Academic notes: 1A Geometry

J. Mak (December 13, 2014) [From notes of R. Gregory, Durham][*]

I. GEOMETRY ON THE PLANE

Vectors in 2-dimensional space $\mathbb{R}^2$ may be written as (x, y), x and y being co-ordinates for vector $\boldsymbol{v}$, encoding how far along the axis they are. Elementary vector operations are addition and multiplication by a scalar, given respectively by

$$\boldsymbol{u} + \boldsymbol{v} = (u_1 + v_1, u_2 + v_2), \qquad \lambda\boldsymbol{v} = (\lambda v_1, \lambda v_2), \tag{1}$$

with $\lambda \in \mathbb{R}$. These operations satisfy the axioms of vector space. For $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^2$ and $\lambda, \mu \in \mathbb{R}$

- closure
  1. $(\boldsymbol{u} + \boldsymbol{v}) \in \mathbb{R}^2$.
  2. $\lambda\boldsymbol{v} \in \mathbb{R}^2$.
- Addition
  1. Associativity: $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$.
  2. Identity: There exists a $\boldsymbol{0} \in \mathbb{R}^2$ where $\boldsymbol{v} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{v} = \boldsymbol{v}$.
  3. Inverse: There exists a $-\boldsymbol{v} \in \mathbb{R}^2$ where $\boldsymbol{v} + (-\boldsymbol{v}) = \boldsymbol{0}$.
  4. Commutativity: $\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{v}$.
- Multiplication,
  1. Associativity: $(\lambda\mu)\boldsymbol{v} = \lambda(\mu\boldsymbol{v})$.
  2. Zero: $0\boldsymbol{v} = \boldsymbol{0}$.
  3. One: $1\boldsymbol{v} = \boldsymbol{v}$.
  4. Distributive: $(\lambda + \mu)\boldsymbol{v} = \lambda\boldsymbol{v} + \mu\boldsymbol{v}$.

It is convenient to write $(x, y) = x(1, 0) + y(0, 1) = x\boldsymbol{e}_1 + y\boldsymbol{e}_2$. Then the set $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ form a basis of $\mathbb{R}^2$, known as the Cartesian basis.

A. Scalar product, lengths and angles

Given $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^2$, the scalar product is defines as

$$\boldsymbol{u}\dot{\boldsymbol{v}} = (u_1 v_1 + u_2 v_2) \in \mathbb{R}. \tag{2}$$

We note that $\boldsymbol{v} \cdot \boldsymbol{v} = v_1^2 + v_2^2 \geq 0$. The modulus of the vector $\boldsymbol{v}$ is defined as

$$|\boldsymbol{v}| = \sqrt{v_1^2 + v_2^2}. \tag{3}$$

The modulus coincides with the length of the vector in this case, under the Euclidean norm (defined by the dot product here).

[*] julian.c.l.mak@googlemail.com

B. Polar co-ordinates

If we let $r = |\boldsymbol{v}|$, then by considering the angle $\boldsymbol{v}$ makes with respect to the base line (e.g., x-axis), we can write a vector $\boldsymbol{v}$ in polar co-ordinates (r, θ), where $r \in \mathbb{R}^+$ and $\theta \in [0, 2\pi)$, so that

$$v_1 = r\cos\theta, \qquad v_2 = r\sin\theta. \tag{4}$$

(The zero vector is not well-defined since any choice of θ will do for $r = 0$).

Lemma I.1 *Suppose $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^2$ and $\theta \in [0.\pi]$ is the angle between the two vectors. Then $\boldsymbol{u}\dot{\boldsymbol{v}} = |\boldsymbol{u}||\boldsymbol{v}|\cos\theta$.*

Proof For $\boldsymbol{u} = (|\boldsymbol{u}|, \phi)$, $\boldsymbol{v} = (|\boldsymbol{v}|, \psi)$, $\theta = |\phi - \psi|$. So

$$\boldsymbol{u}\dot{\boldsymbol{v}} = |\boldsymbol{u}||\boldsymbol{v}|(\cos\phi\cos\psi + \sin\phi\sin\psi) = |\boldsymbol{u}||\boldsymbol{v}|\cos(\phi - \psi) = |\boldsymbol{u}||\boldsymbol{v}|\cos\theta. \tag{5}$$

$\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal iff $\boldsymbol{u} \cdot \boldsymbol{v} = 0$. (This is a more general way for saying two vectors are perpendicular, as we do not necessarily have to restrict ourselves to the Euclidean inner product).

C. Simultaneous equations

Suppose we have the system of equations

$$ax + by = e, \qquad cx + dy = f, \qquad a \cdots f \in \mathbb{R}. \tag{6}$$

Then multiplying and eliminating accordingly gives

$$x = \frac{de - bf}{ad - bc}, \qquad y = \frac{ce - af}{ad - bc}, \tag{7}$$

so unique solution exists if $(ad - bc) \neq 0$. If $(ad - bc) = 0$, then there can be two scenarios:

1. Solution is under-determined. e.g., $2x + 6y = 4$ and $3x + 9y = 6$, and the two equations differ by a constant factor, so all solutions lie on the same line.
2. There is no solution (lines do not intersect in the plane). e.g., $2x + 6y = 6$ and $3x + 9y = 4$.

If we instead write the system of equations in terms of a matrix, for example,

$$\begin{cases} 2x + 6y = 6 \\ 3x + 9y = 4 \end{cases} \qquad \equiv \qquad \begin{pmatrix} 2 & 6 \\ 3 & 9 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}, \tag{8}$$

by using matrix multiplication rules and inverses of matrices, an alternative method can be used to solve simultaneous equations, not necessarily of two variables. In the general case in $\mathbb{R}^2$,

$$\mathsf{A}\boldsymbol{x} = \boldsymbol{b} \to \boldsymbol{x} = \mathsf{A}^{-1}\boldsymbol{b}, \tag{9}$$

where

$$\mathsf{A}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \qquad \mathsf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{10}$$

A^{-1} is the inverse of A, and $|\mathsf{A}| = (ad - bc)$ is known as the determinant of a 2×2 matrix. The adjoint adj$\mathsf{A}$ is the inverse without division by the determinant. The geometric interpretation is then, if $|\mathsf{A}| \neq 0$, two lines intersect uniquely, otherwise there is no intersection, or the lines lie on top of each other.

D. Lines on a plane

We know that $ax + by = c$ is the equation of a straight line. There are several ways to describe lines:

- Parametric form. If $\boldsymbol{v}$ passes through the origin, we can write the line as a collection of all scalar multiples of a direction vector along the line, i.e., $\boldsymbol{x} = \lambda\boldsymbol{v}$. More generally, if $\boldsymbol{a}$ is any point on the line, then $\boldsymbol{x} = \boldsymbol{a} + \lambda\boldsymbol{v}$.
- Normal vector. Let $\boldsymbol{n}$ be a vector orthogonal to $\boldsymbol{v}$, then $\boldsymbol{n} \cdot \boldsymbol{v} = 0$. Thus

  $$\boldsymbol{n} \cdot \boldsymbol{x} = \boldsymbol{n} \cdot \boldsymbol{a} + \lambda\boldsymbol{n} \cdot \boldsymbol{v} = \boldsymbol{n} \cdot \boldsymbol{a}. \tag{11}$$

  A line may therefore be written in terms of $\boldsymbol{a}$ and $\boldsymbol{n}$ as $\boldsymbol{n} \cdot \boldsymbol{x} = n_1x + n_2y = \boldsymbol{n} \cdot \boldsymbol{a}$. It is often convenient to have $|\boldsymbol{n}| = 1$, and denote it $\hat{\boldsymbol{n}}$.

E. Determinants and area

Lemma I.2 *Let $\boldsymbol{u} = (a, c)^T$ and $\boldsymbol{v} = (b, d)^T$. Then $\boldsymbol{u}$ and $\boldsymbol{v}$ are parallel iff $|\mathsf{A}| = 0$.*

Proof If $\boldsymbol{u}$ is parallel to $\boldsymbol{v}$ then $\boldsymbol{u} = \lambda\boldsymbol{v}$. Then $a = \lambda b$ and $c = \lambda = d$. Then $ad - bc = \lambda(bd - bd) = 0$. If $|\mathsf{A}| = 0$, then $ad = bc$, and so $a/b = c/d = \lambda$, and so we arrive at $\boldsymbol{u} = \lambda\boldsymbol{v}$.

Now consider the parallelogram formed from $\boldsymbol{u}$ and $\boldsymbol{v}$, with one vertex on the origin wlog. The claim is that the area of this parallelogram is equal to $|\mathsf{A}|$. The argument goes that the area is base multiplied by the height. Taking $|\boldsymbol{u}|$ to be the case, the height is given by $|\boldsymbol{v}|\sin\theta$, and so the area is $|\boldsymbol{u}||\boldsymbol{v}|\sin\theta$. On the other hand, $\boldsymbol{u}\cdot\boldsymbol{v} = ab + cd = |\boldsymbol{u}||\boldsymbol{v}|\cos\theta$, so squaring both sides gives

$$|\boldsymbol{u}|^2|\boldsymbol{v}|^2\cos^2\theta = (ab + cd)^2, \tag{12}$$

and

$$\text{Area} = |\boldsymbol{u}|^2|\boldsymbol{v}|^2\sin^2\theta = -(ab + cd)^2 + |\boldsymbol{u}|^2|\boldsymbol{v}|^2 = \cdots = (ad - bc)^2 = |\mathsf{A}|^2, \tag{13}$$

from which result follows.

Example Take a parallelogram with vertices

$$\boldsymbol{a} = (1, 3), \qquad \boldsymbol{b} = (4, 4), \qquad , \boldsymbol{c} = (5, 6), \qquad \boldsymbol{d} = (2, 5). \tag{14}$$

Then, noticing that $\boldsymbol{b} - \boldsymbol{a}$ is parallel to $\boldsymbol{c} - \boldsymbol{b}$, we take the spanning vectors as $\boldsymbol{b} - \boldsymbol{a} = (3, 1)$ and $\boldsymbol{c} - \boldsymbol{d} = (1, 2)$, so

$$|\mathsf{A}| = \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = 6 - 1 = 5. \tag{15}$$

From the above, we deduce that the area of a triangle is half of the determinant of the matrix with the spanning vectors of the parallelogram as the entries.

Example For a triangle with vertices

$$\boldsymbol{a} = (-1, 2), \qquad \boldsymbol{b} = (1, 1), \qquad \boldsymbol{c} = (3, 4), \tag{16}$$

we have $\boldsymbol{b} - \boldsymbol{a} = (2, -1)$ and $\boldsymbol{c} - \boldsymbol{d} = (4, 2)$, so

$$\text{Area} = \frac{1}{2}|\mathsf{A}| = \frac{1}{2}\begin{vmatrix} 2 & 4 \\ -1 & 2 \end{vmatrix} = (4 + 4)/2 = 4. \tag{17}$$

F. Curves in the plane

The most familiar curves are graphs of functions. Let $f(x)$ be a real valued function of $x \in \mathbb{R}$, $x \in (a, b)$. The graph $y = f(x)$ is a curve in $\mathbb{R}^2$. The curve is differentiable if we can find a tangent to the graph of f for each $x \in (a, b)$. This is done by taking the limit of chords with

$$y = f(x + \delta x) - f(x), \qquad x = x + \delta x - x, \tag{18}$$

so that the line has gradient $[f(x + \delta x) - f(x)]/\delta x$. By re-scaling and taking the limit as $\delta x \to 0$, this gives the direction vector of the tangent line, or, as a vector,

$$\boldsymbol{t} = \lim_{\delta x \to 0}\begin{pmatrix} 1 \\ [f(x + \delta x) - f(x)]/\delta x \end{pmatrix} = \begin{pmatrix} 1 \\ f'(x) \end{pmatrix}. \tag{19}$$

For fixed x_0, the tangent line is

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ f(x_0) \end{pmatrix} + \boldsymbol{t}\begin{pmatrix} 1 \\ f'(x_0) \end{pmatrix}, \tag{20}$$

with $\boldsymbol{t} = x - x_0$, $y = f(x_0) + \boldsymbol{t}f'(x_0) = f(x_0) + (x - x_0)f'(x_0)$. This is of course the first Taylor expansion approximation of $y = f(x)$ at $x = x_0$.

Example Let $f(x) = x^3 - 6x^2 + 9x$.

1. Find tangent of $y = f(x)$ at $x = 0, 1, 2, 3, 4$.
2. Use this to sketch $y = f(x)$ for $x \in [-1, 5]$.
3. Find the area below the curve and above the x-axis between their two pairs of intersection.

We have $f'(x) = 3x^2 - 12x + 9$. So

x	0	1	2	3	4
$y = f(x)$	0	4	2	0	4
$f'(x)$	9	0	-3	0	9
tangent	$y = 9x$	$y = 4$	$y = 3x + 8$	$y = 0$	$y = 9x - 32$

A sketch of the graph shows a single crossing at $x = 0$ and double crossing at $x = 3$ (can show this by factorising the cubic). So the area under the curve is

$$\int_0^3 (x^3 - 6x^2 + 9x)\,\mathrm{d}x = 27/4. \tag{21}$$

G. Parametric curve

These can deal with places where the gradient becomes infinity, or where there is more than one value of y for a particular x. For example, $(x, y) = (r\cos\theta, r\sin\theta)$ describes a circle centred at the origin, since $x^2 + y^2 = r^2$. A parametric curve in $\mathbb{R}^2$ is a differentiable map

$$\alpha(a, b) \to \mathbb{R}^2, \qquad \alpha(t) = (x(t), y(t)), \qquad t \in (a, b). \tag{22}$$

The curve is differentiable if, for all $t \in (a, b)$, there exists $\alpha'(t) = (x'(t), y'(t))$. Geometrically, $\alpha'(t)$ is the vector tangent to the curve at t. This definition captures the essence of a vector and generalises even to curved spaces. This may be seen via the definition of chords, taking $\delta t \to 0$ for each component.

Example For a circle, $\alpha(t) = (r\cos t, r\sin t)$, and $\alpha'(t) = (-r\sin t, r\cos t)$. More generally, the circle of radius r centred at $\boldsymbol{v}$ is parametrised as $\alpha(t) = (r\cos t + v_1, r\sin t + v_2)$.

For an ellipse centred at the origin, the parametrisation is $\alpha(t) = (a\cos t, b\cos t)$, and $\alpha'(t) = (-a\sin t, b\cos t)$. In Cartesian co-ordinates, this yields

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = c^2, \qquad a, b > 0. \tag{23}$$

The general parametrise form is similar to the circle case.

A parametrised curve is regular is $\alpha'(t)$ exists for $t \in (a, b)$. This means that the parameter is constantly moving along the curve, which allows us to measure distances along the curve. The arc length of a regular smooth parametrised curve, measured from a reference point t_0, is given by

$$s(t) = \int_{t_0}^t |\alpha'(t)|\,\mathrm{d}t = \int_{t_0}^t \sqrt{[x'(t)]^2 + [y'(t)]^2}\,\mathrm{d}t. \tag{24}$$

This is because, for an arc length element,

$$\delta s = \sqrt{\left(\frac{\delta x}{\delta t}\right)^2 + \left(\frac{\delta y}{\delta t}\right)^2}, \qquad \Rightarrow \qquad \sum \delta s = \int_a^b \sqrt{\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2}\,\mathrm{d}t. \tag{25}$$

The curve is said to be parametrised by arc length if $|\alpha'(t)| = 1$. The is equivalent to saying $\delta s \approx \delta t$.

Lemma I.3 *Let $\alpha(t)$ be a regular curve parametrised by arc length. Then $\alpha''(t) = (x''(t), y''(t))$ is normal to $\alpha(t)$.*

Proof We need to show that $\alpha'(t) \cdot \alpha''(t) = 0$. For t the arc length, $|\alpha'(t)|^2 = \alpha'(t) \cdot \alpha'(t) = 1$. So $(\alpha' \cdot \alpha')' = 2\alpha' \cdot \alpha'' = 0$, as required.

Example For a circle with $\alpha(t) = (r\cos t, r\sin t)$, $|\alpha'(t)| = r$, so

$$s(t) = \int_{t_0}^{t} r \, \mathrm{d}t = r(t - t_0), \qquad \sum s = 2\pi r \tag{26}$$

since $t \in [0, 2\pi)$. To parametrise the curve by arc length, we take $t = s/r$, so $\alpha(s) = (r\cos(s/r), r\sin(s/r))$, $\alpha'(s) = (-\sin(s/r), \cos(s/r))$, and $\alpha''(s) = (-1/r)(\cos(s/r), \sin(s/r)) = -(1/r^2)\alpha(s)$. It is easy to see that $\alpha'(s) \cdot \alpha''(s) = 0$. Geometrically, α points away from the circle centre, $\alpha'(s)$ is tangent to the circle, and $\alpha''(s)$ points towards the circle centre. From circular motion, $\alpha = \boldsymbol{x}$, the position vector, $\alpha' = \boldsymbol{v}$ the velocity vector, and $\alpha'' = \boldsymbol{a}$ is the acceleration vector.

Example The <u>cardinod</u> is given by the parametrisation

$$\alpha(t) = 2a(1 - \cos t)(\cos t, \sin t), \tag{27}$$

and looks like a heart shape lying on its side. At $t = 0$, there is a cusp, which will need separate consideration since it is an end point for the cardinod, and the curve is not necessarily smooth there. We have

$$\alpha'(t) = 2a(-\sin t + 2\cos t \sin t, \cos t - \cos^2 t + \sin^2), \tag{28}$$

and, after some algebra (do it yourself), $|\alpha'(t)| = 4a\sin(t/2)$ after using some double angle formulas. The arc length is then

$$s(t) = \int_0^t 4a\sin(t/2) \, \mathrm{d}t = 8a(1 - \cos(t/2)). \tag{29}$$

The total length of a cardinoid is then $s(2\pi) = 16a$.

H. Central conics

Consider a curve $C \in \mathbb{R}^2$, given by the constraint

$$ax^2 = 2bxy + cy^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1. \tag{30}$$

Writing this in polar co-ordinates gives

$$r^2 f(\theta) = 1, \qquad f(\theta) = a\cos^2\theta + 2b\cos\theta\sin\theta + c\sin^2\theta. \tag{31}$$

Since the origin is not within the curve, $r > 0$, so $r = 1/\sqrt{f(\theta)}$. Curve then only exists if $f(\theta) > 0$. Now,

$$\frac{f(\theta)}{\sin^2\theta} = a\cot^2\theta + 2b\cot\theta + c. \tag{32}$$

If $f(\theta) = 0$, $\cot\theta = (-b/a) \pm (\sqrt{b2 - ac}/a)$, and θ is real only if $b^2 - ac \geq 0$. Now, $b^2 - ac = -|\mathsf{A}|$, so if $|\mathsf{A}| > 0$, $f(\theta)$ does not exist. If $ac - b^2 > 0$, then if:

- $a, c > 0 \qquad \Rightarrow \qquad f(\theta) > 0.$
- $a, c < 0 \qquad \Rightarrow \qquad f(\theta) < 0.$

If instead $|\mathsf{A}| = 0$, $ac = b^2$, so multiplying both sides by c, it maybe shown that $c = (bx + cy)^2$, thus $bx + cy = \pm c$, i.e., two parallel straight lines with gradient $-b/c = -a/b$.

In summary:

1. If $|\mathsf{A}| > 0$ and $a, c > 0$, we have a circle, and there are solutions for all θ.
2. If $|\mathsf{A}| > 0$ and $a, c < 0$, there are no solutions since $f(\theta) < 0$ and $r = 1/\sqrt{f(\theta)}$.
3. If $|\mathsf{A}| = 0$, the curve are two parallel straight lines.
4. If $|\mathsf{A}| < 0$, we have hyperbola. There is a solution of F for θ in two open intervals, each of length less than π.

Example For $a = c = 1$, $b = 0$, $|\mathsf{A}| = 1$, and this describes the unit circle centred at the origin. There is a solution for all θ.
for $a = c = 1$, $b = -1/2$, $|\mathsf{A}| = 3/4$, we have the ellipse $x^2 + y^2 - xy = 1$, and may be factorised into

$$\frac{3}{4}(x-y)^2 + \frac{1}{4}(x+y)^2 = 1. \tag{33}$$

There is a solution for all θ.
For $a = b = c = 1$, $|\mathsf{A}| = 0$, and $x + y = \pm 1$, two parallel straight lines.
For $a = 1$, $c = -1$, $b = 0$, $|\mathsf{A}| = -1$, we have $x^2 - y^2 = 1$, and describes hyperbola for $\theta \in [0, \pi/4) \cup (7\pi/4, 2\pi)$ and $\theta \in (3\pi/4, 5\pi/4)$.

It is often useful to find the points of C closest/furthest to/from the origin. These are given by the turning points of C, i.e., $f'(\theta) = (c - a)\sin 2\theta + 2b\cos 2\theta = 0$. An observation we first make is that, for $f'(\theta_0) = 0$, then $f'(\theta_0 + \pi/2) = 0$ also (via straightforward substitution). Then $f'(\theta_0)$ and $f'(\theta_0 + \pi/2)$ are the extremum values of $f(\theta)$. Thus turning points occur at orthogonal lines. The lines $\theta = \theta_0$ and $\theta_0 + \pi/2$ are the principal axes of the conic C.

Example For which values of θ do there exist

$$1 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{34}$$

and what is the smallest value of $r = \sqrt{x^2 + y^2}(= 1/\sqrt{f(\theta)})$ for which solutions exist?
$|\mathsf{A}| = -3$, so solution forms a hyperbola. $f(\theta) = -\cos^2\theta + 4\cos\theta\sin\theta - \sin^2\theta = 4\cos\theta\sin\theta - 1$, and solution exists for $f(\theta) > 0$. Thus we require $\sin\theta > 1/2$, and so $\theta \in (\pi/12, 5\pi/12)$ and $\theta \in (13\pi/12, 17\pi/12)$.
Now, $f'(\theta) = 4(\cos^2\theta - \sin^2\theta)$, so $f' = 0$ if $\cos\theta = |\sin\theta|$. This occurs at $\cos\theta = 1/\sqrt{2}$, and so $\theta = \pi/4, 3\pi/4, 5\pi/4, 17\pi/4$. Taking into account the domain of relevance, $\theta = \pi/4$ and $\theta = 5\pi/4$ are the relevant turning points. It is easy to see the extremum in this case is at $\pi/4$, with $r_{\min} = 1$; since this is a hyperbola, there is no maximum.

Sometimes θ_0 is an extremum point for f, then going back to the relation for $f(\theta_0)$, we notice that

$$\begin{aligned} f(\theta_0)\cos\theta_0 &= a\cos^3\theta_0 + 2b\cos^2\theta_0\sin\theta_0 + c\sin^2\theta_0\cos\theta_0 \\ &= a\cos^3\theta_0 + 2b\cos^2\theta_0 + \sin\theta_0[a\cos\theta_0\sin\theta_0 + b(\sin^2\theta_0 - \cos^2\theta_0)] \\ &= a\cos\theta_0 + b\sin\theta_0 \end{aligned} \tag{35}$$

up on substituting the value of $c\sin 2\theta$ when $f'(\theta_0) = 0$. Similarly, we have

$$f(\theta_0)\sin\theta_0 = b\cos\theta_0 + c\sin\theta_0, \tag{36}$$

so

$$f(\theta_0)\begin{pmatrix} \cos\theta_0 \\ \sin\theta_0 \end{pmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} \cos\theta_0 \\ \sin\theta_0 \end{pmatrix}. \tag{37}$$

For $\hat{\mathbf{r}} = (\cos\theta_0, \sin\theta_0)^T$ (this is a unit vector), we have $\mathsf{A}\hat{\mathbf{r}} = f(\theta_0)\hat{\mathbf{r}}$, and so $\hat{\mathbf{r}}$ is the eigenvector with associated eigenvalue $f(\theta_0)$.
As we will see, matrices are associated with linear mappings, and the eigen-equation says the action of a map on a position vector is to leave it unchanged up to a scale factor λ.

Example The matrix

$$\mathsf{A} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{38}$$

represents a reflection in the y-axis. In this case, the eigenvalues and eigenvectors are

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \lambda_1 = -1, \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \lambda_1 = 1. \tag{39}$$

Going back to conics, θ_0 and $\theta_0 + \pi/2$ gives the extremum values of r, which also corresponds to the direction of the eigenvectors. The eigenvalues gives the distance to the origin via $r = 1/\sqrt{f(\theta)}$, with

$$f(\theta_0)\begin{pmatrix}\cos\theta_0\\ \sin\theta_0\end{pmatrix} = \mathsf{A}\begin{pmatrix}\cos\theta_0\\ \sin\theta_0\end{pmatrix}, \quad f(\theta_0+\pi/2)\begin{pmatrix}\cos\theta_0+\pi/2\\ \sin\theta_0+\pi/2\end{pmatrix} = \mathsf{A}\begin{pmatrix}\cos\theta_0+\pi/2\\ \sin\theta_0+\pi/2\end{pmatrix}. \tag{40}$$

For general θ, we observe that

$$\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} = \begin{pmatrix}\cos(\theta-\theta_0+\theta_0)\\ \sin(\theta-\theta_0+\theta_0)\end{pmatrix} = \begin{pmatrix}\cos(\theta-\theta_0)\cos\theta_0 + \sin\theta_0\sin(\theta-\theta_0)\\ \sin(\theta-\theta_0)\cos\theta_0 + \sin\theta_0\cos(\theta-\theta_0)\end{pmatrix}, \tag{41}$$

so then

$$\begin{aligned}\mathsf{A}\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} &= \cos(\theta-\theta_0)\mathsf{A}\begin{pmatrix}\cos\theta_0\\ \sin\theta_0\end{pmatrix} + \sin(\theta-\theta_0)\mathsf{A}\begin{pmatrix}-\sin\theta_0\\ \cos\theta\end{pmatrix}\\ &= \cos(\theta-\theta_0)f(\theta_0)\begin{pmatrix}\cos\theta_0\\ \sin\theta_0\end{pmatrix} + \sin(\theta-\theta_0)f(\theta_0+\pi/2)\begin{pmatrix}-\sin\theta_0\\ \cos\theta\end{pmatrix}.\end{aligned} \tag{42}$$

Now, we recall that we have

$$1 = r^2(\cos\theta, \sin\theta)\mathsf{A}\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}. \tag{43}$$

Using the above identity, the RHS is

$$\begin{aligned}\text{RHS} =& r^2 f(\theta_0)\cos(\theta-\theta_0)[\cos\theta\cos\theta_0 + \sin\theta\sin\theta_0]\\ &+ r^2 f(\theta_0+\pi/2)\sin(\theta-\theta_0)[-\cos\theta\sin\theta_0 + \sin\theta\cos\theta_0].\end{aligned} \tag{44}$$

Using double angle formulae, we arrive at

$$1 = r^2 f(\theta_0)\cos^2(\theta-\theta_0) + r^2 f(\theta_0+\pi/2)\sin^2(\theta-\theta_0). \tag{45}$$

So we see that if we have (i) an ellipse if both eigenvalues are positive, (ii) a hyperbola if only one of the eigenvalues is positive, (iii) no solution if both eigenvalues are negative.

Example Characterise the conic $1 = 5x^2 + 2\sqrt{3}xy + 3y^2$, an find its largest and smallest distance from the origin.
Now,

$$\mathsf{A} = \begin{pmatrix}5 & \sqrt{3}\\ \sqrt{3} & 3\end{pmatrix}, \tag{46}$$

with $|\mathsf{A}| = 12$, $a, c > 0$, therefore solutions exists for all θ, so we have an ellipse (alternative we could find the eigenvalues). We also have

$$f(\theta) = 5\cos^2\theta + 2\sqrt{3}\cos\theta\sin\theta + 3\sin^2\theta, \qquad f'(\theta) = 2\sqrt{3}\cos 2\theta - 2\sin 2\theta, \tag{47}$$

so extremum values occur where $\tan 2\theta = \sqrt{3}$, and so the principal axes occurs at $\theta = (\pi/6, 2\pi/3, 7\pi/6, 5\pi/3)$. It is seen that $r_{\min} = 1/\sqrt{f(\pi/6)} = 1/\sqrt{6}$ and $r_{\max} = 1/\sqrt{f(2\pi/3)} = \sqrt{2}$ (and similarly with the values another π radians along).

II. GEOMETRY OF COMPLEX NUMBERS

A. Complex numbers and the Argand diagram

We define $\sqrt{-1} = \mathrm{i}$, which is the basic unit imaginary number. A complex number is then a combination of real and imaginary parts $z = a + b\mathrm{i}$, with $a, b \in \mathbb{R}$. The complex numbers $\mathbb{C}$ then obeys the same axioms for addition and multiplication as $\mathbb{R}$ (both are fields).

Consider instead $\mathbb{C}$ as a vector space $z = (x, y)$, where multiplication is defined on $\mathbb{R}^2$ as

$$z_1 \times z_2 = (x_1 x_2 - y_1 y_2, x_1 y_2 - x_2 y_1), \tag{48}$$

and this is commutative. $1 = (1, 0)$ is the identity. So we see that $\mathbb{R}^2$ with this multiplication is a concrete visualisation of $\mathbb{C}$, and is called the Argand diagram.

Given $z = x + \mathrm{i}y$, the conjugate of z is defined to be $\overline{z} = x - \mathrm{i}y$. Geometrically, this represents a reflection of z in the 'real' axis. The real and imaginary part of z is given respectively by

$$\mathrm{Re}(z) = \frac{z + \overline{z}}{2}, \qquad \mathrm{Im}(z) = \frac{z - \overline{z}}{2}. \tag{49}$$

In polar form, $z = r(\cos\theta + \mathrm{i}\sin\theta)$. r is called the modulus of z and is denoted $|z|$, whilst θ is called the argument of z, denoted $\arg(z)$.

B. Geometry of addition and multiplication in $\mathbb{C}$

Addition is as in $\mathbb{R}^2$. From this, we can deduce the triangle inequality.

Lemma II.1 *For $z_1, z_2 \in \mathbb{C}$, $|z_1 + z_2| \leq |z_1| + |z_2|$, and we have an equality iff arg(z_1) = arg(z_2). By corollary, we have $|z_2 + z_2| \geq ||z_1| - |z_2||$.*

Proof Wlog, let $|z_1| > |z_2|$, then $|z_1| = |z_1 + z_2 + (-z_2)| \leq |z_1 + z_2| + |z_2|$ by the triangle inequality for real numbers. So $|z_1| - |z_2| \leq |z_1 + z_2|$, and since $|z_1| > |z_2|$, we have the corollary of the result as required.

For multiplication, we observe that $|z_1 z_2| = |z_1||z_2|$ and $\arg(z_1 z_2) = \arg(z_1) + \arg(z_2)$. Geometrically, this is a spiral scaling. We can use the $\mathbb{C}$-plane to describe various geometrical objects.

Example A circle may be described by $|z - z_0| = a$, where z_0 is the centre of the circle and a is the radius; expanding this accordingly, we see that $a^2 = (x - x_0)^2 + (y - y_0)^2$.

Example The equation $|z - x_0| + |z + x_0| = 2r$ describes an ellipse, where $r > |x_0|$. This may be done via expansion in (x, y). Alternatively, in polar form, we observe that, for $z = a + \mathrm{i}b$, $|z \pm x_0|^2 = (a^2 - b^2)\cos^2\theta \pm 2ax_0\cos\theta + (x_0^2 + b^2)$. If $x_0^2 = (a^2 - b^2)$, then this may be simplified to $|z \pm x_0| = a \pm x_0 \cos\theta$ since $a > x_0$. With this, we obtain $|z - x_0| + |z + x_0| = 2a$, thus, with $x = a\cos\theta$ and $y = b\sin\theta$, this describes an ellipse.

Example The locus of $|z - z_1| = |z - z_2|$ describes the line that is equidistant to the points z_1 and z_2. To see this, expanding everything in x and y and we obtain the equality

$$x(x_2 - x_1) + y(y_2 - y_1) = \frac{y_2^2 - y_1^2}{2} + \frac{x_2^2 - x_1^2}{2}, \tag{50}$$

and the normal to the line is $z_2 - z_1$.

C. de Moivre's theorem

Theorem II.2 *For all $n \in \mathbb{Z}^+$ and angle θ, $\cos n\theta + \mathrm{i}\sin n\theta = (\cos\theta + \mathrm{i}\sin\theta)^n$.*

Proof We do this by induction. The $n = 1$ case is trivial, so, assuming it is true for n, then we observe that

$$\begin{aligned} \cos(n+1)\theta + \mathrm{i}\sin(n+1)\theta &= \cos n\theta\cos\theta + \mathrm{i}^2\sin\theta\sin n\theta + \mathrm{i}\sin n\theta\cos\theta + \mathrm{i}\sin\theta\cos n\theta \\ &= (\cos n\theta + \mathrm{i}\sin n\theta)(\cos\theta + \mathrm{i}\sin\theta) \\ &= (\cos\theta + \mathrm{i}\sin\theta)^{n+1}. \end{aligned} \tag{51}$$

Example

$$\cos 2\theta + \mathrm{i}\sin 2\theta = (\cos\theta + \mathrm{i}\sin\theta)^2 = (\cos^2\theta - \sin^2\theta) + \mathrm{i}(2\sin\theta\cos\theta), \tag{52}$$

and remembering the double angle formulae, the equality agrees. From de Moivre's theorem, we see that

$$\cos n\theta = \mathrm{Re}(\cos\theta + \mathrm{i}\sin\theta)^n, \qquad \sin n\theta = \mathrm{Im}(\cos\theta + \mathrm{i}\sin\theta)^n. \tag{53}$$

We can also use the theorem to find $\sin$ or $\cos$ of rational multiples of π.

Example Express $\sin 4\theta/\cos\theta$ as a polynomial in $\sin\theta$, and hence find $\sin(\pi/4)$.

$$\sin 4\theta = \mathrm{Im}(\cos\theta + \mathrm{i}\sin\theta)^4 = 4\cos^3\theta\sin\theta - 4\cos\theta\sin^3\theta = 4\cos\theta(\sin\theta - 2\sin^3\theta), \tag{54}$$

so $\sin 4\theta/\cos\theta = 4\sin\theta(1-2\sin^\theta)$. Evaluating this $\pi/4$, we see that the LHS is zero. Now, $4\sin(\pi/4) > 0$, so we conclude that $\sin(\pi/4) = 1/\sqrt{2}$, as expected.

Example Find $\cos(k\pi/6)$ for $k = 1, 2, 3, 4, 5$.

Letting $c = \cos\theta$ and $s = \sin\theta$, observe that

$$\sin 6\theta = sc(6c^4 + 6s^4 - 20s^2c^2) = sc(32c^4 - 32c^2 + 6) = 2sc(4c^3 - 3)(4c^2 - 1). \tag{55}$$

Now, $\sin(k\pi) = 0$, so LHS is zero, and since $\sin(k\pi/6) \neq 0$, we have

$$\cos^2(k\pi/6) = 3/4, \quad \cos^2(k\pi/6) = 1/4, \quad \cos\theta = 0 \quad \Rightarrow \qquad \cos(k\pi/6) = \pm\sqrt{3}/2,\ \pm 1/2,\ 0. \tag{56}$$

Since $\cos\theta$ is a decreasing function in $[0,\pi]$, we have

$$\cos(\pi/6) = \sqrt{3}/2, \quad \cos(2\pi/6) = 1/2, \quad \cos(\pi/2) = 0, \quad \cos(2\pi/3) = -1/2, \quad \cos(5\pi/6) = -\sqrt{3}/2. \tag{57}$$

D. Imaginary exponentials

de Moivre's theorem hints at a deeper geometric significance of cosine and sine functions and a way of encoding multiplication by imaginary numbers. Suppose $f(\theta) = \cos\theta + \mathrm{i}\sin\theta$, then we notice that $f'(\theta) = \mathrm{i}f(\theta)$, and, more generally, $f^{(n)}(\theta) = \mathrm{i}^n f(\theta)$. We know that also that the n-th derivative of $\mathrm{e}^{\lambda x}$ is $\lambda^n\mathrm{e}^{\lambda x}$, so this suggests a link with exponential functions; indeed, we have

$$\cos\theta + \mathrm{i}\sin\theta = \mathrm{e}^{\mathrm{i}\theta}. \tag{58}$$

By de Moivre's theroem then,

$$r(\cos n\theta + \mathrm{i}\sin n\theta) = r(\cos\theta + \mathrm{i}\sin\theta)^n = r\mathrm{e}^{\mathrm{i}n\theta}. \tag{59}$$

Lemma II.3 $\mathrm{e}^{\mathrm{i}\pi} + 1 = 0$.

Example Find all the roots of $z^6 + 4z^3 + 8 = 0$.

Factorising the above gives $z^3 = -2 \pm 2\mathrm{i}$. So since $|z^3| = 2\sqrt{2}$, we have $|z| = \sqrt{2}$. Now,

$$\arg(-2+2\mathrm{i}) = \frac{3\pi}{4}, \qquad \arg(-2-2\mathrm{i}) = \frac{5\pi/4}{,} \tag{60}$$

and the argument of the roots z satisfies

$$\arg(z) = \frac{3\pi/4 + 2n\pi}{3}, \qquad \arg(z) = \frac{5\pi/4 + 2n\pi}{3}, \tag{61}$$

where the division by 3 is to take into account the cube root, and the $2n\pi$ factors is to account for all the roots. This eventually yields

$$z = \sqrt{2}(\mathrm{e}^{\mathrm{i}\pi/4}, \mathrm{e}^{5\mathrm{i}\pi/4}, \mathrm{e}^{11\mathrm{i}\pi/12}, \mathrm{e}^{13\mathrm{i}\pi/12}, \mathrm{e}^{19\mathrm{i}\pi/12}, \mathrm{e}^{21\mathrm{i}\pi/21}). \tag{62}$$

III. GEOMETRY OF $\mathbb{R}^3$

A. Vectors and their products

We assume vectors $\boldsymbol{v} \in \mathbb{R}^3$ satisfy the same axioms of addition and multiplication as for $\mathbb{R}^2$. Taking the usual basis, we take the dot product to be

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + u_3 v_3, \tag{63}$$

so that

$$|\boldsymbol{v}| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}} = \sqrt{v_1^2 + v_2^2 + v_3^2}, \qquad \boldsymbol{u} \cdot \boldsymbol{v} = |\boldsymbol{u}||\boldsymbol{v}| \cos\theta. \tag{64}$$

Again, $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

In $\mathbb{R}^3$, we define the vector (cross) product as

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix}. \tag{65}$$

It may be seen that $\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3$, and the sign is kept if the indices are permuted in a cyclic fashion. We also notice that $\boldsymbol{v} \cdot \boldsymbol{u} = -\boldsymbol{u} \cdot \boldsymbol{v}$. In general, $\boldsymbol{u} \times \boldsymbol{v}$ generates a vector orthogonal to both of the vectors, such that $\boldsymbol{u} \times \boldsymbol{v}$ points according to the clockwise/right-hand screw convention.

Lemma III.1 *For non-trivial $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$ and $\theta \in [0, \pi]$ be the angle between them, then $|\boldsymbol{u} \times \boldsymbol{v}| = |\boldsymbol{u}||\boldsymbol{v}| \sin\theta$.*

Proof It may be shown that $|\boldsymbol{u} \times \boldsymbol{v}|^2 = \cdots = |\boldsymbol{u}|^2 |\boldsymbol{v}|^2 - (\boldsymbol{u} \cdot \boldsymbol{v})^2$. Then, using the identity for $\boldsymbol{u} \cdot \boldsymbol{v}$, we obtain the identity.

B. Simultaneous equations in $\mathbb{R}^3$

Suppose we want to solve the simultaneous system of equations

$$\begin{cases} ax + by + cz &= l, \\ dx + ey + fz &= m, \\ gx + hy + jz &= n. \end{cases} \tag{66}$$

A similar approach using matrices results in

$$\mathsf{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & j \end{pmatrix}, \qquad \mathsf{A}\boldsymbol{x} = \boldsymbol{b}. \tag{67}$$

Supposing the inverse A^{-1} exists, then we may solve the system uniquely; the existence of an unique solution again depends on the determinant of the matrix. By brute force or otherwise,

$$|\mathsf{A}| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & j \end{vmatrix} = a \begin{vmatrix} e & f \\ h & j \end{vmatrix} - b \begin{vmatrix} d & f \\ g & j \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}. \tag{68}$$

(This is the expression of the determinant by expanding the first row; it may be done by expanding any column or row, although one needs to take into account of extra minus signs in entries where $i + j = 2n$, with $\mathsf{A} = (A_{ij})$, A_{ij} the entry at the i-th row and j-th column.) With the determinant, A^{-1} may be found by the following steps:

1. Matrix of minors. We find the determinant of each 2×2 matrix, where the element corresponding to that row and column is covered, i.e.,

$$\begin{pmatrix} ej - fh & dj - fg & dh - eg \\ bj - ch & aj - cg & ah - bg \\ bf - ce & af - cd & ae - bd \end{pmatrix}. \tag{69}$$

2. Matrix of co-factors. Change the sign of the places where $i + j = 2n$, where i and j are the row and column number (starting the count from one), i.e.,

$$\begin{pmatrix} ej - fh & -(dj - fg) & dh - eg \\ -(bj - ch) & aj - cg & -(ah - bg) \\ bf - ce & -(af - cd) & ae - bd \end{pmatrix} = \begin{pmatrix} ej - fh & fg - dj & dh - eg \\ ch - bj & aj - cg & bg - ah \\ bf - ce & cd - af & ae - bd \end{pmatrix}. \tag{70}$$

3. Adjoint. Take a transpose of the matrix of co-factors, i.e.,

$$\mathsf{adj}(\mathsf{A}) = \begin{pmatrix} ej - fh & ch - bj & bf - ce \\ fg - dj & aj - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{pmatrix}. \tag{71}$$

(Think of the transpose as reflecting everything about the main diagonal.)

4. Inverse. Divide the adjoint by the determinant, i.e., $\mathsf{A}^{-1} = \mathsf{adj}(\mathsf{A})/|\mathsf{A}|$.

C. Planes in $\mathbb{R}^3$

Consider the constraint $ax + by + cz = l$, which describes a plane in $\mathbb{R}^3$. This is the case because, assuming $c \neq 0$ wlog, we have

$$z = \frac{l - ax - by}{c}, \qquad \Rightarrow \qquad \frac{\mathrm{d}z}{\mathrm{d}x} = -\frac{a}{c}, \qquad \frac{\mathrm{d}z}{\mathrm{d}y} = -\frac{b}{c}, \tag{72}$$

and there exists a unique z for every x and y. The derivatives are constant which implies that $z(x, y)$ traces out a two-dimensional plane in $\mathbb{R}^3$.

An alternative description of the plane can be given in terms of vectors, again as $\boldsymbol{n} \cdot \boldsymbol{r} = \boldsymbol{a} \cdot \boldsymbol{n}$, where $\boldsymbol{n}$ is a normal vector to the plane, $\boldsymbol{r} = (x, y, z)$, and $\boldsymbol{a}$ is a position vector in the plane. As a trivial example, the (x, y) plane is described by the equation with $a, b, l = 0$ and $c = 1$. We have $\boldsymbol{n} = \boldsymbol{e}_z$, $\boldsymbol{a} = \boldsymbol{0}$.

Example For $3x + 4y - 2z = 12$, we have $\boldsymbol{n} = (3, 4, -2)$, and one possibility for $\boldsymbol{a}$ is $(4, 0, 0)$.

The description is as in $\mathbb{R}^2$, except now we require two direction vectors. The equation of the plane is given as $\boldsymbol{r} = \boldsymbol{a} + \lambda \boldsymbol{d}_1 + \mu \boldsymbol{d}_2$. Eliminating λ and μ, it may be shown that $\boldsymbol{r} \cdot (\boldsymbol{d}_1 \times \boldsymbol{d}_2) = \boldsymbol{a} \cdot (\boldsymbol{d}_1 \times \boldsymbol{d}_2)$, i.e., $\boldsymbol{d}_1 \times \boldsymbol{d}_2$ serves as a normal vector to describe the plane.

Example Find the equation of the plane that goes includes the points

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}. \tag{73}$$

We have

$$\boldsymbol{d}_1 = \boldsymbol{b} - \boldsymbol{a} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}, \qquad \boldsymbol{d}_2 = \boldsymbol{c} - \boldsymbol{a} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \qquad \Rightarrow \qquad \boldsymbol{n} = \boldsymbol{d}_1 \times \boldsymbol{d}_2 = \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}, \tag{74}$$

with $\boldsymbol{n} \cdot \boldsymbol{a} = 7$. So one possibility is $4x + y + 2z = 7$.

D. Lines in $\mathbb{R}^3$

A line is given by a point and a direction vector as $\boldsymbol{r} = \boldsymbol{a} + \lambda \boldsymbol{d}$. It is then given by two constraints:

$$\frac{x - a_1}{d_1} = \frac{y - a_2}{d_2} = \frac{z - a_3}{d_3} = \lambda. \tag{75}$$

This line also has two normal vectors. Since only one of these can ever be eliminated, there are two independent solutions. Both

$$\boldsymbol{n} = (d_2, -d_1, 0), \qquad \text{and} \qquad \boldsymbol{n} \times \boldsymbol{d} = (-d_1 d_3, -d_2 d_3, d_1^2 + d_2^2) \tag{76}$$

may serve as normal vectors.

Two planes in $\mathbb{R}^3$ are either parallel or intersect in a line. Suppose the two normal vectors are $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, then if they are parallel, $\boldsymbol{n}_1 = \lambda \boldsymbol{n}_2$, otherwise they intersect in a line $\boldsymbol{r} = \boldsymbol{a} + \lambda \boldsymbol{d}$, with $\boldsymbol{d} = \boldsymbol{n}_1 \times \boldsymbol{n}_2$. To find a point on the line, we may take one of the variables to be zero, and solve for the remaining linear system.

Example Find the intersection of the planes $x - 2y + z = 3$ and $y + 2z = 1$.

For $\boldsymbol{n}_1 = (1, -2, 1)$ and $\boldsymbol{n}_2 = (0, 1, 2)$, $\boldsymbol{d} = (-5, -2, 1)$. Taking $z = 0$, $y = 1$ from the second equation, so $x = 5$, thus $\boldsymbol{r} = (5, 1, 0) = \lambda(-5, -2, 1)$, or, in equation form,

$$\frac{x-2}{-5} = \frac{y-1}{-2} = z. \tag{77}$$

Example Find the parametric form of the line $(x-2)/3 = (y-4)/-1 = (z+5)/-2$.

$\boldsymbol{r} = (2, 4, -5) + \lambda(3, -1, -2)$.

If $\boldsymbol{r} = (a, b, c) + \lambda(d, e, 0)$, then the constraints are

$$\frac{x-a}{d} = \frac{y-b}{e} \qquad \text{and} \qquad z = c. \tag{78}$$

Two lines in $\mathbb{R}^3$ do not necessarily intersect even if they are non-parallel. We now try to find the shortest distance between two lines that do not intersect (if they do, this is of course zero). Suppose $\boldsymbol{r}_1 = \boldsymbol{a}_1 + \lambda \boldsymbol{d}_1$ and $\boldsymbol{r}_2 = \boldsymbol{a}_2 + \lambda \boldsymbol{d}_2$. Then we may have

- Two lines are <u>parallel</u>, with $\boldsymbol{d}_1 = k\boldsymbol{d}_2$. Then the distance is $D = |\boldsymbol{a}_1 - \boldsymbol{a}_2| \sin\theta$, since $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are just position vectors on the line. Using the cross product corollary, we have

$$D = \frac{|\boldsymbol{a}_1 - \boldsymbol{a}_2| \times \boldsymbol{d}_1|}{|\boldsymbol{d}_1|}. \tag{79}$$

- If the two lines are <u>skewed</u>, then $\boldsymbol{n} = \boldsymbol{d}_1 \times \boldsymbol{d}_2 \neq \boldsymbol{0}$ is a normal to the lines. The minimum distance between two points on the line corresponds to the difference vector between the pairs parallel to $\boldsymbol{n}$, i.e.,

$$D\frac{\boldsymbol{n}}{|\boldsymbol{n}|} = \pm(\boldsymbol{a}_1 + \lambda \boldsymbol{d}_1 - \boldsymbol{a}_2 - \mu \boldsymbol{d}_2). \tag{80}$$

  Taking a dot product of the above with $\boldsymbol{n}/|\boldsymbol{n}|$ and noting that $\boldsymbol{n} \cdot \boldsymbol{d}_1 = \boldsymbol{n} \cdot \boldsymbol{d}_2 = 0$, we have

$$D = |\boldsymbol{a}_1 - \boldsymbol{a}_2| \cdot \frac{\boldsymbol{n}}{|\boldsymbol{n}|} = \frac{|(\boldsymbol{a}_1 - \boldsymbol{a}_2) \cdot (\boldsymbol{d}_1 \times \boldsymbol{d}_2)|}{|\boldsymbol{d}_1 \times \boldsymbol{d}_2|}. \tag{81}$$

Example Find the distance between the lines described by

$$\frac{x-2}{-3} = \frac{y-2}{6} = \frac{z+1}{9} \qquad \text{and} \qquad \frac{x+1}{2} = \frac{y}{-4} = \frac{z-2}{-6}. \tag{82}$$

$\boldsymbol{n} = (-3, 6, 9) \times (2, -4, -6) = \boldsymbol{0}$ so two lines are parallel, and using the appropriate formula, we have $D = \sqrt{122/7}$.

Example Find the distance between the two lines described by

$$\boldsymbol{r}_1 = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ 6 \\ 9 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{r}_2 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}. \tag{83}$$

$\boldsymbol{d}_1 \times \boldsymbol{d}_2 = (-72, 0, -24)$, and $\boldsymbol{a}_1 - \boldsymbol{a}_2 = (3, 2, -3)$, so $D = 144/(24\sqrt{10}) = 3\sqrt{10}/5$.

E. Vector products and volumes of polyhedra

Given three vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3$, we define the scalar triple product as $[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}] = \boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$.

Lemma III.2 *The scalar triple product is invariant under cyclic permutations, i.e.,* $[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}] = [\boldsymbol{c}, \boldsymbol{a}, \boldsymbol{c}] = [\boldsymbol{b}, \boldsymbol{c}, \boldsymbol{a}]$. *(May beshown via brute force calculation.)*

Lemma III.3 *The scalar triple product is the determinant of the matrix with* $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ *as its columns. (Thus* $|\mathsf{A}|$ *is unchanged if we cyclically permute its columns.)*

Example The three vectors $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ describes three edges of the unit cube. The scalar triple product of these three vectors is of course 1, since $(\boldsymbol{a}|\boldsymbol{b}|\boldsymbol{c}) = \mathsf{I}$, which coincidentally is the volume of the unit cube.

For a sheared cubed described by $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}3)$, which has the same volume as the cube, the scalar triple product is also 1.

The latter object is a parallelpiped, which is a three-dimensional polyhedron with six quadrilateral faces where opposite faces are parallel. Its relation to the cube is like that of a parallelogram to a square. The volume of the parallelpiped is give by the base area multiplied by the height; these are respectively given by $|\boldsymbol{a}||\boldsymbol{b}|\sin\theta = |\boldsymbol{a} \times \boldsymbol{b}|$ and $|\boldsymbol{c}|\cos\phi = \boldsymbol{c} \cdot \hat{\boldsymbol{n}} = |\boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b})/|\boldsymbol{a} \times \boldsymbol{b}||$. Thus

$$V = (|\boldsymbol{a}||\boldsymbol{b}|\sin\theta)(|\boldsymbol{c}|\cos\phi) = |\boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b})|, \tag{84}$$

and the volume of a parallelpiped spanned by three vectors is given by the determinant of a matrix formed by the three vectors.

Example Find the volume of the parallelpiped with vertices

$$\begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}. \tag{85}$$

The first vector has the lowest value of x, y, z so we take it as the origin of the spanning vectors. We see that three of the vectors are invariant in x, y, z in respect to the first vector, given by the fifth, fourth and second vector respectively. Then the spanning vectors are $(0, 1, 2)$, $(2.0.1)$ and $(1, 2, 0)$ which gives a volume of 9.

For a tetrahedron (triangular bottom pyramid), the volume is a third of the base area multiplied by height; for triangles, this is

$$V = \frac{1}{3}\frac{1}{2}|\boldsymbol{a}||\boldsymbol{b}|\sin\theta|\boldsymbol{c}\cos\phi| = \frac{1}{6}|\boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b})|. \tag{86}$$

Example Find the volume of a tetrahedron with vertex

$$\boldsymbol{a} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{c} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \boldsymbol{d} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}. \tag{87}$$

Taking $\boldsymbol{a}$ to be the origin, the vectors are $\boldsymbol{b} - \boldsymbol{a} = (1, 0, 2)$, $\boldsymbol{c} - \boldsymbol{a} = (2, 1, 3)$, $\boldsymbol{d} - \boldsymbol{a} = (0, 1, 2)$, and the resulting determinant of the matrix is 1, so the volume is $1/6$.

As a final topic, the triple vector product is $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$. This may be checked by brute force, and analogous results exists for any number of vectors.

F. Intersection of planes and simultaneous equations

Two non-parallel planes intersect in a line. Suppose two planes π_1 and π_2 are described by $ax+by+cz = l$ and $dx+ey+fz = m$, then the intersection $\mathcal{C} = \boldsymbol{a} + \lambda(\boldsymbol{n}_1 \times \boldsymbol{n}_2)$, where $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ are the respective normal vectors to the planes. Suppose we now have a third plane π_3 described by $gx + hy + jz = n$, then this plane π_3 will intersect $\mathcal{C}$ at a point if it is no parallel to π_1 or π_2. In equation form, there is intersection if

$$\boldsymbol{r} \cdot \boldsymbol{n}_3 = \boldsymbol{a} \cdot \boldsymbol{n}_3 + \lambda[\boldsymbol{n}_1, \boldsymbol{n}_2, \boldsymbol{n}_3] = n, \tag{88}$$

so that

$$\lambda = \frac{n - \boldsymbol{a}\cdot\boldsymbol{n}_3}{[\boldsymbol{n}_1, \boldsymbol{n}_2, \boldsymbol{n}_3]}, \tag{89}$$

and a solution is well-defined if $[\boldsymbol{n}_1, \boldsymbol{n}_2, \boldsymbol{n}_3] \neq 0$, i.e., the determinant of the matrix formed by the normal vectors is non-zero.

Lemma III.4 (Cramer's rule) *It may be shown that, from algebraic manipulation, if* $\mathsf{A}\boldsymbol{r} = \boldsymbol{b}$*, then*

$$x = |\mathsf{A}_1|/|\mathsf{A}|, \qquad y = |\mathsf{A}_2|/|\mathsf{A}|, \qquad z = |\mathsf{A}_3|/|\mathsf{A}|, \tag{90}$$

where A_j *has the* j^{th} *column replaced by* $\boldsymbol{b}$*.*

Further rules for evaluating the determinant of matrices:

- $|\mathsf{A}| = |\mathsf{A}^T|$.
- If B is formed by multiply a row or column of A by λ, then $|\mathsf{B}| = \lambda|\mathsf{A}|$.
- If B is formed by interchanging two rows/columns of A, then $|\mathsf{B}| = -|\mathsf{A}|$. A cyclic permutation involves two such operation, so the determinant is unchanged in this case.
- If B is formed by adding an arbitrary multiple of a single row/column of A to another row/column, then $\mathsf{B} = \mathsf{A}$. e.g.,

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & j \end{vmatrix} = \begin{vmatrix} a-\lambda d & b-\lambda e & c-\lambda f \\ d & e & f \\ g & h & j \end{vmatrix}. \tag{91}$$

G. Method of row reduction

Using the last rule above, a matrix may be manipulated into an upper-triangular/echelon form, i.e.,

$$\mathsf{A}' = \begin{pmatrix} a' & b' & c' \\ 0 & e' & f' \\ 0 & 0 & j' \end{pmatrix}, \qquad |\mathsf{A}'| = a'e'j'. \tag{92}$$

Example Solve the simultaneous equations

$$\begin{aligned} 2x + 3y + z &= 23, \\ x + 7y + z &= 36, \\ 5x + 4y - 3z &= 16. \end{aligned} \tag{93}$$

$$\begin{aligned} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 23 \\ 1 & 7 & 1 & 36 \\ 5 & 4 & -3 & 10 \end{array}\right) = \left(\begin{array}{ccc|c} 2 & 3 & 1 & 23 \\ 1 & 7 & 1 & 36 \\ 0 & -7/2 & -11/2 & -83/2 \end{array}\right) &= \left(\begin{array}{ccc|c} 2 & 3 & 1 & 23 \\ 0 & 11/2 & 1/2 & 49/2 \\ 0 & -7/2 & -11/ & -83/2 \end{array}\right) \\ &= \left(\begin{array}{ccc|c} 2 & 3 & 1 & 23 \\ 0 & 11/2 & 1/2 & 49/2 \\ 0 & 0 & -57/2/ & -285/11 \end{array}\right). \end{aligned} \tag{94}$$

From this, we can back substitute and see that $z = 5$, $y = 4$ and $x = 3$.

IV. CURVES AND SURFACES IN $\mathbb{R}^3$

A. Parametric curves and surfaces

We are used to functions that go from $\mathbb{R}$ to $\mathbb{R}$. A parametric curve is simple the generalisation of this to the case where a function takes $\mathbb{R}$ to $\mathbb{R}^3$,

$$\gamma:\ \mathbb{R}\to\mathbb{R}^3, \quad t\mapsto (x(t), y(t), z(t)). \tag{95}$$

Example $\lambda\mapsto \boldsymbol{a}+\lambda\boldsymbol{d}$ is the parameteric mapping for a line.

Example $t\mapsto(\cos t, \sin t, t)$ represents a helix gyrating in the xy-plane.

Just as we can differentiate real functions, we can differentiate vector functions as

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t}=\left(\frac{\mathrm{d}x}{\mathrm{d}t}, \frac{\mathrm{d}y}{\mathrm{d}t}, \frac{\mathrm{d}z}{\mathrm{d}t}\right)=\boldsymbol{\tau}. \tag{96}$$

Example In the above two examples, the tangent vectors are respectively $\boldsymbol{d}$ and $(-\sin t, \cos t, 1)$.

The same set of points may be parametrised in different ways. For example, the parabola on the $x=1$ plane may be parametrised by $\gamma=(1,t,t^2)$ or $\beta=(1,\mathrm{e}^{\lambda},\mathrm{e}^{2\lambda})$. However, γ and β are strictly different because they are parametrised using a different variable.

As in $\mathbb{R}^2$, for $\boldsymbol{r}(t)=(x(t),y(t),z(t))$, the arc length s is given by

$$s(t)=\int_{t_0}^{t}\sqrt{\dot{x}^2+\dot{y}^2+\dot{z}^2}\,\mathrm{d}t. \tag{97}$$

We have $\mathrm{d}s/\mathrm{d}t=|\mathrm{d}\boldsymbol{r}/\mathrm{d}t|$, and, at infinitesimal ranges, this is often written as

$$\mathrm{d}s^2=\left|\frac{\mathrm{d}r}{\mathrm{d}t}\right|^2\mathrm{d}t^2=(\dot{x}^2+\dot{y}^2+\dot{z}^2)\mathrm{d}t^2=(\mathrm{d}x)^2+(\mathrm{d}y)^2+(\mathrm{d}z)^2. \tag{98}$$

(cf. 3D version of Pythagoras' theorem.) A curve is said to be parametrised by arc length if $|\mathrm{d}\boldsymbol{r}/\mathrm{d}t|=1$, and in this case we use s instead of t are the parameter. Note that a curve is a collection of points, whilst the tangent is a vector, with a definite direction and length.

Lemma IV.1 *If $\gamma(s)$ is parametrised by arc length with tangent vector $\boldsymbol{\tau}(s)$, then $\mathrm{d}/\mathrm{d}t(\boldsymbol{\tau}(s))$ is normal to the curve γ.*

Proof Since γ is parametrised by arc length, $\boldsymbol{\tau}\cdot\boldsymbol{\tau}=1$. So then

$$\frac{\mathrm{d}}{\mathrm{d}s}(\boldsymbol{\tau}\cdot\boldsymbol{\tau})=2\boldsymbol{\tau}\cdot\frac{\mathrm{d}\boldsymbol{\tau}}{\mathrm{d}s}=0. \tag{99}$$

We can define a surface parametrically, as

$$S:\mathbb{R}^2\to\mathbb{R}^3, \qquad (\mu,\lambda)\mapsto(x(\mu,\lambda),y(\mu,\lambda),z(\mu,\lambda))=\boldsymbol{X}(\mu,\lambda). \tag{100}$$

S then is a function of two variables.

Example $S^2=(\sin\theta\cos\phi,\sin\theta\sin\phi,\cos\theta)$ with $\theta\in[0,\pi]$ and $\phi\in[0,2\pi]$ is the unit 2-sphere.

Example $S=(\sinh\chi,\cos\phi,\sinh\chi\sin\phi,\cosh\chi)$ with $\chi\in\mathbb{R}$ and $\phi\in[0,2\pi]$ gives a hyperboloid.

If $\boldsymbol{X}(\mu,\lambda)$ traces a smooth surface S in $\mathbb{R}^3$, then its tangent vectors at some (μ_0,λ_0) are

$$\boldsymbol{\tau}_1(\mu_0,\lambda_0)=\left.\frac{\partial\boldsymbol{X}}{\partial\mu}\right|_{(\mu_0,\lambda_0)}, \qquad \boldsymbol{\tau}_2(\mu_0,\lambda_0)=\left.\frac{\partial\boldsymbol{X}}{\partial\lambda}\right|_{(\mu_0,\lambda_0)}. \tag{101}$$

The tangent plane of S at (μ_0,λ_0) is then given by

$$\Pi=\{\boldsymbol{r}\in\mathbb{R}^3\mid\boldsymbol{r}=\boldsymbol{X}_0+\alpha\boldsymbol{\tau}_1+\beta\boldsymbol{\tau}_2,\ \alpha,\beta\in\mathbb{R}\} \tag{102}$$

Example Find the tangent plane of the unit 2-sphere at $(\theta_0,\phi_0)=(\pi/2,0)$. [at the equator]

We have

$$\frac{\partial\boldsymbol{X}}{\partial\theta}=(\cos\theta\cos\phi,\cos\theta\sin\phi,-\sin\theta), \qquad \frac{\partial\boldsymbol{X}}{\partial\phi}=(-\sin\theta\sin\phi,\sin\theta\cos\phi,0), \tag{103}$$

and, at the location of interest, $\boldsymbol{X}_0=\boldsymbol{e}_1$, $\boldsymbol{\tau}_1=-\boldsymbol{e}_3$, and $\boldsymbol{\tau}_2=\boldsymbol{e}_2$, so $\pi=\boldsymbol{e}_1-\alpha\boldsymbol{e}_3+\beta\boldsymbol{e}_2$, i.e., the yz-plane translated to $x=1$.

B. Functions, surfaces and gradients

Like planes, surfaces and tangent planes may be written in Cartesian form or using a function f, i.e., as

$$ax + by + cz = d \qquad \text{or} \qquad f(x, y, z) = d. \tag{104}$$

A scalar function (or scalar field) on $\mathbb{R}^3$ is a map

$$f : \mathbb{R}^3 \to \mathbb{R},\ (x, y, z) \mapsto f(x, y, z). \tag{105}$$

f is continuous at $\boldsymbol{a}$ if $|f(\boldsymbol{r}) - f(\boldsymbol{a})| \to 0$ as $|\boldsymbol{r} - \boldsymbol{a}| \to 0$ along any path, and is differentiable is the partial derivatives exists.

Example For $f = x^4 + x^2y^2 + xyz$,

$$\frac{\partial f}{\partial x} = 4x^3 + 2xy^2 + yz, \qquad \frac{\partial f}{\partial y} = 2x^2y + xz, \qquad \frac{\partial f}{\partial z} = xy, \tag{106}$$

and

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \tag{107}$$

is the gradient of the function f at $\boldsymbol{x}_0$.

∇f gives us the steepest rate of change of f, with the length of the vector representing the rate of change.

Example $f(x, y) = 1 - (x^2 + y^2)$ is a cone of height 1 with apex at the origin. $\nabla f = -2(x, y)$, and points inwards and uphill.

The intuition is that ∇f points orthogonally from contours of constraints in f.

Example $f = \sqrt{x^2 + y^2 + z^2} = r$ is the distance from origin. $\nabla f = (x, y, z)/r = \hat{\boldsymbol{r}}$, and ∇f describes a sphere.

Example The hyperboloid $f = x^2 + y^2 - z^2$ has $\nabla f = 2(x, y, -z)$, and $|\nabla f| = 2\sqrt{x^2 + y^2 + z^2}$. This points radially outwards on the xy-plane but down/up in the upper/lower half of $\mathbb{R}^3$.

∇f defines a normal $\boldsymbol{n}$ to the surface $f = \text{const}$, and hence defines the tangent plane via $\boldsymbol{r} \cdot \boldsymbol{n} = \boldsymbol{a} \cdot \boldsymbol{n}$.

Example Find the tangent plane to $x^2 - y^2 - z^2 = 1$ at $(-1, 0, 0)$. [This is a rotated paraboloid with the bottom at $(-1, 0, 0)$.]
$\nabla f = 2(x, -y, -z)$, and so $\boldsymbol{n} \cdot \boldsymbol{r} = -2x$ whilst $\boldsymbol{a} \cdot \boldsymbol{n} = -2$. Thus $x = -1$ is the tangent plane.

Example Find the tangent plane to $x^2 + yz = 1$ at $(0, 1, 1)$. [An inclined hyperboloid.]
$\boldsymbol{n} = \nabla f = (2x, z, y)$, $\boldsymbol{r} = (0, 1, 1)$, so we have $\boldsymbol{n} \cdot \boldsymbol{r} = z + y$, $\boldsymbol{a} \cdot \boldsymbol{n} = 2$, so the tangent at that location satisfies the relation $z + y = 2$.

If $\partial f/\partial x = 0$ at a location, then f has a turning point in the x-direction. In the second example, the turning point in the x-direction is a minimum. Looking at f for fixed y and z shows this also.

A point $P = \boldsymbol{x}_0$ is a critical point of f if $\nabla f = \boldsymbol{0}$ there.

Example $f(x, y, z) = x^2 + yz$, $\nabla f = \boldsymbol{0}$ at $\boldsymbol{x}_0 = \boldsymbol{0}$. If y and z are of the same sign, it is a minimum, whilst if they are of different sign, it is a maximum. This an example of a saddle point; it is a minimum when it is coming down the spine of the hyperboloid, but a maximum when coming up perpendicular to the spine.

We classify extrema by using the second order derivative, which is a matrix called the Hessian, with

$$\mathsf{H}_{ij} = \frac{\partial^2 f}{\partial x^i x^j}. \tag{108}$$

Example Again, with $f = x^2 + yz$, we have

$$\mathsf{H} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{109}$$

Note that H is symmetric assuming f is sufficiently smooth so that derivative order may be interchanged. The eigenvalue of H then give us information about the nature of the extrema.

Example $\mathsf{H}\boldsymbol{r} = \lambda\boldsymbol{r}$ gives $\lambda = \pm 1, 2$. Positive eigenvalues means a local minimum along the direction of the eigenvector, and the opposite is true for negative eigenvalues. If

- All eigenvalues positive, we have an absolute minimum of f.
- All eigenvalues negative, we have an absolute maximum of f.
- A saddle point if two eigenvalues are positive and one is negative.
- Hyperboloid of two sheets if two eigenvalues are negative and one is positive.

So again we confirm that we have a saddle point at the relevant location.

C. Eigenvalue and eigenvectors of 3×3 matrix

Recall that $\boldsymbol{v}$ is an eigenvector of A if $\mathsf{A}\boldsymbol{v} = \lambda = \boldsymbol{v}$, and λ is the eigenvalue associated with the eigenvector.

Lemma IV.2 *If λ is an eigenvalue of A, then $|\mathsf{A} - \lambda\mathsf{I}| = 0$.*

Proof Assuming that we have an eigenvector $\boldsymbol{v}$ associated with λ, suppose that $|\mathsf{A} - \lambda\mathsf{I}| \neq 0$. Then there exists B that is the inverse of $(\mathsf{A} - \lambda\mathsf{I})$. However,

$$\boldsymbol{v} = \mathsf{I}\boldsymbol{v} = \mathsf{B}(\mathsf{A} - \lambda\mathsf{I})\boldsymbol{v} = \mathsf{B}(\mathsf{A}\boldsymbol{v} - \lambda\mathsf{I}), \tag{110}$$

and since $\mathsf{A}\boldsymbol{v} = \lambda\mathsf{I}$, for $\boldsymbol{v} \neq \boldsymbol{0}$, we have a contradiction.

$|\mathsf{A} - \lambda\mathsf{I}|$ gives a polynomial in λ, which is a cubic in when we are concerned with vectors in $\mathbb{R}^3$. This polynomial is called the characteristic polynomial, and $|\mathsf{A} - \lambda\mathsf{I}| = 0$ is the characteristic equation of A.

Example For H above,

$$|\mathsf{H} - \lambda\mathsf{I}| = \begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = (2-\lambda)(\lambda^2 - 1) = 0, \tag{111}$$

so $\lambda = 2, \pm 1$. It may be shown that the associated eigenvectors are

$$\boldsymbol{v}_2 = \boldsymbol{e}_1, \qquad \boldsymbol{v}_1 = \boldsymbol{e}_2 + \boldsymbol{e}_3, \qquad \boldsymbol{v}_{-1} = \boldsymbol{e}_2 - \boldsymbol{e}_3. \tag{112}$$

We notice that the eigenvectors are mutually orthogonal.

Theorem IV.3 *Let A be a real symmetric 3×3 matrix with eigenvalues $\lambda_i \in \mathbb{R}$ and associated eigenvectors $\boldsymbol{v}_i \in \mathbb{R}^3$. Then:*

1. *if $\lambda_i \neq \lambda_j$, then $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$;*
2. *if $\lambda_i = \lambda_j$, we may choose $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$.*

Proof By writing $\boldsymbol{v}_i \cdot \boldsymbol{v}_j$ as $\boldsymbol{v}_i^T \boldsymbol{v}_j = \boldsymbol{v}_i \boldsymbol{v}_j^T$, then we have $\lambda_i \boldsymbol{v}_i \cdot \boldsymbol{v}_j = \lambda_i \boldsymbol{v}_j^T \boldsymbol{v}_i = \boldsymbol{v}_j^T \lambda_i \boldsymbol{v}_i = \boldsymbol{v}_j^T \mathsf{A}\boldsymbol{v}_i$. Taking the transpose does nothing to this because it is a scalar, so $\lambda_i \boldsymbol{v}_i \cdot \boldsymbol{v}_j = (\boldsymbol{v}_j^T \mathsf{A}\boldsymbol{v}_i)^T = \boldsymbol{v}_i^T \mathsf{A}^T \boldsymbol{v}_j$. Now, A is symmetric, so $\lambda_i \boldsymbol{v}_i \cdot \boldsymbol{v}_j = \boldsymbol{v}_i^T \mathsf{A}\boldsymbol{v}_j = \lambda_j \boldsymbol{v}_i^T \boldsymbol{v}_j = \lambda_j \boldsymbol{v}_i \cdot \boldsymbol{v}_j$. Thus we have $(\lambda_j - \lambda_i)\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$, and the first result follows.

Suppose now that $\lambda_i = \lambda_j$. Then $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ span a 2D subspace/plane in which all vectors of A have the same eigenvalue. If $\boldsymbol{v}_i \cdot \boldsymbol{v}_j \neq= 0$, then we may redefine $\boldsymbol{v}_i$ as

$$\boldsymbol{v}_i' = \boldsymbol{v}_j - \frac{(\boldsymbol{v}_i \cdot \boldsymbol{v}_j)\boldsymbol{v}_j}{|\boldsymbol{v}_j|^2}, \tag{113}$$

which is the projection of $\boldsymbol{v}_i$ onto the space/plane that is orthogonal to $\boldsymbol{v}_j$.

(The latter process of using projections to generate an orthogonal basis is know as Gram–Schmidt orthogonalisation.)

D. Quadric surfaces

A quadric is the generalisation of the conic to surfaces in $\mathbb{R}^3$. This has the form $\boldsymbol{r}^T\mathsf{A}\boldsymbol{r} + \boldsymbol{b}\cdot\boldsymbol{r} + c = 0$. A here is symmetric and of the form

$$\mathsf{A} = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}. \tag{114}$$

Example Suppose all cross terms are zero except for the leading diagonal in A, i.e., $ax^2 + dy^2 + fz^2 = 1$. Then we have the following cases:

1. An ellipsoid when $a, d, f > 0$. This may be parametrised as

$$x = \frac{1}{\sqrt{a}}\sin\theta\cos\phi, \qquad y = \frac{1}{\sqrt{d}}\sin\theta\sin\phi, \qquad z = \frac{1}{\sqrt{f}}\cos\phi. \tag{115}$$

2. A hyperboloid of one sheet when $a, d > 0$, $f < 0$. For fixed z, x and y are constrained to be on an ellipse (a circle if $a = d$). This may be parametrised as

$$x = \frac{1}{\sqrt{a}}\cosh\chi\cos\phi, \qquad y = \frac{1}{\sqrt{d}}\sinh\chi\sin\phi, \qquad z = \frac{1}{\sqrt{f}}\sinh\chi. \tag{116}$$

3. A hyperboloid of two sheets when $a, d < 0$, $f > 0$. There is no solution for a limited range of z. This may be parametrised as

$$x = \frac{1}{\sqrt{a}}\sinh\chi\cos\phi, \qquad y = \frac{1}{\sqrt{d}}\sinh\chi\sin\phi, \qquad z = \frac{1}{\sqrt{f}}\cosh\chi. \tag{117}$$

4. No solutions when $a, d, f < 0$.

Note that, generally, we have

$$f(x, y, z) = ax^2 + 2bxy + 2cxz + dy^2 + 2eyz + fz^2 = \text{constant}. \tag{118}$$

As with conics, it is useful to identify the nearest/furthest points from the origin. This will occur when the normal is pointing directly from the origin, i.e., $\boldsymbol{n} = \nabla f \propto \boldsymbol{r}$, or $\nabla f = 2\lambda\boldsymbol{r}$. Since $\nabla f = 2\mathsf{A}\boldsymbol{r}$, we have $\mathsf{A}\boldsymbol{r} = \lambda\boldsymbol{r}$. Thus eigenvectors of A gives the direction of extrema of the quadric surface from the origin, and these principal axes are orthogonal by the previous theorem (since A is symmetric).

Since any vector in $\mathbb{R}^3$ may be written as a sum of three orthogonal eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$, we have $\boldsymbol{r} = r_1\boldsymbol{v}_1 + r_2\boldsymbol{v}_2 + r_3\boldsymbol{v}_3$, so

$$\boldsymbol{r}^T\mathsf{A}\boldsymbol{r} = (r_1\boldsymbol{v}_1^T + r_2\boldsymbol{v}_2^T + r_3\boldsymbol{v}_3^T)\mathsf{A}(r_1\boldsymbol{v}_1 + r_2\boldsymbol{v}_2 + r_3\boldsymbol{v}_3) = r_1^2\lambda_1|\boldsymbol{v}_1|^2 + r_2^2\lambda_2|\boldsymbol{v}_2|^2 + r_3^2\lambda_3|\boldsymbol{v}_3|^2. \tag{119}$$

Choosing $\boldsymbol{v}_1$ to be unit vectors, we have $\boldsymbol{r}^T\mathsf{A}\boldsymbol{r} = r_1^2\lambda_1 + r_2^2\lambda_2 + r_3^2\lambda_3$. Since $\boldsymbol{v}_i$ are orthonormal, we can think of them as the new xyz-axis.

Example Classify the quadric $5x^2 - 6xy + 5y^2 + 9z^2 = 1$.

$$\mathsf{A} = \begin{pmatrix} 5 & -3 & 0 \\ -3 & 5 & 0 \\ 0 & 0 & 9 \end{pmatrix}, \qquad |\mathsf{A} - \lambda\mathsf{I}| = (9-\lambda)(8-\lambda)(2-\lambda) = 0. \tag{120}$$

All eigenvalues are positive, so we have an ellipsoid. It may be shown the principal axes are

$$\boldsymbol{v}_9 = \boldsymbol{e}_3, \qquad \boldsymbol{v}_8 = (\boldsymbol{e}_1 - \boldsymbol{e}_2)/\sqrt{2}, \qquad \boldsymbol{v}_2 = (\boldsymbol{e}_1 + \boldsymbol{e}_2)/\sqrt{2}. \tag{121}$$

V. LINEAR MAPS

A set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\} \in \mathbb{R}^3$ is a basis for $\mathbb{R}^3$ if any vector $\boldsymbol{u} \in \mathbb{R}^3$ may be written uniquely as $\boldsymbol{u} = u_1\boldsymbol{v}_1 + u_2\boldsymbol{v}_2 + u_3\boldsymbol{v}_3$. u_i here are the components of the vector $\boldsymbol{u}$ with respect to the basis $\{\boldsymbol{v}_i\}$.

Suppose $\boldsymbol{r}$ has co-ordinates x_1, x_2, x_3 with respect to $\{\boldsymbol{v}_i\}$. Let $\mathsf{P} = (\boldsymbol{v}_1|\boldsymbol{v}_2|\boldsymbol{v}_3)$, the matrix with $\boldsymbol{v}_i$ as columns. Then, by inspection, $\boldsymbol{v}_i = \mathsf{P}\boldsymbol{e}_i$, where $\boldsymbol{e}_i$ is the standard Cartesian basis; P is the linear map that takes us from the standard basis to the basis $\{\boldsymbol{v}_i\}$.

Theorem V.1 *Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ be a set of three vectors in $\mathbb{R}^3$, and $\mathsf{P} = (\boldsymbol{v}_1|\boldsymbol{v}_2|\boldsymbol{v}_3)$. S is a basis of $\mathbb{R}^3$ if $|\mathsf{P}| \neq 0$.*

Proof Suppose that $|\mathsf{P}| \neq 0$. Then there exists P^{-1}. For an arbitrary vector $\boldsymbol{x} \in \mathbb{R}^3$, $\boldsymbol{x} = \mathsf{P}\mathsf{P}^{-1}\boldsymbol{x} = \boldsymbol{y}$ is unique, with $y = \mathsf{P}^{-1}\boldsymbol{x}$. Now,

$$\boldsymbol{x} = \mathsf{P}\boldsymbol{y} = y_1\mathsf{P}\boldsymbol{e}_1 + y_2\mathsf{P}\boldsymbol{e}_2 + y_3\mathsf{P}\boldsymbol{e}_3 = y_1\boldsymbol{v}_1 + y_2\boldsymbol{v}_2 + y_3\boldsymbol{v}_3, \tag{122}$$

so S is a basis.

Suppose now that S is a basis. Then, in particular, the standard basis has the representation

$$\boldsymbol{e}_i = \sum_{j=1}^{3} e_i^j \boldsymbol{v}_j. \tag{123}$$

The scalar triple product of the standard basis is 1, so

$$1 = [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] = \sum_{i,j,k,=1}^{3} e_1^i e_2^j e_3^k [\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k] \quad \Rightarrow \quad [\boldsymbol{v}_i, \boldsymbol{v}_j, \boldsymbol{v}_k] = |\mathsf{P}| = 1. \tag{124}$$

Thus $|\mathsf{P} \neq 0$ if $i \neq j \neq k$.

Example Which of

$$B_1 = \{(4,1,2),(2,5,1),(0,1,0)\}, \qquad B_2 = \{(4,1,2),(2,5,1),(0,0,1)\} \tag{125}$$

forms a basis in $\mathbb{R}^3$?

$\mathsf{P}_1 = 0$ whilst $\mathsf{P}_2 = 18$, so only B_2 is a basis.

A. Axioms

A linear map $L : \mathbb{R}^3 \to \mathbb{R}^3$ is a function such that, for all $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$ and $\lambda, \mu \in \mathbb{R}$, $L(\lambda\boldsymbol{u} + \mu\boldsymbol{v}) = \lambda L(\boldsymbol{u}) + \mu L(\boldsymbol{v})$. Note then $L(\boldsymbol{0}) = \boldsymbol{0}$.

Example $L(\boldsymbol{u}) = (3u_1 - u_2 + u_3, u_3, 0)$ is linear.

$L(\boldsymbol{x}) = (xy, x, z)$ is not linear since the first component is not a linear function.

$L(\boldsymbol{x}) = (x+1, y, z)$ is not linear since it is not a homogeneous function of degree 1.

$\mathsf{P} = (\boldsymbol{v}_1|\boldsymbol{v}_2|\boldsymbol{v}_3)$ induces a linear map by definition.

Theorem V.2 *A map $L : \mathbb{R}^3 \to \mathbb{R}^3$ is linear iff there exists a 3×3 matrix A where $L(\boldsymbol{v}) = \mathsf{A}\boldsymbol{v}$ for all $\boldsymbol{v} \in \mathbb{R}^3$. Moreover, $\mathsf{A} = (L(\boldsymbol{e}_1)|L(\boldsymbol{e}_2)|L(\boldsymbol{e}_3))$.*

Proof Assuming L is a linear map, then by the linearity property, $L(\boldsymbol{v}) = v_1L(\boldsymbol{e}_1) + v_2L(\boldsymbol{e}_2) + v_3L(\boldsymbol{e}_3) = (L(\boldsymbol{e}_1)|L(\boldsymbol{e}_2)|L(\boldsymbol{e}_3))\boldsymbol{v}$ as required. Assuming we have $L(\boldsymbol{v}) = \mathsf{A}\boldsymbol{v}$, by properties of matrices, A automatically induces a linear map.

Example

$$L\boldsymbol{x} = \begin{pmatrix} 2x - y + z \\ z \\ 0 \end{pmatrix}, \qquad \mathsf{A} = \begin{pmatrix} 3 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{126}$$

$$L\boldsymbol{x} = \begin{pmatrix} 6x - 2y + 5z \\ 2x + 3y - z \\ x + 5y + 2z \end{pmatrix}, \qquad \mathsf{A} = \begin{pmatrix} 6 & -2 & 5 \\ 2 & 3 & -1 \\ 1 & 5 & 2 \end{pmatrix} \tag{127}$$

Suppose L_1 and L_2 are linear maps associated with A_1 and A_2 respectively. Then a combined linear map $L_1 \circ L_2(\boldsymbol{v})$ may be represented by $\mathsf{B} = \mathsf{A}_1\mathsf{A}_2$.

Lemma V.3 *Suppose $L : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear map represented by A. If $|\mathsf{A}| \neq 0$, then there exists a map L^{-1} such that $L^{-1} \circ L(\boldsymbol{v}) = L \circ L^{-1}(\boldsymbol{v}) = \boldsymbol{v}$.*

Proof This follows immediately from the fact that A^{-1} exists when $|\mathsf{A}| \neq 0$.

B. Geometry of linear maps

There are various linear maps with geometric interpretations. For example, a reflection in the yz-plane is represented by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} -x \\ y \\ z \end{pmatrix}, \qquad \mathsf{A} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{128}$$

A projection onto the xy plane is represented by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}, \qquad \mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{129}$$

A rotation of θ radians in the xy-plane is

$$\mathsf{A} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{130}$$

1. *Projection*

Let Π be a plane through $\mathbf{0}$ with normal vector $\boldsymbol{n}$. For $\boldsymbol{x} \in \mathbb{R}^3$,

$$\boldsymbol{x} = \frac{\boldsymbol{x}\cdot\boldsymbol{n}}{|\boldsymbol{n}|^2}\boldsymbol{n} + \left(\boldsymbol{x} - \frac{\boldsymbol{x}\cdot\boldsymbol{n}}{|\boldsymbol{n}|^2}\boldsymbol{n}\right). \tag{131}$$

We note that the first portion is normal to Π, whilst the second part is orthogonal to $\boldsymbol{n}$. Since the equation of Π is $\boldsymbol{r}\cdot\boldsymbol{n} = 0$, the projection of $\boldsymbol{x}$ onto Π is given by

$$P(\boldsymbol{x}) = \boldsymbol{x} - \frac{\boldsymbol{x}\cdot\boldsymbol{n}}{|\boldsymbol{n}|^2}\boldsymbol{n} = \boldsymbol{x} - \boldsymbol{x}\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}}. \tag{132}$$

This is a linear map because

$$P(\lambda\boldsymbol{u} + \boldsymbol{v}) = \lambda\boldsymbol{u} + \boldsymbol{v} - (\lambda\boldsymbol{u} + \boldsymbol{v})\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}} = \lambda(\boldsymbol{u} - \boldsymbol{u}\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}}) + \boldsymbol{v} - \boldsymbol{v}\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}}. \tag{133}$$

2. *Reflection*

Similarly, if $\boldsymbol{x}$ is the sum of $P(\boldsymbol{x})$ and normal to Π, then to reflect $\boldsymbol{x}$ in Π, we subtract twice, so

$$R(\boldsymbol{x}) = \boldsymbol{x} - 2\boldsymbol{x}\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}}. \tag{134}$$

3. Rotation

In $\mathbb{R}^3$, a rotation always leaves a line/axis invariant. Let $\boldsymbol{l}$ be a line through $\mathbf{0}$ with direction vector $\boldsymbol{d}$, then $\boldsymbol{l} = \lambda\boldsymbol{d}$, so

$$\boldsymbol{x} = (\boldsymbol{x} - \boldsymbol{x}\cdot\hat{\boldsymbol{d}}\hat{\boldsymbol{d}}) + \boldsymbol{x}\cdot\hat{\boldsymbol{d}}\hat{\boldsymbol{d}}. \tag{135}$$

The first part is normal to $\boldsymbol{l}$ by construction, so a vector orthogonal to $\boldsymbol{l}$ would be $(\boldsymbol{x}\cdot\hat{\boldsymbol{d}}\hat{\boldsymbol{d}})\times\hat{\boldsymbol{d}} = \boldsymbol{x}\times\hat{\boldsymbol{d}}$. By analogy with rotation around a basis vector, we have

$$R(\boldsymbol{x}) = \boldsymbol{x}\cdot\hat{\boldsymbol{d}}\hat{\boldsymbol{d}} + (\boldsymbol{x} - \boldsymbol{x}\cdot\hat{\boldsymbol{d}}\hat{\boldsymbol{d}})\cos\theta - (\boldsymbol{x}\times\hat{\boldsymbol{d}})\sin\theta. \tag{136}$$

Example Let Π be a plane through $\mathbf{0}$ and $\boldsymbol{n} = (1,1,2)$ Find the matrix of projection onto Π with respect to the standard basis.
$\boldsymbol{n}\cdot\boldsymbol{x} = x+y+2z$, $|\boldsymbol{n}|^2 = 6$, so

$$P(\boldsymbol{x}) = \begin{pmatrix} x-(x+y+2z)/6 \\ y-(x+y+2z)/6 \\ z-(x+y+2z)/6 \end{pmatrix}, \qquad \mathsf{A} = \frac{1}{6}\begin{pmatrix} 5 & -1 & -2 \\ -1 & 5 & -2 \\ -2 & -2 & 2 \end{pmatrix}. \tag{137}$$

Example Let Π be a plane through $\mathbf{0}$ with $\boldsymbol{n} = (1,-1,1)$. Find the matrix of reflection in Π with respect to the standard basis.
$\boldsymbol{n}\cdot\boldsymbol{x} = x-y+z$, $|\boldsymbol{n}|^2 = 3$, so

$$R(\boldsymbol{x}) = \begin{pmatrix} x-(x-y+z)(2/3) \\ y+(x-y+z)(2/3) \\ z-(x-y+z)(2/3) \end{pmatrix}, \qquad \mathsf{A} = \frac{1}{3}\begin{pmatrix} 1 & -2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix}. \tag{138}$$

Example Find the matrix of rotation representing a $-\pi/3$ rotation about $\boldsymbol{l} = \lambda(1,1,1)$.
$\boldsymbol{x}\cdot\boldsymbol{d} = x+y+z$, $|\boldsymbol{d}|^2 = 3$, and

$$\boldsymbol{x} - \boldsymbol{x}\cdot\hat{\boldsymbol{d}}\hat{\boldsymbol{d}} = \frac{1}{3}\begin{pmatrix} 2x-y-z \\ -x+2y-z \\ -x-y+2z \end{pmatrix}, \qquad \boldsymbol{x}\times\hat{\boldsymbol{d}} = \frac{1}{\sqrt{3}}\begin{pmatrix} y-z \\ z-x \\ x-y \end{pmatrix}. \tag{139}$$

Collecting this,

$$R_{-\pi/3}(\boldsymbol{x}) = \frac{1}{3}\begin{pmatrix} x+y+z \\ x+y+z \\ x+y+z \end{pmatrix} + \frac{1}{2}\frac{1}{3}\begin{pmatrix} 2x-y-z \\ -x+2y-z \\ -x-y+2z \end{pmatrix} + \frac{\sqrt{3}}{2}\frac{1}{\sqrt{3}}\begin{pmatrix} y-z \\ z-x \\ x-y \end{pmatrix}, \tag{140}$$

and

$$\mathsf{A} = \frac{1}{6}\begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 2 \\ 2 & -1 & 1 \end{pmatrix}. \tag{141}$$

Example

$$R(R(\boldsymbol{x})) = R(\boldsymbol{x} - 2\boldsymbol{x}\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}}) = R(\boldsymbol{x}) - 2R(\boldsymbol{x})\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}} = \boldsymbol{x} - 2\boldsymbol{x}\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}} - 2(\boldsymbol{x} - 2\boldsymbol{x}\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}})\cdot\hat{\boldsymbol{n}}\hat{\boldsymbol{n}} = \boldsymbol{x}. \tag{142}$$

C. Change of basis

Theorem V.4 *Let L be a linear map given by A with respect to $\{\boldsymbol{e}_i\}$. Let $S = \{\boldsymbol{v}_i\}$ be another basis of $\mathbb{R}^3$, and $\mathsf{P} = (\boldsymbol{v}_1|\boldsymbol{v}_2|\boldsymbol{v}_3)$. Then, with respect to S, $L = \mathsf{P}^{-1}\mathsf{A}\mathsf{P}$.*

Proof From a previous theorem, we have $\boldsymbol{x} = \mathsf{P}\boldsymbol{y}$, where $\boldsymbol{y}$ is expanded in terms of S. Then $L\boldsymbol{y} = L(\mathsf{P}^{-1}\boldsymbol{x})$ for $|\mathsf{P}| \neq 0$. Since L is linear, $L\boldsymbol{y} = \mathsf{P}^{-1}L(\boldsymbol{x}) = \mathsf{P}^{-1}\mathsf{A}\mathsf{P}\mathsf{P}^{-1}\boldsymbol{x} = \mathsf{P}^{-1}\mathsf{A}\mathsf{P}(\boldsymbol{y})$, as required.

Example Suppose L is represented by

$$\mathsf{A} = \begin{pmatrix} 1 & 3 & 3 \\ -2 & 3 & -2 \\ 3 & 2 & 1 \end{pmatrix} \tag{143}$$

with respect to the standard basis. Find A with respect to $S = \{(1,0,0), (-1,1,0), (0,-1,1)\}$.

$$\mathsf{P} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \mathsf{P}^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \mathsf{P}^{-1}\mathsf{AP} = \begin{pmatrix} 2 & 5 & -5 \\ 1 & 4 & -6 \\ 3 & -1 & -1 \end{pmatrix}. \tag{144}$$

Example Find the matrix of linear map L with respect to the basis S, where

$$L\boldsymbol{x} = \begin{pmatrix} 4x - z \\ 2x + 3y - 2z \\ 2x + 2y - z \end{pmatrix}, \qquad S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}. \tag{145}$$

$$\mathsf{A} = \begin{pmatrix} 4 & 0 & -1 \\ 2 & 3 & -2 \\ 2 & 2 & -1 \end{pmatrix}, \qquad \mathsf{P} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}, \qquad \mathsf{P}^{-1} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}, \tag{146}$$

so

$$\mathsf{P}^{-1}\mathsf{AP} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}. \tag{147}$$

Finally, we return to rotations. Since $(R_\theta)^{-1} = R_{-\theta}$, we have

$$\mathsf{A} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \mathsf{A}^{-1} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathsf{A}^T. \tag{148}$$

Now, the scalar product is preserved under rotation (since we are rotating something that only cares about the magnitude). Any transformation that has the property that $\mathsf{A}^T = \mathsf{A}^{-1}$ is called <u>orthogonal</u>, and the set of orthogonal transformations form the <u>group</u> $O(3)$, which includes rotations and reflections. It we restrict $O(3)$ to linear maps where the associated matrix has $|\mathsf{A}| > 0$, then we only have rotations, and the subset is in fact a subgroup called the <u>special orthogonal</u> group, denoted $SO(3)$.