_{C'A'}$, and menelaus points $B^\star \in \ell_{CA}$, $C' \in \ell_{DC}$, $A' \in \ell_{AD}$, we have

$$\frac{B^\star - C}{B^\star - A} \cdot \frac{C' - D}{C' - C} \cdot \frac{A' - A}{A' - D} = 1. \tag{3}$$

We then multiply these three products, resulting in

$$\frac{A^\star - B}{A^\star - C} \cdot \frac{B^\star - C}{B^\star - A} \cdot \frac{C^\star - A}{C^\star - B} = 1,$$

which indicates that $A^\star, B^\star, C^\star$ are collinear as desired. ■

Dual of Desargue's Theorem. *Given two triangles $\triangle ABC$ and $\triangle A'B'C'$. If the intersection points*

$$A^* = \ell_{BC} \cap \ell_{B'C'}, \quad B^* = \ell_{AC} \cap \ell_{A'C'} \quad C^* = \ell_{AB} \cap \ell_{A'B'}$$

are collinear, then the lines $\ell_{AA'}, \ell_{BB'}, \ell_{CC'}$ are concurrent.

2.4. The Euler Line and Nine-Point Circle

Euler Line Theorem. *The orthocenter, barycenter, and circumcenter of a triangle are collinear, and the barycenter is two-thirds the distance from the orthocenter to the circumenter.*

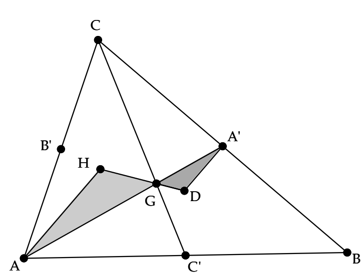


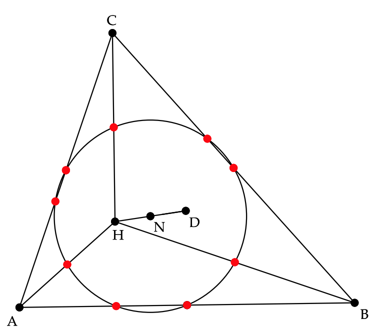
Figure 2.4.1. Euler Line

G is simply the centroid of the triangle obtained by taking the intersection of any two medians of a triangle. H is the orthocenter point obtained by extending the segment DG to twice its length beyond G, and D is the circumcenter obtained by taking the intersection of any two perpendicular bisectors of the sides of a triangle.

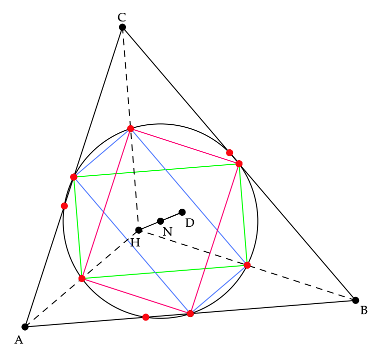
Proof. First, let us find the bisectors of the sides of $\triangle ABC$ and label them A', B', C'. Now, choose a vertex, say A, then draw the triangle $\triangle A'DG$. From the Barycenter Theorem it follows that the length of segment AG is twice that of GA', that is, $|AG| = 2\,|GA'|$. Then extend DG by twice its length to a new point (the orthocenter) called H, and draw the segment HA. Now, notice that the two triangles are similar by virtue of their opposite angles at barycenter G being congruent. Thus $\triangle AGH \approx \triangle A'GD$. Furthermore, the corresponding angles $\angle H$ and $\angle D$ are equal, thus the sides $A'D \parallel AH$. Since the side AH is by construction perpendicular to BC, so is the AH extended, which serves as the altitude on A. Since our choice of A is arbitraty, and H is constructed from D and G independently of choice, it follows that the other altitudes on B and C must go through the orthocenter H. ∎

Nine-Point Circle Theorem. *Given $\triangle ABC$, the three midpoints of its sides, the three feet of the altitudes on each vertex, and the three midpoints of each segment that extends from a vertex to the orthocenter all lie on a circle whose center is the midpoint of the segment connecting the orthocenter with the circumcenter.*

Proof. Choose two vertices of $\triangle ABC$, say B and C, and consider the sides opposite to them. Then the midpoints of these sides together with the midpoints of the segments from the selected vertices (B and C) to the orthocenter H form a rectangle (indicated in blue within Figure 2.3.2b). One pair of sides of this particular quadrilateral is parallel to the side BC, and the other pair is parallel to the line through the vertex A and orthocenter H. This line going through A and H intersects the side BC at foot of the altitude of A. The feet of the altitudes on vertices on B and C lie on the circumcircle of the rectangle we've constructed. Repeating this whole process for the other pairs of vertices, we can construct two more rectangles. In each case, we will obtain two rectangles that share a diagonal. Thus all the vertices of the three rectangles lie on one common circumcircle. ■



(a)



(b)

Figure 2.4.2. Nine Point Circle

3 — Perspective Geometry

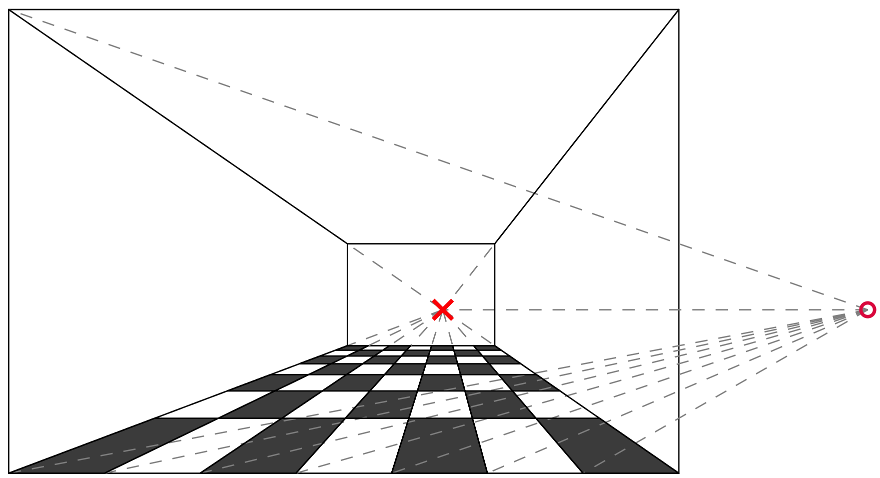


Figure 3.0.1. Perspective Room

3.1. Kepler's Ideal Lines and Points

Definition 3.1.1. An ***ideal poin***t or ***point at infinity*** X is a set of lines parallel to a line ℓ. In the plane, the set of all ideal points comprise the ***ideal line***, ∞. By (ℓX), we mean to say that $\ell \in X$. So, every line ℓ has an ideal point on itself, namely $X = (\ell\infty)$. Two lines are parallel if and only if they meet at infinity (refer to Figure 3.1.1).

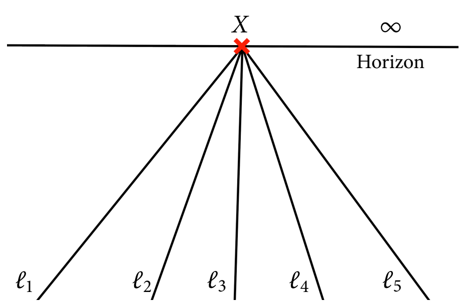


Figure 3.1.1. Ideal line and point: we see that X is an ideal point beloning to the ideal line ∞. Also, by definition, the lines $\ell_1, \ell_2, \ldots, \ell_5$ are parallel, since they meet at the ideal point X.

Remark 3.1.2. In view of our previous definition, we bring attention to some terminology that is common in projective geometry. We may refer to X as the ***pencil*** or ***sheaf of parallel lines***. On the other hand, the ideal line to which X pertains, is commonly referred to as the the ***sheaf of parallel planes*** meeting at the ideal line. In the euclidean plane, we simply take an ordinary line, and pick a point on it, calling it the ideal point or point at infinity. We then have the following properties:

1. Parallel ordinary lines share the same ideal point;
2. Skew or intersecting ordinary lines correspond to different ideal points;
3. All ideal points belonging to the ordinary lines in a given plane form the ideal line of that plane;
4. Parallel ordinary planes share the same ideal line (refer to Figure 3.1.2);
5. Intersecting ordinary planes have distinct ideal lines;
6. All ideal points and ideal lines in space form the ideal plane;
7. Every ideal point is considered to be infinitely far removed from every other ordinary or ideal point.

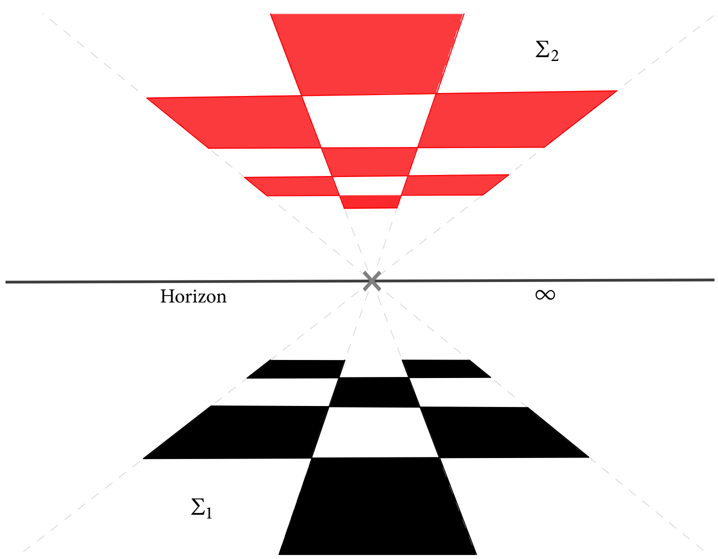


Figure 3.1.2. Two parallel planes sharing the same ideal line ∞. We see that the planes Σ_1 (containing black tiles) and Σ_2 (containing the red tiles) are parallel in the sense that they share the same ideal line ∞. The gray dotted lines are incident at the crossed point along the horizon (ideal line) or line at infinity, hence they're parallel.

3.2. Horizon and Zenith

Some vocabulary:

- The horizon of the plane corresponds to the ideal line of the plane per Kepler's terminology. It is the set of all parallel planes intersecting in a line at infinity. In view of figure 3.1.2, we see that the planes Σ_1 and Σ_2 intersect at the horizon or line at infinity.
- What is a pencil of planes? It is the set of all planes that have the same common line, also called a sheaf of planes. Hence, going back to figure 3.1.2, we see that Σ_1, Σ_2 belong to the pencil of planes intersecting at the horizon or line at infinity.
- The zenith of a plane is the ideal point corresponding to the sheaf or pencil of parallel lines that are perpendicular to the plane.

Theorem 3.2.1. *Given a circle whose diameter is the side of a triangle inscribed within, such a triangle is a right triangle.*

Theorem 3.2.2. *The altitude of a right triangle is the geometric mean of the two segments its foot divides along the hypotenuse.*

Three Problems in Perspective Geometry

- Given an eye point (orthocenter) C, focal distance δ, and the horizon line h, find the corresponding zenith point Z.
- Given a convex quadrilateral in a plane. From where in space does this polygon look square?
- Complete the drawing of a cube in perspective, starting from three vanishing lines (horizons) of its faces, and a line segment representing one edge.

We address the first problem below. First, observe that:

- The hypotenuse is in the picture plane PP;
- The apex of the triangle A is at the eye of the perspectivist;
- And one leg AB is in the sight plane, corresponding to the given horizon line h;
- The altitude AC of the right triangle $\triangle ABC$ connects the eye to the nearest point C on PP;
- The length of the altitude is the focal distance δ.

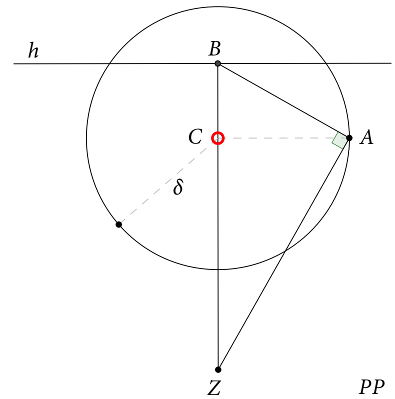


Figure 3.2.1: Thale's Triangle in the Picture Plane

4 — Transformational Geometry

4.1. The Setting: Euclidean Affine Space

We are now concerned with affine spaces $(\mathcal{A}, V)$, where the associated vector space V is endowed with additional structure, namely that of an inner product. A real vector space equipped with an inner product is then called a Euclidean vector space. We now bring this additional structure to our attention as it brings about the notion of length and enables us to characterize special types of point transformations (affine transformations) called isometries. Euclidean geometry is concerned with properties invariant under such transformations. Let us proceed with defining the following spaces.

Definition 4.1.1. A ***Euclidean space*** is a real vector space E endowed with an inner product $\phi : E \times E \to \mathbb{R}$, a function satisfying:

1. $\phi(c_1x_1 + c_2x_2, y) = c_1\phi(x_1, y) + c_2(x_2, y)$ (Bilinear)
2. $\phi(x, y) = \phi(y, x)$ (Symmetric)
3. $\phi(x, x) > 0$ whenever $x \neq 0$. (Positive Definite)

Remark 4.1.2. We also denote $\phi(\cdot,\cdot)$ by $\langle\cdot,\cdot\rangle$. In a real vector space setting, the inner product $\phi(x, y)$ is also referred to as a symmetric bilinear form that is positive definite. By bilinear, we mean ϕ is linear in both arguments. If we're working in a Hermitian space H (complex vector space with inner product), the inner product is a sesquilinear form that is hermitian and positive definite, that is, $\phi : H \times H \to \mathbb{C}$ satisfies:

1. $\phi(x, c_1y_1 + c_2y_2) = \overline{c}_1\phi(x, y_1) + \overline{c}_2(x, y_2)$ (Sesquilinear)
2. $\phi(x, y) = \overline{\phi(y, x)}$ (Hermitian)
3. $\phi(x, x) > 0$ whenever $x \neq 0$. (Positive Definite)

For all $x, y, x_1, x_2 \in E$ and $c_1, c_2 \in \mathbb{C}$, where $c \mapsto \overline{c}$ is a conjugation of the complex field (refer to appendix). By sequilinear, we mean to say that ϕ is linear in the first argument, and conjugate linear in the second.

Lemma 4.1.3. *Let V be a vector space over a field $\mathbb{F}$ equipped with an inner product as defined above. Then for $x, y, z \in V$ and $c \in \mathbb{F}$, we the following statements are true:*

1. $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$
2. $\langle x, cy\rangle = \overline{c}\langle x, y\rangle$
3. $\langle 0_V, x\rangle = \langle x, 0_V\rangle = 0$ *for all* $x \in V$.
4. $\langle x, x\rangle = 0$ *if and only if* $x = 0$.
5. *If* $\langle x, y\rangle = \langle x, z\rangle$ *for all* $x \in V$*, then* $y = z$.

Proof. First observe that V is a inner product space over an arbitrary field $\mathbb{F}$, that being said, we need to accomodate the possibility that it may be Hermitian, and so we proceed to use the properties of sesquilinear form that is hermitian and positive definite. So, for (1), we have:

$$\langle x, y+z\rangle = \overline{\langle y+z, x\rangle} = \overline{\langle y, x\rangle} + \overline{\langle z, x\rangle} = \langle x, y\rangle + \langle x, z\rangle .$$

For (2), we have $\langle x, cy\rangle = \overline{\langle cy, x\rangle} = \overline{c\,\langle y, x\rangle} = \overline{c}\,\langle x, y\rangle$. For (3), we use the trivial fact that $\langle x, x\rangle = \langle x, x\rangle$, then subtracting $\langle x, x\rangle$ from both sides, we obtain

$$\langle x, x\rangle - \langle x, x\rangle = 0 \Leftrightarrow \begin{cases} \langle x + (-x), x\rangle = 0 \\ \langle x, x + (-x)\rangle = 0 \end{cases} \Leftrightarrow \begin{cases} \langle 0_V, x\rangle = 0 \\ \langle x, 0_V\rangle = 0. \end{cases}$$

Thus, $\langle 0_V, x\rangle = \langle x, 0_V\rangle = 0$. For (4), suppose $\langle x, x\rangle = 0$ and $x \neq 0$, then this merely contradicts positive definite property of an inner product, thus x must be 0_V, as desired. Conversely, suppose $x = 0_V$, then $\langle 0_V, 0_V\rangle = 0$ follows directly from (3). Now, for (5), suppose $\langle x, y\rangle = \langle x, z\rangle$ for all $x \in V$, then subtracting $\langle x, z\rangle$ from both sides, we obtain:

$$\langle x, y\rangle - \langle x, z\rangle = 0 \Leftrightarrow \langle x, y - z\rangle = 0.$$

Then, since x is arbitrary, we must have $y - z = 0_V$, in order for $\langle x, y - z\rangle = 0$ to comply with (3). Thus, $y = z$, as desired. ■

Definition 4.1.4. An affine space $(\mathcal{E}, E)$ is a ***Euclidean affine space*** if its associated vector space E is Euclidean.

Remark 4.1.5. Throughout the remainder of the text we will be working in $(\mathcal{E}, (E, \bullet))$. In particular, we will be working in the two-dimensional Affine Euclidean space $\mathcal{E}$=$\mathbb{R}^2$, whose associated vector space is equipped with the dot product or standard inner product as defined below.

Definition 4.1.6. The ***standard inner product*** or ***dot product*** for the Euclidean space $E = \mathbb{R}^n$, is

$$\langle x, y\rangle_S = x \bullet y = \sum_{i=1}^{n} x_i y_i$$

for $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ in E. The ***Euclidean norm*** of $x \in E$, denoted by $\|x\|$, is a function $\|x\|_2 : E \to \mathbb{R}$, defined as

$$\|x\|_2 := \sqrt{\langle x, x\rangle_S} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

The ***distance between two vectors*** $x, y \in E$, is a function $d : E \times E \to \mathbb{R}$ defined as $d(x, y) := \|x - y\|_2$.

The function $d(x, y)$ can generalized into what is called a ***metric*** provided that for any $x, y, z \in M$, the function $d : M \times M \to \mathbb{R}$, satisfies:

1. $d(x, y) > 0$ for $x \neq y$, and $d(x, x) = 0$.
2. $d(x, y) = d(y, x)$.
3. $d(x, z) \leq d(x, y) + d(y, z)$.

Remark 4.1.7. A set M for which such a function exists is called a ***metric space***.

Notation 4.1.8. When the context is clear, we shall simply write $\langle x, y\rangle_S$ as $\langle x, y\rangle$. and $\|x\|_2$ as $\|x\|$.

Remark 4.1.9. The inner product $\phi(x, y) = \langle x, y\rangle$ can be defined in terms of the norm as follows: Define $\Phi(x) = \langle x, x\rangle$, then for $x, y \in E$, we have,

$$\begin{aligned} \Phi(x+y) &= \langle x+y, x+y\rangle = \langle x, x+y\rangle + \langle y, x+y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle \end{aligned}$$

Thus, from this we obtain,

$$\langle x, y\rangle = \frac{1}{2}\left(\langle x+y, x+y\rangle - \langle x, x\rangle - \langle y, y\rangle\right) = \frac{1}{2}\left[\|x+y\|^2 - \|x\|^2 - \|y\|^2\right].$$

Alternatively, if we use $\Phi(x-y)$, we obtain

$$\langle x, y\rangle = \frac{1}{2}\left[\|x\|^2 + \|y\|^2 - \|x-y\|^2\right].$$

Example 4.1.10. Let $X = (x, y)$ and $Y = (u, v)$ be two vectors in the vector space $\mathbb{R}^2$, then the ***dot product*** of these vectors, written as $X \bullet Y$, is an inner product defined as:

$$X \bullet Y = xu + yv.$$

The length of X is then $\|X\| = \sqrt{X \bullet X} = \sqrt{x^2 + y^2}$. The distance between X and Y is

$$\|X - Y\| = \sqrt{(X-Y)\bullet(X-Y)} = \sqrt{(x-u)^2 + (y-v)^2}.$$

In view of our previous remark, we see that

$$X \bullet Y = \frac{1}{2}\left(\|X\|^2 + \|Y\|^2 - \|X-Y\|^2\right).$$

Notation 4.1.11. When using the dot product of two vectors $X \bullet Y$, we will simply write it as XY whenever ambiguity isn't a problem. Hence, $X \bullet X = XX = X^2$.

Theorem 4.1.12. *Let E be a Euclidean space, then for all $X, Y \in E$ and $c \in \mathbb{R}$, the following properties of the norm are true:*

1. $\|cX\| = |c|\,\|X\|$
2. $\|X\| = 0$ *if and only if* $x = 0$. *In any case,* $\|X\| \geq 0$.
3. $|\langle X, Y\rangle| \leq \|X\| \cdot \|Y\|$ (*CBS - Inequality*)
4. $\|X + Y\| \leq \|X\| + \|Y\|$. (*Triangle Inequality*)

Proof. (1) For $c \in \mathbb{R}$ and $X \in E$, we have $\|cX\| = \sqrt{\langle cX, cX\rangle} = \sqrt{c^2\langle X, X\rangle} = \sqrt{c^2}\cdot\sqrt{\langle X, X\rangle} = |c|\,\|X\|$. (2) Suppose $\|X\| = 0$, then $\|X\| = \sqrt{\langle X, X\rangle} = 0$ suggests that $\langle X, X\rangle = 0$, whereby property (4) of the previous lemma, we must have $X = 0_E$. Conversely, if $X = 0_E$, then by property (3) of the previous lemma, must have $\langle 0_E, 0_E\rangle = 0$, thus $\|0_E\| = \sqrt{\langle 0_E, 0_E\rangle} = 0$, as desired. If $X \neq 0$, then positive definiteness of the inner product implies that $\|X\| > 0$. For (3), let us without loss of generality, emplace conditions on Y. If $Y = 0$, then equality holds. Now for $y \neq 0$, it follows by (2) that $\|X - cY\|^2 \geq 0$, which is equivalent to:

$$\begin{aligned}
\langle X - cY, X - cY\rangle &= \langle X, X - cY\rangle + \langle -cY, X - cY\rangle \\
&= \langle X, X\rangle + \langle X, -cY\rangle + \langle -cY, X\rangle + \langle -cY, -cY\rangle \\
&= \langle X, X\rangle - \overline{c}\langle X, Y\rangle - c\langle Y, X\rangle + c\overline{c}\langle Y, Y\rangle \\
&= \langle X, X\rangle - \overline{c}\langle X, Y\rangle - c\langle Y, X\rangle + |c|^2\langle Y, Y\rangle \geq 0.
\end{aligned}$$

Now, applying $c = \langle X, Y\rangle / \langle Y, Y\rangle$, we obtain

$$\langle X, X\rangle - \frac{\overline{\langle X, Y\rangle}}{\overline{\langle Y, Y\rangle}}\langle X, Y\rangle - \frac{\langle X, Y\rangle}{\langle Y, Y\rangle}\langle Y, X\rangle + \frac{|\langle X, Y\rangle|^2}{|\langle Y, Y\rangle|^2}\langle Y, Y\rangle \geq 0.$$

Which is equivalent to

$$\langle X, X\rangle - \frac{|\langle X, Y\rangle|^2}{\langle Y, Y\rangle} - \frac{|\langle X, Y\rangle|^2}{\langle Y, Y\rangle} + \frac{|\langle X, Y\rangle|^2}{\langle Y, Y\rangle} \geq 0.$$

Hence, we have

$$\langle X, X\rangle - \frac{|\langle X, Y\rangle|^2}{\langle Y, Y\rangle} \geq 0 \Leftrightarrow \langle X, X\rangle \langle Y, Y\rangle \geq |\langle X, Y\rangle|^2 \Leftrightarrow \|X\| \cdot \|Y\| \geq |\langle X, Y\rangle|$$

as desired. Now, for (4), we have

$$\begin{aligned} \|X + Y\|^2 &= \langle X + Y, X + Y\rangle = \langle X + Y, X\rangle + \langle X + Y, Y\rangle \\ &= \langle X, X\rangle + \langle Y, X\rangle + \langle X, Y\rangle + \langle Y, Y\rangle \\ &\leq \|X\|^2 + 2\|X\|\|Y\| + \|Y\|^2 = (\|X\| + \|Y\|)^2 . \end{aligned}$$

Thus, taking the square root of both sides we obtain the desired result. ∎

4.2. Affine Transformations

Useful Characterizations of Affine Transformations

Definition 4.2.1. Suppose we have two affine spaces $(\mathcal{A}, V, +)$ and $(\mathcal{A}', V', +')$, a function $f : \mathcal{A} \to \mathcal{A}'$ is an affine transformation if and only if for every family $((\alpha_i, A_i))_{i\in I}$ of weighted points in $\mathcal{A}$, where $\sum_{i\in I} \alpha_i = 1$, we have

$$f\left(\sum_{i\in I} \alpha_i A_i\right) = \sum_{i\in I} \alpha_i f(A_i).$$

That is, barycentres or affine combinations, are preserved.

Definition 4.2.2. An affine transformation of an affine space $(\mathcal{A}, V, +)$ into another affine space $(\mathcal{A}', V', +')$ is a pair of mappings:

$$f : \mathcal{A} \to \mathcal{A}', \quad \mathsf{F} : V \to V'$$

satisfying the two conditions:

1. The mapping $\mathsf{F} : V \to V'$ is a linear transformation of vector spaces $V \to V'$.
2. For every pair of points $X, Y \in \mathcal{A}$, we have

$$\overrightarrow{f(X) f(Y)} = \mathsf{F}\left(\overrightarrow{XY}\right).$$

4.3. Isometries in the Plane

> Euclidean geometry is the study of those properties of geometric figures that are invariant under the group of similarities.
>
> -Felix Klein

Definition 4.3.1. A point transormation α of the plane is an isometry if for every pair of points X and Y in $\mathcal{E}$

$$\|X^\alpha - Y^\alpha\| = \|X - Y\|.$$

Notation 4.3.2. We use the superscript notation X^α to denote $\alpha(X)$.

Proposition 4.3.3. *The composition of two isometries is an isometry.*

Proof. Let α and β be two isometries, then let $\gamma = \beta\alpha$ be their composition. For any pair of points X, Y in $\mathcal{E}$, we have

$$\begin{aligned} \|X^\gamma - Y^\gamma\| &= \|\beta(\alpha(X)) - \beta(\alpha(Y))\| \\ &= \|\beta(X^\alpha) - \beta(Y^\alpha)\| \\ &= \|X^\alpha - Y^\alpha\| && (\beta \text{ is an isometry}) \\ &= \|X - Y\| && (\alpha \text{ is an isometry}), \end{aligned}$$

as desired. ∎

Proposition 4.3.4. *Inverses of isometries are yet again isometries.*

Proof. Let $\omega = \alpha^{-1}$ be the inverse of an isometry α. Then, since $\alpha\alpha^{-1}$is the identity, we have for $X, Y \in \mathcal{E}$, the following

$$\begin{aligned} \|X - Y\| &= \left\|X^{\alpha\alpha^{-1}} - Y^{\alpha\alpha^{-1}}\right\| \\ &= \left\|\alpha X^{\alpha^{-1}} - \alpha Y^{\alpha^{-1}}\right\| \\ &= \|X^\omega - Y^\omega\| && (\alpha \text{ is an isometry}), \end{aligned}$$

as required. ∎

Linear Isometries a.k.a Osometries

Definition 4.3.5. A ***fixed point*** P is that which remains invariant under a point transformation ϕ, that is, $\phi(X) = X = P$.

Definition 4.3.6. *Let $X = X + 0_E$. An **osometry** is an isometry that fixes the origin, that is* $\alpha(O) = O$ or $O^\alpha = O$.

Proposition 4.3.7. *Let β be an osometry, then for vectors $X, Y \in V$, we have:*

1. $\left\|X^\beta\right\| = \|X\|$ (Preservation of distance from choice of origin)
2. $\left\langle X^\beta, Y^\beta\right\rangle = \langle X, Y\rangle$ (Preserves the dot product)
3. $(rX)^\beta = rX^\beta$ for any scalar $r \in \mathbb{R}$. (Preserves scalar products)
4. $X^\beta + Y^\beta = (X+Y)^\beta$ for any X, Y. (Preserves sums)

Proof. Let $X, Y \in \mathcal{E}$ and β be an osometry. Then for (1), if

$$\begin{aligned} \left\|X^\beta - O\right\| &= \left\|X^\beta - O^\beta\right\| && (4.1) \\ &= \|X\|. && (4.2) \end{aligned}$$

Now, for (2), suppose $X \neq Y$, then for $\left\|X^\beta - Y^\beta\right\|^2 = \left\langle X^\beta - Y^\beta, X^\beta - Y^\beta\right\rangle > 0$, we have:

$$\begin{aligned} \left\langle X^\beta - Y^\beta, X^\beta - Y^\beta\right\rangle &= \left\langle X^\beta, X^\beta - Y^\beta\right\rangle - \left\langle Y^\beta, X^\beta - Y^\beta\right\rangle && (4.3) \\ &= \left\langle X^\beta, X^\beta\right\rangle - \left\langle X^\beta, Y^\beta\right\rangle - \left\langle X^\beta, Y^\beta\right\rangle + \left\langle Y^\beta, Y^\beta\right\rangle && (4.4) \\ &= \left\langle X^\beta, X^\beta\right\rangle - 2\left\langle X^\beta, Y^\beta\right\rangle + \left\langle Y^\beta, Y^\beta\right\rangle && (4.5) \\ &= \left\|X^\beta\right\|^2 - 2\left\langle X^\beta, Y^\beta\right\rangle + \left\|Y^\beta\right\|^2 = \|X\|^2 - 2\left\langle X^\beta, Y^\beta\right\rangle + \|Y\|^2. && (4.6) \end{aligned}$$

Now, since β is an isometry, we know $\left\|X^\beta - Y^\beta\right\| = \|X - Y\|$, and so it follows that $\left\|X^\beta - Y^\beta\right\|^2 = \|X - Y\|^2$ if and only if

$$\|X\|^2 - 2\left\langle X^\beta, Y^\beta\right\rangle + \|Y\|^2 = \|X\|^2 - 2\langle X, Y\rangle + \|Y\|^2 . \tag{4.7}$$

Thereby implying that $\left\langle X^\beta, Y^\beta\right\rangle = \langle X, Y\rangle$, as required. For (3), let Z be an arbitrary vector and observe that

$$\left\langle Z, (rX)^\beta\right\rangle = \langle Z, rX\rangle = r\langle Z, X\rangle \tag{4.8}$$

which is equivalent to

$$\left\langle Z, rX^\beta\right\rangle = r\left\langle Z, X^\beta\right\rangle = r\langle Z, X\rangle . \tag{4.9}$$

Hence, $\left\langle Z, (rX)^\beta\right\rangle = \left\langle Z, rX^\beta\right\rangle$, and so by (5) of Lemma 4.1.3, we must have $(rX)^\beta = rX^\beta$. For (4), let Z be an arbitrary vector, and consider $\left\langle Z, X^\beta + Y^\beta\right\rangle$ and $\left\langle Z, (X+Y)^\beta\right\rangle$ Then,

$$\left\langle Z, X^\beta + Y^\beta\right\rangle = \left\langle Z, X^\beta\right\rangle + \left\langle Z, Y^\beta\right\rangle = \langle Z, X\rangle + \langle Z, Y\rangle \tag{4.10}$$

and for $\left\langle Z, (X+Y)^\beta\right\rangle$, we have:

$$\left\langle Z, (X+Y)^\beta\right\rangle = \langle Z, X+Y\rangle \qquad \text{(Preservation of Dot Product)} \tag{4.11}$$

$$= \langle Z, X\rangle + \langle Z, Y\rangle . \tag{4.12}$$

Hence, for arbitrary an arbitrary vector Z, we see that

$$\left\langle Z, X^\beta + Y^\beta\right\rangle = \left\langle Z, (X+Y)^\beta\right\rangle \tag{4.13}$$

whereby (5) of Lemma 4.1.3, it follows that $(X+Y)^\beta = X^\beta + Y^\beta$, as required. ■

Translations

Definition 4.3.8. A **translation** by the vector A in the Euclidean plane is the mapping $\tau_A : \mathcal{E} \to \mathcal{E}$ that associates to a point $X \in \mathcal{E}$ the image point X' uniquely determined as $\tau_A(X) = X' = X + A$.

Proposition 4.3.9. *Translations are isometries.*

Proof. Indeed, let $\tau = \tau_A$, where $A \in E$, then for $X, Y \in \mathcal{E}$, we have

$$\|X^\tau - Y^\tau\| = \|X + A - (Y + A)\| = \|X - Y\| .$$

■

Theorem 4.3.10. *The map $\tau_A : \mathcal{E} \to \mathcal{E}$ is a bijection.*

Proof. Suppose that $\tau_A(X) = \tau_A(X')$, then by definition it follows that

$$X + A = X' + A$$

whereby subtraction of A from both sides of the equation above, we see that $X = X'$, therefore τ_A is injective. Now, let Y be an arbitrary point in $\mathcal{E}^2$, then to see that there exists a unique $X \in \mathcal{E}^2$ such that $\tau_A(X) = Y$, simply solve for X in

$$Y = X + A$$

thereby obtaining the unique solution $X = Y - A$, whence it follows that $\tau_A(X) = Y$, and so τ_A is surjective. ■

Proposition 4.3.11. *Translations preserve displacements.*

Proof. Suppose that the images of X and Y under τ_A are $X' = X + A$ and $Y' = Y + A$ respectively. Then

$$\begin{aligned} X' - Y' &= (X + A) - (Y + A) \\ &= X - Y. \end{aligned}$$

Thus a translation preserves a displacement. ∎

Proposition 4.3.12. *Translations preserve lines.*

Proof. Given $\tau_A : \mathcal{E}^2 \to \mathcal{E}^2$, let ℓ_{PQ} be any line the in the Euclidean plane. For any point $X' \in \ell_{P'Q'}$, it may written as

$$X' = pP' + qQ' \quad \text{with } p + q = 1.$$

Thus,

$$\begin{aligned} X' &= p(P + A) + q(Q + A) \\ &= pP + qQ + (p + q)A \\ &= pP + qQ + A \\ &= \tau_A(X) \end{aligned}$$

whence it follows that a translation preserves a line. ∎

Osometry Factorization Theorem

Theorem 4.3.13. *For any isometry α, there exists a linear isometry β and a translation τ_D such that*

$$\alpha = \tau_D \circ \beta,$$

That is, for every point, $X^\alpha = X^\beta + D$. What this really means is that any isometry can be expressed as the composition of two isometries, a linear isometry followed by a translation.

Proof. We want to show β is a linear isometry. But, first observe that for arbitrary X

$$\tau_D(X) = X + D$$

where D is a vector. Now, to revert $X + D$ back to its pre-image, we need to translate by $-D$, thus the inverse of τ_D is τ_{-D}, and so

$$\begin{aligned} \tau_{-D} \circ \alpha &= \tau_{-D} \circ (\tau_D \circ \beta) \\ &= (\tau_{-D} \circ \tau_D) \circ \beta \\ &= \mathrm{id} \circ \beta = \beta. \end{aligned}$$

So, let $\alpha(O) = D$ and $\beta := \tau_{-D} \circ \alpha$. Now, consider X^β, then we have

$$\begin{aligned} X^\beta &= (\tau_{-D} \circ \alpha)(X) \\ &= \tau_{-D}(X^\alpha) = X^\alpha - D. \end{aligned}$$

Thus, $X^\beta = X^\alpha - D \Leftrightarrow X^\alpha = X^\beta + D$. Now, to see that β is an osometry, observe that

$$\begin{aligned} O^\beta &= (\tau_{-D} \circ \alpha)(O) = \tau_{-D}(O^\alpha) \\ &= \tau_{-D}(D) = D - D = O. \end{aligned}$$

So, we see that β fixes the origin and $X^\beta = X^\alpha + D$, whence we see β is an isometry because it is the composition of two isometries. Also, $X^\alpha = X^\beta + D$, as required. ∎

Proposition 4.3.14. *The factorization of an isometry into the compostion of a linear isometry followed by a translation is unique.*

Proof. Suppose $\tau_D\beta_1 = \tau_E\beta_2$, then for the origin O in the plane, we have

$$\begin{aligned} \tau_D\beta_1(O) &= O^{\beta_1} + D \\ &= O + D \end{aligned}$$

and we also have

$$\begin{aligned} \tau_E\beta_2(O) &= O^{\beta_2} + E \\ &= O + E. \end{aligned}$$

Hence,

$$O + D = O + E \Leftrightarrow O - O = \mathbf{0} = E - D.$$

Thus, $E = D$, which implies that $\tau_D = \tau_E$. Now, since $\tau_D = \tau_E$ is invertible, it follows that multiplying both sides of $\tau_D\beta_1 = \tau_E\beta_2$ by $\tau_{-D} = \tau_{-E}$(the inverse), we have

$$\tau_{-D}\tau_D\beta_1 = \tau_{-E}\tau_E\beta_2 \Leftrightarrow \iota\beta_1 = \iota\beta_2 \Leftrightarrow \beta_1 = \beta_2.$$

Thus, the aforementioned factorization of an isometry is unique. ∎

The Barycenter Theorem

Lemma 4.3.15. *For any two vectors x, y in a euclidean space E, their cosine angle is determined by*

$$\cos(\theta) = \left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle = \frac{\langle x, y\rangle}{\|x\|\,\|y\|}.$$

Proof. TODO, ∎

Angle Preservation Theorem

Theorem 4.3.16. *Let $\phi_{(Q:P,R)}$ denote the angle between vectors $\overrightarrow{QP}$ and $\overrightarrow{QR}$, and $\phi_{\alpha(Q:P,R)} = \phi_{(Q^\alpha:P^\alpha,R^\alpha)}$, denote the angle between $\overrightarrow{\alpha(Q)\alpha(P)}$ and $\overrightarrow{\alpha(Q)\alpha(R)}$. An isometry α preserves angles in the sense that*

$$\cos(\phi_{Q:P,R}) = \cos\left(\phi_{\alpha(Q:P,R)}\right)$$

or

$$\phi_{Q:P,R} = \pm\phi_{\alpha(Q:P,R)}$$

Proof. In view of Lemma 4.3.15, let α be an isometry, and consider:

$$\begin{aligned}\cos\left(\phi_{\alpha(Q:P,R)}\right) &= \frac{\langle P^\alpha - Q^\alpha, R^\alpha - Q^\alpha\rangle}{\|P^\alpha - Q^\alpha\|\,\|R^\alpha - Q^\alpha\|} \\ &= \frac{\langle P^\alpha - Q^\alpha, R^\alpha - Q^\alpha\rangle}{\|P - Q\|\,\|R - Q\|} \qquad (\alpha \text{ is distance preserving})\end{aligned}$$

Now, we must attend to the following cases: α is a translation of the form τ_D, α is an osometry β, α is a composition of a linear isometry followed by a translation (refer to factorization theorem). So, we proceed as follows: let $\alpha = \tau_D$, then

$$\begin{aligned}\tau_D(P) - \tau_D(Q) &= (P + D) - (Q + D) \\ &= P - Q.\end{aligned}$$

Hence, $\langle P^\tau - Q^\tau, R^\tau - Q^\tau\rangle = \langle P - Q, R - Q\rangle$, and so it follows that

$$\cos\left(\phi_{\tau_D(Q:P,R)}\right) = \cos\left(\phi_{(Q:P,R)}\right).$$

Now, if $\alpha = \beta$, an osometry, then

$$\begin{aligned}\left\langle P^\beta - Q^\beta, R^\beta - Q^\beta\right\rangle &= \left\langle (P - Q)^\beta, (R - Q)^\beta\right\rangle \\ &= \langle P - Q, R - Q\rangle.\end{aligned}$$

Hence,

$$\cos\left(\phi_{\beta(Q:P,R)}\right) = \cos\left(\phi_{(Q:P,R)}\right)$$

Now, if $\alpha = \tau_D \circ \beta$, a general isometry, then

$$\begin{aligned}\cos\left(\phi_{(\tau_D\circ\beta)(P:Q,R)}\right) &= \cos\left(\phi_{\tau_D(P^\beta:Q^\beta,R^\beta)}\right) \\ &= \cos\left(\phi_{(P^\beta:Q^\beta,R^\beta)}\right) \\ &= \cos\left(\phi_{\beta(P:Q,R)}\right) = \cos\left(\phi_{(P:Q,R)}\right)\end{aligned}$$

as desired. ■

Isometries Preserve Barycentres

Lemma 4.3.17. *Isometries preserve barycentric coordinates. That, is for a family of weighted points $((A, a), (B, b), (C, c))$ such that $a + b + c$, we have*

$$(aA + bB + cC)^\alpha = aA^\alpha + bB^\alpha + cC^\alpha$$

where α is an isometry.

Proof. Suppose $\alpha = \tau_D$,

$$\begin{aligned}aA^{\tau_D} + bB^{\tau_D} + cC^{\tau_D} &= a(A + D) + b(B + D) + c(C + D) \\ &= aA + bB + cC + (a + b + c)D \\ &= (aA + bB + cC) + D = (aA + bB + cC)^{\tau_D}\end{aligned}$$

as required. Now, if $\alpha = \beta$, an osometry, then

$$\begin{aligned} aA^{\beta} + bB^{\beta} + cC^{\beta} &= (aA)^{\beta} + (bB)^{\beta} + (cC)^{\beta} \quad \text{(Linearity of } \beta) \\ &= (aA + bB + cC)^{\beta} \quad \text{(Linearity of } \beta) \end{aligned}$$

Now, if $\alpha = \tau_D \circ \beta$, then

$$\begin{aligned} a(\tau\beta(A)) + b(\tau\beta(B)) + c(\tau\beta(C)) &= a\left(\tau\left(A^{\beta}\right)\right) + b\left(\tau\left(B^{\beta}\right)\right) + c\left(\tau\left(C^{\beta}\right)\right) \\ &= a\left(\left(A^{\beta}\right)^{\tau}\right) + b\left(\left(B^{\beta}\right)^{\tau}\right) + c\left(\left(C^{\beta}\right)^{\tau}\right) \\ &= \tau\left(aA^{\beta} + bB^{\beta} + cC^{\beta}\right) \\ &= \tau\left((aA + bB + cC)^{\beta}\right) \\ &= \tau\beta(aA + bB + cC) = (aA + bB + cC)^{\tau\beta} \end{aligned}$$

as required. ∎

Fixed Points and Isometries

Lemma 4.3.18. *Let αbe an isometry, then if P and Q are fixed points under α, then every point on the line ℓ_{PQ} is a fixed point of α.*

Proof. Indeed, let $X \in \ell_{PQ}$z ∎

Lemma 4.3.19. *Let α be an isometry. If α has three fixed points P, Q, R which are not collinear, then $\alpha = \iota$, the identity.*

Proof. Indeed, let P, Q, R be non-collinear fixed points under α. Now, it follows by the previous lemma that every point on the lines ℓ_{PQ} and ℓ_{PR} and ℓ_{QR} is a fixed point under α. Let X be different from P and $M = \ell_{QR} \cap \ell_{XP}$. So, in view of our previous lemma, it follows that M is a fixed point and so is every point on the line ℓ_{XP} because X and M are fixed points on this line, hence P is fixed. The same can be argued for every point in the ℓ_{YP} ∎

4.4. Reflections: the Atoms of Isometries

A Primer on Projections

Let $\ell = \{cv \mid v \in E\}$ or $\ell = \text{span}\{v\}$, be a subspace of a Euclidean vector space E. Now, to consider the projection of some vector $a \in E$ onto $w \in \ell$, that is, $\text{proj}_\ell\, a$, we take note that there is some $c_a \in \mathbb{R}$, for which $\text{proj}_\ell(a) = c_a w$, and $w \in \ell$ And, so

$$\left\langle a - \text{proj}_\ell(a), w \right\rangle = 0 \quad \text{for all } w \in \ell \tag{4.14}$$

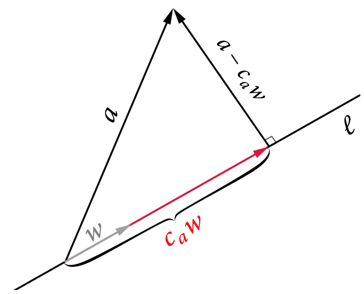


Figure 4.4.1: Vector Projection onto a Subspace

because the vector $a - \text{proj}_\ell(a)$ is perpendicular to any vector $w \in \ell$ (refer to the figure). Thus, we must have

$$\left\langle a - \text{proj}_\ell(a), w \right\rangle = \langle a - c_a w, w \rangle = 0$$

implying that

$$\langle a, w \rangle = c_a \langle w, w \rangle \Leftrightarrow c_a = \frac{\langle a, w \rangle}{\langle w, w \rangle}.$$

Hence,

$$\text{proj}_\ell(a) = \frac{\langle a, w \rangle}{\langle w, w \rangle} w. \tag{4.15}$$

This can be expressed more compactly as

$$\text{proj}_\ell(a) = \langle a, w \rangle \frac{w}{\|w\|^2}. \tag{4.16}$$

Here are some other useful facts:

1. The magnitude of the projection of a in the direction of $w \in \ell$ is determined by the cosine angle formula

$$\cos_{a,w}(\theta) = \frac{\langle a, w \rangle}{\|a\| \|w\|} \tag{4.17}$$

   Indeed, observe that

$$\left\|\text{proj}_\ell(a)\right\| = \frac{\langle a, w \rangle}{\langle w, w \rangle} \|w\| = \frac{\langle a, w \rangle}{\|w\|^2} \|w\| = \frac{\langle a, w \rangle}{\|w\|}. \tag{4.18}$$

   Then, $\|proj_\ell(a)\| = \|a\| \cos_{a,w}(\theta)$, as required. This is called the scalar projection of a onto w, where $w \in \ell$. It is also referred to as the component of a in the direction of w, or the w-component of a, denoted by $\text{comp}_w\, a$.

Nearest Point Theorem

Let W be a subspace of a Euclidean vector space E, and let a be a vector in E. The unique vector in W that is closest to a is $p = \text{proj}_W(a)$, that is, the orthogonal projection of a onto W. Hence,

$$\min_{w \in W} \|a - w\| = \left\|a - \text{proj}_\ell(a)\right\| = d(a, W)$$

which is the orthogonal distance between a and W.

Geometrical Definition of a Reflection

Let ℓ be a line in the plane going through the standard origin[1] we shall call it the mirror of the reflection σ_ℓ. Then, for every point Y on the line m, set $\sigma_m(Y) = Y$, hence every point on the line ℓ is a fixed point under the reflection σ_ℓ. Now, for points not on ℓ, that is X, let X_ℓ be the *nearest point* to X on the line ℓ as illustrated in the figure. In view of the nearest point theorem, X_ℓ is the orthogonal projection of X onto ℓ. Hence

$$X_\ell = \mathrm{proj}_\ell(X) \tag{4.19}$$

[1] We could really use any other point as an origin, but this choice makes the presentation simpler

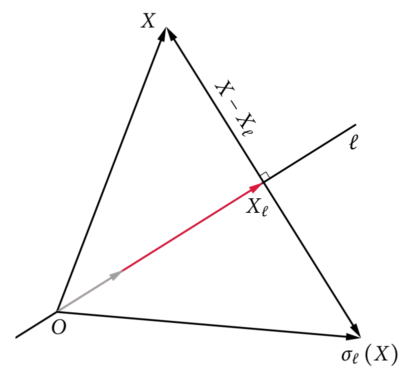


Figure 4.4.2: Reflection about a Line

Moreover, see that the ℓ is the perpendicular bisector of X and $\sigma_\ell(X)$. Hence,

$$X_\ell = \frac{X + \sigma_\ell(X)}{2} \Rightarrow \sigma_\ell(X) = 2X_\ell - X.$$

Now, since $X_\ell = \mathrm{proj}_\ell(X)$, it follows that

$$\begin{aligned}\sigma_\ell(X) &= 2\,\mathrm{proj}_\ell(X) - X\\ &= 2\frac{\langle X, W\rangle}{\|W\|^2}Y - X\end{aligned}$$

for $W \in \ell$. Also, the vector $\overrightarrow{X\sigma_\ell(X)}$ is twice $\overrightarrow{XX_\ell}$, that is,

$$\sigma_\ell(X) - X = 2(X_\ell - X). \tag{4.20}$$

Definition 4.4.1. The reflection $\sigma_\ell$in a line ℓ is the bijection of the plane leaves every point of ℓ fixed and maps every point $P \notin \ell$ to the point $\sigma_\ell(P)$ such that the line ℓ is the perpendicular bisector of the segment joining P and $\sigma_\ell(P)$,moreover $\sigma_\ell^2 = \iota$, which is an involution.

Proposition 4.4.2. *Reflections are Isometries.*

Proof. Indeed, by definition of isometry, we need to show $\|X^\sigma - Y^\sigma\| = \|X - Y\|$. Thus, first observe that

$$X^\sigma - Y^\sigma = (2X_\ell - X) - (2Y_\ell - Y) = 2(X_\ell - Y_\ell) - (X - Y)$$

Use projection...

$$\begin{aligned}&= 2\left(\mathrm{proj}_\ell(X) - \mathrm{proj}_\ell(Y)\right) - (X - Y)\\ &= 2\left(\frac{\langle X, W\rangle W}{\|W\|^2} - \frac{\langle Y, W\rangle W}{\|W\|^2}\right) - (X - Y), \quad W \in \ell\end{aligned}$$

No harm in treating W as a unit vector...

$$\begin{aligned}&= 2(\langle X, W\rangle - \langle Y, W\rangle)W - (X - Y)\\ &= 2\langle X - Y, W\rangle W - (X - Y).\end{aligned}$$

Now, we take the square of the norm of $X^\sigma - Y^\sigma$, and set the vector $Z = X - Y$, hence

$$\begin{aligned}\|X^\sigma - Y^\sigma\|^2 &= \langle 2\langle Z, W\rangle W - Z, 2\langle Z, W\rangle W - Z\rangle\\ &= \langle 2\langle Z, W\rangle W, 2\langle Z, W\rangle W - Z\rangle - \langle Z, 2\langle Z, W\rangle W - Z\rangle\\ &= 4\langle Z, W\rangle^2\langle W, W\rangle - 2\langle Z, W\rangle^2 - 2\langle Z, W\rangle^2 + \langle Z, Z\rangle\\ &= 4\langle Z, W\rangle^2\|W\|^2 - 4\langle Z, W\rangle^2 + \|Z\|^2\end{aligned}$$

Remember W is unit length...

$$= \|Z\|^2 = \|X - Y\|^2.$$

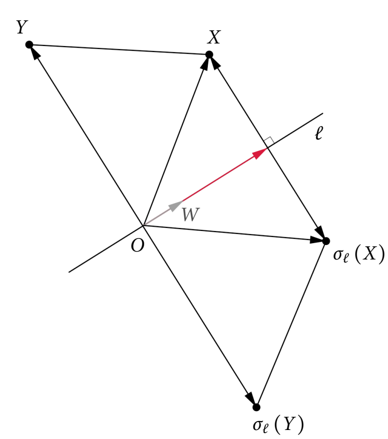


Figure 4.4.3: Reflections are Isometries

Therefore, it follows that $\|X^\sigma - Y^\sigma\| = \|X - Y\|$, and so σ_ℓ is an isometry. ∎

Properties of Reflections and Fixed Points

We list a few properties about fixed points below:

- Every point in the Euclidean Plane is a fixed point of the identity ι.
- No point of the plane is left fixed by a translation τ_D.
- A central reflection σ_Q has a single fixed point, namely Q.
- Every point on the mirror line ℓ of a reflection is a fixed pointed of σ_ℓ.

Properties of Reflections

Lemma 4.4.3. *Let α be a non-trivial isometry, that is, $\alpha \neq \iota$.*

1. *If $\alpha \neq \iota$, that is, there exists X in the plane such that $\alpha(X) \neq X$, then every fixed point of α (if any) lies on the perpendicular bisector of the segment joining X and X^α, we denote the perpendicular bisector as $\ell = \text{perbis}(X, X^\alpha)$.*
2. *If $\alpha \neq i$, has two fixed points, P, Q , then every point on the line ℓ_{PQ} (also denoted (PQ)) is a fixed point as well of α,moreover $\alpha = \sigma_\ell$.*
3. *If an isometry α has exactly one fixed point Q then it factors in to the composition of two reflections in lines through the fixed point, that is there exist lines n and m such that $\alpha = \sigma_n \sigma_m$ and $Q = \ell_{nm}$.*
4. *Every isometry is the composition of at most 4 reflections.*

Proof. (1) Let X be a point on the plane and Q, a point on the perpendicular bisector of the segment joining X and X^α, then by definition of isometry $\|X^\alpha - Q^\alpha\| = \|X - Q\|$. Now, since Q is on the perbis$(X.X^\alpha)$, it follows that Q is equidistant to the both X and its image X^α, thus $Q^\alpha = Q$. (2) Suppose α has two fixed points, P and Q, then every point X on the line ℓ_{PQ} is of the form

$$X = P + t(Q - P) = (1 - t)P + tQ$$

Thus, by the Barytheorem, we know that isometries preserve barycentres, thus

$$X^\alpha = (1 - t)P^\alpha + tQ^\alpha = (1 - t)P + tQ = X.$$

Now, to see why $\alpha$is a reflection, observe that P and Q are fixed points and let $K \notin \ell$. Now, since α is non-trivial isometry, it follows that the image of K under α is not a fixed point. So, by (1), P and Q must belong to the perpendicular bisector (which is ℓ) of some point $K \notin \ell$ and its image under α, namely K^α. Hence ℓ is the line or mirror of reflection such that $\alpha = \sigma_\ell$, and $K^\alpha = \sigma_\ell(K)$. To see why $\alpha = \sigma_\ell$, let $\omega = \sigma_\ell\alpha$, then the points P, Q and K under ω remain fixed, thereby implying that $\omega = \iota$. Indeed, multiplying both sides of $\omega = \sigma_\ell\alpha$ by σ_ℓ,we obtain

$$\sigma_\ell\omega = \sigma_\ell\sigma_\ell\alpha$$

reflection about the line ℓ is its own inverse...

$$= \iota\alpha$$

Thus, $\sigma_\ell \omega = \alpha$ implies that $\omega = \iota$ and that $\sigma_\ell = \alpha$, as required. (3) Suppose Q is the only fixed point under the nontrivial isometry α, then for any other point X, it must be by (1), that Q lies along the perbis of the segment joining X and X^α. Now, the composition $\omega = \sigma_m \alpha$ has two fixed points, namely P and X , where $P \in m$ and $X \notin m$. ■

4.5. The Geometry of Vectors

Affine geometry relies purely on the notion of parellel lines...

A Primer on Vectors

Definition 4.5.1. A mapping $f : \mathcal{E}^2 \to \mathcal{E}^2$ of the euclidean plane is called an **affine mapping** if for any four points A, B, C, D in the euclidean plane and if for every $t \in \mathbb{R}$ it follows from the relation

$$B - A = t(D - C)$$

that the images of these points under f satisfy

$$f(B) - f(A) = t(f(D) - f(C)).$$

Definition 4.5.2. If [2] $A = (a_1, a_2)$ and $B = (b_1, b_2)$ are two points in $\mathbb{R}^2$, then the **vector** $\overrightarrow{AB}$ or $B - A$ is determined by

$$\overrightarrow{AB} = B - A = (b_1 - a_1, b_2 - a_2).$$

[2] Insert illustration.

This is merely an arrow beginning at A and terminating B. We say that two vectors are the same if they are parallel, of the same length, and have the same orientation. So, if we're given points A, B along with points $X = (x_1, x_2)$ and $Y = (y_1, y_2)$, then vectors $B - A = Y - X$ if and only if $b_1 - a_1 = y_1 - x_1$ and $b_2 - a_2 = y_2 - x_2$. This doesn't necessarily mean that the vectors $B - A$ and $Y - X$ coincide. A single point $P = (p_1, p_2)$ in the Cartesian plane can be associated with the vector emanating from the origin[3] $O = (0, 0)$ and terminating at P. Essentially, both $\overrightarrow{OP}$ and P have the same coordinates, thus points in the Cartesian plane cannot only be thought of as ordered pairs, but vectors as well.

[3] O is also referred to as the **zero vector**.

Definition 4.5.3. Given two vectors $A = (a_1, a_2)$ and $B = (b_1, b_2)$, the **vector sum** $A + B$ is given by

$$A + B = (a_1 + b_1, a_2 + b_2)$$

which in itself defines a new vector.

Theorem 4.5.4. *Let A, B, and C be vectors, then the following properties of vector addition hold:*

(a) $A + B = B + A$ *(commutativity)*

(b) $A + (B + C) = (A + B) + C$ *associativity*

(c) $O + A = A + O = A$ *identity*

(d) $A + (-A) = O$ *inverse*

Proof. We prove the above respectively:

(a) Given vectors $A = (a_1, a_2)$ and $B = (b_1, b_2)$, we have

$$\begin{aligned} A + B &= (a_1 + b_1, a_2 + b_2) \\ &= (b_1 + a_1, b_2 + a_2) \quad \text{commutativity in } \mathbb{R} \\ &= B + A. \end{aligned}$$

(b) Given vectors A,B, and C, the sum

$$\begin{aligned} A + (B + C) &= (a_1, a_2) + (b_1 + c_1, b_2 + c_2) \\ &= (a_1 + (b_1 + c_1), a_2 + (b_2 + c_2)) \\ &= ((a_1 + b_1) + c_1, (a_2 + b_2) + c_2) \quad \text{associativity in } \mathbb{R} \\ &= (a_1 + b_1, a_2 + b_2) + (c_1, c_2) \\ &= (A + B) + C. \end{aligned}$$

(c) The proof of this property follows from the proof for **(a)**.

(d) Provided the vector $-A = (-a_1, -a_2)$, this property simply follows from vector addition.

□

Definition 4.5.5. Let $A = (a_1, a_2)$ and $r \in \mathbb{R}$ be a scalar. The **scalar multiple** of A is defined by

$$rA = (ra_1, ra_2).$$

Remark 4.5.6. If the scalar $r = 0$, then $0A = O$, the zero vector. Indeed, by definition $0A = (0 \cdot a_1, 0 \cdot a_2) = (0, 0)$, which is equivalent to O. If $r = 1$, then rA is simply A itself. Refer to the illustration for more examples.

4.6. Other Transformations of the Plane

Homothety

Definition 4.6.1. Let C be an arbitrary but fixed point in the Euclidean Plane $\mathcal{E}^2$, and let $t \in \mathbb{R}$ be a non-zero scalar. The **central dilation** or **homothety** with center of dilation C and dilation factor $t \neq 0$, is the mapping $\delta_{C,t} : \mathcal{E}^2 \to \mathcal{E}^2$ that associates to a point X the image point X', uniquely determined as $\delta_{C,t}(X) = X' = C + t(X - C)$, hence $\overrightarrow{CX'} = t\overrightarrow{CX}$. Conversely, the preimage of X' is the point X satisfying

$$X = C + \frac{1}{t}(X' - C)$$

thus, $\delta_{C,t}$ is a transformation with inverse $\delta_{C,t}^{-1} = \delta_{C,\frac{1}{t}}$.

Theorem 4.6.2. *A homothety is bijective.*

Proof. Let $\delta_{C,t} : \mathcal{E}^2 \to \mathcal{E}^2$ be a homothety in the Euclidean plane where $t \neq 0$. We will first demonstrate that $\delta_{C,t}$ is injective (one-to-one). Suppose $\delta_{C,t}(X) = \delta_{C,t}(X')$ and $t \neq 0$. It follows by definition that $\delta_{C,t}(X) = C + t(X - C)$ and $\delta_{C,t}(X') = C + t(X' - C)$.

A mapping $f : A \to B$ is injective if

$$f(x') = f(x) \Rightarrow x' = x$$

for all $x', x \in A$.

Furthermore, from our hypothesis it follows that

$$C + t(X - C) = C + t(X' - C)$$
$$\Updownarrow$$
$$tX - tC = tX' - tC$$
$$\Updownarrow$$
$$t(X - X') = O.$$

Now, since $t \neq 0$, we conclude that $X = X'$ thereby indicating that a homothety is injective. To see that $\delta_{C,t}$ is a surjective mapping, suppose that Y is an arbitrary point in $\mathcal{E}^2$, then we must show that there exists a point$X$ in $\mathcal{E}^2$ such that $\delta_{C,t}(X) = Y$. Indeed, solving for X in

$$Y = C + t(X - C)$$

we obtain the unique solution

$$X = \frac{1}{t}Y + \frac{t-1}{t}C$$

whence it follows that $\delta_{C,t}(X) = Y$. ■

Proposition 4.6.3. *A homothety scales displacements.*

Proof. Suppose that $X'$and Y' are the image points of X and Y under homothety $\delta_{C,t}$, then displacement

$$\begin{aligned} Y' - X' &= (C + t(Y - C)) - (C + t(X - C)) \\ &= t(Y - X) \end{aligned}$$

thus under a homothety a displacement vector is taken to a multiple of itself. ■

Proposition 4.6.4. *A homothety preserves lines.*

Proof. Given homothety $\delta_{C,t}$, let line ℓ_{AB} be any line in the Euclidean plane, then for any point $X' \in \ell_{A'B'}$, we have

$A' = \delta_{C,t}(A)$ and $B' = \delta_{C,t}(B)$

$$X' = aA' + bB' \quad \text{with } a + b = 1.$$

Thus,

$$\begin{aligned} X' &= a(C + t(A - C)) + b(C + t(B - C)) \\ &= a((1-t)C + tA) + b((1-t)C + tB) \\ &= (a+b)(1-t)C + t(aA + bB) \\ &= (1-t)C + t(aA + bB) \\ &= C + t((aA + bB) - C) = \delta_{C,t}(X). \end{aligned}$$

Therefore every point $X' \in \ell_{A'B'}$ is the image under the central dilation of a point $X \in \ell_{AB}$. Furthermore

$$B' - A' = (C + t(B - C)) - (C + t(A - C)) = t(B - A)$$

whence, we discern that a homothety takes (maps) a line to a line that is parallel to it. ■

Proposition 4.6.5. *A homothety preserves barycentric coordinates.*

☜ This addresses a file card question.

Proof. Suppose $A, B, C \in \mathcal{E}^2$ are noncollinear, then we know there is a point P in the euclidean plane such that

$$P = aA + bB + cC \quad \text{with } a + b + c = 1,$$

So, if P' is the image of P under homothety $\delta_{Q,t}$, then

$$\begin{aligned} P' &= aA' + bB' + cC' \\ &= a\left(Q + t\left(A - Q\right)\right) + b\left(Q + t\left(B - Q\right)\right) + c\left(Q + t\left(C - Q\right)\right) \\ &= a\left(\left(1 - t\right)Q + tA\right) + b\left(\left(1 - t\right)Q + tB\right) + c\left(\left(1 - t\right)Q + tC\right) \\ &= \underbrace{\left(a + b + c\right)}_{1}\left(1 - t\right)Q + t\left(aA + bB + cC\right) \\ &= Q + t\left(\left(aA + bB + cC\right) - Q\right) = \delta_{Q,t}\left(P\right). \end{aligned}$$

Thus barycentric coordinates are preserved under central dilations. ∎

Proposition 4.6.6. *A homothety preserves circles.*

Proof. Suppose we have a circle K that is a circle going through three distinct points A, B, C with center D and radius r. We'll define k as $k := \odot\left(D, r\right) = \odot\left(A, B, C\right)$ and let $\delta_{D,t} : \mathcal{E}^2 \to \mathcal{E}^2$ be a homothety in the Euclidean plane. Then we shall show that

$$\delta_{D,t}\left(K\right) = \bigodot\left(\delta_{D,t}\left(D\right), tr\right) = \bigodot\left(\delta_{D,t}\left(A\right), \delta_{D,t}\left(B\right), \delta_{D,t}\left(C\right)\right).$$

Now, let X be any point on the circumference of K, then the displacement between X and the center D is $X - D$, whereby previous results we know that displacements are preserved under central dilations. Therefore, since X was arbitrarily chosen from the set of loci, we know displacements between any member of the locus of a circle K and its center D are preserved under a central dilation. As a consequence, the image of K, which we shall call K' under homothety $\delta_{D,t}$ is the locus comprised of points whose distance from D is tr. ∎

4.7. Applications of Homotheties and Translations

Proof of Euler's Line Theorem with Central Dilations

Euler Line Theorem. *The orthocenter, barycenter, and circumcenter of a triangle are collinear, and the barycenter is two-thirds the distance from the orthocenter to the circumenter.*

Proof. Given $\triangle ABC$, with orthocenter H and centroid (barycenter) G as the center of dilation for homothety $\delta_{G,-1/2} : \mathcal{E}^2 \to \mathcal{E}^2$. We've chosen a dilation factor of $-1/2$. Under this dilation A maps to A', B to B' , and C to C', thereby obtaining another triangle whose vertices are the images of A, B, C under the given homothety. Notice that $\triangle ABC$ and $\triangle A'B'C'$ share G as their centroid. This is because the medians of $\triangle ABC$ bisect the sides of $\triangle A'B'C'$. Now, the image of H, which we shall call D, under the given homothety is

$$D = \delta_{G,-\frac{1}{2}}\left(H\right) = G + -1/2\left(H - G\right)$$

☜ We've obtained Das the circumcenter.

Thus $D - G = 1/2\left(G - H\right)$. From this we can see that D, G, H are collinear[4]. Furthermore, the barycenter G is one third the distance from the circumcenter D to the orthocenter H, and two thirds the distance vice-versa. ∎

[4] Recall that central dilations preserve lines

Proof of Nine Point Circle

Euler's Trapezoid. *Let $A^\star$ be the foot of the altitude on A, which is a vertex of triangle $\triangle ABC$. Since both $A^\star H$ and $A'D$ are perpendicular to the base BC, and N is the midpoint of the top side of this rectangular trapezoid $A^\star A'DH$, the following distances are equal $|A^\star - N| = |A' - N|$.*

Proof. Suppose we have triangle $\triangle ABC$ with centroid G and a circumcircle K going about the vertices of $\triangle ABC$, that is $K = \odot(A, B, C)$. The image of circle K under homothety $\delta_{G,-1/2}$ is another circle K', where $K' = \odot(A', B', C')$. This triple is is comprised of the the midpoints of the sides of $\triangle ABC$. Now, let A'', B'', C'' be the three midpoints of each segment extending (respectively) from the vertices A, B, C to the orthocenter H, where $H = \delta_{G,-2}(D)$. Also, let $A^\star, B^\star, C^\star$ be the feet of the altitudes emanating from A, B, C respectively. Now, for circumcenter D

$$\delta_{H,\frac{1}{2}}(D) = H + \frac{1}{2}(D - H) = N.$$

Where N is center of circle K'. Now, for $\delta_{H,\frac{1}{2}}(K)$, we have

$$\delta_{H,-\frac{1}{2}}(K) = \odot(A'', B'', C'') = K'.$$

Observe that $|A^\star - N| \ldots$ ∎

Proof of Special Case Desargue's Theorem

Proposition 4.7.1. *Suppose we're given triangles $\triangle ABC$ and $\triangle A'B'C'$ in the euclidean plane where the lines $\ell_{AA'}, \ell_{BB'}, \ell_{CC'}$ concur at a point D. If $\ell_{AB} \parallel \ell_{A'B'}$ and $\ell_{BC} \parallel \ell_{B'C'}$, then $\ell_{AC} \parallel \ell_{A'C'}$.*

Proof. Under a homothety it follows that for every vector Y there is unique $t \in \mathbb{R}$ and vector X, such $\delta_{D,t}(X) = Y$. Thus, for A' there is a unique $t \in \mathbb{R}$ such that $\delta_{D,t}(A) = A'$. Now, since $\ell_{AB} \parallel \ell_{A'B'}$ and $\ell_{BC} \parallel \ell_{B'C'}$, it follows that

$$\delta_{D,t}(\ell_{AB}) = \ell_{A'B'} \Leftrightarrow B' - A' = t(B - A)$$

and

$$\delta_{D,t}(\ell_{BC}) = \ell_{B'C'} \Leftrightarrow B' - C' = t(B - C).$$

Thus, $\delta_{D,t}(B) = B'$ and $\delta_{D,t}(C) = C'$. So, $C' - A' = (D + t(C - D)) - (D + t(A - D)) = t(C - A)$, thereby implying that $\ell_{AC} \parallel \ell_{A'C'}$. ∎

4.8. Exercises in Transformational Geometry

Exercise A. Use the Barytheorem for isometries and other facts about barycentric coordinates to prove that two isometries α and ω that have identical values on 3 non-collinear points, $\triangle ABC$, i.e. , $A^\alpha = A^w, B^\alpha = B^\omega, C^\alpha = C^\omega$ then $\alpha = \omega$. Recall that two transformation are the same if they have the same values on every point in the plane. **Answer:**
Exercise B. Prove that a reflection σ_ℓ is an involution.
Exercise C. Prove that for an isometry α, the identity $\alpha\sigma_\ell = \sigma_l\alpha$ for all lines ℓ implies that $\alpha = \iota$.
Exercise D. Draw a figure and give a geometrical arguement that shows that a central reflection σ_Q factors into the product of two reflections, that is, $\sigma_Q = \sigma_a\sigma_b$, where a, b are any two mirrors intersecting at Q and perp to each other.

Exercise E. Let α be any isometry and σ_Q a central reflection. Use the previous exercise and the conjugacy theorem for reflections to show that its conjugate is given by $\alpha\sigma_Q\alpha^{-1} = \sigma_{Q^\alpha}$.
Exercise F. Prove that for an isometry α, the identity $\alpha\sigma_C = \sigma_C\alpha$ for all points C implies that $\alpha = \iota$.
Exercise G. Use the theory for composing reflections developed in the rotations lession to prove that the product of two reflections in perpepdicular lines it the central reflection in their intersection point.
Exercise H. Let M be a point on a line m, that is, (Mm). Use the conjugacy theorem for reflections to those that σ_M and σ_m commute and find the single reflection equal to their product.
Exercise I.
Exercise J.
Exercise K. Use the conjugacy theorems and other properties of reflection in the notes to prove that $m = \text{perbis}\,(P, Q)$, that is, the perpendicular bisector of a segment, if and only if

$$\sigma_m\sigma_Q\sigma_m\sigma_p = \iota$$

Exercise L. Use the methods of the lessons to that k is a line parallel to the line (AB) if and only if

$$\sigma_B\sigma_k\sigma_B\sigma_A\sigma_k\sigma_A = \iota$$

Exercise M. Use the properties of translatiosn to show that for two parellel lines m and n, we have $\sigma_n\sigma_m = \sigma_m\sigma_n$ iff $m = n$.
Exercise N. Let τ_D be a translation and P a point. Prove that $\tau_D\sigma_P$ is a central reflection σ_A with center $A = P + \frac{1}{2}D$. Similarly, $\sigma_P\tau_D = \sigma_B$ is a central reflection with center $B = P - \frac{1}{2}D$.
Exercise O. Let τ_D be a translation and β an osometry (linear isometry). Use the conjugation theorem for translations to show that in this case we do have $\beta\tau_D\beta^{-1} = \tau_{D^\beta}$.
Exercise P. Given τ_D and a point Q. Find two further points A, B for which $\tau_D = \sigma_Q\sigma_B = \sigma_A\sigma_Q$.
Exercise Q. Consider the triangle $\triangle ABC$ and an arbitrary fourth point G. Prove that G is the centroid of $\triangle ABC$ iff

$$\sigma_G\sigma_C\sigma_G\sigma_B\sigma_G\sigma_A = \iota.$$

Exercise R. Consider a triangle $\triangle ABC$ (oriented counter-clockwise) with positively oriented interior angles α, β, γ and A, B, C. Prove that $\rho_{A,2\alpha}\rho_{B,2\beta}\rho_{C,2\gamma} = \iota$.
Exercise S. Prove that the perpendicular bisectors

$$a = \text{perbis}\,(BC)\,, \quad b = \text{perbis}\,(CA)\,, \quad c = \text{perbis}\,(AB)$$

of the sides $\triangle ABC$ are concurrent, using the classification of isometries.
Exercise T. Consider a triangle $\triangle ABC$ (oriented counter-clockwise) with positively oriented interior angles α, β, γ at A, B, C. Prove that $\rho_{C,2\gamma}\rho_{B,2\beta}\rho_{A,2\alpha}$ is a translation different from the identity.

5 — Appendix

5.1. Rudiments

- A ***class*** is an arbitrary collection of objects. A ***set*** is a class that occurs as a member of some class. A class which is not a set is a ***proper class***.

- Suppose a non-empty set I known as the index set is given such that for each $i \in I$ an object $x_i$is associated, then the ***family*** $(x_i)_{i \in I}$ is the rule which assigns to each element $i \in I$ its object x_i.

  - Consider the following example: families with index set $I = \{2, 4, 6, 8\}$ are tuples of the form (x_2, x_4, x_6, x_8), and families with index set $I = \mathbb{N}$ are sequences of the form $(x_1, x_2, x_3, x_4, \ldots)$. Also, if $I = \{1, 2\}$, then families with this index set are simply ordered pairs of the form (x_1, x_2).

- A set A is called a ***subset*** of a set B if every element of A is an element of B, in which case we write $A \subseteq B$. If A is not a subset of B we write $A \nsubseteq B$. Now, $A \neq B$, but $A \subseteq B$, then we say that A is a ***proper subset*** of B, thereby expressing this as $A \subset B$.

Definition 5.1.1. Let A and B be sets, then:

1. the ***union*** $A \cup B$ of A and B consists of all elements belonging to A or B, or both. We must take note that the "or" statement in mathematics si non-exclusive. Now, if $(A_i)_{i \in I}$ is a family of indexed sets (where I is non-empty index set), then the union $\bigcup_{i \in I} A_i$ consists of all elements belonging to atleast of one of the sets A_i, that is, $x \in \bigcup_{i \in I} A_i$ if and only if there exists an index $i_0 \in I$ such that $x \in A_{i_0}$.

2. the ***intersection*** $A \cap B$ of A and B consists of all elements common to both A and B. Now, if $(A_i)_{i \in I}$ is a family of indexed sets (where I is non-empty index set), then the intersection $\bigcap_{i \in I} A_i$ consists of all elements belonging to all of the sets A_i, that is, $x \in \bigcap_{i \in I} A_i$ if and only if $x \in A_i$ for all $i \in I$.

3. The ***complement*** $A \setminus B$ of B in A consists of all elements in A not belonging to B, hence $A \setminus B$ is obtained from A by removing those elements belonging to B as well, if any.

Definition 5.1.2. Let X be a set, then:

1. Two subsets $A, B \subseteq X$ are disjoint if and only if $A \cap B = \varnothing$. If $(A_i)_{i \in I}$ is a family of indexed subsets of X, then the sets A_i are ***pairwise disjoint*** if and only if $A_i \cap A_j = \varnothing$ whenever $i \neq j$.

2. If $X = \bigcup_{i \in I} A_i$, where A_i are pairwise disjoint, then we say X is the ***disjoint union*** of the sets A_i, and that the sets A_i form a ***partition*** of X, in which case we write $X = \biguplus_{i \in I} A_i$.

Definition 5.1.3. Let A and B be sets, then the ***cartestian product***, denoted by $A \times B$ is set of all pairs of (a, b) for which $a \in A$ and $b \in B$. More generally if $A_1, A_2, \ldots, A_n$ are sets, then the cartesian product $A_1 \times A_2 \times \cdots \times A_n$ of these sets if the set of all n-tuples $(a_1, a_2, \ldots, a_n)$

where $a_i \in A_i$ for $1 \le i \le n$. Even more generally if $(A_i)_{i\in I}$ is any family of indexed sets, then the cartesian product $\prod_{i\in I} A_i$ of the sets A_i is the set of all tuples $(a_i)_{i\in I}$, that is, all families of elements a_i where $a_i \in A_i$ for each $i \in I$.

Definition 5.1.4. Let X and Y be sets, then a ***function***, ***map***, or ***mapping***, of X into Y is a rule which assigns to each $x \in X$ a unique[1] element $y = f(x) \in Y$. We write $f : X \to Y$ to say that f is a function of X into Y. X is called the ***domain*** and Y is called the ***codomain*** of X. If $f(x) = y$, we say that y is the ***value of the argument*** x under the function f. For any set $A \subseteq X$, we call $f(A) := \{f(a) : a \in A\}$ the ***image*** or ***range*** of A under f. For any set $B \subseteq Y$, we call $f^{-1}(B) := \{x \in X : f(x) \in B\}$ the ***pre-image*** of B under f. The set $\Gamma(f) := \{(x, f(x)) : x \in X\}$ is the ***graph*** of f. The set of all functions $f : X \to Y$ is denoted by Y^X. Two functions $f : X \to Y$ and $g : X \to Y$ are ***equal*** if and only if $f(x) = g(x)$ for all $x \in X$.

[1] The assignment of a unique element in the codomain Y to each element in the domain X via the function f is what makes the function f **well-defined**. More, precisely $f : X \to Y$ is well-defined if for each $x \in X$ there is a unique $y \in Y$ such that $f(x) = y$.

Definition 5.1.5. If $A \subseteq B$, then the mapping $\iota : A \to B$ defined by $\iota(a) = a$ for all $a \in A$ is called the ***inclusion map*** of A into B. Let X, Y be sets. If $f : X \to Y$ is a function and if $A \subseteq X$, then the mapping $f\big|_A(a) = f(a)$ for all $a \in A$ is called the ***restriction*** of f to A in$X$. Conversely, if $g : A \to Y$ is a mapping then every function $f : X \to Y$ with $g = f\big|_A$, such that $g(a) = f(a)$ for all $a \in A$, is an ***extension*** of g to X. The ***identity map*** id_X of X is the mapping $f : X \to X$ defined by $f(x) = x$ for all $x \in X$.

Definition 5.1.6. Let X, Y, Z be sets and $f : X \to Y$, $g : Y \to Z$ be functions. Then the ***composition*** $g \circ f : X \to Z$ is the function from X into Z which is defined by $(g \circ f)(x) := g(f(x))$.

Definition 5.1.7. Let $f : X \to Y$ be a function. We say that f is ***injective*** or ***one-to-one*** if for $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$, that is, each element of Y occurs as the image of at most one element in X. We say that f is ***surjective*** or ***onto*** if for each element $y \in Y$ there is an element $x \in X$ such that $f(x) = y$. We say f is ***bijective*** if and only if f is both injective and surjective.

Definition 5.1.8. Let $(A_i)_{i\in I}$ be an indexed family of sets (I is a non-empty index set). The ***product*** of the sets $\{A_i\}_{i\in I}$, written as $\prod_{i\in I} A_i$ consists of all functions a with domain, the indexing set I, having the property that for each $i \in I$, $a(i) \in A_i$. We may express this compactly as:

$$\prod_{i\in I} A_i := \left\{ a : I \to \bigcup_{i\in I} A_i \, : \, a(i) \in A_i \; \forall i \in I \right\}.$$

☜This is rather confusing but let's take a closer look at it. What the set builder notation here is telling us is that if $a \in \prod_{i\in I} A_i$, then $a(i) \in A_i$, so suppose we consider the cartesian product of two sets, namely, $A_1 \times A_2$, which can be expressed as

$$A_1 \times A_2 = \prod_{i\in\{1,2\}} A_i,$$

then an element of this product can be a function $a : \{1,2\} \to A_1 \cup A_2$, where a is a function of $I = \{1,2\}$ into the union $A_1 \cup A_2$, for which $a(1) \in A_1$ and $a(2) \in A_2$. Thus a represents a function from the set $\{1,2\}$ to the set $A_1 \cup A_2$, and the arguments 1 and 2, take on the values A_1 and A_2, respectively under the function a.

Axiom 5.1.9 (Axiom of Choice)**.** *If we have an indexed family of non-empty sets, that is, $(A_i)_{i\in I}$, then we can choose an element from each A_i at a time. More precisely, there exists a* ***choice function*** *$f : I \to \bigcup_{i\in I} A_i$ such that $f(i) \in A_i$ for all i, again meaning that f picks one element from each of the sets A_i.*

Definition 5.1.10. Suppose that a diagram shows arrows between points which represent mappings between sets. Such a diagram is called a commutative diagram if all compositions of mappings leading from one point in the diagram to another coincide.

Example 5.1.11. Consider the following diagrams:

$$\begin{tikzcd} A \arrow[rr, "g"] \arrow[dr, "f"'] & & B \\ & C \arrow[ur, "h"'] & \end{tikzcd} \qquad \begin{tikzcd} \mathcal{V} \arrow[r, "\mathsf{T}"] \arrow[d, "\phi_\beta"'] & \mathcal{W} \arrow[d, "\phi_\gamma"] \\ \mathbb{F}^n \arrow[r, "\mathsf{L}_A"'] & \mathbb{F}^m \end{tikzcd}$$

These commute since for the first diagram, we have $f = h \circ g$ and for the second, we have $\mathsf{L}_A \phi_\beta = \phi_\gamma \mathsf{T}$.

5.2. Fields

Field Axioms

Conjugation on a Field

5.3. Vector Spaces

Characterization of Vector Spaces

Definition 5.3.1. A $\mathbb{F}$-***vector space*** or **vector space over a field** $\mathbb{F}$ is a set $\mathcal{V}$ paired with operations + (addition) and $\cdot$ (multiplication), which are functions defined as such:

$$+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}; \quad (x, y) \mapsto x + y,$$

Whenever we say "let $\mathcal{V}$ be a vector space" we tacitly assume it is a vector space over an arbitrary field $\mathbb{F}$, unless otherwise specified.

and

$$\cdot : \mathbb{F} \times \mathcal{V} \to \mathcal{V}; \quad (c, x) \mapsto c \cdot x.$$

Furthermore, the following axioms (labeled **A1-A10**) must be satisfied:

1. $x + y \in \mathcal{V}$ for any $x, y \in \mathcal{V}$ (Closure under Addition)
2. $x + y = y + x$ for all $x, y \in \mathcal{V}$ (Commutativity under Addition)
3. $(x + y) + z = x + (y + z)$ for all $x, y, z \in \mathcal{V}$ (Associativity under Addition)
4. There exists $0_\mathcal{V} \in \mathcal{V}$ such that $x + 0_\mathcal{V} = x$ for all $x \in \mathcal{V}$ (Existence of Additive Identity)
5. For all $x \in \mathcal{V}$, there exists $y \in \mathcal{V}$ such that $x + y = 0_\mathcal{V}$. (Existence of Additive Inverse)
6. $cx \in \mathcal{V}$ for any $c \in \mathbb{F}, x \in \mathcal{V}$ (Closure under Multiplication)
7. $c(x + y) = cx + cy$ for any $c \in \mathbb{F}, x, y \in \mathcal{V}$ (Distributivity of Scalars amongst Vectors)
8. $(a + b)x = ax + bx$ for any $a, b \in \mathbb{F}, x \in \mathcal{V}$ (Distributivity of Vectors amongst Scalars)
9. $a(bx) = (ab)x$ for any $a, b \in \mathbb{F}, x \in \mathcal{V}$. (Associativity under Multiplication)
10. $1x = x$ for any $x \in \mathcal{V}$. (Multiplicative Identity)

Theorem 5.3.2 (Cancellation Law for Vector Addition). *Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$. If $x, y, z \in \mathcal{V}$ such that $x + z = y + z$, then $x = y$.*

Proof. By **A5**, we know that for $z \in \mathcal{V}$ there exists an additive inverse $u \in \mathcal{V}$ such that $z+u = 0_{\mathcal{V}}$. Thus,

$$\begin{aligned} x &= x + 0_{\mathcal{V}} = x + (z+u) = (x+z) + u \\ &= (y+z) + u = y + (z+u) = y + 0_{\mathcal{V}} = y \end{aligned}$$

follows from **A3** and **A4**. ■

Theorem 5.3.3. *Let $\mathcal{V}$ be a vector space over a field $\mathbb{F}$. Then following statements are true:*

1. *$\mathcal{V}$ has a unique additive identity, that is, the zero vector in $\mathcal{V}$ is unique.*
2. *Every vector $x \in \mathcal{V}$ has a unique additive inverse.*
3. *$0x = 0_{\mathcal{V}}$ for all $x \in \mathcal{V}$, where $0 \in \mathbb{F}$.*
4. *$(-c)x = -(cx) = c(-x)$ for all $x \in \mathcal{V}$ and $c \in \mathbb{F}$.*
5. *$c0_{\mathcal{V}} = 0_{\mathcal{V}}$ for all $c \in \mathbb{F}$.*

Proof. **(1)** Suppose $0_{\mathcal{V}}$ and $0'_{\mathcal{V}}$ are both additive identities in $\mathcal{V}$, then by **A4**, it follows that

$$0_{\mathcal{V}} + 0'_{\mathcal{V}} = 0_{\mathcal{V}} \quad \text{and} \quad 0'_{\mathcal{V}} + 0_{\mathcal{V}} = 0'_{\mathcal{V}}.$$

Now, by **A2**, $0_{\mathcal{V}} + 0'_{\mathcal{V}} = 0'_{\mathcal{V}} + 0_{\mathcal{V}}$; then, by **Theorem 1.1**, $0_{\mathcal{V}} = 0'_{\mathcal{V}}$, thereby establishing uniqueness of the additive identity in $\mathcal{V}$. **(2)** Let $x \in \mathcal{V}$, and suppose y and y' are additive inverses for x, such that $x + y = 0_{\mathcal{V}} = x + y'$. Then, by application of **A4** and**A3**

$$y = y + 0_{\mathcal{V}} = y + (x+y) = y + (x+y') = (y+x) + y' = 0_{\mathcal{V}} + y = y'.$$

Alternatively, we may see that from the hypothesis, $y = y'$ follows immediately from the lemma previously proven. Thus, uniqueness of the additive inverse in $\mathcal{V}$ is asserted. **(3)** By **A8** it follows that

$$0x = \underbrace{(0+0)x = 0x + 0x}_{\mathbf{A8}}.$$

Now, since $0x \in \mathcal{V}$ (why?), we know by **A5** , it has an additive inverse $y \in \mathcal{V}$ such that $0x + y = 0_{\mathcal{V}}$. So, add y to both sides of the equation above, thereby obtaining

$$0_{\mathcal{V}} = 0x + y = \underbrace{(0x+0x) + y = 0x + (0x+y)}_{\mathbf{A3}} = \underbrace{0x + 0_{\mathcal{V}} = 0x}_{\mathbf{A4}} \tag{$\star$}$$

as desired. **(4)** Since $cx \in \mathcal{V}$, then we know it has unique additive inverse in $\mathcal{V}$, namely $-(cx)$, such that

$$cx + [-(cx)] = 0_{\mathcal{V}}.$$

Now, if $cx + (-c)x = 0_{\mathcal{V}}$ as well, uniqueness of the additive inverse suggests that $(-c)x = -(cx)$. Indeed, by **A8** and $(\star)$, it follows that

$$\underbrace{cx + (-c)x = [c + (-c)]x}_{\mathbf{A8}} = \underbrace{0x = 0_{\mathcal{V}}}_{\star}.$$

Thus, since

$$cx + (-c)x = 0_{\mathcal{V}} = cx + [-(cx)]$$

it follows from **A2** (applied to both sides of this equation) and the lemma above that $(-c)\,x = -\,(cx)$. In particular $(-1)\,x = -x$, which in addition to **A9** implies that

$$c\,(-x) = c\,[(-1)\,x] = [c\,(-1)]\,x = (-c)\,x.$$

Consequently, estabilish the chain of equalities $(-c)\,x = -\,(cx) = c\,(-x)$, as desired. **(4)** By **A7**, it follows that

$$c0_{\mathcal{V}} = c\,(0_{\mathcal{V}} + 0_{\mathcal{V}}) = c0_{\mathcal{V}} + c0_{\mathcal{V}},$$

now, add the additive inverse of$c0_{\mathcal{V}}$(call it$y$) to both sides of the equation above, thereby obtaining

$$0_{\mathcal{V}} = c0_{\mathcal{V}} + y = \underbrace{(c0_{\mathcal{V}} + c0_{\mathcal{V}}) + y = c0_{\mathcal{V}} + (c0_{\mathcal{V}} + y)}_{\mathbf{A3}} = \underbrace{c0_{\mathcal{V}} + 0_{\mathcal{V}} = c0_{\mathcal{V}}}_{\mathbf{A4}}.$$

And the proof is complete. ■

5.4. Multilinear Forms

Supplementary Material on 2-forms

Definition 5.4.1. A ***bilinear form*** (2-form) $\varphi : \mathcal{V} \times \mathcal{V} \to F$ is a function $\varphi\,(x, y)$ such that for every fixed $\tilde{x} \in \mathcal{V}$, we have $\varphi\,(\tilde{x}, y)$ as a linear function of $y \in \mathcal{V}$, and for every fixed $\tilde{y} \in \mathcal{V}$, we have$\varphi\,(x, \tilde{y})$ as a linear function of $x \in \mathcal{V}$. Thus, it is linear in each of its arguments, and thereby satisfies both of the following:

$$\varphi\,(cx_1 + x_2, y) = c\varphi\,(x_1, y) + \varphi\,(x_2, y)$$

and

$$\varphi\,(x, cy_1 + y_2) = c\varphi\,(x, y_1) + \varphi\,(x, y_2)$$

for $x, y, x_1, x_2 y_1, y_2 \in \mathcal{V}$ and $c \in F$.

Definition 5.4.2. A bilinear form $\varphi : \mathcal{V} \times \mathcal{V} \to F$ on a vector space $\mathcal{V}$ is:

1. ***symmetric*** if $\varphi\,(x, y) = \varphi\,(y, x)$
2. ***antisymmetric*** if $\varphi\,(x, y) = -\varphi\,(y, x)$.
3. ***alternating*** if $\varphi\,(x, x) = 0$

for all $x, y \in \mathcal{V}$.

6 — Bibliography

[1] M. Audin. *Geometry*. Universitext. Springer Berlin Heidelberg, 2012.

[2] M. Berger, M. Cole, and S. Levy. *Geometry II*. Universitext. Springer Berlin Heidelberg, 2009.

[3] M. Berger and S. Levy. *Geometry I*. Universitext. Springer Berlin Heidelberg, 2009.

7 — Index