STEP Support Programme

STEP 3 Matrices Solutions

Note that several of these solutions contain discussion as well as just the bare solutions, so they are often longer than would be expected in an exam solution.

Question 1

> Express the determinant
> $$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix}$$
> as the product of factors which are linear in a, b, c.

This determinant has a lot of symmetry: each column is identical, except for the variable. We therefore use row and column operations to simplify the determinant and to extract factors. (See the notes on matrices for a discussion of row and column operations.) We could also solve this question by expanding the determinant and extracting factors, but this approach – though more sophisticated – captures the structure of the situation.

There is nothing wrong with expanding the determinant. You should obtain something of the form $abc(bc^2 - b^2c - ac^2 + a^2c + ab^2 - ba^2)$. Factorising further is a bit tricky, but if you substitute $a = b$ into your expression for the determinant you should find that this makes the determinant equal to 0. Hence you should be able to extract a factor of $(a - b)$. This is a little tricky, but can be done by long division, or by using a table to help you (such as below):

	ac	$-c^2$	bc	$-ab$
a	a^2c	$-ac^2$	$\cancel{abc}$	$-a^2b$
$-b$	$\cancel{-abc}$	bc^2	$-b^2c$	ab^2

We note first that column 1 is a multiple of a, column 2 is a multiple of b and column 3 is a multiple of c, so we take these out as factors to begin with:

$$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix} = abc \begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix}.$$

We now note that if $a = b$, then the first two columns are identical, so the determinant is zero in this case. It follows from the factor theorem that $(a - b)$ is a factor of the determinant. (The determinant is a polynomial in a, b and c. Now think of it as a polynomial in a, treating b and c as constants. When $a = b$, the polynomial equals zero, and so $a - b$ is a factor.)

Likewise, if $c = a$ the first and third columns are identical, and if $b = c$ the second and third columns are identical. Therefore $(c - a)$ and $(b - c)$ are also factors of the determinant. But every term in the expansion of the simplified determinant consists of 1 times a, b or c, times a^2, b^2 or c^2, so has total degree 3 (the sum of the powers of all variables appearing in the term). As $(a-b)(b-c)(c-a)$ also has total degree 3, the determinant must be a scalar multiple of this. So

$$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix} = kabc(a - b)(b - c)(c - a)$$

for some k. Looking at the leading diagonal, we see that one of the terms in the determinant is ab^2c^3, and this term appears in the expansion of $kabc(a-b)(b-c)(c-a)$ as kab^2c^3 (we need a b and two c's in the expansion of the brackets, so we get $(-b)(-c)(c) = bc^2$ from that part). Therefore $k = 1$, and the determinant equals $abc(a - b)(b - c)(c - a)$.

Hence, or otherwise, find $x:y:z:u$ if

$$\begin{aligned} x+2\ y+3\ z+4\ u&=0,\\ x+2^2y+3^2z+4^2u&=0,\\ x+2^3y+3^3z+4^3u&=0. \end{aligned}$$

We can write these equations as three simultaneous equations in x, y and z if we regard u as a constant:

$$\begin{aligned} x+2\ y+3\ z&=-4\ u,\\ x+2^2y+3^2z&=-4^2u,\\ x+2^3y+3^3z&=-4^3u. \end{aligned}$$

We also note that the coefficients of x, y, z and u exactly match the entries in the determinant we have just calculated.

We therefore have

$$\det \mathbf{A} = \begin{vmatrix} 1 & 2 & 3 \\ 1^2 & 2^2 & 3^2 \\ 1^3 & 2^3 & 3^3 \end{vmatrix} = 1.2.3.(-1).(-1).2 = 12,$$

and the adjugate matrix is:

$$\begin{pmatrix} 2^2\times 3^3-2^3\times 3^2 & 2^3\times 3-2\times 3^3 & 2\times 3^2-2^2\times 3 \\ 3^2-3^3 & 3^3-3 & 3-3^2 \\ 2^3-2^2 & 2-2^3 & 2^2-2 \end{pmatrix} = \begin{pmatrix} 36 & -30 & 6 \\ -18 & 24 & -6 \\ 4 & -6 & 2 \end{pmatrix}$$

We then have:

$$\begin{aligned} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \frac{1}{12}\begin{pmatrix} 36 & -30 & 6 \\ -18 & 24 & -6 \\ 4 & -6 & 2 \end{pmatrix} \times \begin{pmatrix} -4u \\ -16u \\ -64u \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} 36 & -30 & 6 \\ -18 & 24 & -6 \\ 4 & -6 & 2 \end{pmatrix} \times \begin{pmatrix} -u \\ -4u \\ -16u \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} -36+4\times 30-6\times 16 \\ 18-4\times 24+6\times 16 \\ -4+4\times 6-2\times 16 \end{pmatrix} u \\ &= \frac{1}{3}\begin{pmatrix} 6(-6+4\times 5-16) \\ 6(3-4\times 4+16) \\ 4(-1+6-2\times 4) \end{pmatrix} u \\ &= \frac{1}{3}\begin{pmatrix} 6\times -2 \\ 6\times 3 \\ 4\times -3 \end{pmatrix} u \\ &= \begin{pmatrix} -4u \\ 6u \\ -4u \end{pmatrix} \end{aligned}$$

Thus the ratio $x : y : z : u$ is

$$-4u : 6u : -4u : u = -4 : 6 : -4 : 1 = 4 : -6 : 4 : -1.$$

It is certainly worth us checking our solution: if we take $x = 4$, $y = -6$, $z = 4$ and $u = -1$, our equations become

$$\begin{aligned} 4 + \ 2(-6) + \ 3.4 &= \ -4(-1), \\ 4 + 2^2(-6) + 3^2.4 &= -4^2(-1), \\ 4 + 2^3(-6) + 3^3.4 &= -4^3(-1), \end{aligned}$$

which are all true.

Question 2

Prove that $(a-b)$ and $(x-y)$ are factors of the determinant

$$\begin{vmatrix} (a+x)^2 & (a+y)^2 & (a+z)^2 \\ (b+x)^2 & (b+y)^2 & (b+z)^2 \\ (c+x)^2 & (c+y)^2 & (c+z)^2 \end{vmatrix}$$

and factorise the determinant completely.

This determinant has a lot of symmetry, which suggests that it will have simple factors. If we were to expand this determinant, we would get a polynomial in a, b, c, x, y and z. Every term in the polynomial would have total degree 6 (that is the sum of the powers of every variable appearing in the polynomial). We can also, if we wish, think of the determinant as a polynomial in just a single variable, with coefficients being expressions in the other five variables.

If we let $b=a$, then the first two rows are identical and hence the determinant is zero. So if we think of the determinant as a polynomial in a, we find that $(a-b)$ is a factor of the determinant, using the factor theorem. Likewise, if we let $c=b$ or $a=c$, we again get zero, so $(b-c)$ and $(c-a)$ are also factors. In a similar fashion, if we let $x=y$, the first two columns become identical, so $(x-y)$ is a factor; $(y-z)$ and $(z-x)$ are factors in the same way.

Therefore, the determinant is a multiple of $(a-b)(b-c)(c-a)(x-y)(y-z)(z-x)$. This polynomial has total degree 6, and so the determinant is just a scalar multiple of this expression.

We can find the scalar multiple either by choosing specific values for the variables and evaluating the determinant and this expression, or by considering a specific term and calculating its coefficient in the expansion of the determinant. We will use the second method for illustrative purposes.

If we consider $a^2b^2c^2$, obtained by multiplying the expressions on the leading diagonal, we see that it does not appear in the expansion of $(a-b)(b-c)(c-a)(x-y)(y-z)(z-x)$. This will therefore not help us. Instead, we can consider some term appearing in the expansion of this expression, say a^2bx^2y, which appears with coefficient 1 here. In the expansion of the determinant, as we need a^2x^2, we cannot use the top-left term (as that would give us a total power of 2 for a and x). We must therefore obtain a^2 from $(a+y)^2$ or $(a+z)^2$ and x^2 from $(b+x)^2$ or $(c+x)^2$; we therefore only need to consider the following simplified determinant to identify these terms:

$$\begin{vmatrix} 0 & a^2 & a^2 \\ x^2 & (b+y)^2 & (b+z)^2 \\ x^2 & (c+y)^2 & (c+z)^2 \end{vmatrix} = a^2x^2 \begin{vmatrix} 0 & 1 & 1 \\ 1 & (b+y)^2 & (b+z)^2 \\ 1 & (c+y)^2 & (c+z)^2 \end{vmatrix}$$

Now, we want terms of the form a^2bx^2y, so the determinant part of this final expression must give terms of the form by. There is only one expression capable of doing this, which is $(b+y)^2 = b^2+2by+y^2$, so the coefficient of by is -2, obtained from the product of the three highlighted terms here (paying attention to the sign of this term in the determinant):

$$\begin{vmatrix} 0 & 1 & \boxed{1} \\ 1 & \boxed{(b+y)^2} & (b+z)^2 \\ \boxed{1} & (c+y)^2 & (c+z)^2 \end{vmatrix}$$

Therefore the original determinant factorises as

$$-2(a-b)(b-c)(c-a)(x-y)(y-z)(z-x).$$

Question 3

Given that α, β, γ are the roots of the equation

$$\begin{vmatrix} a & b & x \\ x & c & a \\ c & x & b \end{vmatrix} = 0,$$

prove that

$$\alpha^3 + \beta^3 + \gamma^3 = -6abc.$$

One thing that is striking about this question is that it is connecting the roots of a cubic to a determinant. The determinant does not seem particularly helpful, so we start by evaluating the determinant; this gives

$$x^3 - (a^2 + b^2 + c^2)x + 2abc = 0. \tag{1}$$

Perhaps surprisingly, this is symmetric in a, b and c, even though the form of the determinant does not seem to suggest this.

We could now use what we know about roots of polynomials to solve this problem. We can write this polynomial as

$$(x-\alpha)(x-\beta)(x-\gamma) = x^3 - (\alpha+\beta+\gamma)x^2 + (\alpha\beta+\beta\gamma+\gamma\alpha)x - \alpha\beta\gamma = 0,$$

so comparing the equations gives

$$\begin{aligned} \alpha+\beta+\gamma &= 0 \\ \alpha\beta+\beta\gamma+\gamma\alpha &= -(a^2+b^2+c^2) \\ \alpha\beta\gamma &= -2abc. \end{aligned} \tag{2}$$

We can now calculate

$$(\alpha+\beta+\gamma)^3 = \alpha^3+\beta^3+\gamma^3 + 3(\alpha^2\beta+\beta^2\gamma+\gamma^2\alpha+\alpha^2\gamma+\beta^2\alpha+\gamma^2\beta) + 6\alpha\beta\gamma \tag{3}$$

and

$$(\alpha+\beta+\gamma)(\alpha\beta+\beta\gamma+\gamma\alpha) = (\alpha^2\beta+\beta^2\gamma+\gamma^2\alpha+\alpha^2\gamma+\beta^2\alpha+\gamma^2\beta) + 3\alpha\beta\gamma. \tag{4}$$

We can now substitute our known values for $\alpha+\beta+\gamma$ from equations (2) into equation (4), giving

$$0(-(a^2+b^2+c^2)) = (\alpha^2\beta+\beta^2\gamma+\gamma^2\alpha+\alpha^2\gamma+\beta^2\alpha+\gamma^2\beta) - 6abc,$$

so that

$$\alpha^2\beta+\beta^2\gamma+\gamma^2\alpha+\alpha^2\gamma+\beta^2\alpha+\gamma^2\beta = 6abc.$$

Substituting this result into equation (3) gives

$$0^3 = \alpha^3+\beta^3+\gamma^3 + 3(6abc) + 6(-2abc)$$

and hence

$$\alpha^3+\beta^3+\gamma^3 = -6abc.$$

Another thing we could do, which would be significantly simpler, is to substitute α, β and γ into our original polynomial equation (1). Since these are roots of the polynomial, they satisfy the equation. This gives us three equations:

$$\begin{aligned} \alpha^3 - (a^2+b^2+c^2)\alpha + 2abc &= 0 \\ \beta^3 - (a^2+b^2+c^2)\beta + 2abc &= 0 \\ \gamma^3 - (a^2+b^2+c^2)\gamma + 2abc &= 0 \end{aligned} \tag{5}$$

Adding these together gives

$$(\alpha^3+\beta^3+\gamma^3) - (a^2+b^2+c^2)(\alpha+\beta+\gamma) + 6abc = 0.$$

Now using $\alpha+\beta+\gamma = 0$ from equations (2) gives

$$\alpha^3+\beta^3+\gamma^3 = -6abc.$$

Incidentally, if we first multiplied the equations in (5) by α, β, γ respectively, we could find an expression for $\alpha^4+\beta^4+\gamma^4$ in a similar way (though we would have to first calculate $\alpha^2+\beta^2+\gamma^2 = (\alpha+\beta+\gamma)^2 - 2(\alpha\beta+\beta\gamma+\gamma\alpha)$). By an inductive process, we would likewise be able to find $\alpha^n+\beta^n+\gamma^n$ for any $n \geqslant 3$ in this way.

Question 4

Prove that if the equations

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= 0 \\ a_2x + b_2y + c_2z &= 0 \\ a_3x + b_3y + c_3z &= 0 \end{aligned}\right\}$$

are simultaneously satisfied by values of x, y, z which are not all zero, then

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

We can write the simultaneous equations as $\mathbf{Ax} = \mathbf{0}$, where $\mathbf{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. If $\det \mathbf{A} \neq 0$, then $\mathbf{A}$ is invertible and $\mathbf{x} = \mathbf{A}^{-1}\mathbf{0} = \mathbf{0}$. So if the equations have a solution with $\mathbf{x} \neq \mathbf{0}$, then $\mathbf{A}$ cannot be invertible, and so $\det \mathbf{A} = 0$, as required.

Hence, or otherwise, eliminate x, y, z from the equations

$$a = \frac{x}{y-z}, \quad b = \frac{y}{z-x}, \quad c = \frac{z}{x-y}.$$

[Your answer should **not** be left in determinant form.]

It is not clear how this is related to the previous part of the question, but we can at least express these equations as linear equations in x, y and z. The first equation, on multiplying by $y - z$, gives $a(y - z) = x$, so $x - ay + az = 0$. Doing the same to the other two gives

$$\begin{aligned} x - ay + az &= 0 \\ bx + y - bz &= 0 \\ -cx + cy + z &= 0 \end{aligned}$$

Returning now to our original three formulae for a, b and c, we observe that if we choose distinct values of x, y and z, then this gives us values of a, b and c. For these values of a, b and c, we should be able to obtain our original chosen values of x, y and z from these three simultaneous equations, at least up to a constant factor. (Note that if we replace x, y and z by λx, λy and λz, we obtain the same values of a, b and c.) Therefore the determinant of the coefficients of these three simultaneous equations must be zero, from the previous part.

Thus

$$\begin{vmatrix} 1 & -a & a \\ b & 1 & -b \\ -c & c & 1 \end{vmatrix} = 0,$$

which expands to give

$$ab + bc + ca + 1 = 0,$$

and we have successfully eliminated x, y and z.

Question 5

A matrix $\mathbf{M}$ is said to be transposed into the matrix $\mathbf{M}^{\mathrm{T}}$ if the first row of $\mathbf{M}$ becomes the first column of $\mathbf{M}^{\mathrm{T}}$, the second row of $\mathbf{M}$ becomes the second column of $\mathbf{M}^{\mathrm{T}}$, and so on. Write down the transposes of the matrices

$$\mathbf{M} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \mathbf{T} = \begin{pmatrix} 0 & b & 0 \\ 0 & 0 & c \\ a & 0 & 0 \end{pmatrix}.$$

Calculate the matrix products $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ and $\mathbf{TM}$; show also that $(\mathbf{TM})^{\mathrm{T}} = \mathbf{M}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}$.

The transposes are

$$\mathbf{M}^{\mathrm{T}} = \begin{pmatrix} x & y & z \end{pmatrix}, \qquad \mathbf{T}^{\mathrm{T}} = \begin{pmatrix} 0 & 0 & a \\ b & 0 & 0 \\ 0 & c & 0 \end{pmatrix}.$$

Then

$$\mathbf{M}^{\mathrm{T}}\mathbf{M} = \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^2 + y^2 + z^2 \end{pmatrix},$$

which is a 1×1 matrix or 1-dimensional vector; we could also conveniently think of this as the scalar $x^2 + y^2 + z^2$ (though the next part of the question does talk about "the element of $\mathbf{M}^{\mathrm{T}}\mathbf{M}$").

Also,

$$\mathbf{TM} = \begin{pmatrix} 0 & b & 0 \\ 0 & 0 & c \\ a & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} by \\ cz \\ ax \end{pmatrix}.$$

We can then check that $(\mathbf{TM})^{\mathrm{T}} = \mathbf{M}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}$ by calculating:

$$\begin{aligned} (\mathbf{TM})^{\mathrm{T}} &= \begin{pmatrix} by & cz & ax \end{pmatrix}; \\ \mathbf{M}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}} &= \begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 0 & 0 & a \\ b & 0 & 0 \\ 0 & c & 0 \end{pmatrix} \\ &= \begin{pmatrix} yb & zc & xa \end{pmatrix}, \end{aligned}$$

so they are equal as required.

This is an example of the general rule that for any conformable matrices $\mathbf{A}$ and $\mathbf{B}$, $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$. A proof of this appears in Question 15.

> If the elements of $\mathbf{M}$ are the Cartesian coordinates of a point P, what information is provided by the element of $\mathbf{M}^{\mathrm{T}}\mathbf{M}$?

The single element of $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ is $x^2 + y^2 + z^2$, which is the square of the distance of P from the origin.

> If the matrix $\mathbf{T}$ describes a transformation of the points P of three dimensional space, interpret geometrically the equation
>
> $$(\mathbf{TM})^{\mathrm{T}}(\mathbf{TM}) = \mathbf{M}^{\mathrm{T}}\mathbf{M},$$
>
> and find all appropriate values of a, b and c for which this equation holds for all points P.

$\mathbf{TM}$ is the image of the point P, so by the previous part, $(\mathbf{TM})^{\mathrm{T}}(\mathbf{TM})$ is the square of the distance of the image of P from the origin. Therefore the equation means that the square of the distance of the point P from the origin equals the square of the distance of the image of P from the origin. Since distances are never negative, we can take the square root, so the equation means that the point P and its image are the same distance from the origin.

Algebraically, if the equation holds for all points P, we need

$$(by)^2 + (cz)^2 + (ax)^2 = x^2 + y^2 + z^2$$

for all x, y and z.

Since this has to be true for *all* values of x, y and z, we will help ourselves by choosing really convenient values to start with.

Taking $x = 1$, $y = z = 0$, we get $a^2 = 1$, so we must have $a = \pm 1$. Similarly, we find $b = \pm 1$ and $c = \pm 1$ (where the signs of a, b and c are independent of each other). In this case,

$$(by)^2 + (cz)^2 + (ax)^2 = b^2y^2 + c^2z^2 + a^2x^2 = y^2 + z^2 + x^2$$

so for these values of a, b and c, the equation holds for all points P.

Note that it was not enough to show that $a = \pm 1$, $b = \pm 1$ and $c = \pm 1$ and to stop there: all we have shown at that point is that this is *necessary* for the equation to hold for all points P, but it may not be *sufficient*. To show that this is sufficient, we must also show that when a, b and c take these values, the equation is satisfied for all points P.

Question 6

The functions $t \to \mathbf{P}$, $t \to \mathbf{Q}$ map real numbers t onto matrices $\mathbf{P}$, $\mathbf{Q}$ of fixed dimension; that is, each element of each matrix is a real function of t, and we suppose all the functions to be differentiable. A scalar multiplier s is also a function of t. The derivative of a matrix is defined as a matrix whose elements are the derivatives of the elements of the original matrix. We write $\mathrm{d}\mathbf{P}/\mathrm{d}t$ as $\dot{\mathbf{P}}$, $\mathrm{d}s/\mathrm{d}t$ as $\dot{s}$, and so on.

Let us write p_{ij} for the (i, j)th element of $\mathbf{P}$, and likewise q_{ij} for the (i, j)th element of $\mathbf{Q}$. So

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Then we can differentiate this with respect to t to get

$$\dot{\mathbf{P}} = \begin{pmatrix} \dot{p}_{11} & \dot{p}_{12} & \cdots \\ \dot{p}_{21} & \dot{p}_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix},$$

using the definition of the derivative of a matrix we have been given, so the (i, j)th element of $\dot{\mathbf{P}}$ is $\dot{p}_{ij}$.

(i) Prove that

$$\frac{\mathrm{d}}{\mathrm{d}t}(s\mathbf{P}) = \dot{s}\mathbf{P} + s\dot{\mathbf{P}}.$$

As s is a scalar, we have

$$s\mathbf{P} = \begin{pmatrix} sp_{11} & sp_{12} & \cdots \\ sp_{21} & sp_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Differentiating this matrix with respect to t, by differentiating each element, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}(s\mathbf{P}) = \begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t}(sp_{11}) & \frac{\mathrm{d}}{\mathrm{d}t}(sp_{12}) & \cdots \\ \frac{\mathrm{d}}{\mathrm{d}t}(sp_{21}) & \frac{\mathrm{d}}{\mathrm{d}t}(sp_{22}) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

More efficiently, we could say instead: The (i, j)th element of $s\mathbf{P}$ is sp_{ij}, so the (i, j)th element of $\frac{\mathrm{d}}{\mathrm{d}t}(s\mathbf{P})$ is $\frac{\mathrm{d}}{\mathrm{d}t}(sp_{ij})$.

Now using the product rule, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(sp_{ij}) = \frac{\mathrm{d}}{\mathrm{d}t}(s)p_{ij} + s\frac{\mathrm{d}}{\mathrm{d}t}(p_{ij}) = \dot{s}p_{ij} + s\dot{p}_{ij},$$

which is exactly the (i,j)th element of $\dot{s}\mathbf{P} + s\dot{\mathbf{P}}$. Therefore

$$\frac{\mathrm{d}}{\mathrm{d}t}(s\mathbf{P}) = \dot{s}\mathbf{P} + s\dot{\mathbf{P}}.$$

> (ii) If the product $\mathbf{PQ}$ is defined, prove that
>
> $$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{PQ}) = \dot{\mathbf{P}}\mathbf{Q} + \mathbf{P}\dot{\mathbf{Q}}.$$

Let us write $\mathbf{R} = \mathbf{PQ}$. To prove this, we need an expression for the (i,j)th element of $\mathbf{R}$, so that we can differentiate it. This element is obtained by multiplying the ith row of $\mathbf{P}$ by the jth column of $\mathbf{Q}$:

$$\mathbf{R} = \begin{pmatrix} \ddots & \vdots & \iddots \\ \cdots & r_{ij} & \cdots \\ \iddots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & \vdots & \iddots \\ p_{i1} & p_{i2} & p_{i3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \cdots & q_{1j} & \cdots \\ \cdots & q_{2j} & \cdots \\ \cdots & q_{3j} & \cdots \\ \iddots & \vdots & \ddots \end{pmatrix}$$

so that

$$r_{ij} = p_{i1}q_{1j} + p_{i2}q_{2j} + p_{i3}q_{3j} + \cdots = \sum_k p_{ik}q_{kj}$$

where the sum is from $k = 1$ to the number of columns of $\mathbf{P}$ or the number of rows of $\mathbf{Q}$ (which are equal to each other).

We can now differentiate this expression using the product rule: the (i,j)th element of $\dot{\mathbf{R}}$ is

$$\begin{aligned} \dot{r}_{ij} &= (\dot{p}_{i1}q_{1j} + p_{i1}\dot{q}_{1j}) + (\dot{p}_{i2}q_{2j} + p_{i2}\dot{q}_{2j}) + \dots \\ &= \sum_k (\dot{p}_{ik}q_{kj} + p_{ik}\dot{q}_{kj}) \\ &= \sum_k \dot{p}_{ik}q_{kj} + \sum_k p_{ik}\dot{q}_{kj} \end{aligned}$$

In the final expression, the first sum is the (i,j)th element of $\dot{\mathbf{P}}\mathbf{Q}$ and the second sum is the (i,j)th element of $\mathbf{P}\dot{\mathbf{Q}}$, using exactly the same reasoning as we used to calculate the (i,j)th element of $\mathbf{PQ}$. Thus

$$\dot{\mathbf{R}} = \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{PQ}) = \dot{\mathbf{P}}\mathbf{Q} + \mathbf{P}\dot{\mathbf{Q}}.$$

> (iii) Prove that the derivative of a constant matrix is the zero matrix.

(A constant matrix means that every element is constant with respect to t. It does not necessarily mean that every element is equal to every other element.) If p_{ij} is a constant, then $\dot{p}_{ij} = 0$. So if p_{ij} is constant for every (i,j), then $\dot{\mathbf{P}} = \mathbf{0}$, the zero matrix.

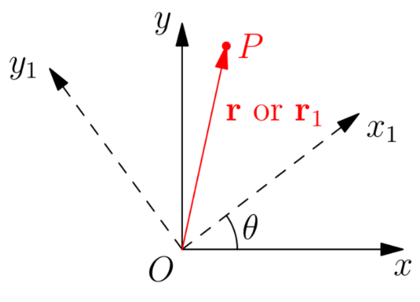
(iv) If $\mathbf{M}$ is the rotation matrix
$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$
θ being a function of t, prove that $\dot{\mathbf{M}} = \mathbf{MJ}\dot{\theta}$, where
$$\mathbf{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

We prove this by calculation, making use of the chain rule.

$$\begin{aligned}
\dot{\mathbf{M}} &= \begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t}(\cos\theta) & \frac{\mathrm{d}}{\mathrm{d}t}(-\sin\theta) \\ \frac{\mathrm{d}}{\mathrm{d}t}(\sin\theta) & \frac{\mathrm{d}}{\mathrm{d}t}(\cos\theta) \end{pmatrix} \\
&= \begin{pmatrix} -\sin\theta.\dot{\theta} & -\cos\theta.\dot{\theta} \\ \cos\theta.\dot{\theta} & -\sin\theta.\dot{\theta} \end{pmatrix} \\
&= \dot{\theta}\begin{pmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix}; \\
\mathbf{MJ} &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix}.
\end{aligned}$$

Therefore $\dot{\mathbf{M}} = \mathbf{MJ}\dot{\theta}$.

The position of a particle in a plane is specified by a vector which can be described either as $\mathbf{r}$ relative to a coordinate system with rectangular axes Ox, Oy, or as $\mathbf{r}_1$ relative to a coordinate system with axes Ox_1, Oy_1, as shown in the diagram. The angle xOx_1, denoted by θ, is a function of t.

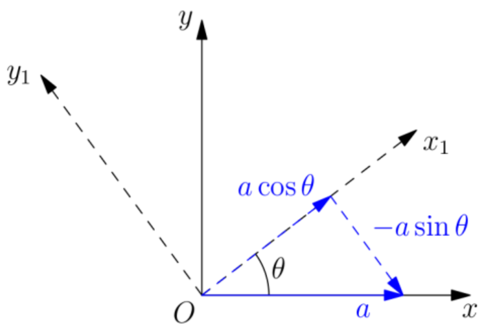


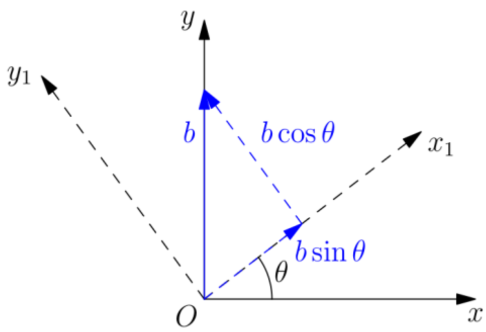
Write an equation connecting $\mathbf{r}$ with $\mathbf{r}_1$, and prove that
$$\dot{\mathbf{r}} = \mathbf{M}(\dot{\mathbf{r}}_1 + \mathbf{J}\mathbf{r}_1\dot{\theta}).$$

We can put coordinates on the system to understand what is going on. If P has coordinates (a, b) in the original coordinate system Oxy, then $\mathbf{r} = \begin{pmatrix} a \\ b \end{pmatrix}$. Now in the Ox_1y_1 coordinate system, we need to calculate the x_1 coordinate of P. This does not seem straightforward, as we do not know the angle that P makes with the Ox axis.

There are two useful approaches we could take to move forward, and we will present both.

The first approach is to realise that we can write $\mathbf{r} = \begin{pmatrix} a \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ b \end{pmatrix}$. Each of these simpler vectors will be easier to write in the new coordinate system, as we can work with them individually. Now $\begin{pmatrix} a \\ 0 \end{pmatrix}$ in the Oxy coordinate system becomes $\begin{pmatrix} a\cos\theta \\ -a\sin\theta \end{pmatrix}$ in the Ox_1y_1 system, and $\begin{pmatrix} 0 \\ b \end{pmatrix}$ in the Oxy system becomes $\begin{pmatrix} b\sin\theta \\ b\cos\theta \end{pmatrix}$ in the Ox_1y_1 system, as we can see from this diagram:





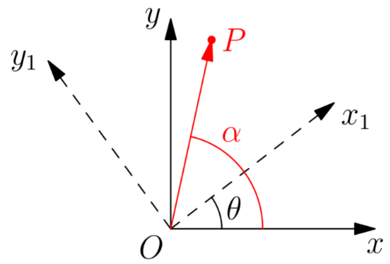
Therefore

$$\begin{aligned}
\mathbf{r}_1 &= \begin{pmatrix} a\cos\theta \\ -a\sin\theta \end{pmatrix} + \begin{pmatrix} b\sin\theta \\ b\cos\theta \end{pmatrix} \\
&= \begin{pmatrix} a\cos\theta + b\sin\theta \\ -a\sin\theta + b\cos\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\
&= \mathbf{M}^{-1}\mathbf{r},
\end{aligned}$$

so that $\mathbf{r} = \mathbf{M}\mathbf{r}_1$.

On reflection, we might realise that we could have equally started with a vector $\mathbf{r}_1$ in the Ox_1y_1 system, and rewritten that in the Oxy system; this would have saved us from having to use the inverse of $\mathbf{M}$. The diagrams and working are very similar, and are left as an exercise for the reader.

A second approach to working with $\mathbf{r}$ and $\mathbf{r}_1$ is to describe P using polar coordinates (magnitude and direction). If OP (or $\mathbf{r}$) makes an angle of α with Ox, and has length p, then it makes an angle of $\alpha - \theta$ with Ox_1, and still has length p:



Therefore $\mathbf{r} = \begin{pmatrix} p\cos\alpha \\ p\sin\alpha \end{pmatrix}$ and $\mathbf{r}_1 = \begin{pmatrix} p\cos(\alpha-\theta) \\ p\sin(\alpha-\theta) \end{pmatrix}$. It follows, using the compound angle formulae, that

$$\begin{aligned}
\mathbf{r}_1 &= \begin{pmatrix} p\cos(\alpha-\theta) \\ p\sin(\alpha-\theta) \end{pmatrix} \\
&= \begin{pmatrix} p\cos\alpha\cos\theta + p\sin\alpha\sin\theta \\ p\sin\alpha\cos\theta - p\cos\alpha\sin\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} p\cos\alpha \\ p\sin\alpha \end{pmatrix} \\
&= \mathbf{M}^{-1}\mathbf{r};
\end{aligned}$$

again, this gives $\mathbf{r} = \mathbf{M}\mathbf{r}_1$, and we could have obtained this directly by starting with OP making an angle of β to the Ox_1 axis, and then calculating $\mathbf{r}$ in terms of $\mathbf{r}_1$.

We now need to prove that $\dot{\mathbf{r}} = \mathbf{M}(\dot{\mathbf{r}}_1 + \mathbf{J}\mathbf{r}_1\dot{\theta})$. Differentiating $\mathbf{r} = \mathbf{M}\mathbf{r}_1$, using the above results, gives

$$\begin{aligned}
\dot{\mathbf{r}} &= \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{M}\mathbf{r}_1) \\
&= \dot{\mathbf{M}}\mathbf{r}_1 + \mathbf{M}\dot{\mathbf{r}}_1 && \text{from (ii)} \\
&= (\mathbf{M}\mathbf{J}\dot{\theta})\mathbf{r}_1 + \mathbf{M}\dot{\mathbf{r}}_1 && \text{from (iv)} \\
&= \mathbf{M}(\dot{\mathbf{r}}_1 + \mathbf{J}\mathbf{r}_1\dot{\theta}).
\end{aligned}$$

Question 7

If

$$x = \mathrm{e}^{kt}(a\cos\lambda t + b\sin\lambda t)$$

show that

$$\dot{x} = \mathrm{e}^{kt}(a'\cos\lambda t + b'\sin\lambda t),$$

where

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} k & \lambda \\ -\lambda & k \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

and dot denotes differentiation with respect to t, and find an expression for $\ddot{x}$ with coefficients given in a similar way.

Differentiating using the product and chain rules gives

$$\begin{aligned}\dot{x} &= k\mathrm{e}^{kt}(a\cos\lambda t + b\sin\lambda t) + \mathrm{e}^{kt}(-a\lambda\sin\lambda t + b\lambda\cos\lambda t)\\ &= \mathrm{e}^{kt}((ak + b\lambda)\cos\lambda t + (kb - a\lambda)\sin\lambda t)\\ &= \mathrm{e}^{kt}(a'\cos\lambda t + b'\sin\lambda t)\end{aligned}$$

where

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} ak + b\lambda \\ kb - a\lambda \end{pmatrix} = \begin{pmatrix} k & \lambda \\ -\lambda & k \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

Since $\dot{x}$ has the same form as x, just with the constants a and b replaced by a' and b', we can differentiate it again using the rule we have just derived:

$$\ddot{x} = \mathrm{e}^{kt}(a''\cos\lambda t + b''\sin\lambda t)$$

where

$$\begin{pmatrix} a'' \\ b'' \end{pmatrix} = \begin{pmatrix} k & \lambda \\ -\lambda & k \end{pmatrix} \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} k & \lambda \\ -\lambda & k \end{pmatrix}^2 \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} k^2 - \lambda^2 & 2k\lambda \\ -2k\lambda & k^2 - \lambda^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

A particular integral solution to

$$\ddot{x} + 2p\dot{x} + q^2 x = \mathrm{e}^{kt}(C\cos\lambda t + D\sin\lambda t)$$

is

$$x = \mathrm{e}^{kt}(a\cos\lambda t + b\sin\lambda t).$$

Show that

$$\begin{pmatrix} C \\ D \end{pmatrix} = \begin{pmatrix} (k^2-\lambda^2)+2pk+q^2 & 2k\lambda+2p\lambda \\ -2k\lambda-2p\lambda & (k^2-\lambda^2)+2pk+q^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}$$

and hence write $\begin{pmatrix} a \\ b \end{pmatrix}$ in terms of $\begin{pmatrix} C \\ D \end{pmatrix}$.

Discuss any particular cases.

The left hand side of the differential equation becomes

$$\mathrm{e}^{kt}\big((a''\cos\lambda t + b''\sin\lambda t) + 2p(a'\cos\lambda t + b'\sin\lambda t) + q^2(a\cos\lambda t + b\sin\lambda t)\big).$$

Equating coefficients of $\mathrm{e}^{kt}\cos\lambda t$ and $\mathrm{e}^{kt}\sin\lambda t$ on the two sides of the equation gives

$$\begin{aligned}\begin{pmatrix} C \\ D \end{pmatrix} &= \begin{pmatrix} a''+2pa'+q^2a \\ b''+2pb'+q^2b \end{pmatrix} \\ &= \begin{pmatrix} k^2-\lambda^2 & 2k\lambda \\ -2k\lambda & k^2-\lambda^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} + 2p\begin{pmatrix} k & \lambda \\ -\lambda & k \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} + q^2\begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} (k^2-\lambda^2)+2pk+q^2 & 2k\lambda+2p\lambda \\ -2k\lambda-2p\lambda & (k^2-\lambda^2)+2pk+q^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}.\end{aligned}$$

Hence

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} (k^2-\lambda^2)+2pk+q^2 & 2k\lambda+2p\lambda \\ -2k\lambda-2p\lambda & (k^2-\lambda^2)+2pk+q^2 \end{pmatrix}^{-1}\begin{pmatrix} C \\ D \end{pmatrix}.$$

We can invert this matrix explicitly if it is non-singular. The determinant is

$$\Delta = \big((k^2-\lambda^2)+2pk+q^2\big)^2 + (2k\lambda+2p\lambda)^2,$$

and so

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{\Delta}\begin{pmatrix} (k^2-\lambda^2)+2pk+q^2 & -2k\lambda-2p\lambda \\ 2k\lambda+2p\lambda & (k^2-\lambda^2)+2pk+q^2 \end{pmatrix}\begin{pmatrix} C \\ D \end{pmatrix}.$$

The only special case which arises is when the determinant is zero. In such a case, we cannot determine C and D from this equation; this is either because no particular integral solution to the equation has this form, or because at least one of the complementary functions has the same form and so there are infinitely many possibilities for a and b.

Since Δ is the sum of two squares, it can only be zero if both of the squares are zero, so we would need

$$(k^2-\lambda^2)+2pk+q^2 = (k+p)\lambda = 0.$$

There are clearly values of k, λ, p and q which satisfy these, even if we exclude the possibility of $\lambda = 0$ or $k = 0$; for example, we may take $p = -k$ and $q^2 = k^2 + \lambda^2$. In this case, the auxiliary quadratic is $u^2 - 2ku + (k^2 + \lambda^2) = 0$, which has roots $k \pm \lambda \mathrm{i}$, giving complementary functions of the form $\mathrm{e}^{kt} \cos \lambda t$ and $\mathrm{e}^{kt} \sin \lambda t$. We would then have to seek a particular integral of the form $x = t\mathrm{e}^{kt}(a \cos \lambda t + b \sin \lambda t)$ instead.

Question 8

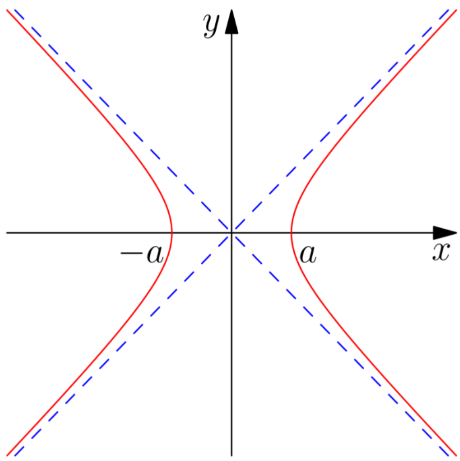
Sketch the graph whose equation is

$$x^2 - y^2 = a^2 \quad (a > 0).$$

This is a (rectangular) hyperbola. As $x \to \pm\infty$, y^2 and x^2 are approximately the same ($y^2 = x^2 - a^2$, so $\frac{y^2}{x^2} = 1 - \frac{a^2}{x^2}$ tends to 1 as $x^2 \to \infty$). Therefore $y = x$ and $y = -x$ are asymptotes.

When $x = 0$, $y^2 = -a^2$, which is impossible, so the graph does not cross the y-axis. As $x^2 = y^2 + a^2$, $x^2 \geqslant a^2$, so $|x| \geqslant a$. When $y = 0$, $x = \pm a$, so these are the x-axis intercepts.

Therefore the graph looks approximately like this:



Prove that the point $P(a\cosh t, a\sinh t)$ lies on this graph for all real values of t; but that there are points of the graph which cannot be expressed in this form.

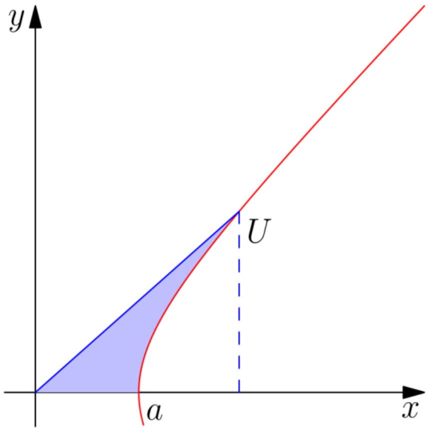
Calculating, we have $x^2 - y^2 = (a\cosh t)^2 - (a\sinh t)^2 = a^2(\cosh^2 t - \sinh^2 t) = a^2$, so P lies on this graph.

Note now that $a\cosh t > 0$ for all real t. However, we have already noted that $(-a, 0)$ lies on the graph. So there are points of the graph which cannot be expressed in this form.

We get all the points on the right-hand branch of this hyperbola using this form of coordinates. We can obtain the left-hand branch as $(-a\cosh t, a\sinh t)$ for real values of t.

If U is the point of the graph for which $t = u$, find the area of the region $\mathscr{R}$ bounded by the curve, the line OU and the x-axis. (You may assume that $u > 0$.)

We first sketch this region:



We see that we can calculate the area of $\mathscr{R}$ by finding the area under the right-angled triangle from the origin to U to $(a\cosh u, 0)$ and subtracting the area under the graph. The area under the triangle is

$$\tfrac{1}{2}a\cosh u . a\sinh u = \tfrac{1}{2}a^2\sinh u\cosh u = \tfrac{1}{4}a^2\sinh 2u,$$

while the area under the curve is

$$\begin{aligned}
A &= \int_0^u y\frac{\mathrm{d}x}{\mathrm{d}t}\,\mathrm{d}t \\
&= \int_0^u a\sinh t . a\sinh t\,\mathrm{d}t \\
&= a^2\int_0^u \sinh^2 t\,\mathrm{d}t \\
&= a^2\int_0^u \tfrac{1}{2}(\cosh 2t - 1)\,\mathrm{d}t \\
&= \tfrac{1}{2}a^2\left[\tfrac{1}{2}\sinh 2t - t\right]_0^u \\
&= \tfrac{1}{2}a^2\left(\tfrac{1}{2}\sinh 2u - u\right) \\
&= \tfrac{1}{4}a^2\sinh 2u - \tfrac{1}{2}a^2u.
\end{aligned}$$

Subtracting these shows that the area of $\mathscr{R}$ is $\frac{1}{2}a^2u$.

A quick reasonableness check is in order: when u is small (which is probably all that is straightforward to get our hands on), the area is approximately a triangle with base a and height $a\sinh u \approx au$, so the area is approximately $\frac{1}{2}a^2u$.

Prove that the transformation with matrix

$$\begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix} \qquad (\alpha \neq 0)$$

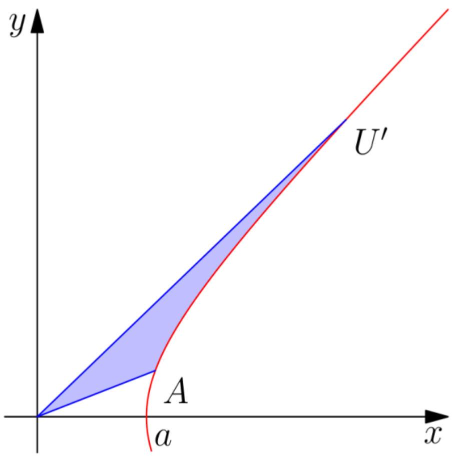
transforms the point P into another point of the curve. Into what region is $\mathscr{R}$ transformed by this? What is the area of the transformed region? Give reasons for your answers.

The point $P(a\cosh t, a\sinh t)$ is transformed into

$$\begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix}\begin{pmatrix} a\cosh t \\ a\sinh t \end{pmatrix} = \begin{pmatrix} a\cosh\alpha\cosh t + a\sinh\alpha\sinh t \\ a\sinh\alpha\cosh t + a\cosh\alpha\sinh t \end{pmatrix} = \begin{pmatrix} a\cosh(t+\alpha) \\ a\sinh(t+\alpha) \end{pmatrix},$$

which is the point P' *on the hyperbola* with parameter $t+\alpha$.

The origin is transformed into the origin, the point $(a, 0)$, corresponding to the parameter 0, is transformed to the point A with parameter α, and the point U is transformed to the point U' with parameter $u+\alpha$. Since linear transformations send straight lines to straight lines (or points), the part of the x-axis bounding the region $\mathscr{R}$ is transformed to the straight line from the origin to $(a\cosh\alpha, a\sinh\alpha)$ and the line segment OU is transformed to the line segment OU', giving the region $\mathscr{R}\,'$:



The area of this region is the area of the region bounded by OU', the curve and the x-axis, minus the area of the region bounded by OA, the curve and the x-axis. Using our above calculation for the area of $\mathscr{R}$ and replacing the parameter, we find the transformed area is $\frac{1}{2}a^2(u+\alpha) - \frac{1}{2}a^2\alpha = \frac{1}{2}a^2u$, which is just the area of $\mathscr{R}$.

Another way of finding the area of the transformed region is to find the determinant of the matrix, which is $\cosh^2\alpha - \sinh^2\alpha = 1$, so the area scale factor is 1. Thus the area of the transformed region is the same as the area of the original region.

We see that this transformation matrix "rotates" the hyperbola, in the same way that $\left(\begin{smallmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{smallmatrix}\right)$ rotates a circle. This is an important idea in mathematical physics and other areas of mathematics.

Question 9

Show that the equations

$$\begin{aligned} 3x + 2y + z &= a - 1, \\ -2x + (a-2)y - az &= 2a, \\ 6x + ay + (a-2)z &= 3a - 6 \end{aligned}$$

have a solution, not necessarily unique, unless $a = \frac{2}{3}$.

Find the complete solution when $a = 0$ and when $a = 4$.

We can write the equations as $\mathbf{Ax} = \mathbf{b}$, where

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 1 \\ -2 & a-2 & -a \\ 6 & a & a-2 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} a-1 \\ 2a \\ 3a-6 \end{pmatrix}.$$

We start by calculating the determinant of $\mathbf{A}$. We can use any technique we like; we'll use some row operations to demonstrate how this approach can be used.

$$\begin{aligned}
\det \mathbf{A} &= \begin{vmatrix} 3 & 2 & 1 \\ -2 & a-2 & -a \\ 6 & a & a-2 \end{vmatrix} \\
&= \begin{vmatrix} 3 & 2 & 1 \\ -2 & a-2 & -a \\ 4 & 2a-2 & -2 \end{vmatrix} && \mathbf{r}_3 \to \mathbf{r}_3 + \mathbf{r}_2 \\
&= \begin{vmatrix} 3 & 2 & 1 \\ 3a-2 & 3a-2 & 0 \\ 4 & 2a-2 & -2 \end{vmatrix} && \mathbf{r}_2 \to \mathbf{r}_2 + a\mathbf{r}_1 \\
&= (3a-2)\begin{vmatrix} 3 & 2 & 1 \\ 1 & 1 & 0 \\ 4 & 2a-2 & -2 \end{vmatrix} && \text{factorise } \mathbf{r}_2 \\
&= (3a-2)\begin{vmatrix} 3 & 2 & 1 \\ 1 & 1 & 0 \\ 10 & 2a+2 & 0 \end{vmatrix} && \mathbf{r}_3 \to \mathbf{r}_3 + 2\mathbf{r}_1 \\
&= (3a-2)\begin{vmatrix} 1 & 1 \\ 10 & 2a+2 \end{vmatrix} && \text{expanding down the third column} \\
&= (3a-2)(2a+2-10) \\
&= (3a-2)(2a-8) \\
&= 2(3a-2)(a-4).
\end{aligned}$$

As there is a unique solution (namely $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$) whenever $\det \mathbf{A} \neq 0$, there is a unique solution for $a \neq \frac{2}{3}$, $a \neq 4$.

When $a = \frac{2}{3}$, the equations become

$$\begin{aligned} 3x + 2y + z &= -\tfrac{1}{3}, \\ -2x - \tfrac{4}{3}y - \tfrac{2}{3}z &= \tfrac{4}{3}, \\ 6x + \tfrac{2}{3}y - \tfrac{4}{3}z &= -4. \end{aligned}$$

For clarity, we can multiply the second and third equations by $\pm\frac{3}{2}$ to clear fractions and remove common factors, giving:

$$\begin{aligned} 3x + 2y + z &= -\tfrac{1}{3}, \\ 3x + 2y + z &= -2, \\ 9x + y - 2z &= -6. \end{aligned}$$

It is now apparent that the first two equations are inconsistent: they have no common solution. Geometrically, they represent two distinct parallel planes.

When $a = 4$, the equations become

$$\begin{aligned} 3x + 2y + z &= 3, \\ -2x + 2y - 4z &= 8, \\ 6x + 4y + 2z &= 6. \end{aligned}$$

Again, we take out common factors from the second and third equations:

$$\begin{aligned} 3x + 2y + z &= 3, \\ x - y + 2z &= -4, \\ 3x + 2y + z &= 3. \end{aligned}$$

We see that the first and third equations are identical and the left hand side of the second equation is not a multiple of the left hand side of the other two. So geometrically, these are two identical planes and a third non-identical plane, and there will therefore be infinitely many solutions.

We can find these solutions by eliminating z from the first two equations (second equation minus 2 times the first) to get $-5x - 5y = -10$, or $x + y = 2$. So if we let $x = t$, $y = 2 - t$, we get $z = 3 - 3x - 2y = 3 - 3t - 2(2 - t) = -1 - t$, so the complete solution is all points of the form $(t, 2 - t, -1 - t)$.

Finally, in the case that $a = 0$, there is a unique solution. The equations become

$$\begin{aligned} 3x + 2y + z &= -1, \\ -2x - 2y &= 0, \\ 6x - 2z &= -6. \end{aligned}$$

We could find the inverse of $\mathbf{A}$ in this case, but it does not seem worth the effort. We remove common factors again to obtain

$$\begin{aligned} 3x + 2y + z &= -1, \\ x + y &= 0, \\ 3x - z &= -3. \end{aligned}$$

The second equation gives $y = -x$, so the other two become

$$\begin{aligned} x + z &= -1, \\ 3x - z &= -3. \end{aligned}$$

Adding these gives $4x = -4$, so $x = -1$, $y = 1$ and $z = 0$.

Question 10

If z is the complex number $x + \mathrm{i}y$, $\mathrm{i} = \sqrt{-1}$, let $\mathbf{M}(z)$ denote the 2×2 matrix

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix}.$$

Prove that

$$\mathbf{M}(z + z') = \mathbf{M}(z) + \mathbf{M}(z')$$

and that

$$\mathbf{M}(zz') = \mathbf{M}(z)\mathbf{M}(z').$$

We let z' be the complex number $x' + \mathrm{i}y'$, so that $\mathbf{M}(z') = \begin{pmatrix} x' & y' \\ -y' & x' \end{pmatrix}$. We then have

$$\begin{aligned}
\mathbf{M}(z + z') &= \mathbf{M}((x + x') + \mathrm{i}(y + y')) \\
&= \begin{pmatrix} x + x' & y + y' \\ -(y + y') & x + x' \end{pmatrix} \\
\mathbf{M}(z) + \mathbf{M}(z') &= \begin{pmatrix} x & y \\ -y & x \end{pmatrix} + \begin{pmatrix} x' & y' \\ -y' & x' \end{pmatrix} \\
&= \begin{pmatrix} x + x' & y + y' \\ -(y + y') & x + x' \end{pmatrix} \\
\mathbf{M}(zz') &= \mathbf{M}((x + \mathrm{i}y)(x' + \mathrm{i}y')) \\
&= \mathbf{M}((xx' - yy') + \mathrm{i}(xy' + yx')) \\
&= \begin{pmatrix} xx' - yy' & xy' + yx' \\ -(xy' + yx') & xx' - yy' \end{pmatrix} \\
\mathbf{M}(z)\mathbf{M}(z') &= \begin{pmatrix} x & y \\ -y & x \end{pmatrix} \begin{pmatrix} x' & y' \\ -y' & x' \end{pmatrix} \\
&= \begin{pmatrix} xx' - yy' & xy' + yx' \\ -(xy' + yx') & xx' - yy' \end{pmatrix}
\end{aligned}$$

Therefore $\mathbf{M}(z + z') = \mathbf{M}(z) + \mathbf{M}(z')$ and $\mathbf{M}(zz') = \mathbf{M}(z)\mathbf{M}(z')$.

In the language of abstract algebra, this (together with $\mathbf{M}(1) = \mathbf{I}$) shows that the function $\mathbf{M}$ gives an *isomorphism* between the (ring of) complex numbers and the (ring of) 2×2 matrices of this form. This idea is studied more extensively in many undergraduate mathematics courses.

Hence, or otherwise, show that

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}^n = \begin{pmatrix} \cos n\theta & \sin n\theta \\ -\sin n\theta & \cos n\theta \end{pmatrix}.$$

We can write the left-hand side using the $\mathbf{M}$ function:

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}^n = (\mathbf{M}(\cos\theta + \mathrm{i}\sin\theta))^n = \mathbf{M}((\cos\theta + \mathrm{i}\sin\theta)^n)$$

using the multiplicative property of $\mathbf{M}$. (To be very careful, we should prove that $\mathbf{M}(z^n) = (\mathbf{M}(z))^n$ by induction, which is clearly true from $\mathbf{M}(zz') = \mathbf{M}(z)\mathbf{M}(z')$. Also, the question has not specified that n is a positive integer; we will assume this to be the case, though the result still holds if n is an arbitrary integer.)

Now we can use de Moivre's theorem to write $(\cos\theta + \mathrm{i}\sin\theta)^n = \cos n\theta + \mathrm{i}\sin n\theta$, so that

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}^n = \mathbf{M}(\cos n\theta + \mathrm{i}\sin n\theta) = \begin{pmatrix} \cos n\theta & \sin n\theta \\ -\sin n\theta & \cos n\theta \end{pmatrix}.$$

This is a potentially circular argument, for how do we prove de Moivre's theorem? One way to do so is by induction, using the addition formulae for sine and cosine. But how do we prove those? One way to do so is by writing down the matrix for rotating by A and the matrix for rotating by B; multiplying these gives the matrix for the composite transformation, which is rotating by $A+B$. But now we are using de Moivre's theorem to prove a result about rotation matrices, which is almost where we obtained de Moivre's theorem from in the first place! One way out of this is to prove de Moivre's theorem and the addition formulae using the series definitions of sine, cosine and the exponential function e^z. But that would take us too far afield here.

Hence find three real 2×2 matrices $\mathbf{A}$ such that $\mathbf{A}^3 = \mathbf{I}$ and a 2×2 matrix $\mathbf{B}$ such that $\mathbf{B}^2 + \mathbf{I} = \mathbf{B}$.

If we have $\mathbf{A} = \mathbf{M}(z)$, then we require $\mathbf{A}^3 = \mathbf{I}$, so $\mathbf{M}(z^3) = \mathbf{M}(1)$. This will be satisfied if and only if $z^3 = 1$, and this has three solutions, $z = 1$, $z = \cos\frac{2\pi}{3} + \mathrm{i}\sin\frac{2\pi}{3} = -\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}$ and $z = \cos\frac{2\pi}{3} - \mathrm{i}\sin\frac{2\pi}{3} = -\frac{1}{2} - \mathrm{i}\frac{\sqrt{3}}{2}$. Therefore three real matrices are

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}; \quad \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

For matrix $\mathbf{B}$, if we take $\mathbf{B} = \mathbf{M}(z)$, then the equation becomes $\mathbf{M}(z^2) + \mathbf{M}(1) = \mathbf{M}(z)$, or $\mathbf{M}(z^2 - z + 1) = \mathbf{M}(0)$. We can solve $z^2 - z + 1 = 0$ to obtain $z = \frac{1 \pm \sqrt{-3}}{2}$, giving two possible matrices $\mathbf{B}$ (though there may be others not of this form):

$$\begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \text{ and } \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}.$$

Question 11

Using any method you wish, calculate the inverse $\mathbf{A}^{-1}$ of the matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 3 \\ 3 & 4 & 6 \end{pmatrix}.$$

We demonstrate how Gaussian elimination (row operations) can be used to find the inverse of a matrix, as mentioned briefly in the notes. We start with an augmented version of $\mathbf{A}$, and perform row operations until the left side of the matrix becomes the identity. In this process, we are also allowed to swap rows (which can be achieved through a sequence of steps involving adding multiples of rows to other rows and multiplying a whole row by a constant).

$$\begin{aligned}
&\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 3 & 3 & 0 & 1 & 0 \\ 3 & 4 & 6 & 0 & 0 & 1 \end{array}\right) \\
\to &\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & -3 & 0 & 1 \end{array}\right) \qquad (\mathbf{r}_2 \to \mathbf{r}_2 - 2\mathbf{r}_1;\ \mathbf{r}_3 \to \mathbf{r}_3 - 3\mathbf{r}_1) \\
\to &\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -3 & 0 & 1 \\ 0 & 1 & -1 & -2 & 1 & 0 \end{array}\right) \qquad (\text{swap } \mathbf{r}_2 \text{ and } \mathbf{r}_3) \\
\to &\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 4 & 0 & -1 \\ 0 & 1 & 0 & -3 & 0 & 1 \\ 0 & 0 & -1 & 1 & 1 & -1 \end{array}\right) \qquad (\mathbf{r}_1 \to \mathbf{r}_1 - \mathbf{r}_2;\ \mathbf{r}_3 \to \mathbf{r}_3 - \mathbf{r}_2) \\
\to &\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6 & 2 & -3 \\ 0 & 1 & 0 & -3 & 0 & 1 \\ 0 & 0 & -1 & 1 & 1 & -1 \end{array}\right) \qquad (\mathbf{r}_1 \to \mathbf{r}_1 + 2\mathbf{r}_3) \\
\to &\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6 & 2 & -3 \\ 0 & 1 & 0 & -3 & 0 & 1 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right) \qquad (\mathbf{r}_3 \to -\mathbf{r}_3)
\end{aligned}$$

Hence

$$\mathbf{A}^{-1} = \begin{pmatrix} 6 & 2 & -3 \\ -3 & 0 & 1 \\ -1 & -1 & 1 \end{pmatrix}.$$

Interpret geometrically in three dimensions the equations

$$\begin{aligned} x + y + 2z &= 4, \\ 2x + 3y + 3z &= 8, \\ 3x + 4y + \lambda z &= 7 + \lambda \end{aligned}$$

in the cases $\lambda = 6$ and $\lambda = 5$.

These three equations are the equations of three planes. They can be written as $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is the above matrix in the case $\lambda = 6$, and the above matrix with the bottom right element replaced by 5 in the case $\lambda = 5$; also, $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 4 \\ 8 \\ 7+\lambda \end{pmatrix}$.

In the case $\lambda = 6$, we have seen that $\mathbf{A}$ is invertible, so the three planes meet at a single point. We can calculate the coordinates of this point as $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$:

$$\mathbf{x} = \begin{pmatrix} 6 & 2 & -3 \\ -3 & 0 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 8 \\ 13 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

so the planes meet at $(1, 1, 1)$, and indeed this satisfies all three equations.

In the case $\lambda = 5$, we have

$$\begin{aligned} x + y + 2z &= 4, \\ 2x + 3y + 3z &= 8, \\ 3x + 4y + 5z &= 12. \end{aligned}$$

It is easy to spot that the third equation is the sum of the first two equations, so if the first two equations are satisfied, the third one will also be. We can eliminate x from the first two equations to get $y - z = 0$, so $y = z$, and thus a general solution is $y = z = t$, $x = 4 - 3t$, which is a parametric equation of a straight line. Geometrically, this means that the three planes all meet along this line, but no two of the planes are identical.

Question 12

Using any method you prefer, and solving the problems in any order you prefer, find

(i) the inverse of the matrix

$$\begin{pmatrix} 1 & 2 & 4 \\ 3 & 5 & 7 \\ 6 & 8 & 3 \end{pmatrix},$$

(ii) the solution of the simultaneous equations

$$\begin{aligned} x + 2y + 4z &= 3, \\ 3x + 5y + 7z &= 12, \\ 6x + 8y + 3z &= 31. \end{aligned}$$

For this question, we will use the adjugate matrix approach. We will calculate the inverse first, and then use this to solve the simultaneous equations.

Writing $\mathbf{A}$ for the matrix in (i), the simultaneous equations become $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 3 \\ 12 \\ 31 \end{pmatrix}$, so that $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

We can calculate the determinant of $\mathbf{A}$ using the forwards and backwards diagonals approach (see the STEP 3 Matrices Topic Notes):

$$\begin{aligned} \det \mathbf{A} &= 1.5.3 + 2.7.6 + 4.3.8 - 1.7.8 - 2.3.3 - 4.5.6 \\ &= 15 + 84 + 96 - 56 - 18 - 120 \\ &= 1. \end{aligned}$$

(How conveniently this matrix was chosen by the examiners!) Since the inverse matrix is $\frac{1}{\det \mathbf{A}}$ times the adjugate matrix, the adjugate matrix $\operatorname{adj} \mathbf{A}$ and the inverse matrix $\mathbf{A}^{-1}$ are the same in this case.

The matrix of cofactors is then

$$\begin{pmatrix} +\begin{vmatrix} 5 & 7 \\ 8 & 3 \end{vmatrix} & -\begin{vmatrix} 3 & 7 \\ 6 & 3 \end{vmatrix} & +\begin{vmatrix} 3 & 5 \\ 6 & 8 \end{vmatrix} \\ -\begin{vmatrix} 2 & 4 \\ 8 & 3 \end{vmatrix} & +\begin{vmatrix} 1 & 4 \\ 6 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 6 & 8 \end{vmatrix} \\ +\begin{vmatrix} 2 & 4 \\ 5 & 7 \end{vmatrix} & -\begin{vmatrix} 1 & 4 \\ 3 & 7 \end{vmatrix} & +\begin{vmatrix} 1 & 2 \\ 3 & 5 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} -41 & 33 & -6 \\ 26 & -21 & 4 \\ -6 & 5 & -1 \end{pmatrix}$$

Therefore the adjugate matrix and inverse are the transpose of this, namely

$$\begin{pmatrix} -41 & 26 & -6 \\ 33 & -21 & 5 \\ -6 & 4 & -1 \end{pmatrix}.$$

(It is certainly worth checking that for our calculated $\mathbf{A}^{-1}$, we have $\mathbf{AA}^{-1} = \mathbf{I}$; this will catch calculational errors.)

Then the solution of the simultaneous equations is given by

$$\mathbf{x} = \begin{pmatrix} -41 & 26 & -6 \\ 33 & -21 & 5 \\ -6 & 4 & -1 \end{pmatrix} \begin{pmatrix} 3 \\ 12 \\ 31 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix},$$

so $x = 3$, $y = 2$, $z = -1$. (Again, these should be substituted into the equations to check the solution is correct.)

> The last of the three equations is replaced by
>
> $$6x + 8y + az = b,$$
>
> and it is found that the first two equations together with the new third one have more than one solution. Find a and b, and state a geometrical interpretation in three dimensions for these equations.

If the equations have more than one solution, then geometrically we must have one of the following scenarios:

- three identical planes, so all three equations are multiples of each other
- two identical planes and one non-parallel plane, so two equations are multiples of each other
- no parallel planes, but all three planes sharing a common line

None of these equations are multiples of each other (we only need to look at the x and y components to see this). Therefore we must be dealing with the third situation.

Algebraically, multiple solutions implies that the equations are not *linearly independent*, meaning that some equation is the sum of multiples of the other equations. The third situation requires that each equation is the sum of multiples of the other two equations. In our case, this means that $6x + 8y + az = b$ is the sum of λ times $x + 2y + 4z = 3$ and μ times $3x + 5y + 7z = 12$ for some λ and μ. So we require

$$\begin{aligned} \lambda + 3\mu &= 6 \\ 2\lambda + 5\mu &= 8 \\ 4\lambda + 7\mu &= a \\ 3\lambda + 12\mu &= b; \end{aligned}$$

we can easily solve the first two equations to obtain $\lambda = -6$ and $\mu = 4$, hence $a = 4$ and $b = 30$.

Question 13

$\mathbf{A}$ is a 3×3 matrix whose elements are 0, 1 or -1 and each row and each column of $\mathbf{A}$ contains exactly one non-zero element. Prove that $\mathbf{A}^2$, $\mathbf{A}^3$, $\ldots$, $\mathbf{A}^n$ are all of the same form and deduce that $\mathbf{A}^h = \mathbf{I}$ for some positive integer $h \leqslant 48$.

We are asked to prove something for a general positive integer n, so induction seems like a natural approach. The most difficult part of this question is finding a way to express the ideas clearly.

We therefore show that if $\mathbf{A}^k$ has this form, then so does $\mathbf{A}^{k+1}$. Now $\mathbf{A}^{k+1} = \mathbf{A}\mathbf{A}^k$. Consider how this multiplication works:

- The first row of $\mathbf{A}^{k+1}$ is obtained by multiplying the first row of $\mathbf{A}$ by each column of $\mathbf{A}^k$. Since the first row of $\mathbf{A}$ has exactly one non-zero element (either 1 or -1), this will result in the first row of $\mathbf{A}^{k+1}$ being:
  - the first/second/third row of $\mathbf{A}^k$ if the first/second/third element of the first row of $\mathbf{A}$ is 1
  - minus the first/second/third row of $\mathbf{A}^k$ if the first/second/third element of the first row of $\mathbf{A}$ is -1

  So the first row of $\mathbf{A}^{k+1}$ will contain exactly one non-zero element, and this element will be 1 or -1.
- The second row of $\mathbf{A}^{k+1}$ is obtained by multiplying the second row of $\mathbf{A}$ by each column of $\mathbf{A}^k$. As with the first row, this means that the second row will contain exactly one non-zero element, which is again 1 or -1.
- The third row behaves in the same way.
- Thus the first row of $\mathbf{A}^k$ becomes the b_1-th row of $\mathbf{A}^{k+1}$, possibly multiplied by -1, where the $(1, b_1)$-th element of $\mathbf{A}$ is non-zero. The second row of $\mathbf{A}^k$ becomes the b_2-th row of $\mathbf{A}^{k+1}$ (again possibly multiplied by -1), where the $(2, b_2)$-th element of $\mathbf{A}$ is non-zero, and similarly for the third row.
- Since each column of $\mathbf{A}$ contains exactly one non-zero element, b_1, b_2 and b_3 are distinct, so they take the values 1, 2 and 3 in some order. This means that the rows of $\mathbf{A}^k$ are simply reordered (the technical term is *permuted*) to obtain $\mathbf{A}^{k+1}$, and some of them may be negated too.
- Therefore each row of $\mathbf{A}^{k+1}$ contains exactly one non-zero element, and each column of $\mathbf{A}^{k+1}$ contains exactly one non-zero element. This also shows that every element of $\mathbf{A}^{k+1}$ is 0, 1 or -1.

Therefore our induction step holds. Since $\mathbf{A}^1$ has the required form, we have a basis for induction, and hence $\mathbf{A}^n$ has this form for all positive integer n.

What follows is a standard proof technique for this type of question. It is well worth paying attention to the structure of this argument, because it is very common in mathematics.

Now consider how many possible such matrices there are. There are 3 choices for b_1 (using the above notation), 2 choices for b_2 (once b_1 has been chosen), and then b_3 is forced, so there are $3! = 6$ choices for the positions of the non-zero terms. Each one is either positive or negative, so there are $2^3 = 8$ possible sign choices for each position. There are therefore 48 distinct matrices of this form.

Then if we look at $\mathbf{A}$, $\mathbf{A}^2$, ..., $\mathbf{A}^{49}$, at least two of these 49 matrices must be the same. Let's say that $\mathbf{A}^j = \mathbf{A}^k$, where $1 \leqslant j < k \leqslant 49$. Then as the determinant of $\mathbf{A}$ is non-zero (it is ± 1 – see below), $\mathbf{A}^j$ is invertible, so we have $\mathbf{I} = \mathbf{A}^k(\mathbf{A}^j)^{-1} = \mathbf{A}^{k-j}$. Since $1 \leqslant k - j \leqslant 48$, we have found $h = k - j$ with $\mathbf{A}^h = \mathbf{I}$ and $h \leqslant 48$, as required.

In the previous paragraph, we claimed that the determinant of $\mathbf{A}$ is non-zero. The reason for this is as follows. Let us write

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Then the full expression for the determinant is the sum/difference of 6 terms:

$$\det \mathbf{A} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

Each of these terms is the product of exactly one element from each row and one element from each column of $\mathbf{A}$, and all 6 such possibilities appear here. Since $\mathbf{A}$ has exactly one non-zero element in each row and in each column, exactly one of these 6 terms is non-zero and the rest are zero. So the determinant is the product of the non-zero elements of $\mathbf{A}$, which is ± 1.

This argument extends to $n \times n$ matrices with this form. There will be $n! \times 2^n$ possible matrices, each of which will have determinant ± 1, and $\mathbf{A}^h = \mathbf{I}$ for some positive integer $h \leqslant n! \times 2^n$.

Interpret the action of $\mathbf{A}$ on a vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ geometrically.

The precise geometric description of $\mathbf{A}$ will clearly depend on the exact matrix. But we can say some general things about the geometric behaviour. Algebraically, the action of $\mathbf{A}$ permutes the components of the vector and possibly negates some of them.

Geometrically, swapping two elements, say x and y, corresponds to reflecting this vector in the plane $x = y$. And sending x to y, y to z and z to x corresponds to rotating it by $\frac{1}{3}$ of a turn, so by an angle of $\frac{2\pi}{3}$, about the line $x = y = z$; sending x to z, z to y and y to x corresponds to a rotation by $-\frac{2\pi}{3}$ about the same line.

Negating x corresponds geometrically to reflecting the vector in the plane $x = 0$, and likewise for y and z.

So the action of $\mathbf{A}$ is one of the identity, a reflection in $x = y$, a reflection in $y = z$, a reflection in $z = x$, or a rotation of $\pm\frac{2\pi}{3}$ about the line $x = y = z$, possibly followed by reflections in some or all of the planes $x = 0$, $y = 0$ or $z = 0$.

We could, if we wished, list the behaviour of all 48 possible matrices separately. These turn out to correspond to the so-called *symmetries of a cube*, that is, all possible transformations of a unit cube centred on the origin and with edges parallel to the axes, such that the image of the cube is the same cube. These can be described as follows:

- the identity transformation (1 transformation)
- rotation by $\pm\frac{\pi}{2}$ about an axis passing through the centres of opposite faces (6 transformations)
- rotation by π about an axis passing through the centres of opposite faces (3 transformations)
- rotation by π about an axis passing through the midpoints of opposite edges (6 transformations)
- rotation by $\pm\frac{2\pi}{3}$ about an axis passing through opposite corners (8 transformations)
- reflection about a plane through the centre of the cube parallel to a face (3 transformations)
- reflection about a plane through the centre of the cube passing through two opposite edges (6 transformations)
- reflection about the centre of the cube (1 transformation)
- rotation by $\pm\frac{\pi}{2}$ about an axis passing through the centres of opposite faces followed by a reflection in the plane through the centre of the cube perpendicular to the rotation axis (6 transformations)
- rotation by π about an axis passing through the centres of opposite faces followed by a reflection in the plane through the centre of the cube perpendicular to the rotation axis: this is actually the same as a reflection about the centre of the cube, so has already been counted
- rotation by $\pm\frac{2\pi}{3}$ about an axis passing through opposite corners followed by a reflection in a plane through the centre of the cube parallel to a face (8 distinct transformations)

Any other composition of reflections and/or rotations is equivalent to one of the 48 in this list.

Question 14

Find all possible solutions to

$$\mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ k \end{pmatrix} \qquad \text{(i)},$$

where

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ 5 & 4 & 1 \end{pmatrix},$$

stating explicitly the value of k that gives these solutions.

We approach this question using Gaussian elimination.

We write the equation (i) as an augmented matrix, and then use row operations:

$$\begin{aligned}
\left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 2 & 1 & 1 & 1 \\ 5 & 4 & 1 & k \end{array}\right) &\to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -3 & 3 & -3 \\ 0 & -6 & 6 & k-10 \end{array}\right) && \mathbf{r}_2 \to \mathbf{r}_2 - 2\mathbf{r}_1;\ \mathbf{r}_3 \to \mathbf{r}_3 - 5\mathbf{r}_1 \\
&\to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & -3 & 3 & -3 \\ 0 & 0 & 0 & k-4 \end{array}\right) && \mathbf{r}_3 \to \mathbf{r}_3 - 2\mathbf{r}_2 \\
&\to \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & k-4 \end{array}\right) && \mathbf{r}_2 \to -\tfrac{1}{3}\mathbf{r}_2
\end{aligned}$$

For the final row to be consistent (it reads: $0x+0y+0z = k-4$), we require $k-4=0$, so $k=4$. In this case, the second row then gives $y-z=1$, so $y=z+1$, and the first row gives $x+2y-z=2$, so $x=-z$. Thus all possible solutions are given by $x=t$, $y=1-t$, $z=-t$ for arbitrary t.

The equations

$$\mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} \qquad \text{and} \qquad \mathbf{A}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

may each be regarded as the equations of three planes; give a geometrical interpretation of your solution of equation (i) in terms of these sets of planes.

No pair of planes involved in either case is parallel, as no two rows of $\mathbf{A}$ are multiples of each other. So in the first case, the three planes all share a common line. In the second case, where there are no solutions, each pair of planes meets in a distinct line, and the three lines are parallel to each other, forming an infinite triangular prism.

Thinking about the original equations, the first two planes (given by the first two rows of the matrix equation) are fixed, whereas the third plane is translated as the value of k changes.

Write down the equation of the line, L, of intersection of the planes

$$x + 2y - z = 2 \quad \text{and} \quad 2x + y + z = 1.$$

From our earlier answers, this is the line parametrised by $x = t$, $y = 1 - t$, $z = -t$. In vector form, we could write this line as $\mathbf{r} = \mathbf{j} + t(\mathbf{i} - \mathbf{j} - \mathbf{k})$.

Find the equation of the plane through $(0, 1, 0)$ perpendicular to L and the coordinates of the points where this plane intersects the lines of intersection of

$$x + 2y - z = 2 \quad \text{and} \quad 5x + 4y + z = 3$$

and

$$2x + y + z = 1 \quad \text{and} \quad 5x + 4y + z = 3.$$

The plane has vector equation $\mathbf{r}.\mathbf{n} = \mathbf{j}.\mathbf{n}$ where $\mathbf{n}$ is a vector normal to the plane; we can take $\mathbf{n} = \mathbf{i} - \mathbf{j} - \mathbf{k}$ (the direction vector of L), so the plane has vector equation $(\mathbf{i} - \mathbf{j} - \mathbf{k}).\mathbf{r} = -1$, and hence cartesian equation $x - y - z = -1$.

To find the points where this plane intersects the lines of intersection of these pairs of planes, we will have to solve three simultaneous equations once more.

More efficiently, we can solve the pair of simultaneous equations $x - y - z = -1$ and $5x + 4y + z = 3$ to obtain a line, and the find the points of intersection of this line with the other two planes.

We can solve

$$\begin{aligned} x - y - z &= -1 \\ 5x + 4y + z &= 3 \end{aligned}$$

by adding the equations to obtain $6x + 3y = 2$, so $x = t$, $y = \frac{1}{3}(2 - 6t)$, $z = x - y + 1 = 3t + \frac{1}{3}$.

Substituting this into $x + 2y - z = 2$ gives

$$t + \tfrac{2}{3}(2 - 6t) - (3t + \tfrac{1}{3}) = 2$$

and so $-6t = 1$, hence $t = -\frac{1}{6}$ and the first intersection point is $(-\frac{1}{6}, 1, -\frac{1}{6})$.

Substituting this into $2x + y + z = 1$ gives

$$2t + \tfrac{1}{3}(2 - 6t) + (3t + \tfrac{1}{3}) = 1$$

and so $3t = 0$, hence $t = 0$ and the second intersection point is $(0, \frac{2}{3}, \frac{1}{3})$.

Question 15

(i) Given $\mathbf{M} = \begin{pmatrix} k & 1 \\ 0 & k \end{pmatrix}$, calculate $\mathbf{M}^2$ and $\mathbf{M}^3$.

Suggest a form for $\mathbf{M}^n$ and confirm your suggestion, using the method of proof by induction.

$$\mathbf{M}^2 = \begin{pmatrix} k^2 & 2k \\ 0 & k^2 \end{pmatrix}$$
$$\mathbf{M}^3 = \begin{pmatrix} k^3 & 3k^2 \\ 0 & k^3 \end{pmatrix}$$

Following this pattern, we could suggest that

$$\mathbf{M}^n = \begin{pmatrix} k^n & nk^{n-1} \\ 0 & k^n \end{pmatrix}$$

We see from our calculations that this holds for $n = 1$, 2 and 3. Assuming that it holds for $n = r$, then we have

$$\mathbf{M}^{r+1} = \begin{pmatrix} k & 1 \\ 0 & k \end{pmatrix} \begin{pmatrix} k^r & rk^{r-1} \\ 0 & k^r \end{pmatrix} = \begin{pmatrix} k^{r+1} & (r+1)k^r \\ 0 & k^{r+1} \end{pmatrix}$$

which shows that it holds for $n = r + 1$. Therefore the suggestion holds for all positive integer n by induction.

(ii) Prove that, for any $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$,

$$\mathbf{AB} = \mathbf{BA}$$

if and only if $(\mathbf{A} - k\mathbf{I})(\mathbf{B} - k\mathbf{I}) = (\mathbf{B} - k\mathbf{I})(\mathbf{A} - k\mathbf{I})$ for all values of the real number k.

This question asks us to prove that something is true *if and only if* something else is true. So we have to prove both directions.

We start by expanding and simplifying the second equation. We have, for any specific value of k,

$$\begin{aligned} & (\mathbf{A} - k\mathbf{I})(\mathbf{B} - k\mathbf{I}) = (\mathbf{B} - k\mathbf{I})(\mathbf{A} - k\mathbf{I}) \\ \iff \quad & \mathbf{AB} - k\mathbf{A} - k\mathbf{B} + k^2\mathbf{I} = \mathbf{BA} - k\mathbf{A} - k\mathbf{B} + k^2\mathbf{I} \\ \iff \quad & \mathbf{AB} = \mathbf{BA}. \end{aligned}$$

So if $(\mathbf{A} - k\mathbf{I})(\mathbf{B} - k\mathbf{I}) = (\mathbf{B} - k\mathbf{I})(\mathbf{A} - k\mathbf{I})$ for *any* value of k, then $\mathbf{AB} = \mathbf{BA}$, so certainly if $(\mathbf{A} - k\mathbf{I})(\mathbf{B} - k\mathbf{I}) = (\mathbf{B} - k\mathbf{I})(\mathbf{A} - k\mathbf{I})$ for *all* values of k, then $\mathbf{AB} = \mathbf{BA}$. Conversely, if $\mathbf{AB} = \mathbf{BA}$, then for each value of k, $(\mathbf{A} - k\mathbf{I})(\mathbf{B} - k\mathbf{I}) = (\mathbf{B} - k\mathbf{I})(\mathbf{A} - k\mathbf{I})$, so this holds for all values of k.

(iii) Prove that, for any $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$, where $\mathbf{A}^{\mathrm{T}}$ is the transpose of $\mathbf{A}$.

We will actually prove this result for any matrices $\mathbf{A}$ and $\mathbf{B}$ which are conformable for multiplication. The specific case of $n \times n$ matrices follows immediately. We will need an expression for the (i, j)th element of the product of two matrices. Since we will be multiplying different matrices, we will use matrices $\mathbf{C}$ and $\mathbf{D}$ for this purpose, where we assume them to be conformable for multiplication.

The element $(\mathbf{CD})_{ij}$ is the result of multiplying the ith row of $\mathbf{C}$ by the jth column of $\mathbf{D}$, so we can write

$$(\mathbf{CD})_{ij} = \sum_k c_{ik} d_{kj}$$

where the sum is from $k = 1$ to the number of columns of $\mathbf{C}$, which equals the number of rows of $\mathbf{D}$. (In the particular case of this question, the sum would be $\sum_{k=1}^{n}$.)

Now the (i, j)th element of $(\mathbf{AB})^{\mathrm{T}}$ is the (j, i)th element of $\mathbf{AB}$, so we have

$$\left((\mathbf{AB})^{\mathrm{T}}\right)_{ij} = (\mathbf{AB})_{ji} = \sum_k a_{jk} b_{ki}.$$

(Note the order of i and j in the final sum, because we are looking at the (j, i)th element of $\mathbf{AB}$.)

Likewise, the (i, j)th element of $\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ is

$$(\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}})_{ij} = \sum_k (\mathbf{B}^{\mathrm{T}})_{ik} (\mathbf{A}^{\mathrm{T}})_{kj} = \sum_k b_{ki} a_{jk}$$

since the (r, s)th element of $\mathbf{A}^{\mathrm{T}}$ is the (s, r)th element of $\mathbf{A}$, which is a_{sr}.

Comparing the expressions for the (i, j)th element of $(\mathbf{AB})^{\mathrm{T}}$ and $\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$, we see that they are equal, and hence $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ as required.

In the general case, we actually need to do one extra thing, which is to check that all the sums are over the same indices, and that these two final matrices have the same dimensions as each other. If $\mathbf{A}$ is an $m \times p$ matrix and $\mathbf{B}$ is a $p \times n$ matrix, then the sum in the expression for $\left((\mathbf{AB})^{\mathrm{T}}\right)_{ij}$ is from $k = 1$ to p, and as $\mathbf{AB}$ is an $m \times n$ matrix, $(\mathbf{AB})^{\mathrm{T}}$ is an $n \times m$ matrix. Now $\mathbf{A}^{\mathrm{T}}$ is a $p \times m$ matrix and $\mathbf{B}^{\mathrm{T}}$ is an $n \times p$ matrix, so $\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$ is an $n \times m$ matrix, and the sums in the expression for $(\mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}})_{ij}$ are from $k = 1$ to p. So everything works.

(iv) Prove that, if $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ symmetric matrices, then $\mathbf{AB}$ is symmetric if and only if $\mathbf{AB} = \mathbf{BA}$.

We recall that a matrix is called symmetric if it equals its transpose. This is another *if and only if* question, so we must be careful to prove the result in both directions.

We are given that $\mathbf{A}$ and $\mathbf{B}$ are symmetric.

If $\mathbf{AB}$ is symmetric, then

$$\begin{aligned} \mathbf{AB} &= (\mathbf{AB})^{\mathrm{T}} && \text{the given assumption} \\ &= \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} && \text{using part (iii)} \\ &= \mathbf{BA} && \text{as } \mathbf{A} \text{ and } \mathbf{B} \text{ are symmetric} \end{aligned}$$

so $\mathbf{AB} = \mathbf{BA}$.

Conversely, if $\mathbf{AB} = \mathbf{BA}$, we have

$$\begin{aligned} (\mathbf{AB})^{\mathrm{T}} &= \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} && \text{using part (iii)} \\ &= \mathbf{BA} && \text{as } \mathbf{A} \text{ and } \mathbf{B} \text{ are symmetric} \\ &= \mathbf{AB} && \text{the given assumption} \end{aligned}$$

so $\mathbf{AB}$ is symmetric.

Question 16

Let $\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

(i) The complex number $a + ib$ is represented by the matrix $a\mathbf{I} + b\mathbf{J}$. Show that if w and z are two complex numbers represented by the matrices $\mathbf{P}$ and $\mathbf{Q}$ respectively, then $w + z$ is represented by $\mathbf{P} + \mathbf{Q}$ and wz is represented by $\mathbf{PQ}$.

See the solution to the first part of question 10; it is identical except that $\mathbf{J}$ has been replaced by $-\mathbf{J}$.

There is actually something quite significant about this: there is no way to distinguish between the two square roots of -1, as long as we are consistent about our choice. So we could replace every occurrence of i by $-$i throughout a piece of mathematics, and it would still be perfectly correct.

(ii) The matrices $\mathbf{A}$ and $\mathbf{B}$ are $\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$ and $\begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix}$ respectively.

Express $\mathbf{A}$, $\mathbf{B}$, $\mathbf{B}^{-1}$ and $\mathbf{AB}^{-1}$ in the form $a\mathbf{I} + b\mathbf{J}$.

Write down the complex numbers represented by $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{AB}$ in the form $re^{i\theta}$. Hence, or otherwise, show that

$$\arctan(\tfrac{1}{2}) + \arctan(\tfrac{1}{3}) = \tfrac{1}{4}\pi.$$

We have $\mathbf{A} = 2\mathbf{I} + \mathbf{J}$ and $\mathbf{B} = 3\mathbf{I} + \mathbf{J}$. We can then calculate

$$\mathbf{B}^{-1} = \frac{1}{10}\begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}$$

so

$$\mathbf{AB}^{-1} = \frac{1}{10}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 7 & -1 \\ 1 & 7 \end{pmatrix}$$

so $\mathbf{B}^{-1} = \frac{1}{10}(3\mathbf{I} - \mathbf{J})$ and $\mathbf{AB}^{-1} = \frac{1}{10}(7\mathbf{I} + \mathbf{J})$.

If the complex number represented by $\mathbf{A}$ is w, then $|w| = \sqrt{2^2 + 1^1} = \sqrt{5}$, and $\arg w = \arctan(\frac{1}{2})$, so $w = \sqrt{5}\,e^{i\theta}$, where $\theta = \arctan(\frac{1}{2})$.

Likewise, if the complex number represented by $\mathbf{B}$ is z, then $z = \sqrt{10}\,e^{i\phi}$, where $\phi = \arctan(\frac{1}{3})$.

Finally, the complex number u represented by $\mathbf{AB}$ is the product of these, so $u = wz = 5\sqrt{2}\,e^{i\psi}$, where $\psi = \arctan(\frac{1}{2}) + \arctan(\frac{1}{3})$.

We can also calculate $\mathbf{AB}$ explicitly:

$$\mathbf{AB} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 5 & -5 \\ 5 & 5 \end{pmatrix}$$

so $\mathbf{AB} = 5(\mathbf{I} + \mathbf{J})$; since $\arctan(1) = \frac{1}{4}\pi$, this matrix represents the complex number $5\sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}$.

Comparing these expressions for the complex number represented by $\mathbf{AB}$, we find that

$$\arctan(\tfrac{1}{2}) + \arctan(\tfrac{1}{3}) = \tfrac{1}{4}\pi.$$

Question 17

Show that a matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

where $a_{11} \neq 0$ and $a_{11}a_{22} - a_{12}a_{21} \neq 0$, may be decomposed into the product of a *lower* triangular form matrix $\mathbf{L}$ and an *upper* triangular form matrix $\mathbf{U}$ such that $\mathbf{A} = \mathbf{LU}$ where

$$\mathbf{L} = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & l & 0 \\ a_{31} & m & n \end{pmatrix} \quad \text{and} \quad \mathbf{U} = \begin{pmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{pmatrix}.$$

If we multiply out $\mathbf{LU}$, we get

$$\mathbf{LU} = \begin{pmatrix} a_{11} & pa_{11} & qa_{11} \\ a_{21} & pa_{21} + l & qa_{21} + lr \\ a_{31} & pa_{31} + m & qa_{31} + mr + n \end{pmatrix}.$$

Comparing this to $\mathbf{A}$, we can choose values for l, m, n, p, q and r so that $\mathbf{A} = \mathbf{LU}$ as follows:

$$\begin{aligned} p &= \frac{a_{12}}{a_{11}} \\ q &= \frac{a_{13}}{a_{11}} \\ l &= a_{22} - pa_{21} \\ m &= a_{32} - pa_{31} \\ r &= \frac{a_{23} - qa_{21}}{l} \\ n &= a_{33} - qa_{31} - mr. \end{aligned}$$

Note that since

$$l = a_{22} - pa_{21} = a_{22} - \frac{a_{12}a_{21}}{a_{11}} = \frac{a_{11}a_{22} - a_{12}a_{21}}{a_{11}} \neq 0,$$

the formula for r does not involve division by zero.

A system of simultaneous linear equations, $\mathbf{Ax} = \mathbf{c}$, may be solved by writing the equations as $\mathbf{LUx} = \mathbf{c}$ and letting $\mathbf{Ux} = \mathbf{u}$. The vector $\mathbf{u}$ is determined from $\mathbf{Lu} = \mathbf{c}$ and the solution $\mathbf{x}$ is then found from $\mathbf{Ux} = \mathbf{u}$. Since both $\mathbf{L}$ and $\mathbf{U}$ are triangular, these two sets of equations can be solved directly by back substitution.

Use this method to solve

$$\begin{aligned} x + y - z &= 2, \\ 3x + 2y + 5z &= 1, \\ 4x - y + 2z &= 0. \end{aligned}$$

We have, in this case,

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & -1 \\ 3 & 2 & 5 \\ 4 & -1 & 2 \end{pmatrix}; \qquad \mathbf{c} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

Using the above calculations, we obtain

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & -1 & 0 \\ 4 & -5 & -34 \end{pmatrix} \quad \text{and} \quad \mathbf{U} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -8 \\ 0 & 0 & 1 \end{pmatrix}.$$

(It is definitely worth checking that $\mathbf{A} = \mathbf{LU}$ at this point, to ensure that our calculations are correct.)

We now solve $\mathbf{Lu} = \mathbf{c}$, so

$$\begin{pmatrix} 1 & 0 & 0 \\ 3 & -1 & 0 \\ 4 & -5 & -34 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

We use back substitution: the first row gives $u = 2$, then the second row gives $3u - v = 1$, so $v = 5$, and finally $4u - 5v - 34w = 0$, giving $w = -\frac{1}{2}$.

We finally solve $\mathbf{Ux} = \mathbf{u}$, so

$$\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -8 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ -\frac{1}{2} \end{pmatrix}.$$

Using back substitution, we have $z = -\frac{1}{2}$, $y - 8z = 5$ so $y = 1$, and $x + y - z = 2$, so $x = \frac{1}{2}$, and we have solved our equations.

It is wise to substitute this solution back into the original equations as a check.

Acknowledgements

The exam questions are reproduced by kind permission of Cambridge Assessment Group Archives. Unless otherwise noted, the questions are reproduced verbatim, except for some differences in convention between the papers as printed and STEP, specifically: the use of j to represent $\sqrt{-1}$ has been replaced by i; variables representing matrices are written in boldface type rather than italics; transpose is denoted $\mathbf{A}^{\mathrm{T}}$ rather than A'; derivatives and integrals are written using a roman d rather than an italic d.

In the list of sources below, the following abbreviations are used:

O&C	Oxford and Cambridge Schools Examination Board
SMP	School Mathematics Project
MEI	Mathematics in Education and Industry
QP	Question paper
Q	Question

1 UCLES, A level Mathematics, 1951, QP 190, Further Mathematics I, Q 2

2 UCLES, A level Mathematics, 1953, QP 188, Further Mathematics I, Q 2

3 UCLES, A level Mathematics, 1954, QP 188, Further Mathematics IV (Scholarship Paper), Q 1

4 UCLES, A level Mathematics, 1958, QP 437/1, Further Mathematics I, Q 1

5 O&C, A level Mathematics (SMP), 1968, QP SMP 34*, Mathematics II, Q 9; editorial change: clarify the last part of the question by adding the 'for all points P' part.

6 O&C, A level Mathematics (SMP), 1968, QP SMP 35, Mathematics III (Special Paper), Q 10; editorial changes here: the term 'order' is replaced by 'dimension'; the term 'frame of reference' has been replaced by 'coordinate system', and a diagram has been introduced to clarify the idea

7 O&C, A level Mathematics (MEI), 1969, QP MEI 32, Applied Mathematics III (Special Paper), Q 13; editorial changes: explained the dot notation, corrected a typographical error in the differential equation (missing dot) and made parentheses consistent

8 O&C, A level Mathematics (SMP), 1970, QP SMP 58, Further Mathematics IV, Q 6

9 UCLES, A level Mathematics, 1971, QP 842/0, Pure Mathematics (Special Paper), Q1

10 O&C, A level Mathematics (MEI), 1971, QP MEI 54, Pure Mathematics II, Q 4

11 UCLES, A level Mathematics, 1972, QP 848/0, Mathematics 0 (Special Paper), Q 12; editorial change: the matrix is called **A** rather than **a**.

12 UCLES, A level Mathematics, 1972, QP 852/0, Further Mathematics 0 (Special Paper), Q 17

13 O&C, A level Mathematics (MEI), 1973, QP MEI 84, Pure Mathematics II, Q 4; editorial change: the row vector has have been replaced by a column vector

14 O&C, A level Mathematics (MEI), 1976, QP 128, Pure Mathematics I, Q 3

15 O&C, A level Mathematics (MEI), 1982, QP 9655/1, Pure Mathematics 1, Q 8

16 O&C, A level Mathematics (MEI), 1985, QP 9658/1, Further Mathematics 1, Q 1(a); editorial change: include a preliminary part showing that we can represent complex numbers in the form $a\mathbf{I} + b\mathbf{J}$ (and then remove the definitions of $\mathbf{I}$ and $\mathbf{J}$ from the main question)

17 O&C, A level Mathematics (MEI), 1988, QP 9658/0, Further Mathematics 0 (Special Paper), Q 9; editorial change: explain what the term 'back substitution' means in a footnote.