PMATH 321: Non-Euclidean Geometry

Instructor: Stephen New

Spring 2021

Table of Contents

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

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 $u \in U = \mathrm{Span}(\mathcal{B})$, 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 $x \in \mathrm{Null}A \iff x \cdot v_i = 0 \text{ for all } i \iff x \in \mathrm{Span}\{v_1, v_2, \cdots, v_k\}^\perp = (\mathrm{Row}A)^\perp$ by part 2.

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

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 $(\mathrm{Null}A)^\perp = \left((\mathrm{Row}A)^\perp\right)^\perp = \mathrm{Row}A$, 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 $\mathrm{Proj}_U(x)$, 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

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

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

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 $A^TA = \mathrm{diag}(|u_1|^2, \cdots, |u_l|^2)$. Since each $u_i \neq 0$ we see that A^TA is invertible. Since $\mathrm{rank}(A) = \mathrm{rank}(A^TA) = l$, 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

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

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

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 $A^TA = \mathrm{diag}(|u_1|^2, \cdots, |u_l|^2)$, which is invertible. Then

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

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

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 $0\leq t\leq 1$ $\big($for $0\leq t\leq\frac14$ rotate the vectors until $U\big(\frac14\big)$ points in the direction of the positive x-axis, then for $\frac14\leq t\leq\frac12$ 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 $\frac12\leq t\leq\frac34$ 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 $\frac34\leq t\leq 1$ scale the two vectors until $U(1) = e_1$ and $V(1) = e_2\big)$. Let $W(t) = U(t)\times V(t)$ for $0\leq t\leq 1$. 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 $U=\mathrm{Span}\{u,v\}$. Then we have

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

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

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

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) $\det(u,v,w)=\det(v,w,u)=\det(w,u,v)$
$\qquad=-\det(u,w,v)=-\det(w,v,u)=-\det(v,u,w)$,
(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 $\big\{\theta+2\pi k\big|k\in\mathbb{Z}\big\}$ 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 $\big\{\theta+2\pi k\big|k\in\mathbb{Z}\big\}$ 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 $\theta_o=\theta_o(u)\in\mathbb{R}/2\pi$ 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

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

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 $L = \{(x,y) \in \mathbb{R}^2 \big| ax + by = c\}$ 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 $L = \{x \in \mathbb{R}^2 \big| x \cdot n = c\}$ 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 $L = \{x \in \mathbb{R}^2 \big| x = p + tu$ 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 $L = \{x \in \mathbb{R}^2 \big| x \cdot n = c\}$, 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 $L = \{x \in \mathbb{R}^2 \big| x = p + tu$ 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 $L = \{u + t(v-u) \big| t \in \mathbb{R}\}$. The **ray** from u through v (or the ray from u in the direction of the vector $v - u$) is the set $R = \{u + t(v-u) \big| 0 \leq t \in \mathbb{R}\}$. The **line segment** between u and v is the set

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

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 $u+\frac{2}{3}\big(\frac{v+w}{2}-u\big)=\frac{1}{3}u+\frac{1}{3}(v+w)=\frac{1}{3}(u+v+w)$. Similarly, the point which lies $\frac{2}{3}$ of the way along the median from v to $\frac{w+u}{2}$ is the point $v=\frac{2}{3}\big(\frac{w+u}{2}-v\big)=\frac{1}{3}(u+v+w)$ and the point which lies $\frac{2}{3}$ of the way from w to $\frac{u+v}{2}$ is the point $w+\frac{2}{3}\big(\frac{u+v}{2}-w\big)=\frac{1}{3}(u+v+w)$. 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

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

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

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

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

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

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 $a = p + \mathrm{Proj}_n(x-p) = p + \frac{(x-p)\cdot n}{|n|^2}\, n$ and so the distance between x and L is given by

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

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

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

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 $i \in C$.

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** $S(u,r) = \big\{x \in \mathbb{R}^n \,\big|\, d_E(x,u) = r\big\}$ and the (closed) **ball** $B(u,r) = \big\{x \in \mathbb{R}^n \,\big|\, d_E(x,u) \leq r\big\}$. 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

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

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 $L = \big\{p + tu \,\big|\, t \in \mathbb{R}\big\}$ 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 $R = \big\{u + t(v-u) \,\big|\, t \geq 0\big\}$, and the **line segment** between u and v is the set

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

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 $P = p + U = \big\{p + u \,\big|\, u \in U\big\}$ 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

$$\left|S(x) - S(y)\right| = |x - y|$$

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 $\left|I(x) - I(y)\right| = |x - y|$ 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

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

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

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

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

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

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

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

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

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

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

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

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

Finally, suppose that $L\cap M = \{p\}$. Say L is in the direction of $\left(\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2}\right)$ and M is in the direction of $\left(\cos\frac{\beta}{2}, \sin\frac{\beta}{2}\right)$. 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 $S \in \mathrm{Isom}(\mathbb{R}^2)$, 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 $S = T_u R_\theta = F_N F_M F_M F_L = F_M F_L$. 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 $\left(\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\right)$ 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 $W_{p,u,\theta}=T_uR_{p,u,\theta}=R_{p,u,\theta}T_u$.
(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 $G_{u,P}=T_uF_P=F_PT_u$.
(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 $H_{p,u,\theta}=R_{p,u,\theta}F_P=F_PR_{p,u,\theta}$ 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.

Chapter 2. Spherical Geometry

The Sphere, Spherical Distance, and Spherical Circles and Lines

2.1 Definition: The (unit) **sphere** is defined to be the set

$$\mathbb{S}^2 = \left\{u \in \mathbb{R}^3 \big| |u| = 1\right\}.$$

If $u \in \mathbb{S}^2$ then we also have $-u \in \mathbb{S}^2$, and we say that the two points $\pm u$ are **antipodal**. Given points $u, v \in \mathbb{S}^2$, we define the (spherical) **distance** between u and v to be

$$d_S(u, v) = \theta(u, v) = \cos^{-1}(u \cdot v).$$

Note that $0 \leq d_S(u, v) \leq \pi$ with $d_S(u, v) = 0$ when $u = v$ and $d_S(u, v) = \pi$ when $v = -u$.

2.2 Theorem: *(Euclidean and Spherical Distance) Given points $u, v \in \mathbb{S}^2$, the spherical distance $d_S(u, v)$ determines and is determined by the Euclidean distance $d_E(u, v)$.*

Proof: First we note that since $u, v \in \mathbb{S}^2$ we have $d_S(u, v) \in [0, \pi]$ and $d_E(u, v) \in [0, 2]$. Applying the Law of Cosines to the triangle in $\mathbb{R}^3$ with vertices at 0, u and v gives

$$d_E(u, v)^2 = 2 - 2\cos\theta(u, v) = 2 - 2\cos d_S(u, v).$$

Thus we have

$$\begin{aligned} d_E(u, v) &= \sqrt{2 - 2\cos d_S(u, v)} \text{ and} \\ d_S(u, v) &= \cos^{-1}\left(1 - \tfrac{1}{2} d_E(u, v)^2\right). \end{aligned}$$

2.3 Theorem: *(Metric Properties of Spherical Distance) For all $u, v, w \in \mathbb{S}^2$ we have*
(1) (Positive Definiteness) $d_S(u, v) \in [0, \pi]$ with $d_S(u, v) = 0$ if and only if $u = v$ and $d_S(u, v) = \pi$ if and only if $u = -v$,
(2) (Symmetry) $d_S(u, v) = d_S(v, u)$, and
(3) (Triangle Inequality) $d_S(u, w) \leq d_S(u, v) + d_S(v, w)$.

Proof: We leave the proofs of Parts (1) and (2) as an exercise. To prove Part (3), note that

$$\begin{aligned} \cos\big(\theta(u, v) + \theta(v, w)\big) &= \cos\theta(u, v)\cos\theta(v, w) - \sin\theta(u, v)\sin\theta(v, w) \\ &= (u \cdot v)(v \cdot w) - |u \times v||v \times w| \\ &\leq (u \cdot v)(v \cdot w) - (u \times v) \cdot (v \times w) \\ &= (u \cdot v)(v \cdot w) - \big((u \cdot v)(v \cdot w) - (u \cdot w)(v \cdot v)\big) \\ &= u \cdot w = \cos\theta(u, w). \end{aligned}$$

Since $\cos\theta$ is decreasing with θ, it follows that $\theta(u, w) \leq \theta(u, v) + \theta(v, w)$.

2.4 Theorem: *(Spherical Area) Given two parallel planes which intersect with $\mathbb{S}^2$, the area of the portion of $\mathbb{S}^2$ which lies between the two planes is equal to $2\pi\Delta$ where Δ is the distance between the two planes. In particular, the total area of $\mathbb{S}^2$ is equal to 4π.*

Proof: Rotate the sphere about the origin so that the two parallel planes have equations $x = a$ and $x = b$ with $-1 \leq a \leq b \leq 1$ and note that the distance between the two planes is $\Delta = b - a$. Recall, from calculus, that the area of the surface which is obtained by revolving the graph of $z = f(x)$ for $a \leq x \leq b$ about the x-axis is equal to

$$A = \int_a^b 2\pi f(x)\sqrt{1 + f'(x)^2}\, dx.$$

We apply this formula with $f(x)=\sqrt{1-x^2}$ and $f'(x)=\frac{-x}{\sqrt{1-x^2}}$ to get

$$A=\int_a^b 2\pi\sqrt{1-x^2}\sqrt{1+\frac{x^2}{1-x^2}}\,dx=\int_a^b 2\pi\sqrt{1-x^2}\frac{1}{\sqrt{1-x^2}}\,dx=\int_a^b 2\pi\,dx=2\pi(b-a).$$

2.5 Definition: For $u\in\mathbb{S}^2$ and $r\in[0,\pi]$, the (spherical) **circle** of radius r centred at u, and the (closed spherical) **disc** of radius r centred at u are the sets

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

Note that when $r=0$ we have $C(u,r)=\{u\}$ and $D(u,r)=\{u\}$ and when $r=\pi$ we have $C(u,r)=\{-u\}$ and $D(u,r)=\mathbb{S}^2$.

2.6 Theorem: *(Spherical Circles) A circle in $\mathbb{S}^2$ is a set of the form $C=\mathbb{S}^2\cap P$ where P is a plane in $\mathbb{R}^3$ whose Euclidean distance from the origin is at most 1.*

Proof: For $u\in\mathbb{S}^2$ and $r\in[0,\pi]$ we have

$$\begin{aligned}C(u,r)&=\big\{x\in\mathbb{S}^2\big|d_S(x,u)=r\big\}=\big\{x\in\mathbb{S}^2\big|\cos d_S(x,u)=\cos r\big\}\\ &=\big\{x\in\mathbb{S}^2\big|x\cdot u=\cos r\big\}=\mathbb{S}^2\cap P\end{aligned}$$

where $P=\big\{x\in\mathbb{R}^3\big|x\cdot u=\cos r\big\}$, which is the plane in $\mathbb{R}^3$ perpendicular to the vector u whose nearest point to the origin is the point $(\cos r)u$.

2.7 Theorem: *(Circumference and Area of Spherical Circles and Discs) Let $u\in\mathbb{S}^2$ and $r\in[0,\pi]$. Then the circumference L of $C(u,r)$ and the area A of $D(u,r)$ are given by*

$$\begin{aligned}L&=2\pi\sin r\text{ , and}\\ A&=2\pi(1-\cos r).\end{aligned}$$

Proof: Let $P=\big\{x\in\mathbb{R}^3\big|x\cdot u=\cos r\big\}$ so that P is the plane whose nearest point to the origin is $(\cos r)u$. Note that for each point $x\in C(u,r)$, the triangle in $\mathbb{R}^3$ with vertices 0, $(\cos r)u$ and x is a right-angle triangle with its right-angle at $(\cos r)u$ whose angle at 0 is equal to $\theta=\theta(x,u)=\cos^{-1}(x\cdot u)=r$. and whose sides are of length $\cos r$, $\sin r$ and 1. The spherical circle $C(u,r)$ is equal to the Euclidean circle in the plane P centred at the point $(\cos r)u$ of Euclidean radius $\sin r$, and so its circumference is $L=2\pi\sin r$.

Let $Q=\big\{x\in\mathbb{R}^3\big|x\cdot u=1\big\}$ so that Q is the plane whose nearest point to the origin is the point u (Q is the tangent plane to $\mathbb{S}^2$ at the point u). Note that the distance between P and Q is $\Delta=1-\cos r$. The spherical disc $D(u,r)$ is equal to the portion of $\mathbb{S}^2$ which lies between the parallel planes P and Q, and so its area is $A=2\pi\Delta=2\pi(1-\cos r)$.

2.8 Definition: A (spherical) **line** (also called a **great circle**) in $\mathbb{S}^2$ is a spherical circle of the form $L=\mathbb{S}^2\cap P$ for some plane P through the origin. Note that each plane P through the origin has two unit normal vectors $\pm u\in\mathbb{S}^2$, and these two normal vectors are called the **poles** of the line $L=\mathbb{S}^2\cap P$. Given $u\in\mathbb{S}^2$, the line in $\mathbb{S}^2$ with poles $\pm u$ is denoted by L_u, so we have

$$L_u=\big\{x\in\mathbb{S}^2\big|x\cdot u=0\big\}=C\big(u,\tfrac{\pi}{2}\big).$$

Two lines in $\mathbb{S}^2$ are said to be **orthogonal** (or **perpendicular**) when their poles are orthogonal (or equivalently when their associated planes are orthogonal).

2.9 Theorem: *(Properties of Spherical Lines)*

(1) Given $u, v \in \mathbb{S}^2$ with $v \neq \pm u$, there is a unique line in $\mathbb{S}^2$ through u and v.
(2) The intersection of any two distinct lines in $\mathbb{S}^2$ consists of two antipodal points.
(3) Given a point $u \in \mathbb{S}^2$ and a line L in $\mathbb{S}^2$ with $L \neq L_u$, there is a unique line in $\mathbb{S}^2$ which passes through u and is perpendicular to L.
(4) There is a unique line which is perpendicular to any two distinct given lines.

Proof: To prove Part 1, let $u, v \in \mathbb{S}^2$ with $u \neq \pm v$. Since $|u| = |v| = 1$ and $u \neq \pm v$, it follows that $\{u, v\}$ is linearly independent, so there is a unique plane P through 0 which contains both u and v, namely $P = \text{Span}\{u, v\}$, so there is a unique (spherical) line in $\mathbb{S}^2$ which contains u and v, namely $L = \mathbb{S}^2 \cap P$. We note that $L = L_w$ with $w = \pm \frac{u \times v}{|u \times v|}$.

To prove Part 2, let L and M be two distinct (spherical) lines in $\mathbb{S}^2$, say $L = \mathbb{S}^2 \cap P$ and $M = \mathbb{S}^2 \cap Q$ where P and Q are distinct planes through 0. The two planes intersect in a line N though 0, and N intersects the unit sphere at two antipodal points, say $\pm w$, and we have $L \cap M = \mathbb{S}^2 \cap P \cap Q = \mathbb{S}^2 \cap N = \{\pm w\}$. Note that if $L = L_u$ and $M = L_v$ so that u and v are unit normal vectors for P and Q, then $u \times v \in P \cap Q = N$ and so $L \cap M = \{\pm w\}$ where $w = \frac{u \times v}{|u \times v|}$.

To prove Part 3, let $u \in \mathbb{S}^2$ and let L be a line in $\mathbb{S}^2$ with $L \neq L_u$, say $L = L_v$ with $v \neq \pm u$. We have $L = \mathbb{S}^2 \cap P$ where P is the plane through 0 with unit normal vector v. The planes Q through 0 which are perpendicular to P are the planes through 0 which pass through v, and so the lines $M = \mathbb{S}^2 \cap Q$ which are perpendicular to $L = \mathbb{S}^2 \cap P = L_v$ are the lines in $\mathbb{S}^2$ through v. Thus the (unique) line which is perpendicular to L and passes through u is the same as the (unique) line through v and u (which exists and is unique by Part a). We note that this unique line is the line L_w with $w = \frac{u \times v}{|u \times v|}$.

To prove Part 4, let L and M be two distinct lines, say $L = L_u$ and $M = L_v$ where $u, v \in \mathbb{S}^2$ with $u \neq \pm v$. As is our proof of Part 3, the lines in $\mathbb{S}^2$ which are perpendicular to $L = L_u$ are precisely the lines in $\mathbb{S}^2$ through u, and the lines in $\mathbb{S}^2$ which are perpendicular to $M = L_v$ are precisely the lines in $\mathbb{S}^2$ through v, and so the (unique) line in $\mathbb{S}^2$ perpendicular to both $L = L_u$ and $M = L_v$ is the (unique) line in $\mathbb{S}^2$ through u and v. We note that the unique line perpendicular to $L = L_u$ and $M = L_v$ is the line $N = L_w$ where $w = \frac{u \times v}{|u \times v|}$.

Oriented Angles

2.10 Definition: For $u \in \mathbb{S}^2$, the **tangent space** to $\mathbb{S}^2$ at u is the 2-dimensional vector space T_u which is orthogonal to u, that is the space

$$T_u = \left\{ x \in \mathbb{R}^3 \middle| x \cdot u = 0 \right\}.$$

2.11 Definition: For $u \in \mathbb{S}^2$ and $0 \neq v, w \in T_u$, we define the **oriented angle** from v to w to be the angle $\theta_o(v,w) \in [0, 2\pi)$ from v counterclockwise to w when looking at the plane T_u with the vector u pointing towards us. The (unoriented) **angle** between u and v is the same as the (unoriented) angle $\theta(u,v) = \cos^{-1} \frac{u \cdot v}{|u|\,|v|}$ between u and v in $\mathbb{R}^3$. We have $\theta(u,v) = \theta_o(u,v)$ if $\theta_o(u,v) \in [0,\pi]$, and $\theta(u,v) = 2\pi - \theta_o(u,v)$ if $\theta_o(u,v) \in [\pi, 2\pi)$.

2.12 Theorem: *(Angle Formula) Let $u \in \mathbb{S}^2$ and let $0 \neq v, w \in T_u$. Then*

$$\cos \theta_o(v,w) = \frac{v \cdot w}{|v||w|} \text{ and}$$
$$\sin \theta_o(v,w) = \frac{\det(u,v,w)}{|v||w|}.$$

Proof: When $v \times w = 0$ so that $\{v, w\}$ is linearly dependent, if $w = tv$ with $t > 0$ then $\theta_o(v,w) = \theta(v,w) = 0$ and if $w = tv$ with $t < 0$ then $\theta_o(v,w) = \theta(v,w) = \pi$, and in both cases we have $\sin \theta_o(v,w) = \det(u,v,w) = 0$.

Suppose that $v \times w \neq 0$. Since $v, w \in T_u$ so that v and w are orthogonal to u it follows that the vector $v \times w$ points either in the direction of u or in the direction of $-u$. When $\theta_o(v,w) \in (0,\pi)$, the Right-Hand Rule implies that $v \times w$ points in the direction of u so that $\theta(u, v \times w) = 0$, and in this case we have $\theta_o(v,w) = \theta(v,w)$. When $\theta_o(v,w) \in (\pi, 2\pi)$, the Right-Hand Rule implies that $v \times w$ points in the direction of $-u$ so that $\theta(u, v \times w) = \pi$, and in this case we have $\theta_o(v,w) = 2\pi - \theta(v,w)$. Note that

$$\cos \theta(u, v \times w) = \frac{u \cdot (v \times w)}{|u||v \times w|} = \frac{\det(u,v,w)}{|v \times w|} \text{ so that}$$
$$\det(u,v,w) = |v \times w| \cos \theta(u, v \times w) = \begin{cases} \ \ |v \times w| \text{ if } \theta_o(v,w) \in (0,\pi), \\ -|v \times w| \text{ if } \theta_o(v,w) \in (\pi, 2\pi). \end{cases}$$

In the case that $\theta_o(v,w) \in (0,\pi)$ we have

$$\cos \theta_o(v,w) = \cos \theta(v,w) = \frac{v \cdot w}{|v||w|} \text{ , and}$$
$$\sin \theta_o(v,w) = \sin \theta(v,w) = \frac{|v \times w|}{|v||w|} = \frac{\det(u,v,w)}{|v||w|}$$

and in the case that $\theta_o(v,w) \in (\pi, 2\pi)$ we have

$$\cos \theta_o(v,w) = \cos\big(2\pi - \theta(v,w)\big) = \cos \theta(v,w) = \frac{v \cdot w}{|v||w|} \text{ , and}$$
$$\sin \theta_o(v,w) = \sin\big(2\pi - \theta(v,w)\big) = -\sin \theta(v,w) = \frac{-|v \times w|}{|v||w|} = \frac{\det(u,v,w)}{|v||w|}.$$

2.13 Remark: Definition 2.11 is not actually a rigorous mathematical definition, and the proof of the above theorem is not actually a rigorous mathematical proof (since the words "counterclockwise" and "right-hand rule" are not rigorously defined). To be rigorous, we would simply take the formulas in Theorem 2.12 as our definition for the oriented angle, and then the above so-called proof can be taken as an informal motivation for the definition.

2.14 Definition: Let $u, v \in \mathbb{S}^2$ with $v \neq \pm u$. The line in $\mathbb{S}^2$ through u and v is cut, at the points u and v, into two arcs from u to v. The shorter of the two arcs is called the (spherical) **line segment** from u to v and is denoted by $[u, v]$. Note that

$$[u, v] = \left\{ x \in \mathbb{S}^2 \middle| x = su + tv \text{ for some } 0 \leq s, t \in \mathbb{R} \right\}.$$

The **vector from** u to v, denoted by u_v, is the unit tangent vector to the arc $[u, v]$ at the point u. By the Right Hand Rule, applied twice, we see that u_v is the unit vector in the direction of the vector $(u \times v) \times u$. Note that since $(u \times v) \cdot u = 0$ so that $\theta(u \times v, u) = \frac{\pi}{2}$, we have $\left|(u \times v) \times u\right| = |u \times v||u| \sin \theta(u \times v, u) = |u \times v|$, and so

$$u_v = \frac{(u \times v) \times u}{|u \times v|} = \frac{v - (u \cdot v)u}{|u \times v|}.$$

Given $u, v, w \in \mathbb{S}^2$ with $v \neq \pm u$ and $w \neq \pm u$, we define the **oriented angle** $\angle_o vuw$ to be

$$\angle_o vuw = \theta_o(u_v, u_w).$$

The (unoriented) **angle** $\angle vuw$ is given by $\angle vuw = \theta(u_v, u_w)$. We have $\angle vuw = \angle_o vuw$ when $\angle_o vuw \in [0, \pi]$, and $\angle vuw = 2\pi - \angle_o vuw$ when $\angle_o vuw \in [\pi, 2\pi)$.

2.15 Theorem: *(Angle Formula) Let $u, v, w \in \mathbb{S}^2$ with $v \neq \pm u$ and $w \neq \pm u$, and let $\alpha_o = \angle_o vuw$. Then*

$$\cos \alpha_o = \frac{(u \times v) \cdot (u \times w)}{|u \times v||u \times w|} = \frac{(v \cdot w) - (u \cdot v)(u \cdot w)}{|u \times v||u \times w|} \text{ and}$$
$$\sin \alpha_o = \frac{\det(u, v, w)}{|u \times v||u \times w|}.$$

Proof: We have

$$\begin{aligned}\cos \alpha_o = \cos \theta_o(u_v, u_w) = u_v \cdot u_w &= \frac{\big(v - (u \cdot v)u\big) \cdot \big(w - (u \cdot w)u\big)}{|u \times v||u \times w|} \\ &= \frac{(v \cdot w) - (u \cdot v)(u \cdot w)}{|u \times v||u \times w|} = \frac{(u \times v) \cdot (u \times w)}{|u \times v||u \times w|}\end{aligned}$$

and

$$\begin{aligned}\sin \alpha_o = \sin \theta_o(u_v, u_w) = \det(u, u_v, u_w) &= \frac{u \cdot \Big(\big(v - (u \cdot v)u\big) \times \big(w - (u \cdot w)u\big) \Big)}{|u \times v||u \times w|} \\ &= \frac{u \cdot \big(v \times w - (u \cdot w)\, v \times u - (u \cdot v)\, u \times w\big)}{|u \times v||u \times w|} = \frac{\det(u, v, w)}{|u \times v||u \times w|}.\end{aligned}$$

Spherical Triangles

2.16 Definition: A (non-degenerate spherical) **triangle** in $\mathbb{S}^2$ is determined by three non-colinear points $u, v, w \in \mathbb{S}^2$, which we call the **vertices** of the triangle. Requiring that u, v and w are non-colinear in $\mathbb{S}^2$ is equivalent to requiring that u, v and w do not all lie on the same plane through the origin in $\mathbb{R}^3$, or equivalently that $\{u, v, w\}$ is linearly independent, or equivalently that $\det(u, v, w) \neq 0$.

We can think of the triangle with vertices u, v and w in several different ways. For example, we can think of the triangle simply as the set of three points $\{u, v, w\}$, or if we wish we can keep track of the the order of the points and think of the triangle as the ordered triple (u, v, w). Alternatively, since u, v and w are non-colinear, no two of the three points are antipodal and so the **edges** $[u, v]$, $[v, w]$ and $[w, u]$ of the triangle are well-defined and, if we want, we can think of the triangle as the union of its three edges. As another alternative, we can think of the triangle as including its interior points, that is we can consider the triangle to be the **solid triangle**

$$[u, v, w] = \big\{x \in \mathbb{S}^2 \big| x = ru + sv + tw \text{ for some } 0 \leq r, s, t \in \mathbb{R}\big\}.$$

An **ordered triangle** in $\mathbb{S}^2$ consists of an ordered triple (u, v, w) of non-colinear points $u, v, w \in \mathbb{S}^2$, together with the set $[u, v, w]$. When $\det(u, v, w) > 0$, so that $\{u, v, w\}$ is a positively oriented basis for $\mathbb{R}^3$, we say that the triangle $[u, v, w]$ is **positively oriented**, and when $\det(u, v, w) < 0$ we say the triangle $[u, v, w]$ is **negatively oriented**.

2.17 Definition: Given an ordered triangle $[u, v, w]$ in $\mathbb{S}^2$, we shall let a, b and c denote the lengths of the edges $[v, w]$, $[w, u]$ and $[u, v]$ respectively, so that

$$a = d_S(v, w)\ ,\ b = d_S(w, u) \ \text{ and } \ c = d_S(u, v)\,,$$

and we shall let α_o, β_o and γ_o, and α, β and γ be the oriented and unoriented angles

$$\alpha_o = \angle_o vuw\ ,\ \beta_o = \angle_o wvu\ ,\ \gamma_o = \angle_o uwv\ ,\ \alpha = \angle vuw\ ,\ \beta = \angle wvu\ ,\ \gamma = \angle uwv.$$

When $[u, v, w]$ is positively oriented, the angles α_o, β_o and γ_o all lie in the interval $(0, \pi)$ so we have $\alpha = \alpha_0$, $\beta = \beta_o$ and $\gamma = \gamma_o$. When $[u, v, w]$ is negatively oriented, the angles α_o, β_o and γ_o all lie in the interval $(\pi, 2\pi)$, so we have $\alpha = 2\pi - \alpha_o$, $\beta = 2\pi - \beta_o$ and $\gamma = 2\pi - \gamma_o$. In either case, the angles α, β and γ are called the **interior angles** of the triangle, and the angles $2\pi - \alpha$, $2\pi - \beta$ and $2\pi - \gamma$ are called the **exterior angles**.

2.18 Theorem: *(Area of Spherical Triangles) Let $[u, v, w]$ be a positively oriented triangle in $\mathbb{S}^2$ with interior angles α, β and γ. Then the area of $[u, v, w]$ is equal to $(\alpha + \beta + \gamma) - \pi$.*

Proof: Let H_α be the hemisphere which contains u whose boundary is the line through v and w, and let $-H_\alpha$ be the antipodal hemisphere. Define H_β, $-H_\beta$, H_γ and $-H_\gamma$ similarly. Let $W_\alpha = \big(H_\beta \cap H_\gamma\big) \cup \big(- H_\beta \cap -H_\gamma\big)$, and define W_β and W_γ similarly. By looking at the sphere with the vector u pointing towards us, we can see that the double wedge W_α covers $\frac{\alpha}{\pi}$ of the entire sphere and so its area is $A_\alpha = \frac{\alpha}{\pi} \cdot 4\pi = 4\alpha$. Similarly the areas of the double wedges W_β and W_γ are equal to $A_\beta = 4\beta$ and $A_\gamma = 4\gamma$. Notice that when we shade each of the double wedges W_α, W_β and W_γ, the triangle $[u, v, w]$ and its antipodal triangle $[-u, -v, -w]$ are each shaded three times while the rest of the sphere is shaded once. It follows that if we let $S = 4\pi$ be the area of the entire sphere and we let T be the area of the triangle $[u, v, w]$ (which is equal to the area of the antipodal triangle $[-u, -v, -w]$) then we have $A_\alpha + A_\beta + A_\gamma = S + 4T$, that is $4\alpha + 4\beta + 4\gamma = 4\pi + 4T$, hence $T = (\alpha + \beta + \gamma) - \pi$.

2.19 Definition: For an ordered triangle $[u,v,w]$, the **polar triangle** of $[u,v,w]$ is the ordered triangle $[u',v',w']$ where

$$u'=\frac{v\times w}{|v\times w|}\ ,\ v'=\frac{w\times u}{|w\times u|}\ \text{ and } w'=\frac{u\times v}{|u\times v|}\,.$$

We denote the side lengths and the angles of the polar triangle $[u',v',w']$ by a', b', c' and α_o', β_o', γ_o', and α', β' and γ'.

2.20 Theorem: *(The Polar Triangle) Let $[u,v,w]$ be a positively oriented triangle in $\mathbb{S}^2$ with polar triangle $[u',v',w']$. Then*
(1) $[u'',v'',w'']=[u,v,w]$,
(2) $[u',v',w']$ is positively oriented,
(3) $a'=\pi-\alpha$, $b'=\pi-\beta$ and $c'=\pi-\gamma$, and
(4) $\alpha'=\pi-a$, $\beta'=\pi-b$ and $\gamma'=\pi-c$.

Proof: We have

$$v'\times w'=\frac{(w\times u)\times(u\times v)}{|w\times u||u\times v|}=\frac{w\cdot(u\times v)\,u-u\cdot(u\times v)\,w}{|w\times u||u\times v|}=\frac{\det(u,v,w)\,u}{|w\times u||u\times v|}.$$

This is a positive multiple of the vector u and so

$$u''=\frac{v'\times w'}{|v'\times w'|}=u.$$

Similarly, we have $v''=v$ and $w''=w$. This shows that $[u'',v'',w'']=[u,v,w]$.

Now let us determine the orientation of $[u',v',w']$. We have

$$\det(u',v',w')=u'\cdot(v'\times w')=\frac{v\times w}{|v\times w|}\cdot\frac{\det(u,v,w)\,u}{|w\times u||u\times v|}=\frac{\det(u,v,w)^2}{|u\times v||v\times w||w\times u|}$$

which is positive. This shows that $[u',v',w']$ is positively oriented.

Next, let us find the lengths of the edges of the polar triangle. We have

$$\cos a'=v'\cdot w'=\frac{(w\times u)\cdot(u\times v)}{|w\times u||u\times v|}=-\cos\alpha.$$

It follows that $a'=\pi\pm\alpha$. Since $[u,v,w]$ is positively oriented, we have $\alpha\in(0,\pi)$ and so we must have $a'=\pi-\alpha$. Similarly, we have $b'=\pi-\beta$ and $c'=\pi-\gamma$.

Finally, let us calculate the angles of the polar triangle. We have

$$\cos\alpha'=\frac{(u'\times v')\cdot(u'\times w')}{|u'\times v'||u'\times w'|}=w''\cdot(-v'')=-w\cdot v=-\cos a.$$

It follows that $\alpha'=\pi\pm a$. But since $[u',v',w']$ is positively oriented so that $\alpha'\in(0,\pi)$, we must have $\alpha'=\pi-a$. Similarly, $\beta'=\pi-b$ and $\gamma'=\pi-c$.

2.21 Exercise: Determine how Parts 1, 2, 3 and 4 of the above theorem must be modified when $[u,v,w]$ is a negatively oriented triangle in $\mathbb{S}^2$.

2.22 Theorem: *(The Sine Law) For any ordered triangle* $[u,v,w]$ *we have*

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

Proof: We have

$$\sin\alpha=\frac{|\det(u,v,w)|}{|u\times v||u\times w|}=\frac{|\det(u,v,w)|}{\sin c\sin b}$$

and similar formulas hold for $\sin\beta$ and $\sin\gamma$, so we obtain

$$|\det(u,v,w)|=\sin\alpha\sin b\sin c=\sin a\sin\beta\sin c=\sin a\sin b\sin\gamma.$$

2.23 Theorem: *(The First Law of Cosines) For any ordered triangle* $[u,v,w]$ *we have*

$$\cos\alpha=\frac{\cos a-\cos b\cos c}{\sin b\sin c}\;,\;\cos\beta=\frac{\cos b-\cos a\cos c}{\sin a\sin c}\;\textit{ and }\;\cos\gamma=\frac{\cos c-\cos a\cos b}{\sin a\sin b}\;.$$

Proof: We have

$$\cos\alpha=\frac{(u\times v)\cdot(u\times w)}{|u\times v||u\times w|}=\frac{(v\cdot w)-(u\cdot w)(v\cdot u)}{|u\times v||u\times w|}=\frac{\cos a-\cos b\cos c}{\sin c\sin b}.$$

The other two formulas may be proven similarly.

2.24 Corollary: *(Side-Side-Side and Side-Angle-Side)*

(1) If we know the lengths of the three sides of a triangle then we can find the three angles.
(2) If we know the lengths of two sides and the angle at the common vertex, then we can find the length of the third side (hence also the other two angles).

2.25 Theorem: *(The Second Law of Cosines) For any ordered triangle* $[u,v,w]$ *we have*

$$\cos a=\frac{\cos\alpha+\cos\beta\cos\gamma}{\sin\beta\sin\gamma}\;,\;\cos b=\frac{\cos\beta+\cos\gamma\cos\alpha}{\sin\gamma\sin\alpha}\;\textit{ and }\;\cos c=\frac{\cos\gamma+\cos\alpha\cos\beta}{\sin\alpha\sin\beta}.$$

Proof: Suppose that $[u,v,w]$ is positively orented, and let $[u',v',w']$ be its polar triangle. Then we have $a'=\pi-\alpha$ and $\alpha'=\pi-a$ and so on. We apply the First Law of Cosines to the polar triangle to get

$$\cos a=-\cos\alpha'=\frac{-\cos a'+\cos b'\cos c'}{\sin b'\sin c'}=\frac{\cos\alpha+\cos\beta\cos\gamma}{\sin\beta\sin\gamma}.$$

The case in which $[u,v,w]$ is negatively oriented is left as an exercise.

2.26 Corollary: *(Angle-Angle-Angle and Angle-Side-Angle)*

(1) If we know the three angles of a triangle then we can find the lengths of the three sides.
(2) If we know the length of one edge of a triangle and the angles at either end of the edge, then we can find the third angle (hence also the lengths of the other two sides).

Isometries

2.27 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 $A \in O_n(\mathbb{R})$.

2.28 Definition: An **isometry** on $\mathbb{S}^2$ is a bijective map $F : \mathbb{S}^2 \to \mathbb{S}^2$ which preserves distance, that is such that for all $u, v \in \mathbb{S}^2$ we have $d_S\big(F(u), F(v)\big) = d_S(u, v)$.

2.29 Theorem: *(Algebraic Classification of Isometries) Every orthogonal map on $\mathbb{R}^3$ restricts to an isometry on $\mathbb{S}^2$, and every isometry on $\mathbb{S}^2$ extends to an orthogonal map on $\mathbb{R}^3$. Thus the group of isometries on $\mathbb{S}^2$ can be identified with the group $O_3(\mathbb{R})$.*

Proof: Let $A \in O_3(\mathbb{R})$ and define $F : \mathbb{R}^3 \to \mathbb{R}^3$ by $F(x) = Ax$. Note that F is bijective with inverse given by $F^{-1}(x) = A^{-1}x = A^T x$. Also note that F preserves Euclidean distance because for all $x, y \in \mathbb{R}^3$ we have

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

Since spherical distance is determined by Euclidean distance, it follows that

$$d_S\big(F(u), F(v)\big) = d_S(u, v) \ \text{ for all } \ u, v \in \mathbb{S}^2$$

and so F restricts to an isometry on $\mathbb{S}^2$.

Now suppose that $F : \mathbb{S}^2 \to \mathbb{S}^2$ is an isometry on $\mathbb{S}^2$. Since Euclidean distance is determined by spherical distance, we have $\big|F(u) - F(v)\big| = |u - v|$ for all $u, v \in \mathbb{S}^2$. By the Polarization Identity, it follows that for all $u, v \in \mathbb{S}^2$

$$\begin{aligned} F(u) \cdot F(v) &= \tfrac{1}{2}\Big(|F(u)|^2 + |F(v)|^2 - \big|F(u) - F(v)\big|^2\Big) = \tfrac{1}{2}\big(1^2 + 1^2 - |u-v|^2\big) \\ &= \tfrac{1}{2}\big(|u|^2 + |v|^2 - |u-v|^2\big) = u \cdot v \end{aligned}$$

and so the map F preserves the dot product between elements of $\mathbb{S}^2$. In particular, if we let $\{e_1, e_2, e_3\}$ be the standard basis for $\mathbb{R}^3$ then we have $F(e_k) \cdot F(e_l) = e_k \cdot e_l = \delta_{kl}$ for all k, l and so the set $\big\{F(e_1), F(e_2), F(e_3)\big\}$ is also an orthonormal basis for $\mathbb{R}^3$.

Given any element $v \in \mathbb{S}^2$, we can write v uniquely in the form $v = \sum_{k=1}^{3} c_k F(e_k)$, and then for each index l we have $v \cdot F(e_l) = \sum_{k=1}^{3} c_k F(e_k) \cdot F(e_l) = \sum_{k=1}^{3} c_k \delta_{kl} = c_l$. Thus for every $v \in \mathbb{S}^2$ we have $v = \sum_{k=1}^{3} (v \cdot F(e_k)) F(e_k)$. Replacing v by $F(u)$, we see that for all $u = (u_1, u_2, u_3) \in \mathbb{S}^2$ we have

$$F(u) = \sum_{k=1}^{3} \big(F(u) \cdot F(e_k)\big) F(e_k) = \sum_{k=1}^{3} (u \cdot e_k) F(e_k) = \sum_{k=1}^{3} u_k F(e_k).$$

Let A be the orthogonal matrix with columns $F(e_1)$, $F(e_2)$ and $F(e_3)$. Then for all $u \in \mathbb{S}^2$ we have $F(u) = \sum_{k=1}^{3} u_k F(e_k) = Au$. Thus the isometry $F : \mathbb{S}^2 \to \mathbb{S}^2$ extends to the orthogonal map $F : \mathbb{R}^3 \to \mathbb{R}^3$ given by $F(x) = Ax$ for all $x \in \mathbb{R}^3$.

2.30 Note: From now on, we shall not distinguish notationally between an orthogonal matrix $A \in O_3(\mathbb{R})$, the corresponding orthogonal map $A : \mathbb{R}^3 \to \mathbb{R}^3$ given by $A(x) = Ax$, and the corresponding isometry $A : \mathbb{S}^2 \to \mathbb{S}^2$ obtained by restricting the map A to $\mathbb{S}^2$.

2.31 Definition: The **inversion** (or **antipodal map** on $\mathbb{S}^2$ is the isometry $N : \mathbb{S}^2 \to \mathbb{S}^2$ given by $N(x) = -x$. Given $u \in \mathbb{S}^2$, the **reflection** on $\mathbb{S}^2$ in the line L_u is the isometry $F_u : \mathbb{S}^2 \to \mathbb{S}^2$ obtained by restricting the orthogonal reflection $F_u : \mathbb{R}^3 \to \mathbb{R}^3$ on $\mathbb{R}^3$ in the plane $x \cdot u = 0$. When $L = L_u$ we also write $F_u = F_L$. Given $u \in \mathbb{S}^2$ and $\theta \in \mathbb{R}$, the **rotation** on $\mathbb{S}^2$ about u by θ is the map $R_{u,\theta} : \mathbb{S}^2 \to \mathbb{S}^2$ obtained by restricting the rotation $R_{u,\theta} : \mathbb{R}^3 \to \mathbb{R}^3$ by the angle θ in the direction of the fingers of the right hand when the thumb is pointing in the direction of the vector u.

2.32 Note: As a matrix, the inversion is given by $N = -I$. Let us describe the maps F_u and $R_{u,\theta}$ as matrices. Given u, choose a unit vector v which is orthogonal to u and then let $w = u \times v$ so that $\{u, v, w\}$ is an orthonormal basis for $\mathbb{R}^3$. Then the rotation $R_{u,\theta}$ is given by $R_{u,\theta}(u) = u$, $R_{u,\theta}(v) = \cos\theta\, v + \sin\theta\, w$ and $R_{u,\theta}(w) = -\sin\theta\, v + \cos\theta\, w$. Thus, as a matrix, we have

$$R_{u,\theta} = PAP^T \quad \text{where} \quad P = (u, v, w) \text{ and } A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Similarly, the reflection F_u is given by $F_u(u) = -u$, $F_u(v) = v$ and $F_u(w) = w$ so that, as a matrix, we have

$$F_u = PAP^T \quad \text{where} \quad P = (u, v, w) \text{ and } A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Alternatively, using the orthogonal projection map given by $\mathrm{Proj}_u(x) = (x{\cdot}u)u$, we have

$$F_u(x) = x - 2\,\mathrm{Proj}_u(x) = x - 2(x{\cdot}u)u = x - 2u(x \cdot u) = x - 2uu^T x = \big(I - 2uu^T\big)x$$

and so, as a matrix, we have

$$F_u = I - 2uu^T.$$

2.33 Theorem: *(The Product of Two Reflections is a Rotation) Let $u \in \mathbb{S}^2$ and let $v, w \in T_u$. Then*

$$F_w F_v = R_{u,2\theta_o(v,w)}.$$

Proof: Let $x = u \times v$ so that $\{u, v, x\}$ is a positively oriented orthonormal basis for $\mathbb{R}^3$. Let $\theta_o = \theta_o(v, w)$ so that

$$w = \cos\theta_o\, v + \sin\theta_o\, x.$$

Then $F_v(u) = u - 2\,u \cdot v\, v = u$, $F_v(v) = v - 2\,v \cdot v\, v = -v$ and $F_v(x) = x - 2\,x \cdot v\, v = x$ and so

$$\begin{aligned}
F_w F_v(u) &= F_w(u) = u - 2\,u \cdot w\, w = u = R_{u,2\theta_o}(u)\,, \\
F_w F_v(v) &= F_w(-v) = -v + 2\,v \cdot w\, w = -v + 2\cos\theta_o(\cos\theta_o v + \sin\theta_o\, x) \\
&= (2\cos\theta_o - 1)v + 2\sin\theta_o \cos\theta_o\, x = \cos 2\theta_o\, v + \sin 2\theta_o\, x = R_{u,2\theta_o}(v), \text{ and} \\
F_w F_v(x) &= F_w(x) = x - 2\,x \cdot w\, w = x - 2\sin\theta_o(\cos\theta_o\, v + \sin\theta_o\, x) \\
&= -2\sin\theta_o \cos\theta_o\, v + (1 - 2\sin\theta_o)x = -\sin 2\theta_o\, v + \cos 2\theta_o\, x = R_{u,2\theta_o}(x).
\end{aligned}$$

2.34 Definition: Given two points $a, b \in \mathbb{R}^3$ with $a \neq b$, recall that the **perpendicular bisector** of a and b in $\mathbb{R}^3$ is the plane P in $\mathbb{R}^3$ which passes through the midpoint $\frac{a+b}{2}$ and is orthogonal to the vector $b-a$, that is the plane

$$P = \left\{ x \in \mathbb{R}^3 \middle| \left(x - \tfrac{a+b}{2}\right) \cdot (b-a) = 0 \right\}.$$

Note that (as in the proof of Theorem 1.60) for $x \in \mathbb{R}^3$ we have

$$x \in P \iff d_E(x,a) = d_E(x,b).$$

Let $u, v \in \mathbb{S}^2$ with $u \neq v$ and let P be the perpendicular bisector of u and v in $\mathbb{R}^3$. Note that since $d_E(0,u) = 1 = d_E(0,v)$, we have $0 \in P$. We define the **perpendicular bisector** of u and v in $\mathbb{S}^2$ to be the line $L = \mathbb{S}^2 \cap P$. Since P is the plane in $\mathbb{R}^3$ through 0 orthogonal to the vector $v-u$, it follows that $L = L_w$ where $w = \frac{v-u}{|v-u|}$. Note that since spherical distance is determined by Euclidean distance, for $x \in \mathbb{S}^2$ we have $x \in L \iff d_S(x,u) = d_S(x,v)$.

2.35 Lemma: *(Reflection in Perpendicular Bisector) Let $u, v \in \mathbb{S}^2$ with $u \neq v$, let P be the perpendicular bisector of u and v in $\mathbb{R}^3$, and let F_P be the orthogonal reflection in P. Then*

(1) $F_P(u) = v$ and $F_P(v) = u$, and
(2) $x \in P \iff F_P(x) = x$.

Proof: Let $w = \frac{v-u}{|v-u|}$. Note that the plane P has equation $x \cdot w = 0$ and the orthogonal reflection F_P is given by $F_P(x) = x - 2\, x{\cdot}w\, w$. We have

$$\begin{aligned} F_P(u) &= u - 2\, u{\cdot}w\, w = u - \frac{2\, u \cdot (v-u)}{|v-u|^2}(v-w) = u - \frac{2\big(u \cdot v - |u|^2\big)}{|u|^2 - 2u{\cdot}v + |v|^2}(v-u) \\ &= u - \frac{2(u{\cdot}v - 1)}{1 - 2u{\cdot}v + 1}(v-u) = u + (v-u) = v \end{aligned}$$

and similarly $F_P(v) = u$. Also, for $x \in \mathbb{R}^3$ we have

$$F_P(x) = x \iff x - 2\, x \cdot w\, w = x \iff 2\, x \cdot w\, w = 0 \iff x \cdot w = 0 \iff x \in P.$$

2.36 Theorem: *(Congruent Triangles and Isometries) Given two ordered triangles $[u,v,w]$ and $[u',v',w']$ with $a = a'$, $b = b'$ and $c = c'$, there exists a unique isometry $F : \mathbb{S}^2 \to \mathbb{S}^2$ with $F(u) = u'$, $F(v) = v'$ and $F(w) = w'$.*

Proof: First we note that if such an isometry exists then it is unique because $\{u,v,w\}$ is a basis for $\mathbb{R}^3$. We now construct the required isometry as a composite of reflections. If $u = u'$ then let $F_1 = I$ and if $u \neq u'$ the let F_1 be the orthogonal reflection in the perpendicular bisector of u and u' so that we have $F_1(u) = u'$. Let $u_1 = F_1(u) = u'$, $v_1 = F_1(v)$ and $w_1 = F_1(w)$. Note that if a_1, b_1 and c_1 are the edge lengths of triangle $[u_1, v_1, w_1]$ then we have $a_1 = a$, $b_1 = b$ and $c_1 = c$ since F_1 is an isometry. If $v_1 = v'$ then let $F_2 = I$ and if $v_1 \neq v'$ then let F_2 be the orthogonal reflection in the perpendicular bisector of v_1 and v' so that we have $F_2(v_1) = v'$. Note that in the case that $v_1 \neq v'$, the point $u_1 = u'$ lies on the perpendicular bisector of v_1 and v' because $|v_1 - u_1| = b = |v' - u'| = |v' - u_1|$, and so we have $F_2(u_1) = u_1 = u'$. Let $u_2 = F_2(u_1) = u'$, $v_2 = F_2(v_1) = v'$ and $w_2 = F_2(w_1)$. Note that if a_2, b_2 and c_2 are the edge lengths of triangle $[u_2, v_2, w_2]$ then we have $a_2 = a$, $b_2 = b$ and $c_2 = c$. If $w_2 = w'$ then let $F_3 = I$ and if $w_2 \neq w'$ then let F_3 be the orthogonal reflection in the perpendicular bisector of w_2 and w' so that $F_3(w_2) = w'$. As above, note that $F_3(u_2) = u_2 = u'$ and $F_3(v_2) = v_2 = v'$. Thus the isometry $F = F_3 F_2 F_1$ satisfies $F(u) = u'$, $F(v) = v'$ and $F(w) = w'$.

2.37 Definition: A **rotary reflection** on $\mathbb{S}^2$ is an isometry on $\mathbb{S}^2$ of the form $F_u R_{u,\theta}$ for some $u \in \mathbb{S}^2$ and $\theta \in \mathbb{R}$. A **rotary inversion** on $\mathbb{S}^2$ is an isometry on $\mathbb{S}^2$ of the form $N\,R_{u,\theta}$ for some $u \in \mathbb{S}^2$ and $\theta \in \mathbb{R}$.

2.38 Theorem: *Let $u \in \mathbb{S}^2$ and $\theta \in \mathbb{R}$. Then*
(1) $N\,R_{u,\theta} = -R_{u,\theta} = R_{u,\theta}N$, and
(2) $F_u R_{u,\theta} = R_{u,\theta}F_u = -R_{u,\theta+\pi}$.

Proof: Part 1 holds because $N\,R_{u,\theta} = (-I)R_{u,\theta} = -R_{u,\theta} = R_{u,\theta}(-I) = R_{u,\theta}N$. To prove Part 2, choose $v, w \in \mathbb{S}^2$ so that $\{u, v, w\}$ is a positively oriented orthonormal basis for $\mathbb{R}^3$. Then we have

$$\begin{aligned}
F_u R_{u,\theta}(u) &= F_u(u) = -u,\\
F_u R_{u,\theta}(v) &= F_u\big((\cos\theta)\,v + (\sin\theta)\,w\big) = (\cos\theta)\,v + (\sin\theta)\,w,\\
F_u R_{u,\theta}(w) &= F_u\big(-(\sin\theta)\,v + (\cos\theta)\,w\big) = -(\sin\theta)\,v + (\cos\theta)\,w,
\end{aligned}$$

and we have

$$\begin{aligned}
R_{u,\theta}F_u(u) &= R_{u,\theta}(-u) = -u,\\
R_{u,\theta}F_u(v) &= R_{u,\theta}(v) = (\cos\theta)\,v + (\sin\theta)\,w,\\
R_{u,\theta}F_u(w) &= R_{u,\theta}(w) = -(\sin\theta)\,v + (\cos\theta)\,w,
\end{aligned}$$

and we have

$$\begin{aligned}
-R_{u,\theta+\pi}(u) &= -u,\\
-R_{u,\theta+\pi}(v) &= -\big(\cos(\theta+\pi)\,v + \sin(\theta+\pi)\,w\big) = (\cos\theta)\,v + (\sin\theta)\,w,\\
-R_{u,\theta+\pi}(w) &= -\big(-\sin(\theta+\pi)\,v + \cos(\theta+\pi)\,w\big) = -(\sin\theta)\,v + (\cos\theta)\,w.
\end{aligned}$$

2.39 Theorem: *(The Geometric Classification of Isometries) Every isometry on $\mathbb{S}^2$ is either a rotation, a reflection, or a rotary inversion.*

Proof: Let S be an isometry on $\mathbb{S}^2$. By Theorem 2.36, S is the unique isometry which sends the standard basis vectors e_1, e_2 and e_3 to the points $S(e_1)$, $S(e_2)$ and $S(e_3)$, and the proof of that theorem shows that S is of the form $S = F_3F_2F_1$ where each F_k is either the identity or a reflection. Thus either S is the identity, or S is a reflection, or S is the product of two reflections, or S is the product of three reflections. By Theorem 2.33, the product of two reflections is a rotation (or the identity, when the two reflections are equal), so it suffices to consider the product of three reflections. Suppose that $S = F_wF_vF_u$ where $u, v, w \in \mathbb{S}^2$. If $u = \pm v$ then $F_vF_u = I$ so that $S = F_w$, which is a reflection. Suppose that $u \neq \pm v$, so we have $L_u \cap L_v = \{\pm p\}$ where $p = \frac{u\times v}{|u\times v|}$. By Theorem 2.33, $F_vF_u = R_{p,\theta}$ where we have $u, v \in T_p$ and $\theta = 2\theta_o(u,v)$, and so $S = F_wF_vF_u = F_wR_{p,\theta}$. If $w = \pm p$ then $S = F_wR_{p,\theta} = F_pR_{p,\theta}$, which is a rotary reflection. Suppose $w \neq \pm p$. Let $w' = w$, let $v' = \frac{p\times w}{|p\times w|} \in T_p$ (so that $L_{v'}$ is the unique line through p which is perpendicular to L_w), and let u' be the point in T_p such that $\theta = 2\theta_o(u', v')$ so that we have $R_{p,\theta} = F_{v'}F_{u'}$. Then $S = F_wR_{p,\theta} = F_{w'}F_{v'}F_{u'} = R_{q,\pi}F_{u'}$ where $q = \frac{v'\times w'}{|v'\times w'|}$ (so that $L_{v'} \cap L_{w'} = \{\pm q\}$). If $u' = \pm q$ then $S = R_{q,\pi}F_{u'} = R_{q,\pi}F_q$, which is a rotary reflection. Suppose that $u' \neq \pm q$. Let $u'' = u'$, let $w'' = \frac{q\times u'}{|q\times u'|}$ (so that $L_{w''}$ is the line through q perpendicular to $L_{u''}$) and let v'' be the vector in T_q such that $\pi = 2\theta_o(v'', w'')$ so that we have $R_{q,\pi} = F_{w''}F_{v''}$. Then $S = R_{q,\pi}F_{u'} = F_{w''}F_{v''}F_{u''}$. Since w'' is perpendicular to both u'' and v'' so that $u'', v'' \in T_{w''}$, we have $F_{v''}F_{u''} = R_{w'',\phi}$ where $\phi = 2\theta_o(u'', v'')$, and $S = F_{w''}R_{w'',\phi}$, which is a rotary reflection. In all cases, either S is the identity, or S is a reflection, or S is a rotary reflection, and every rotary reflection is a rotary inversion, by Theorem 2.38.

Projections

2.40 Definition: Let H be the upper hemisphere $H = \{(x,y,z)\in\mathbb{S}^2 \big| z\geq 0\}$ and let D be the disc $D=\{(u,v))\big|u^2+v^2\leq 1\}$. The **orthogonal projection** from H to D is the map $\phi : H\to D$ given by

$$(u,v)=\phi(x,y,z)=(x,y).$$

Note that this map is invertible and its inverse is the map $\psi : D\to H$ given by

$$(x,y,z)=\psi(u,v)=\big(u,v,\sqrt{1-u^2-v^2}\big).$$

2.41 Note: A curve in H given by $(x,y,z)=\alpha(t)$ for $a\leq t\leq b$ is mapped by the orthogonal projection ϕ to the curve in D given by $(u,v)=\beta(t)=\phi\big(\alpha(t)\big)$. If we are given a formula for $\beta(t)$, say $\beta(t)=\big(u(t),v(t)\big)$, then we can calculate the length L of the curve $(x,y,z)=\alpha(t)=\psi\big(\beta(t)\big)=\Big(u(t),v(t),\sqrt{1-u(t)^2-v(t)^2}\Big)$ using the formula

$$L=\int_{t=a}^{b}\big|\alpha'(t)\big|\,dt.$$

When $\beta(t)$ is given in polar coordinates by $\beta(t)=\big(r(t)\cos\theta(t),r(t)\sin\theta(t)\big)$, we can calculate the length L of the curve $(x,y,z)=\alpha(t)$ as follows.

$$\begin{aligned}
\alpha(t)&=\Big(r\cos\theta\,,\,r\sin\theta\,,\,\sqrt{1-r^2}\Big)\\
\alpha'(t)&=\Big(r'\cos\theta-r\sin\theta\cdot\theta'\,,\,r'\sin\theta+r\cos\theta\,\theta'\,,\,\tfrac{-r}{\sqrt{1-r^2}}\,r'\Big)\\
L&=\int_{t=a}^{b}\sqrt{(r'\cos\theta-r\sin\theta\,\theta')^2+(r'\sin\theta+r\cos\theta\,\theta')^2+\big(\tfrac{r}{\sqrt{1-r^2}}\,r'\big)^2}\,dt\\
&=\int_{t=a}^{b}\sqrt{(r')^2+r^2(\theta')^2+\tfrac{r^2}{1-r^2}(r')^2}\,dt\\
&=\int_{t=a}^{b}\sqrt{\tfrac{1}{1-r^2}(r')^2+r^2(\theta')^2}\,dt.
\end{aligned}$$

2.42 Note: When $R\subseteq\{(r,\theta)\big|0\leq r\leq 1\,,\,0\leq\theta\leq 2\pi\}$ and $\sigma : R\to H$, by using some vector calculus, one can show that the region in H which is given parametrically by $(x,y,z)=\sigma(r,\theta)=\Big(r\cos\theta\,,\,r\sin\theta\,,\,\sqrt{1-r^2}\Big)$ has area

$$\begin{aligned}
A&=\iint_R\big|\sigma_r\times\sigma_\theta\big|\,dr\,d\theta\\
&=\iint_R\Big|\big(\cos\theta\,,\,\sin\theta\,,\,\tfrac{-r}{\sqrt{1-r^2}}\big)\times\big(-r\sin\theta\,,\,r\cos\theta\,,\,0\big)\Big|\,dr\,d\theta\\
&=\iint_R\big(\tfrac{r^2\cos\theta}{\sqrt{1-r^2}}\,,\,\tfrac{r^2\sin\theta}{\sqrt{1-r^2}}\,,\,r\big)\Big|\,dr\,d\theta=\iint_R\sqrt{\tfrac{r^4}{1-r^2}+r^2}\,dr\,d\theta\\
&=\iint_R\tfrac{r}{\sqrt{1-r^2}}\,dr\,d\theta.
\end{aligned}$$

2.43 Remark: We can project orthogonally onto any plane through the origin. When $u\in\mathbb{S}^2$, $H=D(u,\frac{\pi}{2})=\{x\in\mathbb{S}^2\big|x\cdot u\geq 0\}$ and $D=\{x\in\mathbb{R}^3\big|x\cdot u=0,|x|\leq 1\}$, the **orthogonal projection** $\phi : H\to D$ and its inverse $\psi : D\to H$ are given by

$$y=\phi(x)=x-(x\cdot u)u\ \text{ and }\ x=\psi(y)=y+\sqrt{1-|y|^2}\,u\,.$$

2.44 Definition: Let $S = \mathbb{S}^2 \setminus \{\pm(0,0,1)\}$ and $R = \big\{(\theta, z)\big|0 \le \theta < 2\pi, -1 \le z \le 1\big\}$. The **Lambert cylindrical equal-area projection** is the map from S to R which is obtained by first projecting radially outwards from the z-axis to the cylinder $x^2 + y^2 = 1$ and then cutting the cylinder along the line $x = 1$, $y = 0$ and unrolling it into the rectangle R. The map $\phi : S \to R$ is given by

$$\phi\Big(\sqrt{1-z^2}\cos\theta\,,\ \sqrt{1-z^2}\sin\theta\,,\ z\Big) = (\theta, z)$$

and its inverse $\psi : R \to S$ is given by

$$\psi(\theta, z) = \Big(\sqrt{1-z^2}\cos\theta\,,\ \sqrt{1-z^2}\sin\theta\,,\ z\Big).$$

2.45 Theorem: *The Lambert cylindrical equal area projection preserves area.*

Proof: This can be seen to be a consequence of Theorem 2.4 (The Spherical Area Theorem), but we shall provide a proof which uses some vector calculus. The area A of a region $D \subseteq R = \big\{(r, \theta)\big|0 \le \theta < 2\pi, -1 \le z \le 1\big\}$ is given by

$$A = \iint_D 1\, d\theta\, dz.$$

The area of its inverse mage under ϕ is the area of its image under ψ which is

$$\begin{aligned}
B &= \iint_D \big|\psi_\theta \times \psi_z\big|\, d\theta\, dz \\
&= \iint_D \left|\Big(-\sqrt{1-z^2}\sin\theta\,,\ \sqrt{1-z^2}\cos\theta\,,\ 0\Big) \times \Big(\tfrac{-z}{\sqrt{1-z^2}}\cos\theta\,,\ \tfrac{-z}{\sqrt{1-z^2}}\sin\theta\,,\ 1\Big)\right| d\theta\, dz \\
&= \iint_D \Big|\big(\sqrt{1-z^2}\cos\theta\,,\ \sqrt{1-z^2}\sin\theta\,,\ z\big)\Big|\, d\theta\, dz \\
&= \iint_D \sqrt{(1-z^2)+z^2}\, d\theta\, dz = \iint_D 1\, d\theta\, dz = A.
\end{aligned}$$

2.46 Remark: We can obtain an alternate equal-area projection by composing ϕ with a map that scales the rectangle R by a scaling factor of c in the θ direction and by $\frac{1}{c}$ in the z-direction. We can also choose to project radially outwards from any line through the origin to the cylinder centred along that line (we can choose a line other than the z-axis).

2.47 Definition: Let H be the open hemisphere $H = \big\{(x, y, z) \in \mathbb{S}^2\big|z > 0\big\}$. The **gnomic projection** from H to $\mathbb{R}^2$ is the map $\phi : H \to \mathbb{R}^2$ obtained by projecting radially outwards from the origin to the plane $z = 1$ which we identify with $\mathbb{R}^2$. The line through $(0, 0, 0)$ and (x, y, z) meets the plane $z = 1$ at the point $\big(\frac{x}{z}, \frac{y}{z}, 1\big)$, and so the map ϕ is given by

$$(u, v) = \phi(x, y, z) = \Big(\frac{x}{z}\,,\ \frac{y}{z}\Big).$$

Its inverse is the map $\psi : \mathbb{R}^2 \to H$ given by

$$(x, y, z) = \psi(u, v) = \frac{(u, v, 1)}{\big|(u, v, 1)\big|} = \Big(\frac{u}{\sqrt{u^2+v^2+1}}\,,\ \frac{v}{\sqrt{u^2+v^2+1}}\,,\ \frac{1}{\sqrt{u^2+v^2+1}}\Big).$$

2.48 Remark: We could, if we wanted, also define gnomic projections from other open hemispheres, for example from the hemisphere $\big\{(x, y, z) \in \mathbb{S}^2\,\big|\,x > 0\big\}$ or from the hemisphere $\big\{(x, y, z) \in \mathbb{S}^2\,\big|\,x < 0\big\}$.

2.49 Theorem: *The gnomic projection maps great circles on $\mathbb{S}^2$, intersected with the open upper hemisphere H, to lines in $\mathbb{R}^2$.*

Proof: When L is a line in $\mathbb{S}^2$ with pole not equal to $\pm(0,0,1)$, and P is the in $\mathbb{R}^3$ which contains L so that $L=\mathbb{S}^2\cap P$, and M is the intersection of P with the plane $z=1$, it is easy to see with the help of a picture that ϕ maps $L\cap H$ to M. Here is an analytic proof. Let $L=\mathbb{S}^2\cap P$ where P is the plane $ax+by+cz=0$ with $c\neq 0$. The line of intersection of P with the plane $z=1$ is the line given by $ax+by+c=0$, $z=1$. We show that ϕ maps $L\cap H$ to the line M in the uv-plane with equation $au+bv+c=0$. Let $(x,y,z)\in L\cap H$. Then we have $ax+by+cz=0$ and $z>0$. For $(u,v)=\phi(x,y,z)$ we have

$$au+bv+c=a\,\tfrac{x}{z}+b\,\tfrac{y}{z}+c=\tfrac{ax+by+cz}{z}=\tfrac{0}{z}=0$$

and so the point (u,v) lies on the line M. Conversely, let $(u,v)\in M$ so that $au+bv+c=0$ and let $(x,y,z)=\psi(u,v)$. Then $z=\frac{1}{u^2+v^2+1}>0$ so that $(x,y,z)\in H$ and we have

$$ax+by+cz=\tfrac{au}{u^2+v^2+1}+\tfrac{bv}{u^2+v^2+1}+\tfrac{c}{u^2+v^2+1}=\tfrac{au+bv+c}{u^2+v^2+1}=0$$

so the point $(x,y,z)\in L$.

2.50 Definition: Let $S=\mathbb{S}^2\setminus\{(0,0,1)\}$. The **stereographic projection** from S to $\mathbb{R}^2$ is the map $\phi:S\to\mathbb{R}^2$ which sends the point (x,y,z) on the sphere to the point of intersection (u,v) of the line through $(0,0,1)$ and (x,y,z) with the plane $z=0$. Given $(x,y,z)\in S$, the line through $(0,0,1)$ and (x,y,z) is given by

$$(u,v,w)=\alpha(t)=(0,0,1)+t(x,y,z-1)=\big(tx\,,\,ty\,,\,1+t(z-1)\big)\text{ for }t\in\mathbb{R}.$$

The point of intersection of this line with the plane $w=0$ occurs when $1+t(z-1)=0$, that is when $t=1/(1-z)$. The point of intersection is $\alpha\big(\frac{1}{1-z}\big)=\big(\frac{x}{1-z},\frac{y}{1-z},0\big)$, so the map ϕ is given by

$$(u,v)=\phi(x,y,z)=\left(\frac{x}{1-z},\frac{y}{1-z}\right).$$

Given $(u,v)\in\mathbb{R}^2$, the line through $(0,0,1)$ and (u,v) is given by

$$(x,y,z)=\beta(t)=(0,0,1)+t(u,v,-1)=(tu,tv,1-t)\text{ for }t\in\mathbb{R}.$$

The point of intersection with the unit sphere occurs when $|\beta(t)|=1$, that is when we have $(tu)^2+(tv)^2+(1-t)^2=1$, that is $t^2u^2+t^2v^2-2t+t^2=0$, or $t\,(tu^2+tv^2+t-2)=0$. The intersection occurs when $t=\frac{2}{u^2+v^2+1}$, and so the inverse $\psi:\mathbb{R}^2\to S$ is given by

$$(x,y,z)=\psi(u,v)=\left(\frac{2u}{u^2+v^2+1},\frac{2v}{u^2+v^2+1},\frac{u^2+v^2-1}{u^2+v^2+1}\right).$$

2.51 Note: Let $f:U\subseteq\mathbb{R}^n\to\mathbb{R}^m$ be differentiable at $a\in\mathbb{R}^n$ and let $w\in\mathbb{R}^m$. Chose a differentiable map $\alpha(t)$ with $\alpha(0)=a$ and $\alpha'(0)=w$ and let $\beta(t)=f\big(\alpha(t)\big)$. The by the Chain Rule we have $\beta'(t)=Df\big(\beta(t)\big)\beta'(t)$ so in particular $\beta'(0)=Df\big(\alpha(0)\big)\alpha'(0)=Df(a)w$. Thus the map f sends the vector w to the vector $Df(a)w$.

2.52 Definition: Let $f:U\subseteq\mathbb{R}^n\to\mathbb{R}^m$ be differentiable at $a\in U$ and let $c>0$. We say that f is a **local scaling** near the point a of **scaling factor** c when the columns of $Df(a)$ are orthogonal and of length c, in other words when $Df(a)^TDf(a)=c^2I$. We say that f is **conformal** at a when f preserves the angles between vectors at a, that is when

$$\frac{\big(Df(a)w_1\big)\cdot\big(Df(a)w_2\big)}{\big|Df(a)w_1\big|\big|Df(a)w_2\big|}=\frac{w_1\cdot w_2}{|w_1||w_2|}\text{ for all vectors }0\neq w_1,w_2\in\mathbb{R}^n.$$

2.53 Note: When $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is a local scaling near $a \in U$ of scaling factor $c > 0$, f is conformal at a because for $0 \neq w_1, w_2 \in \mathbb{R}^n$ we have

$$\big(Df(a)w_1\big) \boldsymbol{\cdot} \big(Df(a)w_2\big) = \big(Df(a)w_2\big)^T \big(Df(a)w_1\big) = w_2^T(c^2 I)w_1 = c^2 w_1 \boldsymbol{\cdot} w_2$$

hence $\big|Df(a)w_i\big|^2 = \big(Df(a)w_i\big) \boldsymbol{\cdot} \big(Df(a)w_i\big) = c^2|w_i|^2$ so that $\big|Df(a)w_i\big| = c|w_i|$ for $i = 1, 2$, and so

$$\frac{\big(Df(a)w_1\big) \boldsymbol{\cdot} \big(Df(a)w_2\big)}{\big|Df(a)w_1\big|\big|Df(a)w_2\big|} = \frac{c^2\, w_1 \boldsymbol{\cdot} w_2}{c|w_1| \cdot c|w_2|} = \frac{w_1 \boldsymbol{\cdot} w_2}{|w_1||w_2|}.$$

2.54 Theorem: *The inverse stereographic projection map is a local scaling near (u, v) of scaling factor $c = \frac{2}{u^2+v^2+1}$.*

Proof: We have

$$D\psi = \begin{pmatrix} x_u & x_v \\ y_u & y_v \\ z_u & z_v \end{pmatrix} = \frac{2}{(u^2+v^2+1)^2} \begin{pmatrix} -u^2+v^2+1 & -2uv \\ -2uv & u^2-v^2+1 \\ 2u & 2v \end{pmatrix}$$

and a quick calculation yields

$$(D\psi)^T(D\psi) = \frac{4}{(u^2+v^2+1)^2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

2.55 Theorem: *The stereographic projection map circles through $(0, 0, 1)$ in $\mathbb{S}^2$, with the point $(0, 0, 1)$ removed, to lines in $\mathbb{R}^2$, and it sends circles not through $(0, 0, 1)$ in $\mathbb{S}^2$ to circles in $\mathbb{R}^2$.*

Proof: Let $n = (0, 0, 1)$. We leave the first part of the theorem as an exercise, and we show that the image under ϕ of each circle in $\mathbb{S}^2 \setminus \{n\}$ is a circle in $\mathbb{R}^2$. Let C be a circle on $\mathbb{S}^2 \setminus \{n\}$. Say C is the intersection of $\mathbb{S}^2$ with the plane P given by $ax + by + cz + d = 0$. Since $n \notin C$ we have $n \notin P$ and so $c + d \neq 0$. Since the distance from P to the origin is less than 1 we have $\text{dist}(0, P)^2 = \frac{d^2}{a^2+b^2+c^2} < 1$. For $(u, v) \in \mathbb{R}^2$ we have

$$\begin{aligned}
(u,v) \in \phi(C) \iff \psi(u,v) \in P \iff & \Big(\frac{2u}{u^2+v^2+1}, \frac{2v}{u^2+v^2+1}, \frac{u^2+v^2-1}{u^2+v^2-1}\Big) \in P \\
\iff & a \cdot \frac{2u}{u^2+v^2+1} + b \cdot \frac{2v}{u^2+v^+1} + c \cdot \frac{u^2+v^2-1}{u^2+v^2+1} + d = 0 \\
\iff & 2a\,u + 2b\,v + c(u^2+v^2-1) + d(u^2+v^2+1) = 0 \\
\iff & (c+d)u^2 + 2a\,u + (c+d)v^2 + 2b\,v = c - d \\
\iff & u^2 + \tfrac{2a}{c+d}\,u + v^2 + \tfrac{2b}{c+d}\,v = \tfrac{c-d}{c+d} \\
\iff & \big(u + \tfrac{a}{c+d}\big)^2 + \big(v + \tfrac{b}{c+d}\big)^2 = \tfrac{c-d}{c+d} + \tfrac{a^2}{(c+d)^2} + \tfrac{b^2}{(c+d)^2} = \tfrac{a^2+b^2+c^2-d^2}{(c+d)^2} \\
\iff & (u,v) \text{ lies on the circle centred at } \big(\tfrac{-a}{c+d}, \tfrac{-b}{c+d}\big) \text{ of radius } r = \tfrac{\sqrt{a^2+b^2+c^2-d^2}}{|c+d|}.
\end{aligned}$$

2.56 Remark: We defined the stereographic projection from $\mathbb{S}^2 \setminus \{n\}$ to $\mathbb{R}^2$ where n is the north pole $n = (0, 0, 1)$, but we could, if we wanted, also define the stereographic projection from $\mathbb{S}^2 \setminus \{u\}$ to $\mathbb{R}^2$ for any point $u \in \mathbb{S}^2$.

Chapter 3. Projective Geometry

The Projective Plane

3.1 Note: The rudiments of projective geometry were first studied by artists who were interested in perspective drawing. Imagine drawing a picture of a scene as follows. Set up a system of coordinates with the artist's eye at the origin, the z-axis pointing upwards, so that the artist is looking in the direction of the y-axis. Erect a plane of glass along the plane $y = 1$. All points in space which lie along the same ray through the origin will appear to be at the same position relative to the artist's eye. An object at position (x, y, z) with $y > 0$ appears to be at the same position as the point at position $\left(\frac{x}{y}, 1, \frac{z}{y}\right)$ which lies on the pane of glass so the artist draws a spot at this point on the plane of glass.

3.2 Example: An artist views a railway track which lies in the plane $z = -1$ with rails along the lines $x = \pm 1$ and ties along the line segments $-1 \leq x \leq 1$, $y = k$, $k \in \mathbb{Z}^+$. The artist sets up a pane of glass along the plane $y = 1$. The endpoints of the ties, which are at $(x, y, z) = (\pm 1, k, -1)$, will be represented by spots at position $(x, z) = (\pm\frac{1}{k}, -\frac{1}{k})$, $y = 1$. On the pane of glass, the rails are represented by the rays $z = \pm x$, $z < 0$ and the ties are represented by the line segments from $(x, z) = \left(-\frac{1}{k}, -\frac{1}{k}\right)$ to $(x, z) = \left(\frac{1}{k}, -\frac{1}{k}\right)$.

3.3 Note: From the point of view of producing a perspective drawing, all points which lie along the same ray from the origin can be identified as being the same point. Each ray through the origin intersects a unique point on the sphere, so we can identify the set of all such rays with the sphere $\mathbb{S}^2$. The rays which are visible to an artist looking in the direction of the positive y-axis are then identified with the hemisphere $H = \left\{(x, y, z) \in \mathbb{S}^2 \middle| y > 0\right\}$, Notice that the map which sends a ray, or a point in H, to the corresponding point drawn by the artist on the pane of glass at $y = 1$ is the gnomic projection $\phi(x, y, z) = \left(\frac{x}{y}, \frac{z}{y}\right)$.

In projective geometry, rather than considering the set of all *rays* through the origin, we consider instead the set of all *lines* through the origin.

3.4 Definition: The real **projective plane**, denoted by $\mathbb{P}^2$, is the set of lines through the origin in $\mathbb{R}^3$. Given $0 \neq x \in \mathbb{R}^3$ we let $[x]$ denote the line in $\mathbb{R}^3$ through 0 and x, that is $[x] = \text{Span}\{x\}$, so we have

$$\mathbb{P}^2 = \left\{[x] \middle| 0 \neq x \in \mathbb{R}^3\right\}.$$

A line through the origin in $\mathbb{R}^3$ is called a **point** in $\mathbb{P}^2$. Given two lines u and v through the origin in $\mathbb{R}^3$, we define the (projective) **distance** between the points u and v in $\mathbb{P}^2$, denoted by $d_P(u, v)$, to be the angle between the lines u and v in $\mathbb{R}^3$. When $u = [x]$ and $v = [y]$ with $0 \neq x, y \in \mathbb{R}^3$, we have

$$\begin{aligned} d_P(u, v) &= \min\left\{\theta(x, y)\,,\, \theta(x, -y)\right\} = \min\left\{\theta(x, y)\,,\, \pi - \theta(x, y)\right\} \\ &= \cos^{-1}\frac{|x \cdot y|}{|x||y|} = \sin^{-1}\frac{|x \times y|}{|x||y|}. \end{aligned}$$

3.5 Note: A line u through the origin in $\mathbb{R}^3$ intersects the sphere $\mathbb{S}^2$ in two antipodal points $\pm x$, and these points determine the line, indeed $u = [x] = [-x]$. We often identify the line u with the pair of antipodal points $\pm x$ and consider $\mathbb{P}^2$ to be the set of all pairs of antipodal points in $\mathbb{S}^2$, that is

$$\mathbb{P}^2 = \left\{\{\pm x\} \middle| x \in \mathbb{S}^2\right\}.$$

3.6 Theorem: *(Metric Properties of Distance) Let $u, v, w \in \mathbb{P}^2$. Then*

(1) (Positive Definiteness) $d_P(u,v) \in \left[0, \frac{\pi}{2}\right]$ with $d_P(u,v) = 0$ if and only if $u = v$,
(2) (Symmetry) $d(u,v) = d_P(v,u)$, and
(3) (Triangle Inequality) $d_P(u,v) + d_P(v,w) \geq d_P(u,w)$.

Proof: We prove Part (3). Choose $x, y, z \in \mathbb{S}^2$ so that $u = [x]$, $v = [y]$ and $w = [z]$ and also (by replacing x and z by $\pm x$ and $\pm z$ if necessary) such that $x \cdot y \geq 0$ and $y \cdot z \geq 0$. Then

$$\begin{aligned}
\cos\Big(d_P(u,v), d_P(v,w)\Big) &= \cos d_P(u,v) \cos d_P(v,w) - \sin d_P(u,v) \sin d_P(v,w) \\
&= (x \cdot y)(y \cdot z) - |x \times y||y \times z| \text{ , since } x \cdot y \geq 0 \text{ and } y \cdot z \geq 0, \\
&\leq (x \cdot y)(y \cdot z) - \big|(x \times y) \cdot (y \times z)\big| \text{ , by the Cauchy-Schwearz Inequality} \\
&\leq (x \cdot y)(y \cdot z) - (x \times y) \cdot (y \times z) \\
&= (x \cdot y)(y \cdot z) - (x \cdot y)(y \cdot z) + (x \cdot z) \\
&= (x \cdot z) \leq |x \cdot z| = \cos d_P(u,w)
\end{aligned}$$

3.7 Definition: For $u \in \mathbb{P}^2$ and $r \in \left[0, \frac{\pi}{2}\right]$, the (projective) **circle** centred at u of radius r and the (closed projective) disc centred at u of raduis r are the sets

$$\begin{aligned}
C(u,r) &= \left\{v \in \mathbb{P}^2 \middle| d_P(u,v) = r\right\} \text{ and} \\
D(u,r) &= \left\{v \in \mathbb{P}^2 \middle| d_P(u,v) \leq r\right\}.
\end{aligned}$$

3.8 Note: Let $a \in \mathbb{S}^2$ and let $u = [a]$. The union of the lines $v \in C(u,r) \subseteq \mathbb{P}^2$ forms a double cone in $\mathbb{R}^3$ with vertex at the origin, and this double cone intersects $\mathbb{S}^2$ in the pair of antipodal spherical circles $C(a,r)$ and $C(-a,r) = C(a, \pi - r)$. The circumference of the projective circle $C(u,r) \subseteq \mathbb{P}^2$ is equal to that of the spherical circle $C(a,r) \subseteq \mathbb{S}^2$, and the area of the projective disc $D(u,r) \subseteq \mathbb{P}^2$ is equal to that of the spherical disc $D(a,r) \subseteq \mathbb{S}^2$.

3.9 Definition: A (projective) **line** in $\mathbb{P}^2$ is the set of all lines through the origin in R^3 which lie in some given plane through the origin in $\mathbb{R}^3$. Note that a projective line $L \subseteq \mathbb{P}^2$ determines and is determined by a Euclidean plane $P \subseteq \mathbb{R}^3$ through the origin; given a plane P, the corresponding line L is given by $L = \left\{u \in \mathbb{P}^2 \middle| u \subseteq P\right\}$, and given a line L, the corresponding plane P is given by $P = \bigcup_{u \in L} u$. Given a projective line $L \subseteq \mathbb{P}^2$ and its corresponding Euclidean plane $P \subseteq \mathbb{R}^3$, the **pole** of L is the point $u \in \mathbb{P}^2$ which, as a line through 0 in $\mathbb{R}^3$, is perpendicular to the plane P. Given a point $u \in \mathbb{P}^2$, we write L_u to denote the projective line with pole u. We remark that a projective line is the same thing as a projective circle of radius $\frac{\pi}{2}$, indeed for $u \in \mathbb{P}^2$ we have $L_u = C\left(u, \frac{\pi}{2}\right)$.

3.10 Theorem: *(Properties of Projective Lines)*

(1) Given two distinct points $u, v \in \mathbb{P}^2$, there is a unique line $L \subseteq \mathbb{P}^2$ containing u and v.
(2) Given two distinct lines $L, M \subseteq \mathbb{P}^2$ there is a unique point $u \in \mathbb{P}^2$ with $u \in L \cap M$.
(3) Given a point $u \in \mathbb{P}^2$ and a line $L \subseteq \mathbb{P}^2$ with $L \neq L_u$, there is a unique line in $\mathbb{P}^2$ which passes through u and is perpendicular to L.
(4) Given two distinct lines $L, M \subseteq \mathbb{P}^2$ there exists a unique line in $\mathbb{P}^2$ which is perpendicular to L and M.

Proof: All parts of this theorem follow immediately from properties of lines and planes through the origin in $\mathbb{R}^3$.

3.11 Definition: A (projective) **triangle** is determined by a spherical triangle. Three non-colinear points $u, v, w \in \mathbb{S}^2$ determine an ordered spherical triangle $[u, v, w] \subseteq \mathbb{S}^2$. The corresponding solid projective triangle consists of all lines through 0 in $\mathbb{R}^3$ which pass through the solid spherical triangle $[u, v, w] \subseteq \mathbb{S}^2$. Note that the union of all the lines in the projective triangle forms a double cone, with triangular cross-section, which passes through the solid triangle $[u, v, w]$ and also the antipodal triangle $[-u, -v, -w]$. When $[u, v, w]$ is positively oriented, the area, angles, and side lengths of the projective triangle are the same as those of the spherical triangle $[u, v, w]$.

3.12 Note: We do not consider a projective triangle to have an orientation because an ordered spherical triangle $[u, v, w]$ and its antipodal triangle $[-u, -v, -w]$ each determine the same projective triangle, but these two spherical triangles have the opposite orientation.

3.13 Note: For a spherical triangle $[u, v, w]$, the edge lengths are the same as the distances between the vertices, for example $a = d_S(v, w)$, but this is not necessarily the case for the corresponding projective triangle. Indeed when $a = d_S(v, w) > \frac{\pi}{2}$ we find that $d_P\big([u], [v]\big) = d_S(u, -v) = \pi - d_S(u, v) = \pi - a$.

3.14 Definition: An **isometry** on $\mathbb{P}^2$ is a bijective map $F : \mathbb{P}^2 \to \mathbb{P}^2$ which preserves distance, that is such that for all $u, v \in \mathbb{P}^2$ we have $d_P\big(F(u), F(v)\big) = d_P(u, v)$. Note that every isometry on $\mathbb{S}^2$ determines an isometry on $\mathbb{P}^2$ as follows. Given an isometry $F : \mathbb{S}^2 \to \mathbb{S}^2$, extend F to the orthogonal map $F : \mathbb{R}^3 \to \mathbb{R}^3$ and note that $F(tx) = tF(x)$ for all $x \in \mathbb{R}^3$. We define the **induced isometry** $F : \mathbb{P}^2 \to \mathbb{P}^2$ by

$$F\big([x]\big) = \big[F(x)\big].$$

3.15 Theorem: *Every isometry on $\mathbb{P}^2$ is (induced by) a rotation $R_{p,\theta}$ for some $p \in \mathbb{S}^2$ and $\theta \in \mathbb{R}$. The set of isometries on $\mathbb{P}^2$ can be identified with*

$$SO(3, \mathbb{R}) = \big\{A \in M_3(\mathbb{R}) \big| A^T A = I \,,\, \det A = 1\big\}.$$

Proof: Let $F : \mathbb{P}^2 \to \mathbb{P}^2$ be an isometry on $\mathbb{P}^2$. Choose $u_1, u_2, u_3 \in \mathbb{S}^2$ so that we have $F([e_1]) = [u_1]$, $F([e_2]) = [u_2]$ and $F([e_3]) = [u_3]$. Then for all k, l we have

$$|u_k \cdot u_l| = \cos d_P\big([u_1], [u_2]\big) = \cos d_P\big(F([e_k]), F([e_l])\big) = \cos d_P\big([e_k], [e_l]\big) = |e_k \cdot e_l| = \delta_{k,l}.$$

When $k \neq l$ we have $|u_k \cdot u_l| = 0$ so that $u_k \cdot u_l = 0$, and when $k = l$ we have $|u_k \cdot u_l| = 1$ and $u_k \cdot u_l = u_k \cdot u_k \geq 0$ so that $u_k \cdot u_l = 1$. Thus $u_k \cdot u_l = \delta_{k,l}$ for all k, l so that $\{u_1, u_2, u_3\}$ is an orthonormal basis for $\mathbb{R}^3$. Let $x = (x_1, x_2, x_3) \in \mathbb{S}^2$ (so that $[x]$ is an arbitrary element in $\mathbb{P}^2$. Choose $y \in \mathbb{S}^2$ so that $f([x]) = [y]$. For each index k we have

$$|y \cdot u_k| = \cos d_P\big([y], [u_k]\big) = \cos d_P\big(F[x], F[e_k]\big) = \cos d_P\big([x], [e_k]\big) = |x_k \cdot e_k| = |x_k|.$$

Since $\{u_1, u_2, u_3\}$ is orthonormal, we have

$$y = \sum_{k=1}^{3} (y \cdot u_k) u_k = \sum_{k=1}^{3} \pm x_k u_k = Ax$$

where $A = A(x)$ is one of the 8 matrices $\big(\pm u_1, \pm u_2, \pm u_3\big)$. Thus we have

$$F([x]) = [y] = \big[Ax\big]$$

where $A = A(x)$ is one of the 4 matrices $\big(\pm u_1, \pm u_2, \pm u_3\big)$ with determinant equal to 1.

Finally we remark, without providing a rigorous proof, that every isometry is continuous and that the matrix $A = A(x)$ must be constant for all $x \in \mathbb{S}^2$, otherwise F would not be continuous.

Zero Sets of Polynomials

3.16 Definition: For $(x,y,z)\in\mathbb{R}^3$ with $(x,y,z)\neq(0,0,0)$, we write

$$[x,y,z]=\text{Span}\big\{(x,y,z)\big\}\in\mathbb{P}^2.$$

Let $U_1=\big\{[x,y,z]\big|x\neq 0\big\}$, $U_2=\big\{[x,y,z]\big|y\neq 0\big\}$ and $U_3=\big\{[x,y,z]\big|z\neq 0\big\}$. Note that $\mathbb{P}^2=U_1\cup U_2\cup U_3$. We define three **gnomic projections** $\phi_k:U_k\to\mathbb{R}^2$ by

$$\phi_1\big([x,y,z]\big)=\left(\tfrac{y}{x},\tfrac{z}{x}\right)\ ,\ \phi_2\big([x,y,z]\big)=\left(\tfrac{x}{y},\tfrac{z}{y}\right)\ ,\ \phi_3\big([x,y,z]\big)=\left(\tfrac{x}{z},\tfrac{y}{z}\right).$$

Sometimes we identify U_k with $\mathbb{R}^2$ using ϕ_k, and then we can consider $\mathbb{P}^2$ to be a union of three copies of $\mathbb{R}^3$. For each index k, let L_k be the projective line $L_k=\mathbb{P}^2\setminus U_k$ and note that $\mathbb{P}^2$ is the disjoint union $\mathbb{P}^2=U_k\cup L_2$. When we use the gnomic projection ϕ_k, the line L_k is called the **line at infinity**. When we identify U_k with $\mathbb{R}^2$ using ϕ_k, we consider $\mathbb{P}^2$ to be the disjoint union of $\mathbb{R}^2$ with the line at infinity.

We remark that more generally, given any projective line L in $\mathbb{P}^2$ we can define a gnomic projection $\phi:U=\mathbb{P}^2\setminus L\to\mathbb{R}^2$ as follows: let $u\in\mathbb{S}^2$ be a pole for L, choose $v,w\in\mathbb{S}^2$ so that $\{u,v,w\}$ is an orthonormal basis for $\mathbb{R}^3$, then define $\phi(xu+yv+zw)=\left(\frac{y}{x},\frac{z}{x}\right)$.

3.17 Definition: Given a polynomial $f(x,y)$ in two variables, we define the **zero set** of f in $\mathbb{R}^2$ to be the set

$$Z(f)=\big\{(x,y)\in\mathbb{R}^2\big|f(x,y)=0\big\}\subseteq\mathbb{R}^2.$$

Given a polynomial $F(x,y,z)$ in three variables, we define the **zero set** of F in $\mathbb{R}^3$ to be the set

$$Z(F)=\big\{(x,y,z)\in\mathbb{R}^3\big|F(x,y,z)=0\big\}\subseteq\mathbb{R}^3.$$

3.18 Example: When $f(x,y)=y-x^2$, the zero set $Z(f)$ is the parabola $y=x^2$. When $g(x,y)=y-p(x)$, where $p(x)$ is a polynomial in one variable, the zero set $Z(f)$ is the graph $y=p(x)$. When $h(x,y)=x^2+y^2-1$ the zero set $Z(f)$ is the circle $x^2+y^2=1$. When $F(x,y,z)=x^2+y^2+z^2-1$ we have $Z(F)=\mathbb{S}^2$.

3.19 Definition: A polynomial $F(x,y,z)$ is called **homogeneous** of degree n when for every term $c\,x^iy^jz^k$ appearing in F we have $i+j+k=n$. Notice that when F is homogeneous of degree n we have

$$F(tx,ty,tz)=t^nF(x,y,z)\ \text{ for all }\ t\in\mathbb{R}$$

so that for all $(x,y,z)\in\mathbb{R}^3$ we have

$$(x,y,z)\in Z(F)\Longrightarrow t(x,y,z)\in Z(F)\text{ for all }t\in\mathbb{R}\Longrightarrow[x,y,z]\subseteq Z(F).$$

In this case we define the **zero set** of the homogeneous polynomial F in $\mathbb{P}^2$ to be the set

$$Z(F)=\big\{[x,y,z]\in\mathbb{P}^2\big|F(x,y,z)=0\big\}\subseteq\mathbb{P}^2.$$

We do not distinguish notationally between the zero sets $Z(F)\subseteq\mathbb{R}^3$ and $Z(F)\subseteq\mathbb{P}^2$.

3.20 Exercise: Let $F(x,y,z)=x^2+y^2-z^2$. Draw a picture of $Z(F)$.

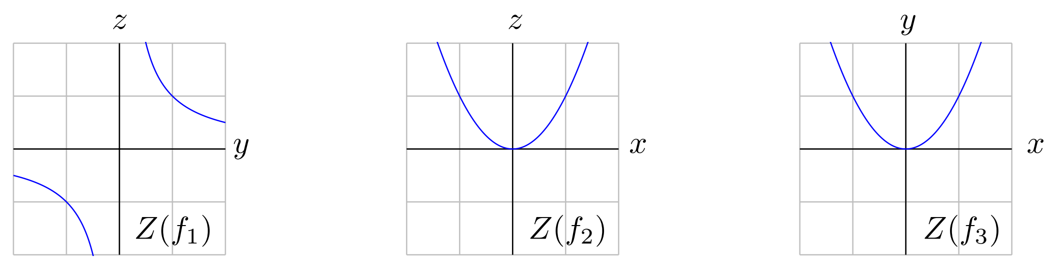
3.21 Definition: Let $F(x,y,z)$ be a homogeneous polynomial. Define

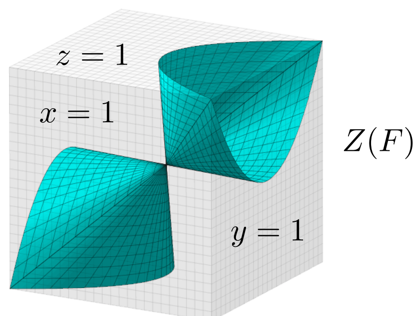
$$f_1(y,z)=F(1,y,z)\ ,\ f_2(x,z)=F(x,1,z)\ \text{ and }\ f_3(x,y)=F(x,y,1).$$

The polynomials f_1, f_2 and f_3 are called the **dehomogenizations** of F.

3.22 Example: Let $F(x,y,z) = yz - x^2$. Use the zero sets $Z(f_1)$, $Z(f_2)$ and $Z(f_3)$ to find the intersection of $Z(F)$ with each of the planes $x = 1$, $y = 1$ and $z = 1$, then draw a picture of $Z(F)$.

Solution: The dehomogenizations are given by $f_1(y,z) = yz - 1$, $f_2(x,z) = z = x^2$ and $f_3(x,y) = y - x^2$. We draw a picture of the zero sets $Z(f_1)$, $Z(f_2)$ and $Z(f_3)$. Then, to draw a picture of $Z(f)$, we draw a cube whose faces are given by $x = \pm 1$, $y = \pm 1$ and $z = \pm 1$, then we draw the zero set $Z(f_1)$ on the face $x = 1$, the zero set $Z(f_2)$ on the face $y = 1$, and the zero set $Z(f_3)$ on the face $z = 1$, in order to obtain a curve on the front, right and top faces of the cube. The points in $Z(f)$ are the points on the lines through 0 which pass through the points on this curve along the surface of the cube.





3.23 Remark: For a homogeneous polynomial $F(x,y,z)$, when we regard $\mathbb{R}^2$ as a subset of $\mathbb{P}^2$ by identifying the point $(x,y) \in \mathbb{R}^2$ with the point $[x,y,1] \in \mathbb{P}^2$, the zero set $Z(f_3) \subseteq \mathbb{R}^2$ is the restriction of the zero set $Z(F) \subseteq \mathbb{P}^2$ to the subset $\mathbb{R}^2 \subseteq \mathbb{P}^2$.

3.24 Definition: Given a polynomial $f(x,y)$ in two variables of degree n, we define the **homogenization** of f to be the homogeneous polynomial $F(x,y,z)$ obtained by replacing each term $c\,x^i y^j$ in f by the term $c\,x^i y^j z^k$ with $k = n - i - j$. Equivalently, we define

$$F(x,y,z) = z^n f\left(\tfrac{x}{z}, \tfrac{y}{z}\right).$$

The zero set $\mathbb{Z}(F) \subseteq \mathbb{P}^2$ is called the **projective completion** of the zero set $Z(f) \subseteq \mathbb{R}^2$. The points of the form $[x,y,0] \in \mathbb{P}^2$ which lie in $Z(F)$ are called the **zeros of f at infinity**.

3.25 Remark: When $F(x,y,z)$ is the homogenization of $f(x,y)$, note that f is equal to the dehomogenization f_3 of F, so $Z(f) \subseteq \mathbb{R}^2$ is the restriction of $Z(F) \subseteq \mathbb{P}^2$ to $\mathbb{R}^2 \subseteq \mathbb{P}^2$.

3.26 Example: Let $f(x,y) = xy - 1$. Find the zeros of f at infinity.

Solution: We homogenize to get $F(x,y,z) = xy - z^2$. The zeros of $f(x,y)$ at infinity are the zeros of $F(x,z,y)$ with $z = 0$. We have $F(x,y,0) = xy$ and $xy = 0$ when $x = 0$ or $y = 0$, and so the zeros at infinity are the lines $x = z = 0$ and $y = z = 0$. Using homogeneous coordinates, the zeros at infinity are the points $[0,1,0]$ and $[1,0,0]$.

3.27 Exercise: Let $f(x,y) = y - x^3$. Draw the projective completion of $Z(f)$ and find the zeros of f at infinity.

Conic Sections

3.28 Definition: For $p \in \mathbb{R}^3$, $u \in \mathbb{S}^2$ and $\phi \in \left(0, \frac{\pi}{2}\right)$, the **double cone** with vertex at p and axis in the direction $u \in \mathbb{S}^2$, which makes the angle ϕ with its axis, is the set

$$V = V(p,u,\phi) = \left\{ x \in \mathbb{R}^3 \,\Big|\, \big|(x-p)\cdot u\big| = |x-p|\cos\phi \right\}.$$

The intersection of a double cone in $\mathbb{R}^3$ with the xy-plane is called a **conic section** in $\mathbb{R}^2$. When the vertex of the cone lies in the xy-plane, the intersection can be a point or a line of a pair of intersecting lines, and these are called **degenerate conic sections**.

3.29 Note: The double cone V with vertex $(a,b,c) \in \mathbb{R}^3$ and axis direction $(u,v,w) \in \mathbb{S}^2$ and angle ϕ is given by the equation

$$\Big((x-a,y-b,z-c)\cdot(u,v,w)\Big)^2 = \big((x-a)^2+(y-b)^2+(z-c)^2\big)\cos^2\phi.$$

This equation is of degree 2 in x, y and z. The curve of intersection of V with the xy-plane is obtained by setting $z = 0$ in the above equation to obtain a degree 2 equation in x and y. Thus every conic section in $\mathbb{R}^2$ is given by a degree 2 equation in x and y.

3.30 Note: There is a theorem in linear algebra which states that every symmetric matrix is orthogonally diagonalizable. It follows from this theorem that the solution set of every degree 2 equation in x and y is either empty, or is a point, a line, a pair of lines, a parabola, a circle, an ellipse, or a hyperbola. By diagonalizing a symmetric matrix, we can find a rotation and a translation of the xy-plane to move the conic section into standard position.

3.31 Example: Diagonalize a symmetric matrix to describe the conic section

$$8x^2 - 12xy + 17y^2 - 36x + 2y = 47.$$

Solution: We can write $8x^2 - 12xy + 17y^2 = (x,y)\,A\begin{pmatrix} x \\ y \end{pmatrix}$ where $A = \begin{pmatrix} 8 & -6 \\ -6 & 17 \end{pmatrix}$. The characteristic polynomial of A is

$$f_A(\lambda) = (8-\lambda)(17-\lambda) - 36 = \lambda^2 - 25\lambda + 100 = (\lambda-5)(\lambda-20)$$

so the eigenvalues are $\lambda = 5$ and $\lambda = 20$. Performing row operations gives

$$A - 5I = \begin{pmatrix} 3 & -6 \\ -6 & 12 \end{pmatrix} \sim \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad A - 20I = \begin{pmatrix} -12 & -6 \\ -6 & 3 \end{pmatrix} \sim \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}$$

and so unit vectors for the eigenvalues 5 and 20 are $\frac{1}{\sqrt{5}}(2,1)^T$ and $\frac{1}{\sqrt{5}}(-1,2)^T$. We change coordinates (scaling by $\sqrt{5}$ and rotating by $\tan^{-1}\frac{1}{2}$) by letting $x = 2u - v$ and $y = u + 2v$. We have

$$\begin{aligned}
& 8x^2 - 12xy + 17y^2 - 36x + 2y = 47 \\
\Longleftrightarrow\ & 8(2u-v)^2 - 12(2u-v)(u+2v) + 17(u+2v)^2 - 36(2u-v) + 2(u+2v) = 47 \\
\Longleftrightarrow\ & 25\,u^2 + 100\,v^2 - 70\,u + 40\,v = 47 \\
\Longleftrightarrow\ & 25\big(u - \tfrac{7}{5}\big)^2 + 100\big(v + \tfrac{1}{5}\big)^2 = 100 \\
\Longleftrightarrow\ & \tfrac{(u-\frac{7}{5})^2}{4} + \tfrac{(v+\frac{1}{5})^2}{1} = 1.
\end{aligned}$$

This is the ellipse in the uv-plane centred at $\left(\frac{7}{5}, \frac{1}{5}\right)$ with vertices at $\left(\frac{7}{5}, \frac{1}{5}\right) \pm (2,0)$ and $\left(\frac{7}{5}, \frac{1}{5}\right) \pm (0,1)$. After changing back to the coordinates x and y (by rotating and scaling), we obtain the ellipse centred at $(x,y) = (3,1)$ with vertices at $(3,1) \pm (4,2)$ and $(3,1) \pm (-1,2)$.

3.32 Lemma: *Consider the double cone* $V(p,u,\phi)$*.*
(1) When $p=(0,0,-h)$ *with* $h>0$ *and* $u=\frac{1}{\sqrt{2}}(0,1,1)$ *and* $\phi=\frac{\pi}{4}$*, the intersection of* $V(p,u,\phi)$ *with the* xy*-plane is the parabola* $y=\frac{1}{2h}\,x^2$*.*
(2) When $p=(0,0,h)$ *with* $h>0$ *and* $u=(1,0,0)$ *and* $\phi\in\left(0,\frac{\pi}{2}\right)$*, the intersection of* $V(p,u,\phi)$ *with the* xy*-plane is the hyperbola* $\frac{x^2}{h^2\cot^2\phi}-\frac{y^2}{h^2}=1$*.*
(3) When $p=(0,0,h)$ *with* $h>0$*, and* $u=(\sin\theta,0,\cos\theta)$ *with* $\theta\in\left[0,\frac{\pi}{4}\right)$*, and* $\phi=\frac{\pi}{4}$*, the intersection of* $V\big(p,u,\phi\big)$ *with the* xy*-plane is the ellipse* $\frac{(x+h\tan 2\theta)^2}{h^2\sec^2 2\theta}+\frac{y^2}{h^2\sec 2\theta}=1$*.*

Proof: To prove Part 1, let $p=(0,0,-h)$ with $h>0$, let $u=\frac{1}{\sqrt{2}}(0,1,1)$, and let $\phi=\frac{\pi}{4}$. The double cone $V(p,u,\phi)$ is given by $\big((x,y,z+h)\cdot\frac{1}{\sqrt{2}}(0,1,1)\big)^2=\big(x^2+y^2+(z+h)^2\big)\big(\frac{1}{\sqrt{2}}\big)^2$, that is $(y+z+h)^2=x^2+y^2+(z+h)^2$. The intersection of $V(p,u,\phi)$ with the xy-plane is given by setting $z=0$ to get $(y+h)^2=x^2+y^2+h^2$, that is $y=\frac{1}{2h}\,x^2$.

To prove Part 2, let $p=(0,0,h)$ with $h>0$, let $u=(1,0,0)$, and let $\phi\in\left(0,\frac{\pi}{2}\right)$. The double cone $V(p,u,\phi)$ is given by $\big((x,y,z-h)\cdot(1,0,0)\big)^2=\big(x^2+y^2+(z-h)^2\big)\cos^2\phi$, that is $x^2=\big(x^2+y^2+(z-h)^2\big)\cos^2\phi$. The intersection of $V(p,u,\phi)$ with the xy-plane is given by setting $z=0$ to get $x^2=(x^2+y^2+h^2)\cos^2\phi$, that is $x^2(1-\cos^2\phi)-y^2\cos^2\phi=h^2\cos^2\phi$. Dividing by $h^2\cos^2\phi$ gives $\frac{x^2\sin^2\phi}{h^2\cos^2\phi}-\frac{y^2}{h^2}=1$, that is $\frac{x^2}{h^2\cot^2\phi}-\frac{y^2}{h^2}=1$.

To prove Part 3, let $p=(0,0,h)$ with $h>0$, let $u=(\sin\theta,0,\cos\theta)$ with $\theta\in\left[0,\frac{\pi}{4}\right)$, and let $\phi=\frac{\pi}{4}$. The double cone $V(p,u,\phi)$ is given by

$$\big((x,y,z-h)\cdot(\sin\theta,0,\cos\theta)\big)^2=\big(x^2+y^2+(z-h)^2\big)\cos^2\phi$$

that is

$$\big(x\sin\theta+(z-h)\cos\theta\big)^2=\big(x^2+y^2+(z-h)^2\big)\cdot\tfrac{1}{2}.$$

The intersection of the cone with the xy-plane is given by setting $z=0$ to obtain

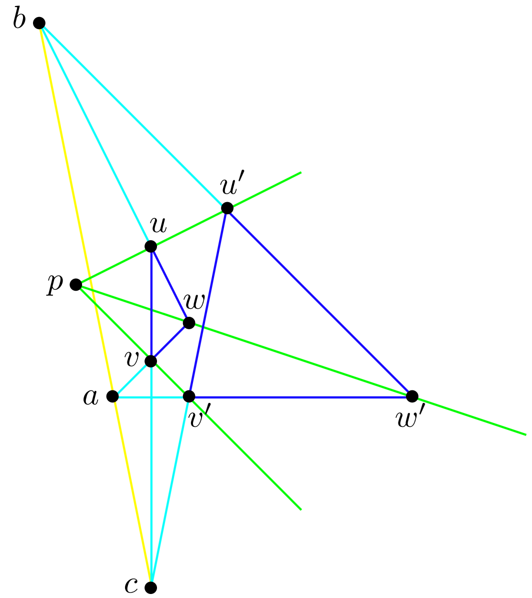
$$\begin{aligned}(x\sin\theta-h\cos\theta)^2&=(x^2+y^2+h^2)\cdot\tfrac{1}{2}\\2\big(x^2\sin^2\theta-2\,xh\sin\theta\cos\theta+h^2\cos^2\theta\big)&=x^2+y^2+h^2\\x^2(1-2\sin^2\theta)+4xh\sin\theta\cos\theta+y^2&=h^2(2\cos^2\theta-1)\\\cos 2\theta\,x^2+2h\sin 2\theta\,x+y^2&=h^2\cos 2\theta\\x^2+2h\tan 2\theta\,x+\sec 2\theta\,y^2&=h^2\\\big(x+h\tan 2\theta\big)^2-h^2\tan^2 2\theta+\sec 2\theta\,y^2&=h^2\\\big(x+h\tan 2\theta\big)^2+\sec 2\theta\,y^2&=h^2\sec^2 2\theta\\\frac{(x+h\tan 2\theta)^2}{h^2\sec^2 2\theta}+\frac{y^2}{h^2\sec 2\theta}&=1\end{aligned}$$

3.33 Corollary: *Given any curve* C *in the* xy*-plane which is a point, a line, a pair of intersecting lines, a parabola, a circle, an ellipse, or a hyperbola, one can find a double cone in* $\mathbb{R}^3$ *whose intersection with the* xy*-plane is equal to the given curve.*

3.34 Definition: A **conic** in $\mathbb{P}^2$ is the zero set of a homogenous polynomial of degree 2 in x, y and z. Note that for at least one of the variables, when we dehomogenize on that variable we obtain a degree 2 polynomial in the other 2 variables, so a conic in $\mathbb{P}^2$ is the projective completion of a conic section in $\mathbb{R}^3$. The completion of the empty set, a point, a line or a pair of lines is a degenerate conic, and the completion of a parabola, a circle, an ellipse, or a hyperbola is a non-degenerate conic.

3.35 Theorem: *(Desargue's Theorem) Let u, v, w, u', v' and w' be distinct points in $\mathbb{P}^2$ with u, v and w noncolinear and with u', v' and w' noncolinear. Suppose that the line u, u', the line v, v' and the line w, w' all intersect at a point p. Let a be the point of intersection of lines v, w and $v'w'$, let b be the point of intersection of lines w, u and $w'u'$, and let c be the point of intersection of lines u, v and $u'v'$. Then the points a, b and c are colinear.*

Proof: Choose a (projective) line M in $\mathbb{P}^2$ which does not pass through any of the given points, and use a gnomic projection from $\mathbb{P}^2 \setminus M$ to $\mathbb{R}^2$ to project all of the given points and lines to corresponding points and lines in the xy-plane. Use the same variables $u, v, w, \cdots$ to denote the corresponding points in the xy-plane. Then $[u, v, w]$ and $[u', v', w']$ are two Euclidean triangles in the xy-plane, and the three lines containing line segments $[u, u']$, $[v, v']$ and $[w, w']$ all intersect at the point p, as shown in the diagram below. Raise the points w and w' vertically out of the xy-plane so that the line in $\mathbb{R}^3$ through w and w' still passes through the point p (which is still in the xy-plane). The diagram remains unchanged when we are looking at the xy-plane with the z-axis pointing towards us, but now the triangles $[u, v, w]$ and $[u', v', w']$ are no longer contained in the xy-plane. Let P and P' be the planes in $\mathbb{R}^3$ which contain $[u, v, w]$ and $[u'v'w']$. Note that the planes P and P' are not equal because, for example, we have $u \in P$ but $u \notin P'$ (the intersection of the plane P' with the xy-plane is the line containing $[u', v']$, which does not contain u). Since the planes P and P' are not equal, they intersect in a line L in $\mathbb{R}^3$. Since the line through $[v, w]$ is contained in P we have $a \in P$, and since the line through $[v', w']$ is contained in P' we have $a \in P'$, and so we have $a \in P \cap P' = L$. Similarly, $b \in L$ and $c \in L$.

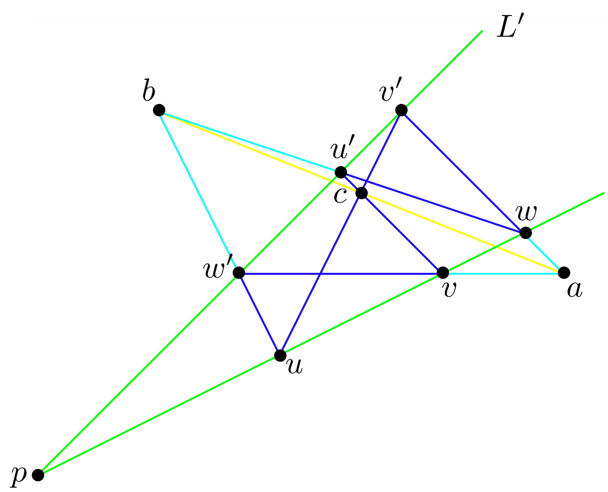


3.36 Theorem: *(Pappus' Theorem) Let L and L' be two distinct lines in $\mathbb{P}^2$. Let u, v and w be points on L and let u', v' and w' be points on L' with all 6 points distinct. Let a be the point of intersection of lines v, w' and w, v', let b be the point of intersection of lines w, u' and u, w', and let c be the point of intersection of lines u, v' and v, u'. Then the points a, b and c are colinear.*

Proof: Let p be the point of intersection of L and L'. Choose a line M in $\mathbb{P}^2$ which does not pass through any of the points $p, u, v, w, u', v', w', a, b, c$ and use a gnomic projection from $\mathbb{P}^2 \setminus M$ to $\mathbb{R}^2$ to project all of the points and lines to corresponding points and lines in the xy-plane. Use the same variables to denote the corresponding points and lines in the xy-plane. Note that we can use a translation to send p to the origin, and then we can use a linear map to send the line L to the x-axis and the line L' to the y-axis, so we may assume that $p = (0,0)$, $u = (r,0)$, $v = (s,0)$, $w = (t,0)$, $u' = (0,k)$, $v' = (0,\ell)$ and $w' = (0,m)$. It is then straightforward (but tedious) to calculate the coordinates of the points a, b and c, and to verify that they are colinear. Here are some of the steps. The line v, w' has equation $y = m - \frac{m}{s}\,x$, or $mx + sy = ms$. The line w, v' has equation $y = \ell - \frac{\ell}{t}\,x$, or $\ell x + ty = \ell t$. The point of intersection of these two lines is $a = \frac{1}{mt-\ell s}\big(st(m-\ell), \ell m(t-s)\big)$. Similarly, we have $b = \frac{1}{kr-mt}\big(rt(k-m), km(r-t)\big)$ and $c = \frac{1}{\ell s-kr}\big(sr(\ell-k), kl(s-r)\big)$. Verify that three points (x_1, y_1), (x_2, y_2) and (x_3, y_3) are colinear when $(y_2 - y_1)(x_3 - x_1) = (y_3 - y_1)(x_2 - x_1)$. In particular, the points a, b and c are colinear when

$$\left(\tfrac{km(r-t)}{kr-mt} - \tfrac{\ell m(t-s)}{mt-\ell s}\right)\left(\tfrac{sr(\ell-k)}{\ell s-kr} - \tfrac{st(m-\ell)}{mt-\ell s}\right) = \left(\tfrac{k\ell(s-r)}{\ell s-kr} - \tfrac{\ell m(t-s)}{mt-\ell s}\right)\left(\tfrac{rt(k-m)}{kr-mt} - \tfrac{st(m-\ell)}{mt-\ell s}\right).$$

By multiplying both sides by $(kr - mt)(mt - \ell s)(\ell s - kr)(mt - \ell s)$ then expanding both sides (which is tedious), one finds that equality does hold and so a, b and c are colinear.

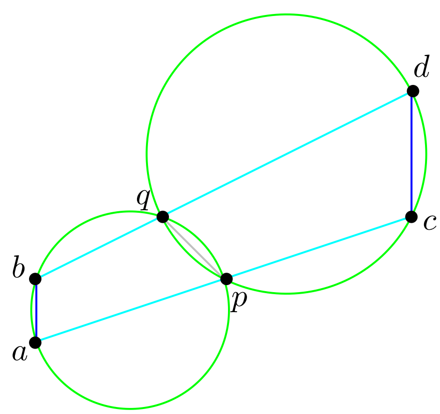


3.37 Note: Let us pause to mention some useful facts about geometry in $\mathbb{R}^2$. Note that given three points $a, b, c \in \mathbb{R}^2$, the points are colinear when $\angle_o abc = 0$ or $\pi \pmod{2\pi}$, that is when $2\angle_o abc = 0 \pmod{2\pi}$. Given four points $a, b, c, d \in \mathbb{R}^2$, note that $[a, b]$ and $[c, d]$ are parallel when $\angle_o cab = \angle_o acd$ or $\angle_o cab = \angle_o acd + \pi \pmod{2\pi}$, that is when $2\angle_o cab = 2\angle_o acd \pmod{2\pi}$. We also recall that given four points a, b, c, p, we can add oriented angles to get $\angle_o apb + \angle_o bpc = \angle_o apc \pmod{2\pi}$. Let us use these facts to prove a pleasing lemma (which we use to prove Pascal's Theorem, below).

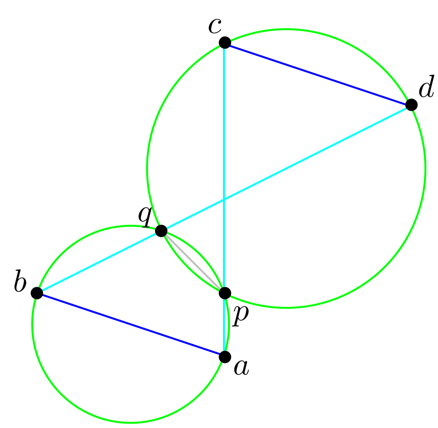
3.38 Lemma: *Let C and D be two circles in $\mathbb{R}^2$ which intersect at p and q. Let $a, b \in C$. Suppose line a, p meets D at c and line b, q meets D at d with all the points a, b, c, d, p, q distinct. Then $[a, b]$ is parallel to $[c, d]$.*

Proof: Working modulo 2π, we have

$$\begin{aligned}
2\,\angle_o cab &= 2\left(\angle_o cap + \angle_o pab\right) &&\text{, by adding angles}\\
&= 2\,\angle_o pab &&\text{, since } c, a, p \text{ are colinear (so } 2\angle_o cap = 0)\\
&= 2\,\angle_o pqb &&\text{, these angles are subtended by the chord } [p, b] \text{ in } C\\
&= 2\left(\angle_o pqd + \angle_o dqb\right) &&\text{, by adding angles}\\
&= 2\,\angle_o pqd &&\text{, since } p, q, b \text{ are colinear}\\
&= 2\,\angle_o pcd &&\text{, these angles are subtended by the chord } [p, d] \text{ in } D\\
&= 2\left(\angle_o pca + \angle_o acd\right) &&\text{by adding angles}\\
&= 2\,\angle_o acd &&\text{, since } p, c, a \text{ are colinear.}
\end{aligned}$$



$\angle_o cab = \pi + \angle_o acd$

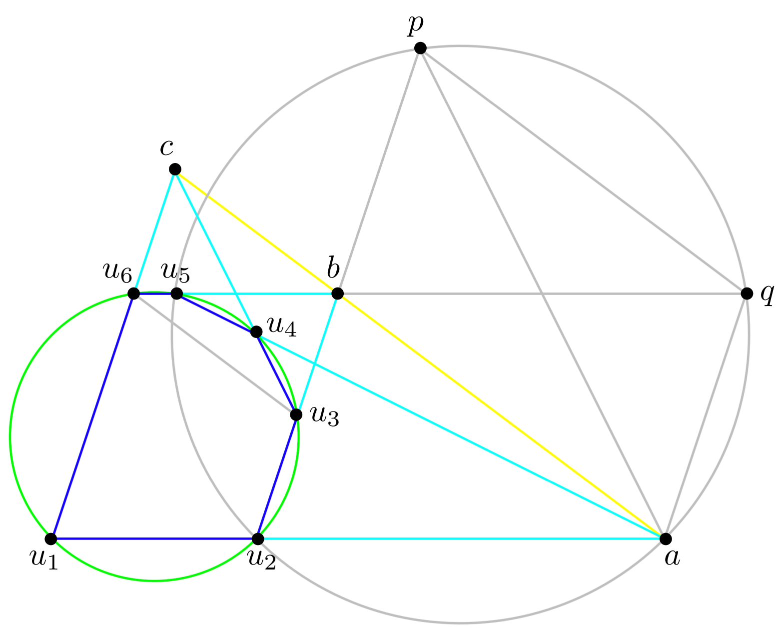


$\angle_o cab = \angle_o acd$

3.39 Theorem: *(Pascal's Theorem) Let C be a conic in $\mathbb{P}^2$. Let $u_1, u_2, \cdots, u_6$ be six distinct points on C. Let a be the point of intersection of lines u_1, u_2 and u_4, u_5, let b be the point of intersection of lines u_2, u_3 and u_5, u_6, and let c be the point of intersection of lines u_3, u_4 and u_6, u_1. Then a, b and c are colinear.*

Proof: The case in which C is empty, or C is a single point cannot occur, the case in which C is a single line is obvious, and the case that C is a pair of lines is Pappus' Theorem. Suppose that C is a non-degenerate conic. Choose a projective line M which does not pass through any of the points u_k, a, b, c, and use a gnomic projection from $\mathbb{P}^2 \setminus M$ to $\mathbb{R}^2$ to project all of the points and lines to corresponding points and lines in $\mathbb{R}^2$. Use the same variables to represent the corresponding curve and points in $\mathbb{R}^2$. The new curve C is a parabola, a circle, an ellipse or a hyperbola in $\mathbb{R}^2$. Choose a double cone V in $\mathbb{R}^3$ whose intersection with the xy-plane is equal to the curve C. Consider the projective space centred at the vertex p of the cone, consisting of all the lines in $\mathbb{R}^3$ through p. Replace C by its projective completion in the projective space centred at p, which consists of all the lines through p in V. Use another gnomic projection, this time from p to a plane which is perpendicular to the axis of the double cone V so that C is replaced by a circle. At this stage, C is a circle in $\mathbb{R}^2$ and $u_1, \cdots, u_6$ are six distinct points on C, the (Euclidean) line through $[u_1, u_2]$ intersects the line through $[u_4, u_5]$ at a, the line through $[u_2, u_3]$ intersects the line through $[u_5, u_6]$ at b, and the line through $[u_3.u_4]$ intersects the line through $[u_6, u_1]$ at c. We need to show that a, b and c are colinear in $\mathbb{R}^2$.

Let D be the circle through u_2, u_5 and a. Say the line u_2, u_3 meets D at u_2 and p, and say the line u_5, u_6 meets D at u_5 and q. By the above lemma, $[u_1, u_6]$ is parallel to $[a, q]$, and $[u_3, u_4]$ is parallel to $[p, a]$, and $[u_3, u_6]$ is parallel to $[p, q]$, so the edges of triangle $[u_3, u_6, c]$ are parallel to the corresponding edges of triangle $[p, a, q]$, and also, the edges of triangle $[u_3, u_6, b]$ are parallel to the corresponding edges of $[p, q, b]$. It follows that $[u_3, u_6, c]$ is similar to $[p, q, a]$ with scaling ratio $r = \frac{|u_3 - u_6|}{|p-q|}$, and that $[u_3, u_6, b]$ is similar to $[p, q, b]$ with the same scaling ratio r. By scaling by the factor r and applying the Side Angle Side Theorem (Corollary 1.54) it follows that $[u_6, b, c]$ is similar to $[q, b, a]$, because the angle at u_6 in $[u_6, b, c]$ is equal to the angle at q in $[q, b, a]$, and the adjacent sides are scaled by $r = \frac{|q-b}{|u_6-b|} = \frac{|q-a|}{|u_6-c|}$. Because u_6, b, q are colinear, and the angle at b in $[u_6, b, c]$ is equal to the angle at b in $[q, b, a]$, it follows that c, b, a are colinear.



3.40 Remark: The Fundamental Theorem of Algebra states that every polynomial $p(x)$ in one variable can be factored over the complex numbers into linear factors, so that if $p(x)$ is of degree n then we have $p(x) = c(x - a_1)^{m_1}(x - a_2)^{m_2} \cdots (x - a_l)^{m_l}$ for some distinct $a_i \in \mathbb{C}$ with $m_1 + m_2 + \cdots + m_l = n$. The numbers $a_i \in \mathbb{C}$ are called the **roots** of p and, for each index i, the positive integer m_i is called the **multiplicity** of the root a_i. Thus every polynomial of degree n has exactly n roots in $\mathbb{C}$ provided that the roots are counted with multiplicity (so the root a_i is counted m_i times). Geometrically, the Fundamental Theorem of Algebra states that given any polynomial p of degree n, the curve $y = p(x)$ intersects the line $y = 0$ at exactly n points, provided the points are allowed to be complex and are counted with multiplicity.

There is a very nice generalization of the Fundamental Theorem of Algebra, called Bézout's Theorem, which states that, given any two polynomials $f(x, y)$ and $g(x, y)$, with no common factors, of degrees m and n, the zero sets $Z(f)$ and $Z(g)$ intersect at exactly nm points, provided that the points are allowed to be complex and are counted with multiplicity, and we also count points at infinity. We shall not prove Bézout's Theorem here, and in fact we shall not even describe precisely what is meant by the term"multiplicity".

Chapter 4. Hyperbolic Geometry

Reflections in Circles

4.1 Definition: When L is the line in $\mathbb{R}^2$ through the point $a \in \mathbb{R}^2$ perpendicular to the vector $0 \neq u \in \mathbb{R}^2$, the **reflection** in the line L is the map $F_L : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$F_L(x) = x - 2\,\mathrm{Proj}_u(x-a) = x - \frac{2(x-a)\cdot u}{|u|^2}\, u.$$

When C is the circle in $\mathbb{R}^2$ centred at the point $a \in \mathbb{R}^2$ of radius $r > 0$, the **reflection** (or **inversion**) in the circle C is the map $F_C : \mathbb{R}^2 \setminus \{a\} \to \mathbb{R}^2 \setminus \{a\}$ given by

$$F_C(x) = a + \frac{r^2}{|x-a|^2}(x-a).$$

4.2 Example: When S is the unit circle $S = \big\{x \in \mathbb{R}^2 \big| |x| = 1\big\}$, we have $F_S(x) = \dfrac{x}{|x|^2}$.

4.3 Note: For any line L we have ${F_L}^2 = I$ and for any circle C we have ${F_C}^2 = I$.

4.4 Note: When C is the circle centred at $a \in \mathbb{R}^2$ of radius $r > 0$ note that for $x \in \mathbb{R}^2 \setminus \{a\}$ we have

(1) $|x-a| < r \iff \big|F_C(x) - a\big| > r$,
(2) $|x-a| > r \iff \big|F_C(x) - a\big| < r$ and
(3) $|x-a| = r \iff \big|F_C(x) - a\big| = r$.

In other words, F_C sends points inside C to points outside C and vice versa, and F_C fixes points on the circle C.

4.5 Note: For $u \in \mathbb{R}^2$, let $T_u : \mathbb{R}^2 \to \mathbb{R}^2$ be the **translation** defined by $T_u(x) = x + u$. For $0 \neq t \in \mathbb{R}$, let $D_t : \mathbb{R}^2 \to \mathbb{R}^2$ be the **dilation** (or **scaling map**) given by $D_t(x) = tx$. Note that when C is the circle of radius $r > 0$ centred at $a \in \mathbb{R}^2$ and S is the unit circle, the reflection F_C is equal to the composite

$$F_C = T_a D_r F_S D_{1/r} T_{-a}$$

because for $x \in \mathbb{R}^2 \setminus \{a\}$ we have

$$\begin{aligned} T_a D_r F_S D_{1/r} T_{-a}(x) &= T_a D_r F_S D_{1/r}(x-a) = T_a D_r F_S\big(\tfrac{1}{r}(x-a)\big) = T_a D_r\Big(\frac{\frac{1}{r}(x-a)}{\frac{1}{r^2}|x-a|^2}\Big) \\ &= T_a D_r\Big(\frac{r}{|x-a|^2}(x-a)\Big) = T_a\Big(\frac{r^2}{|x-a|^2}(x-a)\Big) = a + \frac{r^2}{|x-a|^2}(x-a) = F_C(x). \end{aligned}$$

4.6 Theorem: *(Reflections Preserve Lines and Circles) A reflection in a line or a circle sends lines and circles to lines and circles. When C is the circle centred at $a \in \mathbb{R}^2$ of radius $r > 0$, and $b \in \mathbb{R}^2$ with $b \neq a$ and $c = a + t(b-a)$ with $0 \neq t \in \mathbb{R}$, we have the following.*

(1) Any line through the point a is mapped by F_C to the same line.
(2) The line whose nearest point to a is b is mapped to the circle with diameter $a, F_C(b)$,
(3) The circle with diameter a, b is mapped to the line whose nearest point to a is $F_C(b)$,
(4) The circle with diameter b, c is mapped to the circle with diameter $F(b), F(c)$.

Proof: We prove Parts (2) and (4) and note that (3) follows from (2) because ${F_C}^2 = I$ so that $F_C = {F_C}^{-1}$. To prove Part (2), let L be the line whose nearest point to a is b, and let D be the circle with diameter from a to $F_C(b) = a + \frac{r^2}{|b-a|^2}(b-a)$. For $x \in \mathbb{R}^2 \setminus \{a\}$ and $y = F_C(x) = a + \frac{r^2}{|x-a|^2}(x-a)$ we have

$$x \in L \iff (x-b) \cdot (b-a) = 0$$

and

$$\begin{aligned} y \in D &\iff (y-a) \cdot \big(y - F_C(b)\big) = 0 \\ &\iff \tfrac{r^2}{|x-a|^2}(x-a) \cdot \Big(\tfrac{r^2}{|x-a|^2}(x-a) - \tfrac{r^2}{|b-a|^2}(b-a)\Big) = 0 \\ &\iff \tfrac{r^4}{|x-a|^2} - \tfrac{r^4}{|x-a|^2|b-a|^2}(x-a) \cdot (b-a) = 0 \\ &\iff |b-a|^2 - (x-a) \cdot (b-a) = 0 \iff (b-a) \cdot (b-a) - (x-a) \cdot (b-a) = 0 \\ &\iff \big((b-a) - (x-a)\big) \cdot (b-a) = 0 \iff (b-x) \cdot (b-a) = 0 \iff x \in L. \end{aligned}$$

To prove Part (4), note that since F_C is equal to the composite $F_C = T_a D_r F_S D_{1/r} T_{-a}$, where S is the unit circle $S = \mathbb{S}^1$, it suffices to prove Part (4) in the case that $C = S$. We need to show that when $0 \neq b \in \mathbb{R}^2$ and $0 \neq t \in \mathbb{R}$, the circle D with diameter from b to tb is mapped by F_S to the circle E with diameter from $F_S(b) = \frac{b}{|b|^2}$ to $F_S(tb) = \frac{tb}{|tb|^2} = \frac{b}{t|b|^2}$. When $x \in \mathbb{R}^2 \setminus \{0\}$ and $y = F_S(x) = \frac{x}{|x|^2}$ we have

$$x \in D \iff (x-b) \cdot (x-tb) = 0 \iff |x|^2 - (1+t)x \cdot b + t|b|^2 = 0$$

and

$$\begin{aligned} y \in E &\iff \big(y - F_S(b)\big) \cdot \big(y - F_S(tb)\big) = 0 \\ &\iff \big(\tfrac{x}{|x|^2} - \tfrac{b}{|b|^2}\big) \cdot \big(\tfrac{x}{|x|^2} - \tfrac{b}{t|b|^2}\big) = 0 \\ &\iff \tfrac{1}{|x|^2} - \big(\tfrac{1}{t} + 1\big)\tfrac{x \cdot b}{|x|^2|b|^2} + \tfrac{1}{t|b|^2} = 0 \\ &\iff t|b|^2 - (1+t)x \cdot b + |x|^2 = 0 \\ &\iff x \in D. \end{aligned}$$

4.7 Example: Let C be the circle centred at $a = (2,1)$ of radius $r = \sqrt{10}$. Let L be the line $y = x - 1$, and let M be the line $y = x + 1$. Find the images of the lines L and M under the reflection F_C.

4.8 Example: Let C be the circle centred at $a = (1,2)$ of radius $r = \sqrt{5}$. Let D be the circle with diameter from $(-1,4)$ to $(2,3)$ and let E be the circle $(x-3)^2 + (y-2)^2 = 9$. Find the images of circles D and E under the reflection F_C.

4.9 Theorem: *(Reflections are Conformal) Every reflection in a line or a circle is a conformal map. When C is the circle of radius $r > 0$ centred at $a \in \mathbb{R}^2$, the scaling factor of F_C at the point $x \in \mathbb{R}^2 \setminus \{a\}$ is equal to $\frac{r^2}{|x-a|^2}$.*

Proof: When L is a line in $\mathbb{R}^2$, the reflection F_L is an isometry and hence is a conformal map of scaling factor 1. Change the notation used in the statement of the theorem, and let C be the circle centred at (a, b) of radius $r > 0$. Then

$$\begin{aligned}(u,v) = F_C(x,y) &= (a,b) + \frac{r^2}{\big|(x-a,y-b)\big|^2}\big(x-a,y-b\big)\\ &= \Big(a + \frac{r^2(x-a)}{(x-a)^2+(y-b)^2}\,,\, b + \frac{r^2(y-b)}{(x-a)^2+(y-b)^2}\Big).\end{aligned}$$

Recall that F_C is conformal at (x, y) with scaling factor c when ${DF_C}^T DF_C = c^2 I$. Writing $s = x - a$ and $t = y - b$ we have

$$\begin{aligned}DF_C = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} &= \begin{pmatrix} \frac{r^2\left((x-a)^2+(y-b)^2-2(x-a)^2\right)}{\left((x-a)^2+(y-b)^2\right)^2} & \frac{-2r^2(x-a)(y-b)}{\left((x-a)^2(y-b)^2\right)^2} \\ \frac{-2r^2(x-a)(y-b)}{\left((x-a)^2(y-b)^2\right)^2} & \frac{r^2\left((x-a)^2+(y-b)^2-2(y-b)^2\right)}{\left((x-a)^2+(y-b)^2\right)^2} \end{pmatrix}\\ &= \tfrac{r^2}{(s^2+t^2)^2}\begin{pmatrix} t^2-s^2 & -2st \\ -2st & s^2-t^2 \end{pmatrix}\end{aligned}$$

and so

$${DF_C}^T DF_C = \tfrac{r^4}{(s^2+t^2)^4}\begin{pmatrix} t^2-s^2 & -2st \\ -2st & s^2-t^2 \end{pmatrix}^2 = \tfrac{r^4}{(s^2+t^2)^4}\begin{pmatrix} s^2+t^2 & 0 \\ 0 & s^2+t^2 \end{pmatrix} = \tfrac{r^4}{(s^2+t^2)^2} I.$$

Thus F_C is conformal at (x, y) with scaling factor $\frac{r^2}{s^2+t^2} = \frac{r^2}{(x-a)^2+(y-b)^2}$. Reverting to the notation used in the statement of the theorem, F_C is conformal at x of scaling factor $\frac{r^2}{|x-a|^2}$.

The Poincaré Disc Model of the Hyperbolic Plane

4.10 Definition: We define the **hyperbolic plane** (also called the **Poincaré Disc**) to be the set

$$\mathbb{H}^2 = \left\{x \in \mathbb{R}^2 \middle| |x| < 1\right\} = \left\{(x,y) \in \mathbb{R}^2 \middle| x^2 + y^2 < 1\right\}.$$

The boundary of $\mathbb{H}^2$ is the **unit circle**

$$\mathbb{S}^1 = \left\{x \in \mathbb{R}^2 \middle| |x| = 1\right\} = \left\{(x,y) \in \mathbb{R}^2 \middle| x^2 + y^2 = 1\right\}.$$

and points in $\mathbb{S}^1$ are called **points at infinity** or **asymptotic points**. In the hyperbolic plane, (hyperbolic) **length and area** are measured infinitesimally in terms of Euclidean length and area, by

$$d_H L = \frac{2}{1 - |x|^2}\, d_E L \ , \ d_H A = \frac{4}{(1 - |x|^2)^2}\, d_E A.$$

This means that the (hyperbolic) **length** of a curve given parametrically by $x = \alpha(t)$ where $\alpha : [a,b] \to \mathbb{H}^2$ is given by

$$L = \int_{t=a}^{b} d_H L = \int_{t=a}^{b} \frac{2}{1 - |x|^2}\, d_E L = \int_{t=a}^{b} \frac{2}{1 - |\alpha(t)|^2}\, |\alpha'(t)|\, dt$$

and the (hyperbolic) **area** of a region given in Cartesian coordinates by $x \in D \subseteq \mathbb{H}^2$ or by $(x,y) \in D \subseteq \mathbb{H}^2$, or in polar coordinates by $(r,\theta) \in R$, is equal to

$$\begin{aligned} A &= \iint_D d_H A = \iint_{x \in D} \frac{4}{(1 - |x|^2)^2}\, d_E A \\ &= \iint_{(x,y) \in D} \frac{4}{(1 - x^2 - y^2)^2}\, dx\, dy = \iint_{(r,\theta) \in R} \frac{4r}{(1 - r^2)^2}\, dr\, d\theta. \end{aligned}$$

The (hyperbolic) **angle** between two curves at a point in $\mathbb{H}^2$ and the (hyperbolic) **oriented angle** from one curve to another at a point in $\mathbb{H}^2$ are the same as the Euclidean angle between the two curves and the Euclidean oriented angle from one curve to the other.

4.11 Example: Let $u \in \mathbb{H}^2$. Find the hyperbolic length of the line segment from 0 to u.

Solution: The line segment is given by $x = \alpha(t) = tu$ for $0 \le t \le 1$ and we have $\alpha'(t) = u$, so the hyperbolic length is

$$\begin{aligned} L &= \int_{t=0}^{1} \frac{2\,|\alpha'(t)|}{1 - |\alpha(t)|^2}\, dt = \int_{t=0}^{1} \frac{2\,|u|}{1 - |tu|^2}\, dt = \int_{t=0}^{1} \frac{2\,|u|}{1 - t^2|u|^2}\, dt \\ &= \int_{t=0}^{1} \frac{|u|}{1 + |u|\,t} + \frac{|u|}{1 - |u|\,t}\, dt = \left[\ln \frac{1 + |u|\,t}{1 - |u|\,t}\right]_{t=0}^{1} = \ln \frac{1 + |u|}{1 - |u|}. \end{aligned}$$

4.12 Example: Find the hyperbolic area of the disc $\left\{x \in \mathbb{H}^2 \middle| |x| \le a\right\}$.

Solution: The hyperbolic area is

$$A = \int_{\theta=0}^{2\pi} \int_{r=0}^{a} \frac{4r}{(1 - r^2)^2}\, dr\, d\theta = \int_{\theta=0}^{2\pi} \left[\frac{2}{1 - r^2}\right]_{r=0}^{a} d\theta = 2\pi\left(\frac{2}{1 - a^2} - 2\right) = \frac{4\pi\, a^2}{1 - a^2}.$$

4.13 Definition: A (hyperbolic) **line** in $\mathbb{H}^2$ is set which is either of the form $L = M \cap \mathbb{H}^2$ for some line M in $\mathbb{R}^2$ through 0, or of the form $L = C \cap \mathbb{H}^2$ for some circle C in $\mathbb{R}^2$ which intersects the unit circle $\mathbb{S}^1$ orthogonally.

4.14 Note: Let C be the circle centred at $a \in \mathbb{R}^2$ of radius $r > 0$. The circle C intersects $\mathbb{S}^1$ orthogonally at b when the radius of $\mathbb{S}^1$ from 0 to b is perpendicular to the radius of C from a to b. This occurs when a lies outside $\mathbb{S}^1$ and the triangle $0, a, b$ is a right-angled triangle with hypotenuse of length $|a|$ and legs of length r and 1. Thus C intersects $\mathbb{S}^1$ orthogonally when $|a| > 1$ and $r^2 = |a|^2 - 1$.

4.15 Note: Let C be a circle which intersects $\mathbb{S}^1$ orthogonally. Since F_C sends $\mathbb{S}^1$ to a circle, and F_C fixes the two intersection points in $C \cap \mathbb{S}^1$, and F_C preserves the right angles at the two intersection points, we see that F_C sends $\mathbb{S}^1$ to itself. It follows that the map F_C restricts to a bijection $F_C : \mathbb{H}^2 \to \mathbb{H}^2$.

4.16 Definition: When $L = M \cap \mathbb{H}^2$ where M is a line through 0 in $\mathbb{R}^2$, we define the **reflection** in L to be the bijective map $F_L = F_M : \mathbb{H}^2 \to \mathbb{H}^2$. When $L = C \cap \mathbb{H}^2$ where C is a circle in $\mathbb{R}^2$ which intersects $\mathbb{S}^1$ orthogonally, we define the **reflection** in L to be the bijective map $F_L = F_C : \mathbb{H}^2 \to \mathbb{H}^2$.

4.17 Theorem: *(Reflections are Isometries) Let L be a line in $\mathbb{H}^2$. Then $F_L : \mathbb{H}^2 \to \mathbb{H}^2$ is an isometry on $\mathbb{H}^2$.*

Proof: This is (reasonably) clear when L is given by a line through the origin. Suppose that $L = C \cap \mathbb{H}^2$ where C is the circle centred at $a \in \mathbb{R}^2$ of radius $r = |a|^2 - 1$. The scaling factor of the map F_L at a point x is equal to $\frac{r^2}{|x-a|^2} = \frac{|a|^2-1}{|x-a|^2}$. In order for the map $F_L : \mathbb{H}^2 \to \mathbb{H}^2$ to be an isometry, this scaling factor must exactly compensate for the change in the scaling factor from the point x to the point $F_L(x)$ in the infinitesimal element of hyperbolic arclength, so we need to prove that when $y = F_L(x)$

$$\frac{|a|^2 - 1}{|x-a|^2} = \frac{1 - |y|^2}{1 - |x|^2}.$$

Let $x \in \mathbb{H}^2$ so that $|x| < 1$, and let $y = F_L(x) = a + \frac{r^2}{|x-a|^2}(x-a) = a + \frac{|a|^2-1}{|x-a|^2}(x-a)$. Then

$$\begin{aligned}
1 - |y|^2 &= 1 - \left(a + \tfrac{|a^2|-1}{|x-a|^2}(x-a)\right) \cdot \left(a + \tfrac{|a^2|-1}{|x-a|^2}(x-a)\right) \\
&= 1 - |a|^2 - 2\,\tfrac{|a|^2-1}{|x-a|^2}\, a \cdot (x-a) - \tfrac{(|a|^2-1)^2}{|x-a|^4}|x-a|^2 \\
&= \tfrac{|a|^2-1}{|x-a|^2}\left(-|x-a|^2 - 2\,a \cdot (x-a) - (|a|^2 - 1)\right) \\
&= \tfrac{|a|^2-1}{|x-a|^2}\left(-|x|^2 + 2\,x \cdot a - |a|^2 - 2\,a \cdot x + 2\,|a|^2 - |a|^2 + 1\right) \\
&= \tfrac{|a|^2-1}{|x-a|^2}\left(1 - |x|^2\right),
\end{aligned}$$

as required.

4.18 Theorem: *Given $u, v \in \mathbb{H}^2 \cup \mathbb{S}^1$ with $u \neq v$. there is a unique line L in $\mathbb{H}^2$ which contains (or, in the case of points at infinity, is asymptotic to) u and v.*

Proof: Let $p \in \mathbb{R}^2$ and let C be the circle in $\mathbb{R}^2$ centred at p with radius $r = \sqrt{|p|^2 - 1}$ so that C intersects $\mathbb{S}^1$ orthogonally. Note that

$$u \in C \iff |u - p|^2 = r^2 \iff |u|^2 - 2\,p \cdot u + |p|^2 = |p|^2 - 1 \iff p \cdot u = \tfrac{|u|^2+1}{2}.$$

Similarly, we have $v \in C \iff p \cdot v = \frac{|v|^2+1}{2}$. When $\{u, v\}$ is linearly independent, there is no line in $\mathbb{R}^2$ through 0 which passes through u and v and there is a unique point p such that the above circle C passes through u and v, (namely the point of intersection of the two lines $x \cdot u = \frac{|u|^2+1}{2}$ and $x \cdot v = \frac{|v|^2+1}{2}$, which are are not parallel). Suppose that $\{u, v\}$ is linearly dependent, say $u \neq 0$ and $v = tu$ with $t \in \mathbb{R}$. Then there is a unique line in $\mathbb{R}^2$ through 0 which passes through u and v, namely the line $L = \text{Span}\{u\}$, but we claim that the above circle C cannot pass through both u and v. Suppose, for a contradiction, that $u, v \in C$. Then

$$\frac{t^2|u|^2+1}{2} = \frac{|tu|^2+1}{2} = \frac{|v|^2+1}{2} = p \cdot v = p \cdot (tu) = t(p \cdot u) = t\,\frac{|u|^2+1}{2}$$

so that

$$0 = t^2|u|^2 - t\big(|u|^2 + 1\big) + 1 = \big(|u|^2 t - 1\big)(t - 1)$$

so either $t = 1$ or $t = \frac{1}{|u|^2}$. But if $t = 1$ then $v = tu = u$ and if $t = \frac{1}{|u|^2}$ then $v = tu = \frac{u}{|u|^2}$ so that $|v| = \frac{1}{|u|} > 1$ in which case $v \notin \mathbb{H}^2$.

4.19 Definition: When two hyperbolic lines meet at a point in $\mathbb{H}^2$, we say the lines **intersect**, when two lines are asymptotic at a point in $\mathbb{S}^1$, we say the lines are **asymptotic** (or **critically parallel**), and when two lines do not intersect and are not asymptotic, we say they are **parallel** (or **ultraparallel**).

4.20 Theorem: *(Perpendicular Bisector) Given $u, v \in \mathbb{H}^2$ with $u \neq v$, there is a unique line L in $\mathbb{H}^2$, called the* **perpendicular bisector** *of u and v, for which $F_L(u) = v$.*

Proof: Let $p \in \mathbb{R}^2$ with $|p| > 1$ and let C be the circle in $\mathbb{R}^2$ centred at p with radius $r = \sqrt{|p|^2 - 1}$. In order to have $F_C(u) = v$, the points u and v must lie on the same ray from p, so we must have $p = u + t(v - u)$ for some $t \in \mathbb{R}$ with $t \notin [0, 1]$. When $p = u + t(v - u)$with $t \notin [0, 1]$ we have

$$\begin{aligned}
F_C(u) = v &\iff |u - p||v - p| = r^2 \\
&\iff \big| - t(v - u)\big|\big|(1 - t)(v - u)\big| = |p|^2 - 1 \\
&\iff |t^2 - t||v - u|^2 = \big|u + t(v - u)\big|^2 - 1 \\
&\iff (t^2 - t)|v - u|^2 = |u|^2 + 2t\, u \cdot (v - u) + t^2|v - u|^2 - 1 \\
&\iff -t|v - u|^2 = |u|^2 + 2t\, u \cdot (v - u) - 1 \\
&\iff -t|v|^2 + 2t\, u \cdot v - t|u|^2 = |u|^2 + 2t\, u \cdot v - 2t\,|u|^2 - 1 \\
&\iff t\big(|u|^2 - |v|^2\big) = |u|^2 - 1 \\
&\iff |u| \neq |v| \ \text{ and } \ t = \tfrac{1-|u|^2}{|v|^2-|u|^2}.
\end{aligned}$$

Also note that the unique line M in $\mathbb{R}^2$ for which $F_M(u) = v$ is the perpendicular bisector of u and v. When $|u| = |v|$ there is no point p for which $F_C(u) = v$, but the perpendicular bisector M passes through 0. When $|u| \neq |v|$, there is a unique point p for which $F_C(u) = v$, namely the point $p = u + t(v - u)$ with $t = \frac{1-|u|^2}{|v|^2-|u|^2}$, and the perpendicular bisector M does not pass through 0.

4.21 Example: In the case that $v = 0$, in the above proof we have $t = \frac{1-|u|^2}{|v|^2-|u|^2} = -\frac{1-|u|^2}{|u|^2}$ and $p = u + t(v - u) = u - tu = (1 - t)u = \big(1 + \frac{1-|u|^2}{|u|^2}\big)u = \frac{u}{|u|^2}$. Thus the unique line L in $\mathbb{H}^2$ for which $F_L(u) = 0$ (or equivalently for which $F_L(0) = u$) is the line $L = C \cap \mathbb{H}^2$ where C is the circle centred at $p = \frac{u}{|u|^2}$ of radius $r = \sqrt{|p|^2 - 1}$.

4.22 Remark: Given $u, v \in \mathbb{H}^2$ with $u \neq v$, if L is the line through u and v and M is the perpendicular bisector of u and v, then, because $F_M(u) = v$ and $F_M(v) = u$, it follows that F_M sends the line L to itself, and so the lines L and M intersect orthogonally.

Geodesics and Distance

4.23 Definition: A **geodesic** in $\mathbb{H}^2$ is a smooth curve in $\mathbb{H}^2$ which minimizes the hyperbolic arclength between any two points on the curve.

4.24 Theorem: *The geodesics in $\mathbb{H}^2$ are the hyperbolic lines.*

Proof: First consider a curve C from 0 to u in $\mathbb{H}$. Represent C in polar coordinates by $x = \big(r(t)\cos\theta(t), r(t)\sin\theta(t)\big)$ where $r(t)$ and $\theta(t)$ are smooth functions with $r(t) \geq 0$ for all $t \in [0,1]$. Note that $|\alpha(t)|^2 = r(t)^2$ and we have

$$\begin{aligned}\alpha'(t) &= \big(r'\cos\theta - r\,\sin\theta\cdot\theta'\,,\, r'\sin\theta + r\,\cos\theta\cdot\theta'\big)\\ |\alpha'(t)|^2 &= (r'\cos\theta)^2 - 2\,rr'\,\sin\theta\cos\theta\cdot\theta' + (r\sin\theta\cdot\theta')^2\\ &\quad + (r'\sin\theta)^2 + 2\,rr'\sin\theta\cos\theta\cdot\theta' + (r\,\cos\theta\cdot\theta')^2\\ &= (r')^2 + (r\theta')^2\end{aligned}$$

and so the length of C from 0 to u is

$$\begin{aligned}L(C) &= \int_0^1 \frac{2|\alpha'(t)|}{1-|\alpha(t)|^2}\,dt = \int_{t=0}^1 \frac{2\sqrt{r'(t)^2 + r(t)^2\theta'(t)^2}}{1-r(t)^2}\,dt\\ &\geq \int_{t=0}^1 \frac{2\sqrt{r'(t)^2}}{1-r(t)^2}\,dt = \int_{t=0}^1 \frac{2|r'(t)|}{1-r(t)^2}\,dt \geq \int_{t=0}^1 \frac{2\,r(t)}{1-r(t)^2}\,dt\\ &= \int_{s=0}^{|u|} \frac{2\,ds}{1-s^2} = \int_{s=0}^{|u|} \frac{1}{1+s} + \frac{1}{1-s}\;ds = \left[\ln\frac{1+s}{1-s}\right]_{s=0}^{|u|} = \ln\frac{1+|u|}{1-|u|}.\end{aligned}$$

By comparing this with the result from Example 4.11, we see that the length of C is greater than or equal to the length of the straight line segment from 0 to u. Furthermore, the two inequalities in the above calculation only become equalities in the case that $r(t)\theta'(t) = 0$ and $r'(t) \geq 0$ for all t. When this happens we have $\theta'(t) = 0$ whenever $r(t) > 0$, and so θ is constant for $r > 0$, which implies that C is the straight line segment from 0 to u.

Now consider a smooth curve C from u to v in $\mathbb{H}^2$. Let L be the line such that $F_L(0) = u$. Since ${F_L}^2 = I$ we also have $F_L(u) = 0$. Use the isometry F_L to move the curve C to the curve $D = F_L(C)$ from $0 = F_L(u)$ to $w = F_L(v)$. Let M be the straight line from 0 to w and let $N = F_L(M)$. Notice that N is the unique hyperbolic line through u and v. The hyperbolic arclength along C from u to v is equal to the hyperbolic arclength along $D = F_L(C)$ from 0 to w which is greater than or equal to the hyperbolic arclength along the straight line M from 0 to w, which is equal to the hyperbolic arclength along N from u to v. It follows that the hyperbolic line N through u and v is the geodesic from u to v.

4.25 Definition: For $u, v \in \mathbb{H}$ we define the (hyperbolic) **distance** between u and v, denoted by $d_H(u,v)$, to be equal to the hyperbolic arclength along the (unique) line from u to v. The hyperbolic **line segment** between u and v, that is the arc between u and v along the hyperbolic line through u and v, is denoted by $[u,v]$ (we do not normally distinguish notationally between the Euclidean line segment $[u,v]$ and the hyperbolic line segment $[u,v]$, so it is sometimes necessary to specify).

4.26 Theorem: *Let $u, v \in \mathbb{H}^2$. Then*

$$d_H(u,v) = \cosh^{-1}\left(1 + \frac{2|v-u|^2}{(1-|u|^2)(1-|v|^2)}\right).$$

Proof: Let L be the line in $\mathbb{H}^2$ such that $F_L(u) = 0$ and let $w = F_L(v)$. Since F_L is an isometry, we have

$$d_H(u,v) = d_H\big(F_L(u), F_L(v)\big) = d_H(0,w).$$

From Example 4.11, we have $d_H(0,w) = \ln\frac{1+|w|}{1-|w|}$, and so

$$\begin{aligned}\cosh_H d_H(u,v) = \cosh d_H(0,w) = \cosh\left(\ln\frac{1+|w|}{1-|w|}\right) &= \frac{1}{2}\left(\frac{1+|w|}{1-|w|} + \frac{1-|w|}{1+|w|}\right)\\ = \frac{1}{2}\left(\frac{(1+|w|)^2 + (1-|w|)^2}{1-|w|^2}\right) &= \frac{1+|w|^2}{1-|w|^2} = 1 + \frac{2|w|^2}{1-|w|^2}.\end{aligned}$$

Thus to prove the theorem, it suffices to prove that

$$1 + \frac{2|w|^2}{1-|w|^2} = 1 + \frac{2|u-v|^2}{(1-|u|^2)(1-|v|^2)}.$$

Recall from Example 4.20 that $L = C \cap \mathbb{H}^2$ where C is the circle centred at $p = \frac{u}{|u|^2}$ of radius $r = \sqrt{|p|^2 - 1}$. Recall from Theorem 4.9 that F_L is conformal and scales by the factor $\frac{|p|^2-1}{|v-p|^2}$ when it sends v to w. Recall from Theorem 4.17 that F_L is an isometry because this scaling factor compensates for the change in scaling factor from v to w in the definition of $d_H L$ so we have $\frac{|p|^2-1}{|v-p|^2} = \frac{1-|w|^2}{1-|v|^2}$. Thus we have

$$\begin{aligned}1 - |w|^2 &= \frac{|p|^2-1}{|v-p|^2}(1-|v|^2) = \frac{\frac{1}{|u|^2} - 1}{|v-p|^2}(1-|v|^2) = \frac{(1-|u|^2)(1-|v|^2)}{|u|^2|v-p|^2}\\ |w|^2 &= 1 - \big(1-|w|^2\big) = 1 - \frac{(1-|u|^2)(1-|v|^2)}{|u|^2|v-p|^2} \text{ and}\\ &= \frac{|u|^2|v-p|^2 - (1-|u|^2)(1-|v|^2)}{|u|^2|v-p|^2} = \frac{\big||u|v - \frac{u}{|u|}\big|^2 - (1-|u|^2)(1-|v|^2)}{|u|^2|v-p|^2}\\ &= \frac{\big(|u|^2|v|^2 - 2(u\cdot v) + 1\big) - \big(1 - |u|^2 - |v|^2 + |u|^2|v|^2\big)}{|u|^2|v-p|^2}\\ &= \frac{|u|^2 - 2\,u\cdot v + |v|^2}{|u|^2|v-p|^2} = \frac{|u-v|^2}{|u|^2|v-p|^2}\end{aligned}$$

so that $\dfrac{|w|^2}{1-|w|^2} = \dfrac{|u-v|^2}{(1-|u|^2)(1-|v|^2)}$, as required.

4.27 Definition: For $u \in \mathbb{H}^2$ and $r > 0$, the (hyperbolic) **circle** centred at u of radius r and the (hyperbolic) **disc** centred at u of radius r are the sets

$$C_H(u,r) = \left\{x \in \mathbb{H}^2 \middle| d_H(x,u) = r\right\},$$
$$D_H(u,r) = \left\{x \in \mathbb{H}^2 \middle| d_H(x,u) \leq r\right\}$$

4.28 Note: Every hyperbolic circle is equal to a Euclidean circle (but with a different centre and radius). We can see this as follows. Consider the hyperbolic circle $C_H(u,r)$. Let L be the line in $\mathbb{H}^2$ such that $F_L(u) = 0$. Since F_L is an isometry, the image of $C_H(u,r)$ under F_L is equal to $C_H(0,r)$. By Example 4.11 the hyperbolic circle $C_H(0,r)$ is equal to the Euclidean circle $x^2+y^2=a^2$ where $r = \ln\frac{1+a}{1-a}$. The original circle $C_H(u,r)$ is equal to the image under F_L of the Euclidean circle $x^2+y^2=a^2$, which is also a Euclidean circle by Theorem 4.6.

4.29 Theorem: *Let $u \in \mathbb{H}^2$ and let $r > 0$. The circumference L of the circle $C_H(u,r)$ and the area A of the disc $D_H(u,r)$ are given by*

$$L = 2\pi \sinh r,$$
$$A = 2\pi\big(\cosh r - 1\big).$$

Proof: By the above note, the required circumference and area are the same as the hyperbolic circumference and area of the circle $x^2+y^2=a^2$ with $r = d_H((0,0),(a,0)) = \ln\frac{1+a}{1-a}$. Note that $\cosh r = \frac{1}{2}(e^r+e^{-r}) = \frac{1}{2}\big(\frac{1+a}{1-a}+\frac{1-a}{1+a}\big) = 1+\frac{2a^2}{1-a^2}$, and by Example 4.12, the area of $D_H(u,r)$ is

$$A = \frac{4\pi a^2}{1-a^2} = 2\pi(\cosh r - 1).$$

From $r = \ln\frac{1+a}{1-a}$ we have $e^r = \frac{1+a}{1-a} \Longrightarrow e^r - ae^r = 1+a \Longrightarrow e^r - 1 = a(e^r+1) \Longrightarrow a = \frac{e^r-1}{e^r-1}$. The hyperbolic circumference L of $C_H(u,r)$ is equal to the Euclidean circumference of the circle $x^2+y^2=a^2$ scaled by the factor $\frac{2}{1-a^2}$ to give

$$\begin{aligned} L &= \frac{4\pi\, a}{1-a^2} = \frac{4\pi\big(\frac{e^r-1}{e^r+1}\big)}{1-\big(\frac{e^r-1}{e^r+1}\big)^2} = \frac{4\pi(e^r-1)(e^r+1)}{(e^r+1)^2-(e^r-1)^2} \\ &= \frac{4\pi(e^{2r}-1)}{4e^r} = 2\pi\cdot\frac{e^r-e^{-r}}{2} = 2\pi\sinh r. \end{aligned}$$

4.30 Definition: As mentioned above, a Euclidean circle which is contained in $\mathbb{H}^2$ is also a hyperbolic circle (but with a different centre and radius). When a Euclidean circle E is contained in $\mathbb{H}^2 \cup \mathbb{S}^1$ and is tangent at one point in $\mathbb{S}^1$, the intersection $C = E \cap \mathbb{H}^2$ is called a **horocycle** in $\mathbb{H}^2$. When a Euclidean circle E intersects $\mathbb{S}^1$ at two distinct points, the intersection $C = E \cap \mathbb{H}^2$ is called a **hypercycle** in $\mathbb{H}^2$.

Angles and Triangles

4.31 Definition: Angles between curves in $\mathbb{H}^2$, and oriented angles from one directed curve in $\mathbb{H}^2$ to another, are the same as the corresponding Euclidean angles in $\mathbb{R}^2$. For example, given two smooth parametric curves $x = \alpha(t)$ and $y = \beta(t)$ in $\mathbb{H}^2$ with say $\alpha(0) = \beta(0) = p \in \mathbb{H}^2$, the **oriented angle** at p from the curve $x = \alpha(t)$ to the curve $x = \beta(t)$ is equal to $\theta_o\big(\alpha'(0), \beta'(0)\big) = \theta_o\big(\beta'(0)\big) - \theta_o\big(\alpha'(0)\big) \in [0, 2\pi)$, as in Definition 1.29, and it is determined by

$$\cos\theta_o\big(\alpha'(0), \beta'(0)\big) = \frac{\alpha'(0) \cdot \beta'(0)}{|\alpha'(0)|\,|\beta'(0)|} \quad \text{and} \quad \sin\theta_o\big(\alpha'(0), \beta'(0)\big) = \frac{\det\big(\alpha'(0), \beta'(0)\big)}{|\alpha'(0)|\,|\beta'(0)|},$$

as in Theorem 1.30, and the **unoriented angle** at p between the curves $x = \alpha(t)$ and $x = \beta(t)$ is given by

$$\theta\big(\alpha'(t), \beta'(0)\big) = \cos^{-1}\frac{\alpha'(0) \cdot \beta'(0)}{|\alpha'(0)|\,|\beta'(0)|} \in [0, \pi].$$

Given $u, v, w \in \mathbb{H}^2$ (or more generally, given $u, v, w \in \mathbb{H}^2 \cup \mathbb{S}^1$) with $u \neq v$ and $u \neq w$, we define the oriented and unoriented hyperbolic angles $\angle_o vuw$ and $\angle vuw$ as follows: let u_v be the unit tangent vector (or any tangent vector) at u to the arc along the hyperbolic line from u to v, and let u_w be the unit tangent vector (or any tangent vector) at u to the arc along the hyperbolic line from u to w, and define

$$\angle_o vuw = \theta_o(u_v, u_w) \quad \text{and} \quad \angle vuw = \theta(u_v, u_w).$$

For $u, v, w \in \mathbb{H}^2$ (or, more generally, for $u, v, w \in \mathbb{H}^2 \cup \mathbb{S}^1$) we say that u, v and w are **noncolinear** when there is no hyperbolic line which contains (or is asymptotic) to all three points. A (non-degenerate, hyperbolic) **triangle** in $\mathbb{H}^2$ (or in $\mathbb{H}^2 \cup \mathbb{S}^1$) is determined by three noncolinear points $u, v, w \in \mathbb{H}^2$ (or $u, v, w \in \mathbb{H}^2 \cup \mathbb{S}^1$) which we call the **vertices** of the triangle. When one or more of the vertices of a hyperbolic triangle lies in $\mathbb{S}^1$, the triangle is called an **asymptotic triangle** (we say it is **doubly asymptotic** when two of its vertices lie in $\mathbb{S}^1$ and **triply asymptotic** when all three vertices lie in $\mathbb{S}^1$). As with Euclidean or spherical triangles, we could think of a hyperbolic triangle in several ways: we could think of the triangle as being equal to its set of vertices $\{u, v, w\}$, or we can keep track of the order of the points and think of the triangle as an ordered triple (u, v, w), or we could think of the triangle as being the union of its three hyperbolic **edges** $[v, w]$, $[w, u]$ and $[u, v]$, (where for example $[u, v]$ denotes the arc along the hyperbolic line from u to v), or we can think of the hyperbolic triangle as the region $[u, v, w] \subseteq \mathbb{H}^2 \cup \mathbb{S}^1$ which is bounded by the three edges. We shall agree that an **ordered triangle** in $\mathbb{H}^2$ (or in $\mathbb{H}^2 \cup \mathbb{S}^1$) consists of an ordered triple (u, v, w) of noncolinear points in $\mathbb{H}^2$ (or in $\mathbb{H}^2 \cup \mathbb{S}^1$) together with the region $[u, v, w]$ which is bounded by the three edges $[v, w]$, $[w, u]$ and $[u, v]$. For this triangle, we shall normally denote the hyperbolic **edge lengths** by a, b and c with

$$a = d_H(v, w)\ ,\ b = d_H(w, u)\ ,\ c = d_H(u, v)$$

and we shall normally denote the oriented and unoriented **angles** at the vertices by α_o, β_o and γ_o and α, β and γ with

$$\alpha_o = \angle_o vuw\ ,\ \beta_o = \angle_o wvu\ ,\ \gamma_o = \angle_o uwv\ ,\ \alpha = \angle vuw\ ,\ \beta = \angle wvu\ ,\ \gamma = \angle uwv.$$

The unoriented angles α, β and γ are also called the **interior angles** of the triangle, and the **exterior angles** are given by $\pi - \alpha$, $\pi - \beta$ and $\pi - \gamma$.

4.32 Example: Let $u = \left(\frac{1}{2}, 0\right)$, $v = \left(\frac{1}{2}, \frac{1}{2}\right)$ and $w = \left(-\frac{1}{2}, -\frac{1}{2}\right)$. In the hyperbolic triangle $[u, v, w]$, find the edge length $b = d_H(w, u)$ and the oriented angle $\beta_o = \angle_o wvu$.

Solution: The edge length b is given by

$$b = d_H(w,u) = \cosh^{-1}\left(1 + \tfrac{2|w-u|^2}{(1-|w|^2)(1-|u|^2)}\right) = \cosh^{-1}\left(1 + \tfrac{2\cdot\frac{5}{4}}{\frac{1}{2}\cdot\frac{3}{4}}\right) = \cosh^{-1}\tfrac{23}{3}\,.$$

To find the oriented angle β_o at the vertex v, we shall find v_u and v_w. The hyperbolic line L through v and w is the same as the Euclidean line L through v and w (since it passes through the origin), namely the line $y = x$. Since L has slope 1, we have $v_w = (-1, -1)$ (or some positive multiple of that). Let N be the hyperbolic line through u and w, say $N = C_E(p, r) \cap \mathbb{H}^2$ with $p = (x, y)$. To have $u \in N$, as in the proof of Theorem 4.18 we need $p \bullet u = \frac{|u|^2+1}{2}$, that is $\frac{1}{2}x = \frac{\frac{1}{4}+1}{2} = \frac{5}{8}$ (1). To have $v \in N$ we need $p \bullet v = \frac{|v|^2+1}{2}$, that is $\frac{1}{2}x = \frac{1}{2}y = \frac{\frac{1}{2}+1}{2} = \frac{3}{4}$, or equivalently $x + y = \frac{3}{2}$ (2). Solve Equations (1) and (2) to get $p = (x, y) = \left(\frac{5}{4}, \frac{1}{4}\right)$. We also remark (even though we do not need it for our calculations) that as in Note 4.14, we must have $r = \sqrt{|p|^2 - 1} = \frac{\sqrt{10}}{4}$. Since the radius $v - p = \left(-\frac{3}{4}, \frac{1}{4}\right)$ has slope $-\frac{1}{3}$, the tangent to N at v has slope 3, so we have $v_u = (-1, -3)$ (or any positive multiple of that). Thus $\beta_o = \theta_o(v_w, v_u) = \theta_o\big((-1,-1), (-1,-3)\big)$. Since $\cos\beta_o = \frac{v_w \bullet v_u}{|v_w|\,|v_u|} = \frac{(-1,-1)\bullet(-1,-3)}{|(-1,-1)|\,|(-1,-1)|} = \frac{4}{\sqrt{2}\sqrt{10}} = \frac{2}{\sqrt{5}}$, and $\det(v_w, v_u) = \det\begin{pmatrix} -1 & -1 \\ -1 & -3 \end{pmatrix} = 2$ so that $\sin\beta_o > 0$, we have $\beta_o = \cos^{-1}\frac{2}{\sqrt{5}} = \sin^{-1}\frac{1}{\sqrt{5}} = \tan^{-1}\frac{1}{2}$.

4.33 Theorem: *For a triangle in $\mathbb{H}^2$ with side lengths a, b and c and interior angles α, β and γ, we have*

(1) (The Sine Law) $$\frac{\sinh a}{\sin\alpha} = \frac{\sinh b}{\sin\beta} = \frac{\sinh c}{\sin\gamma},$$

(2) (The First Cosine Law) $$\cos\alpha = \frac{\cosh a - \cosh b\cosh c}{-\sinh b\sinh c}\,, \text{ and}$$

(3) (The Second Cosine Law) $$\cosh a = \frac{\cos\alpha + \cos\beta\cos\gamma}{\sin\beta\sin\gamma}.$$

Similar rules hold with a, b, c and α, β, γ permuted.

Proof: Use a hyperbolic reflection to move the vertex u to the position $(0, 0)$, then use another reflection to move the vertex v to position $(s, 0)$ with $0 < s < 1$, and use a third reflection to move the vertex w to the upper half of $\mathbb{H}^2$ at position $(t\cos\alpha, t\sin\alpha)$ with $0 < t < 1$. Note that the reflections preserve the edge lengths and angles, so it suffices to prove that the Sine Law and Cosine Laws hold for a triangle with vertices at $u = (0, 0)$, $v = (s, 0)$ and $w = (t\cos\alpha, t\sin\alpha)$. For this triangle, α is the angle at u and we have

$$\cosh b = \cosh\big(d_H(w,u)\big) = 1 + \tfrac{2|w-u|^2}{(1-|w|^2)(1-|u|^2)} = 1 + \tfrac{2t^2}{1-t^2} = \tfrac{1+t^2}{1-t^2}$$
$$\sinh b = \sqrt{\cosh^2 b - 1} = \sqrt{\tfrac{1+2t^2+t^4}{1-2t^2+t^4} - 1} = \sqrt{\tfrac{4t^2}{1-2t^2+t^4}} = \tfrac{2t}{1-t^2}$$

and similarly $\cosh c = \frac{1+s^2}{1-s^2}$ and $\sinh c = \frac{2s}{1-s^2}$. Also, by the Euclidean Law of Cosines, we have $|v - w|^2 = s^2 + t^2 - 2st\cos\alpha$ so that

$$\begin{aligned}\cosh a &= \cosh\big(d_H(v,w)\big) = 1 + \tfrac{2|v-w|^2}{(1-|v|^2)(1-|w|^2)} = 1 + \tfrac{2(s^2+t^2-2st\cos\alpha)}{(1-s^2)(1-t^2)} \\ &= \tfrac{(1-s^2)(1-t^2)+2(s^2+t^2)-4st\cos\alpha}{(1-s^2)(1-t^2)} = \tfrac{(1+s^2)(1+t^2)-4st\cos\alpha}{(1-s^2)(1-t^2)}.\end{aligned}$$

Thus we have

$$\frac{\cosh b\cosh c-\cosh a}{\sinh b\sinh c}=\frac{\left(\frac{1+t^2}{1-t^2}\right)\left(\frac{1+s^2}{1-s^2}\right)-\frac{(1+s^2)(1+t^2)-4st\cos\alpha}{(1-s^2)(1-t^2)}}{\left(\frac{2t}{1-t^2}\right)\left(\frac{2s}{1-s^2}\right)}=\cos\alpha$$

proving the First Cosine Law.

To prove the Sine Law, let us find $\sin\beta$ and $\sin\gamma$. Let L be the hyperbolic line through v and w, say $L=C_E(p,r)$ with $p=(x,y)$ and $r=\sqrt{|p|^2-1}$. To get $v\in L$ we need $p\cdot v=\frac{|v|^2+1}{2}$, that is $sx=\frac{s^2+1}{2}$, so we have $x=\frac{1+s^2}{2s}$. To have $w\in L$ we need $p\cdot w=\frac{|w|^2+1}{2}$, that is $(t\cos\alpha)x+(t\sin\alpha)y=\frac{t^2+1}{2}$ so we have $y=\frac{1}{\sin\alpha}\left(\frac{1+t^2}{2t}-x\cos\alpha\right)=\frac{1}{\sin\alpha}\left(\frac{1+t^2}{2t}-\frac{1+s^2}{2s}\cos\alpha\right)$. Thus we have

$$p=(x,y)=\left(\frac{1+s^2}{2s}\,,\,\frac{1}{\sin\alpha}\left(\frac{1+t^2}{2t}-\frac{1+s^2}{2s}\cos\alpha\right)\right)\ \text{ and }\ r=\sqrt{|p|^2-1}.$$

Note also that $x-s=\frac{1+s^2}{2s}-s=\frac{1-s^2}{2s}>0$. The radius vector from v to p is $p-v=(x-s,y)$ and so we have $v_w=(-y,x-s)$ with $|v_w|=|p-v|=r$. We also have $v_u=(-1,0)$ with $|v_u|=1$, so that

$$\cos\beta=\frac{v_w\bullet v_u}{|v_w|\,|v_u|}=\frac{y}{r}.$$

Also note that $r^2=|p-v|^2=(x-s)^2+y^2$, so we have

$$\sin\beta=\sqrt{1-\cos^2\beta}=\sqrt{1-\frac{y^2}{r^2}}=\sqrt{\frac{r^2-y^2}{r^2}}=\sqrt{\frac{(x-s)^2}{r^2}}=\frac{x-s}{r}=\frac{1-s^2}{2rs}.$$

By symmetry, that is by interchanging the roles of v and w, we also have $\sin\gamma=\frac{1-t^2}{2rt}$ so that

$$\frac{\sinh b}{\sin\beta}=\frac{\frac{2t}{1-t^2}}{\frac{1-s^2}{2rs}}=\frac{2rst}{(1-s^2)(1-t^2)}=\frac{\frac{2s}{1-s^2}}{\frac{1-t^2}{2rt}}=\frac{\sinh c}{\sin\gamma}$$

proving the Sine Law.

For the Second Cosine Law, let us also find $\cos\gamma$. The radius vector from p to w is $w-p=\big(t\cos\alpha-x\,,\,t\sin\alpha-y\big)$ so we have $w_v=\big(y-t\sin\alpha,t\cos\alpha-x\big)$ with $|w_v|=|w-p|=r$, and we have $w_u=(-\cos\alpha,-\sin\alpha)$ with $|w_u|=1$, and hence

$$\cos\gamma=\frac{w_u\cdot w_v}{|w_u|\,|w_v|}=\frac{x\sin\alpha-y\cos\alpha}{r}.$$

Thus, making use of many of the above formulas, including the formulas $(x-s)^2+y^2=r^2$ and $y=\frac{1}{\sin\alpha}\left(\frac{1+t^2}{2t}-x\cos\alpha\right)$ and $x=\frac{1+s^2}{2s}$ so that $2sx-s^2=1$, we have

$$\begin{aligned}\frac{\cos\alpha+\cos\beta\cos\gamma}{\sin\beta\sin\gamma}&=\frac{\cos\alpha+\frac{y}{r}\cdot\frac{x\sin\alpha-y\cos\alpha}{r}}{\frac{1-s^2}{2rs}\cdot\frac{1-t^2}{2rt}}=\frac{4r^2st\cos\alpha+4st\,xy\sin\alpha-4sty^2\cos\alpha}{(1-s^2)(1-t^2)}\\&=\frac{4st\Big((x-s)^2\cos\alpha+x\Big(\frac{1+t^2}{2t}-x\cos\alpha\Big)\Big)}{(1-s^2)(1-t^2)}=\frac{4st\Big(x\,\frac{1+t^2}{2t}-(2sx-s^2)\cos\alpha\Big)}{(1-s^2)(1-t^2)}\\&=\frac{4st\Big(\Big(\frac{1+s^2}{2s}\Big)\Big(\frac{1+t^2}{2t}\Big)-\cos\alpha\Big)}{(1-s^2)(1-t^2)}=\frac{(1+s^2)(1+t^2)-4st\cos\alpha}{(1-s^2)(1-t^2)}=\cosh a\end{aligned}$$

proving the Second Cosine Law.

4.34 Exercise: Let $u=\left(\frac{1}{2},0\right)$, $v=\left(\frac{1}{2},\frac{1}{2}\right)$ and $w=\left(-\frac{1}{2},-\frac{1}{2}\right)$ (as in Example 4.32). In the hyperbolic triangle $[u,v,w]$, find a, b and c, then find $\cos\beta$ using the First Cosine Law.

4.35 Lemma: *The area of a doubly asymptotic triangle in $\mathbb{H}^2 \cup \mathbb{S}^1$ with interior angle α at its non-asymptotic vertex is equal to $A = \pi - \alpha$.*

Proof: Consider a doubly asymptotic triangle with angle $\alpha = 2\beta$ at its non-asymptotic vertex. We can use hyperbolic reflections to move the triangle so that the non-asymptotic vertex is at the origin and the asymptotic vertices are $u = (\cos\beta, \sin\beta)$ and $v = (\cos\beta, -\sin\beta)$. Note that the hyperbolic line L which is asymptotic to u and v is equal to $L = C \cap \mathbb{H}^2$ where C is the circle of radius $r = \tan\beta$ centred at $a = (\sec\beta, 0)$. For a point $x = (r\cos\theta, r\sin\theta) \in L$ with $-\beta \le \theta \le \beta$ and $0 \le r \le 1$, the Law of Cosines applied to the triangle with vertices at $0, a, x$ and the Quadratic Formula give

$$\begin{gathered}\tan^2\beta = r^2 + \sec^2\beta - 2r\sec\beta\,\cos\theta \\ r^2 - 2r\sec\beta\cos\theta + 1 = 0 \\ r = \sec\beta\cos\theta \pm \sqrt{\sec^2\beta\cos^2\theta - 1}.\end{gathered}$$

For $-\beta \le \theta \le \beta$ we have $\cos\theta \ge \cos\beta$ so that $\sec\beta\cos\theta \ge 1$, and so in order to have $r \le 1$ we must use the negative sign. Thus the line L is given in polar coordinates by

$$r = \sec\beta\cos\theta - \sqrt{\sec^2\beta\cos^2\theta - 1}.$$

Thus the area of the doubly asymptotic triangle is

$$\begin{aligned}
A &= \int_{\theta=-\beta}^{\beta} \int_{r=0}^{\sec\beta\cos\theta - \sqrt{\sec^2\beta\cos^2\theta-1}} \frac{4r}{(1-r^2)^2}\,dr\,d\theta \\
&= \int_{\theta=-\beta}^{\beta} \left[\frac{2}{1-r^2}\right]_{r=0}^{\sec\beta\cos\theta - \sqrt{\sec^2\beta\cos^2\theta-1}} d\theta \\
&= \int_{\theta=-\beta}^{\beta} \frac{2}{1-\left(\sec\beta\cos\theta - \sqrt{\sec^2\beta\cos^2\theta - 1}\right)^2} - 2\,d\theta \\
&= \int_{\theta=-\beta}^{\beta} \frac{2}{1-\left(2\sec^2\beta\cos^2\theta - 1 - 2\sec\beta\cos\theta\sqrt{\sec^2\beta\cos^2\theta - 1}\right)} - 2\,d\theta \\
&= \int_{\theta=-\beta}^{\beta} \frac{1}{-\left(\sec^2\beta\cos^2\theta - 1\right) + \sec\beta\cos\theta\sqrt{\sec^2\beta\cos^2\theta - 1}} - 2\,d\theta \\
&= \int_{\theta=-\beta}^{\beta} \frac{-(\sec^2\beta\cos^2\theta - 1) - \sec\beta\cos\theta\sqrt{\sec^2\beta\cos^2\theta - 1}}{\left(\sec^2\beta\cos^2\theta - 1\right)^2 - \sec^2\beta\cos^2\theta\left(\sec^2\beta\cos^2\theta - 1\right)} - 2\,d\theta \\
&= \int_{\theta=-\beta}^{\beta} \frac{-\left(\sec^2\beta\cos^2\theta - 1\right) - \sec\beta\cos\theta\sqrt{\sec^2\beta\cos^2\theta - 1}}{\left(\sec^2\beta\cos^2\theta - 1\right)\left(\sec^2\beta\cos^2\theta - 1 - \sec^2\beta\cos^2\theta\right)} - 2\,d\theta \\
&= \int_{\theta=-\beta}^{\beta} 1 + \frac{\sec\beta\cos\theta}{\sqrt{\sec^2\beta\cos^2\theta - 1}} - 2\,d\theta = \int_{\theta=-\beta}^{\beta} \frac{\cos\theta}{\sqrt{\cos^2\theta - \cos^2\beta}} - 1\,d\theta \\
&= \int_{\theta=-\beta}^{\beta} \frac{\cos\theta}{\sqrt{\sin^2\beta - \sin^2\theta}}\,d\theta\ - 2\beta = \int_{\phi=-\pi/2}^{\pi/2} \frac{\sin\beta\cos\phi\,d\phi}{\sin\beta\cos\phi} - \alpha = \pi - \alpha
\end{aligned}$$

where on the last line we made the trigonometric substitution $\sin\beta\sin\phi = \sin\theta$ so that $\sqrt{\sin^2\beta - \sin^2\theta} = \sin\beta\cos\phi$ and $\cos\theta\,d\theta = \sin\beta\cos\phi\,d\phi$.

4.36 Theorem: *The area of a triangle in* $\mathbb{H}^2$ *(or* $\mathbb{H}^2 \cup \mathbb{S}^1$*) with interior angles* α*,* β *and* γ *is equal to*

$$A = \pi - (\alpha + \beta + \gamma).$$

This includes asymptotic triangles with one or more vertices on $\mathbb{S}^1$ *(the interior angles at asymptotic points are equal to zero).*

Proof: This follows from the above lemma because a triply asymptotic triangle can be cut into two doubly asymptotic triangles, and given a singly asymptotic triangle we can add a doubly asymptotic triangle to form a doubly asymptotic triangle, and given a non-asymptotic triangle, we can add three doubly asymptotic triangles to form a triply asymptotic triangle.

4.37 Definition: Let L and M be two distinct lines in $\mathbb{H}^2$. When L and M intersect at a point $p \in \mathbb{H}^2$ and $\theta = 2\varphi$ where φ is the oriented angle from L counterclockwise to M at p, the isometry $R_{p,\theta} = F_M F_L$ is called the **rotation** about p by θ in $\mathbb{H}^2$. When L and M are asyptotic at the point $p \in \mathbb{S}^1$, the isometry $P = F_M F_L$ is called a **horolation** (or a **parallel displacement**) about p in $\mathbb{H}^2$. When L and M do not intersect and are not asymptotic, and N is the unique line which intersects orthogonally with L and M, the isometry $T = F_M F_L$ is called a **translation** along N in $\mathbb{H}^2$, and the isometry $F_N F_M F_L$ is called a **glide reflection** along N in $\mathbb{H}^2$.

4.38 Remark: A rotation on $\mathbb{H}^2$ is also called an **elliptic isometry** on $\mathbb{H}^2$, a horolation on $\mathbb{H}^2$ is also called a **parabolic isometry** on $\mathbb{H}^2$, and a translation on $\mathbb{H}^2$ is also called a **hyperbolic isometry** on $\mathbb{H}^2$.

4.39 Remark: A rotation about the point $p \in \mathbb{H}^2$ moves each point along a (hyperbolic) circle centred at p. A horolation about the point $p \in \mathbb{S}^1$ moves each point along a horocycle at p (that is a Euclidean circle in $\mathbb{H}^2$ which is tangent to $\mathbb{S}^1$ at p). A translation along the line L in $\mathbb{H}^2$ which is asymptotic to $u, v \in \mathbb{S}^1$, moves each point along a hypercycle from u to v (that is along an arc of a Euclidean circle through u and v).

4.40 Theorem: *Let $u, v \in \mathbb{H}^2$ with $u \neq v$. Let L be the perpendicular bisector of u and v in $\mathbb{H}^2$. Then for $x \in \mathbb{H}^2$ we have $d_H(x,u) = d_H(x,v) \iff x \in L$.*

Proof: If $x \in L$ then $F_L(x) = x$ (from Definition 4.1 or from Part 1 of Theorem 4.6) and so, since F_L is an isometry, we have $d_H(x,u) = d_H\big(F_L(x), F_L(u)\big) = d_H(x,v)$.

Recall that (in the statement of Theorem 4.20) we defined the perpendicular bisector L to be the hyperbolic line such that $F_L(u) = v$. Let M be the hyperbolic line through u and v and let m be the point of intersection of L with M. Note that since $m \in L$, we have $d_H(m,u) = d_H(m,v)$ (as shown above), and so m is the hyperbolic midpoint of the hyperbolic line segment $[u,v]$. Since $F_L(u) = v$ and $F_L(m) = m$, we have $F_L(M) = M$ and F_L sends the hyperbolic line segment $[u,m]$ to the hyperbolic line segment $[v,m]$ with $\angle umv = \pi$. Since F_L preserves angles, the angle between L and M at m is equal to $\frac{\pi}{2}$ (for $p \in L$ with $p \neq m$ we have $\angle ump = \angle vmp$ and $\angle ump + \angle vmp = \pi$). This shows that the perpendicular bisector L can also be described as the line through the hyperbolic midpoint m of $[u,v]$ which is orthogonal to $[u,v]$.

Let $x \in \mathbb{H}^2$ with $d_H(x,u) = d_H(x,v)$. Note that in the two hyperbolic triangles $[x,u,m]$ and $[x,v,m]$, the corresponding edge lengths are all equal and hence, by the First Law of Cosines, the corresponding interior angles are all equal. In particular, we have $\angle umx = \angle vmx$. Since m lies between u and v, we also have $\angle umx + \angle vmx = \pi$, and so $\angle umx = \angle vmx = \frac{\pi}{2}$. Thus x lies on the line through m which is orthogonal to $[u,v]$, so $x \in L$, as required.

4.41 Theorem: *Let $[u,v,w]$ be a triangle in $\mathbb{H}^2$ (so the points $u, v, w \in \mathbb{H}^2$ are non-colinear). Then a point $x \in \mathbb{H}^2$ is uniquely determined by the distances $d_H(x,u)$, $d_H(x,v)$ and $d_H(x,w)$.*

Proof: Let $x, y \in \mathbb{H}^2$ with $x \neq y$ and suppose, for a contradiction, that $d_H(x,u) = d_H(y,u)$ and $d_H(x,v) = d_H(y,v)$ and $d_H(x,w) = d_H(y,w)$. Let L be the perpendicular bisector of $[x,y]$ in $\mathbb{H}$. By the above theorem, since $d_H(x,u) = d_H(y,u)$ we have $u \in L$, and since $d_H(x,v) = d_H(y,v)$ we have $v \in L$, and since $d_H(x,w) = d_H(y,w)$ we have $w \in L$, which contradicts the fact that u, v and w are non-colinear.

4.42 Theorem: *Let $[u,v,w]$ and $[u',v',w']$ be ordered triangles in $\mathbb{H}^2$ with corresponding edge lengths equal, that is with $a=a'$, $b=b'$ and $c=c'$. Then there exists a unique isometry F on $\mathbb{H}^2$ such that $F(u)=u'$, $F(v)=v'$ and $F(w)=w'$.*

Proof: The uniqueness of such an isometry follows from the previous theorem. Indeed assuming that such an isometry F exists, then given any point $x\in\mathbb{H}^2$, the point $y=F(x)$ is the unique point $y\in\mathbb{H}^2$ such that $d_H(y,u')=d_H(x,u)$, $d_H(y,v')=d_H(x,v)$ and $d_H(y,w')=d_H(x,w)$.

It remains to show that such an isometry on $\mathbb{H}^2$ exists. If $u=u'$ then let F_1 be the identity map, and if $u\neq u'$ then let F_1 be the hyperbolic reflection in the perpendicular bisector L of u and u'. Let $u_1=F_1(u)=u'$, $v_1=F_1(v)$ and $w_1=F_1(w)$. If $v_1=v'$ then let F_2 be the identity map, and if $v_1\neq v'$ then let F_2 be the hyperbolic reflection in the perpendicular bisector M of v_1 and v'. Note that since $d_H(u_1,v_1)=d_H(u,v)=d_H(u',v')=d_H(u_1,v')$ we have $u_1\in M$ so that $F_M(u_1)=u_1=u'$. Let $u_2=F_2(u_1)=u'$, $v_2=F_2(v_1)=v'$ and $w_2=F_2(w_1)$. If $w_1=w'$ then let F_3 be the identity map, and if $w_2\neq w'$ then let F_3 be the hyperbolic reflection in the perpendicular bisector N of w_2 and w'. As above, since $d_H(u_2,w_2)=d_H(u,w)=d_H(u',w')=d_H(u_2,w')$ we have $u_2\in N$ so that $F_N(u_2)=u_2=u'$, and since $d_H(v_2,w_2)=d_H(v,w)=d_H(v',w')=d_H(v_2,w')$ we have $v_2\in N$ so that $F_N(v_2)=v_2=v'$. Thus we can let F be the composite $F=F_3F_2F_1$ and then we have $F(u)=u'$, $F(v)=v'$ and $F(w)=w'$, as required.

4.43 Theorem: *Every isometry on $\mathbb{H}^2$ is equal to a product of 0, 1, 2 or 3 reflections.*

Proof: Let F be any isometry on $\mathbb{H}^2$. Let $u=(0,0)$, $v=\left(\frac{1}{2},0\right)$ and $w=\left(0,\frac{1}{2}\right)$, and let $u'=F(u)$, $v'=F(v)$ and $w'=F(w)$. The proof of the previous theorem shows that $F=F_3F_2F_1$ where each F_k is equal either to the identity map or to a hyperbolic reflection (the product of zero reflections is the identity map, which occurs when all three of the maps F_k is the identity map).

4.44 Theorem: *Every isometry on $\mathbb{H}^2$ is equal to the identity, a rotation, a translation, a parallel displacement, or a reflection or a glide reflection.*

Proof: Every isometry is the product of 0, 1, 2 or 3 reflections, and the product of 0 reflections is the identity map, the product of 1 reflections is a reflection, and the product of 2 reflections (by definition) is a rotation, a translation or a parallel displacement. It remains to consider the product of 3 reflections. Suppose that $F=F_NF_MF_L$. If $M=L$ then we have $F=F_N$, which is a reflection. Suppose that $M\neq L$. There are three cases to consider: either M and L intersect in $\mathbb{H}^2$, or M and L are asymptotic, or M and L are parallel. We shall consider only the first case, and leave the other two cases as an exercise.

Case 1: suppose that $L\cap M=\{a\}$ and $F_MF_L=R_{a,\theta}$. Let $N'=N$, let M' be the (unique) hyperbolic line through a which is perpendicular to $N=N'$, say M' intersects N' at b, and let L' be the (unique) hyperbolic line through a such that the oriented angle from L' to M' is equal to $\frac{\theta}{2}$ so that $R_{a,\theta}=F_{M'}F_{L'}$. Then we have $F=F_NF_MF_L=F_NR_{a,\theta}=F_{N'}F_{M'}F_{L'}=R_{b,\pi}F_{L'}$. Let $L''=L'$, let N'' be the (unique) hyperbolic line through b perpendicular to L'', and let M'' be the (unique) hyperbolic line through b perpendicular to N'' so that $R_{b,\pi}=F_{N''}F_{M''}$. Then we have $F=R_{b,\pi}F_{L'}=F_{N''}F_{M''}F_{L''}$ where L'' and M'' are both perpendicular to N''. If $L''=M''$ then $F=F_{N''}$ which is a reflection, and if $L''\neq M''$ the F is a glide refection along N''.

For the other cases, first show that given a line L through $u\in\mathbb{S}^1$ and a parallel displacement P about u, there is a line M such that $P=F_MF_M$, and given orthogonal lines L and K and a translation T along K, there is a line M such that $T=F_MF_L$.

The Half-Plane Model, the Minkowski Model, and the Klein Model

We have described the Poincaré disc model of the hyperbolic plane, but there are several other models of the hyperbolic plane that are sometimes used: there are alternate ways of constructing a geometry (a set with a an abstract way of measuring distance between points) in which we can define lines and circles and triangles which have have the same properties and satisfy the same formulas as lines circles and triangles in the Poincaré disc (for example the formula for the area of a disc, the laws of cosines, and the formula for the area of a triangle). Here is a brief description of three such models.

The Poincaré **upper half plane model** of the hyperbolic plane is constructed as follows. Let $\mathbb{U}^2$ be the upper half plane $\mathbb{U}^2 = \{(x,y)\big|y > 0\}$. Let C be the Euclidean circle $C = C_E\big((0,1),\sqrt{2}\big)$ and let L be the Euclidean line $y = 0$ (that is the x-axis). Note that F_C sends the unit circle $\mathbb{S}^1$ (with the point $(0,1)$ removed) to the x-axis with $F_C(1,0) = (1,0)$, $F_C(-1,0) = (-1,0)$, $F_C(0,-1) = (0,0)$ and $F_C(0,1)$ undefined, and F_C sends the disc $\mathbb{H}^2$ to the lower half plane $y < 0$. The composite $S = F_L F_C$ sends the disc $\mathbb{H}^2$ to the upper half plane $\mathbb{U}^2$. The inverse of S is given by $T = S^{-1} = F_C F_L$. In the upper half plane model of the hyperbolic plane, we define the distance between two points in the upper half plane in order to make the map S an isometry, so for $u, v \in \mathbb{H}^2$ we define $d_U\big(u,v\big) = d_H\big(T(u),T(v)\big)$. The maps S and T are conformal (they preserve angles between the curves), so the angles between two curves in $\mathbb{U}^2$ are equal to the Euclidean angles between the curves. The geodesics (lines which minimize distance) in $\mathbb{H}^2$ are mapped by S to the geodesics in $\mathbb{U}^2$. The geodesics in $\mathbb{H}^2$ are the straight lines through 0 and the arcs along circles which are orthogonal to $\mathbb{S}^1$, and the geodesics in $\mathbb{U}^2$ are the vertical lines and the upper half circles which intersect orthogonally with the x–axis.

The **Minkowski model**, also called the **hyperboloid model**, of the hyperbolic plane is one half of a hyperboloid in 3-dimensional **Minkowski space**. The 3-dimensional Minkowski space is the set $\mathbb{R}^3$ using a different norm. The standard Euclidean quadratic form in $\mathbb{R}^3$ (the square of the norm) is given by $N(x,y,z) = x^2+y^2+z^2$, and the Minkowski quadratic form in $\mathbb{R}^3$ is given by $Q(x,y,t) = x^2+y^2-t^2$ (which can take negative values). Let $\mathbb{M}^2$ be the upper sheet of the hyperboloid $Q(x,y,t) = -1$, that is let

$$\mathbb{M}^2 = \big\{(x,y,t) \in \mathbb{R}^3 \,\big|\, x^2+y^2-t^2 = -1\,,\, t > 0\big\}.$$

We can define a projection $S : \mathbb{M}^2 \to \mathbb{H}^2$, similar to the stereographic projection, as follows: given $(x,y,t) \in \mathbb{M}^2$, we let $(u,v) \in \mathbb{H}^2$ be the point such that the line in $\mathbb{R}^3$ through $(0,0,-1)$ and (x,y,t) intersects the xy-plane at the point $(u,v,0)$. In the Minkowski model of the hyperbolic plane, we define the distance between two points in $\mathbb{M}^2$ in order to make the map S an isometry. The geodesics in $\mathbb{M}^2$ are the curves of intersection of $\mathbb{M}^2$ with a plane in $\mathbb{R}^3$ through the origin.

The **Klein model** of the hyperbolic plane is constructed as follows. Let $\mathbb{K}^2$ be the unit disc $\mathbb{K}^2 = \{x,y)\big|\, x^2+y^2 < 1\}$ (so in fact $\mathbb{K}^2 = \mathbb{H}^2$, but the distance between two points in $\mathbb{K}^2$ is not the same as the distance between the same two points in $\mathbb{H}^2$ or between the same two points in $\mathbb{R}^2$). Define another projection $S : \mathbb{M}^2 \to \mathbb{K}^2$, similar to the stereographic projection, as follows: given $(x,y,t) \in \mathbb{M}^2$ we let $(u,v) \in \mathbb{K}^2$ be the point such that the line in $\mathbb{R}^3$ through $(0,0,0)$ and (x,y,t) intersects the plane $z = 1$ at the point $(u,v,1)$. We define distance in $\mathbb{K}^2$ so that the map S is an isometry. The geodesics in $\mathbb{K}^2$ are segments along straight lines (but the hyperbolic angle between two lines in $\mathbb{K}^2$ is not the same as the hyperbolic angle between the same two lines in $\mathbb{R}^2$).

Tilings

This topic will not be covered (but I may include notes later).