Geometry

Richard Earl

Michaelmas Term 2018

SYLLABUS

Euclidean geometry in two and three dimensions approached by vectors and coordinates. Vector addition and scalar multiplication. The scalar product, equations of planes, lines and circles. [3]

The vector product in three dimensions. Use of $\mathbf{a}, \mathbf{b}, \mathbf{a} \wedge \mathbf{b}$ as a basis. $\mathbf{r} \wedge \mathbf{a} = \mathbf{b}$ represents a line. Scalar triple products and vector triple products, vector algebra. [2]

Conics (normal form only), focus and directrix. Showing the locus $Ax^2 + Bxy + Cy^2 = 1$ can be put in normal form via a rotation matrix. Simple examples identifying conics not in normal form. Orthogonal matrices. 2×2 orthogonal matrices and the maps they represent. Orthonormal bases in $\mathbb{R}^3$. Orthogonal change of variable; $A\mathbf{u} \cdot A\mathbf{v} = \mathbf{u} \cdot \mathbf{v}$ and $A(\mathbf{u} \wedge \mathbf{v}) = \pm A\mathbf{u} \wedge A\mathbf{v}$. Statement that a real symmetric matrix can be orthogonally diagonalized. [3]

3×3 orthogonal matrices; $SO(3)$ and rotations; conditions for being a reflection. Isometries of $\mathbb{R}^3$. [2]

Rotating frames in 2 and 3 dimensions. Angular velocity. $\mathbf{v} = \omega \wedge \mathbf{r}$. [1]

Parametrized surfaces, including spheres, cones. Examples of co-ordinate systems including parabolic, spherical and cylindrical polars. Calculating normal as $\mathbf{r}_u \wedge \mathbf{r}_v$. Surface area. Isometries preserve area. [4]

SUGGESTED READING

John Roe, *Elementary Geometry*, Oxford University Press (1993), Chapters 1, 2.2, 3.4, 4, 5.3, 7.1–7.3, 8.1–8.3, 12.1.
Richard Earl, *Towards Higher Mathematics: A Companion,* Cambridge University Press, (2017), Chapters 3.1, 3.2, 3.7, 3.10, 4.2, 4.3

FURTHER READING

3) Roger Fenn, *Geometry,* Springer-Verlag (2007)

1. Euclidean Geometry

1.1 Coordinate Space and the Algebra of Vectors

Definition 1 *By a* ***vector*** *we will mean a list of* n *real numbers* $x_1, x_2, x_3, \ldots, x_n$ *where* n *is a positive integer. Mostly this list will be treated as a* ***row vector*** *and written as*

$$(x_1, x_2, \ldots, x_n).$$

Sometimes (for reasons that will become apparent) the numbers will be arranged as a ***column vector***

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Often we will denote such a vector by a single letter in bold, say $\mathbf{x}$*, and refer to* x_i *as the* i***th co-ordinate*** *of* $\mathbf{x}$*.*

Definition 2 *For a given* n*, we denote the set of all vectors with* n *co-ordinates as* $\mathbb{R}^n$*, and often refer to* $\mathbb{R}^n$ *as* n***-dimensional co-ordinate space*** *or simply as* n***-dimensional space****. If* $n = 2$ *then we commonly use* x *and* y *as co-ordinates and refer to* $\mathbb{R}^2 = \{(x, y) \colon x, y \in \mathbb{R}\}$ *as the* xy***-plane****. If* $n = 3$ *then we commonly use* x, y *and* z *as co-ordinates and refer to* $\mathbb{R}^3 = \{(x, y, z) \colon x, y, z \in \mathbb{R}\}$ *as* xyz***-space****.*

- Note that the order of the co-ordinates matters; so, for example, $(2, 3)$ and $(3, 2)$ are different vectors in $\mathbb{R}^2$.

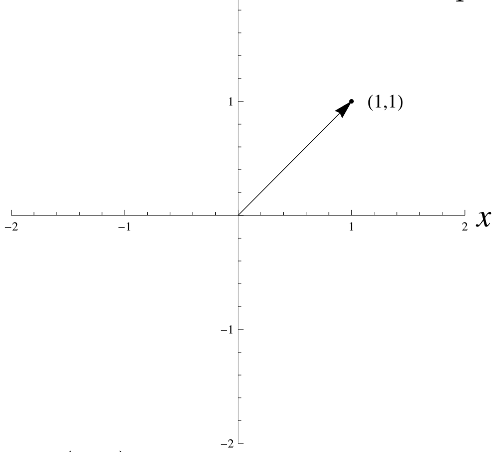
The geometry of this course relates mainly to the plane or three-dimensional space. But it can be useful to develop the theory more generally (especially when it comes to solving systems of linear equations in *Linear Algebra I*).

Definition 3 *There is a special vector* $(0, 0, \ldots, 0)$ *in* $\mathbb{R}^n$ *which we denote as* $\mathbf{0}$ *and refer to as the* ***zero vector.***

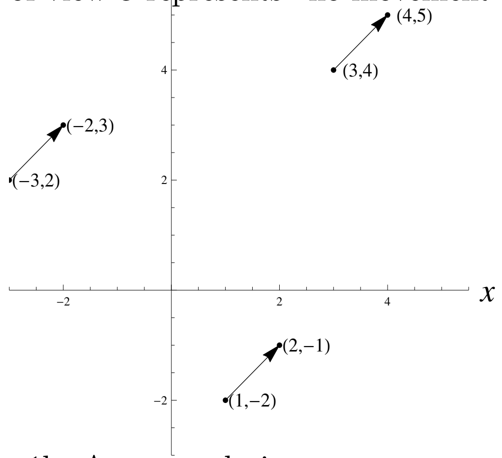
A vector can be thought of in two different ways. For example, in the case of vectors in $\mathbb{R}^2$, the xy-plane:

- From one point of view, a vector can represent the point in $\mathbb{R}^2$ which has co-ordinates x and y. We call this vector the **position vector** of that point. In practice though, rather than referring to the "point with position vector $\mathbf{x}$", we will simply say "the point $\mathbf{x}$" when the meaning is clear. The point $\mathbf{0}$ is referred to as the **origin**.

- From a second point of view, a vector is a 'movement' or translation. For example, to get from the point $(3,4)$ to the point $(4,5)$ we need to move 'one to the right and one up'; this is the same movement as is required to move from $(-3,2)$ to $(-2,3)$ and from $(1,-2)$ to $(2,-1)$. Thinking about vectors from this second point of view, all three of these movements are the same vector, because the same translation 'one right, one up' achieves each of them, even though the 'start' and 'finish' are different in each case. We would write this vector as $(1,1)$. Vectors from this second point of view are sometimes called **translation vectors**. From this point of view $\mathbf{0}$ represents "no movement".



1a. $(1,1)$ as a position vector



1b. As a translation vector

Definition 4 *The points* $(0,0,\ldots,0,x_i,0,\ldots,0)$ *in* $\mathbb{R}^n$*, where* x_i *is a real number, comprise the* x_i***-axis****, with the origin lying at the intersection of all the axes.*

Similarly in three (and likewise higher) dimensions, the triple (x,y,z) can be thought of as the point in $\mathbb{R}^3$ which is x units along the x-axis from the origin, y units parallel to the y-axis and z units parallel to the z-axis, or it can represent the translation which would take the origin to that point.

Definition 5 *Given two vectors* $\mathbf{u} = (u_1,u_2,\ldots,u_n)$ *and* $\mathbf{v} = (v_1,v_2,\ldots,v_n)$ *in* $\mathbb{R}^n$*, we can add and subtract them much as you would expect, by separately adding the corresponding co-ordinates. That is*

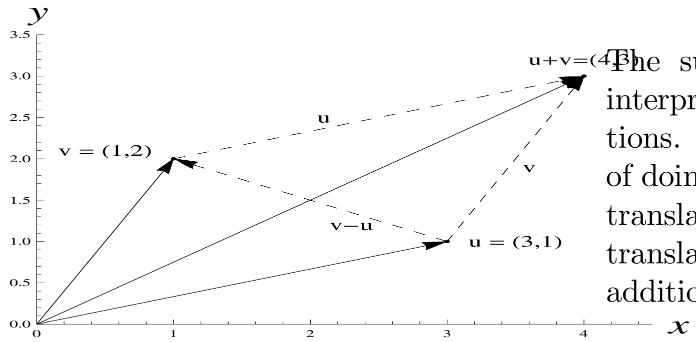
$$\mathbf{u}+\mathbf{v} = (u_1+v_1, u_2+v_2, \ldots, u_n+v_n); \qquad \mathbf{u}-\mathbf{v} = (u_1-v_1, u_2-v_2, \ldots, u_n-v_n).$$

Note that $\mathbf{v}-\mathbf{u}$ *is the vector that translates the point (with position vector)* $\mathbf{u}$ *to the point (with position vector)* $\mathbf{v}$.

- Note that two vectors may be added if and only if they have the same number of co-ordinates. No immediate sense can be made of adding a vector in $\mathbb{R}^2$ to one from $\mathbb{R}^3$, for example.

The diagram to the left shows $\mathbf{u} + \mathbf{v}$ and $\mathbf{v} - \mathbf{u}$ for particular choices of $\mathbf{u}, \mathbf{v}$ in the xy-plane.

The sum of two vectors is perhaps easiest to interpret when we consider the vectors as translations. The translation $\mathbf{u} + \mathbf{v}$ is the overall effect of doing the translation $\mathbf{u}$ first and then doing the translation $\mathbf{v}$ or it can be achieved by doing the translations in the other order – that is, vector addition is *commutative*: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.



2. Showing $\mathbf{v} - \mathbf{u}$ and $\mathbf{u} + \mathbf{v}$

Definition 6 *Given a vector* $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ *and a real number* k *then the* ***scalar multiple*** $k\mathbf{v}$ *is defined as*

$$k\mathbf{v} = (kv_1, kv_2, \ldots, kv_n).$$

- *When* k *is a positive integer, then we can think of* $k\mathbf{v}$ *as the translation achieved when we translate* k *times by the vector* $\mathbf{v}$*.*

- *Note that the points* $k\mathbf{v}$*, as* k *varies through the real numbers, make up the line which passes through the origin and the point* $\mathbf{v}$*. The points* $k\mathbf{v}$*, where* $k > 0$*, lie on one half-line from the origin, the half which includes the point* $\mathbf{v}$*. And the points* $k\mathbf{v}$*, where* $k < 0$*, comprise the remaining half-line.*

- *We write* $-\mathbf{v}$ *for* $(-1)\mathbf{v} = (-v_1, -v_2, \ldots, -v_n)$*. Translating by* $-\mathbf{v}$ *is the inverse operation of translating by* $\mathbf{v}$*.*

Definition 7 *The* n *vectors*

$$(1, 0, \ldots, 0), \quad (0, 1, 0, \ldots, 0), \quad \cdots \quad (0, \ldots, 0, 1, 0), \quad (0, \ldots, 0, 1)$$

in $\mathbb{R}^n$ *are known as the* ***standard*** *(or* ***canonical)*** *basis for* $\mathbb{R}^n$*. We will denote these, respectively, as* $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$*.*

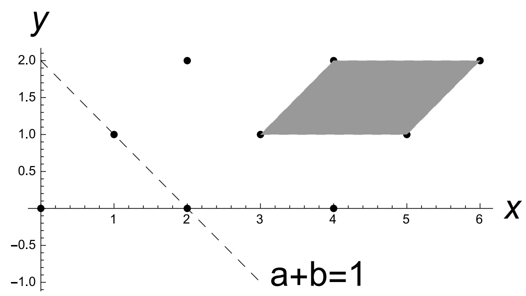
When $n = 2$*, the vectors* $(1, 0)$ *and* $(0, 1)$ *form the standard basis for* $\mathbb{R}^2$*. These are also commonly denoted by the symbols* $\mathbf{i}$ *and* $\mathbf{j}$ *respectively. Note that any vector* $\mathbf{v} = (x, y)$ *can be written uniquely as a linear combination of* $\mathbf{i}$ *and* $\mathbf{j}$*: that is* $(x, y) = x\mathbf{i} + y\mathbf{j}$ *and this is the only way to write* (x, y) *as a sum of scalar multiples of* $\mathbf{i}$ *and* $\mathbf{j}$*. When* $n = 3$*, the vectors* $(1, 0, 0)$*,* $(0, 1, 0)$*,* $(0, 0, 1)$ *form the standard basis for* $\mathbb{R}^3$*, being respectively denoted* $\mathbf{i}$*,* $\mathbf{j}$*,* $\mathbf{k}$*.*

Example 8 *Let* $\mathbf{v} = (1, 1)$ *and* $\mathbf{w} = (2, 0)$ *in* $\mathbb{R}^2$*. (a) Label the nine points* $a\mathbf{v} + b\mathbf{w}$ *where* a *and* b *are one of* $0, 1, 2$*. (b) Sketch the points* $a\mathbf{v} + b\mathbf{w}$ *where* $a + b = 1$*. (c) Shade the region of points* $a\mathbf{v} + b\mathbf{w}$ *where* $1 \leqslant a, b \leqslant 2$*.*

Solution (a) The nine points are $(0,0)$, $(2,0)$, $(4,0)$, $(1,1)$, $(3,1)$, $(5,1)$, $(2,2)$, $(4,2)$, $(6,2)$. Notice that the points $a\mathbf{v}+b\mathbf{w}$ where a and b are integers, make a lattice of parallelograms in $\mathbb{R}^2$

(b) Points $a\mathbf{v}+b\mathbf{w}$ with $a+b=1$ have the form $a\mathbf{v}+(1-a)\mathbf{w}=(2-a,a)$ which is a general point of the line $x+y=2$.

(c) The four edges of the shaded parallelogram lie on the lines $b=1$ (left edge), $b=2$ (right), $a=1$ (bottom), $a=2$ (top).



3. Solution to Example 8

■

1.2 The Geometry of Vectors. Some Geometric Theory.

As vectors represent geometric ideas like points and translations, they have important geometric properties as well as algebraic ones.

Definition 9 *The* ***length*** *(or* ***magnitude****) of a vector* $\mathbf{v}=(v_1,v_2,\ldots,v_n)$*, which is written* $|\mathbf{v}|$*, is defined by*

$$|\mathbf{v}|=\sqrt{(v_1)^2+(v_2)^2+\cdots+(v_n)^2}.$$

We say a vector $\mathbf{v}$ *is a* ***unit vector*** *if it has length* 1*.*

This formula is exactly what you'd expect it to be from Pythagoras' Theorem; we see this is the distance of the point $\mathbf{v}$ from the origin, or equivalently the distance a point moves when it is translated by $\mathbf{v}$. So if $\mathbf{p}$ and $\mathbf{q}$ are points in $\mathbb{R}^n$, then the vector that will translate $\mathbf{p}$ to $\mathbf{q}$ is $\mathbf{q}-\mathbf{p}$, and hence we define:

Definition 10 *The* ***distance*** *between two points* $\mathbf{p}$*,* $\mathbf{q}$ *in* $\mathbb{R}^n$ *is* $|\mathbf{q}-\mathbf{p}|$ *(or equally* $|\mathbf{p}-\mathbf{q}|$*). In terms of their co-ordinates* p_i *and* q_i *we have*

$$\text{distance between } \mathbf{p} \text{ and } \mathbf{q}=\sqrt{\sum_{i=1}^{n}(p_i-q_i)^2}.$$

- Note that $|\mathbf{v}|\geqslant 0$ and that $|\mathbf{v}|=0$ if and only if $\mathbf{v}=\mathbf{0}$.
- Also $|\lambda\mathbf{v}|=|\lambda|\,|\mathbf{v}|$ for any real number λ.

Proposition 11 *(**Triangle Inequality**) Let* $\mathbf{u}, \mathbf{v}$ *vectors in* $\mathbb{R}^n$*. Then*

$$|\mathbf{u}+\mathbf{v}| \leqslant |\mathbf{u}| + |\mathbf{v}| . \tag{1.1}$$

If $\mathbf{v} \neq \mathbf{0}$ *then there is equality in (1.1) if and only if* $\mathbf{u} = \lambda \mathbf{v}$ *for some* $\lambda \geqslant 0$*.*

Remark 12 *Geometrically, this is intuitively obvious. If we review Figure 2 and look at the triangle with vertices* $\mathbf{0}, \mathbf{u}, \mathbf{u}+\mathbf{v}$*, then we see that its sides have lengths* $|\mathbf{u}|, |\mathbf{v}|$ *and* $|\mathbf{u}+\mathbf{v}|$*. So* $|\mathbf{u}+\mathbf{v}|$ *is the distance along the straight line from* $\mathbf{0}$ *to* $\mathbf{u}+\mathbf{v}$*, whereas* $|\mathbf{u}|+|\mathbf{v}|$ *is the combined distance from* $\mathbf{0}$ *to* $\mathbf{u}$ *to* $\mathbf{u}+\mathbf{v}$*. This cannot be shorter and will only be equal if we passed through* $\mathbf{u}$ *on the way to* $\mathbf{u}+\mathbf{v}$*.*

Proof Let $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, $\mathbf{v} = (v_1, v_2, \ldots, v_n)$. The inequality (1.1) is trivial if $\mathbf{v} = \mathbf{0}$, so suppose $\mathbf{v} \neq \mathbf{0}$. Note that for any real number t,

$$0 \leqslant |\mathbf{u}+t\mathbf{v}|^2 = \sum_{i=1}^{n} (u_i + tv_i)^2 = |\mathbf{u}|^2 + 2t \sum_{i=1}^{n} u_i v_i + t^2 |\mathbf{v}|^2 .$$

As $|\mathbf{v}| \neq 0$, the RHS of the above inequality is a quadratic in t which is always non-negative, and so has non-positive discriminant ($b^2 \leqslant 4ac$). Hence

$$4\left(\sum_{i=1}^{n} u_i v_i\right)^2 \leqslant 4 |\mathbf{u}|^2 |\mathbf{v}|^2 \qquad \text{giving} \qquad \left|\sum_{i=1}^{n} u_i v_i\right| \leqslant |\mathbf{u}| |\mathbf{v}| . \tag{1.2}$$

Finally

$$|\mathbf{u}+\mathbf{v}|^2 = |\mathbf{u}|^2 + 2\sum_{i=1}^{n} u_i v_i + |\mathbf{v}|^2 \leqslant |\mathbf{u}|^2 + 2\left|\sum_{i=1}^{n} u_i v_i\right| + |\mathbf{v}|^2 \leqslant |\mathbf{u}|^2 + 2|\mathbf{u}||\mathbf{v}| + |\mathbf{v}|^2 = (|\mathbf{u}|+|\mathbf{v}|)^2$$

to give (1.1). We have equality in $b^2 \leqslant 4ac$ if and only if the quadratic $|\mathbf{u}+t\mathbf{v}|^2 = 0$ has a repeated real solution, say $t = t_0$. So $\mathbf{u} + t_0\mathbf{v} = \mathbf{0}$ and we see that $\mathbf{u}$ and $\mathbf{v}$ are multiples of one another. This is for equality to occur in (1.2). With $\mathbf{u} = -t_0\mathbf{v}$, then equality in

$$\sum_{i=1}^{n} u_i v_i = \left|\sum_{i=1}^{n} u_i v_i\right| \text{ means } -t_0 |\mathbf{v}|^2 = |t_0| |\mathbf{v}|^2$$

which occurs when $-t_0 \geqslant 0$, as $\mathbf{v} \neq \mathbf{0}$. ■

Definition 13 *Given two vectors* $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ *in* $\mathbb{R}^n$*, the* ***scalar product*** $\mathbf{u} \cdot \mathbf{v}$*, also known as the* ***dot product*** *or* ***Euclidean inner product****, is defined as the real number*

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n .$$

We then read $\mathbf{u} \cdot \mathbf{v}$ *as "u dot v"; we also often use 'dot' as a verb in this regard.*

The following properties of the scalar product are easy to verify and left as exercises. Note (e) was proved in (1.2).

Proposition 14 *Let* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *be vectors in* $\mathbb{R}^n$ *and let* λ *be a real number. Then*
(a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$.
(b) $(\lambda \mathbf{u}) \cdot \mathbf{v} = \lambda(\mathbf{u} \cdot \mathbf{v})$.
(c) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$.
(d) $\mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2 \geqslant 0$ *and* $\mathbf{u} \cdot \mathbf{u} = 0$ *if and only if* $\mathbf{u} = \mathbf{0}$.
(e) ***Cauchy-Schwarz Inequality***

$$|\mathbf{u} \cdot \mathbf{v}| \leqslant |\mathbf{u}| \, |\mathbf{v}| \tag{1.3}$$

with equality when one of $\mathbf{u}$ *and* $\mathbf{v}$ *is a multiple of the other.*

We see that the length of $\mathbf{u}$ can be written in terms of the scalar product, namely as

$$|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}}.$$

We can also define the *angle* between two vectors in terms of their scalar product.

Definition 15 *Given two non-zero vectors* $\mathbf{u}, \mathbf{v}$ *in* $\mathbb{R}^n$ *the* ***angle*** *between them is given by the expression*

$$\cos^{-1}\left(\frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}| \, |\mathbf{v}|}\right).$$

- The above formula makes sense as $|\mathbf{u} \cdot \mathbf{v}|/(|\mathbf{u}||\mathbf{v}|) \leqslant 1$ by the Cauchy-Schwarz inequality. If we take the principal values of $\cos^{-1}$ to be in the range $0 \leqslant \theta \leqslant \pi$ the formula measures the smaller angle between the vectors.

- Given two vectors $\mathbf{u}$ and $\mathbf{v}$ with angle θ between them, an equivalent definition of the scalar product $\mathbf{u} \cdot \mathbf{v}$ is then
$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}| \, |\mathbf{v}| \cos \theta. \tag{1.4}$$

- Note that two vectors $\mathbf{u}$ and $\mathbf{v}$ are perpendicular if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.

- There is an obvious concern, that Definition 15 ties in with our usual notion of angle. Given two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$ we might choose xy-co-ordinates in the plane spanned by $\mathbf{u}$ and $\mathbf{v}$ with the x-axis pointing in the direction of $\mathbf{u}$. We then have
$$\mathbf{u} = (|\mathbf{u}|, 0), \qquad \mathbf{v} = (|\mathbf{v}| \cos \theta, |\mathbf{v}| \sin \theta),$$
where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$. Then $\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}| \, |\mathbf{v}| \cos \theta$ as expected.

Example 16 *Let* $\mathbf{u} = (1, 2, -1)$ *and* $\mathbf{v} = (0, 2, 3)$ *in* $\mathbb{R}^3$. *Find the lengths of* $\mathbf{u}$ *and* $\mathbf{v}$ *and the angle* θ *between them.*

Solution We have

$$\begin{aligned} |\mathbf{u}|^2 &= \mathbf{u} \cdot \mathbf{u} = 1^2 + 2^2 + (-1)^2 = 6, \quad \text{giving} \quad |\mathbf{u}| = \sqrt{6}; \\ |\mathbf{v}|^2 &= \mathbf{v} \cdot \mathbf{v} = 0^2 + 2^2 + 3^2 = 13, \quad \text{giving} \quad |\mathbf{v}| = \sqrt{13}; \\ \mathbf{u} \cdot \mathbf{v} &= 1 \times 0 + 2 \times 2 + (-1) \times 3 = 1, \end{aligned}$$

giving $\theta = \cos^{-1}\left(\frac{1}{\sqrt{6}\sqrt{13}}\right) = \cos^{-1} \frac{1}{\sqrt{78}} \approx 1.457$ radians. ■

Theorem 17 *(**Cosine Rule**) Consider a triangle with sides of length a, b, c with opposite angle α, β, γ respectively. Then*

$$a^2 = b^2 + c^2 - 2bc\cos\alpha; \qquad b^2 = a^2 + c^2 - 2ac\cos\beta; \qquad c^2 = a^2 + b^2 - 2ab\cos\gamma.$$

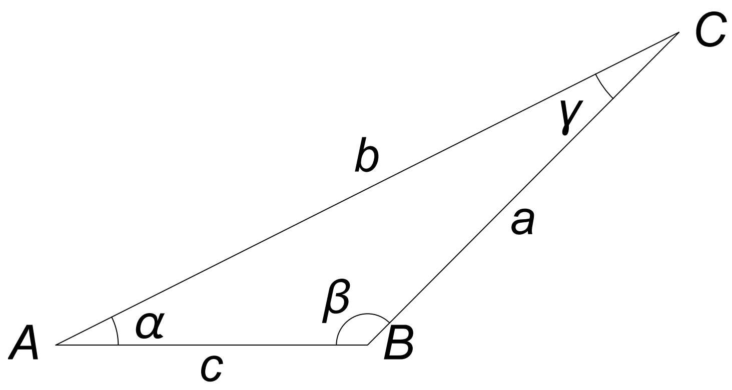
Proof We'll call the vertices at the angles α, β, γ respectively A, B, C. Set

$$\mathbf{u} = \overrightarrow{AB}, \text{ and } \mathbf{v} = \overrightarrow{AC}$$

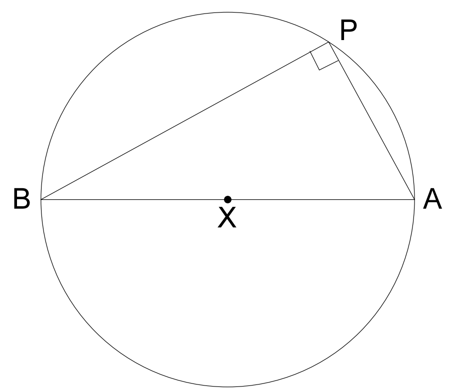
so that $\mathbf{u} - \mathbf{v} = \overrightarrow{BC}$, $c = |\mathbf{u}|$ and $b = |\mathbf{v}|$. Then

$$\begin{aligned} a^2 &= \left|\overrightarrow{BC}\right|^2 = |\mathbf{u}-\mathbf{v}|^2 = (\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}) = \mathbf{u}\cdot\mathbf{u} + \mathbf{v}\cdot\mathbf{v} - 2\mathbf{u}\cdot\mathbf{v} \\ &= |\mathbf{u}|^2 + |\mathbf{v}|^2 - 2|\mathbf{u}||\mathbf{v}|\cos\alpha = \left|\overrightarrow{AB}\right|^2 + \left|\overrightarrow{AC}\right|^2 - 2\left|\overrightarrow{AB}\right|\left|\overrightarrow{AC}\right|\cos\alpha = c^2 + b^2 - 2cb\cos\alpha, \end{aligned}$$

and the other two equations follow similarly. ■



4. Cosine Rule



5. Thales Theorem

Theorem 18 *(**Thales Theorem**). Let A and B be points at opposite ends of the diameter of a circle, and let P be a third point. Then $\angle APB$ is a right angle if and only if P also lies on the circle.*

Proof Let X be the centre of the circle and set $\mathbf{a} = \overrightarrow{XA}$, so that $-\mathbf{a} = \overrightarrow{XB}$; the radius of the circle is then $|\mathbf{a}|$. Also set $\mathbf{p} = \overrightarrow{XP}$. Then the angle $\angle APB$ is a right angle if and only if

$$\begin{aligned} &\overrightarrow{AP} \perp \overrightarrow{BP} \iff (\mathbf{p}-\mathbf{a})\cdot(\mathbf{p}+\mathbf{a}) = 0 \iff \mathbf{p}\cdot\mathbf{p} - \mathbf{a}\cdot\mathbf{a} = 0 \\ &\iff |\mathbf{p}|^2 = |\mathbf{a}|^2 \iff |\mathbf{p}| = |\mathbf{a}| \iff |XP| = \text{circle's radius}, \end{aligned}$$

and the result follows. [In a similar fashion to the proof of the cosine rule, taking position vectors from X, as we did from A previously, simplifies the vector identities involved somewhat.] ■

Theorem 19 *The medians of a triangle are concurrent at its* ***centroid.***

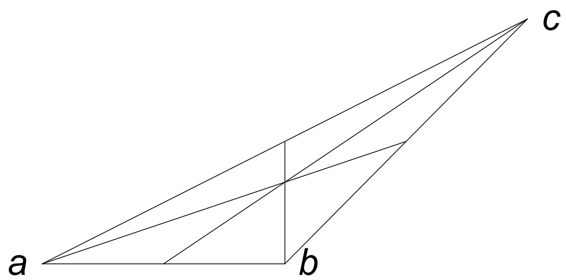
Proof A median of a triangle is a line connecting a vertex to the midpoint of the opposite edge. Let A, B, C denote the vertices of the triangle with position vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ from some origin O. Then the midpoints of the triangle's edges are

$$\frac{\mathbf{a}+\mathbf{b}}{2}, \qquad \frac{\mathbf{b}+\mathbf{c}}{2}, \qquad \frac{\mathbf{c}+\mathbf{a}}{2}.$$

The line connecting two points with position vectors $\mathbf{p}$ and $\mathbf{q}$ consists of those points with position vectors $\lambda\mathbf{p}+\mu\mathbf{q}$ where $\lambda+\mu=1$. So note that

$$\frac{\mathbf{a}+\mathbf{b}+\mathbf{c}}{3} = \frac{1}{3}\mathbf{a}+\frac{2}{3}\left(\frac{\mathbf{b}+\mathbf{c}}{2}\right) = \frac{1}{3}\mathbf{b}+\frac{2}{3}\left(\frac{\mathbf{a}+\mathbf{c}}{2}\right) = \frac{1}{3}\mathbf{c}+\frac{2}{3}\left(\frac{\mathbf{a}+\mathbf{b}}{2}\right).$$

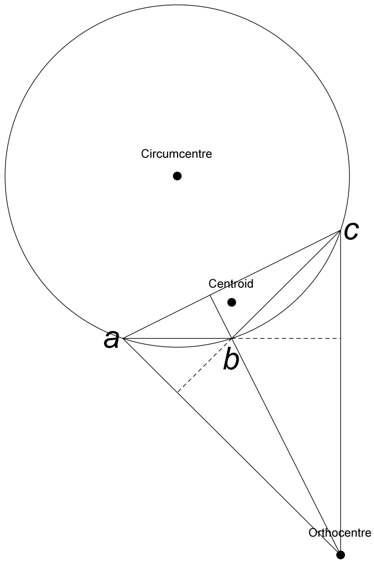
The first equality shows that the triangle's *centroid*, the point with position vector $(\mathbf{a}+\mathbf{b}+\mathbf{c})/3$, lies on the median connecting A and the midpoint of BC with the other equalities showing that the centroid lies on the other two medians. ■



6. The Medians and Centroid

Example 20 *Let ABC be a triangle. Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be the position vectors of A, B, C from an origin O.*
(a) Write down the position vector of the centroid of the triangle.
(b) Let D be the intersection of the altitude from A with the altitude from B. Show that, in fact, all three altitudes intersect at D. The point D is called the ***orthocentre*** *of the triangle ABC.*
(c) Show that if we take D as the origin for the plane, then $\mathbf{a}\cdot\mathbf{b}=\mathbf{b}\cdot\mathbf{c}=\mathbf{c}\cdot\mathbf{a}$.
(d) Show (with the orthocentre as origin) that the triangle's ***circumcentre*** *has position vector $(\mathbf{a}+\mathbf{b}+\mathbf{c})/2$.*

Deduce that the centroid, circumcentre and orthocentre of a triangle are collinear. The line on which they lie is called the ***Euler Line.***



7. Three centres of a triangle

Solution (a) Let ABC be a triangle with $\mathbf{a}, \mathbf{b}, \mathbf{c}$ as the position vectors of A, B, C from an origin O. The centroid has position vector $(\mathbf{a}+\mathbf{b}+\mathbf{c})/3$, whatever the choice of origin.
(b) and (c) Let D be the intersection of the altitude from A with the altitude from B. If we take D as our origin then we note

$$\mathbf{a}\cdot(\mathbf{c}-\mathbf{b}) = 0; \qquad \mathbf{b}\cdot(\mathbf{c}-\mathbf{a}) = 0,$$

as DA is perpendicular to BC and DB is perpendicular to AC. Hence

$$\mathbf{a}\cdot\mathbf{c} = \mathbf{b}\cdot\mathbf{a} = \mathbf{b}\cdot\mathbf{c},$$

and in particular $\mathbf{c}\cdot(\mathbf{b}-\mathbf{a}) = 0$, showing that DC is perpendicular to AB, i.e. that D also lies on the altitude from C.
(d) With the orthocentre still as the origin, we set $\mathbf{p} = (\mathbf{a}+\mathbf{b}+\mathbf{c})/2$ and $\rho = \mathbf{a}\cdot\mathbf{c} = \mathbf{b}\cdot\mathbf{a} = \mathbf{b}\cdot\mathbf{c}$. Then

$$\begin{aligned}
|\mathbf{p}-\mathbf{a}|^2 &= \frac{1}{4}(\mathbf{b}+\mathbf{c}-\mathbf{a})\cdot(\mathbf{b}+\mathbf{c}-\mathbf{a}) \\
&= \frac{|\mathbf{a}|^2+|\mathbf{b}|^2+|\mathbf{c}|^2+2\mathbf{b}\cdot\mathbf{c}-2\mathbf{a}\cdot\mathbf{c}-2\mathbf{a}\cdot\mathbf{b}}{4} \\
&= \frac{|\mathbf{a}|^2+|\mathbf{b}|^2+|\mathbf{c}|^2-2\rho}{4}.
\end{aligned}$$

This expression is symmetric in $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ so that

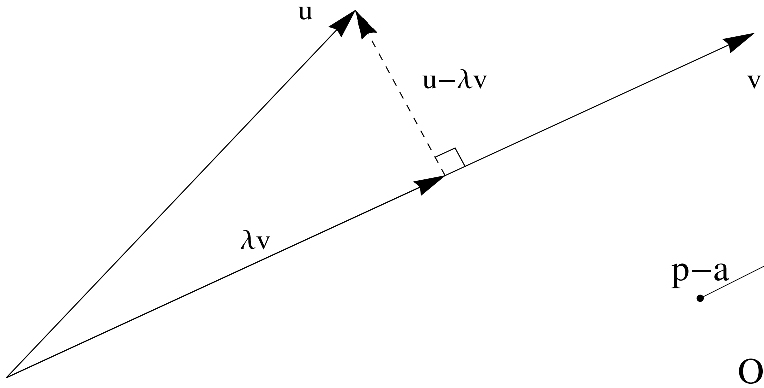
$$|\mathbf{p}-\mathbf{a}|^2 = |\mathbf{p}-\mathbf{b}|^2 = |\mathbf{p}-\mathbf{c}|^2.$$

This means that $\mathbf{p}$ is the circumcentre. So (with the orthocentre as the origin) the circumcentre at $(\mathbf{a}+\mathbf{b}+\mathbf{c})/2$, the centroid at $(\mathbf{a}+\mathbf{b}+\mathbf{c})/3$ and the orthocentre at $\mathbf{0}$ are seen to be collinear.
■

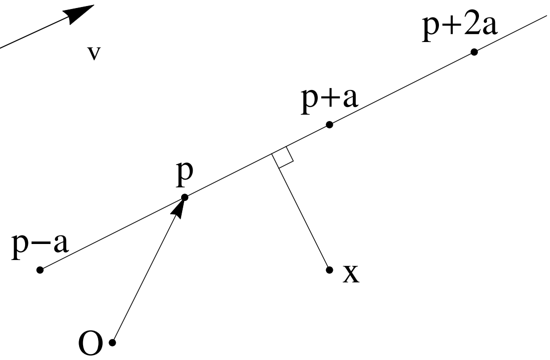
Example 21 *Let* $\mathbf{u}$ *and* $\mathbf{v}$ *be vectors in* $\mathbb{R}^n$ *with* $\mathbf{v} \neq \mathbf{0}$*. Show there is a unique real number* λ *such that* $\mathbf{u} - \lambda\mathbf{v}$ *is perpendicular to* $\mathbf{v}$*. Then*

$$\mathbf{u} = \lambda\mathbf{v} + (\mathbf{u} - \lambda\mathbf{v}),$$

with the vector $\lambda\mathbf{v}$ *being called the* ***component of*** $\mathbf{u}$ ***in the direction of*** $\mathbf{v}$ *and* $\mathbf{u} - \lambda\mathbf{v}$ *being called the* ***component of*** $\mathbf{u}$ ***perpendicular to the direction of*** $\mathbf{v}$.



8. Components of a vector $\mathbf{u}$ relative to $\mathbf{v}$



9. A line described parametrically

Solution We have $\mathbf{u} - \lambda\mathbf{v}$ is perpendicular to $\mathbf{v}$ if and only if

$$(\mathbf{u} - \lambda\mathbf{v}) \cdot \mathbf{v} = 0 \quad \iff \quad \mathbf{u} \cdot \mathbf{v} = \lambda \left|\mathbf{v}\right|^2 \quad \iff \quad \lambda = (\mathbf{u} \cdot \mathbf{v}) / \left|\mathbf{v}\right|^2 .$$

Note that $|\lambda\mathbf{v}| = |\mathbf{u}| \cos\theta$, where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$, as one would expect. ■

Example 22 ***(Parametric Form of a Line)*** *Let* $\mathbf{p}$, $\mathbf{a}$, $\mathbf{x}$ *be vectors in* $\mathbb{R}^n$ *with* $\mathbf{a} \neq \mathbf{0}$*. Explain why the points*

$$\mathbf{r}(\lambda) = \mathbf{p} + \lambda\mathbf{a} \quad \textit{where } \lambda \textit{ varies over all real numbers}$$

comprise a line through $\mathbf{p}$*, parallel to* $\mathbf{a}$*. Show that* $|\mathbf{x} - \mathbf{r}(\lambda)|$ *is minimal when* $\mathbf{x} - \mathbf{r}(\lambda)$ *is perpendicular to* $\mathbf{a}$*.*

Proof As commented following Definition 6, the points $\lambda\mathbf{a}$ comprise the line which passes through the origin and the point $\mathbf{a}$. So the points $\mathbf{p} + \lambda\mathbf{a}$ comprise the translation of that line by the vector $\mathbf{p}$; that is, they comprise the line through the point $\mathbf{p}$ parallel to $\mathbf{a}$. We also have

$$\begin{aligned}
\left|\mathbf{x} - \mathbf{r}(\lambda)\right|^2 &= (\mathbf{x} - \mathbf{r}(\lambda)) \cdot (\mathbf{x} - \mathbf{r}(\lambda)) \\
&= \mathbf{x} \cdot \mathbf{x} - 2\mathbf{r}(\lambda) \cdot \mathbf{x} + \mathbf{r}(\lambda) \cdot \mathbf{r}(\lambda) \\
&= \mathbf{x} \cdot \mathbf{x} - 2(\mathbf{p} + \lambda\mathbf{a}) \cdot \mathbf{x} + (\mathbf{p} + \lambda\mathbf{a}) \cdot (\mathbf{p} + \lambda\mathbf{a}) \\
&= (\mathbf{x} \cdot \mathbf{x} - 2\mathbf{p} \cdot \mathbf{x} + \mathbf{p} \cdot \mathbf{p}) + \lambda(-2\mathbf{a} \cdot \mathbf{x} + 2\mathbf{a} \cdot \mathbf{p}) + \lambda^2(\mathbf{a} \cdot \mathbf{a}).
\end{aligned}$$

At the minimum value of $|\mathbf{x} - \mathbf{r}(\lambda)|$ we have

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\left|\mathbf{x} - \mathbf{r}(\lambda)\right|^2\right) = 2(\mathbf{a} \cdot \mathbf{p} - \mathbf{a} \cdot \mathbf{x}) + 2\lambda(\mathbf{a} \cdot \mathbf{a}) = 0,$$

which is when $0 = (\mathbf{p} + \lambda\mathbf{a} - \mathbf{x}) \cdot \mathbf{a} = (\mathbf{r}(\lambda) - \mathbf{x}) \cdot \mathbf{a}$ as required. ■

Definition 23 *Let* $\mathbf{p}$, $\mathbf{a}$ *be vectors in* $\mathbb{R}^n$ *with* $\mathbf{a} \neq \mathbf{0}$. *Then the equation* $\mathbf{r}(\lambda) = \mathbf{p} + \lambda\mathbf{a}$, *where* λ *is a real number, is the equation of the* ***line*** *through* $\mathbf{p}$, *parallel to* $\mathbf{a}$. *It is said to be in* ***parametric form****, the parameter here being* λ. *The parameter acts as a co-ordinate on the line, uniquely associating to each point on the line a value of* λ.

Example 24 *Show that* (x, y, z) *lies on the line* $\mathbf{r}(\lambda) = (1, 2, -3) + \lambda(2, 1, 4)$ *if and only if* $(x - 1)/2 = y - 2 = (z + 3)/4$.

Solution Note that

$$\mathbf{r}(\lambda) = (x, y, z) = (1, 2, -3) + \lambda\,(2, 1, 4)\,,$$

so that

$$x = 1 + 2\lambda, \qquad y = 2 + \lambda, \qquad z = 4\lambda - 3.$$

In this case

$$\frac{x-1}{2} = y - 2 = \frac{z+3}{4} = \lambda.$$

Conversely if $(x - 1)/2 = y - 2 = (z + 3)/4$, then call their common value λ and hence we can find x, y, z in terms of λ to give the above formula for $\mathbf{r}(\lambda)$. ■

A plane can similarly be described in parametric form. Whereas just one non-zero vector $\mathbf{a}$ was needed to travel along a line $\mathbf{r}(\lambda) = \mathbf{p} + \lambda\mathbf{a}$, we will need two non-zero vectors to move around a plane. However, we need to be a little careful: if we simply considered those points $\mathbf{r}(\lambda, \mu) = \mathbf{p} + \lambda\mathbf{a} + \mu\mathbf{b}$, for non-zero vectors $\mathbf{a}, \mathbf{b}$ and parameters λ, μ, we wouldn't always get a plane. In the case when $\mathbf{a}$ and $\mathbf{b}$ were scalar multiples of one another, so that they had the same or opposite directions, then the points $\mathbf{r}(\lambda, \mu)$ would just comprise the line through $\mathbf{p}$ parallel to $\mathbf{a}$ (or equivalently $\mathbf{b}$). So we make the definitions:

Definition 25 *We say that two vectors in* $\mathbb{R}^n$ *are* ***linearly independent****, or just simply* ***independent****, if neither is a scalar multiple of the other. In particular, this means that both vectors are non-zero. Two vectors which aren't independent are said to be* ***linearly dependent.***

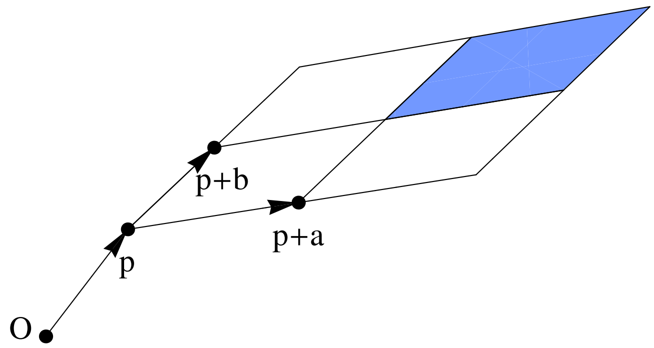
Definition 26 *(****Parametric Form of a Plane****) Let* $\mathbf{p}$, $\mathbf{a}$, $\mathbf{b}$ *be vectors in* $\mathbb{R}^n$ *with* $\mathbf{a}, \mathbf{b}$ *independent. Then*

$$\mathbf{r}(\lambda, \mu) = \mathbf{p} + \lambda\mathbf{a} + \mu\mathbf{b} \qquad \text{where } \lambda, \mu \text{ are real numbers} \tag{1.5}$$

is the equation of the ***plane*** *through* $\mathbf{p}$ *parallel to the vectors* $\mathbf{a}, \mathbf{b}$. *The parameters* λ, μ *act as co-ordinates in the plane, associating to each point of the plane a unique ordered pair* (λ, μ) *for if*

$$\mathbf{p} + \lambda_1\mathbf{a} + \mu_1\mathbf{b} = \mathbf{p} + \lambda_2\mathbf{a} + \mu_2\mathbf{b}$$

then $(\lambda_1 - \lambda_2)\,\mathbf{a} = (\mu_2 - \mu_1)\,\mathbf{b}$ *so that* $\lambda_1 = \lambda_2$ *and* $\mu_1 = \mu_2$ *by independence.*



10. A parametrized plane

Note that $\mathbf{p}$ effectively becomes the origin in the plane with co-ordinates $\lambda = \mu = 0$, that $\mathbf{p} + \mathbf{a}$ has co-ordinates $(\lambda, \mu) = (1, 0)$ and that $\mathbf{p} + \mathbf{b} = \mathbf{r}(0, 1)$.

The shaded area above comprises those points $\mathbf{r}(\lambda, \mu)$ where $1 \leqslant \lambda, \mu \leqslant 2$.

Example 27 *Given three points* $\mathbf{a}, \mathbf{b}, \mathbf{c}$ *in* $\mathbb{R}^n$ *which don't lie in a line, then* $\mathbf{b} - \mathbf{a}, \mathbf{c} - \mathbf{a}$ *are independent and we can parametrize the plane* Π *which contains the points* $\mathbf{a}, \mathbf{b}, \mathbf{c}$ *as*

$$\mathbf{r}(\lambda, \mu) = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a}) + \mu(\mathbf{c} - \mathbf{a}),$$

noting that $\mathbf{a} = \mathbf{r}(0, 0)$, $\mathbf{b} = \mathbf{r}(1, 0)$, $\mathbf{c} = \mathbf{r}(0, 1)$.

A parametric description of a plane in $\mathbb{R}^n$ may be the most natural starting point but, especially in $\mathbb{R}^3$, planes can be easily described by equations in Cartesian co-ordinates.

Proposition 28 *(**Cartesian Equation of a Plane in** $\mathbb{R}^3$) A region* Π *of* $\mathbb{R}^3$ *is a plane if and only if it can be written as*

$$\mathbf{r} \cdot \mathbf{n} = c$$

where $\mathbf{r} = (x, y, z)$, $\mathbf{n} = (n_1, n_2, n_3) \neq \mathbf{0}$ *and* c *is a real number. In terms of the co-ordinates* x, y, z *this equation reads*

$$n_1 x + n_2 y + n_3 z = c. \tag{1.6}$$

The vector $\mathbf{n}$ *is normal (i.e. perpendicular) to the plane.*

Proof Consider the equation $\mathbf{r}(\lambda, \mu) = \mathbf{p} + \lambda\mathbf{a} + \mu\mathbf{b}$ with $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ independent. Separating out the co-ordinates we then have the three scalar equations

$$x = p_1 + \lambda a_1 + \mu b_1, \qquad y = p_2 + \lambda a_2 + \mu b_2, \qquad z = p_3 + \lambda a_3 + \mu b_3,$$

where $\mathbf{r} = (x, y, z)$ and $\mathbf{p} = (p_1, p_2, p_3)$. If we eliminate λ and μ from these three equations we will be left with a single equation involving x, y, z. Omitting some messy algebra here, we eventually arrive at

$$(b_3 a_2 - b_2 a_3) x + (b_1 a_3 - b_3 a_1)\, y + (a_1 b_2 - a_2 b_1)\, z = (b_3 a_2 - b_2 a_3) p_1 + (b_1 a_3 - b_3 a_1)\, p_2 + (a_1 b_2 - a_2 b_1)\, p_3.$$

We can rewrite this much more succinctly as $\mathbf{r} \cdot \mathbf{n} = c$ where

$$\mathbf{n} = (b_3 a_2 - b_2 a_3, b_1 a_3 - b_3 a_1, a_1 b_2 - a_2 b_1) \qquad \text{and} \qquad c = \mathbf{p} \cdot \mathbf{n}.$$

[We shall meet the vector $\mathbf{n}$ again and recognize it later as the vector product $\mathbf{a} \wedge \mathbf{b}$.] Finally we can check directly that

$$\begin{aligned}\mathbf{a} \cdot \mathbf{n} &= a_1(b_3a_2 - b_2a_3) + a_2(b_1a_3 - b_3a_1) + a_3(a_1b_2 - a_2b_1) \\ &= b_1(a_2a_3 - a_3a_2) + b_2(a_3a_1 - a_1a_3) + b_3(a_1a_2 - a_2a_1) = 0,\end{aligned}$$

and similarly $\mathbf{b} \cdot \mathbf{n} = 0$. Hence $\mathbf{n}$ is normal to the plane. ■

Remark 29 *Once you have met more on linear systems you will know that the general solution of an equation*

$$\left(\begin{array}{ccc|c} n_1 & n_2 & n_3 & c \end{array} \right)$$

can be described with two parameters. If $n_1 \neq 0$ then we can assign parameters s and t to y and z respectively and see that

$$(x, y, z) = \left(\frac{c - n_2s - n_3t}{n_1}, s, t \right) = \left(\frac{c}{n_1}, 0, 0 \right) + s\left(-\frac{n_2}{n_1}, 1, 0 \right) + t\left(-\frac{n_3}{n_1}, 0, 1 \right),$$

which is a parametric description of the same plane.

Example 30 *Show that the distance of a point $\mathbf{p}$ from the plane with equation $\mathbf{r} \cdot \mathbf{n} = c$ equals*

$$\frac{|\mathbf{p} \cdot \mathbf{n} - c|}{|\mathbf{n}|}.$$

Solution The normal direction to the plane $\mathbf{r} \cdot \mathbf{n} = c$ is parallel to $\mathbf{n}$. The shortest distance from $\mathbf{p}$ to the plane is measured along the normal – so the nearest point on the plane to $\mathbf{p}$ is the form $\mathbf{p} + \lambda_0\mathbf{n}$ for some λ_0. This value is specified by the equation

$$(\mathbf{p} + \lambda_0\mathbf{n}) \cdot \mathbf{n} = c \quad \implies \quad \lambda_0 = \frac{c - \mathbf{p} \cdot \mathbf{n}}{\mathbf{n} \cdot \mathbf{n}}.$$

The distance from $\mathbf{p} + \lambda_0\mathbf{n}$ to $\mathbf{p}$ is

$$|\lambda_0\mathbf{n}| = \left| \frac{c - \mathbf{p} \cdot \mathbf{n}}{\mathbf{n} \cdot \mathbf{n}} \right| |\mathbf{n}| = \frac{|c - \mathbf{p} \cdot \mathbf{n}|}{|\mathbf{n}|}.$$

■

Example 31 *Find the orthogonal projection of the point $(1, 0, 2)$ onto the plane $2x - y + z = 1$.*

Solution A normal to the plane $\Pi : 2x - y + z = 1$ is $(2, -1, 1)$ and so the normal l to this plane which passes through $(1, 0, 2)$ is given parametrically by

$$\mathbf{r}(\lambda) = (1, 0, 2) + \lambda(2, -1, 1) = (1 + 2\lambda, -\lambda, 2 + \lambda).$$

The orthogonal projection of $(1, 0, 2)$ onto Π is $l \cap \Pi$ and l meets Π when

$$1 = 2(1 + 2\lambda) - (-\lambda) + (2 + \lambda) = 4 + 6\lambda.$$

So $\lambda = -1/2$ at $l \cap \Pi$ and $\mathbf{r}(-1/2) = (0, 1/2, 3/2)$. ■

Example 32 *Let* $(x, y, z) = (2s+t-3, 3s-t+4, s+t-2)$. *Show that, as* s *and* t *vary over the real numbers, the points* (x, y, z) *range over a plane, whose equation in the form* $ax+by+cz = d$, *you should determine.*

Solution Let $(x, y, z) = (2s + t - 3, 3s - t + 4, s + t - 2)$, so that

$$x = 2s + t - 3, \qquad y = 3s - t + 4, \qquad z = s + t - 2.$$

We aim to eliminate s and t from these equations to find a single remaining equation in x, y, z. From the third equation we have $t = z + 2 - s$ and substituting this into the first two we get

$$x = 2s + (z + 2 - s) - 3 = s + z - 1, \qquad y = 3s - (z + 2 - s) + 4 = 4s - z + 2.$$

So $s = x + 1 - z$ and substituting this into the second equation above we find

$$y = 4(x + 1 - z) - z + 2 = 4x + 6 - 5z.$$

Hence the equation of our plane is $-4x + y + 5z = 6$. ■

1.3 The Question Of Consistency

So far the geometry we have described and the definitions we have made have all taken place in $\mathbb{R}^n$. In many ways, though, $\mathbb{R}^n$ is special and so far we have glossed over this fact. Special features of $\mathbb{R}^n$ include:

- a specially chosen point, the origin;
- a set of n axes emanating from this origin;
- a notion of unit length;
- a sense of orientation — for example in $\mathbb{R}^3$, the x-, y-, and z-axes are 'right-handed'.

In the world about us none of the above is present. We might happily use nanometres, light years or feet as our chosen unit of measurement. This choice is determined by whether we wish to discuss the sub-atomic, astronomical or everyday. Similarly the world around us contains no obvious origin nor set of axes — even if there were a centre of the known universe then this would be a useless point of reference for an architect planning a new building. If we consider the scenarios we might wish to model: a camera taking a picture, a planet moving around its star, a spinning gyroscope, then typically there is a special point which stands out as a choice of origin (e.g. the camera, the star) and often there may be obvious choices for one or more of the axes (e.g. along the axis of rotation of the gyroscope). And the geometric properties present in these examples, such as length, angle, area, etc., and the geometric theory we'd wish to reason with, ought not to depend on our choice of origin and axes. For the definitions we have made, and for future ones, we will need to ensure that calculations made with respect to different sets of co-ordinates concur, provided we choose our co-ordinates appropriately. This is an important but subtle point — so in the next few sections we shall 'do' some more geometry before looking into the aspect of consistency further in §3.4.

2. The Vector Product and Vector Algebra

2.1 The Vector Product.

In $\mathbb{R}^3$, *but not generally in other dimensions*[1], together with the scalar product there is also a *vector product.*

Definition 33 *Let* $\mathbf{u} = (u_1, u_2, u_3)$ *and* $\mathbf{v} = (v_1, v_2, v_3)$ *be two vectors in* $\mathbb{R}^3$. *We define their* ***vector product*** *(or* ***cross product****)* $\mathbf{u} \wedge \mathbf{v}$ *as*

$$\mathbf{u} \wedge \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = (u_2v_3 - v_2u_3)\mathbf{i} + (u_3v_1 - v_3u_1)\mathbf{j} + (u_1v_2 - v_1u_2)\mathbf{k}. \tag{2.1}$$

$\wedge$ *is read as "vec". A common alternative notation is* $\mathbf{u} \times \mathbf{v}$ *and hence the alternative name of the cross product.*

- Note firstly that $\mathbf{u} \wedge \mathbf{v}$ is a vector (unlike $\mathbf{u} \cdot \mathbf{v}$ which is a real number). Note also that the vector on the RHS of (2.1) appeared earlier in Proposition 28.
- Note that $\mathbf{i} \wedge \mathbf{j} = \mathbf{k}$, $\mathbf{j} \wedge \mathbf{k} = \mathbf{i}$, $\mathbf{k} \wedge \mathbf{i} = \mathbf{j}$, while $\mathbf{i} \wedge \mathbf{k} = -\mathbf{j}$, $\mathbf{j} \wedge \mathbf{i} = -\mathbf{k}$, $\mathbf{k} \wedge \mathbf{j} = -\mathbf{i}$, and that $\mathbf{u} \wedge \mathbf{v} = \mathbf{0}$ if and only if one of $\mathbf{u}$ and $\mathbf{v}$ is a multiple of the other.

Example 34 *Find* $(2, 1, 3) \wedge (1, 0, -1)$. *Determine all vectors* $(2, 1, 3) \wedge \mathbf{v}$ *as* $\mathbf{v}$ *varies over* $\mathbb{R}^3$.

Solution By definition we have

$$(2, 1, 3) \wedge (1, 0, -1) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & 3 \\ 1 & 0 & -1 \end{vmatrix} = (-1 - 0, 3 - (-2), 0 - 1) = (-1, 5, -1).$$

More generally with $\mathbf{v} = (a, b, c)$ we have

$$(2, 1, 3) \wedge \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & 3 \\ a & b & c \end{vmatrix} = (c - 3b, 3a - 2c, 2b - a).$$

It is easy to check, as a, b, c vary, that this is a general vector in the plane $2x + y + 3z = 0$. ■

Proposition 35 *For* $\mathbf{u}, \mathbf{v}$ *in* $\mathbb{R}^3$, *we have* $|\mathbf{u} \wedge \mathbf{v}|^2 = |\mathbf{u}|^2 |\mathbf{v}|^2 - (\mathbf{u} \cdot \mathbf{v})^2$.

Proof This is a simple algebraic verification and is left as Sheet 2, Exercise 5(i). ■

Corollary 36 *For* $\mathbf{u}, \mathbf{v}$ *in* $\mathbb{R}^3$ *we have* $|\mathbf{u} \wedge \mathbf{v}| = |\mathbf{u}| |\mathbf{v}| \sin\theta$ *where* θ *is the smaller angle between* $\mathbf{u}$ *and* $\mathbf{v}$. *In particular* $\mathbf{u} \wedge \mathbf{v} = \mathbf{0}$ *if and only if* $\mathbf{u}$ *and* $\mathbf{v}$ *are linearly dependent.*

[1] The only other space $\mathbb{R}^n$ for which there is a vector product is $\mathbb{R}^7$. See Fenn's *Geometry* for example.

Proof From (1.4) we have $\mathbf{u}\cdot\mathbf{v} = |\mathbf{u}|\,|\mathbf{v}|\cos\theta$. So

$$|\mathbf{u}\wedge\mathbf{v}|^2 = |\mathbf{u}|^2\,|\mathbf{v}|^2 - (\mathbf{u}\cdot\mathbf{v})^2 = |\mathbf{u}|^2\,|\mathbf{v}|^2\,(1-\cos^2\theta) = |\mathbf{u}|^2\,|\mathbf{v}|^2\sin^2\theta.$$

The result follows as $0 \leqslant \sin\theta$ for $0 \leqslant \theta \leqslant \pi$. ■

Corollary 37 *For* $\mathbf{u},\mathbf{v}$ *in* $\mathbb{R}^3$ *then* $|\mathbf{u}\wedge\mathbf{v}|$ *equals the area of the parallelogram with vertices* $\mathbf{0}$, $\mathbf{u}$, $\mathbf{v}$ *and* $\mathbf{u}+\mathbf{v}$.

Proof If we take $\mathbf{u}$ as the base of the parallelogram then the parallelogram has base of length $|\mathbf{u}|$ and height $|\mathbf{v}|\sin\theta$. ■

Proposition 38 *For* $\mathbf{u},\mathbf{v},\mathbf{w}$ *in* $\mathbb{R}^3$, *and reals* α,β *we have*
(a) $\mathbf{u}\wedge\mathbf{v} = -\mathbf{v}\wedge\mathbf{u}$.
(b) $\mathbf{u}\wedge\mathbf{v}$ *is perpendicular to both* $\mathbf{u}$ *and* $\mathbf{v}$.
(c) $(\alpha\mathbf{u}+\beta\mathbf{v})\wedge\mathbf{w} = \alpha(\mathbf{u}\wedge\mathbf{w}) + \beta(\mathbf{v}\wedge\mathbf{w})$.
(d) If $\mathbf{u},\mathbf{v}$ *are perpendicular unit vectors then* $\mathbf{u}\wedge\mathbf{v}$ *is a unit vector.*
(e) $\mathbf{i}\wedge\mathbf{j} = \mathbf{k}$.

Proof Swapping two rows in a determinant has the effect of changing the determinant's sign, and so (a) follows. And (b) follows from the fact that a determinant with two equal rows is zero; so

$$(\mathbf{u}\wedge\mathbf{v})\cdot\mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \cdot (u_1,u_2,u_3) = \begin{vmatrix} u_1 & u_2 & u_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = 0,$$

and likewise $(\mathbf{u}\wedge\mathbf{v})\cdot\mathbf{v} = 0$. (c) follows from the linearity of the determinant in its rows (specifically the second row here). Finally (d) follows from Corollary 37. ■

Corollary 39 *Given independent vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *in* $\mathbb{R}^3$, *the plane containing the origin* $\mathbf{0}$ *and parallel to* $\mathbf{u}$ *and* $\mathbf{v}$ *has equation* $\mathbf{r}\cdot(\mathbf{u}\wedge\mathbf{v}) = 0$.

The definition of the vector product in (2.1) is somewhat unsatisfactory as it appears to depend upon the choice of xyz-co-ordinates in $\mathbb{R}^3$. If the vector product represents something genuinely geometric – the way, for example, that the scalar product can be written in terms of lengths and angles as in (1.4) – then we should be able to determine the vector product in similarly geometric terms. Now we know that the magnitude of the vector product is determined by the vectors' geometry, and that its direction is perpendicular to those of the vectors. Overall, then, the geometry of the two vectors determines their vector product up to a choice of a minus sign.

We will show below that properties (a)-(d) from Proposition 38, none of which expressly involves co-ordinates, determine the vector product up to a choice of sign. What this essentially means is that there are two different **orientations** of three-dimensional space. The xyz-axes in $\mathbb{R}^3$ are *right-handed* in the sense that $\mathbf{i}\wedge\mathbf{j} = \mathbf{k}$ but we could easily have set up xyz-axes in a *left-handed* fashion instead as in Figure 11.

Proposition 40 *Let $\sqcap$ be a vector product which assigns to any two vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^3$ a vector $\mathbf{u} \sqcap \mathbf{v}$ in $\mathbb{R}^3$ and which satisfies properties (a)-(d) of Proposition 38. Then one of the following holds:*

- *for all $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^3$, $\mathbf{u} \sqcap \mathbf{v} = \mathbf{u} \wedge \mathbf{v}$,*
- *for all $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^3$, $\mathbf{u} \sqcap \mathbf{v} = -\mathbf{u} \wedge \mathbf{v}$.*

Proof By property (d) we must have that

$$\mathbf{i} \sqcap \mathbf{j} = \mathbf{k} \qquad \text{or} \qquad \mathbf{i} \sqcap \mathbf{j} = -\mathbf{k}.$$

For now let us assume that $\mathbf{i} \sqcap \mathbf{j} = \mathbf{k}$. In a similar fashion we must have that

$$\mathbf{j} \sqcap \mathbf{k} = \pm \mathbf{i} \qquad \text{and} \qquad \mathbf{k} \sqcap \mathbf{i} = \pm \mathbf{j}.$$

If we had $\mathbf{i} \sqcap \mathbf{j} = \mathbf{k}$ and $\mathbf{j} \sqcap \mathbf{k} = -\mathbf{i}$ then by property (a) we'd have

$$\mathbf{i} \sqcap \mathbf{j} = \mathbf{k} \qquad \text{and} \qquad \mathbf{k} \sqcap \mathbf{j} = \mathbf{i},$$

but then

$$(\mathbf{i} + \mathbf{k}) \sqcap \mathbf{j} = \mathbf{k} + \mathbf{i}.$$

This contradicts property (b) as $\mathbf{k} + \mathbf{i}$ is not perpendicular to $\mathbf{i} + \mathbf{k}$. Hence we must have that $\mathbf{j} \sqcap \mathbf{k} = \mathbf{i}$ and a similar argument can be made to show $\mathbf{k} \sqcap \mathbf{i} = \mathbf{j}$.

Knowing now that $\mathbf{i} \sqcap \mathbf{j} = \mathbf{k}$, $\mathbf{j} \sqcap \mathbf{k} = \mathbf{i}$, $\mathbf{k} \sqcap \mathbf{i} = \mathbf{j}$, every product $\mathbf{u} \sqcap \mathbf{v}$ is now determined by use of properties (a) and (c), and in fact this product $\sqcap$ agrees with the standard vector product $\wedge$. All the above conclusions were derived from our assumption that $\mathbf{i} \sqcap \mathbf{j} = \mathbf{k}$. If, instead, we had decided that $\mathbf{i} \sqcap \mathbf{j} = -\mathbf{k}$ then we'd have instead found that

$$\mathbf{u} \sqcap \mathbf{v} = -\mathbf{u} \wedge \mathbf{v}$$

for all $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$. ■

An important related product is formed using the dot and vector products.

Definition 41 *Given three vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{R}^3$ we define the* ***scalar triple product*** *as*

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \mathbf{u} \cdot (\mathbf{v} \wedge \mathbf{w}).$$

If $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$ and $\mathbf{w} = (w_1, w_2, w_3)$ then

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}. \tag{2.2}$$

- Consequently

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = [\mathbf{v}, \mathbf{w}, \mathbf{u}] = [\mathbf{w}, \mathbf{u}, \mathbf{v}] = -[\mathbf{u}, \mathbf{w}, \mathbf{v}] = -[\mathbf{v}, \mathbf{u}, \mathbf{w}] = -[\mathbf{w}, \mathbf{v}, \mathbf{u}] \tag{2.3}$$

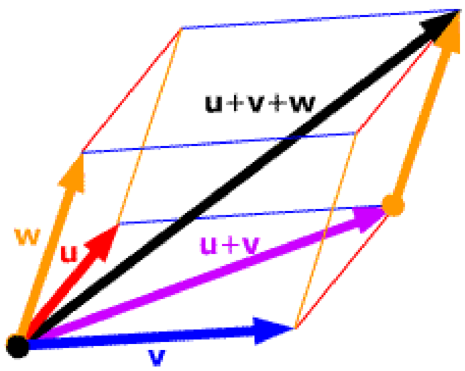
  as swapping two rows of a determinant changes its sign.

- Note that $[\mathbf{u},\mathbf{v},\mathbf{w}]=0$ if and only if $\mathbf{u},\mathbf{v},\mathbf{w}$ are linearly dependent; this is equivalent to the 3×3 matrix with rows $\mathbf{u},\mathbf{v},\mathbf{w}$ being singular.

There is a three-dimensional equivalent of a parallelogram called the **parallelepiped**. This is a three dimensional figure with six parallelograms for faces. Given three vectors $\mathbf{u},\mathbf{v},\mathbf{w}$ in $\mathbb{R}^3$ then they determine a parallelepiped with the eight vertices $\alpha\mathbf{u}+\beta\mathbf{v}+\gamma\mathbf{w}$ where each of α,β,γ is 0 or 1. If we consider $\mathbf{u}$ and $\mathbf{v}$ as determining the base of the parallelepiped then this has area $|\mathbf{u}\wedge\mathbf{v}|$. If θ is the angle between $\mathbf{w}$ and the *normal* to the plane containing $\mathbf{u}$ and $\mathbf{v}$, then the parallelepiped's volume is

$$\text{area of base} \times \text{height} = |\mathbf{u}\wedge\mathbf{v}|\times|\mathbf{w}|\,|\cos\theta| = |(\mathbf{u}\wedge\mathbf{v})\cdot\mathbf{w}| = |[\mathbf{u},\mathbf{v},\mathbf{w}]| \qquad (2.4)$$

as $\mathbf{u}\wedge\mathbf{v}$ is in the direction of the normal of the plane. The volume of the tetrahedron with vertices $\mathbf{0},\mathbf{u},\mathbf{v},\mathbf{w}$ is given by $\frac{1}{6}\,|[\mathbf{u},\mathbf{v},\mathbf{w}]|$ (Sheet 2, Exercise 4(iii)).



12. A parallelepiped

We can likewise also form a vector triple product.

Definition 42 *Given three vectors* $\mathbf{u},\mathbf{v},\mathbf{w}$ *in* $\mathbb{R}^3$ *we define their* ***vector triple product*** *as*

$$\mathbf{u}\wedge(\mathbf{v}\wedge\mathbf{w}).$$

Proposition 43 *For any three vectors* $\mathbf{u},\mathbf{v},\mathbf{w}$ *in* $\mathbb{R}^3$, *then*

$$\mathbf{u}\wedge(\mathbf{v}\wedge\mathbf{w}) = (\mathbf{u}\cdot\mathbf{w})\mathbf{v}-(\mathbf{u}\cdot\mathbf{v})\mathbf{w}.$$

Proof Both the LHS and RHS are linear in $\mathbf{u}$, so it is sufficient to note with $\mathbf{v}=(v_1,v_2,v_3)$ and $\mathbf{w}=(w_1,w_2,w_3)$ that

$$(\mathbf{i}\cdot\mathbf{w})\mathbf{v}-(\mathbf{i}\cdot\mathbf{v})\mathbf{w} = w_1\mathbf{v}-v_1\mathbf{w} = (0, w_1v_2-v_1w_2, w_1v_3-v_1w_3) = \mathbf{i}\wedge(\mathbf{v}\wedge\mathbf{w}) \qquad (2.5)$$

and two similar calculations for $\mathbf{u}=\mathbf{j}$ and $\mathbf{u}=\mathbf{k}$. The result then follows by linearity. In fact, for those comfortable with the comments preceding Proposition 40 that the vector product is entirely determined by geometry, we can choose our x-axis to be in the direction of $\mathbf{u}$ without any loss of generality so that calculation in (2.5) alone is in fact sufficient to verify this proposition. ■

Definition 44 *Given four vectors* $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ *in* $\mathbb{R}^3$, *their* ***scalar quadruple product*** *is*

$$(\mathbf{a} \wedge \mathbf{b}) \cdot (\mathbf{c} \wedge \mathbf{d}).$$

Proposition 45 *For any four vectors* $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ *in* $\mathbb{R}^3$ *then*

$$(\mathbf{a} \wedge \mathbf{b}) \cdot (\mathbf{c} \wedge \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}).$$

Proof Set $\mathbf{e} = \mathbf{c} \wedge \mathbf{d}$. Then

$$\begin{aligned}
& (\mathbf{a} \wedge \mathbf{b}) \cdot (\mathbf{c} \wedge \mathbf{d}) \\
= \;& \mathbf{e} \cdot (\mathbf{a} \wedge \mathbf{b}) \\
= \;& [\mathbf{e}, \mathbf{a}, \mathbf{b}] \\
= \;& [\mathbf{a}, \mathbf{b}, \mathbf{e}] \qquad \text{[by (2.3)]} \\
= \;& \mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{e}) \\
= \;& \mathbf{a} \cdot (\mathbf{b} \wedge (\mathbf{c} \wedge \mathbf{d})) \\
= \;& \mathbf{a} \cdot ((\mathbf{b} \cdot \mathbf{d})\mathbf{c} - (\mathbf{b} \cdot \mathbf{c})\mathbf{d}) \qquad \text{[by the vector triple product]} \\
= \;& (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}).
\end{aligned}$$

■

In many a three-dimensional geometric problem we may be presented with two independent vectors $\mathbf{a}$ and $\mathbf{b}$. In this case it can be useful to know that $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a} \wedge \mathbf{b}$ can be used to place co-ordinates on $\mathbb{R}^3$.

Proposition 46 *Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be linearly independent vectors in* $\mathbb{R}^3$. *Then* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{a} \wedge \mathbf{b}$ *form a basis for* $\mathbb{R}^3$. *This means that for* ***every*** $\mathbf{v}$ *in* $\mathbb{R}^3$ *there are* ***unique*** *real numbers* α, β, γ *such that*

$$\mathbf{v} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{a} \wedge \mathbf{b}. \tag{2.6}$$

We will refer to α, β, γ *as the co-ordinates of* $\mathbf{v}$ *with respect to this basis.*

Solution Note that as $\mathbf{a}$ and $\mathbf{b}$ are independent, then $\mathbf{a} \wedge \mathbf{b} \neq \mathbf{0}$. Given a vector $\mathbf{v}$ in $\mathbb{R}^3$, we define

$$\mathbf{w} = \mathbf{v} - \gamma\mathbf{a} \wedge \mathbf{b} \quad \text{where} \quad \gamma = \frac{\mathbf{v} \cdot (\mathbf{a} \wedge \mathbf{b})}{|\mathbf{a} \wedge \mathbf{b}|^2}.$$

It is a simple check then to note that

$$\mathbf{w} \cdot (\mathbf{a} \wedge \mathbf{b}) = 0,$$

and hence, by Corollary 39, $\mathbf{w}$ lies in the plane through the origin and parallel to $\mathbf{a}$ and $\mathbf{b}$. Hence there are real numbers α, β such that

$$\mathbf{v} - \gamma\mathbf{a} \wedge \mathbf{b} = \mathbf{w} = \alpha\mathbf{a} + \beta\mathbf{b}.$$

Thus we have shown the existence of co-ordinates α, β, γ.

Uniqueness following from the linear independence of the vectors. This means showing that

$$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{a} \wedge \mathbf{b} = \mathbf{0} \qquad \Longrightarrow \qquad \alpha = \beta = \gamma = 0.$$

Dotting the given equation with $\mathbf{a} \wedge \mathbf{b}$ we see $\gamma \left|\mathbf{a} \wedge \mathbf{b}\right|^2 = 0$ and hence $\gamma = 0$. But then $\alpha\mathbf{a} + \beta\mathbf{b} = \mathbf{0}$,and as $\mathbf{a}$ and $\mathbf{b}$ are independent then $\alpha = \beta = 0$. It follows that $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a} \wedge \mathbf{b}$ are independent.

Now, in terms of co-ordinates, the independence of these vectors implies uniqueness: if

$$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{a} \wedge \mathbf{b} = \alpha'\mathbf{a} + \beta'\mathbf{b} + \gamma'\mathbf{a} \wedge \mathbf{b},$$

then

$$(\alpha - \alpha')\mathbf{a} + (\beta - \beta')\mathbf{b} + (\gamma - \gamma')\mathbf{a} \wedge \mathbf{b}$$

and hence, by independence, $\alpha = \alpha', \beta = \beta', \gamma = \gamma'$. ■

The above result is particularly useful in understanding the equation $\mathbf{r} \wedge \mathbf{a} = \mathbf{b}$. Note that if $\mathbf{a} \cdot \mathbf{b} \neq 0$ then this equation can have no solutions – this can be seen by dotting both sides with $\mathbf{a}$.

Proposition 47 *(**Another vector equation for a line.**) Let* $\mathbf{a}, \mathbf{b}$ *vectors in* $\mathbb{R}^3$ *with* $\mathbf{a} \cdot \mathbf{b} = 0$ *and* $\mathbf{a} \neq \mathbf{0}$. *The vectors* $\mathbf{r}$ *in* $\mathbb{R}^3$ *which satisfy* $\mathbf{r} \wedge \mathbf{a} = \mathbf{b}$ *form the line parallel to* $\mathbf{a}$ *which passes through the point* $(\mathbf{a} \wedge \mathbf{b}) / \left|\mathbf{a}\right|^2$.

Proof If $\mathbf{b} = \mathbf{0}$ then we know the equation $\mathbf{r} \wedge \mathbf{a} = \mathbf{0}$ to be satisfied only by scalar multiples of $\mathbf{a}$. So assume that $\mathbf{b} \neq \mathbf{0}$. In this case $\mathbf{a}, \mathbf{b}$ are independent as $\mathbf{a} \cdot \mathbf{b} = 0$ and every vector $\mathbf{r}$ in $\mathbb{R}^3$ can be written uniquely as

$$\mathbf{r} = \lambda\mathbf{a} + \mu\mathbf{b} + \nu\mathbf{a} \wedge \mathbf{b}$$

for reals λ, μ, ν. Then $\mathbf{b} = \mathbf{r} \wedge \mathbf{a}$ if and only if

$$\begin{aligned} \mathbf{b} &= -\mathbf{a} \wedge (\lambda\mathbf{a} + \mu\mathbf{b} + \nu\mathbf{a} \wedge \mathbf{b}) \\ &= -\mu\mathbf{a} \wedge \mathbf{b} - \nu((\mathbf{a} \cdot \mathbf{b})\mathbf{a} - (\mathbf{a} \cdot \mathbf{a})\mathbf{b}) \qquad \text{[vector triple product]} \\ &= -\mu\mathbf{a} \wedge \mathbf{b} + \nu \left|\mathbf{a}\right|^2 \mathbf{b}. \end{aligned}$$

Because of the uniqueness of the co-ordinates, we can compare coefficients and we see that λ may take any value, $\mu = 0$ and $\nu = 1/\left|\mathbf{a}\right|^2$. So $\mathbf{r} \wedge \mathbf{a} = \mathbf{b}$ if and only if

$$\mathbf{r} = \frac{\mathbf{a} \wedge \mathbf{b}}{\left|\mathbf{a}\right|^2} + \lambda\mathbf{a} \quad \text{where } \lambda \text{ is real.}$$

The result follows. ■

Example 48 *Under what conditions do the line* $\mathbf{r} \wedge \mathbf{a} = \mathbf{b}$, *where* $\mathbf{a} \cdot \mathbf{b} = 0$, *and the plane* $\mathbf{r} \cdot \mathbf{n} = c$, *intersect in a unique point?*

Solution One issue with this problem is that neither of the given equations give $\mathbf{r}$ explicitly. So we can't simply substitute one equation into the other. However, from Proposition 47, we know that we can write

$$\mathbf{r} = \mathbf{p} + \lambda\mathbf{a} \quad \text{where } \mathbf{p} = \frac{\mathbf{a} \wedge \mathbf{b}}{|\mathbf{a}|^2}.$$

If we substitute this into the equation of the plane – so that λ is now the variable – we have

$$(\mathbf{p} + \lambda\mathbf{a}) \cdot \mathbf{n} = c.$$

This determines a unique value for λ if and only if $\mathbf{a} \cdot \mathbf{n} \neq 0$. Note that, geometrically, this means that the line is not parallel with the plane. If we had the case that $\mathbf{a} \cdot \mathbf{n} = 0$ then we either have infinitely many points of intersection when $\mathbf{p} \cdot \mathbf{n} = c$ (when the line lies entirely in the plane) or no solutions when $\mathbf{p} \cdot \mathbf{n} \neq c$ (the line is parallel to, but not in, the plane). ■

Remark 49 ***(Properties of Determinants)*** *Here is a list of useful properties of* 2×2 *and* 3×3 *determinants. We shall not make an effort to prove any of these here. Firstly their definitions are:*

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \qquad \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - afh - bdi - ceg.$$

A square matrix M *has a number associated with it called its* ***determinant,*** *denoted* $\det M$ *or* $|M|$*. There are various different ways of introducing determinants, each of which has its advantages but none of which is wholly ideal as will be made clearer below. There is an inductive definition – called "expansion along the first column" – giving* $n \times n$ *determinants in terms of* $(n-1) \times (n-1)$ *determinants which has the merit of being unambiguous but is poorly motivated and computationally nightmarish.*

In the 2×2 *and* 3×3 *cases,* ***but only in these cases,*** *there is a simple way to remember the determinant formula. The* 2×2 *formula is clearly the product of entries on the left-to-right diagonal minus the product of those on the right-to-left diagonal. If, in the* 3×3 *case, we allow diagonals to "wrap around" the vertical sides of the matrix – for example as below*

$$\begin{pmatrix} & \searrow & \\ & & \searrow \\ \searrow & & \end{pmatrix}, \qquad \begin{pmatrix} \swarrow & & \\ & & \swarrow \\ & \swarrow & \end{pmatrix},$$

then from this point of view a 3×3 *matrix has three left-to-right diagonals and three right-to-left. A* 3×3 *determinant then equals the sum of the products of entries on the three left-to-right diagonals minus the products from the three right-to-left diagonals. This method of calculation does* ***not*** *apply to* $n \times n$ *determinants when* $n \geqslant 4$.

Here are some further properties of $n \times n$ *determinants.*

- *(i)* $\det$ *is linear in the rows (or columns) of a matrix.*
- *(ii) if a matrix has two equal rows then its determinant is zero.*
- *(iii)* $\det I_n = 1$.

In fact, these three algebraic properties uniquely characterize a function det *which assigns a number to each* $n \times n$ *matrix. The problem with this approach is that the existence and uniqueness of such a function are still moot.*

- *(a) For any region* S *of* $\mathbb{R}^2$ *and a* 2×2 *matrix* A *we have*

$$\textit{area of } A(S) = |\det A| \times (\textit{area of } S).$$

  Thus $|\det A|$ *is the area-scaling factor of the map* $\mathbf{v} \mapsto A\mathbf{v}$*. A similar result applies to volume in three dimensions.*

- *(b) Still in the plane, the sense of any angle under the map* $\mathbf{v} \mapsto A\mathbf{v}$ *will be reversed when* $\det A < 0$ *but will keep the same sense when* $\det A > 0$*. Again a similar result applies in* $\mathbb{R}^3$ *with a right-handed basis mapping to a left-handed one when* $\det A < 0$

These last two properties (a) and (b) best show the significance of determinants. Thinking along these lines, the following seem natural enough results:

- *(*α*)* $\det AB = \det A \det B$
- *(*β*) a square matrix is singular if and only it has zero determinant.*

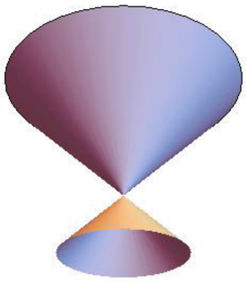
However, whilst these geometric properties might better motivate the importance of determinants, they would be less useful in calculating determinants. Their meaning would also be less clear if we were working in more than three dimensions (at least until we had defined volume and sense/orientation in higher dimensions) or if we were dealing with matrices with complex numbers as entries.
Finally, calculation is difficult and inefficient using the inductive definition – for example, the formula for an $n \times n$ *determinant involves the sum of* $n!$ *separate products. In due course, in Linear Algebra II, you will see that the best way to calculate determinants is via elementary row operations.*

3. Conics. Orthogonal Matrices

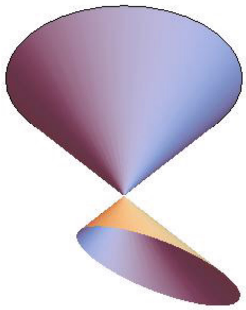
3.1 Conics – The Normal Forms

Remark 50 *The study of the conics dates back to the ancient Greeks. They were studied by Euclid and Archimedes, but much of their work did not survive. It is Apollonius of Perga (c. 262-190 BC), who wrote the 8 volume Conics, who is most associated with the curves and who gave them their modern names. Later Pappus (c. 290-350) defined them in terms of a focus and directrix (Definition 51). With the introduction of analytic geometry (i.e. the use of Cartesian co-ordinates), their study moved into the realm of algebra. From an algebraic point of view, the conics are a natural family of curves to study being defined by degree two equations in two variables (Theorem 59).*

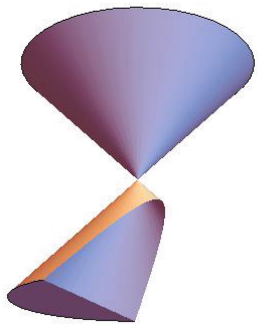
The **conics** or **conic sections** are a family of planar curves. They get their name as they can each be formed by intersecting the double cone $x^2 + y^2 = z^2 \cot^2 \alpha$ with a plane in $\mathbb{R}^3$. For example, intersecting the cone with the plane $z = R$ produces a circle of radius R. The four different possibilities are drawn below.



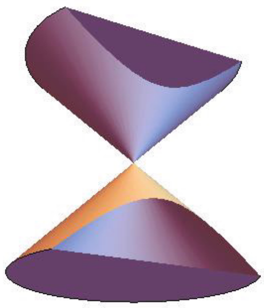
13a. Circle



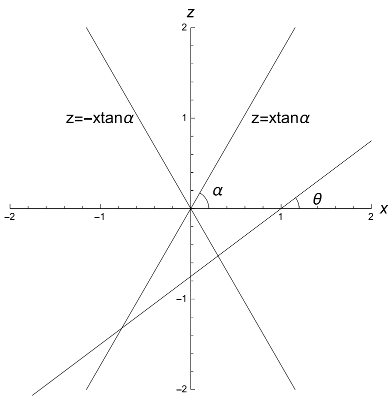
13b. Ellipse



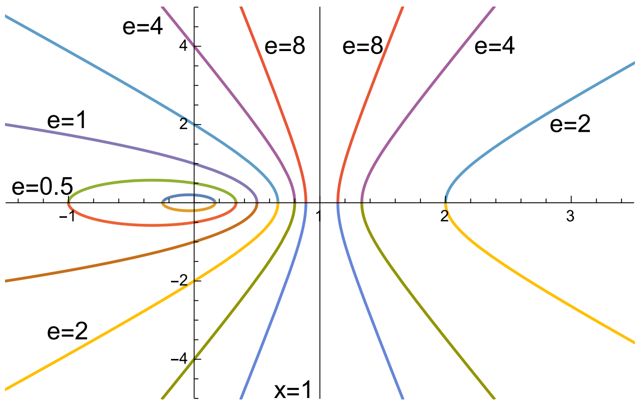
13c. Parabola



13d. Hyperbola



14. A cross-section



15. Varying the eccentricity

In Figure 14 is the cross-sectional view, in the xz-plane, of the intersection of the plane $z = (x - 1) \tan \theta$ with the double cone $x^2 + y^2 = z^2 \cot^2 \alpha$. We see that when $\theta < \alpha$ then the plane intersects only with bottom cone in a bounded curve (which in due course we shall see to be an *ellipse*). When $\theta = \alpha$ it intersects with the lower cone in an unbounded curve (a *parabola*), and when $\theta > \alpha$ we see that the plane intersects with both cones to make two separate unbounded curves (a *hyperbola*).

However, as a first definition, we shall introduce conics using the idea of a *directrix* and *focus*.

Definition 51 *Let D be a line, F a point not on the line D and $e > 0$. Then the **conic** with **directrix** D and **focus** F and **eccentricity** e, is the set of points P (in the plane containing F and D) which satisfy the equation*

$$|PF| = e|PD|$$

where $|PF|$ is the distance of P from the focus and $|PD|$ is the distance of P from the directrix. That is, as the point P moves around the conic, the shortest distance from P to the line D is in constant proportion to the distance of P from the point F.

- *If $0 < e < 1$ then the conic is called an **ellipse**.*
- *If $e = 1$ then the conic is called a **parabola**.*
- *If $e > 1$ then the conic is called a **hyperbola**.*

In the diagram on the right above are sketched, for a fixed focus F (the origin) and fixed directrix D (the line $x = 1$), a selection of conics of varying eccentricity e.

Example 52 *Find the equation of the parabola with focus $(1, 1)$ and directrix $x + 2y = 1$.*

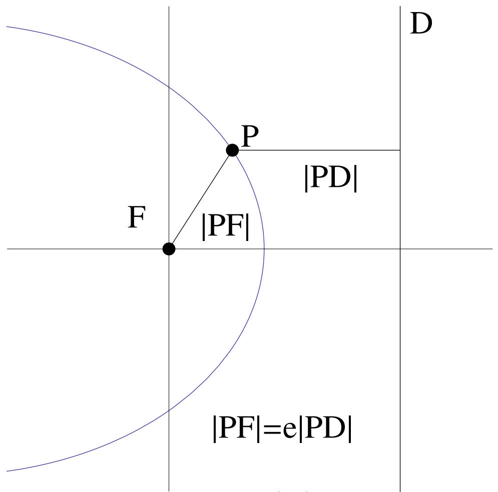
Solution The distance of the point (x_0, y_0) from the line $ax + by + c = 0$ equals

$$\frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

So the given parabola has equation

$$\frac{|x + 2y - 1|}{\sqrt{5}} = \sqrt{(x-1)^2 + (y-1)^2}.$$

With some simplifying this becomes $4x^2 + y^2 - 8x - 6y - 4xy + 9 = 0$. ∎



16. Focus and directrix

It is somewhat more natural to begin describing a conic's equation with polar co-ordinates whose origin is at the focus F. Let C be the conic in the plane with directrix D, focus F and eccentricity e. We may choose polar co-ordinates for the plane in which F is the origin and D is the line $r \cos\theta = k$ (i.e. $x = k$).
Then $|PF| = r$ and $|PD| = k - r\cos\theta$. So we have $r = e(k - r\cos\theta)$ or rearranging

$$r = \frac{ke}{1 + e\cos\theta}. \qquad (3.1)$$

Note that k is purely a scaling factor here and it is e which determines the shape of the conic. Note also that when $0 < e < 1$ then r is well-defined and bounded for all θ. However when $e \geqslant 1$ then r (which must be positive) is unbounded and further undefined when $1+e\cos\theta \leqslant 0$. If we change to Cartesian co-ordinates (u,v) using $u = r\cos\theta$ and $v = r\sin\theta$, we obtain $\sqrt{u^2+v^2} = e(k-u)$ or equivalently

$$(1-e^2)u^2 + 2e^2ku + v^2 = e^2k^2. \tag{3.2}$$

Provided $e \neq 1$, then we can complete the square to obtain

$$(1-e^2)\left(u + \frac{e^2k}{1-e^2}\right)^2 + v^2 = e^2k^2 + \frac{e^4k^2}{1-e^2} = \frac{e^2k^2}{1-e^2}.$$

Introducing new co-ordinates $x = u + e^2k(1-e^2)^{-1}$ and $y = v$, our equation becomes

$$(1-e^2)x^2 + y^2 = \frac{e^2k^2}{1-e^2}. \tag{3.3}$$

Note that this change from uv-co-ordinates to xy-co-ordinates is simply a translation of the origin along the u-axis.

We are now in a position to write down the **normal forms** of a conic's equation whether the conic be an ellipse, a parabola or a hyperbola.

- **Case 1 – the ellipse** $(0 < e < 1)$.

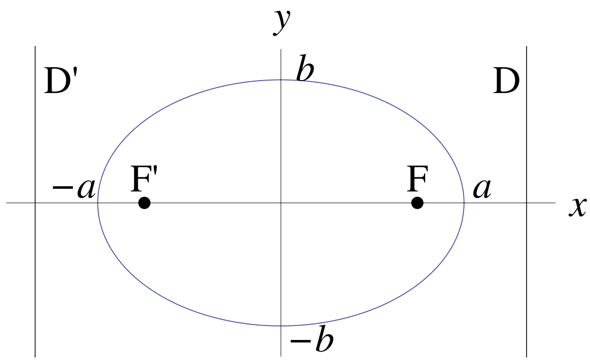
  In this case, we can rewrite equation (3.3) as

  $$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad 0 < b < a$$

  where

  $$a = \frac{ke}{1-e^2}, \qquad b = \frac{ke}{\sqrt{1-e^2}}.$$

  Note the eccentricity is $e = \sqrt{1 - b^2/a^2}$ in terms of a and b.



17. The ellipse

In the xy-co-ordinates, the focus F is at $(ae, 0)$ and the directrix D is the line $x = a/e$. However, it's clear by symmetry that we could have used $F' = (-ae, 0)$ and $D'\colon x = -a/e$ as an alternative focus and directrix and still produce the same ellipse. The area of this ellipse is πab. Further this ellipse can be parametrized either by setting

$$x = a\cos t, \qquad y = b\sin t, \qquad 0 \leqslant t < 2\pi, \tag{3.4}$$

or alternatively as

$$x = a\left(\frac{1-t^2}{1+t^2}\right), \qquad y = b\left(\frac{2t}{1+t^2}\right), \qquad (-\infty < t < \infty). \tag{3.5}$$

Note that this last parametrization omits the point $(-a, 0)$, which can be thought of as corresponding to $t = \infty$.

Remark 53 *The normal form of the ellipse is $x^2/a^2 + y^2/b^2 = 1$ where $a = ke(1-e^2)^{-1}$ and $b = ke(1-e^2)^{-1/2}$. If we keep constant $l = ke$, as we let e become closer to zero we find a and b become closer to l and the ellipse approximates to the circle $x^2 + y^2 = l^2$. As a limit, then, a circle can be thought of as a conic with eccentricity $e = 0$. The two foci $(\pm ae, 0)$ both move to the centre of the circle as e approaches zero and the two directrices $x = \pm a/e$ have both moved towards infinity.*

- **Case 2 – the hyperbola** ($e > 1$).

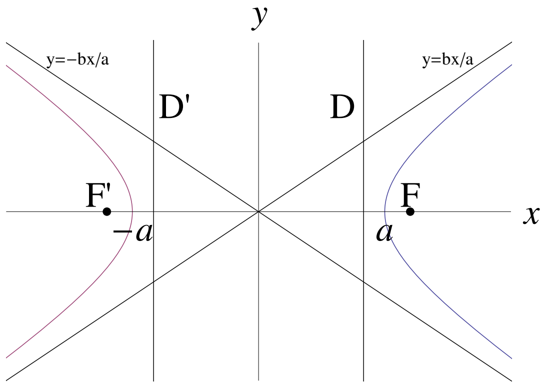
  In this case, we can rewrite equation (3.3) as

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \qquad 0 < a, b$$

where

$$a = \frac{ke}{e^2 - 1}, \qquad b = \frac{ke}{\sqrt{e^2 - 1}}.$$

Note that the eccentricity $e = \sqrt{1 + b^2/a^2}$ in terms of a and b.



18. The hyperbola

In the xy-co-ordinates, the focus $F = (ae, 0)$ and the directrix D is the line $x = a/e$. However it is again clear from symmetry that we could have used $F' = (-ae, 0)$ and D': $x = -a/e$ as a new focus and directrix to produce the same hyperbola. The lines $ay = \pm bx$ are known as the **asymptotes** of the hyperbola; these are, in a sense, the tangents to the hyperbola at its two 'points at infinity'. When $e = \sqrt{2}$ (i.e. when $a = b$) then these asymptotes are perpendicular and C is known as a **right hyperbola.**

In a similar fashion to the ellipse, this hyperbola can be parametrized by

$$x = \pm a \cosh t \qquad y = b \sinh t \qquad (-\infty < t < \infty), \tag{3.6}$$

or alternatively as

$$x = a\frac{1 + t^2}{1 - t^2}, \qquad y = b\frac{2t}{1 - t^2} \qquad (t \neq \pm 1). \tag{3.7}$$

Again this second parametrization misses out the point $(-a, 0)$ which in a sense corresponds to $t = \infty$. Likewise the points corresponding to $t = \pm 1$ can be viewed as the hyperbola's two "points at infinity".

- **Case 3 – the parabola** $(e = 1)$

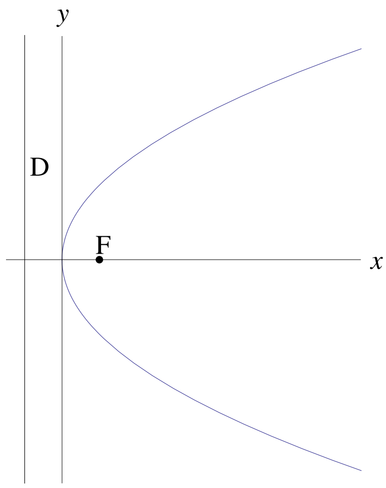
  In our derivation of equation (3.3) we assumed that $e \neq 1$. We can treat this case now by returning to equation (3.2). If we set $e = 1$ then we obtain

  $$2ku + v^2 = k^2.$$

  If we substitute $a = k/2$, $x = a - u$, $y = v$ then we obtain the normal form for a parabola

  $$y^2 = 4ax.$$

  The focus is the point $(a, 0)$ and the directrix is the line $x = -a$. The *vertex* of the parabola is at $(0, 0)$. In this case the conic C may be parametrised by setting $(x, y) = (at^2, 2at)$ where $-\infty < t < \infty$.



19. The parabola

The table below summarizes the details of the circle, ellipse, hyperbola and parabola:

e range	conic type	normal form	e formula	foci	directrices	notes
$e = 0$	circle	$x^2 + y^2 = a^2$	0	$(0, 0)$	at infinity	
$0 < e < 1$	ellipse	$x^2/a^2 + y^2/b^2 = 1$	$\sqrt{1 - b^2/a^2}$	$(\pm ae, 0)$	$x = \pm a/e$	$0 < b < a$
$e = 1$	parabola	$y^2 = 4ax$	1	$(a, 0)$	$x = -a$	vertex: $(0, 0)$
$e > 1$	hyperbola	$x^2/a^2 - y^2/b^2 = 1$	$\sqrt{1 + b^2/a^2}$	$(\pm ae, 0)$	$x = \pm a/e$	asymptotes: $y = \pm bx/a$

Example 54 *Let $0 < \theta, \alpha < \pi/2$. Show that the intersection of the cone $x^2 + y^2 = z^2 \cot^2 \alpha$ with the plane $z = \tan\theta(x - 1)$ is an ellipse, parabola or hyperbola. Determine which type of conic arises in terms of θ.*

Solution We will denote as C the intersection of the cone and plane. In order to properly describe C we need to set up co-ordinates in the plane $z = \tan\theta(x - 1)$. Note that

$$\mathbf{e}_1 = (\cos\theta, 0, \sin\theta), \qquad \mathbf{e}_2 = (0, 1, 0), \qquad \mathbf{e}_3 = (\sin\theta, 0, -\cos\theta),$$

are mutually perpendicular unit vectors in $\mathbb{R}^3$ with $\mathbf{e}_1, \mathbf{e}_2$ being parallel to the plane and $\mathbf{e}_3$ being perpendicular to it. Any point (x, y, z) in the plane can then be written uniquely as

$$(x, y, z) = (1, 0, 0) + X\mathbf{e}_1 + Y\mathbf{e}_2 = (1 + X\cos\theta, Y, X\sin\theta)$$

for some X, Y, so that X and Y then act as the desired co-ordinates in the plane. Substituting the above expression for (x, y, z) into the cone's equation $x^2 + y^2 = z^2 \cot^2 \alpha$ gives

$$(1 + X\cos\theta)^2 + Y^2 = (X\sin\theta)^2 \cot^2 \alpha.$$

This rearranges to

$$(\cos^2\theta - \sin^2\theta\cot^2\alpha)X^2 + 2X\cos\theta + Y^2 = -1.$$

If $\theta \neq \alpha$ then we can complete the square to arrive at

$$\frac{(\cos^2\theta - \sin^2\theta\cot^2\alpha)^2}{\sin^2\theta\cot^2\alpha}\left(X + \frac{\cos\theta}{\cos^2\theta - \sin^2\theta\cot^2\alpha}\right)^2 + \frac{(\cos^2\theta - \sin^2\theta\cot^2\alpha)}{\sin^2\theta\cot^2\alpha}Y^2 = 1.$$

If $\theta < \alpha$ then the coefficients of the squares are positive and we have an ellipse whilst if $\theta > \alpha$ we have a hyperbola as the second coefficient is negative. If $\theta = \alpha$ then our original equation has become

$$2X\cos\theta + Y^2 = -1,$$

which is a parabola. Further calculation shows that the eccentricity of the conic is $\sin\theta/\sin\alpha$.

■

Ellipses also have the following geometric property which means that, if one ties a loose length of string between two fixed points and draws a curve with a pen so as to keep the string taut at all points, then the resulting curve is an ellipse.

Proposition 55 *Let A, B be distinct points in the plane and $r > |AB|$ be a real number. The locus $|AP| + |PB| = r$ is an ellipse with foci at A and B.*

Proof If we consider the ellipse with foci A and B and a point P on the ellipse we have

$$\begin{aligned} |AP| + |PB| &= |F_1P| + |PF_2| \\ &= e\,|D_1P| + e\,|D_2P| \\ &= e\,|D_1D_2| \end{aligned}$$

where D_1D_2 is the perpendicular distance between the two directrices. Thus the value is constant on any ellipse with foci A and B and will take different values for different ellipses as the value of $|AP| + |PB|$ increases as P moves right along the line AB. ■

3.2 The Degree Two Equation in Two Variables

Example 56 *Show that the curve*

$$x^2 + xy + y^2 = 1$$

is an ellipse. Sketch the curve, its foci and directrices, and find its area.

Proof This certainly isn't an ellipse in its normal form $x^2/a^2 + y^2/b^2 = 1$ which involves no mixed term xy. Also we can see that a translation of $\mathbb{R}^2$, which takes the form $(x, y) \mapsto (x + c_1, y + c_2)$, won't eliminate the xy-term for us. We could however try rotating the curve.

A rotation about the origin in $\mathbb{R}^2$ by θ anti-clockwise takes the form

$$\begin{aligned} X &= x\cos\theta + y\sin\theta, \qquad Y = -x\sin\theta + y\cos\theta; \\ x &= X\cos\theta - Y\sin\theta, \qquad y = X\sin\theta + Y\cos\theta. \end{aligned}$$

Writing $c = \cos\theta$ and $s = \sin\theta$, for ease of notation, our equation becomes

$$(Xc - Ys)^2 + (Xc - Ys)(Xs + Yc) + (Xs + Yc)^2 = 1$$

which simplifies to

$$(1 + cs)X^2 + (c^2 - s^2)XY + (1 - cs)Y^2 = 1.$$

So if we wish to eliminate the xy-term then we want

$$\cos 2\theta = c^2 - s^2 = 0$$

which will be the case when $\theta = \pi/4$, say. For this value of θ we have $c = s = 1/\sqrt{2}$ and our equation has become

$$\frac{3}{2}X^2 + \frac{1}{2}Y^2 = 1.$$

This is certainly an ellipse as it can be put into normal form as

$$\left(\frac{X}{\sqrt{2/3}}\right)^2 + \left(\frac{Y}{\sqrt{2}}\right)^2 = 1. \qquad a = \sqrt{\frac{2}{3}}, \quad b = \sqrt{2}.$$

It has eccentricity

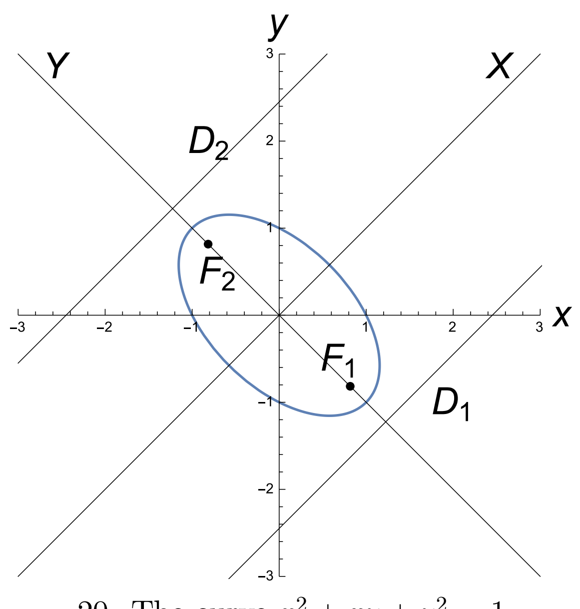
$$e = \sqrt{1 - \frac{a^2}{b^2}} = \sqrt{1 - \frac{2/3}{2}} = \sqrt{\frac{2}{3}}$$

and area $\pi ab = 2\pi/\sqrt{3}$. The foci and directrices, in terms of XY- and xy-co-ordinates are at

foci: $(X, Y) = (0, \pm be) = \left(0, \pm 2/\sqrt{3}\right)$; $\quad x = \mp\sqrt{2/3} \quad y = \pm\sqrt{2/3}.$

directrices: $Y = \pm b/e = \sqrt{3}$; $\quad y = x \pm \sqrt{6}$

A sketch then of the curve is given below, highlighting the X-axis and Y-axis as well.



20. The curve $x^2 + xy + y^2 = 1$

■

More generally we could consider a locus of the form $Ax^2+Bxy+Cy^2=1$ and seek to put it in normal form in a similar way.

Theorem 57 *The solutions of the equation*

$$Ax^2+Bxy+Cy^2=1 \tag{3.8}$$

where A,B,C are real constants, such that A,B,C are not all zero, form one of the following types of loci:

Case (a): If $B^2-4AC<0$ then the solutions form an ellipse or the empty set.
Case (b): If $B^2-4AC=0$ then the solutions form two parallel lines or the empty set.
Case (c): If $B^2-4AC>0$ then the solutions form a hyperbola.

Proof Note that we may assume $A\geqslant C$ without any loss of generality; if this were not the case we could swap the variables x and y. We begin with a rotation of the axes as before to remove the mixed term xy. Set

$$\begin{aligned} X &= x\cos\theta+y\sin\theta, \qquad Y=-x\sin\theta+y\cos\theta;\\ x &= X\cos\theta-Y\sin\theta \qquad y=X\sin\theta+Y\cos\theta. \end{aligned}$$

Again writing $c=\cos\theta$ and $s=\sin\theta$, for ease of notation, our equation becomes

$$A(Xc-Ys)^2+B(Xc-Ys)(Xs+Yc)+C(Xs+Yc)^2=1.$$

The coefficient of the XY term is $-2Acs-Bs^2+Bc^2+2Csc=B\cos 2\theta+(C-A)\sin 2\theta$ which will be zero when

$$\tan 2\theta=\frac{B}{A-C}.$$

If we now choose a solution θ in the range $-\pi/4<\theta\leqslant\pi/4$ then we can simplify our equation further to

$$(Ac^2+Bsc+Cs^2)X^2+(As^2-Bsc+Cc^2)Y^2=1.$$

As $A\geqslant C$ then $\sin 2\theta=B/H$ and $\cos 2\theta=(A-C)/H$ where $H=\sqrt{(A-C)^2+B^2}$. With some further simplification our equation rearranges to

$$\left(\frac{A+C+H}{2}\right)X^2+\left(\frac{A+C-H}{2}\right)Y^2=1.$$

Note that $A+C+H$ and $A+C-H$ will have the same sign if $(A+C)^2>H^2$ which is equivalent to the inequality

$$4AC>B^2.$$

(a) If $4AC>B^2$ then the X^2 and Y^2 coefficients have the same sign and so the equation can be rewritten as $X^2/a^2+Y^2/b^2=\pm 1$ depending on whether these coefficients are both positive or both negative. Thus we either have an ellipse or the empty set.

(c) If $4AC<B^2$ then the X^2 and Y^2 coefficients have different signs and so the equation can be rewritten as $X^2/a^2-Y^2/b^2=\pm 1$. In all cases this represents a hyperbola.

(b) If $4AC = B^2$ then (only) one of the X^2 and Y^2 coefficients is zero. If $A + C + H = 0$ then our equation now reads

$$(A + C)\,Y^2 = 1$$

which is empty as $-A - C = H > 0$. If $A + C - H = 0$ then our equation now reads

$$(A + C)\,X^2 = 1$$

which represents a pair of parallel lines as $A + C = H > 0$. ■

Remark 58 *(Details of this are off-syllabus) More generally, a* ***degree two equation in two variables*** *is one of the form*

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{3.9}$$

where A, B, C, D, E, F are real constants and A, B, C are not all zero. Their loci can again be understood, first by a rotation of axes to eliminate the xy term, and secondly by a translation of the plane (i.e. a change of origin) to get the equation in a normal form. The different cases that can arise are as follows:

Theorem 59 *Case (a): If $B^2 - 4AC < 0$ then the solutions of (3.9) form an ellipse, a single point or the empty set.*

Case (b): If $B^2 - 4AC = 0$ then the solutions of (3.9) form a parabola, two parallel lines, a single line or the empty set.

Case (c): If $B^2 - 4AC > 0$ then the solutions of (3.9) form a hyperbola or two intersecting lines.

Example 60 *Classify the curve with equation*

$$4x^2 - 4xy + y^2 - 8x - 6y + 9 = 0.$$

Solution We met this curve earlier as the parabola with focus $(1, 1)$ and directrix $x + 2y = 1$. However, imagine instead being presented with the given equation and trying to understand its locus. Here we have $A = 4$, $B = -4$, $C = 1$, so we should rotate by θ where

$$\tan 2\theta = \frac{B}{A - C} = \frac{-4}{4 - 1} = \frac{-4}{3}.$$

We then have

$$H = \sqrt{(A - C)^2 + B^2} = \sqrt{3^2 + 4^2} = 5.$$

If $\tan 2\theta = -4/3$ and $-\pi/4 < \theta < 0$ then $\sin 2\theta = -4/5$ and $\cos 2\theta = 3/5$, so that

$$\sin\theta = -\frac{1}{\sqrt{5}} \qquad \cos\theta = \frac{2}{\sqrt{5}}.$$

A rotation by this choice of θ means a change of variable of the form

$$\begin{aligned} X &= \frac{2x}{\sqrt{5}} - \frac{y}{\sqrt{5}}, \qquad Y = \frac{x}{\sqrt{5}} + \frac{2y}{\sqrt{5}}; \\ x &= \frac{2X}{\sqrt{5}} + \frac{Y}{\sqrt{5}}, \qquad y = -\frac{X}{\sqrt{5}} + \frac{2Y}{\sqrt{5}}. \end{aligned}$$

For this choice of θ our equation eventually simplifies to

$$5X^2 - 2\sqrt{5}X - 4\sqrt{5}Y + 9 = 0.$$

If we complete the square we arrive at

$$5\left(X - \frac{1}{\sqrt{5}}\right)^2 - 4\sqrt{5}Y + 8 = 0$$

and we recognize

$$\left(X - \frac{1}{\sqrt{5}}\right)^2 = \frac{4}{\sqrt{5}}\left(Y - \frac{2}{\sqrt{5}}\right)$$

as a parabola. (The normal form is $x^2 = 4ay$ with focus at $(0, a)$ and directrix $y = -a$.) So the above parabola has $a = 1/\sqrt{5}$ and focus at

$$(X, Y) = \left(\frac{1}{\sqrt{5}}, \frac{3}{\sqrt{5}}\right) \quad \text{or equivalently} \quad (x, y) = (1, 1)\,.$$

■

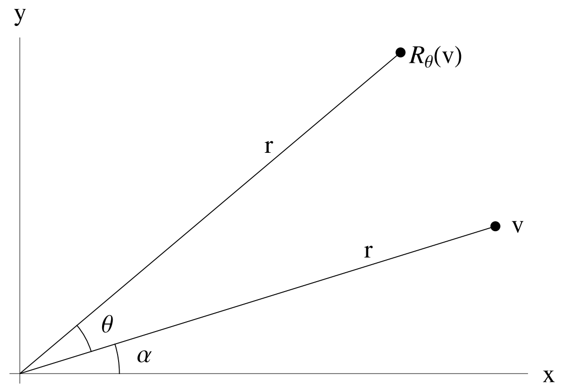
3.3 Orthogonal Matrices and Isometries

Definition 61 *An* ***isometry*** *T from $\mathbb{R}^n$ to $\mathbb{R}^n$ is a distance-preserving map. That is:*

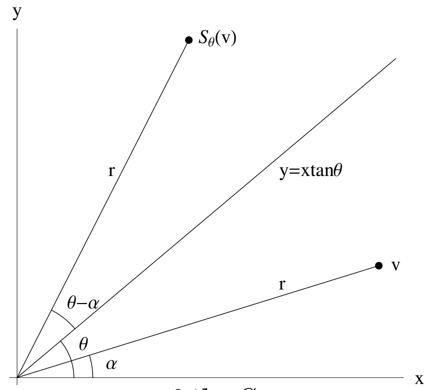
$$|T(\mathbf{x}) - T(\mathbf{y})| = |\mathbf{x} - \mathbf{y}| \qquad \text{for all } \mathbf{x}, \mathbf{y} \text{ in } \mathbb{R}^n.$$

Rotations, reflections and translations are all examples of isometries. We might reasonably ask the question: what is the effect of a rotation or a reflection on a point (x, y) in the plane? We start by assuming we are rotation about the origin, or reflecting in a line through the origin.

Example 62 *(**Rotations and reflections of the plane**) Describe the maps (a) R_θ, which denotes rotation by θ anti-clockwise about the origin; (b) S_θ, which is reflection in the line $y = x\tan\theta$.*



21a. R_θ



21b. S_θ

Solution (a) Given that we are describing a rotation about the origin then polar co-ordinates seem a natural way to help describe the map. Say that $\mathbf{v} = (r\cos\alpha, r\sin\alpha)^T$, as in Figure 21a; then

$$R_\theta(\mathbf{v}) = \begin{pmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta(r\cos\alpha) - \sin\theta(r\sin\alpha) \\ \cos\theta(r\sin\alpha) + \sin\theta(r\cos\alpha) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}(\mathbf{v}).$$

(b) From Figure 21b we see S_θ maps $\mathbf{v}$ to the point $(r\cos(2\theta-\alpha), r\sin(2\theta-\alpha))^T$. So

$$S_\theta(\mathbf{v}) = \begin{pmatrix} r\cos(2\theta-\alpha) \\ r\sin(2\theta-\alpha) \end{pmatrix} = \begin{pmatrix} \cos 2\theta(r\cos\alpha) + \sin 2\theta(r\sin\alpha) \\ \sin 2\theta(r\cos\alpha) - \cos 2\theta(r\sin\alpha) \end{pmatrix} = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}(\mathbf{v}).$$

So in both cases we see that the map can be effected as premultiplication by the matrices

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}; \qquad S_\theta = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}.$$

■

Example 63 *Describe the map in $\mathbb{R}^2$ that is reflection in the line $x+y=2$.*

Solution This certainly won't be a simple matter of multiplication by a matrix as the origin is not fixed by the map. If we had been considering reflection in the line $x+y=0$, a line which can be represented as $y = x\tan(3\pi/4)$ – then we see that the matrix

$$S_{3\pi/4} = \begin{pmatrix} \cos(3\pi/2) & \sin(3\pi/2) \\ \sin(3\pi/2) & -\cos(3\pi/2) \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}. \qquad S_{3\pi/4}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ -x \end{pmatrix}.$$

We could transform our given problem into the latter one by a change of origin. If we set

$$X = x-1, \qquad Y = y-1,$$

then the XY-co-ordinates are a translation away from the xy-co-ordinates. We have simply changed origin; importantly the new origin $(X,Y) = (0,0)$ lies on the line $x+y=2$. Further, in these new co-ordinates, the line of reflection has equation $X+Y=0$. Hence we can resolve our problem as follows:

$$\begin{aligned} \text{point with } xy\text{-co-ordinates } (x,y) = \quad & \text{point with } XY\text{-co-ordinates } (x-1, y-1) \\ \text{transforms to} \quad & \text{point with } XY\text{-co-ordinates } (1-y, 1-x) \\ = \quad & \text{point with } xy\text{-co-ordinates } (2-y, 2-x). \end{aligned}$$

So we cannot write this map as simple pre-multiplication by a matrix – however we can write it as

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 2-y \\ 2-x \end{pmatrix} = \begin{pmatrix} -y \\ -x \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Note that the matrix in this expression is the same matrix $S_{3\pi/4}$ as we have before. ■

Note that the lengths of vectors, and the angles between vectors, are both preserved by maps R_θ and S_θ as one would expect of rotations and reflections. For example, showing that R_θ preserves lengths is equivalent to verifying the identity

$$(x\cos\theta - y\sin\theta)^2 + (x\sin\theta + y\cos\theta)^2 = x^2 + y^2.$$

We might more generally consider what matrices have these properties of preserving lengths and angles? As length and angle are given in terms of the dot product (and conversely as length and angle determine the dot product) then we are interested in those 2×2 matrices such that

$$A\mathbf{v}\cdot A\mathbf{w} = \mathbf{v}\cdot\mathbf{w} \qquad \text{for all column vectors } \mathbf{v},\mathbf{w} \text{ in } \mathbb{R}^2.$$

Now a useful rearrangement of this identity relies on noting

$$\mathbf{v}\cdot\mathbf{w} = v_1w_1 + v_2w_2 = \begin{pmatrix} v_1 & v_2 \end{pmatrix}\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \mathbf{v}^T\mathbf{w}.$$

Using the transpose product rule $(MN)^T = N^TM^T$ we have

$$\begin{aligned} & A\mathbf{v}\cdot A\mathbf{w} = \mathbf{v}\cdot\mathbf{w} \qquad \text{for all column vectors } \mathbf{v},\mathbf{w} \text{ in } \mathbb{R}^2 \\ \iff & (A\mathbf{v})^T(A\mathbf{w}) = \mathbf{v}^T\mathbf{w} \qquad \text{for all column vectors } \mathbf{v},\mathbf{w} \text{ in } \mathbb{R}^2 \\ \iff & \mathbf{v}^TA^TA\mathbf{w} = \mathbf{v}^T\mathbf{w} \qquad \text{for all column vectors } \mathbf{v},\mathbf{w} \text{ in } \mathbb{R}^2. \end{aligned}$$

Now for any 2×2 matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

note that

$$\mathbf{i}^TM\mathbf{i} = a, \qquad \mathbf{i}^TM\mathbf{j} = b, \qquad \mathbf{j}^TM\mathbf{i} = c, \qquad \mathbf{j}^TM\mathbf{j} = d.$$

Hence $\mathbf{v}^TA^TA\mathbf{w} = \mathbf{v}^T\mathbf{w} = \mathbf{v}^TI_2\mathbf{w}$ for all column vectors $\mathbf{v},\mathbf{w}$ in $\mathbb{R}^2$ if and only if

$$A^TA = I_2.$$

This leads us to the following definition.

Definition 64 *A square real matrix A is said to be* ***orthogonal*** *if $A^{-1} = A^T$.*

Arguing as above for two dimensions we can see that:

- The $n\times n$ orthogonal matrices are the linear isometries of $\mathbb{R}^n$.

It is also relatively easy to note at this point that:

Proposition 65 *Let A be an $n\times n$ orthogonal matrix and $\mathbf{b}\in\mathbb{R}^n$. The map*

$$T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

is an isometry of $\mathbb{R}^n$.

Proof As A preserves lengths, being orthogonal, we have

$$|T(\mathbf{x}) - T(\mathbf{y})| = |(A\mathbf{x} + \mathbf{b}) - (A\mathbf{y} + \mathbf{b})| = |A(\mathbf{x} - \mathbf{y})| = |\mathbf{x} - \mathbf{y}|.$$

■

We shall see that, in fact, the converse also applies and that all isometries of $\mathbb{R}^n$ have this form.

And we also note the following, though leave the proofs to the linear algebra course:

Proposition 66 *Let A and B be $n \times n$ orthogonal matrices.*
(a) AB is orthogonal.
(b) A^{-1} is orthogonal.
(c) $\det A = \pm 1$.

What, then, are the 2×2 orthogonal matrices?

Example 67 ***(Orthogonal** 2×2 **Matrices)*** *Let A be a 2×2 orthogonal matrix. If*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{then} \quad I_2 = A^T A = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}.$$

So the orthogonality of A imposes three equations on its entries, namely

$$a^2 + c^2 = 1; \qquad b^2 + d^2 = 1; \qquad ab + cd = 0. \tag{3.10}$$

Note that the first two equations require the columns of A, namely $(a, c)^T$ and $(b, d)^T$, to be of unit length and the third equation requires them to be perpendicular to one another. As $(a, c)^T$ is of unit length then there is unique θ in the range $0 \leqslant \theta < 2\pi$ such that $a = \cos\theta$ and $c = \sin\theta$. Then, as $(b, d)^T$ is also of unit length and perpendicular to $(a, c)^T$, we have two possibilities

$$(b, d)^T = (\cos(\theta \pm \pi/2), \sin(\theta \pm \pi/2)) = (\mp\sin\theta, \pm\cos\theta).$$

Thus we have shown, either

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = R_\theta \qquad \text{or} \qquad A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} = S_{\theta/2}$$

for some unique θ in the range $0 \leqslant \theta < 2\pi$. From Example 62 the former represents rotation anticlockwise by θ about the origin, and the latter represents reflection in the line $y = x\tan(\theta/2)$.

So a 2×2 orthogonal matrix A is either a rotation (when $\det A = 1$) or a reflection (when $\det A = -1$). This is not generally the case for orthogonal matrices in higher dimensions – for example $-I_3$ is orthogonal but not a rotation, as it has determinant -1, but is not a reflection as it fixes only the origin.

3.4 Co-ordinates and Measurement

We saw in the first chapter how a plane Π may be parametrized as

$$\mathbf{r}(\lambda,\mu)=\mathbf{p}+\lambda\mathbf{a}+\mu\mathbf{b} \qquad (\lambda,\mu\in\mathbb{R})$$

where $\mathbf{p}$ is the position vector of some point P in Π and $\mathbf{a}$ and $\mathbf{b}$ are two independent vectors parallel to Π. This parametrization assigns unique co-ordinates λ and μ to each point in Π but it is only one of infinitely many parametrizations for Π. We have essentially chosen P, which is the point $\mathbf{r}(0,0)$, as an origin for Π, a λ-axis (where $\mu=0$) in the direction of $\mathbf{a}$ and a μ-axis (where $\lambda=0$) in the direction of $\mathbf{b}$. The plane Π is featureless and our choice of origin was entirely arbitrary and in choosing axes we simply had to make sure the axes were different lines in the plane.

This is a useful way to 'get a handle' on Π. We can perform calculations with these co-ordinates and describe subsets of Π in terms of them, but it does return us to the question raised earlier – to what extent do our calculations and descriptions of Π depend on our choice of co-ordinates? For example, would two different mathematicians working with Π with different co-ordinates agree on the truth or falsity of statements such as

- the first of two given line segments is the longer;
- the distance from $(1,1)$ to $(4,5)$ is five units long;
- the distance between two given points in one unit long;
- the points (λ,μ) satisfying $\lambda^2+\mu^2=1$ form a circle.

The two ought to agree on the first statement. This is a genuinely geometric problem – one or other of the line segments is longer and the mathematicians ought to agree irrespective of what co-ordinates and units they're using. They will also agree on the second statement although, in general, they will be measuring between different pairs of points – and unless they are using the same units of length the measured distances of '5 units' will, in fact, be different.

There will typically be disagreement, however, for the last two statements. If the two are using different units then the third statement cannot be simultaneously correct for the two. And the locus in the fourth statement will be a circle for a mathematician who took $\mathbf{a}$ and $\mathbf{b}$ to be perpendicular and of the same length but more generally this locus will be an ellipse in Π.

Our problem then is this: the world about us doesn't come with any natural co-ordinates – and even if we have a problem set in $\mathbb{R}^3$, with its given xyz-co-ordinates, it may be in our interest to choose co-ordinates, more naturally relating to the geometric scenario at hand, which simplify calculation and understanding. Will our formulae from Chapter 1 for length and angle – which were in terms of co-ordinates – still be the right ones?

The first definition we need is that of a *basis*. With a basis we can assign co-ordinates uniquely to each vector.

Definition 68 *We say that n vectors $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_n$ are a* ***basis*** *for $\mathbb{R}^n$ if for every $\mathbf{v}\in\mathbb{R}^n$ there exist unique real numbers $\alpha_1,\alpha_2,\ldots,\alpha_n$ such that*

$$\mathbf{v}=\alpha_1\mathbf{v}_1+\alpha_2\mathbf{v}_2+\cdots+\alpha_n\mathbf{v}_n.$$

We shall refer to $\alpha_1, \alpha_2, ..., \alpha_n$ *as the* ***co-ordinates*** *of* $\mathbf{v}$ *with respect to the basis* $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$.

The **standard basis** or **canonical basis** for $\mathbb{R}^n$ is the set of vectors

$$\mathbf{e}_1 = (1, 0, 0, ..., 0), \quad \mathbf{e}_2 = (0, 1, 0, ..., 0), \quad ... \quad \mathbf{e}_n = (0, 0, ..., 0, 1)$$

and we see that every vector $\mathbf{x} = (x_1, ..., x_n) = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$ so that what we've been calling 'the' co-ordinates of $\mathbf{x}$ are the co-ordinates of $\mathbf{x}$ with respect to the standard basis.

Note that if $\mathbf{x}$ and $\mathbf{y}$ respectively have co-ordinates $x_1, x_2, ..., x_n$ and $y_1, y_2, ..., y_n$ with respect to some basis then $\mathbf{x} + \mathbf{y}$ has co-ordinates $x_1 + y_1, ..., x_n + y_n$ and $\lambda\mathbf{x}$ has co-ordinates $\lambda x_1, ..., \lambda x_n$ with respect to the same basis.

- Bases tend to be defined in linear algebra courses as sets that are linearly independent and spanning. From our point of view if a set of vectors didn't span $\mathbb{R}^n$ then we'd be unable to assign co-ordinates to every vector and if the set wasn't linearly independent then some points would have more than one set of co-ordinates associated with them.

So with a choice of basis we can associate unique co-ordinates to each vector – the point being that we need co-ordinates if we're to calculate the lengths of vectors and angles between them. We introduced formulae for length and angle in Chapter 1 in terms of co-ordinates. We'll see that these formulae need not be correct when we use arbitrary bases – we will need to introduce **orthonormal bases.** Note that the formulae for length and angle were in terms of the scalar product and so we're interested in those bases such that

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

holds true for the scalar product of two vectors in terms of their co-ordinates.

Example 69 *The vectors* $\mathbf{v}_1 = (1, 0, 1), \mathbf{v}_2 = (1, 1, 0), \mathbf{v}_3 = (0, 1, 1)$ *form a basis for* $\mathbb{R}^3$ *(check this!). The vectors with co-ordinates* x_i *and* y_i *with respect to this basis have the scalar product*

$$\begin{aligned}\mathbf{x} \cdot \mathbf{y} &= (x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3) \cdot (y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + y_3\mathbf{v}_3) \\ &= \sum_{i,j} x_iy_j\mathbf{v}_i \cdot \mathbf{v}_j \\ &= 2(x_1y_1 + x_2y_2 + x_3y_3) + x_1y_2 + x_2y_1 + x_2y_3 + x_3y_2 + x_3y_1 + x_1y_3.\end{aligned}$$

Proposition 70 *Let* $\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_n$ *be a basis for* $\mathbb{R}^n$*. Then the equation*

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

holds for any two vectors $\mathbf{x}$ *and* $\mathbf{y}$ *which have co-ordinates* x_i *and* y_i *with respect to* $\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_n$ *if and only if*

$$\mathbf{v}_i \cdot \mathbf{v}_i = 1 \text{ for each } i \quad \text{and} \quad \mathbf{v}_i \cdot \mathbf{v}_j = 0 \text{ when } i \neq j.$$

Proof If we suppose that we calculate dot products in terms of co-ordinates with the above formula, then note that ith co-ordinate of $\mathbf{v}_i$ is 1 and each of the others is zero – so $\mathbf{v}_i \cdot \mathbf{v}_i = 1$ for each i and $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ when $i \neq j$ follows immediately by putting these values into the given formula.

Conversely if $\mathbf{v}_i \cdot \mathbf{v}_i = 1$ for each i and $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ when $i \neq j$ then the required expression comes from expanding

$$\begin{aligned} & (x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_n\mathbf{v}_n) \cdot (y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \cdots + y_n\mathbf{v}_n) \\ = & \left(\sum_{i=1}^n x_i\mathbf{v}_i\right) \cdot \left(\sum_{j=1}^n y_j\mathbf{v}_j\right) \\ = & \sum_{i=1}^n \sum_{j=1}^n x_i y_j \mathbf{v}_i \cdot \mathbf{v}_j \\ = & \sum_{i=1}^n \sum_{j=1}^n x_i y_j \delta_{ij} \\ = & x_1y_1 + x_2y_2 + \cdots + x_ny_n. \end{aligned}$$

■

Definition 71 *A basis $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ for $\mathbb{R}^n$ is said to be **orthonormal** if*

$$\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

- n orthonormal vectors in $\mathbb{R}^n$ form a basis. (In fact, orthogonality alone is sufficient to guarantee linear independence.)
- If $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ are an orthonormal basis for $\mathbb{R}^n$ and

  $$\mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$$

  then note that $\alpha_i = \mathbf{x} \cdot \mathbf{v}_i$.

We may now note that orthonormal bases are then intimately related to orthogonal matrices.

Proposition 72 *An $n \times n$ matrix A is orthogonal if and only if its columns form an orthonormal basis for $\mathbb{R}^n$. The same result hold true for the rows of A.*

Proof Say that A has columns $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ then

$$\begin{aligned} A \text{ is orthogonal} \quad & \iff \quad A^T A = I_n \\ & \iff \begin{pmatrix} \longleftarrow & \mathbf{v}_1 & \rightarrow \\ \vdots & \vdots & \vdots \\ \longleftarrow & \mathbf{v}_n & \rightarrow \end{pmatrix} \begin{pmatrix} \uparrow & \cdots & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ \downarrow & \cdots & \downarrow \end{pmatrix} = I_n \\ & \iff \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \cdots & \mathbf{v}_1 \cdot \mathbf{v}_n \\ \vdots & \vdots & \vdots \\ \mathbf{v}_n \cdot \mathbf{v}_1 & \cdots & \mathbf{v}_n \cdot \mathbf{v}_n \end{pmatrix} = I_n \\ & \iff \quad \mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij}. \end{aligned}$$

■

If we return to the basis in Example 69, we can note that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = (x_1 + x_2)\mathbf{i} + (x_2 + x_3)\mathbf{j} + (x_1 + x_3)\mathbf{k}.$$

So if x_1, x_2, x_3 are the co-ordinates associated with the basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and x, y, z are the co-ordinates associated with the standard basis $\mathbf{i}$, $\mathbf{j}, \mathbf{k}$ then these co-ordinate systems are related by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Denote this 3×3 matrix as P, noting that it has the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ as its columns. Now if vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^3$ have co-ordinates $\mathbf{X}$ and $\mathbf{Y}$ with respect to $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, then $\mathbf{x} = P\mathbf{X}$ and $\mathbf{y} = P\mathbf{Y}$. We then see

$$\begin{aligned} \mathbf{x} \cdot \mathbf{y} &= \mathbf{x}^T \mathbf{y} \\ &= (P\mathbf{X})^T (P\mathbf{Y}) \\ &= \mathbf{X} P^T P \mathbf{Y} \\ &= \mathbf{X}^T \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \mathbf{Y} \\ &= (X_1, X_2, X_3)^T \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} \\ &= 2X_1Y_1 + 2X_2Y_2 + 2X_3Y_3 + X_1Y_2 + X_1Y_3 + X_2X_1 + X_2Y_3 + X_3Y_1 + X_3Y_2. \end{aligned}$$

So at this point we can note the following, as found in Example 69, though we will leave the general details to the Linear Algebra I course.

- Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis for $\mathbb{R}^n$ and $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis vectors. The standard co-ordinates $\mathbf{x}$ are related to the co-ordinates $\mathbf{X}$ by

  $$\mathbf{x} = P\mathbf{X}$$

  where P is the $n \times n$ matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

- $\mathbf{x} \cdot \mathbf{y} = \mathbf{X} \cdot \mathbf{Y}$ if and only if P is orthogonal or equivalently the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are an orthonormal basis.

Proposition 73 *Let A be an orthogonal 3×3 matrix, and $\mathbf{x}$, $\mathbf{y}$ be column vectors in $\mathbb{R}^3$.*
(a) If $\det A = 1$ then $A(\mathbf{x} \wedge \mathbf{y}) = A\mathbf{x} \wedge A\mathbf{y}$.
(b) If $\det A = -1$ then $A(\mathbf{x} \wedge \mathbf{y}) = -A\mathbf{x} \wedge A\mathbf{y}$.

Proof For any vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in $\mathbb{R}^3$ we have

$$\begin{aligned} [A\mathbf{x}, A\mathbf{y}, \mathbf{z}] &= \left[A\mathbf{x}, A\mathbf{y}, A(A^{-1}\mathbf{z})\right] \\ &= \det A \times \left[\mathbf{x}, \mathbf{y}, A^{-1}\mathbf{z}\right] \qquad \text{[Sheet 2, Exercise 4]} \\ &= \det A \times \left[\mathbf{x}, \mathbf{y}, A^T\mathbf{z}\right]. \end{aligned}$$

Now $\det A = \pm 1$ and we find

$$\begin{aligned}(A\mathbf{x} \wedge A\mathbf{y}) \cdot \mathbf{z} &= \pm(\mathbf{x} \wedge \mathbf{y}) \cdot A^T\mathbf{z} \\ &= \pm(\mathbf{x} \wedge \mathbf{y})^T A^T \mathbf{z} \\ &= \pm(A(\mathbf{x} \wedge \mathbf{y}))^T \mathbf{z} \\ &= \pm(A(\mathbf{x} \wedge \mathbf{y})) \cdot \mathbf{z}.\end{aligned}$$

By Sheet 1, Exercise 1(iii), $A(\mathbf{x} \wedge \mathbf{y}) = \pm A\mathbf{x} \wedge A\mathbf{y}$. ■

3.5 Orthogonal Change of Variable. Spectral Theorem.

Recall that when we make an invertible linear change of this means our original co-ordinates $\mathbf{x}$ and our new co-ordinates $\mathbf{X}$ are related by a matrix P such that

$$\mathbf{x} = P\mathbf{X}$$

and that the columns of P are the basis vectors for the new co-ordinates. If we want to still determine angles and distances with the usual formulae involving co-ordinates, then we require P to be orthogonal.

We have just seen how an orthogonal change of co-ordinates (specifically a rotation of the axes) can transform the locus of $Ax^2 + Bxy + Cy^2 = 1$ into a normal form. We can connect this more with the matrix approach by rewriting this equation as

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1.$$

Notice, importantly, that the 2×2 matrix is symmetric. We then changed to new co-ordinates X, Y given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} X \\ Y \end{pmatrix}$$

where P is orthogonal. Our equation would then in the new X, Y co-ordinates would read as

$$\begin{pmatrix} X & Y \end{pmatrix} P^T \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} P \begin{pmatrix} X \\ Y \end{pmatrix} = 1$$

and, for the right choice of θ, we saw this equation could be put in the form

$$\begin{pmatrix} X & Y \end{pmatrix} \begin{pmatrix} (A+C+H)/2 & 0 \\ 0 & (A+C-H)/2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = 1.$$

In particular it was the case that

$$P^T \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} P = \begin{pmatrix} (A+C+H)/2 & 0 \\ 0 & (A+C-H)/2 \end{pmatrix}$$

is a *diagonal* matrix. This is our first instance of seeing a very important theorem of mathematics, the *Spectral Theorem.*

Theorem 74 ***(Spectral Theorem – finite dimensional case)*** *Let A be a square real symmetric matrix (so $A^T = A$). Then there is an orthogonal matrix P such that $P^T AP$ is diagonal.*

Proof This will be proved in the Linear Algebra II course. ■

3.6 3×3 Orthogonal Matrices.

Example 75 *Given that one of the matrices below describes a rotation of $\mathbb{R}^3$, one a reflection of $\mathbb{R}^3$, determine which is which. Determine the axis of the rotation, and the invariant plane of the reflection.*

$$A = \frac{1}{25}\begin{pmatrix} 20 & 15 & 0 \\ -12 & 16 & 15 \\ 9 & -12 & 20 \end{pmatrix}; \qquad B = \frac{1}{25}\begin{pmatrix} -7 & 0 & -24 \\ 0 & 25 & 0 \\ -24 & 0 & 7 \end{pmatrix}.$$

Solution If we consider the equations $A\mathbf{x} = \mathbf{x}$ and $B\mathbf{x} = \mathbf{x}$ then the rotation will have a one-dimensional solution space and the reflection's will be two-dimensional. Reduction gives

$$A - I = \frac{1}{25}\begin{pmatrix} -5 & 15 & 0 \\ -12 & -9 & 15 \\ 9 & -12 & -5 \end{pmatrix} \xrightarrow{\text{RRE}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1/3 \\ 0 & 0 & 0 \end{pmatrix};$$

$$B - I = \frac{1}{25}\begin{pmatrix} -32 & 0 & -24 \\ 0 & 0 & 0 \\ -24 & 0 & -18 \end{pmatrix} \xrightarrow{\text{RRE}} \begin{pmatrix} 4 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

So A is the rotation and B is the reflection. The null space of $A - I$ consists of multiples of $(3, 1, 3)^T$, so this is parallel to the line of rotation. We see the invariant plane of the reflection B has equation $4x + 3z = 0$. ■

Example 76 *Let B be the matrix in the previous example. Determine an orthonormal basis $\mathbf{w}_1$, $\mathbf{w}_2$ for the invariant plane of B and extend it to an orthonormal basis $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ for $\mathbb{R}^3$. What is the map B in terms of the co-ordinates associated with $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$?*

Solution The invariant plane of B has equation $4x + 3z = 0$. This has an orthonormal basis

$$\mathbf{w}_1 = \frac{1}{5}\begin{pmatrix} -3 \\ 0 \\ 4 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

This can be extended to an orthonormal basis for $\mathbb{R}^3$ by taking

$$\mathbf{w}_3 = \mathbf{w}_2 \wedge \mathbf{w}_1 = \frac{1}{5}\begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}.$$

Note that $B\mathbf{w}_1 = \mathbf{w}_1$ and $B\mathbf{w}_2 = \mathbf{w}_2$ as $\mathbf{w}_1$ and $\mathbf{w}_2$ are in the (invariant) plane of reflection. Also

$$B\mathbf{w}_3 = \frac{1}{25}\begin{pmatrix} -7 & 0 & -24 \\ 0 & 25 & 0 \\ -24 & 0 & 7 \end{pmatrix}\frac{1}{5}\begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix} = \frac{1}{125}\begin{pmatrix} -100 \\ 0 \\ -75 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -4 \\ 0 \\ -3 \end{pmatrix} = -\mathbf{w}_3.$$

Hence, in terms of co-ordinates X_1, X_2, X_3 associated with $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ we see that

$$X_1\mathbf{w}_1 + X_2\mathbf{w}_2 + X_3\mathbf{w}_3 \overset{B}{\longmapsto} X_1\mathbf{w}_1 + X_2\mathbf{w}_2 - X_3\mathbf{w}_3$$

or if we wished to capture this as a matrix then

$$B\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \\ -X_3 \end{pmatrix} \qquad \Longrightarrow \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

■

When considering a rotation R of $\mathbb{R}^3$ (about an axis through the origin) then we might take a unit vector $\mathbf{v}_1$ parallel to the axis It follows that $R\mathbf{v}_1 = \mathbf{v}_1$. If we extend $\mathbf{v}_1$ to an orthonormal basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ then (from our knowledge of rotations in two dimensions) there exists θ such that

$$R\mathbf{v}_2 = \cos\theta\mathbf{v}_2 + \sin\theta\mathbf{v}_3, \qquad R\mathbf{v}_3 = -\sin\theta\mathbf{v}_2 + \cos\theta\mathbf{v}_3.$$

In terms of co-ordinates X_1, X_2, X_3 associated with $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ we have

$$R\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2\cos\theta - X_3\sin\theta \\ X_2\sin\theta + X_3\cos\theta \end{pmatrix} \qquad \Longrightarrow \qquad R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

The general situation for 3×3 orthogonal matrices is resolved in the theorem below.

Theorem 77 *(**Classifying** 3×3 **orthogonal matrices**) Let A be a 3×3 orthogonal matrix.*
(a) If $\det A = 1$ *then A is a rotation of $\mathbb{R}^3$ about some axis by an angle θ where*

$$\operatorname{trace} A = 1 + 2\cos\theta.$$

(b) If $\det A = -1$ *and* $\operatorname{trace} A = 1$ *then A is a reflection of $\mathbb{R}^3$. The converse also holds.*

Proof (a) We firstly show that when $\det A = 1$ there exists a non-zero vector $\mathbf{x}$ in $\mathbb{R}^3$ such that $A\mathbf{x} = \mathbf{x}$. Note

$$\begin{aligned} \det(A - I) &= \det((A - I)A^T) \qquad [\det A^T = \det A = 1] \\ &= \det(I - A^T) \\ &= \det((I - A)^T) \\ &= \det(I - A) \qquad [\det M = \det M^T] \\ &= (-1)^3 \det(A - I) \qquad [\det(\lambda M) = \lambda^3 \det M] \\ &= -\det(A - I). \end{aligned}$$

Hence $\det(A - I) = 0$, so that $A - I$ is singular and there exists a non-zero vector $\mathbf{x}$ such that $(A - I)\mathbf{x} = \mathbf{0}$ as required. If we set $\mathbf{v}_1 = \mathbf{x}/|\mathbf{x}|$ then we can extend $\mathbf{v}_1$ to an orthonormal basis $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ for $\mathbb{R}^3$.

Let X_1, X_2, X_3 be the co-ordinates associated with $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$. We already have that $A\mathbf{v}_1 = \mathbf{v}_1$ and as A is orthogonal we also have

$$\begin{aligned} A\mathbf{v}_2 \cdot \mathbf{v}_1 &= A\mathbf{v}_2 \cdot A\mathbf{v}_1 = \mathbf{v}_2 \cdot \mathbf{v}_1 = 0, \\ A\mathbf{v}_3 \cdot \mathbf{v}_1 &= A\mathbf{v}_3 \cdot A\mathbf{v}_1 = \mathbf{v}_3 \cdot \mathbf{v}_1 = 0. \end{aligned}$$

So that

$$A\mathbf{v}_1 = \mathbf{v}_1, \qquad A\mathbf{v}_2 = a_{22}\mathbf{v}_2 + a_{32}\mathbf{v}_3, \qquad A\mathbf{v}_3 = q_{23}\mathbf{v}_2 + q_{33}\mathbf{v}_3.$$

Now as A is orthogonal the vectors $A\mathbf{v}_2$ and $A\mathbf{v}_3$ must be of unit length and perpendicular. This means that

$$\tilde{A} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$$

is itself an orthogonal matrix. Further as $\det A = 1$ then $\det \tilde{A} = 1$ and so $\tilde{A} = R_\theta$ for some θ by Example 67. Hence, in terms of the X_1, X_2, X_3 co-ordinates, A is described by the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Note that the trace of this matrix is $1 + 2\cos\theta$ and as trace is invariant under the co-ordinate change (this result is left to Linear Algebra I) then

$$\operatorname{trace} A = 1 + 2\cos\theta.$$

(b) Say now that $\det A = -1$ and $\operatorname{trace} A = 1$ Let $C = -A$. Then C is orthogonal and $\det C = (-1)^3 \det A = 1$. By (a) there exists non-zero $\mathbf{x}$ such that $C\mathbf{x} = \mathbf{x}$ and hence $A\mathbf{x} = -\mathbf{x}$. If we proceed as in part (a) we can introduce orthonormal co-ordinates so that

$$A = -C = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\cos\theta & \sin\theta \\ 0 & -\sin\theta & -\cos\theta \end{pmatrix}$$

for some $0 \leqslant \theta < 2\pi$. Now

$$1 = \operatorname{trace} A = \operatorname{trace} P^T A P = -1 - 2\cos\theta$$

showing that $\theta = \pi$, and hence

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

which describes reflection in the plane $x = 0$. Conversely for a reflection (described by a matrix) there is an orthonormal choice of co-ordinates with respect to which the reflection has the matrix $\operatorname{diag}(-1, 1, 1)$. As both determinant and trace are invariant under such a change of co-ordinates then we see the reflection's determinant equals -1 and its trace equals 1. ■

Example 78 *Let* $\mathbf{n}$ *be a unit vector. Show that reflection in the plane* $\mathbf{r}\cdot\mathbf{n}=c$ *is given by*

$$R_c(\mathbf{v}) = \mathbf{v} - 2(\mathbf{v}\cdot\mathbf{n})\mathbf{n} + 2c\mathbf{n}.$$

Let R_c *denote reflection in the plane* $2x+y+2z=c$. *Show that*

$$R_c\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 1 & -4 & -8 \\ -4 & 7 & -4 \\ -8 & -4 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{2c}{3}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

Solution The orthogonal projection of $\mathbf{v}$ on to the plane $\mathbf{r}\cdot\mathbf{n}=c$ is the vector of the form $\mathbf{v}+\lambda\mathbf{n}$ which lies in the plane. So

$$(\mathbf{v}+\lambda\mathbf{n})\cdot\mathbf{n} = c \qquad \Longrightarrow \qquad \lambda = \frac{c-\mathbf{v}\cdot\mathbf{n}}{\mathbf{n}\cdot\mathbf{n}} = c - \mathbf{v}\cdot\mathbf{n}$$

The reflection of $\mathbf{v}$ in the plane $\mathbf{r}\cdot\mathbf{n}=c$ is $\mathbf{v}+2\lambda\mathbf{n}$ which equals

$$\mathbf{v} + 2\left(c - \mathbf{v}\cdot\mathbf{n}\right)\mathbf{n} = \mathbf{v} - 2(\mathbf{v}\cdot\mathbf{n})\mathbf{n} + 2c\mathbf{n}.$$

Let R_c denote reflection in the plane $2x+y+2z=3c$. We may choose $\mathbf{n}=\frac{1}{3}(2,1,2)$, so that

$$\begin{aligned} R_c\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \left(\frac{4x+2y+4z}{3}\right)\begin{pmatrix} 2/3 \\ 1/3 \\ 2/3 \end{pmatrix} + 2c\begin{pmatrix} 2/3 \\ 1/3 \\ 2/3 \end{pmatrix} \\ &= \frac{1}{9}\begin{pmatrix} 1 & -4 & -8 \\ -4 & 7 & -4 \\ -8 & -4 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \frac{2c}{3}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \end{aligned}$$

as required. Note that R_0 is orthogonal. ■

Remark 79 *In the previous examples and proof we assumed that given one or two orthonormal vectors in* $\mathbb{R}^3$ *these could be extended to an orthonormal basis. But we did not prove that this could always be done. This is in fact the case in* $\mathbb{R}^n$ *and there is a more general result – the* ***Gram-Schmidt orthogonalization process*** *– which can produce an orthonormal set from a linearly independent set. This result is rigorously treated in Linear Algebra II.*

In three dimensions though this is not a hard result to visualize. Given independent vectors $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ *in* $\mathbb{R}^3$ *we can produce a first unit vector* $\mathbf{v}_1$ *as*

$$\mathbf{v}_1 = \frac{\mathbf{w}_1}{|\mathbf{w}_1|}.$$

The vectors $\mathbf{v}_1$ *and* $\mathbf{w}_2$ *span a plane, with* $\mathbf{v}_1$ *spanning a line which splits the plane into two half-planes. We take* $\mathbf{v}_2$ *to be the unit vector perpendicular to* $\mathbf{v}_1$ *which points into the same half-plane as* $\mathbf{w}_2$ *does. Specifically this is the vector* $\mathbf{v}_2 = \mathbf{y}_2/|\mathbf{y}_2|$ *where*

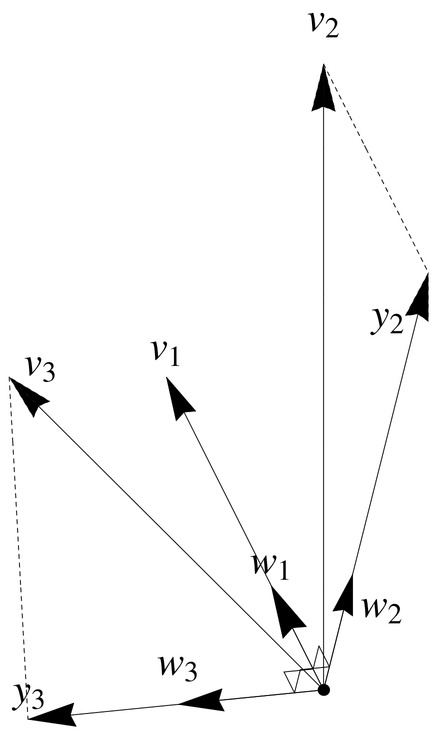
$$\mathbf{y}_2 = \mathbf{w}_2 - (\mathbf{w}_2\cdot\mathbf{v}_1)\mathbf{v}_1.$$

That is $\mathbf{y}_2$ *is the component of* $\mathbf{w}_2$ *perpendicular to* $\mathbf{v}_1$.

Now $\mathbf{v}_1$ *and* $\mathbf{v}_2$ *span a plane which splits* $\mathbb{R}^3$ *into two half-spaces. We take* $\mathbf{v}_3$ *to be the unit vector perpendicular to that plane and which points into the same half-space as* $\mathbf{w}_3$*. Specifically this is the vector* $\mathbf{v}_3 = \mathbf{y}_3 / |\mathbf{y}_3|$ *where*

$$\mathbf{y}_3 = \mathbf{w}_3 - (\mathbf{w}_3 \cdot \mathbf{v}_1)\mathbf{v}_1 - (\mathbf{w}_3 \cdot \mathbf{v}_2)\mathbf{v}_2.$$

That is $\mathbf{y}_3$ *is the component of* $\mathbf{w}_3$ *perpendicular to* $\mathbf{v}_1$ *and* $\mathbf{v}_2$*.*



22. Gram-Schmidt process in 3d

3.7 Isometries of $\mathbb{R}^n$

Definition 80 *A map T from $\mathbb{R}^n$ to $\mathbb{R}^n$ is said to be an* ***isometry*** *if it preserves distances – that is if*

$$|T(\mathbf{v}) - T(\mathbf{w})| = |\mathbf{v} - \mathbf{w}| \qquad \textit{for any } \mathbf{v}, \mathbf{w} \textit{ in } \mathbb{R}^n.$$

Example 81 *(a) If A is an orthogonal matrix then $\mathbf{x} \mapsto A\mathbf{x}$ is an isometry.*
(b) Given $\mathbf{c} \in \mathbb{R}^n$ then ***translation*** *by $\mathbf{c}$, that is the map $T(\mathbf{v}) = \mathbf{v} + \mathbf{c}$, is an isometry as*

$$|T(\mathbf{v}) - T(\mathbf{w})| = |(\mathbf{v} + \mathbf{c}) - (\mathbf{w} + \mathbf{c})| = |\mathbf{v} - \mathbf{w}|.$$

(c) Reflection in the plane $\mathbf{r} \cdot \mathbf{n} = c$ (where $\mathbf{n}$ is a unit vector) is given by

$$R_c(\mathbf{v}) = R_0(\mathbf{v}) + 2c\mathbf{n} = \mathbf{v} - 2(\mathbf{v} \cdot \mathbf{n})\mathbf{n} + 2c\mathbf{n}.$$

We see that this is an isometry as

$$|R_0(\mathbf{v})|^2 = (\mathbf{v} - 2(\mathbf{v} \cdot \mathbf{n})\mathbf{n}) \cdot (\mathbf{v} - 2(\mathbf{v} \cdot \mathbf{n})\mathbf{n}) = |\mathbf{v}|^2 - 4(\mathbf{v} \cdot \mathbf{n})^2 + 4(\mathbf{v} \cdot \mathbf{n})(\mathbf{n} \cdot \mathbf{n}) = |\mathbf{v}|^2,$$

and more generally

$$\begin{aligned} R_c(\mathbf{v}) - R_c(\mathbf{w}) &= (\mathbf{v} - 2(\mathbf{v}\cdot\mathbf{n})\mathbf{n} + 2c\mathbf{n}) - (\mathbf{w} - 2(\mathbf{w}\cdot\mathbf{n})\mathbf{n} + 2c\mathbf{n}) \\ &= (\mathbf{v}-\mathbf{w}) - 2((\mathbf{v}-\mathbf{w})\cdot\mathbf{n})\mathbf{n} \\ &= R_0(\mathbf{v}-\mathbf{w}). \end{aligned}$$

The composition of two isometries is still an isometry, so any map of the form $T(\mathbf{v}) = A\mathbf{v}+\mathbf{b}$ is an isometry where A is an orthogonal matrix and $\mathbf{b} \in \mathbb{R}^n$. We see now that in fact all isometries of $\mathbb{R}^n$ take this form.

Proposition 82 *Let S be an isometry from $\mathbb{R}^n$ to $\mathbb{R}^n$ such that $S(\mathbf{0}) = \mathbf{0}$. Then*
(a) $|S(\mathbf{v})| = |\mathbf{v}|$ for any $\mathbf{v}$ in $\mathbb{R}^n$ and $S(\mathbf{u})\cdot S(\mathbf{v}) = \mathbf{u}\cdot\mathbf{v}$ for any $\mathbf{u}$, $\mathbf{v}$ in $\mathbb{R}^n$.
(b) If $\mathbf{v}_1, ..., \mathbf{v}_n$ is an orthonormal basis for $\mathbb{R}^n$ then so is $S(\mathbf{v}_1), ..., S(\mathbf{v}_n)$.
(c) There exists an orthogonal matrix A such that $S(\mathbf{v}) = A\mathbf{v}$ for each $\mathbf{v}$ in $\mathbb{R}^n$.

Proof (a) Note

$$|S(\mathbf{v})| = |S(\mathbf{v}) - \mathbf{0}| = |S(\mathbf{v}) - S(\mathbf{0})| = |\mathbf{v} - \mathbf{0}| = |\mathbf{v}|$$

as S is an isometry that fixes $\mathbf{0}$. We further have for any $\mathbf{u}$, $\mathbf{v}$ in $\mathbb{R}^n$

$$\begin{aligned} \mathbf{u}\cdot\mathbf{v} &= \frac{1}{2}(|\mathbf{u}|^2 + |\mathbf{v}|^2 - |\mathbf{u}-\mathbf{v}|^2) \\ &= \frac{1}{2}\left(|S(\mathbf{u})|^2 + |S(\mathbf{v})|^2 - |S(\mathbf{u}) - S(\mathbf{v})|^2\right) \\ &= S(\mathbf{u})\cdot S(\mathbf{v}). \end{aligned}$$

(b) This follows immediately from (a).

(c) Let $\mathbf{e}_1, ..., \mathbf{e}_n$ denote the standard basis for $\mathbb{R}^n$ and suppose for now that $S(\mathbf{e}_i) = \mathbf{e}_i$ for each i. For a given $\mathbf{v}$ in $\mathbb{R}^n$ there exist unique λ_i and μ_i such that

$$\mathbf{v} = \lambda_1\mathbf{e}_1 + \cdots + \lambda_n\mathbf{e}_n \quad \text{and} \quad S(\mathbf{v}) = \mu_1\mathbf{e}_1 + \cdots + \mu_n\mathbf{e}_n.$$

Now note by (a) that

$$\mu_i = S(\mathbf{v})\cdot\mathbf{e}_i = S(\mathbf{v})\cdot S(\mathbf{e}_i) = \mathbf{v}\cdot\mathbf{e}_i = \lambda_i.$$

Hence $S(\mathbf{v}) = \mathbf{v}$ for each $\mathbf{v}$ in $\mathbb{R}^n$.

Now, without the initial assumption that $S(\mathbf{e}_i) = \mathbf{e}_i$, let A be the matrix with columns $S(\mathbf{e}_1), S(\mathbf{e}_2), \cdots, S(\mathbf{e}_n)$. As $S(\mathbf{e}_1), ..., S(\mathbf{e}_n)$ is an orthonormal basis for $\mathbb{R}^n$ then A is orthogonal and hence so is A^T. So the map which sends $\mathbf{v}$ to $A^T S(\mathbf{v})$ is an isometry. As the vectors $S(\mathbf{e}_i)$ are orthonormal then

$$(A^T S)(\mathbf{e}_i) = \begin{pmatrix} (S(\mathbf{e}_1))^T \\ \vdots \\ (S(\mathbf{e}_n))^T \end{pmatrix} S(\mathbf{e}_i) = \begin{pmatrix} S(\mathbf{e}_1)\cdot S(\mathbf{e}_i) \\ \vdots \\ S(\mathbf{e}_n)\cdot S(\mathbf{e}_i) \end{pmatrix} = \mathbf{e}_i$$

and hence $A^T S(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v}$ in $\mathbb{R}^n$ by the previous argument. Thus we have

$$S(\mathbf{v}) = (A^T)^{-1}(\mathbf{v}) = A\mathbf{v} \quad \text{for all } \mathbf{v}.$$

We have already noted that A is orthogonal. ■

Theorem 83 ***(Classifying Isometries of*** $\mathbb{R}^n$*)* *Let T be an isometry from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then there is a orthogonal matrix A and a column vector $\mathbf{b}$ such that $T(\mathbf{v}) = A\mathbf{v} + \mathbf{b}$ for all $\mathbf{v}$. Further A and $\mathbf{b}$ are unique in this regard.*

Proof The map $S(\mathbf{v}) = T(\mathbf{v}) - T(\mathbf{0})$ is an isometry which fixes $\mathbf{0}$. Then there is an orthogonal matrix A such that $S(\mathbf{v}) = A\mathbf{v}$ giving $T(\mathbf{v}) = A\mathbf{v} + T(\mathbf{0})$. To show uniqueness, suppose $T(\mathbf{v}) = A_1\mathbf{v} + \mathbf{b}_1 = A_2\mathbf{v} + \mathbf{b}_2$ for all $\mathbf{v}$. Setting $\mathbf{v} = \mathbf{0}$ we see $\mathbf{b}_1 = \mathbf{b}_2$. Then $A_1\mathbf{v} = A_2\mathbf{v}$ for all $\mathbf{v}$ and hence $A_1 = A_2$. ■

Example 84 ***(Euler Angles)*** *Let R denote an orthogonal 3×3 matrix with $\det R = 1$ and let*

$$R(\mathbf{i}, \theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \qquad R(\mathbf{j}, \theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}.$$

(i) Suppose that $R\mathbf{i} = \mathbf{i}$. Show that R is of the form $R(\mathbf{i}, \theta)$ for some θ in the range $-\pi < \theta \leqslant \pi$.
(ii) For general R, show that there exist α, β in the ranges $-\pi < \alpha \leqslant \pi$, $0 \leqslant \beta \leqslant \pi$, and c, d such that $d \geqslant 0$ and $c^2 + d^2 = 1$, with

$$R(\mathbf{i}, \alpha)^{-1} R\mathbf{i} = c\mathbf{i} + d\mathbf{k}, \qquad R(\mathbf{j}, \beta)^{-1} R(\mathbf{i}, \alpha)^{-1} R\mathbf{i} = \mathbf{i}.$$

(iii) Deduce that R can be expressed in the form

$$R = R(\mathbf{i}, \alpha)\, R(\mathbf{j}, \beta)\, R(\mathbf{i}, \gamma)$$

where $-\pi < \alpha \leqslant \pi$, $0 \leqslant \beta \leqslant \pi$ and $-\pi < \gamma \leqslant \pi$.

Solution (i) Suppose that $R\mathbf{i} = \mathbf{i}$. Then the first column of R's matrix is $(1, 0, 0)^T$. Further as the first row of R is a unit vector then R's first row is $(1, 0, 0)$. So we have $R = \operatorname{diag}(1, Q)$ for some 2×2 matrix Q. As the rows and columns of R are orthonormal the same then applies to Q so that Q is orthogonal. Further as $\det R = 1$ then $\det Q = 1$ and so we have and hence

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$

for some θ in the range $-\pi < \theta \leqslant \pi$.

(ii) In the absence of the condition $R\mathbf{i} = \mathbf{i}$, it still remains the case that $R\mathbf{i}$ is a unit vector as R is orthogonal. Say $R\mathbf{i} = (x, y, z)^T$ where $x^2 + y^2 + z^2 = 1$. We wish to find c, d, α such that $R(\mathbf{i}, \alpha)^{-1} R\mathbf{i} = c\mathbf{i} + d\mathbf{k}$ or equivalently

$$\begin{pmatrix} x \\ y\cos\alpha + z\sin\alpha \\ -y\sin\alpha + z\cos\alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} c \\ 0 \\ d \end{pmatrix}.$$

Hence we must set $c = x$. We further see that we need to choose α so that $\tan\alpha = -y/z$ and set $d = -y\sin\alpha + z\cos\alpha$. There are two choices of α in the range $-\pi < \alpha \leqslant \pi$ which differ by

π. Hence the two different α lead to the same value of d save for its sign and we should choose the α that leads to $d > 0$. (The exception to this is when $R\mathbf{i} = \mathbf{i}$ already, in which case any choice of α will do and we would have $d = 0$.)

We now need to determine β such that $R(\mathbf{j}, \beta)^{-1}(c\mathbf{i} + d\mathbf{k}) = \mathbf{i}$. This is equivalent to

$$\begin{pmatrix} c\cos\beta + d\sin\beta \\ 0 \\ -c\sin\beta + d\cos\beta \end{pmatrix} = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} c \\ 0 \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \qquad (3.11)$$

As $c^2 + d^2 = 1$ and $d \geqslant 0$ we see that there is unique β in the range $0 \leqslant \beta \leqslant \pi$ such that $c = \cos\beta$ and $d = \sin\beta$. For this choice of β we see that (3.11) is true.

(iii) For these choices of α and β we have

$$R(\mathbf{j}, \beta)^{-1} R(\mathbf{i}, \alpha)^{-1} R\mathbf{i} = \mathbf{i}.$$

So $R(\mathbf{j}, \beta)^{-1} R(\mathbf{i}, \alpha)^{-1} R$ is an orthogonal, determinant 1 matrix which fixes $\mathbf{i}$. By (i) we know that

$$R(\mathbf{j}, \beta)^{-1} R(\mathbf{i}, \alpha)^{-1} R = R(\mathbf{i}, \gamma)$$

for some γ in the range $-\pi < \gamma \leqslant \pi$ and the required result follows. ■

4. Rotating Frames

Suppose we are considering the motion of a rigid body rotating in space, such as a spinning top or space station. To describe this motion we might choose to take the co-ordinates associated with a fixed right-handed orthonormal basis, but more likely it will suit us to consider co-ordinates associated with an right-handed orthonormal basis which is fixed *relative to the rigid body*.

Let's suppose that the fixed-in-body and fixed-in-space axes share a common origin throughout the motion. Then at a time t the moving axes will be a rotation $A(t)$ from the fixed axes. We then have the equation

$$A(t)A(t)^T = I$$

for all $t \in \mathbb{R}$. Differentiating with respect to t we find that

$$\frac{\mathrm{d}A}{\mathrm{d}t}A^T + A\left(\frac{\mathrm{d}A}{\mathrm{d}t}\right)^T = 0$$

or rewriting this

$$\frac{\mathrm{d}A}{\mathrm{d}t}A^T + \left(\frac{\mathrm{d}A}{\mathrm{d}t}A^T\right)^T = 0.$$

Hence $\frac{\mathrm{d}A}{\mathrm{d}t}A^T$ is an anti-symmetric matrix and so we can write

$$\frac{\mathrm{d}A}{\mathrm{d}t}A^T = \begin{pmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix}$$

for some $\alpha, \beta, \gamma \in \mathbb{R}$ (which of course may still depend on t).

Definition 85 *The vector $\boldsymbol{\omega}(t) = (\alpha, \beta, \gamma)^T$ is known as the **angular velocity** of the body at time t.*

Note that

$$\boldsymbol{\omega} \wedge \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \beta z - \gamma y \\ \gamma x - \alpha z \\ \alpha y - \beta x \end{pmatrix} = \begin{pmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Let $\mathbf{r}(t)$ be the position vector at time t of a point fixed in the body – this is the point's position vector relative to fixed-in-space axes. Then $\mathbf{r}(t) = A(t)\mathbf{r}(0)$. Differentiating this we find

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{\mathrm{d}A}{\mathrm{d}t}\mathbf{r}(0) = \left(\frac{\mathrm{d}A}{\mathrm{d}t}A^T\right)(A\mathbf{r}(0)) = \begin{pmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix} \mathbf{r}(t) = \boldsymbol{\omega}(t) \wedge \mathbf{r}(t).$$

Example 86 *In just two dimensions we have*

$$A(t)=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

where $\theta(t)$ is a function on time. Then

$$A'(t)=\begin{pmatrix}-\dot\theta\sin\theta & -\dot\theta\cos\theta\\ \dot\theta\cos\theta & -\dot\theta\sin\theta\end{pmatrix}=\begin{pmatrix}0 & -\dot\theta\\ \dot\theta & 0\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

and

$$\begin{pmatrix}0 & -\dot\theta\\ \dot\theta & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}-\dot\theta y\\ \dot\theta x\end{pmatrix}=\dot\theta\mathbf{k}\wedge\begin{pmatrix}x\\ y\end{pmatrix}$$

so that the angular velocity is $\boldsymbol{\omega}=\dot\theta\mathbf{k}$.

Example 87 *Let*

$$A(t)=\begin{pmatrix}\cos^2 t & \sin^2 t & \sqrt{2}\sin t\cos t\\ \sin^2 t & \cos^2 t & -\sqrt{2}\sin t\cos t\\ -\sqrt{2}\sin t\cos t & \sqrt{2}\sin t\cos t & \cos^2 t-\sin^2 t\end{pmatrix}.$$

Given that $A(t)$ is an orthogonal matrix for all t, find its angular velocity.

Solution Writing $s=\sin t$ and $c=\cos t$ we see

$$\begin{aligned}A'(t) &= \begin{pmatrix}-2sc & 2sc & \sqrt{2}(c^2-s^2)\\ 2sc & -2sc & -\sqrt{2}(c^2-s^2)\\ -\sqrt{2}(c^2-s^2) & \sqrt{2}(c^2-s^2) & -4sc\end{pmatrix}\\ &= \begin{pmatrix}0 & 0 & \sqrt{2}\\ 0 & 0 & -\sqrt{2}\\ -\sqrt{2} & \sqrt{2} & 0\end{pmatrix}\begin{pmatrix}c^2 & s^2 & \sqrt{2}sc\\ s^2 & c^2 & -\sqrt{2}sc\\ -\sqrt{2}sc & \sqrt{2}sc & c^2-s^2\end{pmatrix}\end{aligned}$$

and hence

$$\boldsymbol{\omega}=\begin{pmatrix}\sqrt{2}\\ \sqrt{2}\\ 0\end{pmatrix}.$$

Note that $\boldsymbol{\omega}$ has magnitude 2. ■

Note that we can better understand the matrix $A(t)$ in Example 87 by choosing the (fixed-in-space) orthonormal basis

$$\mathbf{e}_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\\ 0\end{pmatrix},\qquad \mathbf{e}_2=\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix},\qquad \mathbf{e}_3=\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}.$$

Then

$$\begin{aligned} A(t)\mathbf{e}_1 &= \frac{1}{\sqrt{2}}\begin{pmatrix} c^2+s^2 \\ s^2+c^2 \\ 0 \end{pmatrix} = \mathbf{e}_1; \\ A(t)\mathbf{e}_2 &= \frac{1}{\sqrt{2}}\begin{pmatrix} s^2-c^2 \\ c^2-s^2 \\ 2\sqrt{2}sc \end{pmatrix} = (\cos 2t)\,\mathbf{e}_2 + (\sin 2t)\,\mathbf{e}_3. \\ A(t)\mathbf{e}_3 &= \begin{pmatrix} \sqrt{2}sc \\ -\sqrt{2}sc \\ c^2-s^2 \end{pmatrix} = (-\sin 2t)\,\mathbf{e}_2 + (\cos 2t)\,\mathbf{e}_3. \end{aligned}$$

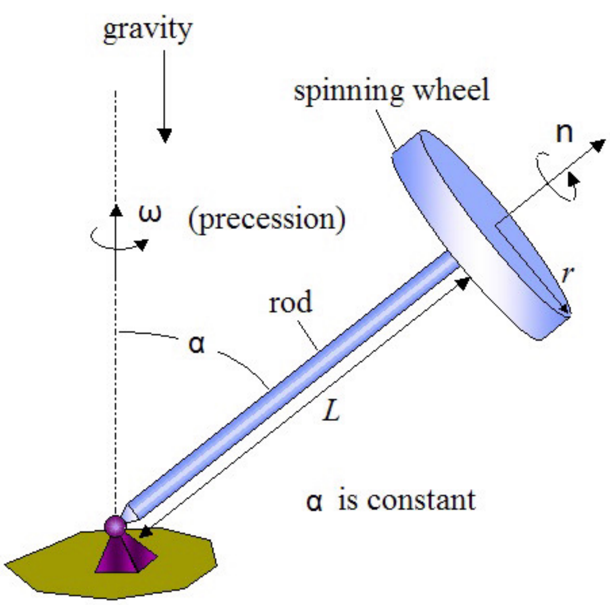
So with respect to the co-ordinates associated with the $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ basis we see

$$A(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos 2t & -\sin 2t \\ 0 & \sin 2t & \cos 2t \end{pmatrix}.$$

Example 88 ***(Precession)*** *When a rigid body (such as a gyroscope) precesses the fixed-in-space and fixed-in-body axes are related by*

$$A(t) = \begin{pmatrix} \sin\alpha\cos\omega t & \cos\alpha\cos nt\cos\omega t - \sin nt\sin\omega t & -\cos\alpha\sin nt\cos\omega t - \cos nt\sin\omega t \\ \sin\alpha\sin\omega t & \cos\alpha\cos nt\sin\omega t + \sin nt\cos\omega t & -\cos\alpha\sin nt\sin\omega t + \cos nt\cos\omega t \\ \cos\alpha & -\sin\alpha\cos nt & \sin\alpha\sin nt \end{pmatrix}$$

where α and ω are constants.



23. Precession

It is a tiresome mess to differentiate $A(t)$ to find the angular velocity, but when we do we find the angular velocity equals

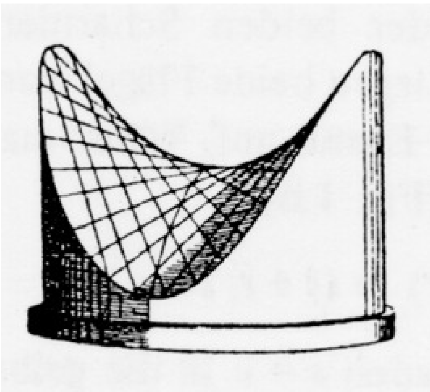
$$\begin{pmatrix} n \sin\alpha \cos\omega t \\ n \sin\alpha \sin\omega t \\ \omega + n\cos\alpha \end{pmatrix} = \omega \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + n \begin{pmatrix} \sin\alpha \cos\omega t \\ \sin\alpha \sin\omega t \\ \cos\alpha \end{pmatrix}.$$

In particular the angular velocity is not constant. Rather it is made up of two angular velocities, one around the fixed z-axis and one around an axis in the body.

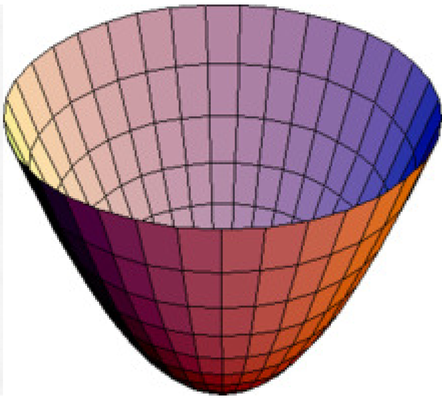
5. Parametrized Surfaces

We probably all feel we know a smooth surface in $\mathbb{R}^3$ when we see one, and this instinct for what a surface is will largely be satisfactory for the purposes of this course. Hopefully it is not surprising that the examples below are all examples of surfaces in $\mathbb{R}^3$.

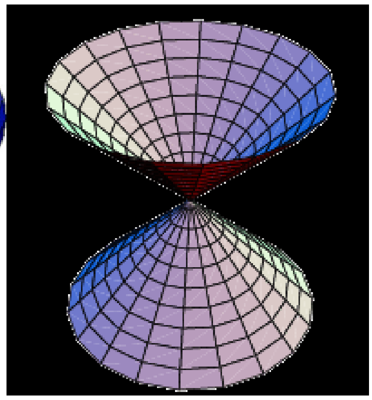
- **Sphere:** $x^2 + y^2 + z^2 = a^2$;
- **Ellipsoid:** $x^2/a^2 + y^2/b^2 + z^2/c^2$;
- **Hyperboloid of One Sheet:** $x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$;
- **Hyperboloid of Two Sheets:** $x^2/a^2 - y^2/b^2 - z^2/c^2 = 1$;
- **Paraboloid:** $z = x^2 + y^2$;
- **Hyperbolic Paraboloid:** $z = x^2 - y^2$;
- **Cone:** $x^2 + y^2 = z^2$ with $z \geqslant 0$.



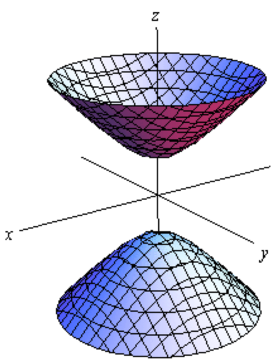
24a. Hyperbolic Paraboloid



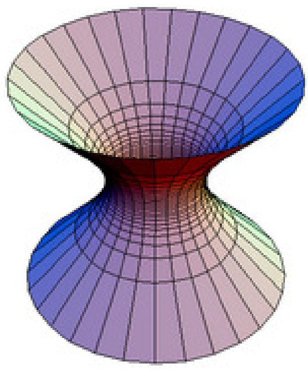
24b. Paraboloid



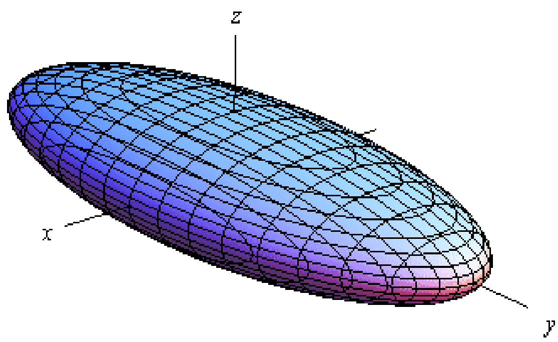
24c. Double Cone



24d. 2-sheet Hyperboloid



24e. 1-sheet Hyperboloid



24f. Ellipsoid

Example 89 *Show that through any point of the hyperbolic paraboloid (see Figure 24a) with equation $z = xy$ pass two lines which are entirely in the surface.*

Solution Consider a general point (a, b, ab) on the surface. A line through this point can be parametrized as

$$\mathbf{r}(\lambda) = (a, b, ab) + \lambda(u, v, w)$$

where (u, v, w) is a unit vector. If this line is to lie entirely inside the hyperbolic paraboloid then we need

$$ab + \lambda w = (a + \lambda u)(b + \lambda v)$$

for all λ. Hence we have

$$w = bu + av, \qquad uv = 0, \qquad u^2 + v^2 + w^2 = 1.$$

So $u = 0$ or $v = 0$. If $u = 0$ then we find

$$(u, v, w) = \pm\frac{(0, 1, a)}{\sqrt{1 + a^2}}$$

and if $v = 0$ we see

$$(u, v, w) = \pm\frac{(1, 0, b)}{\sqrt{1 + b^2}}.$$

So the two lines are

$$\mathbf{r}_1(\lambda) = (a, b + \lambda, ab + \lambda a), \qquad \mathbf{r}_2(\lambda) = (a + \lambda, b, ab + \lambda b).$$

■

Definition 90 *Recall that, given a parametrized curve* $\mathbf{r}\colon [a, b] \to \mathbb{R}^3$ *its* ***arc length*** *is defined to be*

$$\int_a^b |\mathbf{r}'(t)|\ \mathrm{d}t.$$

We might instead ***parametrize the curve by arc length*** *by assigning parameter* s *to the point of the curve which is arc length* s *away from* $\mathbf{r}(a)$.

Example 91 *A* ***cycloid*** *is parametrized as*

$$\mathbf{r}(t) = (t - \sin t, 1 - \cos t) \qquad 0 \leqslant t \leqslant 2\pi.$$

(i) Find the arc length of the cycloid.
(ii) Find the volume bounded by the surface of revolution formed when the cycloid is rotated about the x*-axis.*

Solution (i) Note $\mathbf{r}'(t) = (1 - \cos t, \sin t)$. The arc length is given by

$$\begin{aligned} \mathcal{L} &= \int_0^{2\pi} |\mathbf{r}'(t)|\ \mathrm{d}t \\ &= \int_0^{2\pi} \sqrt{(1 - \cos t)^2 + \sin^2 t}\,\mathrm{d}t \\ &= \int_0^{2\pi} \sqrt{2 - 2\cos t}\,\mathrm{d}t \\ &= \int_0^{2\pi} 2\sin(t/2)\,\mathrm{d}t = 8. \end{aligned}$$

(ii) The volume bounded by the surface of revolution is

$$\begin{aligned} V &= \pi \int_{x=0}^{2\pi} y^2 \, \mathrm{d}x = \pi \int_{t=0}^{2\pi} y(t)^2 \frac{\mathrm{d}x}{\mathrm{d}t} \, \mathrm{d}t \\ &= \pi \int_{t=0}^{2\pi} (1-\cos t)^2 (1-\cos t) \, \mathrm{d}t \\ &= \pi \int_{t=0}^{2\pi} \left(1 - 3\cos t + 3\cos^2 t - \cos^3 t\right) \, \mathrm{d}t \\ &= \pi \int_{t=0}^{2\pi} \left(1 + 3\cos^2 t\right) \, \mathrm{d}t \\ &= \pi\,(2\pi + 3\pi) = 5\pi^2. \end{aligned}$$

■

Example 92 *Find the tangent plane to ellipsoid (Figure 24f)*

$$\frac{x^2}{3} + \frac{y^2}{2} + \frac{z^2}{6} = 1$$

at the point $(1,1,1)$.

Solution The gradient vector ∇f is normal to each level set $f = \text{const.}$. So with

$$f(x,y,z) = \frac{x^2}{3} + \frac{y^2}{2} + \frac{z^2}{6}, \qquad \nabla f = \left(\frac{2x}{3}, y, \frac{z}{3}\right),$$

which equals $(2/3, 1, 1/3)$ at $(1,1,1)$. Hence the tangent plane is given by

$$\frac{2}{3}x + y + \frac{z}{3} = \frac{2}{3} + 1 + \frac{1}{3} = 2$$

or tidying up

$$2x + 3y + z = 6.$$

■

Example 93 *Let $\gamma(s)$ be a curve, parametrized by arc length, on a surface and let $\mathbf{n}$ denote a choice of unit normal to the surface,. Given that the curve has shortest length when*

$$\gamma''(s) \wedge \mathbf{n} = \mathbf{0},$$

(we shall not prove this fact) show that the great circles are the curves of shortest length on a sphere.

Solution Without loss of generality consider the unit sphere centred at the origin and take

$$\gamma(s) = (\cos s, \sin s, 0).$$

Note $\mathbf{n} = \mathbf{r}$ for each point of the sphere. We also have $\gamma''(s) = (-\cos s, -\sin s, 0) = -\gamma(s)$. Hence

$$\gamma''(s) \wedge \mathbf{n}(\gamma(s)) = -\gamma(s) \wedge \gamma(s) = 0.$$

■

For this course we will be happy to work with parametrized surfaces without a more general formal definition of what constitutes a surface.

Definition 94 *A **smooth parametrized surface** is a map* $\mathbf{r}$*, known as the **parametrization***

$$\mathbf{r} : U \to \mathbb{R}^3 \quad \text{given by} \quad \mathbf{r}(u,v) = (x(u,v), y(u,v), z(u,v))$$

*from an (open) **subset** $U \subseteq \mathbb{R}^2$ to $\mathbb{R}^3$ such that:*

- *x, y, z have continuous partial derivatives with respect to u and v of all orders;*
- *$\mathbf{r}$ is a bijection from U to $\mathbf{r}(U)$ with both $\mathbf{r}$ and $\mathbf{r}^{-1}$ being continuous;*
- *(**smoothness condition**) at each point the vectors*

$$\mathbf{r}_u = \frac{\partial \mathbf{r}}{\partial u} \quad \text{and} \quad \mathbf{r}_v = \frac{\partial \mathbf{r}}{\partial v}$$

  are linearly independent (i.e. are not scalar multiples of one another).

Remark 95 *Note that through any point $\mathbf{r}(u_0, v_0)$ run two co-ordinate curves $u = u_0$ and $v = v_0$. As the vector $\mathbf{r}_u$ is calculated by holding v constant, it is a tangent vector in the direction of the $v = v_0$ co-ordinate curve. Likewise $\mathbf{r}_v$ is in the direction of the $u = u_0$ co-ordinate curve.*

We will not be looking to treat the above definition with any generality. Rather we shall just look to parametrize some standard and familiar surfaces and calculate some tangents and normals.

Example 96 *The cone $x^2 + y^2 = z^2$ with $z \geqslant 0$ can be parametrized as*

$$\mathbf{r}(x,y) = \left(x, y, \sqrt{x^2+y^2}\right), \qquad x, y \in \mathbb{R}.$$

For $(x,y) \neq (0,0)$ we then have that

$$\mathbf{r}_x = \left(1, 0, \frac{x}{\sqrt{x^2+y^2}}\right), \qquad \mathbf{r}_y = \left(0, 1, \frac{y}{\sqrt{x^2+y^2}}\right),$$

*but neither vector $\mathbf{r}_x$ nor $\mathbf{r}_y$ is defined at $(0,0,0)$. The point $(0,0,0)$ is said to be a **singular** point of the cone.*

Two co-ordinate systems which lend themselves to many surfaces are:

Definition 97 *(**Cylindrical Polar Co-ordinates**) These are given by*

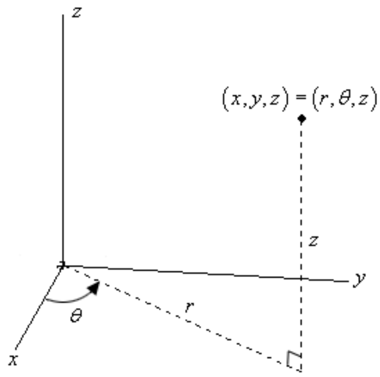
$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z,$$

with $r > 0, -\pi < \theta < \pi$, $z \in \mathbb{R}$.

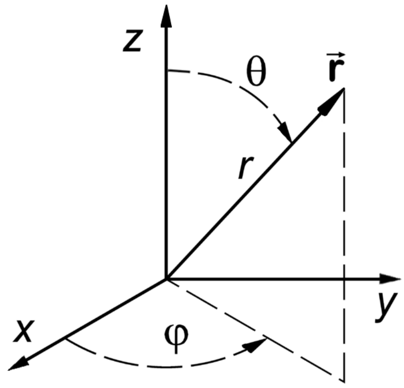
Definition 98 *(**Spherical Polar Co-ordinates**) These are given by*

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta,$$

with $r > 0, 0 < \theta < \pi$, $-\pi < \phi < \pi$.



25. Cylindrical Polars



26. Spherical Polars

Example 99 *A second – perhaps more natural – parametrization for the cone uses cylindrical polar co-ordinates. We could define*

$$\mathbf{r}(\theta, z) = (z\cos\theta, z\sin\theta, z) \qquad 0 < \theta < 2\pi,\ z > 0.$$

Note that the parametrization misses one meridian of the cone.

Definition 100 *Let $\mathbf{r}\colon U \to \mathbb{R}^3$ be a smooth parametrized surface and let $\mathbf{p}$ be a point on the surface. The plane containing $\mathbf{p}$ and which is parallel to the vectors*

$$\mathbf{r}_u(\mathbf{p}) = \frac{\partial \mathbf{r}}{\partial u}(\mathbf{p}) \quad \text{and} \quad \mathbf{r}_v(\mathbf{p}) = \frac{\partial \mathbf{r}}{\partial v}(\mathbf{p})$$

is called the ***tangent plane*** *to $\mathbf{r}(U)$ at $\mathbf{p}$. Because $\mathbf{r}_u$ and $\mathbf{r}_v$ are independent the tangent plane is well-defined.*

Definition 101 *Any vector in the direction*

$$\frac{\partial \mathbf{r}}{\partial u}(\mathbf{p}) \wedge \frac{\partial \mathbf{r}}{\partial v}(\mathbf{p})$$

is said to be ***normal*** *to the surface at $\mathbf{p}$. Thus there are two* ***unit normals*** *of length one.*

Example 102 *Spherical polar co-ordinates unsurprisingly give a natural parametrization for the sphere $x^2 + y^2 + z^2 = a^2$ with*

$$\mathbf{r}(\phi, \theta) = (a\sin\theta\cos\phi, a\sin\theta\sin\phi, a\cos\theta), \qquad -\pi < \phi < \pi,\ 0 < \theta < \pi.$$

We already know the outward-pointing unit normal at $\mathbf{r}(\theta, \phi)$ is $\mathbf{r}(\theta, \phi)/a$ but let's verify this with the previous definitions and find the tangent plane. We have

$$\begin{aligned} \frac{\partial \mathbf{r}}{\partial \phi} &= (-a\sin\theta\sin\phi, a\sin\theta\cos\phi, 0); \\ \frac{\partial \mathbf{r}}{\partial \theta} &= (a\cos\theta\cos\phi, a\cos\theta\sin\phi, -a\sin\theta). \end{aligned}$$

Hence

$$\begin{aligned}
\frac{\partial \mathbf{r}}{\partial \phi} \wedge \frac{\partial \mathbf{r}}{\partial \theta} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -a\sin\theta\sin\phi & a\sin\theta\cos\phi & 0 \\ a\cos\theta\cos\phi & a\cos\theta\sin\phi & -a\sin\theta \end{vmatrix} \\
&= a^2\sin\theta \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sin\phi & \cos\phi & 0 \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \end{vmatrix} \\
&= a^2\sin\theta \begin{pmatrix} -\sin\theta\cos\phi \\ -\sin\theta\sin\phi \\ -\cos\theta \end{pmatrix},
\end{aligned}$$

and so the two unit normals are $\pm(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. *The tangent plane at* $\mathbf{r}(\phi,\theta)$ *is then*

$$\mathbf{r}\cdot(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) = a.$$

Example 103 *Find the normal and tangent plane to the point* (X,Y,Z) *on the hyperbolic paraboloid (Figure 24a) with equation* $z = x^2 - y^2$.

Solution This surface has a simple choice of parametrization as there is exactly one point lying above, or below, the point $(x,y,0)$. So we can take a parametrization

$$\mathbf{r}(x,y) = \left(x, y, x^2 - y^2\right), \qquad x, y \in \mathbb{R}.$$

We then have

$$\frac{\partial \mathbf{r}}{\partial x} = (1,0,2x), \qquad \frac{\partial \mathbf{r}}{\partial y} = (0,1,-2y).$$

So

$$\frac{\partial \mathbf{r}}{\partial x} \wedge \frac{\partial \mathbf{r}}{\partial y} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 2x \\ 0 & 1 & -2y \end{vmatrix} = \begin{pmatrix} -2x \\ 2y \\ 1 \end{pmatrix}.$$

A normal vector to the surface at (X,Y,Z) is then $(-2X, 2Y, 1)$ and we see that the equation of the tangent plane is

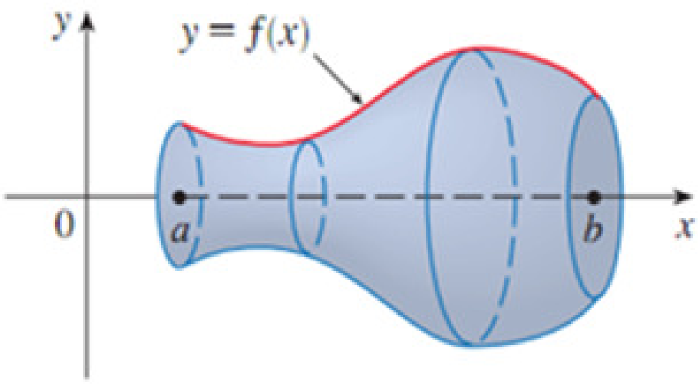
$$\begin{aligned}
\mathbf{r}\cdot(-2X,2Y,1) &= \left(X, Y, X^2 - Y^2\right)\cdot(-2X,2Y,1) \\
&= -2X^2 + 2Y^2 + X^2 - Y^2 \\
&= Y^2 - X^2 = -Z
\end{aligned}$$

or equivalently

$$2Xx - 2Yy + z = Z.$$

■

Example 104 ***(Surfaces of Revolution)***



27. Surface of Revolution

We can form a surface of revolution by rotating the graph $y = f(x)$, where $f(x) > 0$, about the x-axis. There is then a fairly natural parametrization for the surface of revolution with cylindrical polar co-ordinates:

$$\mathbf{r}(x,\theta) = (x, f(x)\cos\theta, f(x)\sin\theta) \qquad -\pi < \theta < \pi,\ a < x < b.$$

We can calculate the normals to the surface by determining

$$\mathbf{r}_x = (1, f'(x)\cos\theta, f'(x)\sin\theta)\,, \qquad \mathbf{r}_\theta = (0, -f(x)\sin\theta, f(x)\cos\theta)\,,$$

and so

$$\mathbf{r}_x \wedge \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & f'(x)\cos\theta & f'(x)\sin\theta \\ 0 & -f(x)\sin\theta & f(x)\cos\theta \end{vmatrix} = \begin{pmatrix} f'(x)f(x) \\ -f(x)\cos\theta \\ -f(x)\sin\theta \end{pmatrix}.$$

The outward-pointing unit normal is then

$$\mathbf{n}(x,\theta) = \frac{(-f'(x), \cos\theta, \sin\theta)}{\sqrt{1+f'(x)^2}}.$$

Note that the surface can also be described as the level set

$$y^2 + z^2 - f(x)^2 = 0$$

and so the gradient vector at $\mathbf{r}(x,\theta)$ will also be normal to the surface – this equals

$$(-2f(x)f'(x), 2y, 2z) = 2f(x)\,(-f'(x), \cos\theta, \sin\theta)\,.$$

Example 105 *Show that a meridian on a surface of revolution is a curve of shortest length.*

Solution Recall that a curve $\gamma(s)$, parametrize by arc length, will be one of shortest length if $\gamma'' \wedge \mathbf{n} = \mathbf{0}$ (though we have not proved this fact). If our generating curve is parametrized by arc length s then we can write it as

$$\gamma(s) = (f(s), g(s))$$

in the xy-plane, and if s is arc length then $|\gamma'(s)| = 1$ so that

$$f'(s)^2 + g'(s)^2 = 1.$$

So we can parametrize the surface of revolution as

$$\mathbf{r}(s, \theta) = (f(s), g(s)\cos\theta, g(s)\sin\theta)$$

and a meridian would be of the form $\gamma(s) = \mathbf{r}(s, \alpha)$ where α is a constant.

We then have

$$\mathbf{r}_s = (f', g'\cos\theta, g'\sin\theta), \qquad \mathbf{r}_\theta = (0, -g\sin\theta, g\cos\theta)$$

and the normal is parallel to

$$\mathbf{n} = \mathbf{r}_s \wedge \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ f' & g'\cos\theta & g'\sin\theta \\ 0 & -g\sin\theta & g\cos\theta \end{vmatrix} = \begin{pmatrix} gg' \\ -gf'\cos\theta \\ -gf'\sin\theta \end{pmatrix}.$$

Finally the acceleration vector γ'' equals

$$\gamma''(s) = (f'', g''\cos\alpha, g''\sin\alpha).$$

So at the point $\gamma(s) = \mathbf{r}(s, \alpha)$ we see that

$$\gamma''(s) \wedge \mathbf{n} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ f'' & g''\cos\alpha & g''\sin\alpha \\ gg' & -gf'\cos\alpha & -gf'\sin\alpha \end{vmatrix} = \begin{pmatrix} 0 \\ g(g'g'' + f'f'')\sin\alpha \\ g(g'g'' + f'f'')\cos\alpha \end{pmatrix}.$$

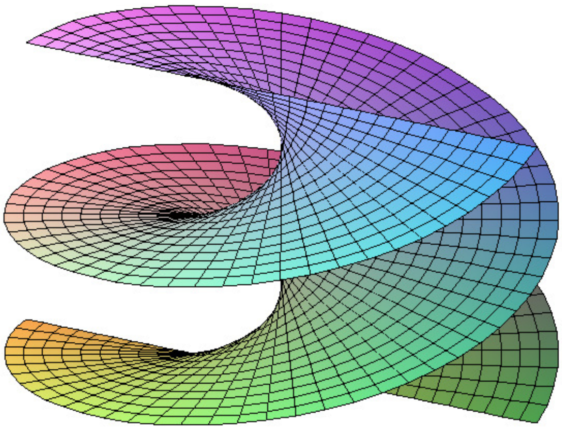
At first glance this does not look to be zero. But recalling $(f')^2 + (g')^2 = 1$ we can differentiate this to get

$$2f'f'' + 2g'g'' = 0$$

and so it is indeed the case that $\gamma''(s) \wedge \mathbf{n} = \mathbf{0}$ on a meridian. ■

Example 106 *The* ***catenoid*** *is the surface of revolution formed by rotating the curve* $y = \cosh x$*, known as a catenary, about the* x*-axis. So we can parametrize it as*

$$\mathbf{r}(x, \theta) = (x, \cosh x\cos\theta, \cosh x\sin\theta), \qquad -\pi < \theta < \pi,\ x \in \mathbb{R}.$$



28. Helicoid

*The **helicoid** is formed in a "propeller-like" fashion by pushing the x-axis up the z-axis while spinning the x-axis at a constant angular velocity. So we can parametrize it as*

$$\mathbf{s}(X,Z) = (X\cos Z, X\sin Z, Z).$$

[The need for differing notation – re x and X – will become apparent in due course.]

As some preliminary calculations we note that

$$\begin{aligned} \mathbf{r}_x &= (1, \sinh x\cos\theta, \sinh x\sin\theta), \qquad \mathbf{r}_\theta = (0, -\cosh x\sin\theta, \cosh x\cos\theta), \\ \mathbf{r}_x\cdot\mathbf{r}_x &= \cosh^2 x, \qquad \mathbf{r}_x\cdot\mathbf{r}_\theta = 0, \qquad \mathbf{r}_\theta\cdot\mathbf{r}_\theta = \cosh^2 x. \end{aligned}$$

And for the helicoid we have

$$\begin{aligned} \mathbf{s}_X &= (\cos Z, \sin Z, 0), \qquad \mathbf{s}_Z = (-X\sin Z, X\cos Z, 1) \\ \mathbf{s}_X\cdot\mathbf{s}_X &= 1, \qquad s_X\cdot s_Z = 0, \qquad \mathbf{s}_Z\cdot\mathbf{s}_Z = 1+X^2. \end{aligned}$$

If we consider the curve

$$\gamma(t) = \mathbf{r}(x(t),\theta(t)) \qquad a\leqslant t\leqslant b$$

in the catenoid then $\gamma' = x'\mathbf{r}_x + \theta'\mathbf{r}_\theta$ and so

$$|\gamma'|^2 = (\mathbf{r}_x\cdot\mathbf{r}_x)(x')^2 + 2(\mathbf{r}_x\cdot\mathbf{r}_\theta)x'\theta' + (\mathbf{r}_\theta\cdot\mathbf{r}_\theta)(\theta')^2 = \cosh^2 x((x')^2+(\theta')^2).$$

For the curve

$$\Gamma(t) = \mathbf{s}(X(t),Z(t)) \qquad c\leqslant t\leqslant d$$

in the helicoid we similarly have

$$|\Gamma'|^2 = (X')^2 + (1+X^2)(Z')^2.$$

Now consider the map from the catenoid to the helicoid given by

$$\mathbf{r}(x,\theta) \mapsto \mathbf{s}(\sinh x,\theta).$$

A curve $\mathbf{r}(x(t), \theta(t))$ *where* $a \leqslant t \leqslant b$ *has length*

$$\int_a^b |\gamma'| \,\mathrm{d}t = \int_a^b \sqrt{\cosh^2 x((x')^2 + (\theta')^2)} \,\mathrm{d}t = \int_a^b \cosh x \sqrt{(x')^2 + (\theta')^2} \,\mathrm{d}t.$$

For the image of the curve $\mathbf{s}(\sinh x, z)$ *in the helicoid we have*

$$\int_a^b |\Gamma'| \,\mathrm{d}t = \int_a^b \sqrt{(X')^2 + (1 + X^2)(Z')^2)} \,\mathrm{d}t.$$

Now $X = \sinh x$ *and* $Z = \theta$ *so that the above equals*

$$\begin{aligned} & \int_a^b \sqrt{\cosh^2 x(x')^2 + (1 + \sinh^2 x)(\theta')^2} \,\mathrm{d}t \\ = & \int_a^b \cosh x \sqrt{(x')^2 + (\theta')^2} \,\mathrm{d}t \\ = & \int_a^b |\gamma'| \,\mathrm{d}t. \end{aligned}$$

So for any curve γ *in the catenoid, its image, under the map to the helicoid, has the same length. This means that the map is an isometry between the surfaces (when distances are measured within the surface).*

6. Surface Area

Let $\mathbf{r} : U \to \mathbb{R}^3$ be a smooth parametrized surface with

$$\mathbf{r}(u,v) = (x(u,v), y(u,v), z(u,v))$$

and consider the small rectangle of the plane that is bounded by the co-ordinate lines $u = u_0$ and $u = u_0 + \delta u$ and $v = v_0$ and $v = v_0 + \delta v$. Then $\mathbf{r}$ maps this to a small region of the surface $\mathbf{r}(U)$ and we are interested in calculating the surface area of this small region, which is approximately that of a parallelogram. Note

$$\begin{aligned} \mathbf{r}(u+\delta u, v) - \mathbf{r}(u,v) &\approx \frac{\partial \mathbf{r}}{\partial u}(u,v)\, \delta u, \\ \mathbf{r}(u, v+\delta v) - \mathbf{r}(u,v) &\approx \frac{\partial \mathbf{r}}{\partial v}(u,v)\, \delta v. \end{aligned}$$

Recall that the area of a parallelogram with sides $\mathbf{a}$ and $\mathbf{b}$ is $|\mathbf{a} \wedge \mathbf{b}|$. So the element of surface area we are considering is approximately

$$\left| \frac{\partial \mathbf{r}}{\partial u} \delta u \wedge \frac{\partial \mathbf{r}}{\partial v} \delta v \right| = \left| \frac{\partial \mathbf{r}}{\partial u} \wedge \frac{\partial \mathbf{r}}{\partial v} \right| \delta u\, \delta v.$$

This motivates the following definitions.

Definition 107 *Let $\mathbf{r} : U \to \mathbb{R}^3$ be a smooth parametrized surface. Then the **surface area** (or simply **area**) of $\mathbf{r}(U)$ is defined to be*

$$\iint_U \left| \frac{\partial \mathbf{r}}{\partial u} \wedge \frac{\partial \mathbf{r}}{\partial v} \right| \mathrm{d}u\, \mathrm{d}v.$$

Definition 108 *We will often write*

$$\mathrm{d}S = \left| \frac{\partial \mathbf{r}}{\partial u} \wedge \frac{\partial \mathbf{r}}{\partial v} \right| \mathrm{d}u\, \mathrm{d}v$$

to denote an infinitesimal part of surface area.

Proposition 109 *The surface area of $\mathbf{r}(U)$ is independent of the choice of parametrization.*

Proof Let $\Sigma = \mathbf{r}(U) = \mathbf{s}(W)$ be two different parametrizations of a surface X; take u, v as the co-ordinates on U and p, q as the co-ordinates on W. Let $f = (f_1, f_2) : U \to W$ be the co-ordinate change map – i.e. for any $(u,v) \in U$ we have

$$\mathbf{r}(u,v) = \mathbf{s}(f(u,v)) = \mathbf{s}(f_1(u,v), f_2(u,v)).$$

Then

$$\frac{\partial \mathbf{r}}{\partial u} = \frac{\partial \mathbf{s}}{\partial p}\frac{\partial f_1}{\partial u} + \frac{\partial \mathbf{s}}{\partial q}\frac{\partial f_2}{\partial u}, \qquad \frac{\partial \mathbf{r}}{\partial v} = \frac{\partial \mathbf{s}}{\partial p}\frac{\partial f_1}{\partial v} + \frac{\partial \mathbf{s}}{\partial q}\frac{\partial f_2}{\partial v}.$$

Hence

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial u} \wedge \frac{\partial \mathbf{r}}{\partial v} &= \frac{\partial \mathbf{s}}{\partial p}\frac{\partial f_1}{\partial u} \wedge \frac{\partial \mathbf{s}}{\partial q}\frac{\partial f_2}{\partial v} + \frac{\partial \mathbf{s}}{\partial q}\frac{\partial f_2}{\partial u} \wedge \frac{\partial \mathbf{s}}{\partial p}\frac{\partial f_1}{\partial v} \\ &= \left(\frac{\partial f_1}{\partial u}\frac{\partial f_2}{\partial v} - \frac{\partial f_1}{\partial v}\frac{\partial f_2}{\partial u}\right)\frac{\partial \mathbf{s}}{\partial p} \wedge \frac{\partial \mathbf{s}}{\partial q} \\ &= \frac{\partial (p,q)}{\partial (u,v)}\frac{\partial \mathbf{s}}{\partial p} \wedge \frac{\partial \mathbf{s}}{\partial q}.\end{aligned}$$

Finally

$$\begin{aligned}\iint_U \left|\frac{\partial \mathbf{r}}{\partial u} \wedge \frac{\partial \mathbf{r}}{\partial v}\right| \mathrm{d}u\,\mathrm{d}v &= \iint_U \left|\frac{\partial (p,q)}{\partial (u,v)}\frac{\partial \mathbf{s}}{\partial p} \wedge \frac{\partial \mathbf{s}}{\partial q}\right| \mathrm{d}u\,\mathrm{d}v \\ &= \iint_U \left|\frac{\partial \mathbf{s}}{\partial p} \wedge \frac{\partial \mathbf{s}}{\partial q}\right| \left|\frac{\partial (p,q)}{\partial (u,v)}\right| \mathrm{d}u\,\mathrm{d}v \\ &= \iint_W \left|\frac{\partial \mathbf{s}}{\partial p} \wedge \frac{\partial \mathbf{s}}{\partial q}\right| \mathrm{d}p\,\mathrm{d}q\end{aligned}$$

by the two-dimensional substitution rule (Theorem 130 in the Institute notes *Mods Calculus*.)
■

Example 110 *Find the surface area of the cone*

$$x^2 + y^2 = z^2 \cot^2 \alpha \qquad 0 \leqslant z \leqslant h.$$

Solution We can parametrize the cone as

$$\mathbf{r}(z,\theta) = (z\cot\alpha\cos\theta, z\cot\alpha\sin\theta, z), \qquad 0 < \theta < 2\pi,\ 0 < z < h.$$

We have

$$\mathbf{r}_z = (\cot\alpha\cos\theta, \cot\alpha\sin\theta, 1), \qquad \mathbf{r}_\theta = (-z\cot\alpha\sin\theta, z\cot\alpha\cos\theta, 0).$$

So

$$\mathbf{r}_z \wedge \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cot\alpha\cos\theta & \cot\alpha\sin\theta & 1 \\ -z\cot\alpha\sin\theta & z\cot\alpha\cos\theta & 0 \end{vmatrix} = \begin{pmatrix} -z\cot\alpha\cos\theta \\ -z\cot\alpha\sin\theta \\ z\cot^2\alpha \end{pmatrix}.$$

Thus the cone has surface area

$$\begin{aligned}&\int_{\theta=0}^{2\pi}\int_{z=0}^{h} \sqrt{z^2\cot^2\alpha\cos^2\theta + z^2\cot^2\alpha\sin^2\theta + z^2\cot^4\alpha}\,\mathrm{d}z\,\mathrm{d}\theta \\ =& \int_{\theta=0}^{2\pi}\int_{z=0}^{h} z\cot\alpha\sqrt{1+\cot^2\alpha}\,\mathrm{d}z\,\mathrm{d}\theta \\ =& \ 2\pi\int_{z=0}^{h} z\cot\alpha\csc\alpha\,\mathrm{d}z \\ =& \ 2\pi \times \frac{\cos\alpha}{\sin^2\alpha} \times \left[\frac{z^2}{2}\right]_0^h \\ =& \ \frac{\pi h^2\cos\alpha}{\sin^2\alpha}.\end{aligned}$$

Note that as $\alpha \to 0$ this area tends to infinity as the cone transforms into the plane and the area tends to zero as $\alpha \to \pi/2$. ■

Example 111 *Calculate the area of a sphere of radius a using spherical polar co-ordinates.*

Solution We can parametrize the sphere by

$$\mathbf{r}\left(\theta,\phi\right) = \left(a\sin\theta\cos\phi, a\sin\theta\sin\phi, a\cos\theta\right) \qquad 0<\theta<\pi,\ 0<\phi<2\pi.$$

Then

$$\begin{aligned}\frac{\partial \mathbf{r}}{\partial\theta}\wedge\frac{\partial \mathbf{r}}{\partial\phi} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial x}{\partial\theta} & \frac{\partial y}{\partial\theta} & \frac{\partial z}{\partial\theta} \\ \frac{\partial x}{\partial\phi} & \frac{\partial y}{\partial\phi} & \frac{\partial z}{\partial\phi}\end{vmatrix} \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a\cos\theta\cos\phi & a\cos\theta\sin\phi & -a\sin\theta \\ -a\sin\theta\sin\phi & a\sin\theta\cos\phi & 0\end{vmatrix} \\ &= a^2\left(\sin^2\theta\cos\phi, \sin^2\theta\sin\phi, \sin\theta\cos\theta\right) \\ &= a^2\sin\theta\left(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\right).\end{aligned}$$

Hence

$$\mathrm{d}S = \left|a^2\sin\theta\left(-\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\right)\right|\,\mathrm{d}\theta\,\mathrm{d}\phi = a^2\left|\sin\theta\right|\,\mathrm{d}\theta\,\mathrm{d}\phi.$$

Finally

$$\begin{aligned} A &= \int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} a^2\left|\sin\theta\right|\,\mathrm{d}\phi\,\mathrm{d}\theta \\ &= 2\pi a^2\int_{\theta=0}^{\pi}\left|\sin\theta\right|\,\mathrm{d}\theta \\ &= 4\pi a^2.\end{aligned}$$

■

Example 112 *Let $0<a<b$. Find the area of the torus obtained by revolving the circle $(x-b)^2+z^2=a^2$ in the xz-plane about the z-axis.*

Solution We can parametrize the torus as

$$\mathbf{r}\left(\theta,\phi\right) = \left(\left(b+a\sin\theta\right)\cos\phi,\ \left(b+a\sin\theta\right)\sin\phi, a\cos\theta\right) \qquad 0<\theta,\phi<2\pi.$$

We have

$$\mathbf{r}_\theta = \left(a\cos\theta\cos\phi, a\cos\theta\sin\phi, -a\sin\theta\right), \qquad \mathbf{r}_\phi = \left(-\left(b+a\sin\theta\right)\sin\phi, \left(b+a\sin\theta\right)\cos\phi, 0\right)$$

and

$$\begin{aligned}
\mathbf{r}_\theta \wedge \mathbf{r}_\phi &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a\cos\theta\cos\phi & a\cos\theta\sin\phi & -a\sin\theta \\ -(b+a\sin\theta)\sin\phi & (b+a\sin\theta)\cos\phi & 0 \end{vmatrix} \\
&= a(b+a\sin\theta)\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{vmatrix} \\
&= a(b+a\sin\theta)(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).
\end{aligned}$$

The surface area of the torus is

$$\begin{aligned}
& \int_{\theta=0}^{2\pi}\int_{\phi=0}^{2\pi} a(b+a\sin\theta)\sqrt{\sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta}\,\mathrm{d}\phi\,\mathrm{d}\theta \\
= & \ 2\pi a\int_{\theta=0}^{2\pi}(b+a\sin\theta)\,\mathrm{d}\theta \\
= & \ 4\pi^2 ab.
\end{aligned}$$

■

Proposition 113 ***(Surface Area of a Graph)*** *Let $z = f(x,y)$ denote the graph of a function f defined on a subset S of the xy-plane. Show that the graph has surface area*

$$\iint_S \sqrt{1 + (f_x)^2 + (f_y)^2}\,\mathrm{d}x\,\mathrm{d}y.$$

Proof We can parametrize the surface as

$$\mathbf{r}(x,y) = (x, y, f(x,y)) \qquad (x,y) \in S.$$

Then

$$\mathbf{r}_x \wedge \mathbf{r}_y = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} = (-f_x, -f_y, 1).$$

Hence the graph has surface area

$$\iint_S |\mathbf{r}_x \wedge \mathbf{r}_y|\,\mathrm{d}x\,\mathrm{d}y = \iint_S \sqrt{1 + (f_x)^2 + (f_y)^2}\,\mathrm{d}x\,\mathrm{d}y.$$

■

Example 114 *Use Proposition 113 to show that a sphere of radius a has surface area $4\pi a^2$.*

Solution We can calculate the area of a hemisphere of radius a by setting

$$f(x,y) = \sqrt{a^2 - x^2 - y^2} \qquad x^2 + y^2 < a^2.$$

We then have

$$f_x = \frac{-x}{\sqrt{a^2 - x^2 - y^2}}, \qquad f_y = \frac{-y}{\sqrt{a^2 - x^2 - y^2}}$$

and so the hemisphere's area is

$$\begin{aligned}
& \iint_{x^2+y^2<a^2} \sqrt{1 + \frac{x^2+y^2}{a^2-x^2-y^2}}\,\mathrm{d}A \\
=& \iint_{x^2+y^2<a^2} \frac{a}{\sqrt{a^2-x^2-y^2}}\,\mathrm{d}A \\
=& \int_{r=0}^{a}\int_{\theta=0}^{2\pi} \frac{a}{\sqrt{a^2-r^2}}\, r\,\mathrm{d}\theta\,\mathrm{d}r \\
=& \; 2\pi\left[-a\sqrt{a^2-r^2}\right]_0^a \\
=& \; 2\pi a^2.
\end{aligned}$$

Hence the area of the whole sphere is $4\pi a^2$. ■

Example 115 *Find the area of the paraboloid $z = x^2 + y^2$ that lies below the plane $z = 4$.*

Solution By Proposition 113 the desired area equals

$$A = \iint_R \sqrt{1 + (2x)^2 + (2y)^2}\,\mathrm{d}A$$

where R is the disc $x^2 + y^2 \leqslant 4$ in the xy-plane. We can parametrize R using polar co-ordinates

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad 0 < r < 2, \quad 0 < \theta < 2\pi,$$

and then we have that

$$\begin{aligned}
A &= \int_{\theta=0}^{2\pi}\int_{r=0}^{2} \sqrt{1 + (2r\cos\theta)^2 + (2r\sin\theta)^2}\, r\,\mathrm{d}r\,\mathrm{d}\theta \\
&= \int_{\theta=0}^{2\pi}\int_{r=0}^{2} \sqrt{1+4r^2}\, r\,\mathrm{d}r\,\mathrm{d}\theta \\
&= 2\pi \int_{r=0}^{2} \sqrt{1+4r^2}\, r\,\mathrm{d}r \\
&= 2\pi \times \frac{1}{8} \times \frac{2}{3} \times \left[\left(1+4r^2\right)^{3/2}\right]_{r=0}^{2} \\
&= \frac{\pi}{6}\left[17^{3/2} - 1\right].
\end{aligned}$$

■

Proposition 116 ***(Surfaces of Revolution)*** *A surface S is formed by rotating the graph of*

$$y = f(x) \qquad a < x < b,$$

about the x-axis. (Here $f(x) > 0$ for all x.) The surface area of S equals

$$\textit{Area}(S) = 2\pi \int_{x=a}^{x=b} f(x)\,\frac{\mathrm{d}s}{\mathrm{d}x}\,\mathrm{d}x.$$

Proof From the parametrization

$$\mathbf{r}(x,\theta) = (x, f(x)\cos\theta, f(x)\sin\theta) \qquad -\pi < \theta < \pi,\ a < x < b$$

we calculated in Example 104 that

$$\mathbf{r}_x \wedge \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & f'(x)\cos\theta & f'(x)\sin\theta \\ 0 & -f(x)\sin\theta & f(x)\cos\theta \end{vmatrix} = \begin{pmatrix} f'(x)f(x) \\ -f(x)\cos\theta \\ -f(x)\sin\theta \end{pmatrix}.$$

So

$$|\mathbf{r}_x \wedge \mathbf{r}_\theta|^2 = f(x)^2 f'(x)^2 + f(x)^2 = f(x)^2(1 + f'(x)^2) = f(x)^2\left(\frac{\mathrm{d}s}{\mathrm{d}x}\right)^2.$$

The result follows. ■

Example 117 *Rederive the area of the paraboloid $z = x^2 + y^2$ that lies below the plane $z = 4$, by thinking of the paraboloid as a surface of revolution.*

Solution We can consider the paraboloid as a rotation of the curve $x = \sqrt{z}$ about the z-axis where $0 < z < 4$. We then have

$$\left(\frac{\mathrm{d}s}{\mathrm{d}z}\right)^2 = 1 + \left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2 = 1 + \left(\frac{1}{2\sqrt{z}}\right)^2 = 1 + \frac{1}{4z}.$$

Hence

$$\begin{aligned} A &= 2\pi \int_{z=0}^{4} x\frac{\mathrm{d}s}{\mathrm{d}z}\,\mathrm{d}z \\ &= 2\pi \int_{z=0}^{4} \sqrt{z}\sqrt{1 + \frac{1}{4z}}\,\mathrm{d}z \\ &= 2\pi \int_{z=0}^{4} \sqrt{z + \frac{1}{4}}\,\mathrm{d}z \\ &= 2\pi \left[\frac{2}{3}\left(z + \frac{1}{4}\right)^{3/2}\right]_0^4 \\ &= \frac{4\pi}{3}\left[\left(\frac{17}{4}\right)^{3/2} - \left(\frac{1}{4}\right)^{3/2}\right] \\ &= \frac{\pi}{6}\left[17^{3/2} - 1\right]. \end{aligned}$$

■

Proposition 118 *Isometries preserve area.*

Proof An isometry is a bijection between surfaces which preserves the lengths of curves. Say that $\mathbf{r}\colon U \to \mathbb{R}^3$ is a parametrization of a smooth surface $X = \mathbf{r}(U)$ and $f\colon \mathbf{r}(U) \to Y$ is an isometry from X to another smooth surface Y. Then the map

$$\mathbf{s} = f \circ \mathbf{r} : U \to Y$$

is a parametrization of Y also using co-ordinates from U.

Consider a curve

$$\gamma(t) = \mathbf{r}(u(t), v(t)) \qquad a \leqslant t \leqslant b$$

in X. By the chain rule

$$\gamma' = u'\mathbf{r}_u + v'\mathbf{r}_v$$

and

$$|\gamma'|^2 = E(u')^2 + 2Fu'v' + G(v')^2$$

where

$$E = \mathbf{r}_u \cdot \mathbf{r}_u, \qquad F = \mathbf{r}_u \cdot \mathbf{r}_v, \qquad G = \mathbf{r}_v \cdot \mathbf{r}_v.$$

The length of γ equals is

$$\mathcal{L}(\gamma) = \int_{t=a}^{t=b} |\gamma'(t)| \, \mathrm{d}t = \int_{t=a}^{t=b} \sqrt{E(u')^2 + 2Fu'v' + G(v')^2} \, \mathrm{d}t.$$

In a similar fashion the length of the curve $f(\gamma)$ equals

$$\mathcal{L}(f(\gamma)) = \int_{t=a}^{t=b} \sqrt{\tilde{E}(u')^2 + 2\tilde{F}u'v' + \tilde{G}(v')^2} \, \mathrm{d}t$$

where

$$\tilde{E} = \mathbf{s}_u \cdot \mathbf{s}_u, \qquad \tilde{F} = \mathbf{s}_u \cdot \mathbf{s}_v, \qquad \tilde{G} = \mathbf{s}_v \cdot \mathbf{s}_v.$$

As f is an isometry then

$$\int_{t=a}^{t=b} \sqrt{E(u')^2 + 2Fu'v' + G(v')^2} \, \mathrm{d}t = \int_{t=a}^{t=b} \sqrt{\tilde{E}(u')^2 + 2\tilde{F}u'v' + \tilde{G}(v')^2} \, \mathrm{d}t.$$

Further as this is true for all b it must follow that

$$E(u')^2 + 2Fu'v' + G(v')^2 = \tilde{E}(u')^2 + 2\tilde{F}u'v' + \tilde{G}(v')^2$$

for all values of t and all functions u, v. By choosing $u = t, v = 0$, we find $E = \tilde{E}$ and we also obtain $G = \tilde{G}$ by setting $u = 0, v = t$. It follows then that $F = \tilde{F}$ as well.

Now the area of a subset $\mathbf{r}(V)$ of X is given by

$$\iint_V |\mathbf{r}_u \wedge \mathbf{r}_v| \, \mathrm{d}u \, \mathrm{d}v.$$

However one can show (Sheet 7, Exercise 5) that

$$|\mathbf{r}_u \wedge \mathbf{r}_v| = \sqrt{EG - F^2}.$$

As

$$|\mathbf{s}_u \wedge \mathbf{s}_v| = \sqrt{\tilde{E}\tilde{G} - \tilde{F}^2} = \sqrt{EG - F^2} = |\mathbf{r}_u \wedge \mathbf{r}_v|$$

then the area of $f(\mathbf{r}(V))$ equals

$$\iint_V |\mathbf{s}_u \wedge \mathbf{s}_v| \, \mathrm{d}u \, \mathrm{d}v = \iint_V |\mathbf{r}_u \wedge \mathbf{r}_v| \, \mathrm{d}u \, \mathrm{d}v$$

and we see that isometries preserve areas. ■

Remark 119 *As angles between curves can similarly be written in terms of E, F, G and the curves' co-ordinates, then isometries also preserve angles.*