Chapter 1. Euclidean Geometry

The Dot Product

1.1 Definition: For vectors $u, v \in \mathbb{R}^n$ we define the **dot product** of u and v to be

$$u \cdot v = \sum_{i=1}^{n} u_i v_i \,.$$

1.2 Theorem: *(Properties of the Dot Product) For all $u, v, w \in \mathbb{R}^n$ and all $t \in \mathbb{R}$ we have*
(1) (Bilinearity) $(u+v)\cdot w = u\cdot w + v\cdot w \ , \ (tu)\cdot v = t(u\cdot v)$
$u\cdot(v+w) = u\cdot v + u\cdot w \ , \ u\cdot(tv) = t(u\cdot v)$,
(2) (Symmetry) $u\cdot v = v\cdot u$, *and*
(3) (Positive Definiteness) $u\cdot u \geq 0$ *with* $u\cdot u = 0$ *if and only if* $u = 0$.

Proof: The proof is left as an exercise.

1.3 Definition: For a vector $u \in \mathbb{R}^n$, we define the **length** (or **norm**) of u to be

$$|u| = \sqrt{u\cdot u} = \sqrt{\sum_{i=1}^{n} {u_i}^2}.$$

We say that u is a **unit vector** when $|u| = 1$.

1.4 Theorem: *(Properties of Length) Let $u, v \in \mathbb{R}^n$ and let $t \in \mathbb{R}$. Then*
(1) (Positive Definiteness) $|u| \geq 0$ *with* $|u| = 0$ *if and only if* $u = 0$,
(2) (Scaling) $|tu| = |t||u|$,
(3) $|u \pm v|^2 = |u|^2 \pm 2(u\cdot v) + |v|^2$.
(4) (The Polarization Identities) $u\cdot v = \frac{1}{2}\big(|u+v|^2 - |u|^2 - |v|^2\big) = \frac{1}{4}\big(|u+v|^2 - |u-v|^2\big)$,
(5) (The Cauchy-Schwarz Inequality) $|u\cdot v| \leq |u|\,|v|$ *with* $|u\cdot v| = |u|\,|v|$ *if and only if the set* $\{u, v\}$ *is linearly dependent, and*
(6) (The Triangle Inequality) $|u+v| \leq |u| + |v|$.

Proof: We leave the proofs of Parts 1, 2 and 3 as an exercise, and we note that 4 follows immediately from 3. To prove part 5, suppose first that $\{u, v\}$ is linearly dependent. Then one of u and v is a multiple of the other, say $v = tu$ with $t \in \mathbb{R}$. Then

$$|u\cdot v| = |u\cdot(tu)| = |t(u\cdot u)| = |t|\,|u|^2 = |u|\,|tu| = |u|\,|v|.$$

Suppose next that $\{u, v\}$ is linearly independent. Then for all $t \in \mathbb{R}$ we have $u + tv \neq 0$ and so

$$0 \neq |u+tv|^2 = (u+tv)\cdot(u+tv) = |u|^2 + 2t(u\cdot v) + t^2|v|^2.$$

Since the quadratic on the right is non-zero for all $t \in \mathbb{R}$, it follows that the discriminant of the quadratic must be negative, that is

$$4(u\cdot v)^2 - 4|u|^2|v|^2 < 0.$$

Thus $(u\cdot v)^2 < |u|^2|v|^2$ and hence $|u\cdot v| < |u|\,|v|$. This proves part 5.

Using part 5 note that

$$|u+v|^2 = |u|^2 + 2(u\cdot v) + |v|^2 \leq |u|^2 + 2|u\cdot v| + |v|^2 \leq |u|^2 + 2|u|\,|v| + |v|^2 = \big(|u| + |v|\big)^2$$

and so $|u+v| \leq |u| + |v|$, which proves part 6.

1.5 Definition: For points $u, v \in \mathbb{R}^n$, we define the (Euclidean) **distance** between u and v to be

$$d_E(u,v) = |v-u|.$$

1.6 Theorem: *(Metric Properties of Euclidean Distance) Let $u, v, w \in \mathbb{R}^n$. Then*
(1) (Positive Definiteness) $d_E(u,v) \geq 0$ with $d_E(u,v) = 0$ if and only if $u = v$,
(2) (Symmetry) $d_E(u,v) = d_E(v,u)$, and
(3) (The Triangle Inequality) $d_E(u,w) \leq d_E(u,v) + d_E(v,w)$.

Proof: The proof is left as an exercise.

1.7 Definition: For nonzero vectors $0 \neq u, v \in \mathbb{R}^n$, we define the (unoriented) **angle** between u and v to be the angle $\theta(u,v) \in [0,\pi]$ such that

$$\cos\theta(u,v) = \frac{u \cdot v}{|u|\,|v|}.$$

Note that $\theta(u,v) = \frac{\pi}{2}$ if and only if $u \cdot v = 0$. For vectors $u, v \in \mathbb{R}^n$, we say that u and v are **orthogonal** when $u \cdot v = 0$.

1.8 Theorem: *(Properties of Angle) Let $0 \neq u, v \in \mathbb{R}^n$. Then*
(1) $\theta(u,v) \in [0,\pi]$ with $\begin{cases} \theta(u,v) = 0 \text{ if and only if } v = tu \text{ for some } t > 0, \text{ and} \\ \theta(u,v) = \pi \text{ if and only if } v = tu \text{ for some } t < 0, \end{cases}$
(2) (Symmetry) $\theta(u,v) = \theta(v,u)$,
(3) (Scaling) $\theta(tu,v) = \theta(u,tv) = \begin{cases} \theta(u,v) & \text{if } 0 < t \in \mathbb{R}, \\ \pi - \theta(u,v) & \text{if } 0 > t \in \mathbb{R}, \end{cases}$
(4) (The Law of Cosines) $|v-u|^2 = |u|^2 + |v|^2 - 2|u|\,|v|\cos\theta(u,v)$,
(5) (Pythagoras' Theorem) $\theta(u,v) = \frac{\pi}{2}$ if and only if $|v-u|^2 = |u|^2 + |v|^2$, and
(6) (Trigonometric Ratios) if $(v-u) \cdot u = 0$ then $\cos\theta(u,v) = \frac{|u|}{|v|}$ and $\sin\theta(u,v) = \frac{|v-u|}{|v|}$.

Proof: We leave the proofs of Parts 1-5 as an exercise. Note that the Law of Cosines follows from the identity $|v-u|^2 = |v|^2 - 2(v \cdot u) + |u|^2$ and the definition of $\theta(u,v)$. Pythagoras' Theorem is a special case of the Law of Cosines. We Prove Part (6). Let $0 \neq u, v \in \mathbb{R}^n$ and write $\theta = \theta(u,v)$. Suppose that $(v-u) \cdot u = 0$. Then we have $v \cdot u - u \cdot u = 0$ so that $u \cdot v = |u|^2$, and so we have

$$\cos\theta = \frac{u \cdot v}{|u|\,|v|} = \frac{|u|^2}{|u|\,|v|} = \frac{|u|}{|v|}\,.$$

Also, by Pythagoras' Theorem we have $|u|^2 + |v-u|^2 = |v|^2$ so that $|v|^2 - |u|^2 = |v-u|^2$, and so

$$\sin^2\theta = 1 - \cos^2\theta = 1 - \frac{|u|^2}{|v|^2} = \frac{|v|^2 - |u|^2}{|v|^2} = \frac{|v-u|^2}{|v|^2}\,.$$

Since $\theta \in [0,\pi]$ we have $\sin\theta \geq 0$, and so taking the square root on both sides gives

$$\sin\theta = \frac{|v-u|}{|v|}\,.$$

Orthogonal Projections

1.9 Definition: Let $U \subseteq \mathbb{R}^n$ be a subspace. We define the **orthogonal complement** of U in $\mathbb{R}^n$ to be

$$U^\perp = \left\{ x \in \mathbb{R}^n \middle| x \cdot u = 0 \text{ for all } u \in U \right\}.$$

1.10 Theorem: *(Properties of the Orthogonal Complement) Let $U \subseteq \mathbb{R}^n$ be a subspace, let $\mathcal{B} \subseteq U$ and let $A \in M_{k\times n}(\mathbb{R})$. Then*

(1) $U^\perp$ is a vector space,
(2) If $U = \text{Span}(\mathcal{B})$ then $U^\perp = \left\{ x \in \mathbb{R}^n \middle| x \cdot u = 0 \text{ for all } u \in \mathcal{B} \right\}$,
(3) $(\text{Row}A)^\perp = \text{Null}A$.
(4) $\dim(U) + \dim(U^\perp) = n$
(5) $U \oplus U^\perp = \mathbb{R}^n$,
(6) $(U^\perp)^\perp = U$,
(7) $(\text{Null}A)^\perp = \text{Row}A$.

Proof: Note that $0 \in U^\perp$ since $0 \cdot u = 0$ for all $u \in U$. If $x, y \in U^\perp$ so that $x \cdot u = 0$ and $y \cdot u = 0$ for all $u \in U$ then we have $(x + y) \cdot u = x \cdot u + y \cdot u = 0$ for all $u \in U$ and so $x + y \in U^\perp$. If $x \in U^\perp$ so that $x \cdot u = 0$ for all $u \in U$ and $t \in \mathbb{R}$ then we have $(tx) \cdot u = t(x \cdot u) = 0$ for all $u \in U$ and so $tu \in U^\perp$. This shows that $U^\perp$ is a subspace of $\mathbb{R}^n$, proving part 1.

To prove part 2, let $V = \left\{ x \in \mathbb{R}^n \middle| x \cdot u = 0 \text{ for all } u \in \mathcal{B} \right\}$. It is clear that $U^\perp \subseteq V$. Let $x \in V$. Let $u \in U = \text{Span}(\mathcal{B})$, say $u = \sum_{i=1}^{m} t_i u_i$ with each $t_i \in \mathbb{R}$ and each $u_i \in \mathcal{B}$. Then $x \cdot u = x \cdot \sum_{i=1}^{m} t_i u_i = \sum_{i=1}^{m} t_i (x \cdot u_i) = 0$. Thus $x \in U^\perp$ and so we have $V \subseteq U^\perp$.

To prove part 3, let $v_1, v_2, \cdots, v_k$ be the rows of A. Note that $Ax = \begin{pmatrix} x \cdot v_1 \\ \vdots \\ x \cdot v_k \end{pmatrix}$ so we have $x \in \text{Null}A \iff x \cdot v_i = 0 \text{ for all } i \iff x \in \text{Span}\{v_1, v_2, \cdots, v_k\}^\perp = (\text{Row}A)^\perp$ by part 2.

Part 4 follows from part 3 since if we choose A so that $\text{Row}A = U$ then we have $\dim(U) + \dim(U^\perp) = \dim \text{Row}A + \dim(\text{Row}A)^\perp = \dim \text{Row}A + \dim \text{Null}A = n$.

To prove part 5, in light of part 4, it suffices to show that $U \cap U^\perp = \{0\}$. Let $x \in U \cap U^\perp$. Since $x \in U^\perp$ we have $x \cdot u = 0$ for all $u \in U$. In particular, since $x \in U$ we have $x \cdot x = 0$, and hence $x = 0$. Thus $U \cap U^\perp = \{0\}$ and so $U \oplus U^\perp = \mathbb{R}^n$.

To prove part 6, let $x \in U$. By the definition of $U^\perp$ we have $x \cdot v = 0$ for all $v \in U^\perp$. By the definition of $(U^\perp)^\perp$ we see that $x \in (U^\perp)^\perp$. Thus $U \subseteq (U^\perp)^\perp$. By part 4 we know that $\dim U + \dim U^\perp = n$ and also that $\dim U^\perp + \dim(U^\perp)^\perp = n$. It follows that $\dim U = n - \dim U^\perp = \dim(U^\perp)^\perp$. Since $U \subseteq (U^\perp)^\perp$ and $\dim U = \dim(U^\perp)^\perp$ we have $U = (U^\perp)^\perp$, as required.

By parts 3 and 6 we have $(\text{Null}A)^\perp = \left((\text{Row}A)^\perp\right)^\perp = \text{Row}A$, proving Part 7.

1.11 Definition: For a subspace $U \subseteq \mathbb{R}^n$ and a vector $x \in \mathbb{R}^n$, we define the **orthogonal projection** of x onto U, denoted by $\text{Proj}_U(x)$, as follows. Since $\mathbb{R}^n = U \oplus U^\perp$, we can choose unique vectors $u, v \in \mathbb{R}^n$ with $u \in U$, $v \in U^\perp$ and $x = u + v$. We then define

$$\text{Proj}_U(x) = u.$$

Note that since $U = (U^\perp)^\perp$, for u and v as above we have $\text{Proj}_{U^\perp}(x) = v$. When $u \in \mathbb{R}^n$ and $U = \text{Span}\{u\}$, we also write $\text{Proj}_u(x) = \text{Proj}_U(x)$ and $\text{Proj}_{u^\perp}(x) = \text{Proj}_{U^\perp}(x)$.

1.12 Theorem: *Let $U \subseteq \mathbb{R}^n$ be a subspace and let $x \in \mathbb{R}^n$. Then $\text{Proj}_U(x)$ is the unique point in U which is nearest to x.*

Proof: Let $u, v \in \mathbb{R}^n$ with $u \in U$, $v \in U^\perp$ and $u + v = x$ so that $\text{Proj}_U(x) = u$. Let $w \in U$ with $w \neq u$. Since $v \in U^\perp$ and $u, w \in U$ we have $v \cdot u = v \cdot w = 0$ and so $v \cdot (w - u) = v \cdot w - v \cdot u = 0$. Thus we have

$$\begin{aligned}|x-w|^2 &= |u+v-w|^2 = |v-(w-u)|^2 = \big(v-(w-u)\big)\cdot\big(v-(w-u)\big)\\ &= |v|^2 - 2\,v\cdot(w-u) + |w-u|^2 = |v|^2 + |w-u|^2 = |x-u|^2 + |w-u|^2 .\end{aligned}$$

Since $w \neq u$ we have $|w-u| > 0$ and so $|x-w|^2 > |x-u|^2$. Thus $|x-w| > |x-u|$, that is $d_E(x,w) > d_E(x,u)$, so u is the vector in U nearest to x, as required.

1.13 Theorem: *For any matrix $A \in M_{n\times l}(\mathbb{R})$ we have $\text{Null}(A^TA) = \text{Null}(A)$ and $\text{Col}(A^TA) = \text{Col}(A^T)$ so that $\text{nullity}(A^TA) = \text{nullity}(A)$ and $\text{rank}(A^TA) = \text{rank}(A)$.*

Proof: If $x \in \text{Null}(A)$ then $Ax = 0$ so $A^TAx = 0$ hence $x \in \text{Null}(A^TA)$. This shows that $\text{Null}(A) \subseteq \text{Null}(A^TA)$. If $x \in \text{Null}(A^TA)$ then we have $A^TAx = 0$ which implies that $|Ax|^2 = (Ax)^T(Ax) = x^TA^TAx = 0$ and so $Ax = 0$. This shows that $\text{Null}(A^TA) \subseteq \text{Null}(A)$. Thus we have $\text{Null}(A^TA) = \text{Null}(A)$. It then follows that

$$\text{Col}(A^T) = \text{Row}(A) = \text{Null}(A)^\perp = \text{Null}(A^TA)^\perp = \text{Row}(A^TA) = \text{Col}\big((A^TA)^T\big) = \text{Col}(A^TA).$$

1.14 Theorem: *Let $A \in M_{n\times l}(\mathbb{R})$, let $U = \text{Col}(A)$ and let $x \in \mathbb{R}^n$. Then*

(1) the matrix equation $A^TA\,t = A^Tx$ has a solution $t \in \mathbb{R}^l$, and for any solution t we have

$$\text{Proj}_U(x) = At,$$

(2) if $\text{rank}(A) = l$ then A^TA is invertible and

$$\text{Proj}_U(x) = A(A^TA)^{-1}A^Tx.$$

Proof: Note that $U^\perp = (\text{Col}A)^\perp = \text{Row}(A^T)^\perp = \text{Null}(A^T)$. Let $u, v \in \mathbb{R}^n$ with $u \in U$, $v \in U^\perp$ and $u + v = x$ so that $\text{Proj}_U(x) = u$. Since $u \in U = \text{Col}A$ we can choose $t \in \mathbb{R}^l$ so that $u = At$. Then we have $x = u + v = At + v$. Multiply by A^T to get $A^Tx = A^TAt + A^Tv$. Since $v \in U^\perp = \text{Null}(A^T)$ we have $A^Tv = 0$ so $A^TA\,t = A^Tx$. Thus the matrix equation $A^TA\,t = A^Tx$ does have a solution $t \in \mathbb{R}^l$.

Now let $t \in \mathbb{R}^l$ be any solution to $A^TA\,t = A^tx$. Let $u = At$ and $v = x - u$. Note that $x = u + v$, $u = At \in \text{Col}(A) = U$, and $A^Tv = A^T(x-u) = A^T(x - At) = A^Tx - A^TA\,t = 0$ so that $v \in \text{Null}(A^T) = U^\perp$. Thus $\text{Proj}_U(x) = u = At$, proving part (1).

Now suppose that $\text{rank}(A) = l$. Since $A^TA \in M_{l\times l}(\mathbb{R})$ with $\text{rank}(A^TA) = \text{rank}(A) = l$, the matrix A^TA is invertible. Since A^TA is invertible, the unique solution $t \in \mathbb{R}^l$ to the matrix equation $A^TA\,t = A^Tx$ is the vector $t = (A^TA)^{-1}A^Tx$, and so from Part (1) we have $\text{Proj}_U(x) = At = A(A^TA)^{-1}A^Tx$, proving Part (2).

1.15 Definition: Let $\mathcal{B} \subseteq \mathbb{R}^n$. We say $\mathcal{B}$ is **orthogonal** when $x \cdot y = 0$ for all $x, y \in \mathcal{B}$ with $x \neq y$. We say $\mathcal{B}$ is **orthonormal** when $\mathcal{B}$ is orthogonal and $|x| = 1$ for every $x \in \mathcal{B}$.

1.16 Note: When $u_1, \cdots, u_l \in \mathbb{R}^n$, $\mathcal{B} = \{u_1, \cdots, u_l\}$ and $A = (u_1, \cdots, u_l) \in M_{n\times l}(\mathbb{R})$, we have

$$A^TA = \begin{pmatrix} {u_1}^T \\ \vdots \\ {u_l}^T \end{pmatrix} (u_1, \cdots, u_l) = \begin{pmatrix} u_1\cdot u_1 & u_1\cdot u_2 & \cdots & u_1\cdot u_l \\ \vdots & \vdots & & \vdots \\ u_l\cdot u_1 & u_l\cdot u_2 & \cdots & u_l\cdot u_l \end{pmatrix}.$$

It follows that $\mathcal{B}$ is orthogonal if and only if A^TA is diagonal, in which case we have $A^TA = \text{diag}(|u_1|^2, |u_2|^2, \cdots, |u_l|^2)$, and $\mathcal{B}$ is orthonormal if and only if $A^TA = I$.

1.17 Note: Recall that when $\mathcal{B} = \{u_1, u_2, \cdots, u_l\}$ is a basis for a vector space $U \subseteq \mathbb{R}^n$, a vector $x \in U$ can be written uniquely as a linear combination $x = \sum_{i=1}^{l} t_i u_i$ with each $t_i \in \mathbb{R}$, and then we define the coordinate vector of x with respect to $\mathcal{B}$ to be

$$[x]_{\mathcal{B}} = t = \big(t_1, t_2, \cdots, t_l\big)^T \in \mathbb{R}^l.$$

1.18 Theorem: *Let $u_1, u_2, \cdots, u_l \in \mathbb{R}^n$, $\mathcal{B} = \{u_1, \cdots, u_l\}$ and $U = \mathrm{Span}(\mathcal{B})$, and let $x \in \mathbb{R}^n$.*
(1) If $\mathcal{B}$ is orthogonal and each $u_i \neq 0$ then $\mathcal{B}$ is a basis for U and $[x]_{\mathcal{B}} = \big(\frac{x \cdot u_1}{|u_1|^2}, \cdots, \frac{x \cdot u_l}{|u_l|^2}\big)^T$.
(2) If $\mathcal{B}$ is orthonormal then $\mathcal{B}$ is a basis for U and $[x]_{\mathcal{B}} = \big(x \cdot u_1, x \cdot u_2, \cdots, x \cdot u_l\big)^T$.

Proof: Suppose $\mathcal{B}$ is orthogonal with each $u_i \neq 0$. Let $A = (u_1, u_2, \cdots, u_l) \in M_{n \times l}(\mathbb{R})$ so that $U = \mathrm{Col}(A)$. Since $\mathcal{B}$ is orthogonal we have $A^T A = \mathrm{diag}(|u_1|^2, \cdots, |u_l|^2)$. Since each $u_i \neq 0$ we see that $A^T A$ is invertible. Since $\mathrm{rank}(A) = \mathrm{rank}(A^T A) = l$, the columns of A are linearly independent, so $\mathcal{B}$ is a basis for U. Write x as a linear combination $x = \sum_{i=1}^{l} t_i u_i = At$ with $t \in \mathbb{R}^l$. Then we have $A^T x = A^T A\, t$ and so

$$[x]_{\mathcal{B}} = t = (A^T A)^{-1} A^T x = \mathrm{diag}\Big(|u_1|^2, \cdots, |u_l|^2\Big)^{-1} \begin{pmatrix} {u_1}^T \\ \vdots \\ {u_l}^T \end{pmatrix} x$$

$$= \mathrm{diag}\Big(\tfrac{1}{|u_1|^2}, \cdots, \tfrac{1}{|u_l|^2}\Big) \begin{pmatrix} x \cdot u_1 \\ \vdots \\ x \cdot u_l \end{pmatrix} = \begin{pmatrix} \frac{x \cdot u_1}{|u_1|^2} \\ \vdots \\ \frac{x \cdot u_l}{|u_l|^2} \end{pmatrix}$$

This proves Part 1, and Part 2 follows immediately from part 1.

1.19 Theorem: *Let $u_1, u_2, \cdots, u_l \in \mathbb{R}^n$, let $\mathcal{B} = \{u_1, u_2, \cdots, u_l\}$, let $U = \mathrm{Span}\mathcal{B}$, and let $x \in \mathbb{R}^n$.*

(1) If $\mathcal{B}$ is orthogonal with each $u_i \neq 0$ then we have $\mathrm{Proj}_U(x) = \sum_{i=1}^{l} \frac{x \cdot u_i}{|u_i|^2}\, u_i$.

(2) If $\mathcal{B}$ is orthonormal then $\mathrm{Proj}_U(x) = \sum_{i=1}^{l} (x \cdot u_i)\, u_i$.

Proof: Suppose that $\mathcal{B}$ is orthogonal with each $u_i \neq 0$. Let $A = (u_1, u_2, \cdots, u_l) \in M_{n \times l}(\mathbb{R})$ so that $U = \mathrm{Col}(A)$ and we have $A^T A = \mathrm{diag}(|u_1|^2, \cdots, |u_l|^2)$, which is invertible. Then

$$\mathrm{Proj}_U(x) = A\,(A^T A)^{-1} A^T x = \big(u_1, \cdots, u_l\big)\mathrm{diag}\Big(\tfrac{1}{|u_1|^2}, \cdots, \tfrac{1}{|u_l|^2}\Big) \begin{pmatrix} {u_1}^T \\ \vdots \\ {u_l}^T \end{pmatrix} x$$

$$= \Big(\tfrac{u_1}{|u_1|^2}, \cdots, \tfrac{u_l}{|u_l|^2}\Big) \begin{pmatrix} x \cdot u_1 \\ \vdots \\ x \cdot u_l \end{pmatrix} = \tfrac{x \cdot u_1}{|u_1|^2}\, u_1 + \cdots + \tfrac{x \cdot u_l}{|u_l|^2}\, u_l.$$

This proves Part 1, and Part 2 follows immediately from Part 1.

1.20 Remark: Note that as a particular case of Part 1 of the above theorem, when $u \in \mathbb{R}^n$ we have

$$\mathrm{Proj}_u(x) = \tfrac{x \cdot u}{|u|^2}\, u.$$

The Cross Product

1.21 Definition: For vectors $u, v \in \mathbb{R}^3$ we define the **cross product** of u with v to be the vector

$$u \times v = \big(u_2v_3 - u_3v_2\,,\, u_3v_1 - u_1v_3\,,\, u_1v_2 - u_2v_1\big)^T.$$

1.22 Theorem: *(Properties of the Cross Product) For all $u, v, w, x \in \mathbb{R}^3$ and $t \in \mathbb{R}$,*
(1) (Bilinearity) $(u+v)\times w = u\times w + v\times w\,,\ (tu)\times v = t(u\times v)$
$u\times(v+w) = u\times v + u\times w\,,\ u\times(tv) = t(u\times v)$.
(2) (Skew-Symmetry) $u\times v = -v\times u$,
(3) (Cross With Cross) $(u\times v)\times w = (u\cdot w)\,v - (v\cdot w)\,u$,
(4) (Cross With Dot) $(u\times v)\cdot(w\times x) = (u\cdot w)(v\cdot x) - (v\cdot w)(u\cdot x)$,
(5) (Triple Product) $(u\times v)\cdot w = u\cdot(v\times w) = \det(u,v,w)$,
(6) (Angle Sine) When $u\neq 0$ and $v\neq 0$ we have $\sin\theta(u,v) = \dfrac{|u\times v|}{|u||v|}$,
(7) (Degeneracy) $u\times v = 0$ if and only if $\{u,v\}$ is linearly dependent,
(8) (Orthogonality) $(u\times v)\cdot u = 0$ and $(u\times v)\cdot v = 0$,
(9) (Area of Parallelogram) $|u\times v|$ is equal to the area of the parallelogram with vertices 0, u, v and $u+v$, and
(10) (Right-Hand Rule) When $u\times v\neq 0$, the vector $u\times v$ points in the direction of the thumb of the right hand when the fingers point from u towards v.

Proof: Parts 1 to 5 can all be proven, somewhat tediously, by expanding both sides. To prove Part 6, we let $\theta = \theta(u,v)$ and then, using Part 4, we have

$$\begin{aligned}\sin^2\theta = 1-\cos^2\theta = 1 - \frac{(u\cdot v)^2}{|u|^2|v|^2} &= \frac{|u|^2|v|^2-(u\cdot v)^2}{|u|^2|v|^2}\\ = \frac{(u\cdot u)(v\cdot v)-(u\cdot v)(u\cdot v)}{|u|^2|v|^2} &= \frac{(u\times v)\cdot(u\times v)}{|u|^2|v|^2} = \frac{|u\times v|^2}{|u|^2|v|^2}.\end{aligned}$$

Parts 7 and 9 follow easily from Part 6, and Part 8 follows easily from Part 5.

Part 10 is not actually a rigorous mathematical statement because the right hand is not a mathematically defined object, but we can justify the statement informally as follows. Given two linearly independent vectors u and v, we can construct continuous functions $U, V : [0,1]\to\mathbb{R}^3$ with $U(0)=u$, $U(1)=e_1$, $V(0)=v$ and $V(1)=e_2$ so that $\big\{U(t),V(t)\big\}$ is linearly independent for all values of t with $0\le t\le 1$ $\big($for $0\le t\le\frac14$ rotate the vectors until $U\big(\frac14\big)$ points in the direction of the positive x-axis, then for $\frac14\le t\le\frac12$ hold the first vector fixed and rotate the second vector until $V\big(\frac12\big)$ lies in the xy-plane on the same side of the x-axis as the positive y-axis, then for $\frac12\le t\le\frac34$ hold the first vector fixed and alter the angle between the vectors until $V\big(\frac34\big)$ points in the direction of the positive y-axis, then for $\frac34\le t\le 1$ scale the two vectors until $U(1)=e_1$ and $V(1)=e_2\big)$. Let $W(t)=U(t)\times V(t)$ for $0\le t\le 1$. For each value of t, since $W(t)$ is orthogonal to $U(t)$ and $V(t)$, either it points in the direction of the right thumb or it points in the direction of the left thumb when the fingers point from $U(t)$ to $V(t)$. Since $W(t)$ is varies continuously and is never equal to zero, it cannot suddenly jump from one direction to the opposite direction, and so either it points in the direction of the right thumb for all values of t or it points in the direction of the left thumb for all values of t. Since $U(1)=e_1$, $V(1)=e_2$ and $W(1)=e_1\times e_2=e_3$ we see that $W(1)$ points in the direction of the right thumb when the fingers point from $U(1)$ to $V(1)$, and hence $W(t)$ points in the direction of the right thumb for all t. In particular, $u\times v = W(0)$ points in the direction of the right thumb when the fingers point from $u=U(0)$ to $v=V(0)$.

1.23 Example: Let $u,v,x\in\mathbb{R}^3$ and let $U=\text{Span}\{u,v\}$. Then we have

$$\text{Proj}_U(x)=x-\text{Proj}_{U^\perp}(x)=x-\text{Proj}_{u\times v}(x).$$

Also, since $\text{Proj}_U(x)$ is the point in U nearest to x, the **distance** from x to U is equal to

$$d_E(x,U)=d_E\big(x,\text{Proj}_U(x)\big)=\big|x-\text{Proj}_U(x)\big|=\big|\text{Proj}_{u\times v}(x)\big|=\frac{|x\cdot(u\times v)|}{|u\times v|}.$$

1.24 Definition: For vectors $u,v,w\in\mathbb{R}^3$, the (scalar) **triple product** of u, v and w is defined to be

$$\det(u,v,w)=(u\times v)\cdot w=u\cdot(v\times w).$$

1.25 Theorem: *(Properties of the Triple Product) For all $u,v,w\in\mathbb{R}^3$ we have*
(1) (Permutations) $\det(u,v,w)=\det(v,w,u)=\det(w,u,v)$
$=-\det(u,w,v)=-\det(w,v,u)=-\det(v,u,w)$,
(2) (Degeneracy) $\det(u,v,w)=0$ *if and only if* $\{u,v,w\}$ *is linearly dependent, and*
(3) (Volume of Parallelotope) $\big|\det(u,v,w)\big|$ *is equal to the volume of the parallelotope with vertices* 0, u, v, w, $u+v$, $v+w$, $w+u$ *and* $u+v+w$.

Proof: All three parts follow immediately from well-known properties of the determinant. Here is a proof of Part 3. The base of the parallelotope is the parallelogram with vertices at 0, u, v and $u+v$ which has area equal to $A=|u\times v|$. The height of the parallelotope, measured in the direction of $u\times v$ which is orthogonal to the base, is equal to $h=|w||\cos\theta|$ where $\theta=\theta(u\times v,w)$. Thus the volume of the parallelotope is

$$V=Ah=|u\times v|\,|w|\cos\theta=|u\times v||w|\,\frac{|(u\times v)\cdot w|}{|u\times v||w|}=\big|(u\times v)\cdot w\big|=\big|\det(u,v,w)\big|.$$

1.26 Definition: Let $\{u,v,w\}$ be a basis for $\mathbb{R}^3$ and note that $\det(u,v,w)\neq 0$. When $\det(u,v,w)>0$ we say that $\{u,v,w\}$ is a **positively oriented** basis for $\mathbb{R}^3$ and when $\det(u,v,w)<0$ we say that $\{u,v,w\}$ is a **negatively oriented** basis for $\mathbb{R}^3$. One can argue informally, as in our proof of Part 9 of the above theorem, that $\{u,v,w\}$ is positively oriented when the vector w lies on the same side of the plane spanned by u and v as the thumb of the right hand when the fingers point from u to v.

1.27 Notation: Given a vector $x\in\mathbb{R}^n$, from now on we shall often write x as a row vector

$$x=(x_1,x_2,\cdots,x_n)$$

when it should be understood that, strictly speaking, x is the column vector

$$x=(x_1,x_2,\cdots,x_n)^T=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}.$$

It is sometimes important to keep this in mind when formulas involve linear algebra operations (for example, when a formula involves an expression of the form Ax where A is a matrix). It is common to abuse notation in this way simply because row vectors are easier to typeset and easier to read than column vectors.

1.28 Note: For $u=(u_1,u_2)\in\mathbb{R}^2$ and $v=(v_1,v_2)\in\mathbb{R}^2$ we have

$$\begin{aligned} u\cdot v &= u_1v_1+u_2v_2 \\ |u| &= \sqrt{{u_1}^2+{u_2}^2} \\ d_E(u,v) &= \sqrt{(v_1-u_1)^2+(v_2-u_2)^2} \\ \theta(u,v) &= \cos^{-1}\frac{u_1v_1+u_2v_2}{|u||v|} \end{aligned}$$

1.29 Definition: A **real number modulo** 2π is a set of the form $\{\theta+2\pi k \big| k\in\mathbb{Z}\}$ for some $\theta\in\mathbb{R}$. The set of all real numbers modulo 2π is denoted by $\mathbb{R}/2\pi$ (or by $\mathbb{R}/2\pi\mathbb{Z}$). When $\theta\in\mathbb{R}$ we often denote the set $\{\theta+2\pi k\big|k\in\mathbb{Z}\}$ simply as $\theta\in\mathbb{R}/2\pi$. Note that when $\theta\in\mathbb{R}/2\pi$ the trigonometric values $\sin\theta$ and $\cos\theta$ are well-defined, and the element $\theta\in\mathbb{R}/2\pi$ is uniquely determined from the values $\sin\theta$ and $\cos\theta$. For a nonzero vector $0\neq u\in\mathbb{R}^2$, we define the **oriented angle** of u to be the (unique) element $\theta_o=\theta_o(u)\in\mathbb{R}/2\pi$ for which

$$u=\big(|u|\cos\theta_o\,,\,|u|\sin\theta_o\big).$$

For nonzero vectors $0\neq u,v\in\mathbb{R}^2$, the **oriented angle** in $\mathbb{R}^2$ from u to v is

$$\theta_o(u,v)=\theta_o(v)-\theta_o(u).$$

1.30 Theorem: *Let $0\neq u,v\in\mathbb{R}^2$. For $\theta_o=\theta_o(u,v)\in\mathbb{R}/2\pi$ we have*

$$\cos\theta_o=\frac{u\cdot v}{|u||v|}\quad\textit{and}\quad\sin\theta_o=\frac{u_1v_2-u_2v_1}{|u||v|}=\frac{\det(u,v)}{|u||v|}$$

and for $\theta=\theta(u,v)\in[0,\pi]$ we have

$$\cos\theta=\cos\theta_o\quad\textit{and}\quad\sin\theta=\big|\sin\theta_o\big|.$$

Proof: The proof is left as an exercise.

1.31 Definition: Let $u\in\mathbb{R}^2$ and let $0<r\in\mathbb{R}$. The **circle** in $\mathbb{R}^2$ and the (closed) **disc** in $\mathbb{R}^2$ centred at u of radius r are the sets

$$\begin{aligned} C(u,r)&=\{x\in\mathbb{R}^2\big|d_E(x,u)=r\}\text{ and} \\ D(u,r)&=\{x\in\mathbb{R}^2\big|d_E(x,u)\leq r\}. \end{aligned}$$

1.32 Theorem: *(The Circumference of a Circle and the Area of a Disc) Let $u\in\mathbb{R}^2$ and let $0<r\in\mathbb{R}$. An arc along the circle $C(u,r)$ which subtends an angle θ at u has length $L=r\theta$ so, in particular, the circumference of $C(u,r)$ is equal to $L=2\pi r$. A sector of the disc $D(u,r)$ which subtends an angle θ at u has area $A=\frac{1}{2}r^2\theta$ so, in particular, the area of $D(u,r)$ is equal to $A=\pi r^2$.*

Proof: An arc along the circle $C(u,r)$ which subtends an angle θ at u is given parametrically by $(x,y)=u+\big(r\cos t,r\sin t\big)$ with say $\alpha\leq t\leq\alpha+\theta$, and its length is

$$L=\int_{t=\alpha}^{\alpha+\theta}\sqrt{x'(t)^2+y'(t)^2}\,dt=\int_{t=\alpha}^{\alpha+\theta}\sqrt{r^2\cos^2t+r^2\sin^2t}\,dt=\int_{t=\alpha}^{\alpha+\theta}r\,dt=r\theta,$$

Using polar coordinates, a sector of the disc $D(u,r)$ which subtends an angle θ has area

$$A=\int_{t=\alpha}^{\alpha+\theta}\int_{s=0}^{r}s\,ds\,dt=\int_{t=\alpha}^{\alpha+\theta}\tfrac{1}{2}\,r^2\,dt=\tfrac{1}{2}\,r^2\,\theta.$$

1.33 Definition: A **line** in $\mathbb{R}^2$ is a set of the form $L = \{(x,y) \in \mathbb{R}^2 \,|\, ax + by = c\}$ for some $a, b, c \in \mathbb{R}$ with $(a,b) \neq (0,0) \in \mathbb{R}^2$. Equivalently, a line is a set of the form $L = \{x \in \mathbb{R}^2 \,|\, x \cdot n = c\}$ for some number $c \in \mathbb{R}$ and some nonzero vector $0 \neq n \in \mathbb{R}^2$.

1.34 Theorem: *(Parametric Representation of a Line) A line in $\mathbb{R}^2$ is a set of the form $L = \{x \in \mathbb{R}^2 \,|\, x = p + tu \text{ for some } t \in \mathbb{R}\}$ for some point $p \in \mathbb{R}^2$ and some nonzero vector $0 \neq u \in \mathbb{R}^2$.*

Proof: The proof is left as an exercise.

1.35 Definition: When $c \in \mathbb{R}$ and $0 \neq n \in \mathbb{R}^2$ and $L = \{x \in \mathbb{R}^2 \,|\, x \cdot n = c\}$, we say that L is the line $x \cdot n = c$ and we say that n is a **normal vector** for L. When $p \in \mathbb{R}^2$ and $0 \neq u \in \mathbb{R}^2$ and $L = \{x \in \mathbb{R}^2 \,|\, x = p + tu \text{ for some } t \in \mathbb{R}\}$, we say that L is the line $x = p + tu$ and we say that u is a **direction vector** for L.

1.36 Theorem: *Let $p, q \in \mathbb{R}^2$, let $0 \neq u, v \in \mathbb{R}^2$, let L be the line $x = p + tu$ and let M be the line $x = q + tv$. Then $L = M$ if and only if $q \in L$ and $v = su$ for some $0 \neq s \in \mathbb{R}$.*

Proof: The proof is left as an exercise.

1.37 Definition: Let L and M be lines in $\mathbb{R}^2$ with direction vectors u and v. We say that L and M are **parallel** when $L \neq M$ and u and v are parallel, that is $v = su$ for some $0 \neq s \in \mathbb{R}$, and we say that L and M are **perpendicular** (or **orthogonal**) when u and v are perpendicular (or orthogonal), that is $u \cdot v = 0$.

1.38 Theorem: *(Properties of Lines in the Euclidean Plane)*
(1) Given lines L, M in $\mathbb{R}^2$, either $L = M$ or $L \cap M = \emptyset$ or $L \cap M = \{p\}$ for some $p \in \mathbb{R}^2$.
(2) Given points $p, q \in \mathbb{R}^2$ with $p \neq q$, there is a unique line in $\mathbb{R}^2$ through p and q.
(3) Given a point $p \in \mathbb{R}^2$ and a line L in $\mathbb{R}^2$ with $p \notin L$, there is a unique line in $\mathbb{R}^2$ which passes through p and is parallel to L.
(4) Given a point $p \in \mathbb{R}^2$ and a line L in $\mathbb{R}^2$, there is a unique line which passes through p and is perpendicular to L.

Proof: The proof is left as an exercise.

1.39 Definition: Let $u, v \in \mathbb{R}^2$ with $u \neq v$ and note that the line L through u and v is the set $L = \{u + t(v-u) \,|\, t \in \mathbb{R}\}$. The **ray** from u through v (or the ray from u in the direction of the vector $v - u$) is the set $R = \{u + t(v-u) \,|\, 0 \leq t \in \mathbb{R}\}$. The **line segment** between u and v is the set

$$[u,v] = \{u + t(v-u) \,|\, 0 \leq t \leq 1\} = \{su + tv \,|\, 0 \leq s, 0 \leq t, s + t = 1\}.$$

For $x \in \mathbb{R}^2$, we say that x is **between** u and v when $x \in [u,v]$.

1.40 Definition: Let $u, v, w \in \mathbb{R}^2$. Let R be the ray from u through v and let S be the ray from u through w. The **oriented angle** $\angle_o wuv$, also called the oriented angle from $[u,v]$ to $[u,w]$, or the oriented angle from R to S, is defined to be

$$\angle_o wuv = \angle_o\big([u,v],[u,w]\big) = \theta_o(R,S) = \theta_o\big(v-u, w-u\big).$$

The (unoriented) **angle** $\angle wuv$, also called the (unoriented) angle between $[u,v]$ and $[u,w]$, or the (unoriented) angle between R and S is defined to be

$$\angle wuv = \theta\big([u,v],[u,w]\big) = \angle(R,S) = \theta\big(v-u, w-u\big).$$

1.41 Definition: Let $p \in \mathbb{R}^2$ and let $0 \neq u, v \in \mathbb{R}^2$. Let L be the line $x = p + tu$ and let M be the line $x = p + tv$. The **oriented angle** from L to M is defined to be

$$\theta_o(L, M) = \min\Big(\theta_o(u, v), \theta_o(u, -v)\Big) \in [0, \pi)$$

where $\theta_o(u, v)$ and $\theta_o(u, -v)$ are being considered as real numbers in $[0, 2\pi)$, and the (unoriented) **angle** between L and M is defined to be

$$\theta(L, M) = \min\Big(\theta(u, v), \theta(u, -v)\Big) \in \big[0, \tfrac{\pi}{2}\big].$$

1.42 Remark: The oriented angle $\theta_o(L, M)$ can be considered as an element of $\mathbb{R}/\pi$.

1.43 Theorem: *Let $p \in \mathbb{R}^2$ and let $0 \neq u, v \in \mathbb{R}^2$. Let L be the line $x = p + tu$ and let M be the line $x = p + tv$. Then*

$$\theta(L, M) = \cos^{-1} \frac{|u \cdot v|}{|u||v|}.$$

Proof: The proof is left as an exercise.

1.44 Theorem: *(Addition of Angles, Supplementary Angles, and Parallel Lines)*
(1) If R, S and T are rays from $p \in \mathbb{R}^2$ then $\theta_o(R, S) + \theta_o(S, T) = \theta_o(R, T)$.
(2) If $p \in \mathbb{R}^2$ and L and M are lines with $L \cap M = \{p\}$ then $\theta_o(L, M) + \theta_o(M, L) = \pi$.
(3) If $p, q \in \mathbb{R}^2$ with $p \neq q$ and L, M and N are lines with $L \cap N = \{p\}$ and $M \cap N = \{q\}$ then $L \parallel M \iff \theta_o(L, N) = \theta_o(M, N) \iff \theta_o(L, N) + \theta_o(N, M) = \pi$.

Proof: The proof is left as an exercise.

1.45 Definition: A **triangle** in $\mathbb{R}^2$ is determined by 3 non-colinear points $u, v, w \in \mathbb{R}^2$, which we call the **vertices** of the triangle. We can think of the triangle T determined by u, v and w in several different ways. For example, we could consider T to be the set $T = \{u, v, w\}$, or we can keep track of the order of the vertices and consider T to be the ordered triple $T = (u, v, w)$. Alternatively, we could consider T to be the union of its three **edges**, that is $T = [v, w] \cup [w, u] \cup [u, v]$, or we can think of the triangle as including its interior points so that T is the **closed solid triangle**

$$\begin{aligned}[u, v, w] &= \big\{u + s(v - u) + t(w - u)\big|0 \leq s, 0 \leq t, s + t \leq 1\big\} \\ &= \big\{ru + sv + tw\big|0 \leq r, 0 \leq s, 0 \leq t, r + s + t = 1\big\}.\end{aligned}$$

An **ordered triangle** in $\mathbb{R}^2$ consists of an ordered triple (u, v, w) of non-colinear points $u, v, w \in \mathbb{R}^2$, together with the closed solid triangle $[u, v, w]$.

Given an ordered triangle $[u, v, w]$ in $\mathbb{R}^2$, the **edge lengths** of the triangle will be denoted by $a, b, c \in \mathbb{R}$ with

$$a = |w - v| \ , \ b = |u - w| \ , \ c = |v - u| \,,$$

the **oriented angles** of the triangle will be denoted by $\alpha_o, \beta_o, \gamma_o \in \mathbb{R}/2\pi$ with

$$\alpha_o = \theta_o(v - u, w - u) \ , \ \beta_o = \theta_o(w - v, u - v) \ , \ \gamma_o = \theta_o(u - w, v - w) \,,$$

and the (unoriented) **angles** (or the **interior angles**) of the triangle will be denoted by $\alpha, \beta, \gamma \in (0, \pi)$ with

$$\alpha = \theta(v - u, w - u) \ , \ \beta = \theta(w - v, u - v) \ , \ \gamma = \theta(u - w, v - w) \,.$$

1.46 Theorem: *(The Sum of the Angles in a Triangle) The sum of the interior angles in any triangle in $\mathbb{R}^2$ is equal to π.*

Proof: The proof is left as an exercise.

1.47 Theorem: *Let $[u,v,w]$ be an ordered triangle in $\mathbb{R}^2$. Then*

$$\begin{aligned}\sin\alpha_o &= \frac{\det(u,v)+\det(v,w)+\det(w,u)}{|v-u||u-w|}\\ \sin\beta_o &= \frac{\det(u,v)+\det(v,w)+\det(w,u)}{|w-v||v-u|}\\ \sin\gamma_o &= \frac{\det(u,v)+\det(v,w)+\det(w,u)}{|u-w||w-v|}\end{aligned}$$

Proof: The proof is left as an exercise.

1.48 Corollary: *Let $[u,v,w]$ be an ordered triangle in $\mathbb{R}^2$. Then $\sin\alpha_o$, $\sin\beta_o$ and $\sin\gamma_o$ all have the same sign.*

1.49 Definition: Let $[u,v,w]$ be an ordered triangle in $\mathbb{R}^2$. When $\sin\alpha_o$, $\sin\beta_o$ and $\sin\gamma_o$ are all positive, we say that the triangle $[u,v,w]$ is **positively oriented**, and when $\sin\alpha_o$, $\sin\beta_o$ and $\sin\gamma_o$ are all negative, we say that the triangle $[u,v,w]$ is **negatively oriented**.

1.50 Corollary: *(The Sine Law) Let $[u,v,w]$ be an ordered triangle in $\mathbb{R}^2$. Then*

$$\frac{\sin\alpha}{a}=\frac{\sin\beta}{b}=\frac{\sin\gamma}{c}.$$

1.51 Corollary: *(The Area of a Triangle) Then the area of the triangle $[u,v,w]$ in $\mathbb{R}^2$ is*

$$A=\tfrac{1}{2}\Big|\det(u,v)+\det(v,w)+\det(w,u)\Big|.$$

1.52 Corollary: *(Similar Triangles) Let $[u,v,w]$ and $[u',v',w']$ be two ordered triangles in $\mathbb{R}^2$. Suppose that the corresponding angles of the two triangles are equal, that is $\alpha'=\alpha$, $\beta'=\beta$ and $\gamma'=\gamma$. Then there exists $s>0$ such that $a'=sa$, $b'=sb$ and $c'=sc$.*

1.53 Theorem: *(The Cosine Law) Let $[u,v,w]$ be an ordered triangle in $\mathbb{R}^2$. Then*

$$\cos\alpha=\frac{b^2+c^2-a^2}{2bc}\ ,\ \cos\beta=\frac{c^2+a^2-b^2}{2ca}\ \text{ and }\ \cos\gamma=\frac{a^2+b^2-c^2}{2ab}.$$

Proof: The proof is left as an exercise.

1.54 Corollary: *Let $[u,v,w]$ be an ordered triangle in $\mathbb{R}^2$.*
(1) (Side-Side-Side) Given the lengths of the three sides we can determine the angles.
(2) (Side-Angle-Side) Given the lengths of two sides and the angle at the common vertex, we can determine the length of the other side and the other two angles.
(3) (Angle-Side-Angle) Given the length of one edge of the triangle and the angles at both ends of the edge, we can determine the third angle and the lengths of the other two sides.
(4) (The Isoceles Triangle Theorem) We have $\beta=\gamma \iff b=c$.

1.55 Corollary: *(The Angle Subtended by a Chord in a Circle) Let $p\in\mathbb{R}^2$, let C be a circle in $\mathbb{R}^2$ centred at p, and let $[u,v,w]$ be a triangle in $\mathbb{R}^2$ with $u,v,w\in C$. Then*

$$\angle_o vpu = 2\,\angle_o vwu\,.$$

Triangle Centres

1.56 Definition: Let $u, v \in \mathbb{R}^2$ with $u \neq v$. The **midpoint** of the line segment $[u, v]$ is the point $m \in [u, v]$ such that $d_E(m, u) = d_E(m, v)$, that is the point

$$m = u + \tfrac{1}{2}(v - u) = \tfrac{1}{2}(u + v).$$

1.57 Definition: In an triangle in $\mathbb{R}^2$, a **median** is a line from a vertex to the midpoint of the opposite side. In the triangle $[u, v, w]$, the median from u is the line through u and $\frac{1}{2}(v + w)$, the median from v is the line through v and $\frac{1}{2}(w + u)$, and the median from w is the line through w and $\frac{1}{2}(u + v)$.

1.58 Theorem: *(The Centroid) The three medians in a triangle meet at a point g, which is called the* **centroid** *of the triangle. The centroid lies two thirds of the way along each of the medians, from a vertex to the midpoint of the opposite side.*

Proof: Let $[u, v, w]$ be an ordered triangle in $\mathbb{R}^2$. The point which lies $\frac{2}{3}$ of the way along the median from u to $\frac{v+w}{2}$ is the point $u + \frac{2}{3}\left(\frac{v+w}{2} - u\right) = \frac{1}{3}u + \frac{1}{3}(v + w) = \frac{1}{3}(u + v + w)$. Similarly, the point which lies $\frac{2}{3}$ of the way along the median from v to $\frac{w+u}{2}$ is the point $v = \frac{2}{3}\left(\frac{w+u}{2} - v\right) = \frac{1}{3}(u + v + w)$ and the point which lies $\frac{2}{3}$ of the way from w to $\frac{u+v}{2}$ is the point $w + \frac{2}{3}\left(\frac{u+v}{2} - w\right) = \frac{1}{3}(u + v + w)$. Thus the point $g = \frac{1}{3}(u + v + w)$ lies $\frac{2}{3}$ of the way along all 3 medians.

1.59 Definition: Let $u, v \in \mathbb{R}^2$ with $u \neq v$. The **perpendicular bisector** of the line segment $[u, v]$ is the line through the midpoint $\frac{u+v}{2}$ with normal vector $v - u$, that is the line

$$\left(x - \tfrac{u+v}{2}\right) \cdot (v - u) = 0.$$

1.60 Theorem: *Let $u, v \in \mathbb{R}^2$ with $u \neq v$ and let L be the perpendicular bisector of the line segment $[u, v]$. Then for $x \in \mathbb{R}^2$ we have $x \in L \iff d_E(x, u) = d_E(x, v)$.*

Proof: We have

$$\begin{aligned} x \in L &\iff \left(x - \tfrac{u+v}{2}\right) \cdot (v - u) = 0 \iff (2x - (v + u)) \cdot (v - u) = 0 \\ &\iff 2x \cdot (v - u) = (v + u) \cdot (v - u) \iff 2x \cdot (v - u) = |v|^2 - |u|^2 \end{aligned}$$

and

$$\begin{aligned} d_U(x, u) = d_E(x, v) &\iff |x - u| = |x - v| \iff |x - u|^2 = |x - v|^2 \\ &\iff (x - u) \cdot (x - u) = (x - v) \cdot (x - v) \\ &\iff |x|^2 - 2x \cdot u + |u|^2 = |x|^2 - 2x \cdot v + |v|^2 \\ &\iff 2x \cdot (v - u) = |v|^2 - |u|^2 \iff x \in L. \end{aligned}$$

1.61 Theorem: *(The Circumcentre) The three perpendicular bisectors of the edges of a triangle intersect at a point o which is called the* **circumcentre** *of the triangle. The circumcentre is equidistant from the three vertices of the triangle so it is the centre of the circle which passes through the three vertices, which we call the* **circumcircle** *(or the* **circumscribed circle***) of the triangle.*

Proof: Let $[u, v, w]$ be an ordered triangle in $\mathbb{R}^2$. Let L, M and N be the perpendicular bisectors of the edges $[v, w]$, $[w, u]$ and $[u, v]$ respectively. Let o be the point of intersection of L and M. By the previous theorem, since $o \in L$ we have $|o - v| = |o - w|$ and since $o \in M$ we have $|o - w| = |o - u|$. It follows that $|o - u| = |o - v| = |o - w|$. By the previous theorem again, since $|o - u| = |o - v|$ we also have $o \in N$, so the point o lies on all three perpendicular bisectors.

1.62 Definition: In a triangle in $\mathbb{R}^2$, an **altitude** is a line through a vertex which is perpendicular to the opposite side. In the triangle $[u,v,w]$, the altitude from u is the line through u with normal vector $w-v$, the altitude from v is the line through v with normal vector $u-w$, and the altitude from w is the line through w with normal vector $v-u$.

1.63 Theorem: *(The Orthocentre) The three altitudes of a triangle meet at a point h which is called the* **orthocentre** *of the triangle. The points o, g and h lie on a line, called the* **Euler line** *of the triangle, with g lying $\frac{1}{3}$ of the way from o to h.*

Proof: Let $[u,v,w]$ be an ordered triangle in $\mathbb{R}^2$. Let g be the centroid and let o be the circumcentre. Let h be the point on the line through o and g such that g lies $\frac{1}{3}$ of the way from o to h, in other words, let h be the point such that $g = o + \frac{1}{3}(h-o)$. Then we have $h = 3g - 2o$. We need to show that h lies on all three altitudes of $[u,v,w]$. We shall show that h lies on the altitude from w (the proof that h lies on the other two altitudes is similar). The altitude from w is given by the equation $(x-w)\cdot(v-u) = 0$, so we need to show that $(h-w)\cdot(v-u) = 0$. Since $g = \frac{1}{3}(u+v+w)$ we have $3g - w = u + v$ and so

$$\begin{aligned}(h-w)\cdot(v-u) &= \big((3g-2o)-w\big)\cdot(v-u) = (3g-w)\cdot(v-u) - 2o\cdot(v-u)\\ &= (u+v)\cdot(v-u) - 2o\cdot(v-u).\end{aligned}$$

Since o lies on the perpendicular bisector of $[u,v]$ we have $\big(o - \frac{u+v}{2}\big)\cdot(v-u) = 0$ and so $2o\cdot(v-u) = (u+v)\cdot(v-u)$ and hence

$$(h-w)\cdot(v-u) = (u+v)\cdot(v-u) - 2o\cdot(v-u) = 0,$$

as required.

1.64 Remark: Given u, v and w, we can find explicit formulas for the circumcentre o and the orthocentre h of the triangle $[u,v,w]$. Let P_v and P_w be the perpendicular bisectors of the edges $[u,w]$ and $[u,v]$. For $x \in \mathbb{R}^2$ we have

$$x \in P_w \iff \big(x - \tfrac{u+v}{2}\big)\cdot(v-u) = 0 \iff x\cdot(v-u) = \tfrac{1}{2}(u+v)\cdot(v-u).$$

and similarly $x \in P_v \iff x\cdot(w-u) = \frac{1}{2}(u+w)\cdot(w-u)$. It follows that o is the point which satisfies both of these two equations. By writing the pair of equations in matrix form, we obtain $Ao = b$ where A is the 2×2 matrix $A = \big(v-u, w-u\big)^T$ and $b = \big(\frac{1}{2}(u+v)\cdot(v-u), \frac{1}{2}(u+w)\cdot(w-u)\big)^T$. Since $[u,v,w]$ is a triangle, the points u, v and w are non-colinear, and so the vectors $v-u$ and $w-u$ are linearly independent hence the matrix A is invertible. Thus o is given by the formula

$$o = A^{-1}b = \tfrac{1}{2}\begin{pmatrix}(v-u)^T\\(w-u)^T\end{pmatrix}^{-1}\begin{pmatrix}(v+u)\cdot(v-u)\\(w+u)\cdot(w-u)\end{pmatrix}.$$

Let H_v and H_w be the altitudes of the triangle $[u,v,w]$ from v and w. For $x \in \mathbb{R}^2$ we have $x \in H_w \iff (x-w)\cdot(v-u) = 0 \iff x\cdot(v-u) = w\cdot(v-u)$ and similarly $x \in M_v \iff x\cdot(w-u) = v\cdot(w-u)$. It follows that h is the point which satisfies both of these equations, so we have $Ah = c$ where A is as above and $c = \big(w\cdot(v-u), v\cdot(w-u)\big)^T$. Thus the point h is given by the formula

$$h = A^{-1}c = \begin{pmatrix}(v-u)^T\\(w-u)^T\end{pmatrix}^{-1}\begin{pmatrix}w\cdot(v-u)\\v\cdot(w-u)\end{pmatrix}.$$

As an exercise, use these explicit formulas to show that g lies $\frac{1}{3}$ of the way from o to h.

1.65 Definition: Let $p \in \mathbb{R}^2$ and let L and M be lines in $\mathbb{R}^2$ with $L \cap M = \{p\}$. The **angle bisectors** of L and M are the lines A and B through p such that $\theta_o(L,A) = \frac{1}{2}\theta_o(L,M)$ and $\theta_o(M,B) = \frac{1}{2}\theta_o(M,L)$. If the lines L and M have direction vectors u and v with $|u| = |v|$ then the two angle bisectors have direction vectors $u+v$ and $u-v$. Likewise, if the lines L and M have normal vectors ℓ and m with $|\ell\,| = |m|$ then the two angle bisectors have normal vectors $\ell + m$ and $\ell - m$. Note that the two angle bisectors are orthogonal to each other, indeed when $|u| = |v|$ we have $(u+v)\cdot(u-v) = |u|^2 - |v|^2 = 0$.

1.66 Definition: Let $x \in \mathbb{R}^2$ and let L be a line in $\mathbb{R}^2$. The (Euclidean) **distance** between x and L, denoted by $d_E(x,L)$, is the distance between x and the point $a \in L$ which is nearest to x. This point a is the point of intersection of L with the line through x which is perpendicular to L. When L is the line $(x-p)\cdot n = 0$, the point $a \in L$ nearest to x is given by the formula $a = x + \text{Proj}_n(x-p) = x + \frac{(x-p)\cdot n}{|n|^2}\, n$ and so the distance between x and L is given by

$$d_E(x,L) = \Big|\text{Proj}_n(x-p)\Big| = \frac{\big|(x-p)\cdot n\big|}{|n|}.$$

1.67 Exercise: Show that when L is the line $ax + by + c = 0$ we have

$$d_E\big((x,y),L\big) = \frac{|ax+by+c|}{\sqrt{a^2+b^2}}.$$

1.68 Theorem: *Let L and M be lines in $\mathbb{R}^2$ with $L \cap M = \{p\}$. Let A and B be the two angle bisectors of L and M at p. Then for $x \in \mathbb{R}^2$ we have*

$$x \in A \cup B \iff d_E(x,L) = d_E(x,M).$$

Proof: Let L and M have normal vectors ℓ and m with $|\ell\,| = |m|$. Then two angle bisectors A and B have normal vectors $\ell \pm m$ and are given by the equations $(x-p)\cdot(\ell \pm m) = 0$. Since $|\ell\,| = |m|$ we have

$$\begin{aligned}
d_E(x,L) = d_E(x,M) &\iff \frac{\big|(x-p)\cdot\ell\big|}{|\ell|} = \frac{\big|(x-p)\cdot m\big|}{|m|} \\
&\iff |(x-p)\cdot\ell| = |(x-p)\cdot m| \\
&\iff (x-p)\cdot\ell = \pm(x-p)\cdot m \\
&\iff (x-p)\cdot(\ell - m) = 0 \;\text{ or }\; (x-p)\cdot(\ell+m) = 0 \\
&\iff x \in A \text{ or } x \in B.
\end{aligned}$$

1.69 Definition: For a vector $(a,b) \in \mathbb{R}^2$, we write $(a,b)^\times = (-b,a)$. In a triangle $[u,v,w]$, the edges have direction vectors $w-v$, $u-w$ and $v-u$ and they have unit normal vectors $\ell = \frac{(w-v)^\times}{|w-v|}$, $m = \frac{(u-w)^\times}{|u-w|}$ and $n = \frac{(v-u)^\times}{|v-u|}$ When $[u,v,w]$ is positively oriented the vectors ℓ, m and n are called the **inward normal vectors** and the vectors $-\ell$, $-m$ and $-n$ are called the **outward normal vectors**, and when $[u,v,w]$ is negatively oriented the situation is reversed. The three **external angle bisectors** of the triangle $[u,v,w]$ are the line through u with normal vector $m+n$, the line through v with normal vector $n+\ell$, and the line through w with normal vector $\ell+m$, and the three **internal angle bisectors** are the line through u with normal vector $m-n$, the line through v with normal vector $n-\ell$, and the line through w with normal vector $\ell - m$.

1.70 Theorem: *(The Incentre) The three internal angle bisectors of a triangle meet at a point which is called the* **incentre** *of the triangle. The incentre of the triangle is equidistant from the three edges of the triangle, so it is the centre of the circle which lies inside the triangle and is tangent to all three edges of the triangle. This circle is called the* **incircle** *(or the* **inscribed circle***) of the triangle.*

Proof: Let $[u, v, w]$ be an ordered triangle in $\mathbb{R}^2$. Let L, M and N be the lines which contain the edges $[v, w]$, $[w, u]$ and $[u, v]$, respectively. A, B and C be the internal angle bisectors from u, v and w, respectively. Let i be the point of intersection of A and B and note that i lies inside the triangle. Since $i \in A$ we have $d_E(i, M) = d_E(i, N)$ and since $i \in B$ we have $d_E(i, N) = d_E(i, L)$. It follows that $d_E(i, L) = d_E(i, M) = d_E(i, N)$. Since $d_E(i, L) = d_E(i, M)$ it follows that i lies on one of the two angle bisectors of the lines L and M, and since i lies inside the triangle it must lie on the internal angle bisector, that is $i \in C$.

Some Geometry in Euclidean Space

1.71 Remark: We do not intend to provide a detailed study of geometry in Euclidean space, but let us describe briefly how some of the aspects of geometry in the plane can be extended to $\mathbb{R}^3$ or to $\mathbb{R}^n$.

For $0 \neq u, v \in \mathbb{R}^2$, we obtained formulas for both the (unoriented) angle $\theta(u, v)$ between u and v, and also for the oriented angle $\theta_o(u, v)$ from u counterclockwise to v. For vectors $0 \neq u, v \in \mathbb{R}^n$ we consider only the (unoriented) angle $\theta(u, v)$ as there is no natural way to decide on an orientation.

For $u \in \mathbb{R}^2$ and $r > 0$, we have the circle $C(u, r)$ and the disc $D(u, r)$. These extend naturally to $\mathbb{R}^3$, or more generally to $\mathbb{R}^n$. For $u \in \mathbb{R}^n$ and $r > 0$, we have the **sphere** $S(u, r) = \left\{x \in \mathbb{R}^n \,\middle|\, d_E(x, u) = r\right\}$ and the (closed) **ball** $B(u, r) = \left\{x \in \mathbb{R}^n \,\middle|\, d_E(x, u) \leq r\right\}$. Just as we found the circumference of the circle and the area of the disc (in Theorem 1.32), one can calculate the area A of the sphere and the volume V of the ball in $\mathbb{R}^3$ using formulas from calculus for areas and volumes of surfaces and solids of revolution: the sphere $x^2 + y^2 + z^2 = r^2$ is obtained by revolving the curve $y = \sqrt{r^2 - x^2}$, $-r \leq x \leq r$ about the y-axis, so its area is

$$\begin{aligned} A &= \int_{x=-r}^{r} 2\pi y(x)\sqrt{1 + y'(x)^2}\; dx = \int_{x=-r}^{r} 2\pi\sqrt{r^2 - x^2}\sqrt{1 + \left(\tfrac{-x}{\sqrt{r^2-x^2}}\right)^2}\, dx \\ &= \int_{x=-r}^{r} 2\pi\sqrt{r^2 - x^2}\,\tfrac{r}{\sqrt{r^2-x^2}}\, dx = \int_{x=-r}^{r} 2\pi r\, dx = 4\pi r^2 \end{aligned}$$

and the ball $x^2 + y^2 + z^2 \leq r^2$ is obtained by revolving the region given by $0 \leq x \leq r$, $-\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2}$ about the y-axis, so (using the shell method) its volume is

$$\begin{aligned} V &= \int_{x=0}^{r} 2\pi x(y_1(x) - y_2(x))\, dx = \int_{x=-r}^{r} 2\pi x \cdot 2\sqrt{r^2 - x^2}\, dx \\ &= \Big[-\tfrac{4}{3}\pi(r^2 - x^2)^{3/2}\Big]_{x=0}^{r} = \tfrac{4}{3}\pi r^3. \end{aligned}$$

These formulas can be generalized to give higher dimensional volumes for spheres and balls in $\mathbb{R}^n$.

After mentioning circles and spheres in $\mathbb{R}^2$, we discussed lines and rays and line segments and various related angles and oriented angles between them. These all generalize easily enough to $\mathbb{R}^n$, but we only consider unoriented angles. A **line** in $\mathbb{R}^n$ is a set of the form $L = \left\{p + tu \,\middle|\, t \in \mathbb{R}\right\}$ for some $p \in \mathbb{R}^n$ and some $0 \neq u \in \mathbb{R}^n$, and we say that L is the line $x = p + tu$. Given $u, v \in \mathbb{R}^n$ with $u \neq v$, the line in $\mathbb{R}^n$ through u and v is the line $x = u + t(v - u)$. The **ray** in $\mathbb{R}^n$ from u through v is the set $R = \left\{u + t(v - u) \,\middle|\, t \geq 0\right\}$, and the **line segment** between u and v is the set

$$[u, v] = \left\{u + t(v - u) \,\middle|\, 0 \leq t \leq 1\right\} = \left\{su + tv \,\middle|\, 0 \leq s, 0 \leq t\, s + t = 1\right\}.$$

We will leave it as an exercise to determine formulas for the angle between two lines or between two rays or between two line segments.

When working in $\mathbb{R}^3$, in addition to lines we can consider **planes**. More generally, when working in $\mathbb{R}^n$ we can consider **affine spaces**: an affine space in $\mathbb{R}^n$ is a set of the form $P = p + U = \left\{p + u \,\middle|\, u \in U\right\}$ for some $p \in \mathbb{R}^n$ and some vector space $U \subseteq \mathbb{R}^n$, and we say that P is the affine space through p with associated vector space U. We leave it as an entertaining optional exercise to try to come up with a reasonable definition for the angle between two affine spaces in $\mathbb{R}^n$. The angle between the two planes $P = p + U$ and $Q = p + V$ in $\mathbb{R}^n$ is equal to the angle between their associated vector spaces U and V, and when $U \neq V$, this is equal to the angle between the two lines $L = U \cap (U \cap V)^{\perp}$ and $M = V \cap (U \cap V)^{\perp}$.

Next we discussed triangles in $\mathbb{R}^2$ (which have 3 vertices and 3 edges). The analogous object in $\mathbb{R}^3$ is a **tetrahedron** (which has 4 vertices, 6 edges, and 4 triangular faces). Recall that given 3 non-colinear points $u, v, w \in \mathbb{R}^2$, the closed solid triangle with vertices u, v and w is the set $[u, v, w] = \big\{u + s(v-u) + t(w-u) \,\big|\, 0 \leq s, 0 \leq t, s+t \leq 1\big\}$. Similarly, given 4 non-coplanar points $u, v, w, z \in \mathbb{R}^3$, the closed solid tetrahedron in $\mathbb{R}^3$ is the set

$$\begin{aligned}[u, v, w, z] &= \big\{u + r(v-u) + s(w-u) + t(z-u) \,\big|\, r, s, t \geq 0\,,\, r+s+t \leq 1\big\} \\ &= \big\{qu + rv + sw + tz \,\big|\, q, r, s, t \geq 0\,,\, q+r+s+t = 1\big\}.\end{aligned}$$

A triangle in $\mathbb{R}^2$ has 3 internal angles. For a tetrahedron in $\mathbb{R}^3$, we have a richer variety of angles that we can consider. Each triangular face has 3 interior (unoriented) angles. For each pair of faces, the faces meet along an edge, and there is an interior angle between the two faces. Given a face and given one of the 3 edges which is *not* in the face, there is an angle between the edge and the face. There is also another kind of angle, called the **solid angle**, at each vertex of the tetrahedron. The solid angle at the vertex u in the tetrahedron $[u, v, w, z]$ is the area of the portion of the unit sphere $S(u, 1)$ which lies in the solid cone $\big\{u + r(v-u) + s(w-u) + t(z-u) \,\big|\, r, s, t \geq 0\big\}$ (which is the cone obtained by extending the tetrahedron away from u). Such a region on the sphere is called a **spherical triangle**, and we shall find a formula for its area in Chapter 2. Triangles in $\mathbb{R}^2$ and tetrahedra in $\mathbb{R}^3$ are both special cases of a **simplex** in $\mathbb{R}^n$.

Finally, we remark that all of the centres of triangles in $\mathbb{R}^2$, which we discussed above, can be generalized to obtain various centres of tetrahedra in $\mathbb{R}^3$ (and, more generally, centres of simplices in $\mathbb{R}^n$). For example, if we define a **medial line** in a tetrahedron to be a line from a vertex to the centroid of the opposite face, then one can show that the 4 medial lines of a tetrahedron meet at a point, which we call the **centroid**. Alternatively, we can define a **medial plane** in a tetrahedron to be a plane which contains one of the 6 edges and passes through the midpoint of the opposite edge, and then one can show that the 6 medial planes all intersect at the centroid. As another example, we can define the **perpendicular bisector** of the line segment $[u, v]$ in $\mathbb{R}^3$ to be the plane through the midpoint $\frac{u+v}{2}$ which is perpendicular to the vector $v - u$, and then one can show that the 6 perpendicular bisectors of the edges of a tetrahedron all intersect at a point, called the **circumcentre**, which is equidistant from each of the vertices.

Isometries

1.72 Definition: An $n \times n$ matrix $A \in M_n(\mathbb{R})$ is called **orthogonal** when $A^T A = I$ or equivalently, when its columns form an orthonormal basis for $\mathbb{R}^n$. The set of all orthogonal $n \times n$ matrices is denoted by $O_n(\mathbb{R})$. An **orthogonal map** on $\mathbb{R}^n$ is a map $F : \mathbb{R}^n \to \mathbb{R}^n$ of the form $F(x) = Ax$ for some $A \in O_n(\mathbb{R})$.

1.73 Definition: For a map $S : \mathbb{R}^n \to \mathbb{R}^n$, we say that S **preserves distance** when

$$\big|S(x) - S(y)\big| = |x - y|$$

for all $x, y \in \mathbb{R}^n$. An **isometry** on $\mathbb{R}^n$ is an invertible map $S : \mathbb{R}^n \to \mathbb{R}^n$ which preserves distance. The set of all isometries on $\mathbb{R}^n$ is denoted by $\mathrm{Isom}(\mathbb{R}^n)$.

1.74 Theorem: $\mathrm{Isom}(\mathbb{R}^n)$ *is a group. This means that the identity map is an isometry, the composite of two isometries is an isometry, and the inverse of an isometry is an isometry.*

Proof: The identity map $I : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry because $\big|I(x) - I(y)\big| = |x - y|$ for all $x, y \in \mathbb{R}^n$. Note that if $S, T \in \mathrm{Isom}(\mathbb{R}^n)$ then we have $ST \in \mathrm{Isom}(\mathbb{R}^n)$ because for $x, y \in \mathbb{R}^n$ we have

$$\big|S(T(x)) - S(T(y))\big| = \big|T(x) - T(y)\big| = |x - y|.$$

Finally, note that if $S \in \mathrm{Isom}(\mathbb{R}^n)$ then $S^{-1} \in \mathrm{Isom}(\mathbb{R}^n)$ because given $u, v \in \mathbb{R}^n$, if we let $x = S^{-1}(u)$ and $y = S^{-1}(v)$ so that $u = S(x)$ and $v = S(y)$ then we have

$$\big|S^{-1}(u) - S^{-1}(v)\big| = |x - y| = |S(x) - S(y)| = |u - v|.$$

1.75 Example: For a vector $u \in \mathbb{R}^n$, the **translation** by u is the map $T_u : \mathbb{R}^n \to \mathbb{R}^n$ given by $T_u(x) = x + u$. Note that T_u is an isometry on $\mathbb{R}^n$ because

$$\big|T_u(x) - T_u(y)\big| = \big|(u + x) - (u + y)\big| = |x - y|\,.$$

1.76 Example: If $A \in O_n(\mathbb{R})$, so that $A^T A = I$, then the map $S : \mathbb{R}^n \to \mathbb{R}^n$ given by $S(x) = Ax$ is an isometry because for $x, y \in \mathbb{R}^n$ we have

$$\begin{aligned}\big|Ax - Ay\big|^2 &= \big|A(x - y)\big|^2 = \big(A(x - y)\big)^T \big(A(x - y)\big) \\ &= (x - y)^T A^T A(x - y) = (x - y)^T (x - y) = |x - y|^2.\end{aligned}$$

1.77 Example: For a vector space U in $\mathbb{R}^n$, the **reflection** in U is the map $F_U : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$F_U(x) = x - 2\,\mathrm{Proj}_{U^\perp}(x)$$

where $\mathrm{Proj}_{U^\perp}(x)$ is the orthogonal projection of x onto $U^\perp$. When $\{u_1, u_2, \cdots, u_k\}$ is an orthonormal basis for $U^\perp$ and $A = \big(u_1, u_2, \cdots, u_k\big) \in M_{n\times k}(\mathbb{R})$, we have

$$\mathrm{Proj}_{U^\perp}(x) = \sum_{i=1}^{n} (x \cdot u_i) u_i = AA^T x\,,$$
$$F_U(x) = x - 2\,AA^T x = (I - 2AA^T)x.$$

Note that since $\{u_1, u_2, \cdots, u_k\}$ is orthonormal, we have $A \in O_n(\mathbb{R})$, that is $A^T A = I$, and it follows that $(I - 2AA^T) \in O_n(\mathbb{R})$ because

$$\begin{aligned}(I - 2AA^T)^T (I - 2AA^T) &= I - 2AA^T - 2AA^T + 4AA^T AA^T \\ &= I - 4AA^T + 4A(A^T A)A^T = I - 4AA^T + 4AA^T = I.\end{aligned}$$

This shows that $F_U \in O_n(\mathbb{R})$ and hence $F_U \in \text{Isom}(\mathbb{R}^n)$. In particular, when U is a hyperspace (that is a vector space of dimension $n-1$) and u is a non-zero vector in $U^\perp$, we have

$$\text{Proj}_{U^\perp}(x) = \text{Proj}_u(x) = \frac{x \cdot u}{|u|^2}\, u \;\text{ and }\; F_U(x) = x - 2\frac{x \cdot u}{|u|^2}\, u.$$

1.78 Example: For the affine space $P = p + U = \{p + u \,|\, u \in U\}$, where $p \in \mathbb{R}^n$ and $U \subseteq \mathbb{R}^n$ is a subspace, the **reflection** in P is the map $F_P : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$F_P(x) = p + F_U(x - p).$$

Note that $F_P \in \text{Isom}(\mathbb{R}^n)$ because F_P is equal to the composite $F_P = T_p F_U T_{-p}$.

1.79 Theorem: *(The Algebraic Classification of Isometries on $\mathbb{R}^n$) A map $S : \mathbb{R}^n \to \mathbb{R}^n$ preserves distance if and only if S is of the form $S(x) = Ax + b$ for some $A \in O_n(\mathbb{R})$ and some $b \in \mathbb{R}^n$.*

Proof: First note that if $S(x) = Ax + b$ where $A \in O_n(\mathbb{R})$ and $b \in \mathbb{R}^n$, then S is the composite $S = T_b A$, which is an isometry.

Conversely, suppose that S is an isometry. Let $b = S(0)$ and define $L : \mathbb{R} \to \mathbb{R}$ by $L(x) = S(x) - b$. Note that $L(0) = 0$ and that for $x \in \mathbb{R}^n$ we have

$$\big|L(x)\big| = \big|L(x) - L(0)\big| = \big|(S(x) - b) - (S(0) - b)\big| = |S(x) - S(0)| = |x - 0| = |x|.$$

For $x, y \in \mathbb{R}^n$, we have

$$|x - y|^2 = (x - y) \cdot (x - y) = x \cdot x - x \cdot y - y \cdot x + y \cdot y = |x|^2 - 2\, x \cdot y + |y^2|$$

from which we obtain the following version of the Polarization Identity:

$$x \cdot y = \tfrac{1}{2}\big(|x|^2 + |y|^2 - |x - y|^2\big).$$

For $x, y \in \mathbb{R}^n$, using the Polarization Identity, we have

$$L(x) \cdot L(y) = \tfrac{1}{2}\big(|L(x)|^2 + |L(y)|^2 - |L(x) - L(y)|^2\big) = \tfrac{1}{2}\big(|x|^2 + |y|^2 - |x - y|^2\big) = x \cdot y.$$

In particular, $L(e_i) \cdot L(e_j) = e_i \cdot e_j = \delta_{i,j}$ for all i, j, so the set $\{L(e_1), L(e_2), \cdots, L(e_n)\}$ is an orthonormal basis for $\mathbb{R}^n$. For $x \in \mathbb{R}^n$, if we write $x = \sum\limits_{i=1}^{n} x_i e_i$ and $L(x) = \sum\limits_{i=1}^{n} t_i L(e_i)$ then we have

$$t_k = L(x) \cdot L(e_k) = x \cdot e_k = x_k$$

and so we have $L(x) = \sum x_k L(e_k) = Ax$ where $A = \big(L(e_1), L(e_2), \cdots, L(e_n)\big) \in M_n(\mathbb{R})$. Since $\{L(e_1), L(e_2), \cdots, L(e_n)\}$ is an orthonormal set, it follows that $A^T A = I$ so we have $A \in O_n(\mathbb{R})$. Thus $S(x) = Ax + b$ with $A \in O_n(\mathbb{R})$ and $b \in \mathbb{R}^n$, as required.

1.80 Corollary: *Every distance preserving map $S : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry.*

Proof: If $S : \mathbb{R}^n \to \mathbb{R}^n$ preserves distance then S is invertible; indeed if S is given by $S(x) = Ax + b$ with $A \in O_n(\mathbb{R})$ and $b \in \mathbb{R}^n$ then S^{-1} is given by $S^{-1}(x) = A^{-1}x - A^{-1}b$.

1.81 Definition: Let $S \in \text{Isom}(\mathbb{R}^n)$, say $S(x) = Ax + b$ with $A \in O_n(\mathbb{R})$ and $b \in \mathbb{R}^n$. Note that since $A^T A = I$ we have $\det(A) = \pm 1$. We say that S **preserves orientation** when $\det(A) = 1$, and we say that S **reverses orientation** when $\det(A) = -1$.

1.82 Example: The following maps are all isometries on $\mathbb{R}^2$.

(1) The **identity** map is the map $I:\mathbb{R}^2\to\mathbb{R}^2$ given by $I(x)=x$.
(2) For $u\in\mathbb{R}^2$, the **translation** by u is the map $T_u:\mathbb{R}^2\to\mathbb{R}^2$ given by $T_u(x)=x+u$.
(3) For $p\in\mathbb{R}^2$ and $\theta\in\mathbb{R}$, the **rotation** about p by θ is the map $R_{p,\theta}:\mathbb{R}^2\to\mathbb{R}^2$ given by

$$R_{p,\theta}(x)=p+R_\theta(x-p)\ \text{ where }\ R_\theta=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

(4) For a line L in $\mathbb{R}^2$, the **reflection** in L is the map $F_L:\mathbb{R}^2\to\mathbb{R}^2$ which is given by any of the following three equivalent formulas. When L is the line in $\mathbb{R}^2$ through p perpendicular to u, we have

$$F_L(x)=x-\frac{2(x-p)\cdot u}{|u|^2}\,u.$$

When L is the line $ax+by+c=0$, the above formula becomes

$$F_L(x,y)=(x,y)-\frac{2(ax+by+c)}{a^2+b^2}\,(a,b).$$

When L is the line through p in the direction of the vector $\left(\cos\frac{\theta}{2},\sin\frac{\theta}{2}\right)$, F_L is given by

$$F_L(x)=p+F_\theta(x-p)\ \text{ where }\ F_\theta=\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}.$$

(5) For a vector $u\in\mathbb{R}^2$ and a line L in $\mathbb{R}^2$ which is parallel to u, the **glide reflection** $G_{u,L}:\mathbb{R}^2\to\mathbb{R}^2$ is the composite

$$G_{u,L}=T_uF_L=F_LT_u$$

(when L is not parallel to u, the composites T_uF_L and F_LT_u are not equal, and they are not called glide reflections).

Of the above examples, the maps I, T_u and $R_{p,\theta}$ all preserve orientation, while the maps F_L and $G_{u,L}$ reverse orientation.

1.83 Theorem: *(Composites of Reflections in $\mathbb{R}^2$) Let L and M be lines in $\mathbb{R}^2$.*

(1) If $L=M$ then $F_MF_L=I$.
(2) If L is parallel to M then $F_MF_L=T_{2u}$ where u is the vector from L orthogonally to M.
(3) If $L\cap M=\{p\}$ then $F_MF_L=R_{p,2\theta}$ where θ is the angle from L counterclockwise to M.

Proof: Suppose first that $L=M$. Say L is the line through p in the direction of $\left(\cos\frac{\theta}{2},\sin\frac{\theta}{2}\right)$ so that $F_L(x)=p+F_\theta(x-p)$. Then for all $x\in\mathbb{R}^2$ we have

$$F_LF_L(x)=F_L\big(p+F_\theta(x-p)\big)=p+F_\theta\big(F_\theta(x-p)\big)=p+x-p=x=I(x).$$

Next, suppose that L is parallel to M. Let u be the vector from L orthogonally to M, let $p\in L$ and let $q=p+u\in M$. Then for $x,y\in\mathbb{R}^2$ we have $F_L(x)=x-\frac{2(x-p)\cdot u}{|u|^2}\,u$ and

$F_M(y) = y - \frac{2(y-p-u)\bullet u}{|u|^2}\,u$ and so

$$\begin{aligned} F_M F_L(x) &= F_M\Big(x - \tfrac{2(x-p)\bullet u}{|u|^2}\,u\Big) \\ &= \Big(x - \tfrac{2(x-p)\bullet u}{|u|^2}\,u\Big) - \frac{2\Big(x-p-u-\frac{2(x-p)\bullet u}{|u|^2}\,u\Big)\bullet u}{|u|^2}\,u \\ &= x - \tfrac{2(x-p)\bullet u}{|u|^2}\,u - \tfrac{2(x-p)\bullet u}{|u|^2}\,u + \tfrac{2u\bullet u}{|u|^2}\,u + \frac{4\big((x-p)\bullet u\big)(u\bullet u)}{|u|^4}\,u \\ &= x + 2u = T_{2u}(x). \end{aligned}$$

Finally, suppose that $L \cap M = \{p\}$. Say L is in the direction of $\big(\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2}\big)$ and M is in the direction of $\big(\cos\frac{\beta}{2}, \sin\frac{\beta}{2}\big)$. Then for $x, y \in \mathbb{R}^2$ we have $F_L(x) = p + F_\alpha(x-p)$ and $F_M(y) = p + F_\beta(y-p)$ and so

$$F_M F_L(x) = F_M\big(p + F_\alpha(x-p)\big) = p + F_\beta\big(F_\alpha(x-p)\big) = p + R_{\beta-\alpha}(x-p) = R_{p,2\theta}(x)$$

where $\theta = \frac{\beta}{2} - \frac{\alpha}{2}$, which is the angle from L to M.

1.84 Theorem: *(The Geometric Classification of Isometries on $\mathbb{R}^2$) Every isometry on $\mathbb{R}^2$ is equal to one of the maps I , T_u , $R_{p,\theta}$, F_L , $G_{u,L}$.*

Proof: Let $S \in \text{Isom}(\mathbb{R}^2)$, say $S(x) = Ax + b$ with $A \in O_2(\mathbb{R})$ and $b \in \mathbb{R}^2$. Recall that the elements in $O_2(\mathbb{R})$ are the rotation and reflection matrices R_θ and F_θ, and so with $S = T_u R_\theta$ or $S = T_u F_\theta$ where $u = -b$. First suppose that $S = T_u R_\theta$. Let M be the line through the origin perpendicular to u. Let $L = R_{-\theta/2}(M)$ so that $F_M F_L = R_\theta$. Let $N = T_{u/2}(M)$ so that $T_u = F_N F_M$. Then $S = T_u R_\theta = F_N F_M F_M F_L = F_N F_L$. By the above theorem, S is equal to the identity, a translation, or a rotation.

Now suppose that $S = T_u F_\theta$. Let L be the line through the origin in the direction of $\big(\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\big)$ so that $F_\theta = F_L$. Let M be the line through the origin which is perpendicular to u and let $N = T_{u/2}(M)$ so that $T_u = F_n F_M$. Then we have $S = F_N F_M F_L$. Note that $F_M F_L = R_{2\alpha}$ where α is the angle from L to M. Let $N' = N$, let M' be the line through $(0,0)$ which is perpendicular to N', and let $L' = R_{-\alpha}(M')$ so that $F_{M'} F_{L'} = R_{2\alpha}$. Then $S = F_N F_M F_L = F_N R_{2\alpha} = F_{N'} F_{M'} F_{L'} = R_{p,\pi} F_{L'}$ where p is the point of intersection of M' and N' (which are perpendicular). Let $L'' = L'$, let M'' be the line through p parallel to L' and let $N'' = R_{p.\pi/2}(M'')$ so that $R_{p,\pi} = F_{N''} F_{M''}$. Then we have $S = R_{p,\pi} F_{L'} = F_{N''} F_{M''} F_{L''}$. Since L'' is parallel to M'' we have $F_{M''} F_{L''} = T_{2v}$ where v is the vector from L'' to M''. Since L'' and M'' are perpendicular to N'', the vector v is parallel to N'' and so $S = F_{N''} T_v$ is a glide reflection (or a reflection in the case that $v = 0$).

1.85 Example: The following maps are all isometries on $\mathbb{R}^3$.
(1) the **identity** map is the map $I : \mathbb{R}^3 \to \mathbb{R}^3$ given by $I(x) = x$.
(2) For $u \in \mathbb{R}^3$, the **translation** by u is the map $T_u : \mathbb{R}^3 \to \mathbb{R}^3$ given by $T_u(x) = x + u$.
(3) For a point $p \in \mathbb{R}^3$, a nonzero vector $0 \neq u \in \mathbb{R}^3$ and an angle $\theta \in \mathbb{R}$ the **rotation** $R_{p,u,\theta} : \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$R_{p,u,\theta}(x) = p + R_{u,\theta}(x - p)$$

where $R_{u,\theta}$ is the rotation in $\mathbb{R}^3$ about the vector u by the angle θ; if $\{u, v, w\}$ is a positively oriented orthogonal basis for $\mathbb{R}^3$ with all three vectors u, v and w of the same length, then $R = R_{u,\theta}$ is given by $R(u) = u$, $R(v) = (\cos\theta)v + (\sin\theta)w$ and $R(w) = -(\sin\theta)v + (\cos\theta)w$.
(4) For a point $p \in \mathbb{R}^3$, a nonzero vector $0 \neq u \in \mathbb{R}^3$ and an angle $\theta \in \mathbb{R}$ the **twist** $W_{p,u,\theta} : \mathbb{R}^3 \to \mathbb{R}^3$ is the composite $W_{p,u,\theta} = T_u R_{p,u,\theta} = R_{p,u,\theta} T_u$.
(5) For a plane P in $\mathbb{R}^3$, the **reflection** in P is the map $F_P : \mathbb{R}^3 \to \mathbb{R}^3$ described in Example 1.78.
(6) For a vector $u \in \mathbb{R}^3$ and a plane P in $\mathbb{R}^3$ which is parallel to u, the **glide reflection** $G_{u,P} : \mathbb{R}^3 \to \mathbb{R}^3$ is the composite $G_{u,P} = T_u F_P = F_P T_u$.
(7) For a point $p \in \mathbb{R}^3$, a nonzero vector $0 \neq u \in \mathbb{R}^3$ and an angle $\theta \in \mathbb{R}$, the **rotary reflection** $H_{p,u,\theta} : \mathbb{R}^3 \to \mathbb{R}^3$ is the composite $H_{p,u,\theta} = R_{p,u,\theta} F_P = F_P R_{p,u,\theta}$ where P is the plane through p perpendicular to u.

1.86 Theorem: *(The Geometric Classification of Isometries in $\mathbb{R}^3$) Every isometry on $\mathbb{R}^3$ is equal one of the following*

$$I\,,\ T_u\,,\ R_{p,u,\theta}\,,\ W_{p,u,\theta}\,,\ F_P\,,\ G_{u,P}\,,\ H_{p,u,\theta}\,.$$

Proof: We omit the proof.