Sketch Solutions for Exercises in the Main Text of A First Look at Vectors

T. W. Körner

Introduction

Here are what I believe to be sketch solutions to the bulk of exercises in the main text the book (i.e. those not in the "Further Exercises"). I have written in haste in the hope that others will help me correct at leisure. I am sure that they are stuffed with errors ranging from the TEXtual through to the arithmetical and not excluding serious mathematical mistakes. I would appreciate the opportunity to correct at least some of these problems. Please tell me of any errors, unbridgeable gaps, misnumberings etc. I welcome suggestions for additions.

ALL COMMENTS GRATEFULLY RECEIVED.

If you can, please use LATEX 2_ε or its relatives for mathematics. If not, please use plain text. My e-mail is **twk@dpmms.cam.ac.uk**. You may safely assume that I am both lazy and stupid so that a message saying 'Presumably you have already realised the mistake in Exercise Z' is less useful than one which says 'I think you have made a mistake in Exercise Z because you have have assumed that the sum is necessarily larger than the integral. One way round this problem is to assume that f is decreasing.'

When I was young, I used to be surprised when the answer in the back of the book was wrong. I could not believe that the wise and gifted people who wrote textbooks could possibly make mistakes. I am no longer surprised.

It may be easiest to navigate this document by using the table of contents which follow on the next few pages. To avoid disappointment, observe that those exercises marked ★ have no solution given.

Contents

Exercise 1.1.2

Starting with

$$\begin{aligned} x + y + z &= 1 \\ x + 2y + 3z &= 2 \\ x + 4y + 9z &= 6. \end{aligned}$$

We subtract the first equation from the second and from the third to get

$$\begin{aligned} x + y + z &= 1 \\ y + 2z &= 1 \\ 3y + 8z &= 5. \end{aligned}$$

We now solve the equations

$$\begin{aligned} y + 2z &= 1 \\ 3y + 8z &= 5, \end{aligned}$$

for z by subtracting 3 times the first equation from the second to obtain

$$\begin{aligned} y + 2z &= 2 \\ 2z &= 2. \end{aligned}$$

Thus $z = 1$ and, working backwards, $y = 1 - 2z = -1$, whence $x = 1 - y - z = 1$.

Exercise 1.2.1

STEP 1 If $a_{ij} = 0$ for all i and j, then our equations have the form

$$0 = y_i \qquad [1 \leq i \leq m].$$

Our equations are inconsistent unless $y_1 = y_2 = \ldots = y_m = 0$. If $y_1 = y_2 = \ldots = y_m = 0$ the equations impose no constraints on x_1, x_2, $\ldots$, x_n which can take any value we want.

STEP 2 If the condition of STEP 1 does not hold, we can arrange,by reordering the equations and the unknowns, if necessary, that $a_{11} \neq 0$. We now subtract a_{i1}/a_{11} times the first equation from the ith equation $[2 \leq i \leq m]$ to obtain

$$\bigstar\bigstar \qquad \sum_{j=2}^{n} b_{ij}x_j = z_i \qquad [2 \leq i \leq m].$$

where

$$b_{ij} = \frac{a_{11}a_{ij} - a_{i1}a_{1j}}{a_{11}} \text{ and } z_i = \frac{a_{11}y_i - a_{i1}y_1}{a_{11}}.$$

STEP 3 If the new set of equations ★★ has no solution, then our old set ★ has no solution. If our new set of equations ★★ has a solution $x_i = x_i'$ for $2 \leq i \leq m$, then our old set ★ has the solution

$$x_1 = \frac{1}{a_{11}}\left(y_1 - \sum_{j=2}^{n} a_{1j}x_j'\right)$$

$$x_i = x_i' \qquad [2 \leq i \leq n].$$

(Continued in Exercise 1.2.2.)

Exercise 1.2.2

(Continued from Exercise 1.2.1.)

This means that, if ★★ has exactly one solution, then ★ has exactly one solution and, if ★★ has infinitely many solutions, then ★ has infinitely many solutions. We have already remarked that, if ★★ has no solutions, then ★ has no solutions.

EXERCISE 1.2.3

(i)(a) If $a_i = 0$, but $y_i \neq 0$ for some i, then there can be no solution, since the equation $0 = y_i$ cannot be satisfied.

(b) If $y_r = 0$ whenever $a_r = 0$ and there exists an i such that $a_i = 0$, then there are an infinity of solutions with $x_s = a_s^{-1} y_s$ when $a_s \neq 0$ and x_s chosen freely otherwise.

(c) If $a_i \neq 0$ for all i, there is a unique solution $x_i = a_i^{-1} y_i$ for all i.

(ii) (a) If $a_j = 0$ for all j then either $b = 0$ and every choice of x_j gives a solution (so there is an infinity of solutions or $b \neq 0$ and there is no solution.

(b) If $a_k \neq 0$ for some k, then, choosing x_i freely for $i \neq k$ and setting

$$x_k = a_k^{-1} \left(b - \sum_{j \neq k} a_j x_j \right)$$

gives an infinity of solutions.

Exercise 1.2.5

(i) If $a = 1$, $b = 2$, $c = d = 4$ the first and second equations are incompatible, so there is no solution.

(ii) If $a = 1$, $b = 2$, $c = d = 4$, then the third equation gives the same information as the first equation, so the system reduces to

$$x + y = 2$$
$$x + 2y = 4$$

Subtracting the first equation from the second we see that $y = 2$. Thus $x = 0$. By inspection this is a solution.

(iii) If $a = b = 2$, $c = d = 4$, then the second and third equation give the same information as the first equation so the system reduces to

$$x + y = 1$$

with the infinite set of solutions given by choosing x arbitrarily and setting $y = 1 - x$.

Exercise 1.2.6

The two equations are incompatible if $a = 1$. There is then no solution.

If $a \neq 1$, then subtracting the first equation from the second we get

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

so $z = 2/(a-1)$. Knowing the value of z the system reduces to

$$x + y = 2 - \frac{2}{a-1} = 2\frac{a-2}{a-1}$$

so x may be chosen freely and then

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

There are an infinity of solutions.

Exercise 1.2.7

(i) Observe that, if $n \geq 4$

$$n^3 \geq \sum_{r=1}^{n} n^2 \geq \sum_{r=1}^{n} r^2 \geq \sum_{n \geq r \geq n/2}^{n} (n/2)^2 \geq (n/4) \times (n/2)^2 = n^3/16.$$

(ii) Let $f(x) = x^2$. Then f is increasing so

$$f(r) \leq f(x) \leq f(r+1)$$

for $r \leq x \leq r+1$.

Integrating we get

$$f(r) = \int_r^{r+1} f(r)\,dx \leq \int_r^{r+1} f(x)\,dx \leq \int_r^{r+1} f(r+1)\,dx = f(r+1)$$

Thus, summing,

$$\sum_{r=1}^{n-1} f(r) \leq \int_1^n f(x)\,dx \leq \sum_{r=2}^{n} f(r).$$

In other words

$$\sum_{r=1}^{n-1} r^2 \leq (n^3-1)/3 \leq \sum_{r=2}^{n} r^2$$

so

$$\left(\sum_{r=1}^{n} r^2\right) - n^2 \leq (n^3-1)/3 \leq \left(\sum_{r=1}^{n} r^2\right) - 1$$

Thus

$$(n^3-1)/3 \leq \left(\sum_{r=1}^{n} r^2\right) - 1 \leq (n^3-1)/3 + n^2.$$

Dividing by n^3 and allowing $n \to \infty$, we obtain

$$n^{-3} \sum_{r=1}^{n} r^2 \to 1/3.$$

Exercise 1.2.8

If we have the triangular system of size n, one operation is needed to get x_n from the nth equation. Substitution of the value of x_n in the remaining $n-1$ equations reduces them to a triangular system of size $n-1$ in about $2(n-1)$ operations. We can repeat this process to obtain the complete solution in about

$$\sum_{r=1}^{n} 2(r-1)+1 \approx 2\sum_{r=1}^{n-1}(r-1) \approx n^2$$

operations.

Exercise 1.2.9★

Rather you than me.

Exercise 1.2.10

The number of operations is about An^3. I reckon $A \approx 1$, so I need about 1000 operations. Reckoning about 100 operations an hour and a 5 hour day gives 2 days. The reader may disagree about everything, but we should still have about the same order of magnitude for the task.

Exercise 1.3.5

Row operations

Subtract twice first row from second row.

$$\begin{pmatrix} 1 & -1 & 3 \\ 2 & 5 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 3 \\ 0 & 6 & -4 \end{pmatrix}$$

Divide second row by 6.

$$\begin{pmatrix} 1 & -1 & 3 \\ 0 & 6 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & -\frac{4}{3} \end{pmatrix}$$

Add second row to first row.

$$\begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & -\frac{4}{3} \end{pmatrix} \to \begin{pmatrix} 1 & 0 & \frac{5}{3} \\ 0 & 1 & -\frac{4}{3} \end{pmatrix}$$

Column operations.

Add first column to second. Subtract three times first column from third.

$$\begin{pmatrix} 1 & -1 & 3 \\ 2 & 5 & 2 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & -10 \end{pmatrix}$$

Divide second column by 7.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & -10 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{10}{7} \end{pmatrix}$$

Add 10/7 times second column to third.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{10}{7} \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Exercise 1.3.9

(i) Subtract 2 times first row from second. Subtract 4 times first row from third.

$$\begin{pmatrix}1 & -1 & 3\\ 2 & 5 & 2\\ 4 & 3 & 8\end{pmatrix} \to \begin{pmatrix}1 & -1 & 3\\ 0 & 7 & -4\\ 0 & 7 & -4\end{pmatrix}$$

Subtract second row from third.

$$\begin{pmatrix}1 & -1 & 3\\ 0 & 7 & -4\\ 0 & 7 & -4\end{pmatrix} \to \begin{pmatrix}1 & -1 & 3\\ 0 & 7 & -4\\ 0 & 0 & 0\end{pmatrix}$$

Add first column to second. Subtract 3 times first column from third.

$$\begin{pmatrix}1 & -1 & 3\\ 0 & 7 & -4\\ 0 & 0 & 0\end{pmatrix} \to \begin{pmatrix}1 & 0 & 0\\ 0 & 7 & -4\\ 0 & 0 & 0\end{pmatrix}$$

Add 4/7 times second column to third. Divide second column by 7.

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 7 & -4\\ 0 & 0 & 0\end{pmatrix} \to \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}$$

(ii) Subtract first row from second and interchange first and second row.

$$\begin{pmatrix}2 & 4 & 5\\ 3 & 2 & 1\\ 4 & 1 & 3\end{pmatrix} \to \begin{pmatrix}2 & 4 & 5\\ 1 & -2 & -4\\ 4 & 1 & 3\end{pmatrix} \to \begin{pmatrix}1 & -2 & -4\\ 2 & 4 & 5\\ 4 & 1 & 3\end{pmatrix}$$

Subtract 2 times first row from second and 4 times first row from third.

$$\begin{pmatrix}1 & -2 & -4\\ 2 & 4 & 5\\ 4 & 1 & 3\end{pmatrix} \to \begin{pmatrix}1 & -2 & -4\\ 0 & 8 & 13\\ 0 & 9 & 19\end{pmatrix}$$

Subtract second row from third, interchange second and third.

$$\begin{pmatrix}1 & -2 & -4\\ 0 & 8 & 13\\ 0 & 9 & 19\end{pmatrix} \to \begin{pmatrix}1 & -2 & -4\\ 0 & 8 & 13\\ 0 & 1 & 6\end{pmatrix} \to \begin{pmatrix}1 & -2 & -4\\ 0 & 1 & 6\\ 0 & 8 & 13\end{pmatrix}$$

Add twice second row to first and subtract 8 times second row from third.

$$\begin{pmatrix}1 & -2 & -4\\ 0 & 1 & 6\\ 0 & 8 & 13\end{pmatrix} \to \begin{pmatrix}1 & 0 & 8\\ 0 & 1 & 6\\ 0 & 0 & -35\end{pmatrix}$$

Divide third row by -35. Subtract 6 times third row from second and 8 times third row from first.

$$\begin{pmatrix}1 & 0 & 8\\ 0 & 1 & 6\\ 0 & 0 & -35\end{pmatrix} \to \begin{pmatrix}1 & 0 & 8\\ 0 & 1 & 6\\ 0 & 0 & 1\end{pmatrix} \to \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

Subtract first row from second and interchange first and second row.

$$\begin{array}{ccc} 2x+4y+5z=-3 & 2x+4y+5z=-3 & x-2y-4z=5\\ 3x+2y+z=2 & x-2y-4z=5 & 2x+4y+5z=-3\\ 4x+y+3z=1 & 4x+y+3z=1 & 4x+y+3z=1 \end{array}$$

Subtract 2 times first row from second and 4 times first row from third. Subtract second row from third, interchange second and third

$$\begin{array}{ccc} x-2y-4z=5 & x-2y-4z=5 & x-2y-4z=5\\ 8y+13z=-13 & 8y+13z=-13 & y+6z=-6\\ 9y+19z=-19 & y+6z=-6 & 8y+13z=-13 \end{array}$$

Add twice second row to first and subtract 8 times second row from third. Divide third row by -35. Subtract 6 times third row from second and 8 times third row from first.

$$\begin{array}{ccc} x \quad +8z=-7 & x \quad +8z=-7 & x=1\\ y+6z=-6 & y+6z=-6 & y=0\\ -35z=35 & z=-1 & z=-1 \end{array}$$

Exercise 1.3.10

By subtracting the second row from the first and the third from the second, we see that the matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

has rank 3. By adding the first row to the second and the first row to the third, we obtain a matrix of the same rank 3 with all entries non-zero.

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{pmatrix}.$$

The same arguments, starting with

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

yields matrices of rank 2 and 1 with all entries non-zero

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

Not possible. All elementary operations (which change a matrix) involve a non-zero row or column which remains non-zero (though possibly moved) after the operation. Thus a rank 0 matrix must be the zero matrix.

Exercise 1.4.3

(i) Observe that

$$\pi_i\big((\mathbf{x}+\mathbf{y})+\mathbf{x}\big) = (x_i+y_i)+z_i = x_i+(y_i+z_i) = \pi_i\big(\mathbf{x}+(\mathbf{y}+\mathbf{x})\big).$$

(ii) Observe that

$$\pi_i(\mathbf{x}+\mathbf{y}) = x_i+y_i = y_i+x_i = \pi_i(\mathbf{y}+\mathbf{x}).$$

(iii) Observe that

$$\pi_i(\mathbf{x}+\mathbf{0}) = x_i+0 = x_i = \pi_i(\mathbf{x}).$$

(v) Observe that

$$\pi_i\big((\lambda+\mu)\mathbf{x}\big) = (\lambda+\mu)x_i = \lambda x_i+\mu x_i = \pi_i(\lambda\mathbf{x}+\mu\mathbf{y}).$$

(vi) Observe that

$$\pi_i\big((\lambda\mu)\mathbf{x}\big) = (\lambda\mu)x_i = \lambda(\mu x_i) = \pi_i\big(\lambda(\mu\mathbf{y})\big).$$

(vii) Observe that

$$\pi_i(1\mathbf{x}) = 1x_i = x_i = \pi_i(\mathbf{x})$$

and

$$\pi_i(0\mathbf{x}) = 0x_i = 0 = \pi_i(\mathbf{0}).$$

(viii) Can do as above or

$$\mathbf{x}-\mathbf{x} = 1\mathbf{x}+(-1)\mathbf{x} = \big(1+(-1)\big)\mathbf{x} = 0\mathbf{x} = \mathbf{0}.$$

Exercise 2.1.3

If $c \neq 0$, then $\mathbf{u} = (a/c, 0)$ and $\mathbf{v} = (0, b/c)$ are distinct vectors representing points on the line.

If $c = 0$, then suppose, without loss of generality that $a \neq 0$. If $b = 0$, then $\mathbf{u} = (0, 0)$ and $\mathbf{v} = (0, 1)$ are distinct vectors representing points on the line. If $b \neq 0$, then $\mathbf{u} = (0, 0)$ and $\mathbf{v} = (1, -a/b)$ are distinct vectors representing points on the line.

Exercise 2.1.4

(i) Suppose that

$$\{\mathbf{v} + \lambda\mathbf{w} \,:\, \lambda \in \mathbb{R}\} \cap \{\mathbf{v}' + \mu\mathbf{w} \,:\, \mu \in \mathbb{R}\} \neq \emptyset.$$

Then there exist λ_0, $\mu_0 \in \mathbb{R}$ such that

$$\mathbf{v} + \lambda_0\mathbf{w} = \mathbf{v}' + \mu_0\mathbf{w}$$

and so

$$\mathbf{v} + \lambda\mathbf{w} = \mathbf{v}' + (\lambda - \lambda_0 + \mu_0)\mathbf{w}.$$

Thus

$$\{\mathbf{v} + \lambda\mathbf{w} \,:\, \lambda \in \mathbb{R}\} \subseteq \{\mathbf{v}' + \mu\mathbf{w} \,:\, \mu \in \mathbb{R}\}$$

and, similarly

$$\{\mathbf{v}' + \mu\mathbf{w} \,:\, \mu \in \mathbb{R}\} \subseteq \{\mathbf{v} + \lambda\mathbf{w} \,:\, \lambda \in \mathbb{R}\}$$

so

$$\{\mathbf{v} + \lambda\mathbf{w} \,:\, \lambda \in \mathbb{R}\} = \{\mathbf{v}' + \mu\mathbf{w} \,:\, \mu \in \mathbb{R}\}.$$

(ii) The line joining $\mathbf{u}$ to $\mathbf{u}'$ is

$$\{\mathbf{u} + \lambda\mathbf{w} \,:\, \lambda \in \mathbb{R}\}$$

with $\mathbf{w} = \mathbf{u} - \mathbf{u}'$. Since $\mathbf{w} = (\sigma/\tau)\mathbf{v} - \mathbf{v}'$, the line joining $\mathbf{v}$ to $\mathbf{v}'$ is

$$\{\mathbf{v} + \mu\mathbf{w} \,:\, \mu \in \mathbb{R}\}.$$

Thus part (ii) follows from part (i).

(iii) Observe that

$$\mathbf{u}'' = -\frac{1}{\mu''}(\mu\mathbf{u} + \mu'\mathbf{u}') = \frac{1}{\mu + \mu'}(\mu\mathbf{u} + \mu'\mathbf{u}') = \mathbf{u} + \frac{\mu'}{\mu + \mu'}(\mathbf{u}' - \mathbf{u}).$$

Exercise 2.1.6★

Exercise 2.1.7★

Exercise 2.1.9

$$\frac{\lambda}{1-\lambda} = -\frac{\alpha}{\beta} \Rightarrow \alpha\lambda = -\alpha + \lambda\alpha \Rightarrow \lambda = \frac{\alpha}{\alpha - \beta}.$$

Interchanging α and $-(1-\alpha)$ and β and $-(1-\beta)$, yields

$$\lambda' = \frac{1-\alpha}{\beta-\alpha}.$$

Exercise 2.2.4★

Exercise 2.2.5★

Exercise 2.2.8

Let $\mathbf{y}$ be the centroid of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_q$.

Observe that
$$\frac{q-1}{q}\mathbf{y}_j + \frac{1}{q}\mathbf{x}_j = \frac{1}{q}\sum_{i \neq j}\mathbf{x}_i + \frac{1}{q}\mathbf{x}_j = \mathbf{y},$$
so $\mathbf{y}$ lies on the line joining $\mathbf{y}_j$ and $\mathbf{x}_j$.

Exercise 2.2.10

(i) Let $\mathbf{y}$ be centre of mass of all the points $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_q$ and let $M = \sum_{j=1}^{q} m_j$.

Observe that

$$\frac{m_j}{M}\mathbf{y}_j + \frac{M - m_j}{M}\mathbf{x}_j = \frac{1}{M}\sum_{i \neq j} m_i \mathbf{x}_i + \frac{m_j}{M}\mathbf{x}_j = \mathbf{y}$$

so $\mathbf{y}$ lies on the line joining $\mathbf{y}_j$ and $\mathbf{x}_j$.

(ii) We might have $m_1 + m_2 + \ldots + m_q = 0$ and we cannot divide by 0.

Exercise 2.3.3

(i) We take the positive square root.

(ii) $\mathbf{x} = 0 \Rightarrow \|\mathbf{x}\|^2 = 0 \Rightarrow \sum_{j=1}^n x_j^2 = 0 \Rightarrow x_j = 0$ $\forall$j$\Rightarrow$ $\mathbf{x} = \mathbf{0}$.

(iii) $\|\lambda\mathbf{x}\|^2 = \sum_{j=1}^n (\lambda x_j)^2 = \lambda^2 \sum_{j=1}^n x_j^2$, so $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$.

Exercise 2.3.4★

Exercise 2.3.11

The result is trivial if $\mathbf{a} = \mathbf{b}$, or $\mathbf{b} = \mathbf{c}$, so we assume this is not the case.

Setting $\mathbf{x} = \mathbf{a} - \mathbf{b}$, $\mathbf{y} = \mathbf{b} - \mathbf{c}$, we see that Theorem 2.3.10 gives

$$\begin{aligned}\|\mathbf{a} - \mathbf{b}\| + \|\mathbf{b} - \mathbf{c}\| &= \|\mathbf{x}\| + \|\mathbf{y}\|\\ &\geq \|\mathbf{x} + \mathbf{y}\|\\ &= \|\mathbf{a} - \mathbf{c}\|\end{aligned}$$

with equality if and only if $\lambda(\mathbf{a} - \mathbf{b}) = \mu(\mathbf{b} - \mathbf{c})$ for some λ, $\mu > 0$ i.e. if and only if $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ lie in order along a line.

The length of one side of a triangle AB is less than or equal to the sum of the lengths of the other two sides BC and CA with equality if and only if the triangle is degenerate with A, B, C lying in order along a straight line.

Exercise 2.3.12

If $\mathbf{x}$ lies on the line through $\mathbf{a}$ and $\mathbf{b}$ then

$$\mathbf{x} = t\mathbf{a} + (1-t)\mathbf{b}$$

for some $t \in \mathbb{R}$. The condition

$$\|\mathbf{x} - \mathbf{a}\| = \|\mathbf{x} - \mathbf{b}\|$$

yields

$$\|(1-t)(\mathbf{a} - \mathbf{b})\| = \|t(\mathbf{a} - \mathbf{b})\|$$

that is to say

$$|1-t|\|\mathbf{b} - \mathbf{a}\| = |t|\|\mathbf{b} - \mathbf{a}\|$$

so $|t| = |1-t|$ so $t = 1-t$ or $t = t-1$. The second equation is insoluble so $t = 1/2$ and $\mathbf{x} = \frac{1}{2}(\mathbf{a} + \mathbf{b})$.

We have proved uniqueness. Direct substitution shows we have a solution.

Exercise 2.3.13

(i) Choose X on BC so that AX is perpendicular to BC. By the theorem of Pythagoras and the geometric definition of the cosine and sine

$$\begin{aligned}|AC|^2 &= |AX|^2 + |CX|^2 = |AX|^2 + (|BC| - |BX|)^2\\ &= |BA|^2 \sin^2\theta + (|BC| - |BA|\cos\theta)^2\\ &= |BA|^2(\cos^2\theta + \sin^2\theta) + |BC|^2 - 2|BA||BC|\cos\theta\\ &= |BC|^2 + |BA|^2 - 2|BC| \times |BA|\cos\theta\end{aligned}$$

(If the reader is a very careful mathematician, she will not be entirely happy with this argument. She will ask, for example, how we choose the sign of some of the terms. However in this book the algebra is *primary* and the geometry is *illustrative*.)

(ii) We have

$$\|\mathbf{a} - \mathbf{c}\|^2 = (\mathbf{a} - \mathbf{c}) \cdot (\mathbf{a} - \mathbf{c}) = \|\mathbf{a}\|^2 + \|\mathbf{c}\|^2 - 2\mathbf{a} \cdot \mathbf{c}.$$

(iii) The formula in (i) can be rewritten

$$\|\mathbf{a}\|^2 + \|\mathbf{c}\|^2 - 2\|\mathbf{a}\| \times \|\mathbf{c}\|\cos\theta = \|\mathbf{a} - \mathbf{c}\|^2$$

giving

$$\mathbf{a} \cdot \mathbf{c} = \|\mathbf{a}\|\|\mathbf{c}\|\cos\theta.$$

Exercise 2.3.14

Since
$$\frac{\mathbf{v}\cdot\mathbf{u}}{\|\mathbf{u}\|\|\mathbf{v}\|}=\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|},$$
the angle between $\mathbf{v}$ and $\mathbf{u}$ is the same as the the angle between $\mathbf{u}$ and $\mathbf{v}$

If θ is the angle between $\mathbf{u}$ and $\mathbf{v}$ and ϕ is the angle between $\mathbf{u}$ and $-\mathbf{v}$, then
$$\cos\phi=\frac{\mathbf{u}\cdot(-\mathbf{v})}{\|\mathbf{u}\|\|\mathbf{v}\|}=-\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}=-\cos\theta$$
and $0\le\theta,\phi\le\pi$, so $\phi=\pi-\theta$

Exercise 2.3.16

Let us put the vertices at

$$\mathbf{a} = (a_1, a_2, \ldots, a_n).$$

The diagonals join vertices $\mathbf{a}$ and $-\mathbf{a}$ (call this diagonal $d_{\mathbf{a}}$).

Diagonals $d_{\mathbf{a}}$, $d_{\mathbf{b}}$ are perpendicular if and only if

$$0 = \mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i$$

Since $a_i b_i = \pm 1$, $\sum_{i=1}^{n} a_i b_i$ is odd if n is odd and even if n is even. (Proof by induction or modular arithmetic.) Thus if n is odd, $\mathbf{a} \cdot \mathbf{b} \neq 0$ and no diagonals are perpendicular.

Now suppose $n = 4$. In finding the possible angles θ we may suppose without loss of generality that $\mathbf{a} = (1, 1, 1, 1)$. The possible angles are given by

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|} \in \{r/4 \,:\, -4 \leq r \leq 4\}$$

and all the possibilities occur (note that we may get different angles according to the direct assigned to the diagonals). The possible angles are

$$0,\ \cos^{-1}\tfrac{1}{4},\ \tfrac{\pi}{6},\ \cos^{-1}\tfrac{3}{4},\ \tfrac{\pi}{2},\ \pi - \cos^{-1}\tfrac{3}{4},\ \tfrac{5\pi}{6},\ \pi - \cos^{-1}\tfrac{1}{4}, \pi$$

the angle 0 and π being the angle between the diagonal and itself and the angle between the diagonal and itself reversed.

Exercise 2.3.17

(i) Always true

$$\mathbf{u} \perp \mathbf{v} \Rightarrow \mathbf{u} \cdot \mathbf{v} = 0 \Rightarrow \mathbf{v} \cdot \mathbf{u} = 0 \Rightarrow \mathbf{v} \perp \mathbf{u}.$$

(ii) Sometimes false. Take $\mathbf{u} = \mathbf{w} = (1, 0, 0)$ and $\mathbf{v} = (0, 1, 0)$.

(iii) Always true

$$\mathbf{u} \perp \mathbf{u} \Rightarrow \mathbf{u} \cdot \mathbf{u} = 0 \Rightarrow \mathbf{u} = \mathbf{0}.$$

Exercise 2.3.18

(i) We have

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v})\\
&= \mathbf{u}\cdot\mathbf{u}+\mathbf{u}\cdot\mathbf{v}+\mathbf{v}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v}\\
&= \|\mathbf{u}\|^2+0+0+\|\mathbf{v}\|^2 = \|\mathbf{u}\|^2+\|\mathbf{v}\|^2
\end{aligned}$$

Consider the right angled triangle OUV with O at $\mathbf{0}$, U at $\mathbf{u}$ and V at $\mathbf{v}$. Then $\|\mathbf{u}+\mathbf{v}\|$ is the length of the hypotenuse.

(ii) We have

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}+\mathbf{w}\|^2 &= (\mathbf{u}+\mathbf{v}+\mathbf{w})\cdot(\mathbf{u}+\mathbf{v}+\mathbf{w})\\
&= \mathbf{u}\cdot\mathbf{u}+\mathbf{v}\cdot\mathbf{v}+\mathbf{w}\cdot\mathbf{w}+2\mathbf{u}\cdot\mathbf{v}+2\mathbf{v}\cdot\mathbf{w}+2\mathbf{w}\cdot\mathbf{u}\\
&= \|\mathbf{u}\|^2+\|\mathbf{v}\|^2+\|\mathbf{w}\|^2+0+0+0 = \|\mathbf{u}\|^2+\|\mathbf{v}\|^2+\|\mathbf{w}\|^2
\end{aligned}$$

(iii) If $\mathbf{u}_j \in \mathbb{R}^4$ and $\mathbf{u}_j \perp \mathbf{u}_k$ for $k \neq j$, we have

$$\begin{aligned}
\left\|\sum_{j=1}^{4}\mathbf{u}_j\right\|^2 &= \left(\sum_{j=1}^{4}\mathbf{u}_j\right)\cdot\left(\sum_{j=1}^{4}\mathbf{u}_j\right)\\
&= \sum_{j=1}^{4}\sum_{k=1}^{4}\mathbf{u}_j\cdot\mathbf{u}_k\\
&= \sum_{j=1}^{4}\mathbf{u}_j\cdot\mathbf{u}_j = \sum_{j=1}^{4}\|\mathbf{u}_j\|^2.
\end{aligned}$$

Exercise 2.4.1

$$\begin{aligned}\|\mathbf{a}+\mathbf{b}\|^2+\|\mathbf{a}-\mathbf{b}\|^2 &= (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b})+(\mathbf{a}-\mathbf{b})\cdot(\mathbf{a}-\mathbf{b})\\ &= (\|\mathbf{a}\|^2+2\mathbf{a}\cdot\mathbf{b}+\|\mathbf{b}\|^2)+(\|\mathbf{a}\|^2-2\mathbf{a}\cdot\mathbf{b}+\|\mathbf{b}\|^2)\\ &= 2(\|\mathbf{a}\|^2+\|\mathbf{b}\|^2)\end{aligned}$$

The sum of the squares of the lengths of the two diagonals of a parallelogram equals the sum of the lengths of the four sides.

Exercise 2.4.2

(i) Consider the parallelogram $OACB$ with vertex 0 at $\mathbf{0}$, vertex A at $\mathbf{a}$, vertex C at $\mathbf{a}+\mathbf{b}$ and vertex B at $\mathbf{b}$. The midpoint of the diagonal OC is at

$$\mathbf{x} = \tfrac{1}{2}\mathbf{0} + \tfrac{1}{2}(\mathbf{a}+\mathbf{b}) = \tfrac{1}{2}(\mathbf{a}+\mathbf{b}).$$

The midpoint of the diagonal AB is at

$$\mathbf{y} = \tfrac{1}{2}\mathbf{a} + \tfrac{1}{2}\mathbf{b} = \tfrac{1}{2}(\mathbf{a}+\mathbf{b}).$$

Thus $\mathbf{x} = \mathbf{y}$ and we are done.

The midpoint M of the side CA opposite 0 is at

$$\mathbf{m} = \tfrac{1}{2}(\mathbf{a}+\mathbf{b}) + \tfrac{1}{2}\mathbf{b} = \tfrac{1}{2}\mathbf{a} + \mathbf{b}.$$

The point Z given by

$$\mathbf{z} = \tfrac{1}{3}\mathbf{0} + \tfrac{2}{3}\mathbf{m} = \tfrac{1}{3}\mathbf{a} + \tfrac{2}{3}\mathbf{b}$$

trisects OM and AB, so we are done.

Exercise 2.4.4★

Exercise 2.4.6

(i) If (n, m) is a unit vector perpendicular to (u, v) we have $nu = -mv$ so $n = -vx$, $m = ux$ and $1 = (u^2 + v^2)x^2 = x^2$ thus

$$(n, m) = (-v, u) \text{ or } (n, m) = (v, -u).$$

(ii) Let $\mathbf{n}$ be a unit vector perpendicular to $\|\mathbf{c}\|^{-1}\mathbf{c}$ as found in (i). Let $p = \mathbf{n} \cdot \mathbf{a}$.

Then

$$\mathbf{x} = \mathbf{a} + t\mathbf{c} \Rightarrow \mathbf{x} \cdot \mathbf{n} = p$$

and if $\mathbf{x} \cdot \mathbf{n} = p$ we have $(\mathbf{x} - \mathbf{a}) \cdot \mathbf{n} = 0$ so $\mathbf{x} - \mathbf{a} = t\mathbf{c}$ for some t.

(iii) Let $\mathbf{a} = p\mathbf{n}$ and let $\mathbf{c} = \mathbf{u}$ be a unit vector perpendicular to $\mathbf{n}$.

$$\mathbf{x} = p\mathbf{n} + t\mathbf{u} \Rightarrow \mathbf{x} \cdot \mathbf{n} = p$$

and, if $\mathbf{x} \cdot \mathbf{n} = p$, we have $(\mathbf{x} - p\mathbf{n}) \cdot \mathbf{u} = 0$ so $\mathbf{x} - p\mathbf{n} = t\mathbf{u}$ for some t.

(Or use a geometric argument.)

Exercise 2.4.9

We have

$$\begin{aligned}
\pi_1 &= \{(0, x_2, x_3) \,:\, x_2,\, x_3 \in \mathbb{R}\},\\
\pi_2 &= \{(x_1, 0, x_3) \,:\, x_2,\, x_3 \in \mathbb{R}\},\\
\pi_3 &= \{(x_1, x_2, x_3) \,: x_1 + x_2 = 1\}\, x_j \in \mathbb{R}\}.
\end{aligned}$$

π_1 and π_2 meet in the line $x_1 = x_2 = 0$.

π_1 and π_3 meet in the line $x_1 = 0$, $x_2 = 1$.

π_2 and π_3 meet in the line $x_1 = 1$, $x_2 = 0$.

The lines have no point in common.

(i) Now let π_j be given by

$$\mathbf{n}_j \cdot \mathbf{x} = p_j,$$

with

$$\mathbf{n}_1 = \mathbf{n}_2 = \mathbf{n}_3 = (1, 0, 0)$$

and $p_1 = -1$, $p_2 = 0$, $p_3 = 1$. By inspection no two planes meet.

(ii) Finally let π_j be given by

$$\mathbf{n}_j \cdot \mathbf{x} = p_j,$$

with

$$\mathbf{n}_2 = \mathbf{n}_3 = (1, 0, 0),\ \ \mathbf{n}_1 = (0, 1, 0)$$

and $p_1 = p_2 = 0$, $p_3 = 1$. We have that π_1 and π_2 meet in the line given by $x_1 = x_2 = 0$ and π_1 and π_2 meet in the line given by $x_1 = 1$, $x_2 = 0$, but π_2 and π_3 do not meet.

Exercise 2.4.10

$$\{\mathbf{x} \in \mathbb{R}^2 \,:\, \|\mathbf{x} - \mathbf{a}\| = r\}$$

is the singleton $\{\mathbf{a}\}$ if $r = 0$ and the empty set $\emptyset$ if $r < 0$.

EXERCISE 2.4.11

(i) We have

$$\mathbf{y}\big(\mathbf{y}(\mathbf{x})\big) = \frac{1}{\|\mathbf{y}(\mathbf{x})\|^2}\mathbf{y}(\mathbf{x}) = \frac{\|\mathbf{x}\|^4}{\|\mathbf{x}\|^2}\frac{1}{\|\mathbf{x}\|^2}\mathbf{x} = \mathbf{x}.$$

(ii) Observe that, if $\|\mathbf{a}\| \neq r$, then, if $\mathbf{x} \neq \mathbf{0}$, setting $C = (\|\mathbf{a}\|^2 - r^2)$,

$$\begin{aligned}\|\mathbf{x}-\mathbf{a}\|^2 = r^2 &\Leftrightarrow \|\mathbf{x}\|^2 - 2\mathbf{x}\cdot\mathbf{a} + (\|\mathbf{a}\|^2 - r^2) = 0\\ &\leftrightarrow 1 - 2\mathbf{y}\cdot\mathbf{a} + (\|\mathbf{a}\|^2 - r^2)\|\mathbf{y}\|^2\\ &\Leftrightarrow \|\mathbf{y}\|^2 - 2\mathbf{y}\cdot(C^{-1}\mathbf{a}) + C^{-1} = 0\\ &\Leftrightarrow \|\mathbf{y} - C^{-1}\mathbf{a}\| = C^{-1}(C^{-2}\|\mathbf{a}\|^2 - 1)\end{aligned}$$

so we transform the circle into a circle centre $C^{-1}\mathbf{a}$ radius

$$\big((C^{-1}(C^{-2}\|\mathbf{a}\|^2 - 1)\big)^{1/2}.$$

(Note that we are taking the square root of a positive number.)

(iii) Observe that, if $\|\mathbf{a}\| = r > 0$, then, if $\mathbf{x} \neq \mathbf{0}$,

$$\begin{aligned}\|\mathbf{x}-\mathbf{a}\|^2 = r^2 &\Leftrightarrow \|\mathbf{x}\|^2 - 2\mathbf{x}\cdot\mathbf{a} = 0\\ &\Leftrightarrow 1 - 2\mathbf{y}\cdot\mathbf{a} = 0\end{aligned}$$

so we transform the circle into a line perpendicular to $\mathbf{a}$ closest distance to the origin $(2\|\mathbf{a}\|)^{-1}$.

(iv) Goes through without change, replacing ‘circle’ by ‘sphere’ and ‘line’ by plane’

EXERCISE 2.4.12

$$\begin{aligned}\left\|\frac{\mathbf{x}}{\|\mathbf{x}\|^2}-\frac{\mathbf{y}}{\|\mathbf{y}\|^2}\right\|^2&=\left(\frac{\mathbf{x}}{\|\mathbf{x}\|^2}-\frac{\mathbf{y}}{\|\mathbf{y}\|^2}\right)\cdot\left(\frac{\mathbf{x}}{\|\mathbf{x}\|^2}-\frac{\mathbf{y}}{\|\mathbf{y}\|^2}\right)\\&=\frac{\mathbf{x}\cdot\mathbf{x}}{\|\mathbf{x}\|^4}-2\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2}+\frac{\mathbf{y}\cdot\mathbf{y}}{\|\mathbf{y}\|^4}\\&=\frac{1}{\|\mathbf{x}\|^2}-2\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2}+\frac{1}{\|\mathbf{y}\|^2}\\&=\frac{\|\mathbf{x}\|^2-2\mathbf{x}\cdot\mathbf{y}+\|\mathbf{y}\|^2}{\|\mathbf{x}\|^2\|\mathbf{y}\|^2}\\&=\left(\frac{\|\mathbf{x}-\mathbf{y}\|}{\|\mathbf{x}\|\|\mathbf{y}\|}\right)^2.\end{aligned}$$

Thus

$$\left\|\frac{\mathbf{x}}{\|\mathbf{x}\|^2}-\frac{\mathbf{y}}{\|\mathbf{y}\|^2}\right\|=\frac{\|\mathbf{x}-\mathbf{y}\|}{\|\mathbf{x}\|\|\mathbf{y}\|}.$$

Using this result, we have, by the triangle inequality, (if $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}\neq\mathbf{0}$)

$$\begin{aligned}(\|\mathbf{x}\|\|\mathbf{y}\|\|\mathbf{z}\|)^{-1}(\|\mathbf{y}\|\|\mathbf{z}-\mathbf{x}\|+\|\mathbf{x}\|\|\mathbf{y}-\mathbf{z}\|)&=\frac{\|\mathbf{z}-\mathbf{x}\|}{\|\mathbf{z}\|\|\mathbf{x}\|}+\frac{\|\mathbf{y}-\mathbf{z}\|}{\|\mathbf{y}\|\|\mathbf{z}\|}\\&=\left\|\frac{\mathbf{z}}{\|\mathbf{z}\|^2}-\frac{\mathbf{x}}{\|\mathbf{x}\|^2}\right\|+\left\|\frac{\mathbf{y}}{\|\mathbf{y}\|^2}-\frac{\mathbf{z}}{\|\mathbf{z}\|^2}\right\|\\&\geq\left\|\frac{\mathbf{x}}{\|\mathbf{x}\|^2}-\frac{\mathbf{y}}{\|\mathbf{y}\|^2}\right\|\\&=(\|\mathbf{x}\|\|\mathbf{y}\|\|\mathbf{z}\|)^{-1}\|\mathbf{z}\|\|\mathbf{y}-\mathbf{x}\|\end{aligned}$$

with equality if and only if $\|\mathbf{x}\|^{-2}\mathbf{x}-\|\mathbf{y}\|^{-2}\mathbf{y}$ and $\|\mathbf{z}\|^{-2}\mathbf{z}-\|\mathbf{y}\|^{-2}\mathbf{y}$ are scalar multiples of each other (i.e. $\|\mathbf{x}\|^{-2}$, $\|\mathbf{y}\|^{-2}\mathbf{y}$ $\|\mathbf{z}\|^{-2}\mathbf{z}$ lie on a straight line). Thus

$$\|\mathbf{z}\|\|\mathbf{x}-\mathbf{y}\|\leq\|\mathbf{y}\|\|\mathbf{z}-\mathbf{x}\|+\|\mathbf{x}\|\mathbf{y}\|$$

if $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}\neq\mathbf{0}$. If at least one of the vectors is zero the result is trivial.

To obtain the Euclidean result place D at the origin , let A have position vector $\mathbf{x}$, B have position vector $\mathbf{y}$ and C position vector $\mathbf{z}$.

We now use the notation of Exercise 2.4.11. The condition $\|\mathbf{x}\|^{-2}\mathbf{x}$, $\|\mathbf{y}\|^{-2}\mathbf{y}$ $\|\mathbf{z}\|^{-2}\mathbf{z}$ are scalar multiples of each other is the same as saying that $f(\mathbf{x})$, $f(\mathbf{y})$, $f(\mathbf{y})$ lie on a straight line so $\mathbf{x}=f^2(\mathbf{x})$, $\mathbf{y}=f^2(\mathbf{y})$ $\mathbf{z}=f^2(\mathbf{z})$ lie on a circle through the origin i.e. A, B, C, D are concyclic.

Exercise 3.2.1

$$z_i = a_{ik}y_k = a_{ik}(b_{kj}x_j) = (a_{ik}b_{kj})x_j = c_{ij}x_j$$

EXERCISE 3.3.4

There are many proofs. We can work coordinatewise e.g.

$$a_{ij} + (b_{ij} + c_{ij}) = (a_{ij} + b_{ij}) + c_{ij}$$

so $A + (B + C) = (A + B) + C$

Or direct from definition.

$$\begin{aligned}(A + (B + C))\mathbf{x} &= A\mathbf{x} + (B + C)\mathbf{x} = A\mathbf{x} + (B\mathbf{x} + C\mathbf{x}) \\ &= (A\mathbf{x} + B\mathbf{x}) + C\mathbf{x}) = (A + B)\mathbf{x} + C\mathbf{x} \\ &= ((A + B) + C))\mathbf{x}\end{aligned}$$

for all $\mathbf{x}$, so $(A + B) + C = A + (B + C)$.

Exercise 3.3.6

Suppose $C\mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$. Taking $\mathbf{x}$ to be the column vector with 1 in k th place and zero elsewhere, we get

$$c_{ik} = \delta_{ik}.$$

Conversely, if $\mathbf{y} \in \mathbb{R}^n$

$$\sum_{j=1}^{n} \delta_{ij} y_i = y_j$$

so $I\mathbf{y} = \mathbf{y}$.

EXERCISE 3.3.9

(iii) By definition. Observe that

$$\begin{aligned}\big((B+C)A\big)\mathbf{x} &= (B+C)(A\mathbf{x}) = B(A\mathbf{x}) + C(A\mathbf{x})\\ &= (BA)\mathbf{x} + (CA)\mathbf{x} = (BA+CA)\mathbf{x}\end{aligned}$$

for all $\mathbf{x}$ so $(B+C)A = BA + CA$.

(iii) By calculation. We have

$$\sum_{j=1}^{n}(b_{ij}+c_{ij})a_{jk} = \sum_{j=1}^{n} b_{ij}a_{jk} + c_{ij}a_{jk} = \sum_{j=1}^{n} b_{ij}a_{jk} + \sum_{j=1}^{n} c_{ij}a_{jk}$$

(iii) By summation convention. Observe that

$$(b_{ij}+c_{ij})a_{jk} = b_{ij}a_{jk} + c_{ij}a_{jk}$$

(iv) By definition. Observe that

$$\begin{aligned}\big((\lambda A)B\big)\mathbf{x} &= (\lambda A)(B\mathbf{x}) = \lambda\big(A(B\mathbf{x})\big)\\ &= \lambda\big((AB)\mathbf{x}\big) = \big(\lambda(AB)\big)\mathbf{x}\end{aligned}$$

for all $\mathbf{x}$ so $(\lambda A)B = \lambda(AB)$.

Again

$$\begin{aligned}\big(A(\lambda B)\big)\mathbf{x} &= A\big((\lambda B)\mathbf{x}\big) = \big(A(\lambda(B\mathbf{x}))\big)\\ &= \lambda\big(A((B\mathbf{x})\big) = \lambda\big((AB)\mathbf{x}\big) = \big(\lambda(BA)\big)\mathbf{x}\end{aligned}$$

for all $\mathbf{x}$ so $A(\lambda B) = \lambda(AB)$.

(iv) By calculation. We have

$$\sum_{j=1}^{n}(\lambda a_{ij})b_{jk} = \lambda\sum_{j=1}^{n} a_{ij}b_{jk} = \sum_{j=1}^{n} a_{ij}(\lambda b_{jk})$$

(iv) By summation convention. Observe that

$$(\lambda a_{ij})b_{jk} = \lambda(a_{ij}b_{jk}) = a_{ij}(\lambda b_{jk})$$

(v) By definition. Observe that

$$(IA)\mathbf{x} = I(A\mathbf{x}) = A\mathbf{x}$$

and

$$(AI)\mathbf{x} = A(I\mathbf{x}) = A\mathbf{x}$$

for all $\mathbf{x}$ so $IA = AI$.

(v) By calculation. We have

$$\sum_{j=1}^{n} \delta_{ij} a_{jk} = a_{ik} = \sum_{j=1}^{n} a_{ij} \delta_{jk} = \sum_{j=1}^{n} b_{ij} a_{jk} + c_{ij} a_{jk}.$$

(v) By summation convention. Observe that

$$\delta_{ij} a_{jk} = a_{ik} = a_{ij} \delta_{jk}.$$

Exercise 3.3.10

We have
$$BA = A \ \forall A \Rightarrow BI = I \Rightarrow B = I.$$

Exercise 3.3.12

If A is an $m \times n$ matrix and $\lambda \in \mathbb{R}$, then $\lambda A = C$ where C is the $m \times n$ matrix such that

$$\lambda(A\mathbf{x}) = C\mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^n$.

If A and B are $m \times n$ matrices, then $A + B = C$ where C is the $m \times n$ matrix such that

$$A\mathbf{x} + B\mathbf{x} = C\mathbf{x}$$

If A is an $m \times n$ matrices and B is an $m \times p$ matrix, then $AB = C$ where C is the $n \times p$ matrix such that

$$A(B\mathbf{x}) = C\mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^n$

Remark The matrices defined above are certainly unique. The definitions do not by themselves show existence, but the existence may be easily checked.

In (i) take $c_{ij} = \lambda a_{ij}$. In (ii) take $c_{ij} = \lambda a_{ij} + b_{ij}$. In (iii) take $c_{ij} = \sum_{r=1}^{m} a_{ir} b_{rj}$.

Exercise 3.3.13

(i) Observe that

$$\begin{aligned}\big((A+B)+C)\mathbf{x} &= (A+B)\mathbf{x}+C\mathbf{x} = (A\mathbf{x}+B\mathbf{x})+C\mathbf{x}\\ &= A\mathbf{x}+(B\mathbf{x}+C\mathbf{x}) = A\mathbf{x}+(B+C)\mathbf{x}\\ &= (\big(A+(B+C)\big)\mathbf{x}\end{aligned}$$

for all $\mathbf{x}\in\mathbb{R}^n$ so $(A+B)+C = A+(B+C)$.

(ii) Observe that

$$(A+B)\mathbf{x} = A\mathbf{x}+B\mathbf{x} = B\mathbf{x}+A\mathbf{x} = (B+A)\mathbf{x}$$

for all $\mathbf{x}\in\mathbb{R}^n$ so $A+B = B+A)$.

(iii) Observe that

$$(A+0)\mathbf{x} = A\mathbf{x}+0\mathbf{x} = A\mathbf{x}+\mathbf{0} = A\mathbf{x}$$

for all $\mathbf{x}\in\mathbb{R}^n$ so $A+0 = A$.

(iv) Observe that

$$\begin{aligned}\big(\lambda(A+B)\big)\mathbf{x} &= \lambda\big((A+B)\mathbf{x}\big) = \lambda(A\mathbf{x}+B\mathbf{x})\\ &= \lambda(A\mathbf{x})+\lambda(B\mathbf{x}) = (\lambda A)\mathbf{x}+(\lambda B)\mathbf{x}\\ &= (\lambda A+\lambda B)\mathbf{x}\end{aligned}$$

for all $\mathbf{x}\in\mathbb{R}^n$ so $\lambda(A+B) = \lambda A+\lambda B$.

(v) Observe that

$$\begin{aligned}\big((\lambda+\mu)A\big)\mathbf{x} &= (\lambda+\mu)(A\mathbf{x}) = \lambda(A\mathbf{x})+\mu(A\mathbf{x})\\ &= (\lambda A)\mathbf{x}+(\mu A)\mathbf{x} = (\lambda A+\mu A)\mathbf{x}\end{aligned}$$

for all $\mathbf{x}\in\mathbb{R}^n$ so $(\lambda+\mu) = \lambda A+\mu A$.

(vi) Observe that

$$\big((\lambda\mu)A\big)\mathbf{x} = (\lambda\mu)A\mathbf{x} = \lambda(\mu A\mathbf{x}) = \big(\lambda(\mu A)\big)\mathbf{x}$$

for all $\mathbf{x}\in\mathbb{R}^n$ so $(\lambda\mu)A = \lambda(\mu A)$.

(vii) Observe that

$$(0A)\mathbf{x} = 0(A\mathbf{x}) = \mathbf{0} = \underline{0}\mathbf{x}$$

for all $\mathbf{x}\in\mathbb{R}^n$ so $0A = \underline{0}$.

(ix) We have

$$A-A = (1-1)A = 0A = \underline{0}.$$

Exercise 3.3.14

(i) Observe that

$$\big((AC)F\big)\mathbf{x} = (AC)(F\mathbf{x}) = A\big(C(F\mathbf{x})\big) = A\big((CF)\mathbf{x})\big) = \big((ACF)\big)\mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^n$, so $(AC)F = A(CF)$.

(ii) Observe that

$$\begin{aligned}\big(G(A+B)\big)\mathbf{x} &= G\big((A+B)\mathbf{x}\big) = G(A\mathbf{x} + B\mathbf{x})\\ &= G(A\mathbf{x}) + G(B\mathbf{x}) = (GA)\mathbf{x} + (GB)\mathbf{x}\\ &= (GA + GB)\mathbf{x}\end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n$, so $G(A+B) = GA + GB$.

(iii) Observe that

$$\begin{aligned}\big((A+B)C\big)\mathbf{x} &= (A+B)(C\mathbf{x}) = A(C\mathbf{x}) + A(B\mathbf{x})\\ &= (AC)\mathbf{x} + (BC)\mathbf{x}\\ &= (AC + BC)\mathbf{x}\end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n$, so $(A+B)C = AC + BC$.

(iv) Observe that

$$\big((\lambda A)C\big)\mathbf{x} = (\lambda A)(C\mathbf{x}) == \lambda\big(A(C\mathbf{x})\big) = \lambda\big((AC)\mathbf{x})\big) = \big(\lambda(AC)\big)\mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^n$, so $(\lambda A)C = \lambda(AC)$.

Observe that

$$\begin{aligned}\big(A(\lambda C)\big)\mathbf{x} &= A\big((\lambda C)\mathbf{x}\big) = A\big(\lambda(C\mathbf{x})\big) = \lambda\big(A(C\mathbf{x})\big)\\ &= \lambda\big((AC)\mathbf{x})\big) = \big(\lambda(AC)\big)\mathbf{x}\end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^n$ so $A(\lambda C) = \lambda(AC)$.

Exercise 3.4.5

The conjecture is false.

The matrices A and B given by

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

are shear matrices, but

$$AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

is not.

Exercise 3.4.7

Since $r \neq s$, we have $\delta_{rs} = 0$ and

$$\begin{aligned}
\sum_{j=1}^{n} (\delta_{ij} + \lambda\delta_{ir}\delta_{js})(\delta_{jk} + \mu\delta_{jr}\delta_{ks}) \\
&= \sum_{j=1}^{n} (\delta_{ij}\delta_{jk} + \lambda\delta_{ir}\delta_{js}\delta_{jk} + \mu\delta_{jr}\delta_{ks}\delta_{ij} + \lambda\mu\delta_{ir}\delta_{js}\delta_{jr}\delta_{ks}) \\
&= \delta_{ik} + \lambda\delta_{ir}\delta_{ks} + \mu\delta_{ir}\delta_{ks} + \lambda\mu\delta_{sr}\delta_{ir}\delta_{ks} \\
&= \delta_{ik} + (\lambda + \mu)\delta_{ir}\delta_{ks}
\end{aligned}$$

The result is also obvious geometrically.

Exercise 3.5.1

$$(L_k L_{k-1} \ldots L_1 I)A = L_k L_{k-1} \ldots L_1 A = I,$$

so A is invertible with

$$L_k L_{k-1} \ldots L_1 I = A^{-1}.$$

Thus the same set of elementary operations applied in the same order which reduce A to I will transform I to A^{-1}.

Exercise 3.5.2

If we use row operations and column exchanges we can reduce A to I. If we omit column exchanges then the same row operations in the same order produce a matrix with exactly one 1 in each row and column and the remaining entries 0. By exchanging rows we can ensure that the first row is $(1, 0, 0, \ldots, 0)$, the second row $(0, 1, 0, \ldots, 0)$ and so on, i.e. we can produce I.

Exercise 4.1.1

(i) The vertices ABC, BCA, CAB are described in one sense (say anti clockwise) and the same vertices ACB, BAC, CBA in the opposite sense (say clockwise).

(ii) (Of course a bit hand waving.) $OAXB$, $0PQB$, BXQ and $0AP$ anticlockwise but $AXPQ$ clockwise. Writing $|\,\mathrm{area}\,\Sigma|$ for the absolute value of the area we have

$$
\begin{aligned}
&|\,\text{area } OPQB| \\
&\qquad = |\,\mathrm{area}\, OAXB| - |\,\mathrm{area}\, APQX| + |\,\mathrm{area}\, OAP| - |\,\mathrm{area}\, BXQ| \\
&\qquad = |\,\mathrm{area}\, OAXB| - |\,\mathrm{area}\, APQX|.
\end{aligned}
$$

Exercise 4.1.2

$$\mathcal{D}(\mathbf{a}, \mathbf{b}) + \mathcal{D}(\mathbf{b}, \mathbf{a}) = \mathcal{D}(\mathbf{a} + \mathbf{b}, \mathbf{b}) + \mathcal{D}(\mathbf{a} + \mathbf{b}, \mathbf{a})$$

(By adding second entry to first.)

$$\mathcal{D}(\mathbf{a} + \mathbf{b}, \mathbf{b}) + \mathcal{D}(\mathbf{a} + \mathbf{b}, \mathbf{a}) = \mathcal{D}(\mathbf{a} + \mathbf{b}, \mathbf{a} + \mathbf{b})$$

(Since first entry the same, we can add as shown.)

$$\mathcal{D}(\mathbf{a} + \mathbf{b}, \mathbf{a} + \mathbf{b}) = \mathcal{D}(\mathbf{a} + \mathbf{b} - (\mathbf{a} + \mathbf{b}), \mathbf{a} + \mathbf{b})$$

(By subtracting second entry from first entry.)

$$\mathcal{D}(\mathbf{a} + \mathbf{b} - (\mathbf{a} + \mathbf{b}), \mathbf{a} + \mathbf{b}) = \mathcal{D}(\mathbf{0}, \mathbf{a} + \mathbf{b})$$

(Just do the calculation.)

$$\mathcal{D}(\mathbf{0}, \mathbf{a} + \mathbf{b}) = 0$$

(Area degenerate parallelogram.) $\mathcal{D}(\mathbf{a}, \mathbf{b})$ is the area of the parallelogram with vertices $\mathbf{0}$, $\mathbf{a}$, $\mathbf{a} + \mathbf{b}$, $\mathbf{b}$ described in that order. $\mathcal{D}(\mathbf{b}, \mathbf{a})$ is the area of the same parallelogram but with the vertices laid out in the opposite sense $\mathbf{0}$, $\mathbf{b}$, $\mathbf{a} + \mathbf{b}$, $\mathbf{a}$.

Exercise 4.1.3

$$\mathcal{D}\left(\begin{pmatrix}a_1\\a_2\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}\right)=\mathcal{D}\left(\begin{pmatrix}a_1\\0\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}\right)+\mathcal{D}\left(\begin{pmatrix}0\\a_2\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}\right)$$

(Since second column the same, can add first columns.)

$$\begin{aligned}&\mathcal{D}\left(\begin{pmatrix}a_1\\0\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}\right)+\mathcal{D}\left(\begin{pmatrix}0\\a_2\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}\right)\\&\quad=\mathcal{D}\left(\begin{pmatrix}a_1\\0\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}-\frac{b_1}{a_1}\begin{pmatrix}a_1\\0\end{pmatrix}\right)+\mathcal{D}\left(\begin{pmatrix}0\\a_2\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}-\frac{b_2}{a_2}\begin{pmatrix}0\\a_2\end{pmatrix}\right)\end{aligned}$$

(Since we may subtract multiples of the first column from the second.)

$$\begin{aligned}&\mathcal{D}\left(\begin{pmatrix}a_1\\0\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}-\frac{b_1}{a_1}\begin{pmatrix}a_1\\0\end{pmatrix}\right)+\mathcal{D}\left(\begin{pmatrix}0\\a_2\end{pmatrix},\begin{pmatrix}b_1\\b_2\end{pmatrix}-\frac{b_2}{a_2}\begin{pmatrix}0\\a_2\end{pmatrix}\right)\\&\quad=\mathcal{D}\left(\begin{pmatrix}a_1\\0\end{pmatrix},\begin{pmatrix}0\\b_2\end{pmatrix}\right)+\mathcal{D}\left(\begin{pmatrix}0\\a_2\end{pmatrix},\begin{pmatrix}b_2\\0\end{pmatrix}\right)\end{aligned}$$

(Just doing the calculation.)

$$\begin{aligned}&\mathcal{D}\left(\begin{pmatrix}a_1\\0\end{pmatrix},\begin{pmatrix}0\\b_2\end{pmatrix}\right)+\mathcal{D}\left(\begin{pmatrix}0\\a_2\end{pmatrix},\begin{pmatrix}b_2\\0\end{pmatrix}\right)\\&\quad=\mathcal{D}\left(\begin{pmatrix}a_1\\0\end{pmatrix},\begin{pmatrix}0\\b_2\end{pmatrix}\right)-\mathcal{D}\left(\begin{pmatrix}b_2\\0\end{pmatrix},\begin{pmatrix}0\\a_2\end{pmatrix}\right)\end{aligned}$$

(Interchanging columns.)

$$\begin{aligned}&\mathcal{D}\left(\begin{pmatrix}a_1\\0\end{pmatrix},\begin{pmatrix}0\\b_2\end{pmatrix}\right)-\mathcal{D}\left(\begin{pmatrix}b_2\\0\end{pmatrix},\begin{pmatrix}0\\a_2\end{pmatrix}\right)\\&\quad=(a_1b_2-a_2b_1)\mathcal{D}\left(\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right)=a_1b_2-a_2b_1.\end{aligned}$$

(Using the rules $\mathcal{D}(\lambda\mathbf{a},\mathbf{b})=\mathcal{D}(\mathbf{a},\lambda\mathbf{b})=\lambda\mathcal{D}(\mathbf{a},\mathbf{b})$ and the fact that that the area of a unit square is 1.)

Exercise 4.2.1

(Not a good idea, but possible as follows.)

$$\begin{aligned}\mathcal{D}(AB) &= \mathcal{D}\begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix} \\ &= (a_{11}b_{11}+a_{12}b_{21})(a_{21}b_{12}+a_{22}b_{22}) \\ &\qquad - (a_{11}b_{12}+a_{12}b_{22})(a_{21}b_{11}+a_{22}b_{21}) \\ &= a_{11}a_{22}(b_{11}b_{22}-b_{12}b_{21}) - a_{12}a_{21}(b_{11}b_{22}-b_{12}b_{21}) \\ &= \mathcal{D}(A)\mathcal{D}(B).\end{aligned}$$

Exercise 4.2.2

(i) We have

$$\mathcal{D}I = \mathcal{D}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = 1\times 1 - 0\times 0 = 1.$$

The area of a unit square is 1.

(i) We have

$$\mathcal{D}E = \mathcal{D}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = 0\times 0 - 1\times 1 = -1,$$

and

$$E\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}0x+1y\\ 1x+0y\end{pmatrix} = \begin{pmatrix}y\\ x\end{pmatrix}.$$

We observe that

$$E\begin{pmatrix}\cos t\\ \sin t\end{pmatrix} = \begin{pmatrix}\sin t\\ \cos t\end{pmatrix} = \begin{pmatrix}\cos(\frac{\pi}{2}-t)\\ \sin(\frac{\pi}{2}-t)\end{pmatrix}$$

runs clockwise from $(0,1)$ back to $(0,1)$ as $(\cos t, \sin t)^T$ runs anticlockwise from $(1,0)$ back to $(1,0)$.

(iii) We have

$$\mathcal{D}\begin{pmatrix}1 & \lambda\\ 0 & 1\end{pmatrix} = 1\times 1 - \lambda\times 0 = 1$$

and

$$\mathcal{D}\begin{pmatrix}1 & 0\\ \lambda & 1\end{pmatrix} = 1\times 1 - 0\times \lambda = 1.$$

(iv) We have

$$\mathcal{D}\begin{pmatrix}a & 0\\ 0 & b\end{pmatrix} = a\times b - 0\times 0 = ab$$

corresponding to the fact that a rectangle with sides of length a and b has area ab.

Exercise 4.2.3★

Exercise 4.3.2

As we hoped,

$$\epsilon_{11}a_{11}a_{12} + \epsilon_{12}a_{11}a_{22}\epsilon_{21}a_{21}a_{12} + \epsilon_{22}a_{21}a_{22} = a_{11}a_{22} - a_{21}a_{12}.$$

Exercise 4.3.3

The result is automatic if r, s and t are not distinct.

When r, s and$t$ are distinct, we check the case $r = 1$ when either $s = 2$, $t = 3$ or $s = 3$, $t = 2$ in Definition 4.3.1. We now do the same for $r = 2$ and $r = 3$.

Exercise 4.3.10

Observe that writing $C = AB$ $D = C^T$, $A' = A^T$, $B' = B^T$

$$d_{ij} = c_{ji} = \sum_{r=1}^{m} a_{jr} b_{ri} = \sum_{r=1}^{m} a'_{rj} b'_{ir} = \sum_{r=1}^{m} b'_{ir} a'_{rj}$$

(for $1 \leq i \leq p$, $1 \leq j \leq n$). Thus $(AB)^T = B^T A^T$.

EXERCISE 4.3.13

$\mathbf{x} \mapsto E_{r,s,\lambda}\mathbf{x}$ is a shear which leaves volume unchanged.

Let D have ith diagonal entry d_i. $\mathbf{x} \mapsto D\mathbf{x}$ (where D is a diagonal matrix) stretches by $|d_i|$ in the $0x_i$ direction and reflects if $d_{ii} < 0$. Thus the volume is multiplied by $\prod_{i=1}^{n} d_i = \det D$.

$\mathbf{x} \mapsto P(\sigma)\mathbf{x}$ exchanges the handedness of coordinates (or equivalently is a reflection in a particular plane) which multiplies volume by -1.

If A is any 3 matrix then $A = a_1 A_2 \ldots A_k$ with the A_j elementary matrices. Since

$$A\mathbf{x} = A_1(A_2(A_3 \ \ldots(A_k\mathbf{x})\ldots))$$

the transformation $\mathbf{x} \mapsto A\mathbf{x}$ rescales by $\prod_{j=1}^{k} \det A_j = \det A$.

Exercise 4.3.14

The map $\mathbf{x} \mapsto M\mathbf{x}$ changes the length scale by λ and the volume scale by λ^3. (This is just a special case of the map $\mathbf{x} \mapsto D\mathbf{x}$ considered in the previous question.)

Exercise 4.3.15

Let

$$F(r,s,t)=\epsilon_{ijk}a_{ir}a_{js}a_{kt}.$$

We observe that $F(r,s,t)$ is the determinant of the matrix with first row the rth row of A, second row the sth row of A and third row the tth row of A.

Thus interchanging any two of r, s, t multiplies $F(r,s,t)$ by -1 and so

$$F(r,s,t)=\epsilon_{rst}K$$

for some constant K.

Since $F(1,2,3)=\det A$ we have

$$\epsilon_{ijk}a_{ir}a_{js}a_{kt}=F(r,s,t)=\epsilon_{rst}\det A.$$

(ii) Writing $C=AB$, we have

$$\begin{aligned}\epsilon_{rst}\det AB&=\epsilon_{rst}\det C=\epsilon_{ijk}c_{ir}c_{js}c_{kt}\\&=\epsilon_{ijk}a_{iu}b_{ur}a_{jv}b_{vs}a_{kw}b_{wt}=\epsilon_{ijk}a_{iu}a_{jv}a_{kw}b_{ur}b_{vs}b_{wt}\\&=\epsilon_{uvw}\det Ab_{ur}b_{vs}b_{wt}=\epsilon_{rst}\det\det B\end{aligned}$$

Taking $r=1$, $s=2$, $t=3$ we get $\det AB=\det A\det B$.

Exercise 4.3.16★

Exercise 4.4.1

There is no non-trivial χ.

Observe that

$$\chi_{ijkrs} = -\chi_{kijrs} = \chi_{jkirs} = -\chi_{ijkrs},$$

so $\chi_{ijkrs} = 0$.

Exercise 4.4.4

We have

$$\zeta(\sigma) = \frac{(\sigma 2 - \sigma 1)(\sigma 3 - \sigma 1)(\sigma 4 - \sigma 1)(\sigma 3 - \sigma 2)(\sigma 4 - \sigma 2)(\sigma 4 - \sigma 3)}{(2-1)(3-1)(4-1)(3-2)(4-2)(4-3)},$$

$$\zeta(\tau) = \frac{(3-2)(1-2)(4-2)(1-3)(4-3)(4-1)}{(2-1)(3-1)(4-1)(3-2)(4-2)(4-3)} = 1,$$

$$\zeta(\rho) = \frac{(3-2)(4-2)(1-2)(4-3)(1-3)(1-4)}{(2-1)(3-1)(4-1)(3-2)(4-2)(4-3)} = -1.$$

Exercise 4.4.6

(i) If $t \notin \{1,\, 2,\, i\}$

$$\begin{aligned}
\alpha\rho\alpha 1 &= \alpha\rho 2 = \alpha 1 = i = \tau 1\\
\alpha\rho\alpha 2 &= \alpha\rho 1 = \alpha 2 = 2 = \tau 2\\
\alpha\rho\alpha i &= \alpha\rho i = \alpha i = 1 = \tau i\\
\alpha\rho\alpha t &= \alpha\rho t = \alpha t = t = \tau t
\end{aligned}$$

(ii) If $j = 1$, then the result follows from part (iv) of the lemma. If not, let τ be as in part (iv) of the lemma and $\beta \in S_n$ interchange 1 and j leaving the remaining integers unchanged. Then, by inspection,

$$\beta\tau\beta(r) = \kappa(r)$$

for all $1 \le r \le n$ and so $\beta\tau\beta = \kappa$. It follows that

$$\zeta(\kappa) = \zeta(\beta)\zeta(\tau)\zeta(\beta) = \zeta(\beta)^2\zeta(\tau) = -1.$$

Exercise 4.4.7

If all the suffices i, j, k, l are distinct, then $\sigma(1) = i$, $\sigma(2) = j$, $\sigma(3) = k$, $\sigma(4) = l$ gives a unique $\sigma \in S_4$, so

$$\epsilon_{ijkl} = \epsilon_{\sigma(1)\sigma(2)\sigma(3)\sigma(4)} = \zeta(\sigma)$$

does define ϵ_{ijkl}. If τ interchanges two suffices and leaves the rest unchanged

$$\epsilon_{\tau(i)\tau(j)\tau(k)\tau(l)} = \epsilon_{\tau\sigma(1)\tau\sigma(2)\tau\sigma(3)\tau\sigma(4)} = \zeta(\tau\sigma) = \zeta(\tau)\zeta(\sigma) = -\zeta(\sigma) = -\epsilon_{ijkl}$$

If i, j, k, l are not all distinct and i', j', k', l' is some rearrangement

$$\epsilon_{ijkl} = 0 = -\epsilon_{i'j'k'l'}.$$

Finally $\epsilon_{1234} = \zeta(\iota) = 1$.

Exercise 4.4.8

Observe that $1^{-1} = 1$ and $(-1)^{-1} = -1$.

Since $\zeta(\sigma)\zeta(\sigma^{-1}) = \zeta(\sigma\sigma^{-1}) = \zeta(\iota) = 1$, it follows that $\zeta(\sigma) = \zeta(\sigma^{-1})$.

Thus

$$\begin{aligned}
\det A^T &= \sum_{\sigma \in S_n} \zeta(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} \zeta(\sigma) a_{\sigma^{-1}(1)1} a_{\sigma^{-1}(2)2} \dots a_{\sigma^{-1}(n)n} \\
&= \sum_{\sigma \in S_n} \zeta(\sigma^{-1}) a_{\sigma^{-1}(1)1} a_{\sigma^{-1}(2)2} \dots a_{\sigma^{-1}(n)n} \\
&= \sum_{\tau \in S_n} \zeta(\tau) a_{\tau(1)1} a_{\tau(2)2} \dots a_{\tau(n)n} \\
&= \det A
\end{aligned}$$

Exercise 4.4.9

(i) We have

$$\det\begin{pmatrix}1 & 1\\ x & y\end{pmatrix} = y - x.$$

(ii) Each term $\zeta(\sigma)a_{\sigma(1)1}a_{\sigma(2)2}a_{\sigma(3)3}$ in the standard expansion is a multinomial of degree 3.

$$F(x,x,z) = \det\begin{pmatrix}1 & 1 & 1\\ x & x & z\\ x^2 & x^2 & z^2\end{pmatrix} = 0,$$

because two columns of the matrix are identical. Thus $y-x$ must be be a factor of $F(x,y,z)$. Similarly $z-y$ and $z-x$ are factors. Since F has degree 3 and $(y-x)(z-y)(z-x)$ has degree 3,

$$F(x,y,z) = A(y-x)(z-y)(z-x)$$

for some constant A. By considering the terms in the standard determinant expansion, we know that the coefficient of yz^2 is 1. Thus $A = 1$ and

$$F(x,y,z) = (y-x)(z-y)(z-x).$$

(iii) Since each term in the standard expansion of $\det V$ is a multinomial of degree $0+1+\ldots+(n-1) = (n-1)n/2$, F is a multinomial of degree $(n-1)n/2$.

$$F(x_1, x_1, x_3, x_4, \ldots, x_n) = 0$$

because two columns of the associated matrix are identical. Thus $x_2 - x_1$ must be be a factor of F. Similarly $x_j - x_k$ is a factor for each $j > k$. Since F has degree $(n-1)n/2$ and $\prod_{i>j}(x_i - x_j)$ has the same degree,

$$F(x_1, x_2, \ldots, x_n) = A\prod_{i>j}(x_i - x_j)$$

for some constant A.

By considering the terms in the standard determinant expansion, we know that the coefficient of $\prod_{j=2}^{n} x_j^{j-1}$ is 1. Thus $A = 1$ and

$$F(x_1, x_2, \ldots, x_n) = \prod_{i>j}(x_i - x_j).$$

(iv) We have

$$\tilde{\zeta}_{\mathbf{x}}(\sigma) = \frac{F(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})}{F(x_1, x_2, \ldots, x_n)} = \prod_{i>j}\frac{x_{\sigma(i)} - x_{\sigma j}}{x_i - x_j}.$$

Now
$$\prod_{i>j}(x_{\sigma(i)} - x_{\sigma j}) = B\sigma \prod_{i>j}(x_i - x_j)$$
where B depends only on σ (and not on $\mathbf{x}$). Thus
$$\tilde{\zeta}_{\mathbf{x}}(\sigma) = \tilde{\zeta}_{\mathbf{y}}(\sigma)$$
with $y_j = j$ and so
$$\tilde{\zeta}_{\mathbf{x}}(\sigma) = \zeta(\sigma).$$
Thus $\zeta = \tilde{\zeta}$.

Exercise 4.5.1

Each of the $n!$ terms involves n multiplications (in addition to finding the appropriate sign) so we need $n \times n!$ multiplications. We can either use a Stirling approximation or an electronic calculator (which probably uses some version of Stirling's formula) to obtain an estimate of about 36 000 000 for $n = 10$ and about 4.9×10^{19} for $n = 20$.

Exercise 4.5.3

If $A = (a_{ij})$ is an $n \times n$ matrix which is both upper and lower triangular then $a_{ij} = 0$ if $i < j$ or if $j < i$ so A is diagonal.Each of the $n!$ terms involves n multiplications (in addition to finding the appropriate sign) so we need $nn!$ multiplications. We can either use a Stirling approximation or an electronic calculator (which probably uses some version of Stirling's formula) to obtain an estimate of about $36\,000\,000$ for $n = 10$ and about 4.9×10^{19} for $n = 20$.

EXERCISE 4.5.3

If $A = (a_{ij})$ is an $n \times n$ matrix which is both upper and lower triangular then $a_{ij} = 0$ if $i < j$ or if $j < i$ so A is diagonal.

Exercise 4.5.5

(i) By row and column operations on the first r rows and columns we can reduce

$$C = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \text{ to } C' = \begin{pmatrix} A' & 0 \\ 0 & B \end{pmatrix}$$

with $\det C = K \det C'$, $\det A' = K \det A$ and A' lower triangular.

By row and column operations on the last s rows and columns we can reduce C' to C'' with

$$C'' = \begin{pmatrix} A' & 0 \\ 0 & B' \end{pmatrix}$$

with $\det C'' = K' \det C'$, $\det B' = K' \det B$ and B' lower triangular.

We now have

$$\det C = KK' \det C'' = KK' \det A' \det B' = \det A \det B.$$

(ii) FALSE. Consider

$$F = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

With the suggested notation

$$\det A = \det B = \det C = \det D = 0$$

but

$$\det F = 1 \neq 0 = \det A \det D - \det B \det C.$$

Exercise 4.5.6

Dividing the first row by 2,

$$\det\begin{pmatrix}2&4&6\\3&1&2\\5&2&3\end{pmatrix}=2\det\begin{pmatrix}1&2&3\\3&1&2\\5&2&3\end{pmatrix}.$$

Subtracting multiples of the first row from the second and third row,

$$2\det\begin{pmatrix}1&2&3\\3&1&2\\5&2&3\end{pmatrix}=2\det\begin{pmatrix}1&2&3\\0&-5&-7\\0&-8&-12\end{pmatrix}.$$

Expanding by first row,

$$2\det\begin{pmatrix}1&2&3\\0&-5&-7\\0&-8&-12\end{pmatrix}=2\det\begin{pmatrix}-5&-7\\-8&-12\end{pmatrix}.$$

Dividing the first row by -1 and the second row by -4,

$$\begin{pmatrix}-5&-7\\-8&-12\end{pmatrix}=8\det\begin{pmatrix}5&7\\2&3\end{pmatrix}.$$

Subtracting two times the second row from the first (hardly necessary, we could have finished the calculation here),

$$8\det\begin{pmatrix}5&2\\2&1\end{pmatrix}=8\det\begin{pmatrix}1&0\\2&1\end{pmatrix}.$$

From the definition,

$$8\det\begin{pmatrix}1&0\\2&1\end{pmatrix}=8.$$

Exercise 4.5.7

(i) Observe that

$$\det\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix} = \det\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}^T$$

$$\begin{aligned}
&= \sum_{\sigma\in S_3}\prod_{i=1}^{3} a_{i\sigma(i)}\\
&= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33}\\
&\qquad\qquad + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}\\
&= a_{11}\det\begin{pmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{pmatrix} - a_{12}\det\begin{pmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{pmatrix} + a_{31}\det\begin{pmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{pmatrix}.
\end{aligned}$$

(ii) We have

$$\begin{aligned}
\det\begin{pmatrix}2 & 4 & 6\\ 3 & 1 & 2\\ 5 & 2 & 3\end{pmatrix} &= 2\det\begin{pmatrix}1 & 2\\ 2 & 3\end{pmatrix} - 4\det\begin{pmatrix}3 & 2\\ 5 & 3\end{pmatrix} + 6\det\begin{pmatrix}3 & 1\\ 5 & 2\end{pmatrix}\\
&= 2\times(-1) - 4(-1) + 6\times 1 = 8.
\end{aligned}$$

Exercise 4.5.8

(i) We look at each of the four terms $a_{r1} \det B_{r1}$ say. If $r \neq 1, j$ then $\det B_{r1}$ changes sign to give $-a_{r1} \det B_{r1}$. $a_{11} \det B_{11}$ changes to $-a_{j1} \det B_{j1}$ and $a_{j1} \det B_{j1}$ changes to $-a_{11} \det B_{11}$.

Thus
$$\mathcal{F}(\tilde{A}) = -\mathcal{F}(A).$$

(iii) The result is already proved if i or j takes the value 1. If not, interchanging row 1 with row i, then interchanging row 1 of the new matrix with row j of the new matrix and finally interchanging row 1 and i again transforms A to $\tilde{A}$. By part (i)
$$\mathcal{F}(\tilde{A}) = -\mathcal{F}(A).$$

If rows i and j are the same this gives $\mathcal{F}(A) = -\mathcal{F}(A)$ and so $\mathcal{F}(A) = 0$.

(iv) By inspection
$$\mathcal{F}(\tilde{A}) = \mathcal{F}(A) + \mathcal{F}(B)$$
where B is the matrix A with the first row replaced by the ith row. By part (iii), $\mathcal{F}(B) = 0$ so
$$\mathcal{F}(\tilde{A}) = \mathcal{F}(A).$$

By considering the effect of interchanging the first and jth row, we get the more general result
$$\mathcal{F}(\bar{A}) = \mathcal{F}(A).$$

(v) If we now carry out the diagonalisation procedure of Theorem 3.4.8 on A using the rules above and observe that
$$\mathcal{F}(I) = 1 = \det I,$$
we get
$$\mathcal{F}(A) = \det A.$$

Exercise 4.5.10

(i) The argument goes through essentially word for word (replacing 4 by n). The formula

$$\sum_{j=1}^{n} a_{1j} A_{1j} = \det A.$$

is the row expansion formula so obtained.

(ii) If $i = 1$ there is nothing to prove. If $i \neq 1$, we argue as follows.

If B is the matrix obtained from A by interchanging row 1 and row i.

$$\det A = -\det B = -\sum_{j=1}^{n} b_{1j} B_{1j} = \sum_{j=1}^{n} a_{ij} A_{ij}.$$

(iii) If $i \neq k$, then

$$\sum_{j=1}^{n} a_{ij} A_{kj} = \det C$$

where C is an $n \times n$ matrix with ith and kth rows the same. Thus

$$\sum_{j=1}^{n} a_{ij} A_{kj} = 0.$$

(iv) Conditions (ii) and (iii) together give

$$\sum_{j=1}^{n} a_{ij} A_{kj} = \delta_{kj} \det A$$

Exercise 4.5.13

We have

$$\det A \det A^{-1} = \det AA^{-1} = \det I = 1$$

so $\det A^{-1} = (\det A)^{-1}$.

Exercise 4.5.14

The formula of Exercise 4.5.10 (v) shows that $\sum_{k=1}^n b_k A_{kj}$ is the determinant of a matrix obtained from A by replacing the jth column of A by $\mathbf{b}$.

Thus

$$\begin{aligned}\sum_{j=1}^n a_{ij}\det B_j &= \sum_{j=1}^n a_{ij}\sum_{k=1}^n b_k A_{kj}\\ &= \sum_{j=1}^n\sum_{k=1}^n a_{ij}b_kA_{kj}\\ &= \sum_{k=1}^n\sum_{j=1}^n a_{ij}b_kA_{kj}\\ &= \sum_{k=1}^n b_k\sum_{j=1}^n a_{ij}A_{kj}\\ &= \sum_{k=1}^n b_k\delta_{ik}\det A = b_i\det A\end{aligned}$$

Thus, if $\det A \neq 0$,

$$x_j = \frac{\det B_j}{\det A}$$

gives a solution of $A\mathbf{x} = \mathbf{b}$ (and, since there is only one solution, it is the solution).

Exercise 4.5.15

The statement is true but you still need to find the determinant of two $n \times n$ matrices. Unless we deal with very small systems of equations (corresponding to a 3×3 matrix, say) the labour involved in computing $\det A$ (at least by the methods given in this book or any I am aware of) is, at best, comparable with the effort of finding all the solutions of $A\mathbf{x} = \mathbf{b}$

Exercise 4.5.16

(i) Observe that, since S_n is a group

$$\begin{aligned}\operatorname{perm} A^T &= \sum_{\sigma\in S_n}\prod_{i=1}^n a_{i\sigma(i)}\\ &= \sum_{\sigma\in S_n}\prod_{i=1}^n a_{\sigma^{-1}(i)i}\\ &= \sum_{\sigma\in S_n}\prod_{i=1}^n a_{\tau(i)i} = \operatorname{perm} A\end{aligned}$$

(ii) Both statements false. If we set

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

then $\operatorname{perm} A = 4 \neq 0$, but $\det A = 0$, and $\det B = -2$, but $\operatorname{perm} B = 0$.

(iii) Write A_{ji} for the $n-1\times n-1$ matrix obtained by removing the ith row and jth column from A. We have

$$\operatorname{perm} A = \sum_{i=1}^n a_{1i}\operatorname{perm} A_{i1}$$

(iv) By (iii),

$$\operatorname{perm} A \leq nK \max_i A_{i1}$$

so, by induction, $|\operatorname{perm} A| \leq n!K^n$.

If we take $A(n)$ to be the $n\times n$ matrix with all entries K (where $K\geq 0$), then

$$\operatorname{perm} A(n) = nK\operatorname{perm} A(n-1)$$

so, by induction, $\operatorname{perm} A(n) = n!K^n$.

(v) If B is the $n\times n$ matrix with $b_{ij} = |a_{ij}|$

$$|\det A| \leq \operatorname{perm} B \leq n!K^n.$$

Hadamard's inequality is proved later.

Exercise 4.5.17

(i) True. If A is antisymmetric,

$$\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix},$$

and $\det A = b^2 \neq 0$ if $A \neq 0$.

(ii) False. Consider

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(iii) True.

$$\det A = \det A^T = \det(-A) = (-1)^n \det A = -\det A.$$

Exercise 5.1.1

If we work over $\mathbb{Z}$ we cannot always divide.

Thus the equation $2x = 1$ has no solution, although the 1×1 matrix (2) has non-zero determinant.

Exercise 5.1.2★

EXERCISE 5.2.7

I would be inclined to pick (iv) and (vi).

(ii) $(f+g)(x) = f(x) = g(x) = g(x) + f(x) = (g+f)(x)$ for all x so $f + g = g + f$.

(iii) $(f+0)(x) = f(x)+0(x) = f(x)+0 = f(x)$ for all x so $f+0 = f$.

(iv) $(\lambda(f+g))(x) = \lambda((f+g)(x)) = \lambda(f(x)+g(x)) = \lambda f(x)+\lambda g(x) = (\lambda f)(x) + (\lambda g)(x) = (\lambda f + \lambda g)(x)$ for all x so $\lambda(f + g) = \lambda f + \lambda g$.

(v) $((\lambda + \mu)f)(x) = (\lambda + \mu)(f(x)) = \lambda f(x) + \mu f(x) = (\lambda f)(x) + (\mu g)(x) + (\lambda f + \mu g)(x)$ for all x so $(\lambda + \mu)f = \lambda f + \mu f$.

(vi) $((\lambda\mu)f)(x) = (\lambda\mu)(f(x)) = (\lambda\mu)f(x) = \lambda(\mu f(x)) = \lambda((\mu f)(x)) = (\lambda(\mu f))(x)$ for all x so $(\lambda\mu)f = \lambda(\mu f)$.

(vii) $(1f)(x) = 1 \times f(x) = f(x)$ and $(0f)(x) = 0 \times f(x) = 0 = 0(x)$ for all x and so $1f = f$ and $0f = 0$.

Exercise 5.2.8

The correspondence $\mathbb{F}^X \leftrightarrow \mathbb{F}^n$ given by

$$f \leftrightarrow (f(1), f(2), f(3), \ldots, f(n))$$

identifies $\mathbb{F}^X$ with the known vector space $\mathbb{F}^n$.

Exercise 5.2.11

Observe that all these sets are subsets of the vector space $\mathbb{R}^{\mathbb{R}}$ so we may use Lemma 5.2.10.

(i) Subspace. If f and g are 3 times differentiable so is $\lambda f + \mu g$.

(ii) Not a vector space. Let $f(t) = t^2$. Then f is in the set but $(-1)f$ is not.

(iii) Not a vector space. Let $P(t) = t$. Then P is in the set but $(-1)P$ is not.

(iv) Subspace. If P, Q in set, then $\lambda P + \mu Q$ is a polynomial and $(\lambda P + \mu Q)'(1) = \lambda P'(1) + \mu Q'(1) = 0$

(v) Subspace. If P, Q in set, then $\lambda P + \mu Q$ is a polynomial and

$$\begin{aligned}\int_0^1 (\lambda P + \mu Q)(t)\,dt &= \int_0^1 (\lambda P(t) + \mu Q(t)\,dt \\ &= \lambda \int_0^1 P(t)\,dt + \mu \int_0^1 q(t)\,dt = 0.\end{aligned}$$

(vi) Not a vector space. Let

$$h(t) = \max\{(1 - 10|t|), 0\}$$

and observe that $\int h(t)^3\,dt = A \neq 0$. If $f(t) = h(t) - h(t + 1/5)$ and $g(t) = h(t) - h(t + 2/5)$, then

$$\int_{-1}^1 f(t)^3\,dt = \int_{-1}^1 g(t)^3\,dt = A - A = 0$$

but

$$\int_{-1}^1 (f(t) + g(t))^3\,dt = 8A - 2A \neq 0.$$

(vii) Not a vector space. If $f(t) = -g(t) = t^3$ then f and g have degree exactly 3 but $f + g = 0$ does not.

(viii) Not a vector space. If $P(t) = t^2 - t$, $Q(t) = -t^2$, then P and Q are in the set, but not $P + Q$.

Exercise 5.3.2

$$T(\mathbf{0}) = T(0\mathbf{0}) = 0T(\mathbf{0}) = \mathbf{0}.$$

Exercise 5.3.3

Let $f,\ g \in \mathcal{D}$

(i) We have
$$\delta(\lambda f + \mu g) = (\lambda f + \mu g)(0) = \lambda f(0) + \mu g(0) = \lambda \delta f + \mu \delta g.$$

(ii) We have
$$D(\lambda f + \mu g) = (\lambda f + \mu g)' = \lambda f' + \mu g' = \lambda D f + \mu D g.$$

(iii) We have
$$\begin{aligned}\big(K(\lambda f + \mu g)\big)(x) &= (x^2+1)(\lambda f + \mu g)(x) = (x^2+1)(\lambda f(x) + \mu g(x))\\ &= \lambda(x^2+1)f(x) + \mu(x^2+1)g(x)\\ &= \lambda(Kf)(x) + \mu(Kg)(x) = (\lambda Kf + \mu Kg)(x)\end{aligned}$$
for all x and so
$$K(\lambda f + \mu g) = \lambda K f + \mu K g.$$

(iv) We have
$$\begin{aligned}\big(J(\lambda f + \mu g)\big)(x) &= \int_0^x (\lambda f(t) + \mu g(t))\,dt\\ &= \lambda \int_0^x f(t)\,dt + \mu \int_0^x g(t)\,dt = (\lambda Jf + \mu Jg)(x)\end{aligned}$$
for all x and so
$$J(\lambda f + \mu g) = \lambda J f + \mu J g.$$

Exercise 5.3.10

(i) We have

$$\iota(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\mathbf{x} + \mu\mathbf{y} = \lambda\iota\mathbf{x} + \mu\iota\mathbf{y}$$

so ι is linear

(ii) Just a restatement of definitions.

(iii) The fundamental theorem of the calculus states that $DJ = \iota$. However $JD1 = J0 = 0 \neq 1$. To see that J is injective observe that

$$Jf = Jg \Rightarrow DJf = DJg \Rightarrow f = g$$

However $Jf(0) = 0$ so $1 \notin J\mathcal{D}$ so J is not surjective.

To see that D is surjective observe that $D(Jf) = f$. However $D0 = D1 = 0$ so D is not injective.

(iv) Observe that

$$(\alpha\beta)(\beta^{-1}\alpha^{-1}) = \alpha(\beta\beta^{-1})\alpha^{-1} = \alpha\iota\alpha^{-1} = \alpha^{-1}\alpha = \iota$$

and similarly

$$(\beta^{-1}\alpha^{-1})(\alpha\beta) = \iota$$

Exercise 5.3.11

The only if part is trivial

To prove the if part suppose

$$T\mathbf{u} = \mathbf{0} \Rightarrow \mathbf{u} = \mathbf{0}.$$

Then

$$\begin{aligned} T\mathbf{x} = T\mathbf{y} &\Rightarrow T\mathbf{x} - T\mathbf{y} = \mathbf{0} \\ &\Rightarrow T(\mathbf{x} - \mathbf{y}) = \mathbf{0} \\ &\Rightarrow \mathbf{x} - \mathbf{y} = \mathbf{0} \\ &\Rightarrow \mathbf{x} = \mathbf{y}. \end{aligned}$$

Exercise 5.3.14

(i) If α, $\beta \in GL(U)$, then α and β are bijective, so $\alpha\beta$ is, so $\alpha\beta \in GL(U)$.

(ii) Observe that if α, β, $\gamma \in GL(U)$, then

$$\big((\alpha\beta)\gamma\big)\mathbf{u} = (\alpha\beta)(\gamma\mathbf{u}) = \Bigg(\alpha\big(\beta(\gamma\mathbf{u})\big)\Bigg) = \alpha\big(\beta(\gamma\mathbf{u})\big) = \alpha\big((\beta\gamma)\mathbf{u}\big) = \big(\alpha(\beta\gamma)\big)\mathbf{u}$$

for all $\mathbf{u} \in U$ so $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.

(iii) $\iota \in GL(U)$ and $\alpha\iota = \iota\alpha = \alpha$ for all $\alpha \in GL(U)$.

(iv) Use the definition of $GL(U)$.

Let $\alpha(x, y) = (x + y, x - y)$, $\beta(x, y) = (y, x)$. Then $\alpha\beta(x, y) = (x + y, y - x)$, $\beta\alpha(x, y) = (x - y, x + y)$. (Could use matrices.)

Exercise 5.3.16

(i) A subgroup. If $\alpha(\mathbf{a}) = \beta\mathbf{a} = \mathbf{a}$, then $\iota\mathbf{a} = \mathbf{a}$, $\alpha^{-1}(\mathbf{a}) = \mathbf{a}$ and

$$(\alpha\beta)\mathbf{a} = \alpha(\beta\mathbf{a}) = \alpha\mathbf{a} = \mathbf{a}.$$

(ii) Not necessarily a group. Work in $\mathbb{R}^2$. Let $\alpha(x, y) = (x, 0)$, $\mathbf{a} = (1, 0)$, $\beta(x, y) = (0, y)$, $\mathbf{b} = (0, 1)$.

(iii) Not necessarily a group. Work in $\mathbb{R}^3$. Let *alpha* be a rotation through π about the x axis, β a rotation through π about an axis running through $(0, 0, 0)$ and $(1, 1, 1)$ $\alpha^2 = \beta^2 = \iota$ but by considering $(\alpha\beta\alpha\beta(1, 1, 1)$ we see that $(\alpha\beta)^2 \neq \iota$.

Exercise 5.4.3

Observe that

$$\begin{aligned}\sum_{j=1}^{n}\lambda_j(\mathbf{e}_j+\mathbf{y})=\mathbf{0} &\Rightarrow \sum_{j=1}^{n}\lambda_j\mathbf{e}_j+\sum_{j=1}^{n}\sum_{k=1}^{n}\lambda_j a_k\mathbf{e}_k=\mathbf{0}\\ &\Rightarrow \sum_{j=1}^{n}\lambda_j\mathbf{e}_j+\sum_{j=1}^{n}\sum_{k=1}^{n}\lambda_k a_j\mathbf{e}_j=\mathbf{0}\\ &\Rightarrow \sum_{j=1}^{n}\left(\lambda_j+a_j\sum_{r=1}^{n}\lambda_r\right)\mathbf{e}_j=\mathbf{0}\\ &\Rightarrow \lambda_j+a_j\sum_{r=1}^{n}\lambda_r=0\ \forall j\end{aligned}$$

Thus, if

$$\sum_{j=1}^{n}\lambda_j\mathbf{e}_j+\mathbf{y}=\mathbf{0}$$

and we write $K=-\sum_{r=1}^{n}\lambda_r$,

$$\lambda_j=Ka_j.$$

Summing we obtain

$$-K=\sum_{j=1}^{n}\lambda_j=K\sum_{j=1}^{n}a_j$$

so, $a_1+a_2+\ldots+a_n+1\neq 0$, $K=0$ and $\lambda_j=0$ showing that the vectors $\mathbf{e}_1+\mathbf{y}$ are linearly independent.

If $a_1+a_2+\ldots+a_n+1=0$, then

$$\sum_{j=1}^{n}a_j(\mathbf{e}_j+\mathbf{y})=\mathbf{0}$$

and (since $\mathbf{y}\neq\mathbf{0}$) not all the a_j are zero, so we do not have linear independence.

Exercise 5.4.11

U and V are the null spaces of linear maps from $\mathbb{R}^4$ to $\mathbf{R}^2$.

The elements of $U \cap V$ are given by

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

Subtracting the 4th equation from the 2nd yields $x = 0$ so the system becomes

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

Adding the 1st and 3rd equation reveals that the 4th equation is superfluous so the system is

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

or equivalently

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

so $x = 0$, $z = 2t$, $y = -2t$. Thus $U \cap V$ has basis $\mathbf{e}_1 = (0, -2, 2, 1)^T$.

The equations for U yield

$$\begin{aligned} x + y - 2z + t &= 0 \\ 2y - z - 2t &= 0 \end{aligned}$$

so by inspection $\mathbf{e}_2 = (1, 1, 2, 0)^T$ is in U. Since $\mathbf{e}_1$ and $\mathbf{e}_2$ are linearly independent and $\dim U = 2$ we have basis for U

(To see that $\dim U = 2$ either quote general theorems or observe that x and y determine z and t.)

The equations for v yield

$$\begin{aligned} x - 2y + z + 2t &= 0 \\ y + z - 2t &= 0 \end{aligned}$$

so, by inspection, $\mathbf{e}_3 = (3, 1, 2, 0)^T$ is in U. Since $\mathbf{e}_1$ and $\mathbf{e}_3$ are linearly independent and $\dim V = 2$ (argue as before) we have basis for V.

The proof of Lemma 5.4.10 now shows that $\mathbf{e}_1$, $\mathbf{e}_2$ form a basis for $\mathbf{e}_3$

Exercise 5.4.12

(i) Since $U \supseteq V + W \supseteq V, W$ we have

$$\dim U \geq \dim(V+W) \geq \dim V, \dim W$$

By Lemma 5.4.10

$$\dim(V+W) = \dim V + \dim W - \dim(V \cap W) \geq \dim V + \dim W.$$

Putting these two results together, we get

$$\min\{\dim U, \dim V + \dim W\} \geq \dim(V+W) \geq \max\{\dim V, \dim W\}.$$

(ii) Consider a basis $\mathbf{e}_1,\ \mathbf{e}_2,\ \ldots,\ \mathbf{e}_n$ for U.

Let

$$V = \operatorname{span}\{\mathbf{e}_1,\ \mathbf{e}_2,\ \ldots,\ \mathbf{e}_r\}$$
$$W = \operatorname{span}\{\mathbf{e}_{t-s+1},\ \mathbf{e}_{t-s+2},\ \ldots,\ \mathbf{e}_t\}$$

Then

$$V + W = \operatorname{span}\{\mathbf{e}_1,\ \mathbf{e}_2,\ \ldots,\ \mathbf{e}_t\}$$

and

$$\dim V = r,\ \dim W = s,\ \dim(V+W) = t.$$

Exercise 5.4.13

We use row vectors.

E is the null space of a linear map from $\mathbb{R}^3$ to $\mathbb{R}$ so a subspace.

Let $\mathbf{e}_1 = (1, -1, 0)$, $\mathbf{e}_2 = (1, 0, -1)$. We have $\mathbf{e}_1$, $\mathbf{e}_2 \in U$.

$$x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = \mathbf{0} \Rightarrow (x_1 + x_2, -x_1, -x_2) = 0 \Rightarrow x_1 = x_2 = 0$$

so $\mathbf{e}_1$, $\mathbf{e}_2$ are independent.

If $\mathbf{x} \in U$ then

$$\mathbf{x} = (x_1, x_2, x_3) \Rightarrow \mathbf{x} = (-x_2 - x_3, x_2, x_3) \Rightarrow \mathbf{x} = x_2\mathbf{e}_1 + x_3\mathbf{e}_2.$$

Thus $\mathbf{e}_1$, $\mathbf{e}_2$ span U and so form a basis.

I do not think everybody will choose the basis and so there cannot be a ‘standard basis’ in this case.

Exercise 5.4.14

(i) $\mathbf{0} \in V$ for all $V \in \mathcal{V}$ and so $\mathbf{0} \in \bigcap_{V \in \mathcal{V}} V$.

Further, if λ, $\mu \in \mathbb{F}$,

$$\begin{aligned}\mathbf{u},\, \mathbf{v} \in \bigcap_{V \in \mathcal{V}} V &\Rightarrow \mathbf{u},\, \mathbf{v} \in V \ \forall V \in \mathcal{V}\\ &\Rightarrow \lambda\mathbf{u} + \mu\mathbf{v} \in V \ \forall V \in \mathcal{V}\\ &\Rightarrow \lambda\mathbf{u} + \mu\mathbf{v} \in \bigcap_{V \in \mathcal{V}} V.\end{aligned}$$

Thus $\bigcap_{V \in \mathcal{V}} V$ is a subspace of U.

(ii) By (i) W is a subspace of U. Since $E \subseteq V$ for all $V \in \mathcal{V}$, we have $E \subseteq W$. If W' is as stated $W' \in \mathcal{V}$ so $W' \supseteq W$.

(iii) Let

$$W' = \left\{ \sum_{j=1}^{n} \lambda_j \mathbf{e}_j \ : \ \lambda_j \in \mathbb{F} \text{ for } 1 \leq j \leq n \right\}.$$

Observe that $E \subseteq W'$ (take $\lambda_j = \delta_{ij}$). and W is a subspace of U since

$$\lambda \sum_{j=1}^{n} \lambda_j \mathbf{e}_j + \mu \sum_{j=1}^{n} \mu_j \mathbf{e}_j = \sum_{j=1}^{n} (\lambda\lambda_j + \mu\mu_j \mathbf{e}_j.$$

Any subspace of U containing E must certainly contain W'. Thus $W = W'$.

Exercise 5.5.3

Always false. $\mathbf{0} \notin \alpha^{-1}\mathbf{v}$.

Exercise 5.5.5

We use the rank-nullity theorem.

$$\begin{aligned}\alpha \text{ injective} &\Leftrightarrow \alpha^{-1}(\mathbf{0}) = \mathbf{0}\\ &\Leftrightarrow \dim \alpha^{-1}(\mathbf{0}) = 0\\ &\Leftrightarrow n - \dim \alpha^{-1}(\mathbf{0}) = n\\ &\Leftrightarrow \dim \alpha(U) = n\\ &\Leftrightarrow \alpha \text{ surjective}\end{aligned}$$

We now note that bijective means injective and surjective and that a map is bijective if and only if it is invertible.

Exercise 5.5.6

Adding the second and third row to the top row, we get

$$\det\begin{pmatrix}a&b&b\\b&a&b\\b&b&a\end{pmatrix}=\det\begin{pmatrix}a+2b&a+2b&a+2b\\b&a&b\\b&b&a\end{pmatrix}$$

$$=(a+2b)\begin{pmatrix}1&1&1\\b&a&b\\b&b&a\end{pmatrix}=(2a+b)\begin{pmatrix}1&1&1\\0&a-b&0\\0&0&a-b\end{pmatrix}=(2a+b)(a-b)^2.$$

Thus, if $a\neq b$ and $a\neq -2b$, α has rank 3 $\alpha^{-1}(\mathbf{0})=\mathbf{0}$, has empty basis and $\alpha(U)=U$ has basis $(1,0,0)^T$, $(0,1,0)^T$, $(0,0,1)^T$.

If $a=b$, there are two cases. If $a=0$, α has rank 3 $\alpha(U)=\mathbf{0}$ has empty basis and $\alpha^{-1}(\mathbf{0})=U$ has basis $(1,0,0)^T$, $(0,1,0)^T$, $(0,0,1)^T$.

If $a=b\neq 0$, then

$$(x,y,z)\in\alpha^{-1}(\mathbf{0})\Leftrightarrow x+y+z=0$$

so α has rank $3-2=1$ $\alpha^{-1}(\mathbf{0})$ has basis $(1,0,-1)^T$, $(0,1,-1)^T$. U has basis $(1,0,-1)^T$, $(0,1,-1)^T$, $(0,0,1)^T$, so $\alpha(U)$ has basis $\alpha(0,0,1)^T=(a,a,a)^T$ so basis $(1,1,1)^T$.

If $a=-2b\neq 0$, then

$$(x,y,z)\in\alpha^{-1}(\mathbf{0}\Leftrightarrow x-2y+z=0,\ x+y-2z=0\Leftrightarrow x=y=z$$

so α has rank $3-1=2$ $\alpha^{-1}(\mathbf{0})$ has basis $(1,1,1)^T$. U has basis $(1,1,1)^T$, $(0,1,0)^T$, $(0,0,1)^T$ so $\alpha(U)$ has basis $\alpha(0,1,0)^T=(b,a,b)^T$ $\alpha(0,1,0)^T=(b,b,a)^T$ so basis $(1,-2,1)^T$, $(1,1,-2)$.

Exercise 5.5.12★

Exercise 5.5.13

(i) If A is an $m \times n$ matrix with rows $\mathbf{a}_1$, $\mathbf{a}_2$, $\dots \mathbf{a}_m$ then the row rank of A is dimension of

$$\operatorname{span}\{\mathbf{a}_1,\, \mathbf{a}_2,\, \dots,\, \mathbf{a}_m\}.$$

(ii) If B is a non-singular $m \times m$ matrix then if $\mathbf{x}$ is a row vector with m entries

$$\mathbf{x}B = \mathbf{0} \Rightarrow \mathbf{x}BB^{-1} = \mathbf{0}B^{-1} \Rightarrow \mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x}B = \mathbf{0}$$

so

$$\mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{x}B = \mathbf{0}.$$

It follows that

$$\begin{aligned}\sum_{j=1}^{m} \lambda_j \mathbf{a}_j B = \mathbf{0} &\Leftrightarrow \left(\sum_{j=1}^{m} \lambda_j \mathbf{a}_j\right) B = \mathbf{0}\\ &\Leftrightarrow \sum_{j=1}^{m} \lambda_j \mathbf{a}_j = \mathbf{0},\end{aligned}$$

so

$$\dim \operatorname{span}\{\mathbf{a}_1 B,\, \mathbf{a}_2 B,\, \dots,\, \mathbf{a}_m B\} = \dim \operatorname{span}\{\mathbf{a}_1,\, \mathbf{a}_2,\, \dots,\, \mathbf{a}_m\}$$

and the row rank of A and AB are the same.

Since the elementary column operations correspond to post multiplication by invertible matrices, the row rank is unaltered by elementary column operations. A similar argument with pre-multiplication replacing post-multiplication shows that row rank is unaltered by elementary row operations. Similarly the column rank is unaltered by elementary row and column operations.

(iii) By elementary row and column operations any matrix A can be transformed to a matrix $C = (c_{ij})$ with $c_{ij} = 0$ if $i \neq j$. By inspection, C has row and column rank equal to the number of non-zero entries c_{ii}. Thus, by (ii), the row rank and column rank of A are the same.

Exercise 5.5.14

If $a = b = 0$, then $r = 0$.

If $a \neq 0$ and $b = 0$, the matrix is

$$\begin{pmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ a & 0 & 0 & 0 \\ 0 & 0 & 0 & a \end{pmatrix}$$

The first and third rows are identical, so $r \leq 3$. However, the matrix obtained by removing the third row and first column

$$\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}$$

is non-singular, so $r = 3$.

If $a = 0$, $b \neq 0$, we get

$$\begin{pmatrix} 0 & 0 & b & 0 \\ 0 & 0 & 0 & b \\ 0 & 0 & 0 & b \\ 0 & b & 0 & 0 \end{pmatrix}.$$

The second and third rows are identical, so $r \leq 3$. However, the matrix obtained by removing the third row and first column

$$\begin{pmatrix} 0 & b & 0 \\ 0 & 0 & b \\ b & 0 & 0 \end{pmatrix}$$

is non-singular (for example, because the determinant is non-zero), so $r = 3$.

If $a = b \neq 0$, the matrix is

$$\begin{pmatrix} a & 0 & a & 0 \\ 0 & a & 0 & a \\ a & 0 & 0 & a \\ 0 & a & 0 & a \end{pmatrix}.$$

The second and fourth rows are identical so $r \leq 3$. However, the matrix obtained by removing the fourth row row and column

$$\begin{pmatrix} a & 0 & a \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}$$

is non-singular (for example, because the determinant is non-zero) so $r = 3$.

If a, $b \neq 0$ and $a \neq b$, then subtracting the first row from the third row and the a/b times the second row from the fourth

$$\begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ a & 0 & 0 & b \\ 0 & b & 0 & a \end{pmatrix} \to \begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ 0 & 0 & -b & b \\ 0 & b & 0 & a \end{pmatrix} \to \begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ 0 & 0 & -b & b \\ 0 & 0 & 0 & a - b^2/a \end{pmatrix}$$

yields a non-singular matrix (e.g. by looking at the determinant), so $r = 4$ and the original matrix was non-singular.

Exercise 5.5.16

Consider the linear maps α and β whose matrices with respect to a given basis are (a_{ij}) with $a_{ii} = 1$ if $1 \leq i \leq r$, $a_{ij} = 0$, otherwise, $b_{ii} = 1$ if $r - t + 1 \leq i \leq r + s - t$, $b_{ij} = 0$, otherwise. (Note that we use the facts that r, $s \geq t$ and $r + s - t \leq n$.) Then AB has matrix (c_{ij}) with $c_{ii} = 1$ if $r - t + 1 \leq i \leq r$, $c_{ij} = 0$, otherwise. Then α has rank r, β rank s and γ rank t.

Exercise 5.5.17

By noting that $(1, -1, 0, 0, \ldots, 0)^T$, $(1, 0, -1, 0, \ldots, 0) \ldots$, $(1, 0, 0, 0, \ldots, -1)^T$ are $n - 1$ linearly independent eigenvectors with eigenvalue 0 and $(1, 1, \ldots, 1)^T$ is an eigenvector with eigenvalue n for the $n \times n$ matrix A with all entries 1, we see that the characteristic polynomial of A is $\det(tI - A) = t^{n-1}(t - n)$. Thus $A + sI$ is invertible if and only if $s \neq 0, -n$.

Let P be as stated. Then $PP^T = Q$ with

$$q_{ij} = \sum_{r=1}^{n} p_{ir}p_{jr} = \begin{cases} 1 & \text{if } i \neq j \\ k & \text{if } i = j \end{cases}$$

Thus $PP^T = A + (k - 1)I$ and, since $k \geq 2$ we have PP^T invertible so P is invertible so of full rank n.

If $k = 1$, then there can only be one party $P = (1\ 1\ 1\ \ldots\ 1)^T$ and has rank 1. The proof fails because $PP^T = A$ is not of full rank.

Exercise 5.5.18

(i) Observe that $(\alpha+\beta)U \subseteq V$, so $\dim V \geq \dim(\alpha+\beta)U$.

Since $(\alpha+\beta)\mathbf{u} = \alpha\mathbf{u}+\beta\mathbf{u} \in \alpha U+\beta U$, we have $(\alpha+\beta)U \subseteq \alpha U+\beta U$, so, by Lemma 5.4.10,

$$\dim(\alpha+\beta)U \leq \dim(\alpha U+\beta U) \leq \dim \alpha U + \dim \beta U.$$

Thus

$$\min\{\dim V, \operatorname{rank}\alpha + \operatorname{rank}\beta\} \geq \operatorname{rank}(\alpha+\beta).$$

(ii) By (i),

$$\operatorname{rank}(\alpha+\beta)+\operatorname{rank}\beta = \operatorname{rank}(\alpha+\beta)+\operatorname{rank}(-\beta) \geq \operatorname{rank}\big((\alpha+\beta)-\beta\big) = \operatorname{rank}\alpha.$$

Thus

$$\operatorname{rank}(\alpha+\beta) \geq \operatorname{rank}\alpha - \operatorname{rank}\beta$$

and

$$\operatorname{rank}(\alpha+\beta) = \operatorname{rank}(\beta+\alpha) \geq \operatorname{rank}\beta - \operatorname{rank}\alpha.$$

Thus

$$\operatorname{rank}(\alpha+\beta) \geq |\operatorname{rank}\alpha - \operatorname{rank}\beta|.$$

(iii) Since $\alpha+\beta = \beta+\alpha$, there is no loss of generality in assuming $r \geq s$.

Fix bases for U and V. If $t \geq r$, consider the linear map α given by the matrix (a_{ij}) with $a_{ii} = 1$ if $1 \leq i \leq r$, $a_{ij} = 0$ otherwise, and the linear map β given by the matrix (b_{ij}) with $b_{ii} = 1$ if $t-s+1 \leq i \leq t$, $b_{ij} = 0$ otherwise. Then $\alpha+\beta$ has matrix (c_{ij}) with $c_{ii} = 1$ if $1 \leq i \leq t-s$ $c_{ii} = 2$ if $t-s+1 \leq i \leq r$, $c_{ii} = 1$ if $r+1 \leq i \leq t$, $c_{ij} = 0$ otherwise. Thus $\operatorname{rank}\alpha = r$, $\operatorname{rank}\beta = s$ $\operatorname{rank}(\alpha+\beta) = t$.

If $t \leq r-1$, consider the linear map α given by the matrix (a_{ij}) with $a_{ii} = 1$ if $1 \leq i \leq r$, $a_{ij} = 0$ otherwise, and the linear map β given by the matrix (b_{ij}) with $b_{ii} = -1$ if $1 \leq i \leq r-t$, $b_{ii} = 1$ if $r-t+1 \leq i \leq s$, $b_{ij} = 0$ otherwise. Then $\alpha+\beta$ has matrix (c_{ij}) with $c_{ii} = 2$ if $r-t+1 \leq i \leq s$ $c_{ii} = 1$ if $s+1 \leq i \leq r$, $c_{ij} = 0$ otherwise. Thus $\operatorname{rank}\alpha = r$, $\operatorname{rank}\beta = s$ $\operatorname{rank}(\alpha+\beta) = t$.

Exercise 5.5.19

To avoid repetition, we go straight to (iii). We consider $n\times n$ matrices and say $A \in \Gamma$ if and only if there is a K such that $\sum_{r=1}^{n} a_{rj} = K$ for all j and $\sum_{r=1}^{n} a_{ir} = K$ for all i.

(i)′ Observe that

$$\begin{aligned} A,\, B \in \Gamma &\Rightarrow \sum_{r=1}^{n} a_{rj} = \sum_{r=1}^{n} a_{ir},\ \sum_{i=1}^{n} b_{ij} = \sum_{j=1}^{n} b_{ij}\ \forall i,\, j \\ &\Rightarrow \sum_{r=1}^{n} \lambda a_{rj} + \mu b_{rj} = \sum_{r=1}^{n} \lambda_{ir} + \mu b_{ir}\ \forall i,\, j \\ &\lambda A + \mu B \in \Gamma. \end{aligned}$$

Since $0 \in \Gamma$, we have γ a vector space.

(ii)′ Let Γ_0 be the subset of Γ such that

$$\sum_{r=1}^{n} a_{rj} = \sum_{r=1}^{n} a_{ir} = 0.$$

We observe that Γ_0 is a subspace of Γ since $0 \in \Gamma_0$ and

$$\begin{aligned} A,\, B \in \Gamma &\Rightarrow \sum_{r=1}^{n} a_{rj} = \sum_{r=1}^{n} a_{ir} = 0,\ \sum_{i=1}^{n} b_{ij} = \sum_{j=1}^{n} b_{ij} = 0\ \forall i,\, j \\ &\Rightarrow \sum_{r=1}^{n} \lambda a_{rj} + \mu b_{rj} = \sum_{r=1}^{3} \lambda_{ir} + \mu b_{ir} = 0\ \forall i,\, j \\ &\lambda A + \mu B \in \Gamma_0. \end{aligned}$$

We note that if E is the $n \times n$ matrix all of whose entries are 1 then any basis of Γ_0 together with E gives a basis for Γ.

If $1 \leq i, j \leq n-1$, let $E(i,j)$ be the $n \times n$ matrix with entry 1 in the (i,j)th place and (n,n)th place and entry -1 in the (i,n)th place and (n,j)th place. We observe that $E(i,j) \in \Gamma_0$.

We observe by looking at the (r,s)th entry that

$$\sum_{1\leq i,j\leq n-1} a_{ij} E(i,j) = 0 \Rightarrow a_{rs} = 0\ \forall 1 \leq r, s \leq n-1,$$

so the E_{rs} are linearly independent.

If $A = (a_{ij})$ then

$$A - \sum_{1\leq i,j\leq n-1} a_{ij} E(i,j) = 0$$

We check the (n,n)th entry by observing that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} = 0$$

Thus the $E(i,j)$ form a basis for Γ_0 and the $E_{i,j}$ $[1 \leq i,j \leq n-1]$ together with E form a basis for Γ. Thus Γ has dimension $(n-1)^2+1 = n^2-2n+1$

I cannot see any natural basis.

Exercise 5.6.2

We have

$$\begin{aligned} P(r) &\equiv b_0 + b_1 c_r \\ P(r) &\equiv 2 + 2c_r \\ P(1) &\equiv 2 + 2 \times 2 \equiv 6 \\ P(2) &\equiv 2 + 2 \times 4 \equiv 3 \\ P(3) &\equiv 2 + 2 \times 5 \equiv 5 \end{aligned}$$

(all modulo 7).

$(P(1), c(1))$ and $(P(2), c(2))$ yield

$$\begin{aligned} 6 &\equiv b_0 + 2b_1 \\ 3 &\equiv b_0 + 4b_1 \end{aligned}$$

Subtracting the first equation from the second yields $3 \equiv -2b_1$. Euclid's algorithm (or inspection) tells us to multiply by 3 to recover $b_1 = 2$. Substitution yields $b_0 = 2$.

$(P(1), c(1))$ and $(P(3), c(3))$ yield

$$\begin{aligned} 6 &\equiv b_0 + 2b_1 \\ 5 &\equiv b_0 + 5b_1 \end{aligned}$$

Subtracting the first equation from the second yields $-1 \equiv -3b_1$. Euclid's algorithm (or inspection) tells us to multiply by 5 to recover $b_1 = 2$. Substitution yields $b_0 = 2$.

Exercise 5.6.3

If $c_j = c_k$ with $j \neq k$, then two people have the same secret. More people are needed to find the full secret.

The Vandermonde determinant

$$\det \begin{pmatrix} 1 & c_{r(1)} & c_{r(1)}^2 & \dots & c_{r(1)}^{k-1} \\ 1 & c_{r(2)} & c_{r(2)}^2 & \dots & c_{r(2)}^{k-1} \\ 1 & c_{r(3)} & c_{r(3)}^2 & \dots & c_{r(3)}^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & c_{r(k)} & c_{r(k)}^2 & \dots & c_{r(k)}^{k-1} \end{pmatrix}$$

does not vanish (and so the system is soluble) if and only if the $c_{r(j)}$ are distinct.

If $c_j = 0$ then $P(j) = b_0$ and j knows the secret.

The proof that fewer cannot find the secret depended on

$$\det \begin{pmatrix} c_{r(1)} & c_{r(1)}^2 & \dots & c_{r(1)}^{k-1} & \\ c_{r(2)} & c_{r(2)}^2 & \dots & c_{r(2)}^{k-1} & \\ c_{r(3)} & c_{r(3)}^2 & \dots & c_{r(3)}^{k-1} & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{r(k-1)} & c_{r(k-1)}^2 & \dots & c_{r(k-1)}^{k-1} & \end{pmatrix} \equiv c_{r(1)} c_{r(2)} \dots c_{r(k-1)} \prod_{1 \le j < i \le k-1} (c_{r(i)} - c_{r(j)}) \not\equiv 0$$

and this needs the $c_{r(j)}$ distinct and non-zero.

Exercise 5.6.4

$$
\begin{aligned}
P(r) &\equiv b_0 + b_1 c_r \\
P(r) &\equiv 1 + c_r \\
P(1) &\equiv 1 + 1 \times 1 \equiv 2 \\
P(2) &\equiv 1 + 1 \times 4 \equiv 5
\end{aligned}
$$

all modulo 6. The recovery equations are

$$
\begin{aligned}
2 &\equiv b_0 + b_1 \\
5 &\equiv b_0 + 4b_1
\end{aligned}
$$

There are two solutions $(b_0, b_1)) = (1, 1)$ and $(b_0, b_1)) = (5, 3)$.

If p is not a prime, $mn \equiv k$ may have more than one solution in n even if $k \not\equiv 0$.

Exercise 5.6.5

(i) The secret remains safe. The arguments continue to work if the c_j are known.

(ii) If the b_s are known for $r \neq 0$ the holder of any pair $(c_r, P(r))$ can compute b_0 directly as

$$b_0 \equiv P(r) - \sum_{s=1} b_s c_r^s.$$

Exercise 5.6.6

With k rings, they can say nothing with high probability since one of the rings may be a fake so they would have only $k-1$ truth telling rings.

With $k+1$ rings, either every set of k rings gives the same answer in which case all the rings are true and the answer is correct or each set of k rings will give a different answer (one of which corresponding to the true k rings will be correct). All they can say is that one ring is fake but they do not know which one or what the Integer of Power is.

With $k+2$ rings, there are $(k+2)(k+1)/2$ different sets of k rings. Either each set tells the same story in which case all the rings are true and the answer is correct or $k+1$ sets will tell the same story in which case they are the $k+1$ sets containing k true rings and their answer is correct and $k(k-1)/2$ sets giving different answers. The heroes can identify the $k+1$ correct rings since they belong to the truth telling sets.

One prime to rule them all
One prime to bind them
In the land of composites
Where the factors are

Exercise 6.1.2

Observe that, writing $A = (a_{ij})$, we have

$$\alpha \mathbf{e}_j = \sum_{i=1}^{n} a_{ij}\mathbf{e}_i = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}.$$

Exercise 6.1.3

Let $A = (a_{ij})$, $B = (b_{ij})$ and let $\alpha\beta$ have matrix $C = (c_{ij})$.

Then

$$\begin{aligned}\sum_{i=1}^{n} c_{ij}\mathbf{e}_i &= \alpha\beta\mathbf{e}_j = \alpha(\beta\mathbf{e}_j)\\ &= \alpha\left(\sum_{k=1}^{n} b_{kj}\mathbf{e}_k\right)\\ &= \sum_{k=1}^{n} b_{kj}\alpha\mathbf{e}_k = \sum_{k=1}^{n} b_{kj}\sum_{r=1}^{n} a_{rk}\mathbf{e}_r \quad = \sum_{r=1}^{n}\left(\sum_{k=1}^{n} a_{rk}b_{kj}\right)\mathbf{e}_r\end{aligned}$$

so, by the uniqueness of the basis expansion,

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

i.e. $C = AB$.

Exercise 6.1.5

$$\begin{aligned} b_{ij}\mathbf{f}_i &= \alpha(\mathbf{f}_j) = \alpha(p_{rj}\mathbf{e}_r) \\ &= p_{rj}\alpha\mathbf{e}_r = p_{rj}a_{sr}\mathbf{e}_s \\ &= p_{rj}a_{sr}q_{is}\mathbf{f}_i = q_{is}a_{sr}p_{rj}\mathbf{f}_i. \end{aligned}$$

Thus $b_{ij} = q_{is}a_{sr}p_{rj}$.

Exercise 6.1.7

(i) Theorem 6.1.5 shows that, if A and B represent the same map with respect to two bases, then A and B are similar.

If A and B are similar then $B = P^{-1}AP$. Let α be the linear map with matrix A corresponding to a basis

$$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n.$$

Let $\mathbf{f}_j = \sum_{i=1}^{n} p_{ij}\mathbf{e}_i$. Since P has rank n, the $\mathbf{f}_j$ form a basis and Theorem 6.1.5 tells us that α has matrix B with respect to this basis.

(ii) Follows at once from Theorem 6.1.5 and part (i).

(iii) Write $A \sim B$ if $B = P^{-1}AP$ for some non-singular P.

$I^{-1}AI = IAI = A$ so $A \sim A$.

If $A \sim B$ then $B = P^{-1}AP$ for some non-singular P and writing $Q = P^{-1}$ we have Q non-singular and $A = Q^{-1}BQ$ so $B \sim A$.

If $A \sim B$, $B \sim C$ then $B = P^{-1}AP$, $C = Q^{-1}CQ$ for some non-singular P, Q. Then PQ is non-singular and $C = (PQ)^{-1}A(PQ)$ so $C \sim A$.

(iv) If A represents α with respect to one basis then A represents α with respect to the same basis. Thus $A \sim A$.

If $A \sim B$ then there exist two bases with respect to which A and B represent the same map so $B \sim A$.

By (i), if there exist two bases with respect to which A and B represent the same map then given any basis $\mathcal{E}_1$ there exists a basis $\mathcal{E}_2$ so that A and B represent the same map with respect to $\mathcal{E}_1$ and $\mathcal{E}_2$. Thus if $A \sim B$, $B \sim C$ we can find bases $\mathcal{E}_j$ such that A and B represent the same map with respect to $\mathcal{E}_1$ and $\mathcal{E}_2$, and b and C represent the same map with respect to $\mathcal{E}_2$ and $\mathcal{E}_3$, so A and C represent the same map with respect to $\mathcal{E}_1$ and $\mathcal{E}_3$ so $A \sim B$.

Exercise 6.1.8

We have

$$\mathbf{f}_j = \sum_{i=1}^{n} p_{ij}\mathbf{e}_j = (p_{i1}, p_{i2}, \ldots, p_{in})^T.$$

Exercise 6.1.9

Observe that $\mathbf{e}_1 = \mathbf{f}_1$, $\mathbf{e}_2 = \mathbf{f}_2 - \mathbf{f}_1$, $\mathbf{e}_3 = \mathbf{f}_3 - \mathbf{f}_2$,

$$\alpha(\mathbf{f}_1) = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \mathbf{e}_1 - \mathbf{e}_2 = 2\mathbf{f}_1 - \mathbf{f}_2,$$

$$\begin{aligned} \alpha(\mathbf{f}_2) &= \begin{pmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \\ &= 2\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 = \mathbf{f}_1 + \mathbf{f}_3, \end{aligned}$$

$$\begin{aligned} \alpha(\mathbf{f}_3) &= \begin{pmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 4 \end{pmatrix} \\ &= 4\mathbf{e}_1 + 2\mathbf{e}_2 + 4\mathbf{e}_3 = 2\mathbf{f}_1 - 2\mathbf{f}_2 + 4\mathbf{f}_3. \end{aligned}$$

The new matrix is

$$B = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 0 & 1 \\ 2 & -2 & 4 \end{pmatrix}.$$

Exercise 6.1.12

Without Theorem 6.1.10, our definition would depend on the choice of basis.

Exercise 6.2.3

(i) Observe that

$$\begin{aligned}\det\left(t\begin{pmatrix}1&0\\0&1\end{pmatrix}-\begin{pmatrix}a&b\\c&d\end{pmatrix}\right)&=\det\begin{pmatrix}t-a&-b\\-c&t-d\end{pmatrix}\\&=(t-a)(t-d)-(b)(-c)\\&=t^2-(a+d)+(ad-bc)\\&=t^2-(\operatorname{Tr}A)t+\det A.\end{aligned}$$

(ii) We have

$$\det(tI-A)=\sum_{\sigma\in S_3}\zeta(\sigma)\prod_{i=1}^{3}(t\delta_{i\sigma i}-a_{i\sigma i})=ut^3+vt^2+ct+w.$$

Looking at the coefficient of t^3, we have $u=1$. Looking at the coefficient of t^2, we have $v=-a_{11}-a_{22}-a_{33}$. Setting $t=0$

$$w=\det(-A)=(-1)^3\det A=-\det A.$$

Exercise 6.2.4

Write ∂P for the degree of a polynomial P.

Observe that $\partial b_{i\sigma(i)} = 0$ if $\sigma(i) \neq i$ and $\partial b_{i\sigma(i)} = 1$ if $\sigma(i) = i$.

Thus if $\sigma \neq \iota$ we know that $\sigma(i) \neq i$ for at least two distinct values of i so

$$\partial \prod_{i=1}^{n} b_{i\sigma(i)} \leq \sum_{i=1}^{n} \partial b_{i\sigma(i)} \leq n-2.$$

It follows that

$$\det(tI - A) = \sum_{\sigma \in S_n} \zeta(\sigma) b_{i\sigma(i)}(t) = \prod_{i=1}^{n} b_{ii}(t) \sum_{\sigma \neq \iota} \zeta(\sigma) b_{i\sigma(i)}(t) = \prod_{i=1}^{n} (t - a_{ii}) + Q(t),$$

where Q is a polynomial of degree at most $n-2$.

Thus the coefficient of t^{n-1} in $\det(tI - A)$ is the coefficient of t^{n-1} in $\prod_{i=1}^{n}(t - a_{ii})$, that is to say, $-\operatorname{Tr} A$.

The constant term of a polynomial P is $P(0)$ so the constant term in $\det(tI - A)$ is $\det((-1)A) = (-1)^n \det A$.

Exercise 6.2.9

We show that any linear map $\alpha : \mathbb{R}^{2n+1} \to \mathbb{R}^3$ has an eigenvector. It follows that there exists some line l through $\mathbf{0}$ with $\alpha(l) \subseteq l$.

Proof. Since $\det(t\iota - \alpha)$ is a real polynomial of odd degree, the equation $\det(t\iota - \alpha) = 0$ has a real root λ say. We know that λ is an eigenvalue and so has an associated eigenvector $\mathbf{u}$, say. Let

$$l = \{s\mathbf{u} \,:\, s \in \mathbb{R}\}.$$

□

Exercise 6.2.11

$$\begin{aligned}\det(t\iota - \rho_\theta) &= \det\begin{pmatrix} t-\cos\theta & \sin\theta \\ -\sin\theta & t-\cos\theta\end{pmatrix}\\ &= (t-\cos\theta)^2 + \sin^2\theta\\ &= t^2 - 2\cos\theta + 1.\end{aligned}$$

The equation $t^2 - 2\cos\theta + 1 = 0$ has a real root if and only if

$$4\cos^2\theta \geq 4,$$

so if and only if $\cos\theta = \pm 1$, i.e. if and only if $\theta \equiv 0 \mod \pi$.

If $\theta \equiv 0 \mod 2\pi$, then $\rho_\theta = \iota$ and every non-zero vector is an eigenvector with eigenvalue 1. If $\theta \equiv \pi \mod 2\pi$, then $\rho_\theta = -\iota$ and every non-zero vector is an eigenvector with eigenvalue -1.

Exercise 6.2.14

(i) $P_A(t) = \det(tI + A) = (-1)^n \det(-tI - A) = (-1)^n P_A(-t)$.

(ii) Let δ be as in Lemma 6.2.13. Let $t_n = \delta/2n$.

Exercise 6.2.15

$$\begin{aligned}\chi_{AB}(t) &= \det(tI - AB) = \det(A(tA^{-1} - B))\\ &= \det A \det(tA^{-1} - B) = \det(tA^{-1} - B)\det A\\ &= \det((tA^{-1} - B)A) = \det(tI - BA) = \chi_{BA}(t).\end{aligned}$$

In general, we can find $s_n \to 0$ such that $s_n I + A$ is invertible. Thus if t is fixed

$$\chi_{(s_n I + A)B}(t) = \chi_{B(s_n - A)}(t).$$

But $\chi_{(s_n I+A)B}(t) \to \chi_{AB}(t)$ and $\chi_{B(s_n I+A)}(t) \to \chi_{BA}(t)$ so

$$\chi_{AB}(t) = \chi_{BA}(t).$$

Allowing t to vary, we see that $\chi_{AB} = \chi_{BA}$.

Observe that $-\operatorname{Tr} C$ is the coefficient of t^{n-1} in $\chi_C(t)$. Thus $\chi_A = \chi_B \Rightarrow \operatorname{Tr} A = \operatorname{Tr} B$. By the previous result this gives $\operatorname{Tr} AB = \operatorname{Tr} BA$.

(i) (False in general.) We have

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},\ C = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

then

$$ABC = (AB)C = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

and

$$ACB = (AC)B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

Thus $\operatorname{Tr}(ABC) = 1 \neq 0 = \operatorname{Tr}(ACB)$ and $\chi_{ABC} \neq \chi_{ACB}$.

(ii) (True.) $\chi_{ABC} = \chi_{A(BC)} = \chi_{(BC)A} = \chi_{BCA}$.

Exercise 6.2.16

Let $\mathbf{e} = (1, 1, 1, \ldots, 1)^T$. Observe that A is magic with constant κ if and only if

$$A\mathbf{e} = A^T\mathbf{e} = \kappa\mathbf{e}.$$

In particular $\mathbf{e}$ is an eigenvector of A.

If A is magic with constant κ and $AB = BA = I$, then

$$\mathbf{e} = BA\mathbf{e} = \kappa B\mathbf{e},$$

so $\kappa \neq 0$ and $B\mathbf{e} = \kappa^{-1}\mathbf{e}$. Since $B^T A^T = (AB)^T = I^T = I$, the argument of the previous sentence gives $B^T\mathbf{e} = \kappa^{-1}\mathbf{e}$. Thus B is magic with constant κ^{-1}.

Suppose A is magic with constant κ. If A is invertible $\operatorname{Adj} A = (\det A)A^{-1}$ so, by the previous paragraph with $B = A^{-1}$, $\operatorname{Adj} A$ is magic with constant $\kappa^{-1}\det A$.

Chose $t_n \to 0$ with $A + t_n I$ invertible. If we write $\operatorname{Adj}(A + t_n I) = (c_{ij}(n))$, $\operatorname{Adj} A = (c_{ij})$ then $c_{ij}(n) \to c_{ij}$ as $n \to 0$. Thus

$$0 = \sum_{r=1}^{n} c_{ir}(n) - \sum_{r=1}^{n} c_{kr}(n) \to \sum_{r=1}^{n} c_{ir} - \sum_{r=1}^{n} c_{kr}$$

and $\sum_{r=1}^{n} c_{ir} - \sum_{r=1}^{n} c_{kr}$ for all i and k. Similarly $\sum_{r=1}^{n} c_{ir} = \sum_{r=1}^{n} c_{rk}$. Thus $\operatorname{Adj} A$ is magic.

Exercise 6.3.2

(i) If
$$x_1\mathbf{e}_1 + x_2\mathbf{e}_2 = \mathbf{0},$$
then
$$\mathbf{0} = \alpha(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = x_1\alpha\mathbf{e}_1 + x_2\alpha\mathbf{e}_2 = \lambda_1 x_1\mathbf{e}_1 + \lambda_2 x_2\mathbf{e}_2$$
Thus
$$\begin{aligned}\mathbf{0} &= \lambda_1(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) - (\lambda_1 x_1\mathbf{e}_1 + \lambda_2 x_2\mathbf{e}_2)\\ &= (\lambda_1 - \lambda_2)x_2\mathbf{e}_2,\end{aligned}$$
so that
$$(\lambda_1 - \lambda_2)x_2 = 0$$
and $x_2 = 0$. Similarly $x_1 = 0$, so $\mathbf{e}_1$ and $\mathbf{e}_2$ are linearly independent.

(ii) Observe that
$$\mathbf{0} = (\alpha - \lambda_2\iota)(\mathbf{0}) = (\alpha - \lambda_2\iota)(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = (\lambda_1 - \lambda_2)x_2\mathbf{e}_2$$
and proceed as before.

(iii) Observe that
$$\begin{aligned}\mathbf{0} &= (\alpha - \lambda_2\iota)(\alpha - \lambda_3\iota)\mathbf{0}\\ &= (\alpha - \lambda_2\iota)(\alpha - \lambda_3\iota)(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3)\\ &= (\alpha - \lambda_2\iota)\big((\lambda_1 - \lambda_3)x_1\mathbf{e}_1 + (\lambda_2 - \lambda_3)x_1\mathbf{e}_2\big)\\ &= (\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)x_1\mathbf{e}_1\end{aligned}$$
so that
$$(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)x_1 = 0$$
and $x_1 = 0$. Similarly $x_2 = x_3 = 0$ so $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are linearly independent.

EXERCISE 6.4.2

(i) Observe that

$$\det(t\iota - \beta) = \det \begin{pmatrix} t & -1 \\ 0 & t \end{pmatrix} = t^2.$$

Thus the characteristic equation of β only has 0 as a root and β only has zero as an eigenvalue.

(ii) If $\mathbf{x} = (x, y)^T$ then

$$\beta\mathbf{x} = 0\mathbf{x} \Leftrightarrow (y, 0)^T = (0, 0)^T \Leftrightarrow y = 0$$

Thus the eigenvectors of β are the non-zero vectors of the form $(x, 0)^T$ and these do not span $\mathbb{R}^2$.

Exercise 6.4.5

(i) We have

$$Q(t) = \det \begin{pmatrix} t-\lambda & 0 \\ 0 & t-\mu \end{pmatrix} = (t-\lambda)(t-\mu) = t^2 - (\lambda+\mu)t + \lambda\mu,$$

so

$$Q(A) = \begin{pmatrix} \lambda^2 & 0 \\ 0 & \mu^2 \end{pmatrix} - \begin{pmatrix} (\lambda+\mu)\lambda & 0 \\ 0 & (\lambda+\mu)\mu \end{pmatrix} + \begin{pmatrix} \lambda\mu & 0 \\ 0 & \lambda\mu \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus $Q(\alpha) = 0$.

(ii) We have

$$Q(t) = \det \begin{pmatrix} t-\lambda & -1 \\ 0 & t-\lambda \end{pmatrix} = (t-\lambda)^2 = t^2 - 2\lambda t + \lambda^2,$$

so

$$Q(A) = \begin{pmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{pmatrix} - \begin{pmatrix} 2\lambda^2 & 0 \\ 0 & 2\lambda^2 \end{pmatrix} + \begin{pmatrix} \lambda^2 & 0 \\ 0 & \lambda^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Thus $q(\alpha) = 0$.

(iii) We know that there exists a basis with respect to which α has one of the two forms discussed, so the result of Example 6.4.4 is correct.

Exercise 6.4.7

Let P be as in Theorem 6.4.6. Since P is non-singular, $\det P \neq 0$, so (working in $\mathbb{C}$) we can find a κ with $\kappa^2 = \det P$. Set $\nu = \kappa^{-1}$ and $Q = \nu P$.

Then $\det Q = \nu^2 \det P = 1$ and

$$Q^{-1}AQ = \nu^{-1}\nu P^{-1}AP = P^{-1}AP.$$

Exercise 6.4.8

(i) Since

$$\frac{d}{dt}(e^{-\lambda t}x(t)) = -\lambda e^{-\lambda t}x(t) + e^{-\lambda t}\dot{x}(t) = e^{-\lambda t}(-\lambda x(t) + \dot{x}(t)) = 0,$$

the mean value theorem tells us that $e^{-\lambda t}x(t) = C$ and so $x(t) = Ce^{-\lambda t}$ for some constant C.

(ii) Since

$$\frac{d}{dt}\big(e^{-\lambda t}x(t) - Kt\big) = e^{-\lambda t}(-\lambda x(t) + \dot{x}(t) - Ke^{\lambda t}) = 0,$$

the mean value theorem tells us that $e^{-\lambda t}x(t) - Kt = C$ and so $x(t) = (C + Kt)e^{-\lambda t}$ for some constant C.

Exercise 6.4.9

If $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$ then $\dot{x}_1(t) = x_2(t)$ automatically and $\dot{x}_2(t) = -bx_1(t) - ax_2(t) \Leftrightarrow \ddot{x} = -a\dot{x} - bx \Leftrightarrow \ddot{x} + a\dot{x} + bx = 0$

$$\det(\lambda I - A) = \lambda(\lambda + a) + b = \lambda^2 + a\lambda + b$$

EXERCISE 6.5.5

By definition,

$$\begin{aligned}
Q^{-1}R_\theta Q &= \frac{1}{2}\begin{pmatrix}1 & i\\ i & 1\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1 & -i\\ -i & 1\end{pmatrix}\\
&= \frac{1}{2}\begin{pmatrix}\cos\theta + i\sin\theta & -\sin\theta + i\cos\theta\\ \sin\theta + i\cos\theta & \cos\theta - i\sin\theta\end{pmatrix}\begin{pmatrix}1 & -i\\ -i & 1\end{pmatrix}\\
&= \frac{1}{2}\begin{pmatrix}e^{i\theta} & ie^{i\theta}\\ ie^{-i\theta} & e^{-i\theta}\end{pmatrix}\begin{pmatrix}1 & -i\\ -i & 1\end{pmatrix}\\
&= \frac{1}{2}\begin{pmatrix}2e^{i\theta} & 0\\ 0 & 2e^{-i\theta}\end{pmatrix} = \begin{pmatrix}e^{i\theta} & 0\\ 0 & e^{-i\theta}\end{pmatrix}.
\end{aligned}$$

Exercise 6.5.2

Since A has distinct eigenvalues, it is diagonalisable and we can find a an invertible matrix B such that

$$BAB^{-1} = D,$$

where D is the diagonal matrix with jth diagonal entry λ_j. If we set $\mathbf{y} = B\mathbf{x}$ (using column vectors) we obtain

$$\dot{\mathbf{y}} = B\dot{\mathbf{x}} = BA\mathbf{x} = BAB^{-1}\mathbf{y} = D\mathbf{y}.$$

Thus $\dot{y}_j = \lambda_j y_j$ and $y_j = c_j e^{\lambda_j t}$ for $[1 \leq j \leq n]$ and so, writing $F = B^{-1}$, $F = (f_{ij})$

$$x_1 = \sum_{j=1}^{n} f_{1j} y_j = \sum_{j=1}^{n} f_{1j} c_j \exp(i\lambda_j t).$$

If we set $\mu_j = f_{1j} c_j$, we obtain the required result.

If $A = D$, then $\mathbf{x} = \mathbf{y}$ and $x_1 = c_1 \exp(i\lambda_1 t)$ and, with the notation of the question, we have $\mu_j = 0$ for $2 \leq j \leq n$.

Exercise 6.6.2

D^q is the diagonal matrix with diagonal entries the qth powers of the diagonal entries of D.

(ii) A and D represent the same linear map α with respect to two bases $\mathcal{E}_1$ and $\mathcal{E}_2$ with basis change rule $U = PVP^{-1}$. A^q and D^q represent the same linear map α^q with respect to the two bases $\mathcal{E}_1$ and $\mathcal{E}_2$ so $A^q = PD^qP^{-1}$.

Exercise 6.6.7

(i) We have
$$\lambda^{-q}\alpha^q(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = x_1\mathbf{e}_1 + x_2(\mu/\lambda)^n x_2\mathbf{e}_2 \to x_1\mathbf{e}_1$$
coordinate wise.

(i)′ We have
$$\lambda^{-q}\alpha^q(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = x_1\mathbf{e}_1 + x_2(\mu/\lambda)^n x_2\mathbf{e}_2,$$
but $(\mu/\lambda)^n$ does not tend to a limit, so $\lambda^{-q}\alpha^q(x_1\mathbf{e}_1 + x_2\mathbf{e}_2)$ converges coordinatewise if and only if $x_2 = 0$.

(ii)′ Nothing to prove.

(iii) We have
$$\begin{aligned}\alpha\big((\lambda^q x_1 + q\lambda^{q-1}x_2)\mathbf{e}_1 + \lambda^q x_2\mathbf{e}_2\big) &= \lambda(\lambda^q x_1 + q\lambda^{q-1}x_2)\mathbf{e}_1 + \lambda^q(x_2\mathbf{e}_1 + \lambda\mathbf{e}_2)\\ &= \big(\lambda^{q+1}x_1 + (q+1)\lambda^q x_2\big)\mathbf{e}_1 + \lambda^{q+1}x_2\mathbf{e}_2\end{aligned}$$
so, by induction
$$\alpha^q(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = (\lambda^q x_1 + q\lambda^{q-1}x_2)\mathbf{e}_1 + \lambda^q x_2\mathbf{e}_2$$
for all $q \geq 0$.

Thus
$$q^{-1}\lambda^{-q}\alpha^q(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) = (q^{-1}x_1 + \lambda^{-1}x_2)\mathbf{e}_1 + q^{-1}x_2\mathbf{e}_2 \to \lambda^{-1}x_2)\mathbf{e}_1$$
coordinatewise.

(iv) Immediate.

Exercise 6.6.8

$F_1 = F_2 = 1$, $F_3 = 2$, $F_4 = 3$, $F_5 = 5$, $F_6 = 8$, $F_7 = 13$, $F_8 = 21$, $F_9 = 34$, $F_{10} = 55$.

$$F_0 + 1 = F_0 + F_1 = F_2 = 1,$$

so $F_0 = 0$.

$$F_{-1} + 0 = F_{-1} + F_0 = F_1 = 1,$$

so $F_{-1} = 1$.

Write $G_{-n} = (-1)^{n+1}F_n$. Then

$$G_{-n+1} + G_{-n} = (-1)^n(F_{n-1} - F_n) = (-1)^{n-1}F_{n-2} = G_{-n+2}$$

and $G_{-1} = F_{-1}$, $G(0) = F(0)$ so, by induction, $F_{-n} = G_{-n}$ for all $n \geq 0$. Thus $F_{-n} = (-1)^{n+1}F_n$ for all n.

Exercise 6.6.9

If $b = 0$, but $a \neq 0$, then $u_n = (-a)^n u_0$.

If $a = b = 0$, then $u_n = 0$ for all n.

Exercise 6.6.10

We look at vectors

$$\mathbf{u}(n) = \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix}.$$

We then have

$$\mathbf{u}(n+1) = A\mathbf{u}(n), \text{ where } A = \begin{pmatrix} 0 & 1 \\ -a & -b \end{pmatrix}$$

and so

$$\begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} = A^n \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}.$$

Now A is not a multiple of the identity matrix so (since the roots are equal) there is a basis $\mathbf{e}_1$ $\mathbf{e}_2$ with $A\mathbf{e}_1 = \lambda\mathbf{e}_1 = \lambda\mathbf{e}_1$ and $A\mathbf{e}_2 = \lambda\mathbf{e}_1 + \lambda\mathbf{e}_2$.

Now

$$A^n(x_1\mathbf{e}_1 + x_2\mathbf{e}_1) = (\lambda^n x_1 + n\lambda^{n-1}x_2) + \lambda^n x_2\mathbf{e}_2$$

so, since

$$\mathbf{u}(n) = A^n\mathbf{u}(0),$$

we have

$$u_n = (c + c'n)\lambda^n$$

for some constants c and c' depending on u_0 and u_1.

Exercise 6.6.11

We look at vectors

$$\mathbf{u}(n) = (u_n, u_{n+1}, \dots, u_{n+m-1})^T.$$

We then have

$$\mathbf{u}(n+1) = A\mathbf{u}(n), \text{ where } A = \begin{pmatrix} 0 & 1 & 0 \dots & 0 & \\ 0 & 0 & 1 \dots & 0 & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 \dots & 1 & \\ -a_{m-1} & -a_{m-2} & -a_{m-3} & \dots & -a_0 \end{pmatrix}$$

The characteristic polynomial of A is $P(t)$, so, since all the roots of P are distinct and non-zero, the eigenvectors $\mathbf{e}_j$ (corresponding to an eigenvalue λ_j) of A form a basis.

Setting $\mathbf{u}_0 = \sum_{k=1}^m b_k \mathbf{e}_k$ we get

$$\mathbf{u}_n = A^n \mathbf{u}_0 = \sum_{k=1}^m b_k \lambda_k \mathbf{e}_k$$

and so looking at the first entry in the vectors

$$u_n = \sum_{k=1}^m c_k \lambda_k^n$$

for some constants c_j.

Conversely, if

$$u_n = \sum_{k=1}^m c_k \lambda_k^n,$$

then

$$u_r + \sum_{j=0}^{m-1} a_j u_{j-m+r} = \sum_{k=1}^m \lambda_k^{n-m} P(\lambda_k) = 0.$$

Observe that, if $|\lambda_1| > |\lambda_k|$ for $2 \leq k \leq m$, then

$$\lambda_1^{-n} u_n = \sum_{k=1}^m c_k (\lambda_1^{-1} \lambda_k)^n \to c_1$$

as $n \to \infty$. Thus, if $c_0 \neq 0$, we have $u_{r-1} \neq 0$ for r large and

$$\frac{u_r}{u_{r-1}} \to \lambda_1$$

as $r \to \infty$.

A very similar proof shows that, if $|\lambda_k| > |\lambda_m|$ for $1 \leq k \leq m-1$ and $c_m \neq 0$, then $u_n \neq 0$ for $-n$ large and

$$\frac{u_{r-1}}{u_r} \to \lambda_m.$$

Exercise 6.6.12

(i) Observe that
$$\frac{1+\sqrt{5}}{2} \times \frac{-1+\sqrt{5}}{2} = \frac{(\sqrt{5})^2 - 1^2}{4} = 1.$$

(ii) We have
$$\det(tI - A) = t(t-1) - 1 = t^2 - t - 1 = (t-\tau)(t+\tau^{-1}),$$
so the eigenvalues are τ and $-\tau^{-1}$.

If $(x, y)^T$ is an eigenvector with eigenvalue τ,
$$\begin{aligned} y &= \tau x \\ x + y &= \tau y, \end{aligned}$$
so $\mathbf{e}_1 = (1, \tau)^T$ is an eigenvector with eigenvalue τ.

A similar argument shows that $\mathbf{e}_2 = (-1, \tau^{-1})^T$ is an eigenvector with eigenvalue τ^{-1}.

By inspection (or solving two simultaneous linear equations)
$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{5}}(\mathbf{e}_1 + \mathbf{e}_2).$$

(iii) It follows that
$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = A^n \begin{pmatrix} F_0 \\ F_1 \end{pmatrix} = 5^{-1/2}(\tau^n \mathbf{e}_1 + (-\tau)^{-n}\mathbf{e}_2)$$
and so
$$F_n = 5^{-1/2}(\tau^n + (-\tau)^{-n})$$

(vi) Since F_n is an integer and
$$|5^{-1/2}(-\tau)^{-n}| \leq 5^{-1/2} < 1/2$$
for $n \geq 0$ we know that F_n is the closest integer to $\tau^n/5^{1/2}$.

Both methods give $F_{20} = 6765$.

(v) Observe that A^n has first column
$$A^{n-1}(F_0, F_1)^T = (F_{n-1}, F_n)^T$$
and second column
$$A^{n-1}(F_1, F_2)^T = (F_n, F_{n+1})^T.$$
(Or use induction.)

$$(-1)^n = (\det A)^n = \det A^n = F_{n-1}F_{n+1} - F_n^2.$$

(vi)True for all n because $F_n = (-1)^{n+1}F_{-n}$.

(vi) We have $A^{2n} = A^n A^n$, so

$$\begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix} \begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix} = \begin{pmatrix} F_{2n-1} & F_{2n} \\ F_{2n} & F_{2n+1} \end{pmatrix}$$

so, evaluating the upper left hand term of the matrix, we get

$$F_n^2 + F_{n-1}^2 = F_{2n-1}.$$

Exercise 6.6.13

(i) Let $e_{ij}^{(n)}$ be the number of routes from i to j involving exactly n flights. We claim that $e_{ij}^{(n)} = d_{ij}^{(n)}$. We use induction on n. The result is clearly true if $n = 1$. Suppose it is true for $n = r$.

$$\begin{aligned}
e_{ij}^{(r+1)} &= \text{number of journeys with exactly } r+1 \text{ flights from } i \text{ to } j \\
&= \sum_{k=1}^{m} \text{number of journeys with exactly } r+1 \text{ flights} \\
&\qquad\quad \text{from } i \text{ to } j \text{ ending with a flight from } k \text{ to } j \\
&= \sum_{d_{kj} \neq 0} \text{number of journeys with exactly } r \text{ flights from } i \text{ to } k \\
&= \sum_{k=1}^{m} d_{ik}^{(r)} d_{kj} = d_{ij}^{(r+1)}
\end{aligned}$$

The result follows.

In particular, there is a journey from i to j which involve exactly n flights if and only if $d_{ij}^{(n)} > 0$.

(ii) We can produce an argument like (i) if we make the convention that a flight includes just staying at an airport (so a flight from i to i is allowed).

In particular there is a journey from i to j which involve n flights or less (between different airports) if and only if $\tilde{d}_{ij}^{(n)} > 0$.

(iii) We write down the safe states for example

$$i = (G|PWC)$$

goat on one side, peasant, wolf and cabbage on other and write $d_{ij} = 1$ if the peasant can change the situation from i to j in one crossing. Otherwise $d_{ij} = 0$. The least number of crossings is the least N for which

$$d^N_{(GPWC|\emptyset),(\emptyset|GPWC)} \neq 0.$$

Observe that, since the peasant will not repeat a state, if there are M safe states the problem is soluble, if at all, with at most $M - 1$ crossings so the method will reveal if the problem is insoluble.

(The cabbage survived but was so distressed by its narrow escape that it turned over a new leaf).

EXERCISE 6.7.1

(i) Suppose that $A = (a_{ij})_{1 \le i \le n}^{1 \le i \le n}$ is lower triangular. Then, by row expansion,

$$\det A = a_{11} \det B$$

where $B = (a_{i+1,j+1})_{1 \le i \le n-1}^{1 \le i \le n-1}$. Since B is lower triangular with diagonal entries $a_{i+1,i+1}$ $[1 \le i \le n-1]$ the required result follows by induction.

(ii) We have

$$L \text{ invertible} \Leftrightarrow \det L \neq 0 \Leftrightarrow \prod_{j=1}^{n} l_{jj} \neq 0 \Leftrightarrow l_{jj} \neq 0 \ \forall j$$

(iii) Since $tI - L$ is lower triangular

$$\det(tI - L) = \prod_{j=1}^{n} (t - l_{jj})$$

so the roots of the characteristic equation are the diagonal entries l_{jj} (multiple roots occurring with the correct multiplicity).

Observe that $(0, 0, 0, \dots, 0, 1)^T$ is an eigenvector with eigenvalue l_{nn}.

Exercise 6.7.2

Just repeat the earlier discussion mutatis mutandis.

The following is merely a variation.

If the system is

$$\sum_{j=1}^{n} u_{ij}x_j = y_i$$

then setting $x'_j = x_{n-j}$, $y'_i = y_{n-i}$ $l_{ij} = u_{n-i,n-j}$ gives the lower triangular system

$$\sum_{j=1}^{n} l_{ij}x'_j = y'_i.$$

Exercise 6.7.3

(i) Observe first that the product of two lower triangular matrices is triangular since if a_{ij}, $b_{ij} = 0$ for $i + 1 \geq j$, then

$$\sum_{r=1}^{n} a_{ir} b_{rj} = \sum_{j \leq r \leq i} a_{ir} b_{rj} = 0$$

for $i + 1 \geq j$.

We use induction to show that the inverse of an invertible $n \times n$ lower triangular matrix is triangular.

Since every 1×1 matrix is lower triangular, the result is certainly true for $n = 1$. Suppose it is true for $n = m$ and A is an $m+1 \times m+1$ invertible lower triangular matrix. Since A is invertible $a_{11} \neq 0$. If $B = (b_{ij})$ is the $m + 1 \times m + 1$ matrix with $b_{11} = a_{11}^{-1}$, $b_{i1} = -a_{i1} a_{11}^{-1}$ $b_{ii} = 1$ for $2 \leq i \leq m + 1$ and $b_{ij} = 0$, otherwise, then

$$BA = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & L \end{pmatrix}$$

where L is an $m \times m$ invertible lower triangular matrix and $\mathbf{0} \in \mathbb{R}^m$ is the zero column vector.

If we now set

$$C = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & L^{-1} \end{pmatrix}$$

then since L^{-1} is lower triangular (by the inductive hypothesis), we know that C^{-1} is lower triangular so $C^{-1}B$ is lower triangular. But

$$(C^{-1}B)A = C^{-1}(BA) = C^{-1}C = I$$

so A has lower triangular inverse and the induction is complete.

(ii) By the first paragraph of (i),

$$A, B \in \mathcal{L} \Rightarrow AB \in \mathcal{L}.$$

By inspection $I \in \mathcal{L}$ and, by part (i),

$$A \in \mathcal{L} \Rightarrow A^{-1} \in \mathcal{L}.$$

EXERCISE 6.7.6

Observe that

$$\begin{aligned} LU &= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2-1 & 2-2 \\ -2 & -1+3 & -1+6-4 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix}. \end{aligned}$$

Exercise 6.7.7

Consider an $n \times n$ matrix A. The first step will be impossible only if the entire first row consists of zeros in which case it is clear that the $n \times n$ matrix A is not invertible. If the first step can be done, then we repeat the process with a $n-1 \times n-1$ square matrix B and the same remarks apply to B.

Thus either the process will stop because the matrix with which we deal at the given stage has top row all zeros or the process will not stop in which case we obtain an upper triangular matrix U and a lower triangular matrix L both with all diagonal entries non-zero so invertible. Since the product of invertible matrices is invertible $A = LU$ is invertible.

Thus the process works if and only if A is invertible.

Exercise 6.7.9

Suppose that
$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \begin{pmatrix} u & v \\ w & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$
Then
$$\begin{pmatrix} au & av \\ cu & cv + dw \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$
It follows that $au = 0$ and $av = 1$, so $u = 0$, which is incompatible with $cu = 1$. Our initial formula cannot hold.

Exercise 6.7.10

We use induction on n. The result is trivial for $n = 1$ since

$$(1)(u_1) = (1)(u_2) \Rightarrow (u_1) = (u_2) \Rightarrow u_1 = u_2.$$

Now suppose the result is true for $n = m$ and L_1, L_2 are lower triangular (with diagonal entries 1) and U_1, U_2 upper triangular $m + 1 \times m + 1$ matrices with $L_1U_1 = L_2U_2$. Observe that we can write

$$L_j = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{l}_j & \tilde{L}_j \end{pmatrix} \text{ and } L_j = \begin{pmatrix} u(j) & \mathbf{u}_j^T \\ \mathbf{0} & \tilde{U}_j \end{pmatrix}$$

where the $\tilde{L}_j$, L_2 lower triangular (with diagonal entries 1) and the $\tilde{U}_j$ upper triangular $m \times m$ matrices, $u(j) \in \mathbb{F}$ and $\mathbf{l}_j$, $\mathbf{u}_j \in \mathbb{F}^m$ are column vectors. Equating first rows in the equation $L_1U_1 = L_2U_2$ we get

$$(u(1), \mathbf{u}_1^T) = (u(2), \mathbf{u}_2^T)$$

and equating first columns we get

$$(1, \mathbf{l}_1^T)^T = (1, \mathbf{l}_2^T)^T$$

so $u(1) = u(2)$, $\mathbf{u}_1 = \mathbf{u}_2$, $\mathbf{l}_1 = \mathbf{l}_2$.

The condition $L_1U_1 = L_2U_2$ now gives

$$\tilde{L}_1\tilde{U}_1 = \tilde{L}_2\tilde{U}_2$$

so by the inductive hypothesis

$$\tilde{L}_1 = \tilde{L}_2,\ \tilde{U}_1 = \tilde{U}_2.$$

We have shown that

$$L_1 = L_2,\ U_1 = U_2$$

and this completes the induction

Exercise 6.7.11

Since $\det LU = \det L \det U$ and $\det L$, $\det U$ are the product of the diagonal entries, $\det LU$ is cheap to calculate.

By solving the n sets of simultaneous linear equations

$$\sum_{r=1}^{n} l_{ir}x_{rj} = \delta_{ij}$$

each of which requires of the order of n^2 operations we can compute $X = L^{-1}$ in the order of n^3 operations and we can compute U^{-1} similarly. We now compute $(LU)^{-1} = U^{-1}L^{-1}$ again in the order of n^3 operations.

Exercise 6.7.12

(i) It is probably most instructive to do this ab initio, but quicker to use known results.

We know that (after reordering columns) $A = L\tilde{U}$ with L lower triangular and all diagonal entries 1 and $\tilde{U}$ upper triangular. Since A is non-singular $\tilde{U}$ is, so all its diagonal entries are non-zero. Let D be the diagonal matrix with diagonal entries the diagonal entries of $\tilde{U}$. Set $U = D^{-1}\tilde{U}$. Then U is upper triangular with all diagonal entries 1, D is non-singular and

$$A = LDU.$$

Suppose that L_j is a lower triangular $n \times n$ matrix with all diagonal entries 1, U_j is an upper triangular $n \times n$ matrix with all diagonal entries 1 and D_j is a non-singular diagonal $n \times n$ matrix $[j = 1, 2]$. We claim that, if

$$L_1 D_1 U_1 = L_2 D_2 U_2,$$

then $L_1 = L_2$, $U_1 = U_2$ and $D_1 = D_2$.

To see this observe that $D_j U_j$ is non singular upper triangular so, by Exercise 6.7.10, $L_1 = L_2$ and

$$D_1 U_1 = D_2 U_2.$$

By looking at the diagonal entries of both sides of the equation we get $D_1 = D_2$ so

$$U_1 = D_1^{-1}(D_1 U_1) = D_2^{-1}(D_2 U_2) = U_2.$$

(ii) Yes. Either repeat the argument mutatis mutandis or argue as follows.

If $G = (g_{ij})$ is an $n \times n$ matrix, write $\tilde{G} = (g_{n+1-i,n+1-j})$. If $\tilde{G} = LU$ with L lower triangular, U upper triangular, then $G = \tilde{L}\tilde{U}$ and $\tilde{L}$ is upper triangular whilst $\tilde{U}$ is lower triangular.

(iii) No. Observe that

$$\begin{pmatrix} a & b \\ c & 0 \end{pmatrix} \begin{pmatrix} x & 0 \\ y & z \end{pmatrix} = \begin{pmatrix} ax + by & bz \\ cx & 0 \end{pmatrix}.$$

Exercise 7.1.4

(i) Any set of orthonormal vectors is linearly independent. Any set of k linearly independent vectors forms a basis for U.

(ii) By (i), we can find λ_j such that

$$\mathbf{x} = \sum_{j=1}^{k} \lambda_j \mathbf{e}_j.$$

Now observe that

$$\langle \mathbf{x}, \mathbf{e}_i \rangle = \sum_{j=1}^{k} \lambda_j \langle \mathbf{e}_j, \mathbf{e}_i \rangle = \lambda_i.$$

Exercise 7.1.6

If we take $\mathbf{e}_1 = 3^{-1/2}(1,1,1)$, $\mathbf{e}_1 = 2^{-1/2}(1,-1,0)$, then, by inspection, they form an orthonormal system.

We now use Gramm-Schmidt for $\mathbf{x} = (1,0,0)$.

$$\begin{aligned}\mathbf{x} - \langle \mathbf{x}, \mathbf{e}_1\rangle \mathbf{e}_1 - \langle \mathbf{x}, \mathbf{e}_2\rangle \mathbf{e}_2 &= (1,0,0) - 3^{-1}(1,1,1) - 2^{-1}(1,-1,0)\\ &= (1/6, 1/6, -1/3) = \mathbf{a},\end{aligned}$$

say, Setting

$$\mathbf{e}_3 = \|\mathbf{a}\|^{-1}\mathbf{a} = 6^{-1/2}(1,1,-2)$$

we have $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ orthonormal.

It follows at one that $\mathbf{e}_1 = 3^{-1/2}(1,1,1), 2^{-1/2}(0,-1,1), 6^{-1/2}(-2,1,1)$ is another orthonormal system.

Exercise 7.1.8

(Dimension 2) If a point B does not lie in a plane π, then there exists a unique line l' perpendicular to π passing through B. The point of intersection of l with π is the closest point in π to B. More briefly, the foot of the perpendicular dropped from B to π is the closest point in π to B.

(Dimension 1) If a point B does not lie on a line l, then there exists a unique line l' perpendicular to l passing through B. The point of intersection of l with l' is the closest point in l to B. More briefly, the foot of the perpendicular dropped from B to l is the closest point in l to B.

Exercise 7.1.9

(i) We have

$$\begin{aligned}\left\|\mathbf{x}-\sum_{j=1}^{k}\lambda_j\mathbf{e}_j\right\|^2 & \left\langle \mathbf{x}-\sum_{j=1}^{k}\lambda_j\mathbf{e}_j, \mathbf{x}-\sum_{j=1}^{k}\lambda_j\mathbf{e}_j\right\rangle \\ &= \|\mathbf{x}\|^2 - 2\sum_{j=1}^{k}\lambda_j\langle\mathbf{x},\mathbf{e}_j\rangle + \sum_{j=1}^{k}\lambda_j^2 \\ &= \|\mathbf{x}\|^2 - \sum_{j=1}^{k}\langle\mathbf{x},\mathbf{e}_j\rangle^2 + \sum_{j=1}^{k}(\lambda_j - \langle\mathbf{x},\mathbf{e}_j\rangle)^2 \\ &\geq \|\mathbf{x}\|^2 - \sum_{j=1}^{k}\langle\mathbf{x},\mathbf{e}_j\rangle^2\end{aligned}$$

with equality if and only if $\lambda_j = \langle\mathbf{x},\mathbf{e}_j\rangle$.

(ii) It follows from (i) that

$$\|\mathbf{x}\|^2 \geq \sum_{j=1}^{k}\langle\mathbf{x},\mathbf{e}_j\rangle^2$$

with equality if and only if $\lambda_j = \langle\mathbf{x},\mathbf{e}_j\rangle$ and

$$\left\|\mathbf{x}-\sum_{j=1}^{k}\lambda_j\mathbf{e}_j\right\| = 0,$$

ie

$$\mathbf{x} = \sum_{j=1}^{k}\lambda_j\mathbf{e}_j$$

and this occurs if and only if

$$\mathbf{x} \in \operatorname{span}\{\mathbf{e}_1,\, \mathbf{e}_2,\, \ldots,\, \mathbf{e}_k\}.$$

Exercise 7.1.10

Observe first that

$$\langle \mathbf{0}, \mathbf{u} \rangle = 0$$

for all $\mathbf{u}$ so $\mathbf{0} \in U^\perp$.

Next observe that if $\lambda_1, \lambda_2 \in \mathbb{R}$

$\mathbf{v}_1,\ \mathbf{v}_2 \in U^\perp$

$\Rightarrow \langle \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \mathbf{u} \rangle = \lambda_1 \langle \mathbf{v}_1, \mathbf{u} \rangle + \lambda_2 \langle \mathbf{v}_2, \mathbf{u} \rangle = 0 + 0 = 0$ for all $\mathbf{u} \in U$

$\Rightarrow \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 \in U^\perp$.

If we take $\mathbf{a} - \mathbf{b} = \mathbf{v}$ in Theorem 7.1.7, we see that $\mathbf{a} \in \mathbf{R}^n$ can be written in one and only one way as $\mathbf{x} = \mathbf{u} + \mathbf{v}$ with $\mathbf{u} \in U$, $\mathbf{v} \in U^\perp$.

If $\mathbf{e}_1, \ldots, \mathbf{e}_k$ is a basis for U and $\mathbf{e}_{k+1}, \ldots, \mathbf{e}_l$ is a basis for $U^\perp$, the previous paragraph tells us that $\mathbf{e}_1, \ldots, \mathbf{e}_k, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_l$ is a basis for $\mathbb{R}^n$ so $l = n$ and

$$\dim U + \dim U^\perp = n.$$

Exercise 7.2.2

$$\begin{aligned}\langle \alpha\mathbf{x}, \mathbf{y}\rangle &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j y_i\right) = \sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}x_j y_i\right)\\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} c_{ji}x_j y_i\right) = \langle \mathbf{x}, \alpha^*\mathbf{y}\rangle.\end{aligned}$$

Exercise 7.2.6

Suppose that A, B represent α, β with respect to some basis. Lemma 7.2.4 enables us to translate Lemma 7.2.5 into matrix form

(i) Since $(\alpha\beta)^* = \beta^*\alpha^*$, we have $(AB)^T = B^T A^T$.

(ii) Since $\alpha^{**} = (\alpha^*)^* = \alpha$, we have $A^{TT} = (A^T)^T + A$.

(iii) Since $(\lambda\alpha+\mu\beta)^* = \lambda\alpha^*+\mu\beta^*$, we have $(\lambda A+\mu B)^T = \lambda A^T+\mu B^T$.

(vi) Since $\iota^* = \iota$, we have $I^T = I$.

In coordinates

(i) If $AB = C$, then
$$c_{ij} = \sum_{r=1}^{n} a_{ir} b_{rj}$$
so
$$c_{ji} = \sum_{r=1}^{n} b_{rj} a_{ir}$$
and $(AB)^T = C^T = B^T A^T$.

(ii) $A^{TT} = (a_{ji})^T = (a_{ij}) = A$.

(iii) $(\lambda A + \mu B)^T = (\lambda a_{ji} + \mu b_{ji}) = \lambda(a_{ji} + \mu b_{ji}) = \lambda A^T + \mu B^T$.

(iv) $\delta_{ij} = \delta_{ji}$ so $I = I^T$.

Exercise 7.2.10

Let the jth row of A be $\mathbf{a}_j$.

$$A \in O(\mathbb{R}^n) \Leftrightarrow AA^T = I \Leftrightarrow \mathbf{a}_i\mathbf{a}_j^T = \delta_{ij} \Leftrightarrow \mathbf{a}_i \cdot \mathbf{a}_j = \delta_{ij} \Leftrightarrow \text{the } \mathbf{a}_j \text{ are o.n.}$$

Thus (i)⇔(ii). Since

$$A \in O(\mathbb{R}^n) \Rightarrow A^{TT}A^T = AA^T = I \Rightarrow A^T \in O(\mathbb{R}^n)$$

transposition gives (i)⇔(iii).

We turn to the statements.

(i) FALSE. Consider

$$\begin{pmatrix} 1 & -1 \\ 2 & 2 \end{pmatrix}.$$

(ii) FALSE. Consider

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}.$$

(iii) FALSE If $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}, \mathbf{a}_n$ are the rows of an orthogonal matrix, so are $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}, -\mathbf{a}_n$ (and $\mathbf{a}_n \neq \mathbf{0}$).

(iv) TRUE If $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_{n-1}$ are orthonormal in $\mathbb{R}^n$ then there are exactly two choices $\pm\mathbf{a}_n$ giving an orthonormal set and of these exactly one will give $\det A = 1$.

Exercise 7.2.13

Take
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}.$$
Note that the columns in each matrix are not orthogonal.

Exercise 7.3.2

Any pair of orthonormal vectors form a basis for $\mathbb{R}^2$. We observe that

$$\langle \mathbf{e}_1, \mathbf{e}_1 \rangle = 1,\ \langle \mathbf{e}_1, -\mathbf{e}_2 \rangle = -\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = 0,\ \langle -\mathbf{e}_2, -\mathbf{e}_2 \rangle = \langle \mathbf{e}_2, \mathbf{e}_2 \rangle = 1.$$

Observe that

$$\begin{aligned} \alpha \mathbf{e}_1 &= \cos\theta \mathbf{e}_1 + \sin\theta \mathbf{e}_2 = \cos\theta \mathbf{e}_1 - \sin\theta(-\mathbf{e}_2) \\ &= \alpha(-\mathbf{e}_2) = -\alpha \mathbf{e}_2 = \sin\theta \mathbf{e}_1 - \cos\theta \mathbf{e}_2 = \sin\theta \mathbf{e}_1 + \cos\theta(-\mathbf{e}_2) \end{aligned}$$

so, with respect to the new basis, α has matrix

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

Exercise 7.3.4

(i) We have
$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}x\cos\theta-y\sin\theta\\ x\sin\theta+y\cos\theta\end{pmatrix}$$
which I was told in school was a rotation through θ

We have
$$\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}-x\\ y\end{pmatrix}$$
which certainly looks like a reflection in the y axis to me.

Since
$$\langle \mathbf{e}_1,\mathbf{u}_1\rangle^2+\langle \mathbf{e}_1,\mathbf{u}_2\rangle^2=1,$$
we may take $\langle \mathbf{e}_1,\mathbf{u}_1\rangle=\cos\phi$, $\langle \mathbf{e}_1,\mathbf{u}_2\rangle=\sin\phi$. We have
$$\mathbf{e}_1=\cos\phi\mathbf{u}_1+\sin\phi\mathbf{u}_2$$
so
$$\begin{aligned}\mathbf{e}_1&=-\alpha\mathbf{e}_1\\ &=-\cos\phi\alpha\mathbf{u}_1-\sin\phi\alpha\mathbf{u}_2\\ &=-\cos\phi(\cos\theta\mathbf{u}_1+\sin\theta\mathbf{u}_2)-\sin\phi(\sin\theta\mathbf{u}_1+\cos\theta\mathbf{u}_2)\end{aligned}$$
Taking the inner product with $\mathbf{u}_1$ we get
$$\cos\phi=-\cos\phi\cos\theta-\sin\phi\sin\theta=-\cos(\theta-\phi)$$
so
$$\phi=\pi+\theta-\phi$$
and $\theta=2\phi-\pi$.

Exercise 7.3.6

Since a reflection in a plane has perpendicular eigenvectors with eigenvalues -1, 1, 1, α is a reflection only if $\det\alpha = -1$ so, by our discussion, α has matrix

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

so

$$\det(t\iota - \alpha) = \det\begin{pmatrix} t+1 & 0 & 0 \\ 0 & t-\cos\theta & \sin\theta \\ 0 & -\sin\theta & t-\cos\theta \end{pmatrix} = (t-1)(t^2-2t\cos\theta+1)$$

and we can only have the correct eigenvalues if $\cos\theta = 1$. In this case we clearly have a reflection in the plane spanned by $\mathbf{e}_2$ and $\mathbf{e}_3$.

By definition α is a reflection in the origin if and only if $\alpha = -\iota$ so if and only if $\cos\theta = -1$.

EXERCISE 7.3.7

$$\begin{aligned}\det(t\iota-\alpha) &= \det\begin{pmatrix} t-\cos\theta & \sin\theta \\ -\sin\theta & t-\cos\theta\end{pmatrix}\det\begin{pmatrix} t-\cos\phi & \sin\phi \\ -\sin\phi & t-\cos\phi\end{pmatrix} \\ &= (t^2-2t\cos\theta+1)(t^2-2t\cos\phi+1)\end{aligned}$$

with no real roots unless $\cos\theta=\cos\phi=1$ i.e. $\theta\equiv\phi\equiv 0 \pmod{\pi}$.

Exercise 7.4.3

Let $\mathbf{u} = \langle \mathbf{x}, \mathbf{n} \rangle \mathbf{n}$ and $\mathbf{v} = \mathbf{x} - \mathbf{u}$. Then

$$\langle \mathbf{v}, \mathbf{n} \rangle = \langle \mathbf{x} - \mathbf{u}, \mathbf{n} \rangle = \langle \mathbf{x} - \langle \mathbf{x}, \mathbf{n} \rangle \mathbf{n}, \mathbf{n} \rangle = \langle \mathbf{x}, \mathbf{n} \rangle - \langle \mathbf{x}, \mathbf{n} \rangle \langle \mathbf{n}, \mathbf{n} \rangle = 0$$

as required.

We observe that

$$\rho(x) = \mathbf{u} + \mathbf{v} - 2(\langle \mathbf{u}, \mathbf{n} \rangle + \langle \mathbf{v}, \mathbf{n} \rangle)\mathbf{n} = \mathbf{u} + \mathbf{v} - 2\langle \mathbf{x}, \mathbf{n} \rangle \mathbf{n} = \mathbf{v} - \mathbf{u}.$$

Exercise 7.4.6

Let A be the point given by the vector $\mathbf{a}$ and B be the point given by the vector $\mathbf{b}$. Let OM be an angle bisector of the line OA, OB of length 1. and let OM be given by the vector $\mathbf{m}$. Let $\mathbf{n}$ be a unit vector perpendicular to $\mathbf{m}$.

Exercise 7.4.9

The product of two reflections in lines OA, OB through the origin making an angle $\angle AOB = \theta$ is a rotation through 2θ.

To prove this, observe that, if ρ_1 and ρ_2 are reflections, then $\rho_2\rho_1$ preserves lengths and has determinant 1, so is a rotation. Now let $\mathbf{e}_1$ and $\mathbf{e}_2$ be orthonormal and let $\mathbf{n} = -\mathbf{e}_1$, $\mathbf{m} = -\cos\theta\mathbf{e}_1 - \sin\theta\mathbf{e}_2$. If

$$\begin{aligned}\rho_1(\mathbf{x}) &= \mathbf{x} - 2\langle\mathbf{x},\mathbf{n}\rangle\mathbf{n}\\ \rho_2(\mathbf{x}) &= \mathbf{x} - 2\langle\mathbf{x},\mathbf{m}\rangle\mathbf{m},\end{aligned}$$

then

$$\begin{aligned}\rho_2\rho_1\mathbf{e}_1 = \rho_2(\mathbf{e}_1) &= \mathbf{e}_1 - 2\langle\mathbf{e}_1,\mathbf{m}\rangle\mathbf{m}\\ &= \mathbf{e}_1 - 2\cos\theta(\cos\theta_1 + \sin\theta_2)\\ &= \cos 2\theta_1 + \sin 2\theta_2.\end{aligned}$$

Thus the rotations through angle θ are obtained from two reflections in lines at angle $\theta/2$ (in the correct sense).

Exercise 7.5.1

(i) Observe that

$$\begin{aligned}\|A\mathbf{x}-\mathbf{b}\|^2 &= \sum_{j=1}^{n}(x-b_j)^2\\ &= nx^2-2x\sum_{j=1}^{n}b_j+\sum_{j=1}^{n}b_j^2\\ &= n\left(x-n^{-1}\sum_{j=1}^{n}b_j\right)^2+\sum_{j=1}^{n}b_j^2-n\left(n^{-1}\sum_{j=1}^{n}b_j\right)^2\end{aligned}$$

is minimised by taking $x=n^{-1}\sum_{j=1}^{n}b_j$.

In the case of the mountain this amounts to the reasonable choice of the average of the measured heights.

(ii) Observe that

$$\|A\mathbf{x}-\mathbf{b}\|^2=f(x_1,x_2)$$

with

$$f(x_1,x_2)=\sum_{i=1}^{n}(x_1-v_ix_2-b_i)^2$$

We observe that $f(x_1,x_2)\to\infty$ as $|x_1|$, $|x_2|\to\infty$.

Since

$$\frac{\partial f}{\partial x_1}(x_1,x_2)=\sum_{i=1}^{n}2(x_1-v_ix_2-b_i)=2\left(nx_1-\sum_{i=1}^{n}b_i\right)$$

and

$$\frac{\partial f}{\partial x_2}(x_1,x_2)=-\sum_{i=1}^{n}2v_i(x_1-v_ix_2-b_i)=-2\left(x_2\sum_{i=1}^{n}v_i^2-\sum_{i=1}^{n}b_iv_i\right),$$

we see that f and therefore $\|A\mathbf{x}-\mathbf{b}\|$ has a unique minimum when $(x_1,x_2)=(\mu,\kappa)$.

We have sought to find a line $u+x_1v=x_2$ 'close to' the points (u_i,v_i) by minimising

$$\sum_{i=1}^{n}(u_i-x_1-x_2v_i)^2.$$

Exercise 7.5.2

Observe that

$$f(\mathbf{x})\sum_{i=1}^{n}\left(\sum_{j=1}^{m}a_{ij}x_j-b_i\right)^2$$

is a smooth function of $\mathbf{x}$ with $f(\mathbf{x})\rightarrow\infty$ as $\|\mathbf{x}\|\rightarrow\infty$. Differentiating and solving the resulting n linear equations in n unknowns will, in general (to see possible problems, take $a_{ij}=0$ for all i and j) give a unique stationary point which will be a minimum.

The other two penalty functions cannot (in general) be minimised by calculus techniques, but could be attacked as linear programming problems.

Exercise 7.5.3

(i) If we write the *columns* of A as *column vectors* $\mathbf{a}_1$, $\mathbf{a}_2$, $\ldots \mathbf{a}_m$, then (possibly after reordering the columns of A), the Gram–Schmidt method gives us orthonormal row vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\ldots$, $\mathbf{e}_k$ such that

$$\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{e}_1\\
\mathbf{a}_2 &= r_{12}\mathbf{e}_1 + r_{22}\mathbf{e}_2\\
\mathbf{a}_3 &= r_{13}\mathbf{e}_1 + r_{23}\mathbf{e}_2 + r_{33}\mathbf{e}_3\\
&\;\vdots\\
\mathbf{a}_k &= r_{1k}\mathbf{e}_1 + r_{2k}\mathbf{e}_2 + r_{3k}\mathbf{e}_3 + \ldots + r_{kk}\mathbf{e}_{kk}
\end{aligned}$$

for some r_{ij} $[1 \leq j \leq i \leq k]$ with $r_{ii} \neq 0$ for $1 \leq i \leq k.$ and

$$\mathbf{a}_p = r_{1p}\mathbf{e}_1 + r_{2p}\mathbf{e}_2 + r_{3p}\mathbf{e}_3 + \ldots + r_{pk}\mathbf{e}_k$$

for r_{ij} $[1 \leq i \leq k,\ k+1 \leq p \leq m]$. We set $r_{ij} = 0$ in all the cases with $1 \leq i \leq n$ and $1 \leq j \leq m$ where this has not previously been defined.

Using the Gram–Schmidt method again, we can now find $\mathbf{e}_{k+1}$, $\mathbf{e}_{k+2}$, $\ldots$, $\mathbf{e}_n$ so that the vectors $\mathbf{e}_j$ with $1 \leq j \leq n$ form an orthonormal basis for the space $\mathbb{R}^n$ of column vectors. If we take Q to be the $n \times n$ matrix with jth column $\mathbf{e}_j$, then Q is orthogonal and condition ★ gives

$$A = QR.$$

(ii) We have

$$\begin{aligned}
r_{ii} \neq 0 \ \forall\ 1 \leq i \leq m &\Leftrightarrow \operatorname{rank} R = m\\
&\Leftrightarrow \operatorname{rank} Q^T A = m\\
&\Leftrightarrow \operatorname{rank} A = m
\end{aligned}$$

since the linear map corresponding to Q^T is an isomorphism.

Exercise 7.5.4

(i) We know that
$$\pi = \{A\mathbf{t} : \mathbf{t} \in \mathbb{R}^n\}$$
is a subspace of $\mathbb{R}^m$. Thus (e.g. by Theorem 7.1.7) it contains a closest point $\mathbf{y}$ to $\mathbf{b}$. Choose $\mathbf{x}_0$ with $A\mathbf{x}_0 = \mathbf{y}$. to obtain a minimiser for our problem.

Since A has rank strictly less than m we can find a non zero $\mathbf{u} \in \mathbb{R}^m$ with $A\mathbf{u} = \mathbf{0}$. Then every choice $\mathbf{x} = \mathbf{x}_0 + \lambda\mathbf{z}$ minimises $\|A\mathbf{x} - \mathbf{b}\|$.

(ii) The QL result is just the QR result with a renumbering of the basis.

If A is an $m \times n$ matrix, Q an $n \times n$ orthogonal matrix and R an $n \times m$ upper triangular matrix with $A^T = QR$, then $A = R^TQ^T$, Q^T is orthogonal and R^T a lower triangular matrix. Thus the QR results imply the corresponding LQ results and the QL results imply the corresponding RQ results.

We could have used the QL in the discussion. However, the RQ and LQ results give Q in the wrong place when we derive the opening formulae. In addition (at least if we follow the obvious path) we factorise an $n \times m$ matrix with $n \geq m$ and we are not usually interested in cases where we have more unknowns than equations.

Exercise 7.5.5

Since ρ is norm preserving,

$$\rho(\mathbf{a}) = \mathbf{b} \Rightarrow \|\mathbf{a}\| = \|\rho(\mathbf{a})\| = \|\mathbf{b}\|.$$

Now take $\|\mathbf{a}\| = \|\mathbf{b}\|$ with $\mathbf{b} \neq \pm\mathbf{a}$ and set $\mathbf{c} = (\mathbf{a} - \mathbf{b})/2$. Suppose that

$$\rho\mathbf{x} = \lambda\mathbf{x} + \mu\langle\mathbf{c}, \mathbf{x}\rangle\mathbf{c}$$

describes a reflection with $\rho\mathbf{a} = \mathbf{b}$ and $\rho\mathbf{x} = \mathbf{x}$ whenever $\langle\mathbf{x}, \mathbf{a}\rangle = \langle\mathbf{x}, \mathbf{b}\rangle$. If $\mathbf{x} \perp \mathbf{c}$ then

$$\mathbf{x} = \rho(\mathbf{x}) = \lambda\mathbf{x}$$

so $\lambda = 1$.

Since

$$\langle\mathbf{c}, \mathbf{a}\rangle = \tfrac{1}{2}(\|\mathbf{a}\|^2 + \langle\mathbf{a}, \mathbf{b}\rangle) = \tfrac{1}{2}(\|\mathbf{b}\|^2 + \langle\mathbf{b}, \mathbf{a}\rangle),$$

we have $\rho(\mathbf{c}) = \mathbf{c}$, so

$$\mathbf{c} = \mathbf{c} + \mu\langle\mathbf{c}, \mathbf{c}\rangle\mathbf{c}$$

and $\mu = -2\|\mathbf{c}\|^2$. Thus

$$\rho\mathbf{x} = \mathbf{x} - 2\frac{\langle\mathbf{c}, \mathbf{x}\rangle}{\|\mathbf{c}\|^2}\mathbf{c}$$

Conversely, if

$$\rho\mathbf{x} = \mathbf{x} - 2\frac{\langle\mathbf{c}, \mathbf{x}\rangle}{\|\mathbf{c}\|^2}\mathbf{c},$$

then $\rho\mathbf{c} = -\mathbf{c}$ and

$$\mathbf{d} \perp \mathbf{c} \Rightarrow \rho\mathbf{d} = \mathbf{d},$$

so ρ is a reflection. Further

$$\langle\mathbf{x}, \mathbf{a}\rangle = -\langle\mathbf{x}, \mathbf{b}\rangle \Leftrightarrow \langle\mathbf{x}, \mathbf{c}\rangle = 0 \Leftrightarrow \rho\mathbf{x} = -\mathbf{x}$$

and

$$\begin{aligned}
\rho\mathbf{a} &= \mathbf{a} - 2\frac{\langle\mathbf{c}, \mathbf{a}\rangle}{\|\mathbf{c}\|^2}\mathbf{c} \\
&= \mathbf{a} - \frac{\langle\mathbf{c}, \mathbf{a}\rangle + \langle\mathbf{c}, \mathbf{b}\rangle}{\|\mathbf{c}\|^2}\mathbf{c} \\
&= \mathbf{a} - 2\frac{\langle\mathbf{c}, \mathbf{c}\rangle}{\|\mathbf{c}\|^2}\mathbf{c} \\
&= \mathbf{a} - 2\mathbf{c} = \mathbf{b}
\end{aligned}$$

and, since ρ is a reflection $\rho\mathbf{b} = \rho\mathbf{a}$.

We have

$$T_{ij}x_j = x_i - 2c_jx_j\|\mathbf{c}\|^{-2}c_i = (\delta_{ij} - 2(c_kc_k)^{-2}c_ic_j)x_j,$$

so

$$T_{ij} = \delta_{ij} - 2(c_kc_k)^{-2}c_ic_j.$$

If $\mathbf{a} = (\pm\|\mathbf{a}\|, 0, 0, \dots, 0)$, take $T_1 = I$. Otherwise, take $T_1 = T$ with T as above.

If C is an $n \times m$ matrix with $c_{ij} = 0$ for all $j > i$ when $i \leq r$, consider the $(n-r) \times (m-r)$ matrix A formed by the c_{ij} with $n \geq i \geq n-r+1$ $m \geq j \geq m-r+1$. We can find a $(n-r) \times (n-r)$ reflection matrix (or the identity) T with TA having first column zero except possibly the first entry. If S is the $n \times n$ matrix given by

$$S = \begin{pmatrix} I & 0 \\ 0 & A \end{pmatrix},$$

then $H = SC$ is an $n \times m$ matrix with $h_{ij} = 0$ for all $j > i$ when $i \leq r+1$.

Thus we can find reflection (or identity) matrices T_j such that

$$T_{n-1}T_{n-2} \dots T_1 A = R$$

is upper triangular. The matrix $Q = T_1 T_2 \dots T_{n-1}$ is the product of reflection (or identity) so of orthonormal matrices so (since $O(\mathbb{R}^n)$ is a group under multiplication) an orthonormal matrix.

We have

$$A = (T_{n-1}T_{n-2} \dots T_1)^{-1} R = T_1^{-1} T_2^{-1} \dots T_n^{-1} R = T_1 T_2 \dots T_n R = QR.$$

Exercise 7.5.6

We are interested in the vector $\mathbf{a} = (1,2,2)^T$ of length 3 which we wish to reflect to the vector $\mathbf{b} = (3,0,0)^T$. We thus want a reflection in the plane perpendicular to

$$\mathbf{c} = (\mathbf{a} - \mathbf{b})/2 = (-1,1,1)^T.$$

This reflection is given by

$$\rho(\mathbf{x}) = \mathbf{x} - 2\|\mathbf{c}\|^{-2}(\mathbf{c}\cdot\mathbf{x})\mathbf{c}$$

with matrix

$$\begin{aligned}T &= I - 2\|\mathbf{c}\|^2\mathbf{c}^T\mathbf{c}\\ &= \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} - \frac{2}{3}\begin{pmatrix}1&-1&-1\\-1&1&1\\-1&1&1\end{pmatrix}\\ &= \begin{pmatrix}\frac{1}{3}&\frac{2}{3}&\frac{2}{3}\\ \frac{2}{3}&\frac{1}{3}&-\frac{2}{3}\\ \frac{2}{3}&-\frac{2}{3}&\frac{1}{3}\end{pmatrix}\end{aligned}$$

Now

$$TA = \begin{pmatrix}3&\frac{11}{3}&\frac{4}{3}\\0&0&1\\0&1&-\frac{1}{3}\end{pmatrix}$$

so (interchanging the second and third row) TA is upper triangular.

Exercise 7.5.7

We are interested in the vector $\mathbf{a} = (2,2,-1)^T$ of length 3 which we wish to reflect to the vector $\mathbf{b} = (3,0,0)^T$. We thus want a reflection in the plane perpendicular to

$$\mathbf{c} = (\mathbf{a}-\mathbf{b})/2 = (-\tfrac{1}{2}, 1, -\tfrac{1}{2})^T.$$

Setting $\mathbf{d} = 2\mathbf{c} = (-1,2,-1)^T$, reflection is given by

$$\rho(\mathbf{y}) = \mathbf{y} - 2\|\mathbf{d}\|^2\mathbf{d}\cdot\mathbf{y}\mathbf{d}$$

with matrix

$$\begin{aligned} T &= I - 2\|\mathbf{c}\|^2\mathbf{d}^T\mathbf{d} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix} \end{aligned}$$

Setting

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}.$$

We see that (since Q is orthogonal) we seek the least squares fit for

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

that is to say

$$\begin{pmatrix} 1 & 3 \\ 0 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 4 \\ 3 \\ 0 \\ 3 \end{pmatrix}$$

so we require $x_2 = 1$, $x_1 = 1$.

Let f be the square of the norm of

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

Then $f(x_1, x_2) = (x_1+3x_2-4)^2+(2x_2-1)^2+(2x_2-4)^2+(x_2+1)^2$ and the unique stationary value of f (which must be minimum) is given by

$$0 = \frac{\partial f}{\partial x_1} = 2(x_1 + 3x_2 - 4)$$
$$0 = \frac{\partial f}{\partial x_2} = 2\big(3(x_1 + 3x_2 - 4) + 2(2x_2 - 1) + 2(2x_2 - 4) + (x_2 + 1)\big)$$

that is to say, by

$$\begin{aligned} 4 &= x_1 + 3x_2 \\ 21 &= 3x_1 + 18x_2, \end{aligned}$$

that is to say, by

$$\begin{aligned} 4 &= x_1 + 3x_2 \\ 7 &= x_1 + 6x_2, \end{aligned}$$

so we require $x_2 = 1$, $x_1 = 1$, as stated earlier.

Exercise 8.1.7

(i) We have

$$A^T = A \Rightarrow (P^T AP)^T = P^T A^T P^{TT} = P^T AP$$

and

$$\begin{aligned}(P^T AP)^T = P^T AP &\Rightarrow P^T A^T P = P^T AP\\ &\Rightarrow P(P^T A^T P)P^T = P((P^T AP)P^T\\ &\Rightarrow A^T = A.\end{aligned}$$

(ii) If $P \in SO(\mathbb{R}^n)$, then set $Q = P$.

If $P \notin SO(\mathbb{R}^n)$, then set $Q = PD$. We have

$$\det Q = \det P \det D = (-1)^2 = 1$$

and $Q^T Q = D^T P^T PD = DID = I$, so $Q \in SO(\mathbb{R}^n)$. Further, since $P^T AP$ is diagonal,

$$Q^T AQ = D^T P^T APD = D(P^T AP)D = P^T AP$$

and $Q^T AQ$ is diagonal.

Exercise 8.2.2

If f $a_{kj} = a_{jk}$ for all j and k, $\sum_{j=1}^n a_{kj}u_j = \lambda u_k$ and $\sum_{j=1}^n a_{kj}v_j = \mu v_k$ for all k, but $\lambda \neq \mu$, then

$$\begin{aligned}
\lambda \sum_{k=1}^n u_k v_k &= \sum_{k=1}^n \lambda u_k v_k \\
&= \sum_{k=1}^n \sum_{j=1}^n a_{kj} u_j v_k = \sum_{k=1}^n \sum_{j=1}^n a_{jk} u_j v_k \\
&= \sum_{j=1}^n \sum_{k=1}^n u_j a_{jk} v_k = \sum_{j=1}^n \mu u_j v_j \\
&= \mu \sum_{j=1}^n u_j v_j = \mu \sum_{k=1}^n u_k v_k,
\end{aligned}$$

so $\sum_{k=1}^n u_k v_k = 0$.

Exercise 8.2.6

We have (by Cramer's rule or direct calculation),

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

so

$$PAP^{-1} = P(AP^{-1}) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 2 & -1 \end{pmatrix}.$$

Observe that P is not orthogonal.

Exercise 8.2.7

Let
$$A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$$
so $A^T = A$.
$$\det(tI - A) = (t-1)(t+1) + 1 = t^2$$
so all the eigenvalues of A are 0 and, if A were diagonalisable, we would have $A = 0$ which is not the case.

Exercise 8.2.9

Interchanging x_2 and x_3 coordinates is the same as reflecting in a plane passing through the x_1 axis and a line bisecting the angle between the x_2 and x_3 axes (so perpendicular to $(0, 1, -1)^T$).

EXERCISE 8.2.10

If there are n distinct eigenvalues λ_j with associated eigenvectors $\mathbf{e}_j$ then

$$P^T AP = D$$

with P orthogonal and D diagonal if and only if the columns of P consist of the eigenvectors $\pm\mathbf{e}_j$ (with choice of signs) in some order so there are exactly $2^n n!$ such matrices.

If A has less than n distinct eigenvalues then $\mathbb{R}^n$ has a basis of eigenvectors $\mathbf{e}_j$ with $\mathbf{e}_{n-1}$ and $\mathbf{e}_n$ having the same eigenvalues, then, if P has rth column $\mathbf{e}_r$ for $1 \leq r \leq n-2$ $n-1$th column $\cos\theta\mathbf{e}_{n-1} - \sin\theta\mathbf{e}_n$ and nth column $\sin\theta\mathbf{e}_{n-1} + \cos\theta\mathbf{e}_n$, we have P orthogonal and

$$P^T AP = D$$

so case (ii) occurs.

Exercise 8.3.2

Observe that there exists a diagonal matrix D and an orthogonal matrix P with
$$D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \text{ and } A = P^T A P$$
where λ_1, λ_2 are the eigenvalues of A. Parts (i), (ii) and (iii) can be read off from the observation that $\det A = \det D = \lambda_1 \lambda_2$ and $\operatorname{Tr} A = \operatorname{Tr} D = \lambda_1 + \lambda_2$.

(iv) $\det A = uw - v^2 \leq uw$, so $\det A > 0 \Rightarrow u \neq 0$.

If $u > 0$ and $\det A > 0$, then $w > 0$ so $\operatorname{Tr} A > 0$, so, by (iii), the eigenvalues of A are strictly positive.

If $u < 0$ and $\det A > 0$, then $w < 0$, so $\operatorname{Tr} A < 0$, so, by applying (iii) to A, the eigenvalues of A are strictly positive.

Exercise 8.3.3★

Exercise 8.3.4

By an orthogonal transformation we may reduce the equation $ax^2 + 2bxy + cy^2 = d$ to $Ax^2 + By^2 = K$. By multiplying by -1 if necessary we may assume $K \geq 0$ and by interchanging x and y if necessary we may assume that $A = 0 \Rightarrow B = 0$ and if A and B are non-zero with opposite signs then $A > 0 > B$

The cases may now be enumerated as follows.

(1) $K > 0$, $A > 0$, $B > 0$ ellipse.

(2) $K > 0$, $A < 0$, $B < 0$ empty set.

(3) $K > 0$, $A > 0$, $B < 0$ hyperbola.

(4) $K = 0$, $A > 0$, $B > 0$ single point $\{\mathbf{0}\}$.

(5) $K = 0$, $A < 0$, $B < 0$ single point $\{\mathbf{0}\}$.

(6) $K = 0$, $A > 0$, $B < 0$ pair of lines meeting at $\mathbf{0}$.

(7) $K > 0$, $A > 0$ $B = 0$ pair of parallel lines.

(8) $K > 0$, $A < 0$, $B = 0$ empty set.

(7) $K = 0$, $A > 0$ $B = 0$ single line through $\mathbf{0}$.

(8) $K = 0$, $A < 0$, $B = 0$ single line through $\mathbf{0}$.

(9) $K > 0$, $A = 0$ $B = 0$ empty set.

(10) $K = 0$ $A = 0$, $B = 0$ whole plane.

Exercise 8.3.5

Let

$$\begin{pmatrix} 8 & 6^{1/2} \\ 6^{1/2} & 7 \end{pmatrix}$$

Then

$$\det(t - A) = (t-8)(t-7) - 6 = t^2 - 15t + 50 = (t-10)(t-5)$$

so A has eigenvalues 5 and 10.

$$A \begin{pmatrix} x \\ y \end{pmatrix} = 5 \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \begin{cases} 3x & = 6^{1/2} y \\ 2y & = 6^{1/2} x \end{cases}$$

so the eigenvector $(1, (3/2)^{1/2})^T$ gives one axis of symmetry.

$$A \begin{pmatrix} x \\ y \end{pmatrix} = 10 \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \begin{cases} -2x & = 6^{1/2} y \\ -3y & = 6^{1/2} x \end{cases}$$

so the eigenvector $(1, -(2/3)^{1/2})^T$ gives the other axis of symmetry.

Exercise 8.4.2

(i) $\langle \mathbf{z}, \mathbf{z} \rangle = \sum_{j=1}^{n} |z_j|^2$ is always real and positive.

(ii) We have

$$\begin{aligned}\langle \mathbf{z}, \mathbf{z} \rangle = 0 &\Leftrightarrow \sum_{j=1}^{n} |z_j|^2 = 0 \\ &\Leftrightarrow |z_j|^2 = 0 \text{ for } 1 \leq j \leq n \\ &\Leftrightarrow |z_j| = 0 \text{ for } 1 \leq j \leq n \\ &\Leftrightarrow \mathbf{z} = \mathbf{0}.\end{aligned}$$

(iii) We have

$$\begin{aligned}\langle \lambda \mathbf{z}, \mathbf{w} \rangle &= \sum_{j=1}^{n} (\lambda z_j) w_j^* \\ &= \lambda \sum_{j=1}^{n} z_j w_j^* = \lambda \langle \mathbf{z}, \mathbf{w} \rangle.\end{aligned}$$

(iv) We have

$$\begin{aligned}\langle \mathbf{z} + \mathbf{u}, \mathbf{w} \rangle &= \sum_{j=1}^{n} (z_j + u_j) w_j^* \\ &= \sum_{j=1}^{n} z_j w_j^* + u_j w_j^* = \langle \mathbf{z}, \mathbf{w} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle.\end{aligned}$$

(v) $\langle \mathbf{z}, \mathbf{w} \rangle^* = \left(\sum_{j=1}^{n} z_j w_j^* \right)^* = \sum_{j=1}^{n} z_j^* w_j^{**} = \sum_{j=1}^{n} z_j^* w_j = \langle \mathbf{w}, \mathbf{z} \rangle$.

Exercise 8.4.3

The result is trivial if $\mathbf{z} = \mathbf{0}$ or $\mathbf{w} = \mathbf{0}$, so we need only consider the case when $\mathbf{z}$, $\mathbf{w} \neq \mathbf{0}$.

Suppose first that $\langle \mathbf{z}, \mathbf{w} \rangle$ is real and positive. Then, if λ is real,

$$\begin{aligned}\langle (\lambda\mathbf{z} + \mathbf{w}), (\lambda\mathbf{z} + \mathbf{w}) \rangle &= \lambda^2 \langle \mathbf{z}, \mathbf{z} \rangle + 2\lambda \langle \mathbf{z}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \\ &= \left(\lambda \langle \mathbf{z}, \mathbf{z} \rangle^{1/2} + \frac{\langle \mathbf{z}, \mathbf{w} \rangle}{\langle \mathbf{z}, \mathbf{z} \rangle^{1/2}} \right) + \langle \mathbf{w}, \mathbf{w} \rangle - \frac{\langle \mathbf{z}, \mathbf{w} \rangle^2}{\langle \mathbf{z}, \mathbf{z} \rangle}\end{aligned}$$

Thus, taking $\lambda = \lambda_0$, with $\lambda_0 = (\langle \mathbf{z}, \mathbf{w} \rangle)^{1/2} (\langle \mathbf{z}, \mathbf{z} \rangle)^{-1/2}$ we see that

$$\langle \mathbf{w}, \mathbf{w} \rangle - \frac{\langle \mathbf{z}, \mathbf{w} \rangle^2}{\langle \mathbf{z}, \mathbf{z} \rangle} \geq 0$$

with equality if and only if

$$\langle (\lambda_0\mathbf{z} + \mathbf{w}), (\lambda_0\mathbf{z} + \mathbf{w}) \rangle = 0$$

and so if and only if

$$\lambda_0\mathbf{z} + \mathbf{w} = \mathbf{0}$$

In general, we may choose θ so that

$$e^{i\theta} \langle \mathbf{z}, \mathbf{w} \rangle$$

is real and positive. Then the result above, applied to $e^{i\theta}\mathbf{z}$, yields

$$\langle \mathbf{w}, \mathbf{w} \rangle - e^{2i\theta} \frac{\langle \mathbf{z}, \mathbf{w} \rangle^2}{\langle \mathbf{z}, \mathbf{z} \rangle} \geq 0$$

so

$$|\langle \mathbf{z}, \mathbf{w} \rangle| \leq \|\mathbf{z}\| \|\mathbf{w}\|.$$

with equality only possible if

$$\lambda_1\mathbf{z} + \mathbf{w} = 0$$

for some $\lambda_1 \in \mathbb{C}$. By inspection, if this last condition holds, we do have equality.

Exercise 8.4.4

(i) $\|\mathbf{z}\| \geq 0$ since we take the positive square root.

(ii) $\|\mathbf{z}\| = 0 \Leftrightarrow \langle \mathbf{z}, \mathbf{z} \rangle = 0 \Leftrightarrow \mathbf{z} = \mathbf{0}$.

(iii) $\|\lambda\mathbf{z}\|^2 = \langle \lambda\mathbf{z}, \lambda\mathbf{z} \rangle = \lambda\lambda^*\langle \mathbf{z}, \mathbf{z} \rangle = (|\lambda|\|\mathbf{z}\|)^2$, so $\|\lambda\mathbf{z}\| = |\lambda|\|\mathbf{z}\|$.

(iv) We have

$$\begin{aligned}
\|\mathbf{z}+\mathbf{w}\|^2 &= \langle \mathbf{z}+\mathbf{w}, \mathbf{z}+\mathbf{w} \rangle \\
&= \|\mathbf{z}\|^2 + \langle \mathbf{z}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{z} \rangle + \|\mathbf{w}\|^2 \\
&= \|\mathbf{z}\|^2 + 2\Re\langle \mathbf{z}, \mathbf{w} \rangle + \|\mathbf{w}\|^2 \\
&\leq \|\mathbf{z}\|^2 + 2\|\mathbf{z}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2 \\
&= (\|\mathbf{z}\| + \|\mathbf{w}\|)^2
\end{aligned}$$

The result follows on taking square roots.

Exercise 8.4.6

(i) Observe that

$$\sum_{j=1}^{n} \lambda_j \mathbf{e}_j = 0 \Rightarrow \left\langle \sum_{j=1}^{n} \lambda_j \mathbf{e}_j, \mathbf{e}_k \right\rangle = 0 \quad \forall k$$
$$\Rightarrow \sum_{j=1}^{n} \lambda_j \delta_{jk} = 0 \quad \forall k \Rightarrow \lambda_k = 0 \quad \forall k,$$

so we have a set of n linearly independent vectors in a space of dimension n which is thus a basis.

(ii) By (i),

$$\mathbf{z} = \sum_{j=1}^{n} \lambda_j \mathbf{e}_j$$

for some λ_j. We observe that

$$\lambda_j = \sum_{j=1}^{n} \lambda_j \delta_{jk} = \left\langle \sum_{j=1}^{n} \lambda_j \mathbf{e}_j, \mathbf{e}_k \right\rangle = \langle \mathbf{z}, \mathbf{e}_k \rangle$$

(iii) False. If $n > m$, m vectors cannot span $\mathbb{C}^m$.

Exercise 8.4.7

If $k = q$, we are done, since a linearly independent set forms a basis if and only if the number of elements in the basis equals the dimension of the space.

If not, then $k < q$ and there exists a

$$\mathbf{u} \in U \setminus \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k\}$$

so

$$\mathbf{v} = \mathbf{u} - \sum_{j=1}^{k} \langle \mathbf{u}, \mathbf{e}_j \rangle \mathbf{e}_j$$

is a non-zero element of U. Setting $\mathbf{e}_{k+1} = \|\mathbf{v}\|^{-1}\mathbf{v}$ we have $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_{k+1}$ orthonormal in U.

After repeating the process at most q times we obtain a basis for U.

Exercise 8.4.8

(i) Observe that

$$\begin{aligned}\left\|\mathbf{z}-\sum_{j=1}^{k}\lambda_j\mathbf{e}_j,\right\|^2 &= \|\mathbf{z}\|^2-\sum_{j=1}^{k}\lambda_j\langle\mathbf{e}_j,\mathbf{z}\rangle^*-\sum_{j=1}^{k}\lambda_j^*\langle\mathbf{e}_j,\mathbf{z}\rangle+\sum_{j=1}^{k}|\lambda_j|^2\\ &=\sum_{j=1}^{k}(\lambda_j-\langle\mathbf{e}_j,\mathbf{z}\rangle)(\lambda_j-\langle\mathbf{e}_j,\mathbf{z}\rangle)^*+\|\mathbf{z}\|^2-\sum_{j=1}^{k}|\langle\mathbf{e}_j,\mathbf{z}\rangle|^2\\ &=\sum_{j=1}^{k}|\lambda_j-\langle\mathbf{e}_j,\mathbf{z}\rangle|^2+\|\mathbf{z}\|^2-\sum_{j=1}^{k}|\langle\mathbf{e}_j,\mathbf{z}\rangle|^2\\ &\geq\|\mathbf{z}\|^2-\sum_{j=1}^{k}|\langle\mathbf{e}_j,\mathbf{z}\rangle|^2\end{aligned}$$

with equality if and only if $\lambda_j=\langle\mathbf{z},\mathbf{e}_j\rangle$ for all j.

(ii) Since

$$\mathbf{z}=\sum_{j=1}^{k}\langle\mathbf{z},\mathbf{e}_j\rangle\mathbf{e}_j$$

if and only if $\mathbf{z}\in\operatorname{span}\{\mathbf{e}_1,\,\mathbf{e}_2,\,\dots,\,\mathbf{e}_k\}$, it follows from (i) that

$$\|\mathbf{z}\|^2\geq\sum_{j=1}^{k}\langle\mathbf{z},\mathbf{e}_k\rangle^2,$$

with equality if and only if $\mathbf{z}\in\operatorname{span}\{\mathbf{e}_1,\,\mathbf{e}_2,\,\dots,\,\mathbf{e}_k\}$.

Exercise 8.4.9

We have

$$\begin{aligned}\|\mathbf{z}+\mathbf{w}\|^2-\|\mathbf{z}-\mathbf{w}\|^2 &= \langle \mathbf{z}+\mathbf{w},\mathbf{z}+\mathbf{w}\rangle - \langle \mathbf{z}-\mathbf{w},\mathbf{z}-\mathbf{w}\rangle\\ &= \|\mathbf{z}\|^2+\langle \mathbf{z},\mathbf{w}\rangle-\langle \mathbf{z},\mathbf{w}\rangle^*+\|\mathbf{w}\|^2-\|\mathbf{z}\|^2+\langle \mathbf{z},\mathbf{w}\rangle+\langle \mathbf{z},\mathbf{w}\rangle^*-\|\mathbf{w}\|^2\\ &= 4\Re\langle \mathbf{z},\mathbf{w}\rangle.\end{aligned}$$

Thus

$$\|\mathbf{z}+i\mathbf{w}\|^2-\|\mathbf{z}-i\mathbf{w}\|^2 = 4\Re\langle \mathbf{z},i\mathbf{w}\rangle = \Im\langle \mathbf{z},\mathbf{w}\rangle$$

and

$$\|\mathbf{z}+\mathbf{w}\|^2-\|\mathbf{z}-\mathbf{w}\|^2+i\|\mathbf{z}+i\mathbf{w}\|^2-i\|\mathbf{z}-i\mathbf{w}\|^2 = 4\langle \mathbf{z},\mathbf{w}\rangle.$$

Exercise 8.4.10

Existence Choose an orthonormal basis $\mathbf{e}_j$. If α has matrix $A = (a_{ij})$ with respect to this basis, let α^* be the linear map with matrix $A^* = (b_{ij})$ where $b_{ij} = a^*_{ji}$. Then

$$\begin{aligned}\left\langle \alpha \sum_{j=1}^{n} z_j \mathbf{e}_j, \sum_{r=1}^{n} w_j \mathbf{e}_r \right\rangle &= \left\langle \sum_{j=1}^{n} z_j \sum_{i=1}^{n} a_{ij}\mathbf{e}_i, \sum_{r=1}^{n} w_r \mathbf{e}_r \right\rangle \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n}\sum_{r=1}^{n} z_j w_r a_{ij}\delta_{ir} \\ &= \sum_{j=1}^{n}\sum_{i=1}^{n} z_j w_i a_{ij} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} z_j w_i b^*_{ji} \\ &= \left\langle \sum_{j=1}^{n} z_j \mathbf{e}_j, \alpha^* \sum_{r=1}^{n} w_j \mathbf{e}_r \right\rangle,\end{aligned}$$

so

$$\langle \alpha \mathbf{z}, \mathbf{w} \rangle = \langle \mathbf{z}, \alpha^* \mathbf{w} \rangle$$

for all $\mathbf{z}$, $\mathbf{w} \in \mathbb{C}^n$.

Uniqueness Observe that, if β and γ are linear maps with

$$\langle \alpha \mathbf{z}, \mathbf{w} \rangle = \langle \mathbf{z}, \beta \mathbf{w} \rangle = \langle \mathbf{z}, \gamma \mathbf{w} \rangle$$

for all $\mathbf{z}$, $\mathbf{w} \in \mathbb{C}^n$, then

$$(\beta - \gamma)\mathbf{w} = 0$$

for all $\mathbf{w} \in \mathbb{C}^n$ (set $\mathbf{z} = (\beta - \gamma)\mathbf{w}$), so $\beta = \gamma$.

The required formula $A^* = (b_{ij})$ with $b_{ij} = a^*_{ji}$ follows from the results already established in the question.

Finally

$$\begin{aligned}
\det \alpha^* = \det A^* &= \sum_{\sigma \in S_n} \zeta(\sigma) \prod_{i=1}^{n} a^*_{\sigma(i)i} \\
&= \sum_{\sigma \in S_n} \zeta(\sigma) \prod_{j=1}^{n} a^*_{j\sigma^{-1}(j)} \\
&= \sum_{\tau \in S_n} \zeta(\sigma) \prod_{j=1}^{n} a^*_{j\tau(j)} \\
&= \left(\sum_{\tau \in S_n} \zeta(\sigma) \prod_{j=1}^{n} a_{j\tau(j)} \right)^* \\
&= (\det A)^* = (\det \alpha)^*
\end{aligned}$$

Exercise 8.4.11

(i)⇒(ii) If (i) holds, then using the polarisation identity

$$\begin{aligned}
4\langle\alpha\mathbf{z},\alpha\mathbf{w}\rangle &= \|\alpha\mathbf{z}+\alpha\mathbf{w}\|^2-\|\alpha\mathbf{z}-\alpha\mathbf{w}\|^2+i\|\alpha\mathbf{z}+i\alpha\mathbf{w}\|^2-i\|\alpha\mathbf{z}-i\alpha\mathbf{w}\|^2\\
&= \|\alpha(\mathbf{z}+\mathbf{w})\|^2-\|\alpha(\mathbf{z}-\mathbf{w})\|^2+i\|\alpha(\mathbf{z}+i\mathbf{w})\|^2-i\|\alpha(\mathbf{z}-i\mathbf{w})\|^2\\
&= \|\mathbf{z}+\mathbf{w}\|^2-\|\mathbf{z}-\mathbf{w}\|^2+i\|\mathbf{z}+i\mathbf{w}\|^2-i\|\mathbf{z}-i\mathbf{w}\|^2\\
&= 4\langle\mathbf{z},\mathbf{w}\rangle,
\end{aligned}$$

so (ii) holds.

(ii)⇒(iii) if (ii) holds

$$\langle\alpha^*\alpha\mathbf{z},\mathbf{w}\rangle=\langle\mathbf{w},\alpha^*\alpha\mathbf{z}\rangle^*=\langle\alpha\mathbf{w},\alpha\mathbf{z}\rangle^*=\langle\alpha\mathbf{z},\alpha\mathbf{w}\rangle=\langle\mathbf{z},\mathbf{w}\rangle$$

for all $\mathbf{w}$ so $\alpha^*\alpha\mathbf{z}=\mathbf{z}$ for all $\mathbf{z}$ so $\alpha^*\alpha=\iota$.

(iii)⇒(iv) Immediate.

(iv)⇔(v)⇔(vi) Immediate.

(iv)⇒(i) If (iv) holds,

$$\|\alpha\mathbf{z}\|^2=\langle\alpha\mathbf{z},\alpha\mathbf{z}\rangle=\langle\mathbf{z},\alpha^*\alpha\mathbf{z}\rangle=\langle\mathbf{z},\mathbf{z}\rangle=\|\mathbf{z}\|^2.$$

Exercise 8.4.12

(i) $U(\mathbb{C}^n)$ is a subset of the group $GL(\mathbb{C}^n)$. We have

$$\begin{aligned}\alpha,\, \beta \in U(\mathbb{C}^n) \Rightarrow (\alpha\beta)^*(\alpha\beta) &= (\beta^*\alpha^*)(\alpha\beta)\\ &= \beta^*(\alpha^*\alpha)\beta = \beta^*\iota\beta = \beta^*\beta = \iota\\ \Rightarrow \alpha\beta \in U(\mathbb{C}^n)\end{aligned}$$

and

$$\alpha \in U(\mathbb{C}^n) \Rightarrow (\alpha^{-1})^* = \alpha^{**} = \alpha = (\alpha^{-1})^{-1} \Rightarrow \alpha^{-1} \in U(\mathbb{C}^n)$$

whilst

$$\iota^*\iota = \iota\iota = \iota$$

so $\iota \in U(\mathbb{C}^n)$,

Thus $U(\mathbb{C}^n)$ is a subgroup of $GL(\mathbb{C}^n)$.

(ii) $1 = \det \iota = \det \alpha\alpha^* = \det \alpha \det \alpha^* = \det \alpha (\det \alpha)^* = |\det \alpha|^2$.

The converse is false. If

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

then $\det A = 1$, but

$$A^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \neq A^*.$$

Exercise 8.4.13

If D is a real diagonal matrix with diagonal entries d_j,

$$D \in O(\mathbb{R}^n) \Leftrightarrow DD^T = I \Leftrightarrow d_j^2 = 1 \ \forall j \Leftrightarrow d_j = \pm 1 \ \forall j.$$

If D is a complex diagonal matrix with diagonal entries d_j,

$$D \in U(\mathbb{C}^n) \Leftrightarrow DD^* = I \Leftrightarrow d_j d_j^* = 1 \ \forall j \Leftrightarrow |d_j| = \pm 1 \ \forall j.$$

The diagonal matrix with first diagonal entry $e^{i\theta}$ and all other diagonal entries 1 is unitary with determinant $e^{i\theta}$.

Exercise 8.4.14

$SU(\mathbb{C}^n)$ is a subset of the group $SU(\mathbb{C}^n)$ containing ι.

$$\alpha,\, \beta \in SU(\mathbb{C}^n) \Rightarrow \det \alpha\beta = \det \alpha \det \beta = 1 \times 1 = 1 \Rightarrow \alpha\beta = \in SU(\mathbb{C}^n).$$

$$\alpha \in SU(\mathbb{C}^n) \Rightarrow \det \alpha^{-1} = (\det \alpha)^{-1} = 1 \Rightarrow \alpha^{-1} = \in SU(\mathbb{C}^n).$$

Thus $SU(\mathbb{C}^n)$ is a subgroup of $U(\mathbb{C}^n)$.

Exercise 8.4.15

If $\mathbf{e}_1$, $\mathbf{e}_2$, $\ldots$, $\mathbf{e}_n$ is an orthonormal basis for $\mathbb{C}^n$ and α has matrix A with respect to this basis.

$$\begin{aligned} A = A^* \Leftrightarrow \alpha = \alpha^* &\Leftrightarrow \langle \mathbf{z}, \alpha\mathbf{w}\rangle = \langle \mathbf{z}, \alpha\mathbf{w}\rangle \ \forall \mathbf{w}, \mathbf{z} \in \mathbb{C}^n \\ &\Leftrightarrow \langle \mathbf{z}, \alpha^*\mathbf{w}\rangle = \langle \mathbf{z}, \alpha\mathbf{w}\rangle \ \forall \mathbf{w}, \mathbf{z} \in \mathbb{C}^n \end{aligned}$$

Exercise 8.4.16

If A is Hermitian,

$$\det A = \det A^* = (\det A)^*$$

If

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

then $\det A = 1$ is real, but A is not Hermitian.

Exercise 8.4.17

(i) If $\mathbf{u}$ is an eigenvector with eigenvalue λ then

$$\lambda\|\mathbf{u}\|^2 = \langle\lambda\mathbf{u},\mathbf{u}\rangle = \langle\alpha\mathbf{u},\mathbf{u}\rangle = \langle\mathbf{u},\alpha\mathbf{u}\rangle = \langle\mathbf{u},\lambda\mathbf{u}\rangle = \lambda^*\|\mathbf{u}\|^2$$

so $\lambda = \lambda^*$.

(ii) If $\mathbf{u}$, $\mathbf{v}$ are eigenvectors with distinct eigenvalues λ and μ then

$$\lambda\langle\mathbf{u},\mathbf{v}\rangle = \langle\lambda\mathbf{u},\mathbf{v}\rangle = \langle\alpha\mathbf{u},\mathbf{v}\rangle = \langle\mathbf{u},\alpha\mathbf{v}\rangle = \langle\mathbf{u},\mu\mathbf{v}\rangle = \mu^*\langle\mathbf{u},\mathbf{v}\rangle = \mu\langle\mathbf{u},\mathbf{v}\rangle$$

so $\langle\mathbf{u},\mathbf{v}\rangle = 0$.

Exercise 8.4.18

Part (ii) is equivalent to part (i) by the basis change rule.

We prove part (i) by induction on n.

If $n = 1$, then, since every 1×1 matrix is diagonal, the result is trivial.

Suppose now that the result is true for $n = m$ and that $\alpha : \mathbb{C}^{m+1} \to \mathbb{C}^{m+1}$ is a symmetric linear map. We know that the characteristic polynomial must have a root and that all its roots are real. Thus we can can find an eigenvalue $\lambda_1 \in \mathbb{R}$ and a corresponding eigenvector $\mathbf{e}_1$ of norm 1. Consider the subspace

$$\mathbf{e}_1^{\perp} = \{\mathbf{u} : \langle \mathbf{e}_1, \mathbf{u} \rangle = 0\}.$$

We observe (and this is the key to the proof) that

$$\mathbf{u} \in \mathbf{e}_1^{\perp} \Rightarrow \langle \mathbf{e}_1, \alpha\mathbf{u} \rangle = \langle \alpha\mathbf{e}_1, \mathbf{u} \rangle = \lambda_1 \langle \mathbf{e}_1, \mathbf{u} \rangle = 0 \Rightarrow \alpha\mathbf{u} \in \mathbf{e}_1^{\perp}.$$

Thus we can define $\alpha|_{\mathbf{e}_1^{\perp}} : \mathbf{e}_1^{\perp} \to \mathbf{e}_1^{\perp}$ to be the restriction of α to $\mathbf{e}_1^{\perp}$. We observe that $\alpha|_{\mathbf{e}_1^{\perp}}$ is symmetric and $\mathbf{e}_1^{\perp}$ has dimension m so, by the inductive hypothesis, we can find m orthonormal eigenvectors of $\alpha|_{\mathbf{e}_1^{\perp}}$ in $\mathbf{e}_1^{\perp}$. Let us call them $\mathbf{e}_2$, $\mathbf{e}_3$, $\dots$,$\mathbf{e}_{m+1}$. We observe that $\mathbf{e}_1$, $\mathbf{e}_2$, $\dots$, $\mathbf{e}_{m+1}$ are orthonormal eigenvectors of α and so α is diagonalisable. The induction is complete.

Exercise 8.4.19

We note that

$$\det(tI - A) = (t-5)(t-2) - 4 = t^2 - 7t + 6,$$

so the eigenvalues of A are

$$\frac{7 \pm 5i}{2}.$$

Eigenvectors $(z, w)^T$ corresponding to the eigenvalue $\frac{7+5i}{2}$ are given by

$$\begin{aligned} 10z + 4iw &= (7+5i)z \\ -4iz + 4w &= (7+5i)w, \end{aligned}$$

that is to say

$$\begin{aligned} (3-5i)z + 4iw &= 0 \\ -4iz - (3+5i)w &= 0, \end{aligned}$$

so an eigenvector of norm 1 is given by $2^{-1/2}5^{-1}(4i, 5i-3)^T$.

Eigenvectors $(z, w)^T$ corresponding to the eigenvalue $\frac{7-5i}{2}$ are given by

$$\begin{aligned} 10z + 4iw &= (7-5i)z \\ -4iz + 4w &= (7-5i)w, \end{aligned}$$

that is to say,

$$\begin{aligned} (3+5i)z + 4iw &= 0 \\ -4iz - (3-5i)w &= 0, \end{aligned}$$

so an eigenvector of norm 1 is given by $2^{-1/2}5^{-1}(5i-3, -4i)^T$.

Thus

$$U = 2^{-1/2}5^{-1}\begin{pmatrix} 4i & 5i-3 \\ 5i-3 & -4i \end{pmatrix}$$

is a unitary matrix with

$$U^*AU = \begin{pmatrix} (7+5i)/2 & 0 \\ 0 & (7-5i)/2 \end{pmatrix}$$

Exercise 8.4.20

We can do this using matrices, but it is nicer to use maps.

If γ is unitary, set $\alpha = (\gamma+\gamma^*)/2$, $\beta = (\gamma-\gamma^*)/(2i)$. Then $\gamma = \alpha+i\beta$ and

$$\alpha^* = (\gamma^* + \gamma^{**})/2 = (\gamma^* + \gamma)/2 = \alpha$$
$$\beta^* = -(\gamma^* + \gamma^{**})/(2i) = -(\gamma^* - \gamma)/(2i) = \beta$$

so α and β are Hermitian.

To prove uniqueness, suppose that γ is unitary, α and β Hermitian and $\gamma = \alpha + i\beta$.

Then

$$\alpha - i\beta = \alpha^* - i\beta^* = (\alpha + i\beta)^* = \gamma^* = \gamma$$

so

$$2\alpha = (\alpha - i\beta) + (\alpha + i\beta) = 2\gamma$$

and $\alpha = (\gamma + \gamma^*)/2$, $i\beta = \gamma - \alpha$.

We observe that

$$\alpha\beta = (\gamma + \gamma^*)(\gamma - \gamma^*)/(4i) = (\gamma^2 - 2\iota + (\gamma^*)^2)/(4i)$$
$$(\gamma - \gamma^*)(\gamma + \gamma^*)/(4i) = \beta\alpha$$

and

$$\alpha^2 + \beta^2 = 4^{-1}((\gamma + \gamma^*)^2 - (\gamma - \gamma^*)^2) = 4^{-1}(2\gamma\gamma^* + 2\gamma^*\gamma) = \iota.$$

If we look at 1×1 matrices we get $G = (e^{i\theta})$, $A = (\cos\theta)$, $B = (i\sin\theta)$ and recover Euler's formula.

Exercise 9.1.2

We have
$$LL^T = I \text{ and } \det L = \det \begin{pmatrix} 2^{-1/2} & -2^{-1/2} \\ 2^{-1/2} & 2^{-1/2} \end{pmatrix} = 1$$
so $L \in SO(\mathbb{R}^3)$.

However, if we set $\mathbf{x} = (0,1,0)^T$, we have $\mathbf{x}' = (0, 2^{-1/2}, 2^{-1/2})$, so
$$x_2'^4 = 4^{-1} \neq 2^{-1/2} = l_2 x_2^4 \sum_{j=1}^{3} l_{2j} x_j^4$$

Exercise 9.1.4

(i) We have
$$\dot{u}'_i = \frac{d}{dt}u'_i = \frac{d}{dt}l_{ij}u_j = l_{ij}\dot{u}_j$$
so $\dot{\mathsf{u}}$ is a Cartesian tensor of order 1.

(ii) x is a Cartesian tensor of order 1, so $\dot{\mathsf{x}}$ is, so $\ddot{\mathsf{x}}$ is, so $\mathsf{F} = m\ddot{\mathsf{x}}$ is.

Exercise 9.2.1

(i) Observe that

$$a_{ij} = u_i'v_j' = l_{ir}u_r l_{js}v_s = l_{ir}l_{js}a_{rs}.$$

(ii) Observe that

$$a_{ij}' = \frac{\partial u_j'}{\partial x_i'} = \frac{\partial u_j'}{\partial x_r}\frac{\partial x_r}{\partial x_i'} = l_{ir}\frac{\partial l_{jk}u_k}{\partial x_r} = l_{ir}l_{jk}\frac{\partial u_k}{\partial x_r} = l_{ir}l_{jk}a_{rk}.$$

(iii) Observe that

$$a_{ij}' = \frac{\partial^2 \phi}{\partial x_i' \partial x_j'} = \frac{\partial^2 \phi}{\partial x_r \partial x_s}\frac{\partial x_r}{\partial x_i'}\frac{\partial x_s}{\partial x_j'} = l_{ir}l_{js}a_{rs}.$$

Exercise 9.3.1

(i) If u_i, v_j w_k are Cartesian tensors of order 1 then

$$(u_i v_j w_k)' = u'_i v'_j w'_k = l_{ir} l_{js} l_{kp} u_r v_s w_p$$

so $u_i v_j w_k$ is a Cartesian tensor of order 3.

(ii) If u_{ij} is a Cartesian tensor of order 2 then

$$\frac{\partial u'_{ij}}{\partial x'_k} = \frac{\partial l_{ir} l_{js} u'_{rs}}{\partial x_p} \frac{\partial x_p}{\partial x'_k} = l_{ir} l_{js} l_{kp} \frac{\partial u_{rs}}{\partial x_p}$$

so

$$\frac{\partial u_{ij}}{\partial x_k}$$

is a Cartesian tensor of order 3.

(iii) If ϕ is a smooth tensor of order 0

$$\frac{\partial^2 \phi}{\partial x'_i \partial x'_j \partial x'_k} = \frac{\partial^3 \phi}{\partial x_r \partial x_s \partial x_p} \frac{\partial x_r}{\partial x'_i} \frac{\partial x_s}{\partial x'_j} \frac{\partial x_p}{\partial x'_k} = l_{ir} l_{js} l_{kp} \frac{\partial^3 \phi}{\partial x_r \partial x_s \partial x_p}$$

so

$$\frac{\partial^2 \phi}{\partial x_i \partial x_j \partial x_k}$$

is a Cartesian tensor of order 3.

Exercise 9.3.5

(i) Observe that, in our standard notation

$$l_{ik}a'_i b_k = a'_i b'_i = (a_i b_i)' = a_i b_i = a_k b_k.$$

and so

$$(l_{ik}a'_i - a_k)b_k = 0.$$

Since we can assign b_k any values we please in a particular coordinate system, we must have

$$l_{ik}a'_i - a_k = 0,$$

so

$$\delta_{ir}a'_i - l_{rk}a_k = l_{rk}(l_{ik}a'_i - a_k) = 0$$

and we have shown that $a'_r = l_{rk}a_k$ and a_i is a Cartesian tensor of order 1.

(ii) By (i), $a_{ij}b_i$ is a Cartesian tensor of rank 1 whenever b_i is a Cartesian tensor of order 1 and so, by Theorem 9.3.4, a_{ij} is a Cartesian tensor of order 2.

(iii) Apply (ii) with $u_{ij} = b_i c_j$.

(iv) Observe that, in our standard notation

$$l_{kr}l_{ms}a'_{ijkm}b_{rs} = a'_{ijkm}b'_{km} = (a_{ijkm}b_{km})' = l_{ip}l_{jq}(a_{pqkm}b_{km}) = l_{ip}l_{jq}(a_{pqrs}b_{rs})$$

and so

$$(l_{kr}l_{ms}a'_{ijkm} - l_{ip}l_{jq}a_{pqrs})b_{rs} = 0.$$

Since we can assign b_{rs} any values we please in a particular coordinate system, we must have

$$l_{kr}l_{ms}a'_{ijkm} - l_{ip}l_{jq}a_{pqrs} = 0.$$

so

$$l_{er}l_{fs}(l_{kr}l_{ms}a'_{ijkm} - l_{ip}ljq)a_{pqrs} = 0$$

whence $a'_{ijef} = l_{er}l_{fs}l_{ip}ljqa_{pqrs}$ and we have shown that a_{ijkl} is a Cartesian tensor.

Exercise 9.3.6

(i) We have

$$\begin{aligned} c_{ijkm}e_{km} &= \tfrac{1}{2}(\tilde{c}_{ijkm} + \tilde{c}_{ijmk})e_{km} = \tfrac{1}{2}(\tilde{c}_{ijkm}e_{km} + \tilde{c}_{ijmk}e_{km}) \\ &= \tfrac{1}{2}(\tilde{c}_{ijkm}e_{km} + \tilde{c}_{ijmk}e_{mk}) = p_{ij} \end{aligned}$$

and

$$c_{ijmk} = \tfrac{1}{2}(\tilde{c}_{ijmk} + \tilde{c}_{ijkm}) = \tfrac{1}{2}(\tilde{c}_{ijkm} + \tilde{c}_{ijmk}) = c_{ijkm}.$$

(ii) Set $e_{mk} = \frac{1}{2}(b_{km} + b_{mk})$ and $f_{mk} = \frac{1}{2}(b_{km} - b_{mk})$.

(iii) We have

$$c_{ijmk}f_{mk} = c_{ijkm}f_{km} = -c_{ijmk}f_{mk}$$

so $c_{ijmk}f_{mk} = 0$ and

$$c_{ijkm}e_{km} = c_{ijmk}b_{mk}.$$

Thus $c_{ijmk}b_{mk}$ is a Cartesian tensor of order 2 whenever b_{mk} is, so, by the quotient rule given in Exercise 9.3.5 (iv), c_{ijmk} is a Cartesian tensor of order 4.

Exercise 9.3.8

If we work with a fixed set of coordinates, Cartesian tensors of rank n may be identified with the space of functions

$$f : \mathbb{R}^{3n} \to \mathbb{R}$$

with pointwise multiplication and multiplication by scalars. They therefore form a vector space of dimension 3^n.

EXERCISE 9.4.3

Using the Levi-Civita identity, we have

$$\epsilon_{ijk}\epsilon_{klm}\epsilon_{mni} = \epsilon_{kij}\epsilon_{klm}\epsilon_{mni} = (\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl})\epsilon_{mni} = \epsilon_{jnl}-\epsilon_{mnm} = \epsilon_{jnl} = \epsilon_{ljn}.$$

Exercise 9.4.4 Observe that

$$\epsilon_{ijk\ldots p}\epsilon_{ijk\ldots p} = \sum_{\sigma \in S_n} \zeta(\sigma)^2 = \sum_{\sigma \in S_n} 1 = \operatorname{card} S_n = n!$$

(See the definition of the determinant on page 87 if necessary.)

Exercise 9.4.5

(i) We have

$$\lambda(\epsilon_{ijk}a_j b_k) = \epsilon_{ijk}(\lambda a_j)b_k = \epsilon_{ijk}a_j(\lambda b_k),$$

so

$$\lambda(\mathbf{a}\times\mathbf{b}) = (\lambda\mathbf{a})\times\mathbf{b} = \mathbf{a}\times(\lambda\mathbf{b}).$$

(ii) We have

$$\epsilon_{ijk}a_j(b_k + c_k) = \epsilon_{ijk}a_j b_k + \epsilon_{ijk}a_j c_k,$$

so

$$\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = \mathbf{a}\times\mathbf{b} + \mathbf{a}\times\mathbf{c}.$$

(iii) We have

$$\epsilon_{ijk}a_j b_k = -\epsilon_{ikj}a_j b_k = -\epsilon_{ikj}b_k a_j,$$

so

$$\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}.$$

(iv) Putting $\mathbf{b}=\mathbf{a}$ in (iii), we get

$$\mathbf{a}\times\mathbf{a} = -\mathbf{a}\times\mathbf{a}$$

so $\mathbf{a}\times\mathbf{a} = 0$.

Exercise 9.4.6

There is no identity. For suppose $\mathbf{e}$ was an identity. Then

$$\mathbf{e} = \mathbf{e} \times \mathbf{e} = \mathbf{0},$$

but

$$\mathbf{0} \times \mathbf{a} = \mathbf{0}$$

for all $\mathbf{a}$ so (since $\mathbb{R}^3$ has more than one element) $\mathbf{0}$ is not an identity.

Exercise 9.4.7

(i) We have
$$\lambda(a_j b_j) = (\lambda a_j) b_j = a_j(\lambda b_j),$$
so
$$\lambda(\mathbf{a} \cdot \mathbf{b}) = (\lambda \mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\lambda \mathbf{b}).$$

(ii) We have
$$a_k(b_k + c_k) = a_k b_k + a_k c_k,$$
so
$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}.$$

(iii) We have
$$a_j b_j = b_j a_j,$$
so
$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}.$$

Exercise 9.4.9

(i) We observe that

$$
\begin{aligned}
&\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (a_1,a_2,a_3)\times(b_2c_3-b_3c_2,b_3c_1-b_1c_3,b_1c_2-b_2c_1)\\
&=\big(a_2(b_1c_2-b_2c_1)-a_3(b_3c_1-b_1c_3),a_3(b_2c_3-b_3c_2)-a_1(b_1c_2-b_2c_1),\\
&\qquad\quad a_1(b_3c_1-b_1c_3)-a_2(b_2c_3-b_3c_2)\big)\\
&=\big((a_2c_2+a_3c_3)b_1,(a_3c_3+a_1c_1)b_2,(a_1c_1+a_2c_2)b_3\big)\\
&\qquad\quad-\big((a_2b_2+a_3b_3)c_1,(a_3b_3+a_1b_1)c_2,(a_1b_1+a_2b_2)c_3\big)\\
&=\big((a_2c_2+a_3c_3+a_1c_1)b_1,(a_3c_3+a_1c_1+a_2c_2)b_2,(a_1c_1+a_2c_2+a_3c_3)b_3\big)\\
&-\big((a_2b_2+a_3b_3+a_1b_1)c_1,(a_3b_3+a_1b_1+a_2b_2)c_2,(a_1b_1+a_2b_2+a_3b_3)c_3\big)\\
&=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{a}\cdot\mathbf{b})\mathbf{c}.
\end{aligned}
$$

(ii) ‘Search me, guv’ner!’ There is a geometric proof in the American Mathematical Monthly (Vol 72, Feb 1965) by K. Bishop, but I doubt if the reader will call it simple.

Exercise 9.4.10

We have

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = -\mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = -\big((\mathbf{c} \cdot \mathbf{b})\mathbf{a} - (\mathbf{c} \cdot \mathbf{a})\mathbf{b}\big) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}.$$

as required.

If $\mathbf{x} = (1, 0, 0)$, $\mathbf{y} = \mathbf{z} = (0, 1, 0)$ then

$$\begin{aligned}(\mathbf{x} \times \mathbf{y}) \times \mathbf{z} &= (0, 0, 1) \times (0, 1, 0) = (-1, 0, 0) \neq (0, 0, 0)\\ &= (0, 0, 1) \times (0, 0, 0) = \mathbf{x} \times (\mathbf{y} \times \mathbf{z}).\end{aligned}$$

Exercise 9.4.11

We have

$$\begin{aligned}
\mathbf{a}\times(\mathbf{b}\times\mathbf{c})+\mathbf{b}\times(\mathbf{c}\times\mathbf{a})+\mathbf{c}\times(\mathbf{a}\times\mathbf{b})& \\
&=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{a}\cdot\mathbf{b})\mathbf{c}+(\mathbf{b}\cdot\mathbf{a})\mathbf{c}-(\mathbf{b}\cdot\mathbf{c})\mathbf{a}+(\mathbf{c}\cdot\mathbf{b})\mathbf{a}-(\mathbf{c}\cdot\mathbf{a})\mathbf{b}\\
&=\mathbf{0}.
\end{aligned}$$

Exercise 9.4.12

We have

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{a}\times\mathbf{b})+(\mathbf{a}\cdot\mathbf{b})^2 &= \epsilon_{ijk}a_jb_k\epsilon_{irs}a_rb_s+a_ib_ia_jb_j\\ &= (\delta_{jr}\delta_{ks}-\delta_{js}\delta_{kr})a_jb_ka_rb_s+a_ib_ia_jb_j\\ &= a_ja_jb_kb_k-a_jb_ja_kb_k+a_ib_ia_jb_j\\ &= a_ja_jb_kb_k=\|\mathbf{a}\|^2\|\mathbf{b}\|^2.\end{aligned}$$

If we use the geometric formulae for $\mathbf{a}\times\mathbf{b}$ and $\mathbf{a}\cdot\mathbf{b}$, our formula becomes

$$a^2b^2\cos^2\theta+a^2b^2\sin^2\theta=a^2b^2$$

where θ is the angle between $\mathbf{a}$ and $\mathbf{b}$, $a=\|\mathbf{a}\|$ and $b=\|\mathbf{b}\|$. We thus recover the formula

$$\cos^2\theta+\sin^2\theta=1.$$

Exercise 9.4.13

(i) We have

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = \epsilon_{ijk} a_j b_k a_i = -\epsilon_{jik} a_j b_k a_i = -(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a}$$

so

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = 0$$

(ii) $\mathbf{a} \times \mathbf{b} \perp \mathbf{a}$ so $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = 0$

(iii) The volume of a parallelepiped with one vertex at $\mathbf{0}$ and adjacent vertices at $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{a}$ is 0 so

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = [\mathbf{a}, \mathbf{b}, \mathbf{a}] = 0.$$

Exercise 9.4.14

We have

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})\times(\mathbf{c}\times\mathbf{d}) &= ((\mathbf{a}\times\mathbf{b})\cdot\mathbf{d})\mathbf{c} - ((\mathbf{a}\times\mathbf{b})\cdot\mathbf{c})\mathbf{d}\\ &= [\mathbf{a},\mathbf{b},\mathbf{d}]\mathbf{c} - [\mathbf{a},\mathbf{b},\mathbf{c}]\mathbf{d}.\end{aligned}$$

We also have

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})\times(\mathbf{c}\times\mathbf{d}) &= -(\mathbf{c}\times\mathbf{d})\times(\mathbf{a}\times\mathbf{b})\\ &= (\mathbf{c}\times\mathbf{d})\cdot\mathbf{a}\mathbf{b} - \mathbf{c}\times\mathbf{d})\cdot\mathbf{b}\mathbf{a}\\ &= [\mathbf{c},\mathbf{d},\mathbf{a}]\mathbf{b} - [\mathbf{c},\mathbf{d},\mathbf{b}]\mathbf{a}\\ &= -[\mathbf{c},\mathbf{a},\mathbf{d}]\mathbf{b} - [\mathbf{b},\mathbf{c},\mathbf{d}]\mathbf{a}.\end{aligned}$$

Thus

$$\mathbf{a},\mathbf{b},\mathbf{d}]\mathbf{c} - [\mathbf{a},\mathbf{b},\mathbf{c}]\mathbf{d} = [\mathbf{c},\mathbf{a},\mathbf{d}]\mathbf{b} - [\mathbf{b},\mathbf{c},\mathbf{d}]\mathbf{a}$$

and the required result follows.

Our formulae also give

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})\times(\mathbf{a}\times\mathbf{c}) &= ((\mathbf{a}\times\mathbf{b})\cdot\mathbf{c})\mathbf{a} - ((\mathbf{a}\times\mathbf{b})\cdot\mathbf{a})\mathbf{c}\\ &= [\mathbf{a},\mathbf{b},\mathbf{c}]\mathbf{a} - 0 = [\mathbf{a},\mathbf{b},\mathbf{c}]\mathbf{a}\end{aligned}$$

since $(\mathbf{a}\times\mathbf{b})\perp\mathbf{a}$

Exercise 9.4.15

Since $\mathbf{b} \times \mathbf{c} \perp \mathbf{b}$, $\mathbf{c}$ we have

$$\mathbf{x} \cdot (\mathbf{b} \times \mathbf{c} = (\lambda\mathbf{a} + \mu\mathbf{b} + \nu\mathbf{c}) \cdot (\mathbf{b} \times \mathbf{c}) = \lambda\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$$

so

$$\lambda = \frac{[\mathbf{x}, \mathbf{b}, \mathbf{c}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}.$$

Similarly

$$\mu = \frac{[\mathbf{a}, \mathbf{x}, \mathbf{b}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]} \text{ and } \nu = \frac{[\mathbf{a}, \mathbf{b}, \mathbf{x}]}{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}.$$

Exercise 9.4.16

(i) We work with row vectors

$$
\begin{aligned}
[\mathbf{a},\mathbf{b},\mathbf{c}]^2 &= \det\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\\ \mathbf{c}\end{pmatrix}^2\\
&= \det\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\\ \mathbf{c}\end{pmatrix}\det\begin{pmatrix}\mathbf{a}^T & \mathbf{b}^T & \mathbf{c}^T)\end{pmatrix}\\
&= \det\begin{pmatrix}\mathbf{a}\mathbf{a}^T & \mathbf{a}\mathbf{b}^T & \mathbf{a}\mathbf{c}^T\\ \mathbf{b}\mathbf{a}^T & \mathbf{b}\mathbf{b}^T & \mathbf{b}\mathbf{c}^T\\ \mathbf{c}\mathbf{a}^T & \mathbf{c}\mathbf{b}^T & \mathbf{c}\mathbf{c}^T\end{pmatrix}\\
&= \det\begin{pmatrix}\mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b} & \mathbf{a}\cdot\mathbf{c}\\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b} & \mathbf{b}\cdot\mathbf{c}\\ \mathbf{c}\cdot\mathbf{a} & \mathbf{c}\cdot\mathbf{b} & \mathbf{c}\cdot\mathbf{c}\end{pmatrix}
\end{aligned}
$$

(ii) In the case given we first observe that

$$r^2 = \|\mathbf{a}+\mathbf{b}+\mathbf{c}\|^2 = \|\mathbf{a}\|^2+\|\mathbf{b}\|^2+\|\mathbf{c}\|^2+2(\mathbf{a}\cdot\mathbf{b}+\mathbf{b}\cdot\mathbf{c}+\mathbf{a}\cdot\mathbf{b}$$

so

$$\mathbf{a}\cdot\mathbf{b}+\mathbf{b}\cdot\mathbf{c}+\mathbf{a}\cdot\mathbf{b} = -r^2$$

$$
\begin{aligned}
[\mathbf{a},\mathbf{b},\mathbf{c}]^2 &= \det\begin{pmatrix}\mathbf{a}\cdot\mathbf{a} & \mathbf{a}\cdot\mathbf{b} & \mathbf{a}\cdot\mathbf{c}\\ \mathbf{b}\cdot\mathbf{a} & \mathbf{b}\cdot\mathbf{b} & \mathbf{b}\cdot\mathbf{c}\\ \mathbf{c}\cdot\mathbf{a} & \mathbf{c}\cdot\mathbf{b} & \mathbf{c}\cdot\mathbf{c}\end{pmatrix}\\
&= \det\begin{pmatrix}r^2 & \mathbf{a}\cdot\mathbf{b} & \mathbf{a}\cdot\mathbf{c}\\ \mathbf{b}\cdot\mathbf{a} & r^2 & \mathbf{b}\cdot\mathbf{c}\\ \mathbf{c}\cdot\mathbf{a} & \mathbf{c}\cdot\mathbf{b} & r^2\end{pmatrix}\\
&= r^2(r^4-(\mathbf{b}\cdot\mathbf{c})^2)-\mathbf{a}\cdot\mathbf{b}(r^2\mathbf{a}\cdot\mathbf{b}-(\mathbf{b}\cdot\mathbf{c})(\mathbf{a}\cdot\mathbf{c}))\\
&\quad+\mathbf{a}\cdot\mathbf{c}((\mathbf{b}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{c})-r^2\mathbf{b}\cdot\mathbf{a})\\
&= 2r^6+2r^4(\mathbf{a}\cdot\mathbf{b}+\mathbf{b}\cdot\mathbf{c}+\mathbf{a}\cdot\mathbf{b})\\
&\quad+r^2(\mathbf{a}\cdot\mathbf{b}+\mathbf{b}\cdot\mathbf{c}+\mathbf{a}\cdot\mathbf{b})^2-r^2((\mathbf{a}\cdot\mathbf{b})^2+(\mathbf{b}\cdot\mathbf{c})^2+(\mathbf{c}\cdot\mathbf{a})^2)\\
&\quad+2(\mathbf{a}\cdot\mathbf{b})(\mathbf{b}\cdot\mathbf{c})(\mathbf{c}\cdot\mathbf{a})\\
&= 2(r^2+\mathbf{a}\cdot\mathbf{b})(r^2+\mathbf{b}\cdot\mathbf{c})(r^2+\mathbf{c}\cdot\mathbf{a}).
\end{aligned}
$$

Exercise 9.4.17

Observe, using the summation convention, that

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})\cdot(\mathbf{c}\times\mathbf{d}) &= \epsilon_{ijk}a_jb_k\epsilon_{irs}c_rd_s = (\delta_{jr}\delta_{ks}-\delta_{kr}\delta_{js})a_jb_kc_rd_s\\ &= a_jc_jb_kd_k - a_jd_jb_kc_k = (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d})-(\mathbf{a}\cdot\mathbf{d})(\mathbf{c}\cdot\mathbf{d})\end{aligned}$$

(Or we could use the formula for the triple vector product.)

Thus

$$\begin{aligned}(\mathbf{a}\times\mathbf{b})&\cdot(\mathbf{c}\times\mathbf{d})+(\mathbf{a}\times\mathbf{c})\cdot(\mathbf{d}\times\mathbf{b})+(\mathbf{a}\times\mathbf{d})\cdot(\mathbf{b}\times\mathbf{c})\\ &= (\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d})-(\mathbf{a}\cdot\mathbf{d})(\mathbf{c}\cdot\mathbf{d})\\ &\quad+(\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c})-(\mathbf{a}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{d})\\ &\quad+(\mathbf{a}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{d})-(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d})\\ &= 0.\end{aligned}$$

Exercise 9.4.19

(i) Using the summation convention, we have

$$\left(\frac{d}{dt}\phi\mathbf{a}\right)_i = \frac{d}{dt}(\phi a_i) = \dot{\phi}a_i + \phi\dot{a}_i = \left(\dot{\phi}\mathbf{a} + \phi\dot{\mathbf{a}}\right)_i.$$

(ii) Using the summation convention, we have

$$\frac{d}{dt}\mathbf{a}\cdot\mathbf{b} = \frac{d}{dt}a_ib_i = \dot{a}_ib_i + a_i\dot{b}_i = \dot{\mathbf{a}}\cdot\mathbf{b} + \mathbf{a}\cdot\dot{\mathbf{b}}.$$

(iii) By Lemma 9.4.18,

$$\frac{d}{dt}\mathbf{a}\times\dot{\mathbf{a}} = \dot{\mathbf{a}}\times\dot{\mathbf{a}} + \mathbf{a}\times\ddot{\mathbf{a}} = \mathbf{a}\times\ddot{\mathbf{a}}.$$

Exercise 9.4.21

(i) Using the summation convention, we have

$$\begin{aligned}\big(\nabla\times(\phi\mathbf{u})\big)_i &= \epsilon_{ijk}\frac{d}{dx_j}\phi u_k\\ &= \epsilon_{ijk}\frac{d\phi}{dx_j}u_k+\epsilon_{ijk}\frac{du_k}{dx_j}\phi \quad = \big((\nabla\phi)\times\mathbf{u}+\phi\nabla\times\mathbf{u}\big)_i.\end{aligned}$$

(ii) Using the summation convention, we have

$$\nabla\cdot(\nabla\times\mathbf{u})=\frac{d}{dx_i}\epsilon_{ijk}\frac{du_k}{dx_j}=\epsilon_{ijk}\frac{d^2u_k}{dx_idx_j}=\epsilon_{ijk}\frac{d^2u_k}{dx_jdx_i}=-\nabla\cdot(\nabla\times\mathbf{u})$$

so

$$\nabla\cdot(\nabla\times\mathbf{u})=0.$$

(ii) Using the summation convention, we have

$$\begin{aligned}\big(\nabla\times(\mathbf{u}\times\mathbf{v})\big)_i &= \epsilon_{ijk}\frac{d}{dx_j}\epsilon_{krs}u_rv_s\\ &= \epsilon_{kij}\epsilon_{krs}\frac{d}{dx_j}u_rv_s\\ &= (\delta_{ir}\delta_{js}-\delta_{is}\delta_{jr})\frac{d}{dx_j}u_rv_s\\ &= \frac{d}{dx_j}u_iv_j-\frac{d}{dx_j}u_jv_i\\ &= v_j\frac{du_i}{dx_j}+u_i\frac{dv_j}{dx_j}-v_i\frac{du_j}{dx_j}-u_j\frac{dv_i}{dx_j}\\ &= \big((\nabla\cdot\mathbf{u})\mathbf{v}+\mathbf{u}\cdot\nabla\mathbf{v}-(\nabla\cdot\mathbf{v})\mathbf{u}-\mathbf{v}\cdot\nabla\mathbf{u}\big)_i.\end{aligned}$$

Exercise 9.4.22

(i) Observe that the cancellation law continues to hold, so it is sufficient to show that the analogue of the scalar vector product (think of the 4 dimensional volume) is a scalar.

Let $a_i = \epsilon_{ijkl}x_j y_k z_l$. Then if b_i is an vector (ie order 1 Cartesian tensor),

$$\begin{aligned}(a_i b_i)' &= \epsilon_{ijkl}x_j' y_k' z_l' b_i' = \epsilon_{ijkl}l_{ju}l_{kv}l_{lv}l_{ip}x_u y_v z_w b_p\\ &= (\det L)\epsilon_{puvw}x_u y_v z_w b_p\\ &= \epsilon_{puvw}x_u y_v z_w b_p = a_i b_i\end{aligned}$$

since $L \in SO(\mathbb{R}^4)$.

Thus, by the cancellation law, $\mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z}$ is a vector

(ii) We have

$$\begin{aligned}\big((\lambda\mathbf{x} + \mu\mathbf{u}) \wedge \mathbf{y} \wedge \mathbf{z}\big)_i &= \epsilon_{ijkl}(\lambda x_j + \mu u_j)y_k z_l\\ &= \lambda\epsilon_{ijkl}x_j y_k z_l + \mu\epsilon_{ijkl}u_j y_k z_l\\ &= (\lambda\mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z} + \mu\mathbf{u} \wedge \mathbf{y} \wedge \mathbf{z})_i.\end{aligned}$$

$$(\mathbf{y} \wedge \mathbf{x} \wedge \mathbf{z})_i = \epsilon_{ijkl}y_j x_k z_l = -\epsilon_{ikjl}x_k y_j z_l = (\mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z})_i$$

so

$$\mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z} = -\mathbf{y} \wedge \mathbf{x} \wedge \mathbf{z}.$$

The proof that

$$\mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z} = -\mathbf{z} \wedge \mathbf{y} \wedge \mathbf{x}$$

follows the same lines.

(iii) Using the summation convention with range 1, 2, 3, 4, 5, we define

$$\mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z} \wedge \mathbf{w} = \mathbf{a}$$

with

$$a_i = \epsilon_{ijklu}x_j y_k z_l w_u.$$

EXERCISE 10.1.4

Since the material is isotropic so is a_{ijk} thus $a_{ijk} = \kappa\epsilon_{ijk}$ and

$$\sigma_{ij} = a_{ijk}B_k = \kappa\epsilon_{ijk}B_k = -\kappa\epsilon_{jik}B_k = -\sigma_{ji} = -\sigma_{ij}.$$

Thus $\sigma_{ij} = 0$.

Exercise 10.1.2

It is sufficient to show that $\delta_{ij}\delta_{kl}$ is isotropic, since it then follows that $\delta_{ik}\delta_{jl}$ and $\delta_{il}\delta_{jk}$ and all linear combinations are.

But $(\delta_{ij}\delta_{kl})' = \delta'_{ij}\delta'_{kl} = \delta_{ij}\delta_{kl}$ so we are done.

If
$$\alpha\delta_{ij}\delta_{kl} + \beta\delta_{ik}\delta_{jl} + \gamma\delta_{il}\delta_{jk} = 0,$$
then
$$0 = \alpha\delta_{ij}\delta_{kk} + \beta\delta_{ik}\delta_{jk} + \gamma\delta_{ik}\delta jk = (3\alpha + \beta + \gamma)\delta_{ij}.$$
Thus $3\alpha + \beta + \gamma = 0$ and similarly $\alpha + 3\beta + \gamma = 0$, $\alpha + \beta + 3\gamma = 0$. Thus
$$2\alpha = 2\beta = 2\gamma = -(\alpha + \beta + \gamma)$$
and $\alpha = \beta = \gamma = 0$. (Alternatively, just examine particular entries in the tensor $\alpha\delta_{ij}\delta_{kl} + \beta\delta_{ik}\delta_{jl} + \gamma\delta_{il}\delta_{jk}$.)

For the last part, observe that, since ϵ_{ijk} and ϵ_{ist} are isotropic, so is $\epsilon_{ijk}\epsilon_{ist}$. Thus
$$\epsilon_{ijk}\epsilon_{ist} = \alpha\delta_{jk}\delta_{st} + \beta\delta_{js}\delta_{kt} + \gamma\delta_{jt}\delta ks.$$

We observe that
$$\alpha\delta_{jk}\delta_{st}+\beta\delta_{js}\delta_{kt}+\gamma\delta_{jt}\delta ks = \epsilon_{ijk}\epsilon_{ist} = -\epsilon_{ikj}\epsilon_{ist} = -\alpha\delta_{jk}\delta_{st}+\gamma\delta_{js}\delta_{kt}+\beta\delta_{jt}\delta ks$$
so, by the linear independence proved earlier, $\beta = -\gamma$.

Next we observe that
$$0 = \epsilon_{ijj}\epsilon_{ist} = (3\alpha + \beta + \gamma)\delta_{st}$$
so $3\alpha + \beta + \gamma = 0$ and $\alpha = 0$ Finally
$$6 = \epsilon_{ijk}\epsilon_{ijk} = 3\alpha + 9\beta + 3\gamma$$
so $\alpha = 0$, $\beta = 1$ and $\gamma = -1$. We have shown that
$$\epsilon_{ijk}\epsilon_{ist} = \delta_{js}\delta_{kt} - \delta_{ks}\delta_{jt}.$$

Exercise 10.1.3

$$a'_{ij} = l_{ir}l_{js}a_{rs} = l_{ir}l_{js}a_{sr} = a'_{ji}.$$

Exercise 10.1.6

(i) We have

$$a'_{ij} = l_{ir}l_{js}a_{rs} = -l_{ir}l_{js}a_{sr} = -a'_{ji}.$$

(ii) We have

$$b_{ij} = \tfrac{1}{2}(b_{ij} + b_{ji}) + \tfrac{1}{2}(b_{ij} - b_{ji})$$

and

$$\begin{aligned}\tfrac{1}{2}(b_{ji} + b_{ij}) &= \tfrac{1}{2}(b_{ij} + b_{ji})\\ \tfrac{1}{2}(b_{ji} - b_{ij}) &= -\tfrac{1}{2}(b_{ij} - b_{ji}).\end{aligned}$$

(iii) Observe that the zero order two tensor is both symmetric and antisymmetric. Further

$$a_{ij} = a_{ji},\ b_{ij} = b_{ji} \Rightarrow \lambda a_{ij} + \mu b_{ij} = \lambda a_{ji} + \mu b_{ji}$$

and

$$a_{ij} = -a_{ji},\ b_{ij} = -b_{ji} \Rightarrow \lambda a_{ij} + \mu b_{ij} = -(\lambda a_{ji} + \mu b_{ji})$$

Thus the symmetric tensors form a subspace and the antisymmetric tensors form a subspace.

We do not use the summation convention for α and β.

Let $\mathsf{e}(\alpha, \beta)$ be the second order Cartesian tensor with

$$e_{ij}(\alpha, \beta) = \delta_{\alpha i}\delta_{\beta j} + \delta_{\alpha i}\delta_{\beta j}$$

in a particular coordinate system S $[1 \leq \alpha < \beta \leq 3]$ and

$$e_{ij}(\alpha, \alpha) = \delta_{\alpha i}\delta_{\alpha j}$$

for $1 \leq \alpha \leq 3$

Let $\mathsf{f}(\alpha, \beta)$ be the second orderCartesian tensor with

$$f_{ij}(\alpha, \beta) = \delta_{\alpha i}\delta_{\beta j} - \delta_{\alpha i}\delta_{\beta j}$$

in the same coordinate system S $[1 \leq \alpha < \beta \leq 3]$.

Then in that coordinate system $\mathsf{e}(\alpha, \beta)$ is symmetric, $\mathsf{f}(\alpha, \beta)$ is antisymmetric. If a is a second order symmetric tensor then in the coordinate system S,

$$\mathsf{a} = \sum_{1 \leq \alpha \leq \beta \leq 3} a\alpha\beta\mathsf{e}(\alpha, \beta)$$

and if b is an antisymmetric tensor then in the coordinate system S,

$$\mathsf{b} = \sum_{1 \leq \alpha < \beta \leq 3} b_{\alpha\beta}\mathsf{f}(\alpha, \beta).$$

Since the result is true in one system it is true in all.

Working in S we see that

$$\sum_{1\leq\alpha\leq\beta\leq 3} a_{\alpha\beta}\mathsf{e}(\alpha,\beta) = 0 \Rightarrow a_{\alpha\beta} = 0 \ \ [1 \leq \alpha \leq \beta \leq 3]$$

and

$$\sum_{1\leq\alpha<\beta\leq 3} b_{\alpha\beta}\mathsf{e}(\alpha,\beta) = 0 \Rightarrow b_{\alpha\beta} = 0 \ \ [1 \leq \alpha < \beta \leq 3].$$

Thus the $\mathsf{e}(\alpha,\beta)$ give a basis for the subspace of symmetric tensors which thus has dimension 6 the $\mathsf{f}(\alpha,\beta)$ give a basis for the subspace of antisymmetric tensors which thus has dimension 3.

Exercise 10.1.8

Since
$$\epsilon_{ijk}\omega_j x_k = (\boldsymbol{\omega} \times \mathbf{x})_i$$
and $\omega \times \mathbf{x} \perp \mathbf{x}$, the only possible eigenvalue is 0 and this corresponds to eigenvectors $t\boldsymbol{\omega}$ with $t \neq 0$.

EXERCISE 10.2.1

Observe first that

$$\sum_\alpha m_\alpha \mathbf{r}_\alpha = \sum_\alpha m_\alpha \mathbf{x}_\alpha - M\mathbf{x}_G = \mathbf{0}.$$

We have

$$\begin{aligned}
\mathbf{H} &= \sum_\alpha m_\alpha \mathbf{x}_\alpha \times \dot{\mathbf{x}}_\alpha \\
&= \sum_\alpha m_\alpha (\mathbf{r}_\alpha + \mathbf{x}_G) \times (\dot{\mathbf{r}}_\alpha + \dot{\mathbf{x}}_G) \\
&= \sum_\alpha m_\alpha \mathbf{r}_\alpha \times \dot{\mathbf{r}}_\alpha + \mathbf{x}_G \times \sum_\alpha m_\alpha \dot{\mathbf{r}}_\alpha + \sum_\alpha m_\alpha \mathbf{r}_\alpha \times \dot{\mathbf{r}}_\alpha \\
&\qquad\qquad + \left(\sum_\alpha m_\alpha \mathbf{r}_\alpha\right) \times \mathbf{x}_G \\
&= \sum_\alpha m_\alpha \mathbf{r}_\alpha \times \dot{\mathbf{r}}_\alpha + \mathbf{x}_G \times \frac{d}{dt}\sum_\alpha m_\alpha \mathbf{r}_\alpha + \sum_\alpha m_\alpha \mathbf{r}_\alpha \times \dot{\mathbf{r}}_\alpha + \mathbf{0} \times \mathbf{x}_G \\
&= \sum_\alpha m_\alpha \mathbf{r}_\alpha \times \dot{\mathbf{r}}_\alpha + \mathbf{x}_G \times \frac{d}{dt}\mathbf{0} + \sum_\alpha m_\alpha \mathbf{r}_\alpha \\
&= M\mathbf{x}_G \times \dot{\mathbf{x}}_G + \sum_\alpha m_\alpha \mathbf{r}_\alpha \times \dot{\mathbf{r}}_\alpha
\end{aligned}$$

The total angular momentum about the origin is the angular momentum of the system around the centre of mass plus the angular momentum of a point mass equal to the mass of the system at the centre of gravity around the origin.

Exercise 10.2.2

We have

$$
\begin{aligned}
\dot{\mathbf{H}} &= \frac{d}{dt}\sum_\alpha m_\alpha \mathbf{x}_\alpha \times \dot{\mathbf{x}}_\alpha \\
&= \sum_\alpha m_\alpha(\dot{\mathbf{x}}_\alpha \times \dot{\mathbf{x}}_\alpha + \mathbf{x}_\alpha \times \ddot{\mathbf{x}}_\alpha) \\
&= \sum_\alpha m_\alpha \mathbf{x}_\alpha \times \ddot{\mathbf{x}}_\alpha \\
&= \sum_\alpha \mathbf{x}_\alpha \times m_\alpha \ddot{\mathbf{x}}_\alpha \\
&= \sum_\alpha \mathbf{x}_\alpha \times \mathbf{F}_\alpha \\
&= -k\sum_\alpha m_\alpha \mathbf{x}_\alpha \times \dot{\mathbf{x}}_\alpha \\
&= -k\mathbf{H}
\end{aligned}
$$

Thus (you can work with the summation convention, if you prefer)

$$\frac{d}{dt}e^{kt}\mathbf{H}(t) = 0$$

so $\mathbf{H}(t)e^{kt}$ is a constant vector $\mathbf{H}_0$ and we are done.

Exercise 10.2.3

Since the first lump has centre of mass at $\mathbf{0}$ we have

$$\int x_i\rho(\mathbf{x})\,dV(\mathbf{x}) = 0$$

By a simple change of variable

$$\begin{aligned}
J_{ij} &= \int \big((x_k+a_k)(x_k+a_k)\delta_{ij} - (x_i+a_i)(x_j+a_j)\big)\rho(\mathbf{x})\,dV(\mathbf{x})\\
&= \int (x_kx_k\delta_{ij} - x_ix_j)\rho(\mathbf{x})\,dV(\mathbf{x}) + 2a_k\delta_{ij}\int x_k\rho(\mathbf{x})\,dV(\mathbf{x})\\
&\qquad + (a_ka_k\delta_{ij} - a_ia_j)\int \rho(\mathbf{x})\,dV(\mathbf{x}) + a_i\int x_j\rho(\mathbf{x})\,dV(\mathbf{x})\\
&\qquad + a_j\int x_i\rho(\mathbf{x})\,dV(\mathbf{x})\\
&= I_{ij} + M(a_ka_k\delta_{ij} - a_ia_j).
\end{aligned}$$

Exercise 10.2.4

In the coordinate system chosen, taking $x_1 = x$, $x_2 = y$, $x_3 = z$,

$$\begin{aligned} A &= I_{11} \\ &= \int (x_k x_k \delta_{11} - x_1 x_1)\rho(\mathbf{x})\, dV(\mathbf{x}) \\ &= \int (x_2 x_2 + x_3 x_3)\rho(\mathbf{x})\, dV(\mathbf{x}) \\ &= \iiint (y^2 + z^2)\rho(x, y, z)\, dx\, dy\, dz \geq 0 \end{aligned}$$

We also have

$$\begin{aligned} B &= \iiint (z^2 + x^2)\rho(x, y, z)\, dx\, dy\, dz \geq 0, \\ C &= \iiint (x^2 + y^2)\rho(x, y, z)\, dx\, dy\, dz \geq 0. \end{aligned}$$

Exercise 10.3.2

(i) We have

$$l_{ir}l_{js}\delta_{rs} = l_{ir}l_{jr} = \delta_{ij}$$

since $LL^T = I$.

(ii) If a_{ijk} is an isotropic extended Cartesian tensor, then a_{ijk} is an isotropic Cartesian tensor so $a_{ijk} = A\epsilon_{ijk}$ so $a_{ijk} = 0$.

(iii) We have

$$(\epsilon_{rst}a_r b_s c_t)' = l_{ri}l_{sj}l_{tk}\epsilon_{rst}a_i b_j c_k = (\det L)\epsilon_{ijk}a_i b_j c_k = -\epsilon_{ijk}a_i b_j c_k.$$

(iv) We have

$$(a_r b_r)' = a'_r b'_r = l_{ri}l_{rj}a_i b_j = \delta_{ij}a_i b_j.$$

(v) If $\mathbf{a} \times \mathbf{b}$ was an extended Cartesian tensor then by (iv) so would be $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$. Now choose $\mathbf{c}$ non-zero and orthogonal to $\mathbf{a}$ and $\mathbf{b}$.

Exercise 10.4.1

(i) We have

$$(\lambda u^i + \mu v^i)^- = (\lambda \bar{u}^i + \mu \bar{v}^i) = (\lambda l^i_j u^j + \mu l^i_j v^j) = l^i_j(\lambda u^j + \mu v^j).$$

(ii) We have

$$(\bar{u}')^i = (l^i_j u^i)' = l^i_j \dot{u}^j.$$

Exercise 10.4.3

$$m_j^i = l_j^i \Leftrightarrow L = (L^T)^{-1} \Leftrightarrow L^T = L^{-1} \Leftrightarrow L \in O(\mathbb{R}^3).$$

Exercise 10.4.6

(i) We have
$$l_u^i m_j^v \delta_i^j = l_u^i m_i^v = \delta_u^v = \bar{\delta}_u^v.$$

Choose an L such that $L^2 \neq I$, for example
$$L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{5} & -\frac{3}{5} \\ 0 & \frac{3}{5} & \frac{4}{5} \end{pmatrix}.$$

Then
$$l_i^u l_j^v \delta_{uv} = l_i^u l_u^v \neq \delta_{ij} = \bar{\delta}_{ij}.$$

(ii) The rule is
$$\bar{a}_{ij}^{kn} = l_i^u l_j^v m_p^k m_q^n a_{uv}^{pq}.$$

(iii) We have
$$l_i^u l_j^v m_p^k m_q^n a_{uv}^{pq} = l_i^u l_j^v m_p^k m_q^n \delta_i^k \delta_j^n = l_i^u l_j^v m_p^i m_q^j = l_i^u m_p^i l_j^v m_q^j = \delta_p^u \delta_q^v = \bar{a}_{pq}^{uv},$$
so a_{ij}^{kn} is a 4th rank tensor.

Next observe that
$$a_{ij}^{in} = \delta_i^i \delta_j^n = 3\delta_j^n$$
a second rank tensor.

However
$$a_{ii}^{kn} = \delta_i^k \delta_i^n = \delta_{kn}$$
is not a second rank tensor . The same argument shows that a_{ij}^{nn} is not.

(iv) The rule is
$$\bar{a}_{ijk}^{n} = l_i^u l_j^v l_k^p m_q^n a_{uvp}^{q}.$$

We observe that
$$\bar{a}_{ijn}^{n} = l_i^u l_j^v l_n^p m_q^n a_{uvp}^{q} = l_i^u l_j^v \delta_q^p a_{uvp}^{q} = l_i^u l_j^v a_{uvp}^{p}$$
so a_{ijn}^{n} is a second rank tensor.

Exercise 11.1.2

Observe that

$$(\alpha+\beta)(\mathbf{e}_j)=\alpha(\mathbf{e}_j)+\beta(\mathbf{e}_j)=\sum_{i=1}^{n}a_{ij}\mathbf{f}_i+\sum_{i=1}^{n}b_{ij}\mathbf{f}_i=\sum_{i=1}^{n}(a_{ij}+b_{ij})\mathbf{f}_i$$

and

$$\lambda\alpha(\mathbf{e}_j)=\lambda\sum_{i=1}^{n}a_{ij}\mathbf{f}_i=\sum_{i=1}^{n}\lambda a_{ij}\mathbf{f}_i.$$

Exercise 11.1.3

If

$$\alpha(\mathbf{f}_j) = \sum_{i=1}^{p} a_{ij}\mathbf{g}_i$$

$$\beta(\mathbf{e}_r) = \sum_{s=1}^{n} b_{sr}\mathbf{f}_s,$$

then

$$\alpha\beta(\mathbf{e}_r) = \sum_{s=1}^{n} b_{sr}\alpha\mathbf{f}_s = \sum_{s=1}^{n} b_{sr}\sum_{i=1}^{p} a_{is}\mathbf{g}_i = \sum_{i=1}^{p}\left(\sum_{s=1}^{n} a_{is}b_{sr}\right)\mathbf{g}_i,$$

so $\alpha\beta$ has matrix AB with respect to the given bases.

Exercise 11.1.4

(i) Observe that, by definition,

$$(\alpha+\beta)\gamma\mathbf{u} = \alpha(\gamma\mathbf{u}) + \beta(\gamma\mathbf{u}) = (\alpha\gamma)\mathbf{u} + (\beta\gamma)\mathbf{u} = (\alpha\beta+\alpha\gamma)\mathbf{u}$$

for all $\mathbf{u} \in U$ so, by definition

$$(\alpha+\beta)\gamma = \alpha\gamma + \beta\gamma.$$

(ii) and (iii) follow at once from Exercises 11.1.2 and 11.1.3.

(iv) Let U, V and W be (not necessarily finite dimensional) vector spaces over $\mathbb{F}$ and let α, $\beta \in \mathcal{L}(U,V)$, $\gamma \in \mathcal{L}(V,W)$. Then

$$\gamma(\alpha+\beta) = \gamma\alpha + \gamma\beta.$$

If A and B are $n \times m$ matrices over $\mathbb{F}$ and C is an $p \times n$ matrix over $\mathbb{F}$, we have

$$C(A+B) = CA + CB.$$

Exercise 11.1.5

Since the $\mathbf{f}_i$ are a basis for V, we can find $p_{ij} \in \mathbb{F}$ such that

$$\mathbf{f}'_j = \sum_{i=1}^{n} p_{ij}\mathbf{f}_i \ [1 \leq j \leq n]$$

and, similarly, we can find m_{ij} and q_{rs} such that

$$\mathbf{f}_j = \sum_{i=1}^{n} m_{ij}\mathbf{f}_i \ [1 \leq j \leq n]$$

and

$$\mathbf{e}'_s = \sum_{r=1}^{m} q_{rs}\mathbf{e}_r \quad [1 \leq s \leq m].$$

We have

$$\mathbf{f}_j = \sum_{i=1}^{n} m_{ij}\mathbf{f}'_i = \sum_{i=1}^{n} m_{ij} \sum_{k=1}^{n} p_{ki}\mathbf{f}_k$$

so by uniqueness

$$\delta_{kj} = \sum_{i=1}^{n} p_{ki}m_{ij}$$

that is to say $PM = I$ so P is invertible with $M = P^{-1}$.

Similarly

$$\begin{aligned}
\alpha\mathbf{e}'_s &= \alpha \sum_{r=1}^{m} q_{rs}\mathbf{e}_r \\
&= \sum_{r=1}^{m} q_{rs}\alpha\mathbf{e}_r \\
&= \sum_{r=1}^{m} q_{rs} \sum_{p=1}^{n} a_{pr}\mathbf{f}'_p \\
&= \sum_{r=1}^{m} q_{rs} \sum_{p=1}^{n} a_{pr} \sum_{i=1}^{n} m_{ip}\mathbf{f}'_i \\
&\sum_{i=1}^{m} \left(\sum_{p=1}^{n} \sum_{r=1}^{m} m_{ip}a_{pr}q_{rs} \right) \mathbf{f}'_p,
\end{aligned}$$

so

$$B = MAQ = P^{-1}AQ.$$

Exercise 11.1.9

(i) Since $I \in G$,

$$A \in X \Rightarrow A,\, A \in X, A = AI \Rightarrow A \sim_1 A.$$

If $A \sim_1 B$ then A, $B \in X$ and we can find a $P \in G$ such that $B = PA$. But $P^{-1} \in G$ and $A = P^{-1}B$ so $B \sim_1 A$.

If $A \sim_1 B$, $B \sim_1 C$, then A, B, $C \in X$ and we can find a P, $Q \in G$ such that $B = PA$, $C = QB$. Since $QP \in G$ and $C = (QP)A$ we have $A \sim_1 C$. Thus $\sim_1$ is an equivalence relation on X.

(ii) Since $I \in G$,

$$A \in X \Rightarrow A,\, A \in X, A = I^{-1}AI \Rightarrow A \sim_2 A.$$

If $A \sim_2 B$, then A, $B \in X$ and we can find P, $Q \in G$ such that $B = P^{-1}AQ$. But P^{-1}, $Q \in G$ and $A = (P^{-1})^{-1}BQ^{-1}$, so $B \sim_2 A$.

If $A \sim_2 B$, $B \sim_2 C$, then A, B, $C \in X$ and we can find P, Q, R, $S \in G$ such that $B = P^{-1}AQ$, $C = R^{-1}BS$. Since PR, $SQ \in G$ and $C = (PR)^{-1}A(SQ)$, we have $A \sim_2 C$. Thus $\sim_2$ is an equivalence relation on X.

(iii) If $A \sim_2 B$, then A, $B \in X$ and we can find P, $Q \in GL(\mathbb{F}^n)$ with $B = P^{-1}AQ$ and so $\operatorname{rank} B = \operatorname{rank} A$. If $\operatorname{rank} B = \operatorname{rank} A = k$, then we can find P, Q, R, $S \in GL(\mathbb{F}^n)$ such that $D = P^{-1}AQ$ and $D = R^{-1}BS$ where D is the $n \times n$ diagonal matrix with first k diagonal entries 1 and remaining entries 0 so $A \sim_2 D$, $D \sim_2 B$ and $A \sim_2 B$. Thus there are $n+1$ equivalence classes corresponding to the matrices of rank k $[0 \leq k \leq n]$.

(iv) Since $I \in G$.

$$A \in X \Rightarrow A,\, A \in X, A = I^{-1}AI \Rightarrow A \sim_3 A.$$

If $A \sim_3 B$ then A, $B \in X$ and we can find $P \in G$ such that $B = P^{-1}AP$. But $P^{-1} \in G$ and $A = (P^{-1})^{-1}BP^{-1}$ so $B \sim_3 A$.

If $A \sim_3 B$, $B \sim_3 C$, then A, B, $C \in X$ and we can find P, $Q \in G$ such that $B = P^{-1}AP$, $C = Q^{-1}BQ$. Since $PQ \in G$ and $C = (PQ)^{-1}A(PQ)$ we have $A \sim_3 C$. Thus $\sim_3$ is an equivalence relation on X.

(v) Since A and $P^{-1}AP$ have the same characteristic polynomial

$$A \sim_3 B \Rightarrow \chi_A = \chi_B$$

Since whenever A is symmetric there exists an orthogonal matrix P such that $P^{-1}AP = P^TAP = D$ where D is a diagonal matrix with the same characteristic polynomial

$$\chi_A = \chi_B \Rightarrow A \sim_3 B.$$

Thus

$$A \sim_3 B \Leftrightarrow \chi_A = \chi_B.$$

If D_λ is the diagonal matrix with first diagonal entry λ and all others 0 then

$$D_\lambda \sim_3 D_\mu \Leftrightarrow \chi_{D_\lambda} = \chi_{D_\mu} \Leftrightarrow \lambda = \mu.$$

Thus there are infinitely many equivalence classes for $\sim_3$.

(vi) Since $I \in G$

$$A \in X \Rightarrow A,\, A \in X, A = I^T AI \Rightarrow A \sim_3 A.$$

If $A \sim_4 B$, then A, $B \in X$ and we can find $P \in G$ such that $B = P^T AP$. But $P^{-1} \in G$ and $A = (P^{-1})^T BP^{-1}$, so $B \sim_4 A$.

If $A \sim_4 B$, $B \sim_4 C$, then A, B, $C \in X$ and we can find P, $Q \in G$ such that $B = P^T AP$, $C = Q^T BQ$. Since $PQ \in G$ and $C = (PQ)^T A(PQ)$ we have $A \sim_4 C$. Thus $\sim_4$ is an equivalence relation on X.

(vii) If A is a real symmetric $n \times n$ matrix we can find $P \in O(\mathbb{R}^n)$ such that $D = P^T AP$ with $D = (d_{ij})$ a diagonal matrix with $d_{ii} > 0$ for $1 \leq i < u$, $d_{ii} < 0$ for $u \leq i < v$ and $d_{ii} = 0$ otherwise. Let $E = (e_{ij})$ a diagonal matrix with $e_{ii} = d_{ii}^{-1/2}$ for $1 \leq i < u$, $e_{ii} = (-d_{ii})^{-1/2}$ for $u \leq i < v$ and $e_{ii} = 1$ otherwise. Then $PE \in GL(\mathbb{R}^n)$ and $(PE)^T A(PE)$ is a diagonal matrix with entries 1, -1 and 0. Thus there can only be finitely many equivalence classes.

[The reason that this does not complete the discussion is that we have not shown (as we will in Section 16.2) that different (u, v) correspond to different equivalence classes.]

Exercise 11.1.15

Let U and V be finite dimensional vector spaces. and let $\alpha : U \to V$ be linear.

By Theorem 11.1.13,

$$\dim \operatorname{im} \alpha = \dim(U/\ker \alpha)$$

and, by Lemma 11.1.14,

$$\dim U = \dim \ker \alpha + \dim(U/\ker \alpha),$$

so

$$\dim U = \dim \ker \alpha + \dim(\operatorname{im} \alpha).$$

Exercise 11.1.16

Observe that

$$\dim H_j = \dim B_j - \dim Z_j = \operatorname{rank}(\alpha_{j+1}) - (\dim C_j - \operatorname{rank}(\alpha_j))$$

and so

$$\dim H_j + \dim C_j = \operatorname{rank}(\alpha_{j+1}) + \operatorname{rank}(\alpha_j)$$

for $1 \leq j \leq n$

Now $\operatorname{rank}(\alpha_{n+1}) = \operatorname{rank}(\alpha_0) = 0$ since $C_{n+1} = C_0 = \{\mathbf{0}\}$ so

$$\sum_{j=1}^{n} (-1)^j (\dim H_j + \dim C_j) = 0.$$

Exercise 11.2.1

The sum of two polynomials is polynomial and the product of a polynomial with a real number is a polynomial, so we have a subspace of the vector space of functions $f : \mathbb{R} \to \mathbb{R}$ with pointwise addition and scalar multiplication.

(i) Let $\lambda,\ \mu \in \mathbb{R},\ P,\ Q \in \mathcal{P}$.

We have
$$D(\lambda P + \mu Q) = \lambda DP + \mu DQ,$$
by the standard rules for differentiation.

We have
$$M(\lambda P+\mu Q)(t) = t\big(\lambda P(t)+\mu Q(t)\big) = \lambda tP(t)+\mu tQ(t) = (\lambda MP+\mu MQ)(t),$$
so
$$M(\lambda P + \mu Q) = \lambda MP + \mu MQ.$$

We have
$$\begin{aligned} E_h(\lambda P + \mu Q)(t) &= (\lambda P(t+h) + \mu Q(t+h)) = \lambda E_h P(t) + \mu E_h Q(t) \\ &= (\lambda E_h P + \mu E_h Q)(t), \end{aligned}$$
so
$$M(\lambda P + \mu Q) = \lambda MP + \mu MQ.$$

(ii) We have
$$\big((DM - MD)\big)P(t) = \frac{d}{dt}tP(t) - tP'(t) = P(t),$$
so $(DM - MD)P = P$ for all $P \in \mathcal{P}$ and $DM - MD = \iota$ the identity map.

(iii) Well defined since for each P we only consider a finite sum and if $\alpha_j P = 0$ for all $j \geq N$
$$\sum_{j=0}^{N} \alpha_j P = \sum_{j=0}^{M} \alpha_j P$$
for all $M \geq N$.

If $\lambda,\ \mu \in \mathbb{R},\ P,\ Q \in \mathcal{P}.$ we can find an N such that $\alpha_j P = \alpha_j Q = 0$ for all $j \geq N$, so
$$\begin{aligned} \sum_{j=0}^{\infty} \alpha_j(\lambda P + \mu Q) &= \sum_{j=0}^{N} \alpha_j(\lambda P + \mu Q) = \sum_{j=0}^{N} (\lambda \alpha_j P + \mu \alpha_j Q) \\ &= \lambda \sum_{j=0}^{N} \alpha_j P + \mu \sum_{j=0}^{N} \alpha_j Q = \lambda \sum_{j=0}^{\infty} \alpha_j P + \mu \sum_{j=0}^{\infty} \alpha_j Q. \end{aligned}$$

Thus $\sum_{j=0}^{\infty}\alpha_j \in \mathcal{L}(\mathcal{P},\mathcal{P})$.

(iv) If P is a polynomial of degree N, then $D^jP=0$ for $j\geq N+1$.

If $e_n=t^n$, then $M^je_0=e_j\neq 0$, so M does not have the required property.

Since $E_h^je_0=e_j\neq 0$, E_h does not have the required property.

$(\iota-E_h)e_k$ is a polynomial of degree $k-1$ if $k\geq 1$. and is zero if $k=0$. Thus, if P is a polynomial of degree n, $(\iota-E_h)e_k$ is a polynomial of degree at most $n-1$ if $n\geq 1$. and is zero if $n=0$. Thus if P is a polynomial of degree N $(\iota-E_h)^{N+1}P=0$ and we have desired property.

(v) Part (iii) tells us that $\exp(\alpha)$ and $\log(\iota-\alpha)$ are well defined members of $\mathcal{L}(\mathcal{P},\mathcal{P})$.

Observe that

$$\exp\alpha-\iota=\alpha\sum_{j=1}^{\infty}\alpha^{j-1}/(j-1)=\alpha\beta$$

where β is an endomorphism which commutes with α. Thus

$$(\exp\alpha-\iota)^N=\alpha^N\beta^N=0$$

if $\alpha^N=0$. It follows, by part (iii), that $\log(\exp\alpha)$ is well defined.

By the standard rules on Taylor expansions, we have

$$\log(\exp t)=\sum_{j=0}^{\infty}c_jt^j$$

for $|t|$ small and (using those standard rules to compute c_j)

$$\log(\exp\alpha)=\sum_{j=0}^{\infty}c_j\alpha^j.$$

However we know from calculus that

$$\log(\exp t)=t$$

for $|t|$ small and so (by the uniqueness of Taylor expansions) $c_1=1$ and $c_j=0$ otherwise. Thus

$$\log(\exp\alpha)=\alpha.$$

(vi) We know that polynomials are their own Taylor series so

$$\big((\exp hD)P\big)=\sum_{j=0}^{\infty}\frac{h^j}{j!}P^{(j)}(t)=P(t+h)=(E_hP)(t)$$

and thus

$$\exp hD=E_h.$$

We know, from part (iv), that, for each $P \in \mathcal{P}$, we can find an $N(P)$ such that $\triangle_h^j(P) = 0$ for all $j > N(P)$. Thus (v) tells us that

$$\log E_h = hD,$$

that is to say

$$hD = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} \triangle_h^j.$$

(vii If P has degree N,

$$(\iota - \lambda D) \sum_{j=0}^{\infty} \lambda^j D^r P = (\iota - \lambda D) \sum_{j=N+1}^{\infty} \lambda^j D^r P = (\iota - \lambda^{N+2} D^{N+2}) P = P = \iota P$$

Thus

$$(\iota - \lambda D) \sum_{j=0}^{\infty} \lambda^j D^j = \iota$$

and similarly

$$\left(\sum_{j=0}^{\infty} \lambda^j D^j \right) (\iota - \lambda D) = \iota$$

It follows that

$$(\iota - \lambda D)P = Q \Rightarrow P = \left(\sum_{j=0}^{\infty} \lambda^j D^j \right) (\iota - \lambda D) P = \left(\sum_{j=0}^{\infty} \lambda^j D^j \right) Q.$$

We also have

$$(\iota - \lambda D) \left(\sum_{j=0}^{\infty} \lambda^j D^j \right) P = P.$$

Thus the unique polynomial solution of

$$(\iota - \lambda D)P = Q$$

is

$$Q = \sum_{j=0}^{\infty} \lambda^j D^j P = \sum_{j=0}^{\infty} \lambda^j P^{(j)}.$$

In particular, the unique polynomial solution of

$$f'(x) - f(x) = x^2$$

ie of

$$(\iota - D) f = -e_2$$

is

$$\sum_{j=0}^{\infty} \frac{d^j}{dx^j} (-x^2) = -x^2 - 2x - 2.$$

There is only one *polynomial* solution. (The general, not necessarily polynomial, solution is $Ae^x - x^2 - 2x - 2$ with A chosen freely.)

Exercise 11.2.3

If $\operatorname{rank}\alpha^n \neq 0$, then, since $\operatorname{rank}\alpha^m = 0$ for some m,

$$\operatorname{rank}\alpha^n - \operatorname{rank}\alpha^{n+1} \geq 1$$

Thus if $n \geq r \geq 0$

$$\operatorname{rank}\alpha^r - \operatorname{rank}\alpha^{r+1} \geq \operatorname{rank}\alpha^n - \operatorname{rank}\alpha^{n+1} \geq 1$$

so

$$\operatorname{rank}\alpha^n \leq n - \sum_{r=0}^{n-1} \operatorname{rank}\alpha^r - \operatorname{rank}\alpha^{r+1} \leq 0$$

which contradicts our original hypothesis.

By reductio ad absurdum, $\operatorname{rank}\alpha^n = 0$.

Now suppose that α has rank r and $\alpha^m = 0$, We have

$$n - r = n - \operatorname{rank}\alpha \geq \operatorname{rank}\alpha^j - \operatorname{rank}\alpha^{j+1}$$

so

$$m(n-r) \geq \sum_{j=0}^{m} \operatorname{rank}\alpha^j - \operatorname{rank}\alpha^{j+1} = n - \operatorname{rank}\alpha^m = n$$

so $mn - n \geq mr$ and $r \leq n(1 - m^{-1})$.

EXERCISE 11.2.4

Observe that if $A_k = \big(a_{ij}(k)\big)$ is the $k \times k$ matrix with $a_{i,i+1}(k) = 1$ for $1 \leq i \leq k-1$ $a_{ij}(k) = 0$ otherwise, then the associated linear map $\alpha_k : \mathbb{F}^k \to \mathbb{F}^k$ has $\operatorname{rank} \alpha^u(k) = \max(k, u-1)$. Consider the $n \times n$ matrix A which has $\big(s_j - s_{j+1}\big) - \big(s_{j+1} - s_{j+2}\big)$ copies of A_j and one $s_m \times s_m$ identity matrix along the diagonal and all other entries zero. If α has matrix A with respect to some basis, then α will have the desired property.

Exercise 11.2.5

If we write $r_k = \operatorname{rank} \alpha^k$, we know from general theorems that
$$r_0 - r_1 \geq r_1 - r_2 \geq r_2 - r_3 \geq \ldots \geq r_{n-1} - r_n \geq 0.$$
We have $r_0 - r_1 = 2$, so $2 \geq r_j - r_{j+1} \geq 0$. Since
$$2n = r_0 - r_n = \sum_{j=0}^{n-1} r_j - r_{j+1},$$
it follows that $2 = r_j - r_{j+1}$ for $0 \leq j \leq n-1$ and $\operatorname{rank} \alpha^j = 2n - 2j$ for $0 \leq j \leq n$.

Exercise 11.3.1

(i) We have
$$\delta_a(\lambda + \mu g) = (\lambda f + \mu g)(a) = \lambda f(a) + \mu g(a) = \lambda \delta_a f + \mu \delta_a g.$$

(ii) We have
$$\delta'_a(\lambda + \mu g) = -(\lambda f + \mu g)'(a) = -\lambda f'(a) - \mu g'(a) = \lambda \delta'_a f + \mu \delta'_a g.$$

(iii) We have
$$\begin{aligned} J(\lambda f + \mu g) &= \int_0^1 (\lambda f(x) + \mu g(x))\, dx \\ &= \lambda \int_0^1 f(x)\, dx + \mu \int_0^1 g(x)\, dx \\ &= \lambda J f + \mu J g. \end{aligned}$$

Exercise 11.3.6

$\Phi(\alpha) = 0 \Rightarrow \alpha' = 0$ by definition.

$\alpha' = 0 \Rightarrow \alpha'(\mathbf{v}') = 0\ \forall \mathbf{v}' \in V'$ by definition.

$\alpha'(\mathbf{v}') = 0\ \forall \mathbf{v}' \in V' \Rightarrow \alpha'(\mathbf{v}')\mathbf{u} = \mathbf{0}\ \forall \mathbf{v}' \in V'\ \forall \mathbf{u} \in U$ by definition.

$\alpha'(\mathbf{v}')\mathbf{u} = \mathbf{0}\ \forall \mathbf{v}' \in V'\ \forall \mathbf{u} \in U \Rightarrow \mathbf{v}'\alpha\mathbf{u} = \mathbf{0}\ \forall \mathbf{v}' \in V'\ \forall \mathbf{u} \in U$ since, by definition, $\alpha'(\mathbf{v}')\mathbf{u} = \mathbf{v}'\alpha(\mathbf{u})$.

$\mathbf{v}'\alpha\mathbf{u} = \mathbf{0}\ \forall \mathbf{v}' \in V'\ \forall \mathbf{u} \in U \Rightarrow \alpha\mathbf{u} = \mathbf{0}\ \forall \mathbf{u} \in U$ since V' separates V.

$\alpha\mathbf{u} = \mathbf{0}\ \forall \mathbf{u} \in U \Rightarrow \alpha = 0$ by definition.

Exercise 11.3.7

(i) We have

$$\begin{aligned}\big((\beta\alpha)'\mathbf{w}'\big)\mathbf{u} &= \mathbf{w}'\big((\beta\alpha)\mathbf{u}\big)\\ &= \mathbf{w}'\big(\beta(\alpha\mathbf{u})\big)\\ &= (\beta'\mathbf{w}')(\alpha\mathbf{u})\\ &= \big(\alpha'(\beta'\mathbf{w}')\big)\mathbf{u}\\ &= \big((\alpha'\beta')\mathbf{w}'\big)\mathbf{u}\end{aligned}$$

for all $\mathbf{u} \in U$, so

$$(\beta\alpha)'\mathbf{w}' = (\alpha'\beta')\mathbf{w}'$$

for all $\mathbf{w}' \in W'$, so

$$(\beta\alpha)' = \alpha'\beta'.$$

(ii) We have

$$(\iota_U'\mathbf{u}')\mathbf{v} = \mathbf{u}'\iota_u\mathbf{v} = \mathbf{u}'\mathbf{v}$$

for all $\mathbf{v} \in U$,so

$$\iota_U'\mathbf{u}' = \mathbf{u}'$$

for all $\mathbf{u}' \in U'$, so

$$\iota_U' = \iota_{U'}.$$

(iii) If α is invertible, then $\alpha\alpha^{-1} = \iota$ so, by the previous parts, $(\alpha^{-1})'\alpha' = \iota$ so α' is invertible and $(\alpha')^{-1} = (\alpha^{-1})'$.

Exercise 11.3.8

We have

$$\begin{aligned}\alpha''(\Theta\mathbf{u})\mathbf{v}' &= \Theta\mathbf{u}(\alpha'\mathbf{v}')\\ &= (\alpha'\mathbf{v}')\mathbf{u}\\ &= \mathbf{v}'\alpha\mathbf{u}\\ &= \big(\Theta(\alpha\mathbf{u})\big)\mathbf{v}'\end{aligned}$$

for all $\mathbf{v}' \in V$. Since V'' is separated by V',

$$\alpha''(\Theta\mathbf{u}) = \Theta(\alpha\mathbf{u})$$

for all $\mathbf{u} \in U$.

Exercise 11.3.9

(i) Certainly the zero sequence $\mathbf{0} \in c_{00}$.

If $\mathbf{a}$, $\mathbf{b} \in c_{00}$ we can find n and m such that $a_r = 0$ for $r \geq n$ and $b_r = 0$ for $r \geq m$. Thus if λ, $\mu \in \mathbb{R}$ we have $\lambda a_r + \mu b_r = 0$ for all $r \geq \max\{n, m\}$ so $\lambda\mathbf{a} + \mu\mathbf{b} \in c_{00}$. Thus c_{00} is a subspace of s.

(ii) If $\mathbf{x} \in c_{00}$, then we can find an n such that $x_r = 0$ for $r \geq n$. We have

$$\mathbf{x} = \sum_{r=1}^{n-1} x_r \mathbf{e}_r$$

and

$$T\mathbf{x} = \sum_{r=1}^{n-1} x_r T\mathbf{e}_r = \sum_{r=1}^{n-1} a_r x_r = \sum_{r=1}^{\infty} a_r x_r.$$

(iii) If $\mathbf{x} \in c_{00}$, then we can find an n such that $x_r = 0$ for $r \geq n$. Thus

$$T_{\mathbf{a}}\mathbf{x} = \sum_{r=1}^{n-1} a_r x_r \sum_{r=1}^{\infty} a_r x_r$$

is well defined.

If $\mathbf{x}$, $\mathbf{y} \in c_{00}$ and λ, $\mu \in \mathbb{R}$ we can find an n such that $x_r = y_r = 0$ for $r \geq n$. Thus

$$T\mathbf{a}(\lambda\mathbf{x}+\mu\curvearrowleft) = \sum_{r=1}^{n-1} a_r(\lambda x_r+\mu y_r) = \lambda\sum_{r=1}^{n-1} a_r x_r+\mu\sum_{r=1}^{n-1} a_r y_r = \lambda T_{\mathbf{a}}\mathbf{x}+\mu T_{\mathbf{a}}\mathbf{y}.$$

We have shown that $T_{\mathbf{a}} \in c_{00}$.

If $\mathbf{x} \in c_{00}$ and $\mathbf{x} \neq 0$ then $T_{\mathbf{x}}\mathbf{x} \neq 0$. Thus c'_{00} separates c_{00}.

(iv) Let $\theta : s \to c'_{00}$ be given by $\theta\mathbf{a} = T_{\mathbf{a}}$. By (iii), θ is well defined and, by (ii), θ is surjective.

If $\mathbf{x} \in c_{00}$, then we can find an n such that $x_r = 0$ for all $r \geq n$.

$$\begin{aligned} T_{\lambda\mathbf{a}+\mu\mathbf{b}}\mathbf{x} &= \sum_{r=1}^{n} x_r(\lambda a_r + \mu b_r) = \lambda\sum_{r=1}^{n} x_r a_r + \mu\sum_{r=1}^{n} x_r b_r \\ &= \lambda T_{\mathbf{a}}\mathbf{x} + \mu T_{\mathbf{b}}\mathbf{x} = (\lambda T_{\mathbf{a}} + \mu T_{\mathbf{b}})\mathbf{x} \end{aligned}$$

Thus $T_{\lambda\mathbf{a}+\mu\mathbf{b}} = \lambda T_{\mathbf{a}} + \mu T_{\mathbf{b}}$ and θ is linear. Further,

$$T_{\mathbf{a}} = 0 \Leftrightarrow T_{\mathbf{a}}\mathbf{e}_j = 0\ \forall j \Leftrightarrow a_j = 0\ \forall j \Leftrightarrow \mathbf{a} = \mathbf{0},$$

so θ is injective.

Thus c'_{00} is isomorphic to s.

Exercise 11.3.10

(i) If $\mathbf{x} \in c_{00}$, then we can find an n such that $x_r = 0$ for $r \geq n$. We have

$$\mathbf{x} = \sum_{r=1}^{n-1} x_r \mathbf{e}_r$$

(ii) Let If $\mathbf{w}_j \in \mathbb{R}^{n+1}$ be the row vector whose kth entry is the $m+k$ th entry of $\mathbf{f}_j$. Since n vectors cannot span a space of dimension $n+1$ there exists a vector $\mathbf{u} \in \mathbb{R}^{n+1}$ with

$$\mathbf{u} \notin \operatorname{span}\{\mathbf{w}_1,\, \mathbf{w}_2,\, \ldots,\, \mathbf{w}_{n+1}\}$$

Set $b_{m+r} = u_r$.

(iii) Let $m(1) = 0$ and let $m(n) = \sum_{r=1}^{n-1}(r+1)$ for $n \geq 2$. For each $n \geq 1$ choose a_r with $m(n) + 1 \leq r \leq m(n+1)$ so that

$$\sum_{j=1}^{n} \lambda_j \mathbf{f}_j = \mathbf{a}$$

has no solution with $\lambda_j \in \mathbb{R}$ $[1 \leq j \leq n]$.

By construction, $\mathbf{a}$ does not lie in the span of the $\mathbf{f}_j$. Thus s cannot be spanned by a countable set and is not isomorphic to c_{00} which can. Thus c_{00} which is isomorphic to s is not isomorphic to c_{00}'.

Exercise 11.4.2

We use the fact that $e_j(x_k) = \delta_{jk}$ repeatedly.

To see that the e_j are linearly independent, observe that

$$\sum_{j=0}^{n} \lambda_j e_j = 0 \Rightarrow \sum_{j=0}^{n} \lambda_j e_j(x_k) = 0 \ \forall k \Rightarrow \lambda_k = 0 \ \forall k.$$

To see that the e_j span, observe that, if $P \in \mathcal{P}_n$, then

$$P - \sum_{j=0}^{n} P(x_j) e_j$$

is a polynomial of degree at most n which vanishes at the $n+1$ points x_k and is this identically zero. Thus

$$P = \sum_{j=0}^{n} P(x_j) e_j.$$

We have $e_j(x_k) = \delta_{jk} = \hat{e}_k e_j$ so

$$\hat{e}_k \left(\sum_{j=0}^{n} \lambda_j e_j \right) = \sum_{j=0}^{n} \lambda_j \hat{e}_k e_j = \sum_{j=0}^{n} \lambda_j e_j(x_k)$$

and $\hat{e}_k P = P(x_k)$.

Exercise 11.4.4

Observe that
$$\mathbf{e}_j\hat{\mathbf{e}}_k = \hat{\mathbf{e}}_k\mathbf{e}_j = \delta_{jk}$$
for all k so $\hat{\hat{\mathbf{e}}}_j = \mathbf{e}_j$.

Exercise 11.4.5

We have

$$\mathbf{f}_1 = a\mathbf{e}_1 + b\mathbf{e}_2,$$
$$\mathbf{f}_2 = c\mathbf{e}_1 + d\mathbf{e}_2$$

and, since the $\hat{\mathbf{e}}_j$ form a basis,

$$\hat{\mathbf{f}}_1 = A\hat{\mathbf{e}}_1 + B\hat{\mathbf{e}}_2,$$
$$\hat{\mathbf{f}}_2 = B\hat{\mathbf{e}}_1 + D\hat{\mathbf{e}}_2$$

Now

$$1 = \hat{\mathbf{f}}_1\mathbf{f}_1 = (A\hat{\mathbf{e}}_1 + B\hat{\mathbf{e}}_2)(a\mathbf{e}_1 + b\mathbf{e}_2) = Aa + Bb$$

and

$$0 = \hat{\mathbf{f}}_1\mathbf{f}_2 = (A\hat{\mathbf{e}}_1 + B\hat{\mathbf{e}}_2)(c\mathbf{e}_1 + d\mathbf{e}_2) = Ac + Bd.$$

Similarly $Ca + Db = 0$, $Cc + Dd = 0$. In other words

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = I$$

so

$$Q = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}^{-1} = (ad - bc)^{-1} \begin{pmatrix} a & -c \\ -b & d \end{pmatrix}.$$

and

$$\hat{\mathbf{f}}_1 = (ad - bc)^{-1}a\hat{\mathbf{e}}_1 - b(ad - bc)^{-1}\hat{\mathbf{e}}_2,$$
$$\hat{\mathbf{f}}_2 = -c(ad - bc)^{-1}a\hat{\mathbf{e}}_1 + (ad - bc)^{-1}d\hat{\mathbf{e}}_2.$$

Exercise 11.4.6

Using the summation convention,

$$\delta_{js} = \hat{\mathbf{f}}_s \mathbf{f}_j = k_{rs} l_{ij} \hat{\mathbf{e}}_r \mathbf{e}_i = k_{rs} l_{ij} \delta_{ir} = k_{rs} l_{rj}$$

Thus $I = KL^T$ and $K = (L^T)^{-1}$.

Exercise 11.4.9

Let A and B be $n \times n$ matrices and α and β the corresponding linear maps for some fixed basis.

We know that $(\alpha+\beta)' = \alpha' + \beta'$, $(\lambda\alpha)' = \lambda\alpha'$, $\alpha'' = \alpha$ (since we work in a finite dimensional space) $\iota' = \iota$, $(\alpha\beta)' = \beta'\alpha'$ so

$$(A+B)^T = A^T+B^T,\ (\lambda A)^T = \lambda A^T,\ A^{TT} = A,\ I^T = I,\ (AB)^T = B^T A^T.$$

If A is invertible, then α is, so α' is with $(\alpha')^{-1} = (\alpha^{-1})'$, so A^T is invertible with $(A^T)^{-1} = (A^{-1})^T$.

Exercise 11.4.10

Let α have matrix A with respect to some basis. Then α' has matrix A^T with respect to the dual basis so (by the matrix result)

$$\det \alpha' = \det A^T = \det A = \det \alpha.$$

It follows that

$$\det(t\iota - \alpha) = \det(t\iota - \alpha)' = \det(t\iota - \alpha)$$

and α and α' have the same characteristic polynomials and so the same eigenvalues.

The trace is minus the coefficient of t^{n-1} in the characteristic polynomial (supposed of degree n). Thus $\operatorname{Tr} \alpha' = \operatorname{Tr} \alpha$.

Exercise 11.4.12

We certainly have the zero map $\mathbf{0} \in W^0$. If $\lambda,\ \mu \in \mathbb{F}$

$$\begin{aligned}\mathbf{u}',\ \mathbf{v}' \in W^0 &\Rightarrow (\lambda\mathbf{u}' + \mu\mathbf{v}')\mathbf{w} = \lambda\mathbf{u}'\mathbf{w} + \mu\mathbf{v}'\mathbf{w} = \lambda\mathbf{0} + \mu\mathbf{0} = \mathbf{0}\ \forall \mathbf{w} \in W\\ &\Rightarrow \lambda\mathbf{u}' + \mu\mathbf{v}' \in W^0.\end{aligned}$$

Thus W^0 is a subspace of U'.

Exercise 11.4.15

$$\begin{aligned}
\sum_{j=1} x_j\mathbf{e}_j \in W^{00} &\Leftrightarrow \left(\sum_{j=1}^{n} x_j\mathbf{e}_j\right)\sum_{r=k+1}^{n} y_r\mathbf{e}_r = 0\ \forall y_r \in \mathbb{F}\\
&\Leftrightarrow \left(\sum_{j=1}^{n} x_j\mathbf{e}_j\right)\mathbf{e}_r = 0\ \forall k+1 \le r \le n\\
&\Leftrightarrow x_r = 0\ \forall k+1 \le r \le n\\
&\Leftrightarrow \sum_{j=1} x_j\mathbf{e}_j \in W
\end{aligned}$$

Exercise 11.4.19

$$\begin{aligned}\dim\{\mathbf{u} \in U \,:\, \alpha\mathbf{u} = \lambda\mathbf{u}\} &= \dim(\lambda\iota - \alpha)^{-1}\mathbf{0}\\ &= \dim \operatorname{im}(\lambda\iota - \alpha)\\ &= \dim \operatorname{im}(\lambda\iota - \alpha)'\\ &= \dim \operatorname{im}(\lambda\iota - \alpha')\\ &= \dim\{\mathbf{u}' \in U' \,:\, \alpha'\mathbf{u} = \lambda\mathbf{u}'.\}\end{aligned}$$

The eigenvalues and the dimensions of the spaces spanned by the corresponding eigenvectors are the same for α and α'.

Exercise 11.4.20

$$\begin{aligned}\mathbf{v}' \in (\alpha U)^0 &\Rightarrow \mathbf{v}'(\alpha \mathbf{u}) = \mathbf{0}\ \forall \mathbf{u} \in U\\ &\Rightarrow (\alpha' \mathbf{v}')\mathbf{u} = \mathbf{v}'(\alpha \mathbf{u}) = \mathbf{0}\ \forall \mathbf{u} \in U\\ &\Rightarrow \alpha' \mathbf{v}' \in U^0\end{aligned}$$

so $\alpha'(\alpha U^0)$ is a subspace of U. (Remark that $(\alpha U)^0$ is a subspace of V' so $\alpha'(\alpha U)^0$ is a subspace of U'.)

Let V have basis $\mathbf{e}_1$, $\mathbf{e}_2$ and V' dual basis $\hat{\mathbf{e}}_1$ $\hat{\mathbf{e}}_2$. Let α be the endomorphism given by $\alpha(\mathbf{e}_1) = \mathbf{e}_1$, $\alpha(\mathbf{e}_2) = \mathbf{0}$. We observe that $\alpha'(\hat{\mathbf{e}}_1) = \hat{\mathbf{e}}_1$, $\alpha'\hat{(}\mathbf{e}_2) = \mathbf{0}$.

If $U = \operatorname{span}\{\mathbf{e}_1\}$, then $\alpha U = U$, $(\alpha U)^0 = \operatorname{span}\{\hat{\mathbf{e}}_2\}$ and

$$\alpha'(\alpha U)^0 = \{\mathbf{0}\} \neq U^0.$$

If If $W = \operatorname{span}\{\mathbf{e}_2\}$, then $\alpha W = \{\mathbf{0}\}$, $(\alpha W)^0 = V$ and

$$\alpha'(\alpha U)^0 = \operatorname{span}\{\hat{\mathbf{e}}_1\} = U^0$$

Exercise 12.1.2

If $U = U_1 \oplus U_2 \oplus \ldots \oplus U_m$ and $\mathbf{u} \in U$ then, by Definition 12.1.1 (i), we can find $\mathbf{u}_j \in U_j$ such that

$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_m.$$

If $\mathbf{v}_j \in U_j$ and

$$\mathbf{u} = \mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_m,$$

then $\mathbf{u}_j - \mathbf{v}_j \in U_j$ and

$$\mathbf{0} = \mathbf{u} - \mathbf{v} = (\mathbf{u}_1 - \mathbf{v}_1) + (\mathbf{u}_2 - \mathbf{v}_2) + \ldots + (\mathbf{u}_m - \mathbf{v}_m)$$

so, by Definition 12.1.1 (ii), $\mathbf{u}_j - \mathbf{v}_j = \mathbf{0}$ and $\mathbf{u}_j = \mathbf{v}_j$ $\forall j$.

If the equation

$$\boldsymbol{u} = \mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_m$$

has exactly one solution with $\mathbf{u}_j \in U$, then the conditions of Definition 12.1.1 can be read off.

Exercise 12.1.3

(i) If $\mathbf{u} \in U$ then we can find $\mathbf{u}_j \in U_j$ such that

$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_m.$$

For each j we can find $\lambda_{jk} \in \mathbb{F}$ such that

$$\mathbf{u}_j = \sum_{k=1}^{n(j)} \lambda_{jk}\mathbf{e}_{jk}$$

and so

$$\mathbf{u} = \sum_{j=1}^{m}\sum_{k=1}^{n(j)} \lambda_{jk}\mathbf{e}_{jk}.$$

Thus the $\mathbf{e}_{jk}$ span U.

If

$$\mathbf{0} = \sum_{j=1}^{m}\sum_{k=1}^{n(j)} \lambda_{jk}\mathbf{e}_{jk}$$

then setting $\mathbf{u}_j = \sum_{k=1}^{n(j)} \lambda_{jk}\mathbf{e}_{jk}$ we have $\mathbf{u}_j \in U_j$ and, applying the definition of direct sum, we obtain

$$\mathbf{0} = \mathbf{u}_j = \sum_{k=1}^{n(j)} \lambda_{jk}\mathbf{e}_{jk},$$

so $\lambda_{jk} = 0$ for all k and j. Thus the $\mathbf{e}_{jk}$ are linearly independent and so form a basis.

(ii) and (iii) Since every subspace of a finite dimension subspace is finite dimensional

$$U \text{ finite dimensional} \Rightarrow U_j \text{ finite dimensional}$$

If the U_j are finite dimensional the we can choose bases as in (i) and observe that U is finite dimensional with

$$\dim U = \sum_{j=1}^{m} n(j) = \sum_{j=1}^{m} \dim U_j.$$

Exercise 12.1.4

Observe that $B = U \cap V$ is a subspace of a finite dimensional space so has a basis $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$. Since U is a subspace of a finite dimensional space containing B we can extend the basis of B to a basis of U

$$\mathbf{e}_1,\, \mathbf{e}_2,\, \dots,\, \mathbf{e}_k,\, \mathbf{e}_{k+1},\, \mathbf{e}_{k+2},\, \dots,\, \mathbf{e}_{k+l}.$$

Similarly we can extend the basis of B to a basis of v

$$\mathbf{e}_1,\, \mathbf{e}_2,\, \dots,\, \mathbf{e}_k,\, \mathbf{e}_{k+l+1},\, \mathbf{e}_{k+l+2},\, \dots,\, \mathbf{e}_{k+l+r}.$$

Let

$$C = \operatorname{span}\{\mathbf{e}_{k+1},\, \mathbf{e}_{k+2},\, \dots,\, \mathbf{e}_{k+l},\, \mathbf{e}_{k+l+1},\, \mathbf{e}_{k+l+2},\, \dots,\, \mathbf{e}_{k+l+r}\}.$$

We show that the $\mathbf{e}_j$ are linearly independent as follows. If

$$\sum_{j=1}^{k+l+r} \lambda_j \mathbf{e}_j = \mathbf{0}$$

then

$$\sum_{j=1}^{k+l} \lambda_j \mathbf{e}_j = -\sum_{j=k+l+1}^{k+l+r} \lambda_j \mathbf{e}_j \in U \cap V = B$$

Thus

$$\sum_{j=1}^{k+l} \lambda_j \mathbf{e}_j = \sum_{j=1}^{k} \mu_j \mathbf{e}_j$$

for some μ_j. Thus $\lambda_j = 0$ for $1 \leq j \leq k$. Thus

$$\sum_{j=1}^{k} \lambda_j \mathbf{e}_j + \sum_{j=k+l+1}^{k+l+r} \lambda_j \mathbf{e}_j = 0$$

so $\lambda_j = 0$ for $k + l + 1 \leq j \leq k + l + r$ and for $1 \leq j \leq k$.

We can now read off the desired results,

$$U = A \oplus B,\ W = B \oplus C \text{ and } U + W = U \oplus C = W \oplus A.$$

Observe that if

$$U = A' \oplus B',\, W = B' \oplus C', U + W = U \oplus C' = W \oplus A'$$

we have $B' \subseteq U, V$ so

$$B' \subseteq U \cap V.$$

If $\mathbf{v} \in U \cap W$ then we can find $\mathbf{b}_1, \mathbf{b}_2 \in B'$ and $\mathbf{a} \in A$, $\mathbf{c} \in C$ such that

$$\mathbf{v} = \mathbf{a} + \mathbf{b}_1 = \mathbf{c} + \mathbf{b}_2$$

so

$$\mathbf{a} - \mathbf{c} = \mathbf{b}_1 - \mathbf{b}_2 \in B'.$$

Thus

$$\mathbf{a} = \mathbf{c} + (\mathbf{b}_1 - \mathbf{b}_2)$$

with $\mathbf{a} \in A' \subseteq U$ and $\mathbf{b}_1 - \mathbf{b}_2 \in U \cap W \subseteq U$. Since $U + W = U \oplus C'$ we have $\mathbf{c} = \mathbf{0}$. so

$$\mathbf{v} = \mathbf{c} + \mathbf{b}_2 = \mathbf{b}_2 \in B'.$$

Thus

$$B' = U \cap W$$

and we have shown B unique.

Not unique. Take $V = \mathbb{R}^3$ (using row vectors)

$$U = \{(x, y, 0) : x,\, y \in \mathbb{R}\},\ V = \{(x, 0, z) : x,\, z \in \mathbb{R}\}$$
$$B = \{(x, 0, 0) : x \in \mathbb{R}\},$$
$$A = \{(0, y, 0) : y \in \mathbb{R}\},\ A' = \{(y, y, 0) : y \in \mathbb{R}\}$$
$$C = \{(0, 0, z) : y \in \mathbb{R}\},\ C' = \{(z, 0, z) : z \in \mathbb{R}\}.$$

Then

$$U = A \oplus B,\ W = B \oplus C \text{ and } U + W = U \oplus C = W \oplus A$$

and

$$U = A' \oplus B,\ W = B' \oplus C \text{ and } U + W = U \oplus C' = W \oplus A'.$$

Exercise 12.1.6

(i) Can occur. Take $V = V_1 = V_2 = \mathbb{F}^2$.

(ii) Cannot occur. If $\mathbf{e}_1$, $\mathbf{e}_2$, $\ldots \mathbf{e}_k$ is a basis for V_1 and $\mathbf{e}_{k+1}$, $\mathbf{e}_{k+2}$, $\ldots \mathbf{e}_{k+m}$ is a basis for V_2 then $\mathbf{e}_1$, $\mathbf{e}_2$, $\ldots \mathbf{e}_{k+m}$ span $V_1 + V_2$ so

$$\dim(V_1 + V_2) \le k + m = \dim V_1 + \dim V_2$$

(ii) Can occur. Let $\mathbf{e}_1$, $\mathbf{e}_2$ be a basis for $V = \mathbb{F}^2$ and

$$V_1 = V_2 = \operatorname{span}\{\mathbf{e}_1\}.$$

Exercise 12.1.7

We know from Theorem 11.2.2 that there exists an m such that $\alpha^{m+1}U = \alpha^m U$. We thus have $\alpha^m U$ an invariant subspace.

If W is an invariant subspace

$$W = \alpha^m V \subseteq \alpha^m U.$$

Thus $\alpha^m U$ is the unique maximal invariant subspace.

(ii) If $\mathbf{x} \in M \cap N$ then

$$\alpha^m \mathbf{x} = 0$$

But $\alpha|_M : M \to M$ is surjective so invertible so $\alpha|_M^m$ is so $\mathbf{x} = \mathbf{0}$. By the rank-nullity theorem $\dim M + \dim N = \dim U$ so $U + M \oplus V$.

(iii) $\alpha(M) = M \subseteq M$, $\alpha^m \alpha(N) = \alpha\alpha^m N = \mathbf{0}$ so $\alpha N \subseteq N$.

(iv) As we observed in (ii), $\beta : M \to M$ is surjective, so β is invertible, so an isomorphism. We know that $\gamma^m = 0$.

(v) Observe that M' is an invariant subspace so $M' \subseteq M$. If $\mathbf{b} \in N'$ then

$$\alpha^r(\mathbf{b}) = (\gamma')^r \mathbf{b} = \mathbf{0}$$

for r large so $\mathbf{b} \in N$. Thus $N' \subseteq N$. We have $\dim M' \leq \dim M$, $\dim N' \leq \dim N$

$$\dim U = \dim M' + \dim N' \leq \dim M + \dim N = \dim U$$

so $\dim M' = \dim M$, $\dim N' = \dim N$ and $M = M'$, $N = N'$. By considering $\mathbf{b} = \mathbf{0}$, we have $\beta = \beta'$ and similarly, by considering $\mathbf{a} = \mathbf{0}$, we have $\gamma = \gamma'$.

(vi) Let A correspond to a linear map α for some choice of basis for a space U of appropriate dimension. Choose a basis $\mathcal{M}$ for M defined as above and a basis $\mathcal{N}$ for N. Then $\mathcal{E}$, $\mathcal{N}$ is a basis for U with respect to which α has the form

$$\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$$

with B an invertible $r \times r$ matrix and C a nilpotent $n - r \times n - r$ matrix, The result now follows from the change of basis formula.

Exercise 12.1.11

Let $\dim U = n$ and $\dim V = r$. If U is neither V nor $\{\mathbf{0}\}$, then $0 < r < n$. Choose a basis

$$\mathbf{e}_1,\, \mathbf{e}_2,\, \ldots,\, \mathbf{e}_r$$

and extend it to a basis

$$\mathbf{e}_1,\, \mathbf{e}_2,\, \ldots,\, \mathbf{e}_n$$

of V.

If

$$W = \operatorname{span}\{\mathbf{e}_{r+1},\, \mathbf{e}_2,\, \ldots,\, \mathbf{e}_n\}$$
$$W' = \operatorname{span}\{\mathbf{e}_1 + \mathbf{e}_{r+1},\, \mathbf{e}_2,\, \ldots,\, \mathbf{e}_n\},$$

then $V = U \oplus W = U \oplus W'$ but $\mathbf{e}_{r+1} \in W \setminus W'$ so $W \neq W$.

Exercise 12.1.13

(i) It will remain true θ has an $n^2 \times n^2$ diagonal matrix with $\dim U$ diagonal entries taking the value 1 and $\dim V$ diagonal entries taking the value -1, since $\theta|_U = \iota|_U$ and since $\theta|_V = -\iota|_V$.

(ii) $\det \theta = (-1)^{\dim V} = (-1)^{n(n-1)/2}$. Now

$$\frac{(n+4)((n+4)-1)}{2} - \frac{n(n-1)}{2} = \frac{n^2 + 8n + 16 - n - 4 - n^2 + n}{2} = 4n+6$$

is divisible by 2 so $\det \theta$ depends only on the value of n modulo 4. By inspection

$$\begin{aligned} n \equiv 0 &\Rightarrow \det \theta = 1, \\ n \equiv 1 &\Rightarrow \det \theta = 1, \\ n \equiv 2 &\Rightarrow \det \theta = -1, \\ n \equiv 3 &\Rightarrow \det \theta = -1. \end{aligned}$$

Exercise 12.1.14

Automatically $0 \in V$ and, if $\lambda,\ \mu \in \mathbb{R}$,

$$\begin{aligned} f,\ g \in V \Rightarrow (\lambda f + \mu g)(-x) &= \lambda f(-x) + \mu g(-x) \\ &= -\lambda f(x) - \mu g(x) = -(\lambda f + \mu g)(x)\ \forall x \\ \Rightarrow \lambda f + \mu g \in V, & \end{aligned}$$

so V is subspace. A similar argument shows that U is a subspace,

If $f \in C(\mathbb{R})$, then $u(x) = (f(x)+f(-x))/2$, $v(x) = (f(x)-f(-x))/2$ define $u \in U$, $v \in V$. Since $f = u + v$ we have $U + V = C(\mathbb{R})$.

Now if $f \in U \cap V$, then $f(x) = f(-x) = -f(x)$ so $f(x) = 0$ for all $x \in \mathbb{R}$ and $f = 0$. Thus $U \cap V = \{0\}$ and U and V are complementary subspaces.

Exercise 12.1.15

(i)⇒(ii) Let $U = \alpha V$ and $W = (\iota - \alpha)V$.

If $\mathbf{u} \in U$, then $\mathbf{u} = \alpha\mathbf{v}$ for some $\mathbf{v} \in V$ so

$$\alpha\mathbf{u} = \alpha^2\mathbf{v} = \alpha\mathbf{v} = \mathbf{u}$$

so $\alpha|_U = \iota|_U$.

If $\mathbf{u} \in W$, then $\mathbf{u} = (\iota - \alpha)\mathbf{v}$ for some $\mathbf{v} \in V$, so

$$\alpha\mathbf{u} = (\alpha - \alpha^2)\mathbf{v} = \alpha\mathbf{v} - \alpha\mathbf{v} = \mathbf{0},$$

so $\alpha|_W = 0|_W$

Since $\mathbf{v} = \alpha\mathbf{v} + (\iota - \alpha)\mathbf{v}$ for all $\mathbf{v} \in V$ so $V = U + W$.

If $\mathbf{u} \in U \cap W$, then

$$\mathbf{u} = \alpha|_U\mathbf{u} = \alpha\mathbf{u} = \alpha|_W\mathbf{u} = \mathbf{0}.$$

Thus $U \cap W = \{\mathbf{0}\}$ and $V = U \oplus W$.

(ii)⇒(iii) Let $\mathbf{e}_1$, $\mathbf{e}_2$, $\dots$, $\mathbf{e}_r$ be a basis of U and $\mathbf{e}_{r+1}$, $\dots$, $\mathbf{e}_n$ be a basis of W. Then with respect to the basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\dots$, $\mathbf{e}_n$ of V, α has a matrix of stated type.

(iii)⇒(i) Since $A^2 = A$ we have $\alpha^2 = \alpha$.

We have $\alpha_1(x, y) = (x, 0) = \alpha_1^2(x, y)$ so α_1 is a projection.

We have $\alpha_2(x, y) = (0, y) \neq (0, 0) = \alpha_2^2(x, y)$ if $y = 1$, so α_2 is not a projection.

We have $\alpha_3(x, y) = (x, y) \neq (y, x) = \alpha_3^2(x, y)$ if $x = 0$, $y = 1$, so α_3 is not a projection.

We have $\alpha_4(x, y) = (x + y, 0) = (x + y, 0) = \alpha_4^2(x, y)$, so α_4 is a projection.

We have $\alpha_5(x, y) = (x + y, x + y) \neq \big(2(x + y), 2(x + y)\big) = \alpha_5^2(x, y)$ if $x = 0$, $y = 1$, so α_5 is not a projection.

We have $\alpha_6(x, y) = \big(\frac{1}{2}(x + y), \frac{1}{2}(x + y)\big) = \alpha_6^2(x, y)$ so α_6 is a projection.

Exercise 12.1.16

If α is a projection, then
$$(\iota-\alpha)^2=\iota-2\alpha+\alpha^2=\iota-2\alpha+\alpha=\iota-\alpha$$
so $\iota-\alpha$ is a projection.

If $\iota-\alpha$ is a projection, then the first paragraph tells us that $\alpha=\iota-(\iota-\alpha)$ is a projection.

Exercise 12.1.17

If α and β are projections and $\alpha\beta = \beta\alpha$, then
$$(\alpha\beta)^2 = (\alpha\beta)(\alpha\beta) = \alpha(\beta\alpha)\beta = \alpha(\alpha\beta)\beta = \alpha^2\beta^2 = \alpha\beta.$$

We work in $\mathbb{R}^2$ with row vectors. If
$$\beta(x,y) = (x+y,0),\ \alpha(x,y) = (0,y)$$
then $\alpha^2 = \alpha$, $\beta^2 = \alpha$
$$\beta\alpha(x,y) = (y,0),\ \alpha\beta(x,y) = (0,y)$$
so $(\alpha\beta)^2 = (\alpha\beta)$ but $(\beta\alpha)^2 \neq (\beta\alpha)$.

If
$$\beta(x,y) = (x+y,0),\ \alpha(x,y) = (0,x+y)$$
then $\alpha^2 = \alpha$, $\beta^2 = \alpha$
$$\beta\alpha(x,y) = (x+y,0),\ \alpha\beta(x,y) = (0,x+y)$$
so $(\alpha\beta)^2 = (\alpha\beta)$ and $(\beta\alpha)^2 = (\beta\alpha)$, but (looking e.g. at $(x,y) = (1,1)$) $\alpha\beta \neq \alpha\beta$.

Exercise 12.1.18

(i) If $\alpha\beta = -\beta\alpha$, then

$$\alpha\beta = \alpha\alpha\beta = \alpha(-\beta\alpha) = (-\alpha\beta)\alpha = \beta\alpha\alpha = \beta\alpha$$

Thus $\alpha\beta = 0$ and $\beta\alpha = 0$.

(ii) If $(\alpha + \beta)$ is a projection, then

$$\alpha + \alpha\beta + \beta\alpha + \beta = (\alpha + \beta)^2 = \alpha + \beta,$$

so $\alpha\beta = -\beta\alpha$ and $\alpha\beta = \beta\alpha = 0$.

If $\alpha\beta = \beta\alpha = 0$, then

$$(\alpha + \beta)^2 = \alpha + \alpha\beta + \beta\alpha + \beta = \alpha + \beta,$$

so $\alpha + \beta$ is a projection.

(ii) If $(\alpha - \beta)$ is a projection, then

$$\alpha - \alpha\beta - \beta\alpha + \beta = (\alpha - \beta)^2 = \alpha - \beta,$$

so $-\alpha\beta - \beta\alpha = 2\beta$ whence $(\iota - \alpha)\beta = -\beta(\iota - \alpha)$ so, since $\iota - \alpha$ is a projection, $(\iota - \alpha)\beta = \beta(\iota - \alpha) = 0$ and $\alpha\beta = \beta\alpha = \beta$.

If $\alpha\beta = \beta\alpha = \beta$, then

$$(\alpha - \beta)^2 = \alpha - \alpha\beta - \beta\alpha + \beta = \alpha - \beta,$$

so $\alpha - \beta$ is a projection.

Exercise 12.1.19

If α is diagonalisable with distinct eigenvalues λ_1, λ_2, $\dots$, λ_m, then U is the direct sum of the spaces

$$E_j = \{\mathbf{e} : \alpha\mathbf{e} = \lambda_j\mathbf{e}\}.$$

If $\mathbf{u} \in U$, we can write $\mathbf{u}$ uniquely as

$$\mathbf{u} = \mathbf{e}_1 + \mathbf{e}_2 + \dots + \mathbf{e}_m$$

with $\mathbf{e}_j \in E_j$ If we set $\pi_j(\mathbf{e}) = \mathbf{e}_j$ for $1 \leq j \leq m$ then direct verification shows that π_j is linear and π_j is a projection. By inspection $\pi_k\pi_j = 0$ when $k \neq j$ and

$$\alpha = \lambda_1\pi_1 + \lambda_2\pi_2 + \dots + \lambda_m\pi_m.$$

Conversely, if the stated conditions hold and we write $E_j = \pi_j U$, we see that E_j is a subspace and $\pi_i E_j = \pi_i\pi_j E_j = 0$ for $i \neq j$. Since $iota = \pi_1 + \pi_2 + \dots + \pi_m$ we have

$$\mathbf{u} = \iota\mathbf{u} = \sum_{j=1}^{m} \pi_j\mathbf{u} = \sum_{j=1}^{m} \mathbf{u}_j$$

with $\mathbf{u}_j = \pi_j\mathbf{u} \in E_j$. On the other hand, if $\mathbf{e}_j \in E_j$ and

$$\sum_{j=1}^{m} \mathbf{e}_j = \mathbf{0}$$

then applying π_i to both sides we get $\mathbf{e}_i = \mathbf{0}$ for all i. Thus

$$U = E_1 \oplus E_2 \oplus \dots \oplus E_m.$$

Let $\mathcal{E}_j$ be a basis for E_j. The set $\mathcal{E} = \bigcup_{j=1}^{m} \mathcal{E}_j$ is a basis for U. The conditions $\pi_k\pi_j = 0$ when $k \neq j$ and

$$\alpha = \lambda_1\pi_1 + \lambda_2\pi_2 + \dots + \lambda_m\pi_m$$

show that $\mathcal{E}$ is a basis of eigenvectors so α is diagonalisable.

Exercise 12.2.1

Let $E(rs) = (\delta_{ir}\delta_{js})_{1\leq i\leq n}^{1\leq j\leq n}$

Then, if $A = (a_{ij})$,

$$A = \sum_{r=1}^{n}\sum_{s=1}^{n} a_{rs}E(rs)$$

so the $E(r, s)$ span $M_n(\mathbb{F})$.

Also

$$\sum_{r=1}^{n}\sum_{s=1}^{n} b_{rs}E(rs) = 0 \Rightarrow B = 0 \Rightarrow b_{rs} = 0 \ \forall r \ \forall s$$

so the $E(r, s)$ form a basis for $M_n(\mathbb{F})$. Thus $\dim M_n(\mathbb{F}) = n^2$.

It follows that the $n^2 + 1$ elements A^j with $0 \leq j \leq n^2$ must be linearly dependent, ie we can find $a_j \in \mathbb{C}$, not all zero, such that $\sum_{j=0}^{n^2} a_j A^j = 0$.

In other words, there is a non-trivial polynomial P of degree at most n^2 such that $P(A) = 0$.

Exercise 12.2.2

(i) Observe that

$$\chi_D(t) = \prod_{j=1}^{n}(t - \lambda_j).$$

Let D_j be the diagonal matrix with the same entries as D except that the jth diagonal entry is 0. Then

$$\sum_{k=0}^{n} b_k D^k = \prod_{j=1}^{n}(D - \lambda_j I) = \prod_{j=1}^{n} D_j = 0.$$

(ii) We have $A = P^{-1}DP$ for some invertible matrix P.

$$\chi_D = \chi_A$$

and

$$\chi_A(A) = \chi_D(A) = \chi_D(P^{-1}DP) = P^{-1}\chi_D(D)P = P^{-1}0P = 0$$

as stated.

Exercise 12.2.3

(i) Let $C = B^{-1}$. Using the summation convection,

$$\operatorname{Tr} B^{-1}AB = \operatorname{Tr} CAB = c_{ij}a_{jk}b_{ki} = b_{ki}c_{ij}a_{jk} = \delta_{kj}a_{jk} = a_{kk} = \operatorname{Tr} A.$$

(Lots of other proofs available.)

(ii) $Q_A(tI - A) = t\delta_{ii} - a_{ii} = nt - \operatorname{Tr} A$. Part (i) yields $Q_{B^{-1}AB} = Q_A$.

(iii) $Q_A(A) = nA - (\operatorname{Tr} A)I$. If $a_{11} = -a_{22} = 1$ and $a_{ij} = 0$ otherwise, the. $Q_A(A) \neq 0$.

If $n = 1$, $Q_{(a_{11})}(t) = t - a_{11}$ and $Q_{(a_{11})}\big((a_{11})\big) = (a_{11}) - a_{11}(1) = (0)$.

Exercise 12.2.6

The roots of the characteristic polynomial of a triangular matrix are the diagonal entries. If every real $n \times n$ was triangularisable by change of basis, then, since the characteristic equation is unaltered by change of basis, every root of the characteristic equation would be real. Taking $A = (a_{ij})$ with $a_{rr} = 1$ for $r \neq 1, 2$ $a_{12} = -a_{21} = 1$, $a_{ij} = 0$, otherwise, we see that this is not the case.

If $\dim V = 1$ every matrix is triangular.

Exercise 12.2.7

(i) Choose a basis such that α has an upper triangular matrix with respect to that basis. The eigenvalues of α are the diagonal entries.

With respect to the chosen basis α^r has matrix A^r which is upper triangular with diagonal entries the rth powers of the diagonal entries of A. Thus α^r has an eigenvalue μ if and only if α has an eigenvalue λ with $\lambda^r = \mu$.

(ii) If α is invertible then α^r has an eigenvalue μ if and only if α has an eigenvalue λ with $\lambda^r = \mu$.

To prove this observe that if β is invertible with eigenvalue λ and associated eigenvector $\mathbf{e}$

$$\beta^{-1}\beta\mathbf{e} = \mathbf{e}$$

so $\lambda \neq 0$ and β^{-1} has eigenvalue λ^{-1} and associated eigenvector $\mathbf{e}$.

(iii) False. Let

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

Then A has no real eigenvalues but $A^2 = I$ has 1 as eigenvalue.

$$\begin{pmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ 0 & 0 & b_{23} \\ 0 & 0 & b_{33} \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ 0 & c_{22} & c_{23} \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Exercise 12.2.8

(i) We have

$$\begin{pmatrix}0 & a_{12} & a_{13}\\ 0 & a_{22} & a_{23}\\ 0 & 0 & a_{33}\end{pmatrix}\begin{pmatrix}b_{11} & b_{12} & b_{13}\\ 0 & 0 & b_{23}\\ 0 & 0 & b_{33}\end{pmatrix}\begin{pmatrix}c_{11} & c_{12} & c_{13}\\ 0 & c_{22} & c_{23}\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

$$=\begin{pmatrix}0 & a_{12} & a_{13}\\ 0 & a_{22} & a_{23}\\ 0 & 0 & a_{33}\end{pmatrix}\begin{pmatrix}b_{11}c_{11} & b_{11}c_{12}+b_{12}c_{22} & b_{11}c_{13}+b_{12}c_{23}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

$$=\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

(ii) Let T_j be an $n\times n$ upper triangular matrix with the jth diagonal entry 0. Then $T_1T_2\dots T_n=0$

We can prove this by induction on n using matrices or as follows. Let τ_j be the linear map on an n dimensional vector space U which has matrix T_j with respect to some basis $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$. Then writing

$$E_j=\operatorname{span}\{\mathbf{e}_1,\dots,\mathbf{e}_j\}$$

for $1\leq j\leq n$ and $E_0=\{\mathbf{0}\}$ we have $\tau_jE_j\subseteq E_{j-1}$. Thus

$$\tau_1\tau_2\dots\tau_nU=\tau_1\tau_2\dots\tau_nE_n=E_0$$

and $\tau_1\tau_2\dots\tau_n=0$ so $T_1T_2\dots T_n=0$

(iii) Not necessarily true.

$$\begin{pmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}1 & 1 & 1\\ 0 & 0 & 1\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}$$

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

$$=\begin{pmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}3\\ 1\\ 1\end{pmatrix}=\begin{pmatrix}5\\ 2\\ 1\end{pmatrix}$$

so the matrix product is not zero.

Exercise 12.2.9

By induction on r, $\beta\alpha^r = \alpha^r\beta$, so, by induction on s, $\beta^s\alpha^r = \alpha^r\beta^s$. Thus

$$\begin{aligned}\sum_{j=0}^{n} a_j\alpha^j \sum_{k=0}^{m} a_k\beta^k &= \sum_{j=0}^{n}\sum_{k=0}^{m} a_j b_k \alpha^j\beta^k = \sum_{k=0}^{m}\sum_{j=0}^{n} b_k a_j \beta^k\alpha^j \\ &= \sum_{k=0}^{m} a_k\beta^k \sum_{j=0}^{n} a_j\alpha^j.\end{aligned}$$

Exercise 12.2.11

Setting $t = 0$ we have

$$0 \neq \det A = a_0$$

By the Cayley–Hamilton theorem

$$a_0 I = \sum_{j=1}^{n} a_j A^j$$

so, multiplying by $a_0^{-1}A^{-1}$, we get

$$A^{-1} = -a_0^{-1} \sum_{j=1}^{n} a_j A^{j-1}.$$

Does not appear to be a good method. The computation of the a_j appears to require the evaluation of

$$\det(tI - A)$$

which appears to require at least as many operations as inverting a matrix by Gaussian elimination.

Exercise 12.2.12

Let M be the smallest value of m with $\alpha^m = 0$

We know that there exist a_j such that

$$\alpha^n = \sum_{j=0}^{n-1} a_j \alpha^j$$

and so either $\alpha^n = 0$ or there exists a k with $n - 1 \geq k \geq 0$ such that

$$\alpha^n = \sum_{j=k}^{n-1} a_j \alpha^j$$

and $a_k \neq 0$. But then applying α^{M-k-1} to both sides we get

$$0 = a_k \alpha^{M-1}$$

so $\alpha^{M_1} = 0$, contrary to our definition of M.

Exercise 12.3.1

(i) We have
$$\det(tI - A_1) = \det(tI - A_2) = t^2.$$
A_1 and A_2 are not similar since they have different ranks.

(ii) We have
$$\det(tI - A_1) = \det(tI - A_2) = \det(tI - A_3) = t^3.$$
None of A_1, A_2, A_3 are similar since they have different ranks.

(iii) Set
$$A_8 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\ A_9 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$
and
$$A_{10} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
Then
$$\det(tI - A_j) = t^4$$
for $6 \leq j \leq 10$. Now A_6 has rank 1, A_7 and A_9 have rank 2, A_8 has rank 0 and A_9 has rank 3.

$$A_7^2 = 0, A_9^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
so A_7^2 has rank 0 and A_9^2 rank 1. None of the A_j can be similar.

Exercise 12.3.3

(i) Since the characteristic polynomials have the form t^n so do the minimal polynomials. The minimal polynomial of A_k is $m_k(t) = t^r$ where $A_k^r = 0$, $A_k^{r-1} \neq 0$. Thus

$$m_1(t) = t,\ m_2(t) = t^2,\ m_3(t) = t,\ m_4(t) = t^2, m_5(t) = t^3 m_6(t) = t^2,\ m_7(t) = t^2,\ m_8(t) = t,\ m$$

(ii) We have

$$\det(tI - A) = (t-1)^2(t-2)^2 = \det(tI - B)$$

Now $(A-I)(A-2I)^2 \neq 0$ and $(A-I)^2(A-I)^2 \neq 0$ so A has minimal polynomial $(t-1)^2(t-2)^2$.

We also have $(B-I)(B-2I) \neq 0$ and $(B-II)^2 \neq 0$ but $(B-I)^2(B-2I) \neq 0$ so A has minimal polynomial $(t-1)^2(t-2)$.

Exercise 12.3.6

(i) $\prod_{i=1}^{r}(D-\lambda_i I)$ is a diagonal matrix with jth diagonal entry $\prod_{i=1}^{r}(d_{jj}-\lambda_i)=0$ so $\prod_{i=1}^{r}(D-\lambda_i I)=0$.

Now suppose, without loss of generality that $d_{11}=\lambda_1$. Then $\prod_{i=2}^{r}(D-\lambda_i I)$ is a diagonal matrix with 1st diagonal entry $\prod_{i=2}^{r}(d_{11}-\lambda_i)\neq 0$ so $\prod_{i=2}^{r}(D-\lambda_i I)\neq 0$. Thus $\prod_{i=1}^{r}(t-\lambda_i)$ is the minimal polynomial for D.

(ii) Choose a basis of eigenvectors. $\mathbf{e}_1, \mathbf{e}_2, \dots \mathbf{e}_n$. With respect to this basis α has a diagonal matrix with j th diagonal entry the eigenvalue associated with $\mathbf{e}_j$. Thus if λ is associated with k of the basis vectors $\det(tI-D)=\det(t\iota-\alpha)$ has $(t-\lambda)^k$ as a factor but not $(t-\lambda)^{k+1}$

Now writing S for the set of eigenvalues which are not equal to λ

$$\dim U_\lambda = \dim(\prod_{\mu\in S}(\mu\iota-\alpha)U) = \dim\operatorname{span}\{\prod_{\mu\in S}(\mu\iota-\alpha)U)\mathbf{e}_j \,:\, 1\leq j\leq n\} = k$$

Exercise 12.3.7

(i) If P has a repeated root θ, then $P(t) = (t-\theta)^k Q(t)$ with Q a polynomial and $k \geq 2$. We have
$P'(t) = k(t-\theta)^{k-1}Q(t) + (t-\theta)^k Q'(t) = (t-\theta)^{k-1}(kQ(t) + (t-\theta)Q'(t)$
so $t-\theta$ is a factor of both $P(t)$ and $P'(t)$.

The converse is true If $P(t) = (t-\theta)Q(t)$ with Q a polynomial having $Q(\theta) \neq 0$ then

$$P'(t) = Q(t) + (t-\theta)^k Q'(t)$$

so $P'(\theta) \neq 0$.

(ii) Since $A^m = I$ the minimal polynomial m_A of A must divide $t^m - 1$. But $t^m - 1 = 0$ has no repeated roots (by part (i), or otherwise). Thus m_a has no repeated roots so A is diagonalisable.

If A is a real symmetric matrix then its eigenvalues are real so we have $A = LDL^{-1}$ with L invertible (indeed we may take L orthogonal) and D real diagonal

$$D^m = I$$

so $d_{jj}^m = 1$ and d_{jj} is real so $d_{jj} = \pm 1$ so $D^2 + I$ so $A^2 = I$.

Exercise 12.3.9

(i) Let ∂Q denote the degree of Q. If $\mathcal{P} \neq \{0\}$ then the set
$$E = \{\partial P \,:\, P \in \mathcal{P} \setminus \{0\}\}$$
is a non-empty subset of the positive integers and so has a least member N.

Let $Q_0 \in \mathcal{P}$ have degree N and leading coefficient a. If we set $P_0 = a^{-1}Q_0$, then P_0 is a monic polynomial of smallest degree in $\mathcal{P}$.

If $Q \in \mathcal{P}$ then $Q = SP + R$ with S, R polynomials and $\partial R < N$. Now $R = Q - SP_0 \in \mathcal{P}$ so by the definition of N, $R = 0$ and $Q = SP_0$.

(ii) Observe that $\mathcal{P}$ satisfies the hypotheses of (i).

If S divides each P_j then S divides each Q_jP_j and so divides P_0.

(iii) Take $P_j(t) = \sum_{j=1}^{r} Q_j(t) \prod_{i\neq j}(t-\lambda_i)^{m(i)}$ and apply (ii).

Exercise 12.3.11

Write $\alpha = \alpha_1 \oplus \alpha_2 \oplus \ldots \oplus \alpha_r$ The expansion
$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_r$$
with $\mathbf{u}_j \in U_j$ always exists a and is unique so
$$\alpha(\mathbf{u}) == \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_r u_r$$
is well defined.

If $\mathbf{u}$, $\mathbf{v} \in U$ and λ, μ then we can write
$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_r$$
$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_r$$
with $\mathbf{u}_j$, $\mathbf{v}_j \in U_j$. We then have
$$\lambda\mathbf{u} + \mu\mathbf{v} = (\lambda\mathbf{u}_1 + \mu\mathbf{v}_1) + (\lambda\mathbf{u}_2 + \mu\mathbf{v}_2) + \ldots + (\lambda\mathbf{u}_r + \mu\mathbf{v}_r)$$
and $\lambda\mathbf{u}_j + \mu\mathbf{v}_j \in U_j$ so
$$\begin{aligned}\alpha(\lambda\mathbf{u} + \mu\mathbf{v}) &= \alpha_1(\lambda\mathbf{u}_1 + \mu\mathbf{v}_1) + \alpha_2(\lambda\mathbf{u}_2 + \mu\mathbf{v}_2) + \ldots + \alpha_r(\lambda\mathbf{u}_r + \mu\mathbf{v}_r)\\ &= (\lambda\alpha_1\mathbf{u}_1 + \mu\alpha_1\mathbf{v}_1) + (\lambda\alpha_2\mathbf{u}_2 + \mu\alpha_2\mathbf{v}_2) + \ldots + (\lambda\alpha_r\mathbf{u}_1 + \mu\alpha_r\mathbf{v}_1)\\ &= \lambda(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_r\mathbf{u}_r) + \mu(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_r\mathbf{v}_r)\\ &= \lambda\alpha\mathbf{u} + \mu\alpha\mathbf{v}\end{aligned}$$
so α is linear.

EXERCISE 12.3.13

(i) We have

$$\iota = \sum_{j=1}^{r} Q_j(\alpha) \prod_{i\neq j} (\alpha - \lambda_j \iota)^{m(j)},$$

so, if $\mathbf{u} \in U$ we have

$$\mathbf{u} = \sum_{j=1}^{r} Q_j(\alpha) \prod_{i\neq j} (\alpha - \lambda_j \iota)^{m(j)} \mathbf{u} = sum_{j=1}^{r} \mathbf{u}_j$$

where

$$\mathbf{u}_j = Q_j(\alpha) \prod_{i\neq j} (\alpha - \lambda_j \iota)^{m(j)} \mathbf{u}.$$

Since

$$\begin{aligned}(\alpha - \lambda_j \iota)^{m(j)} \mathbf{u}_j &= (\alpha - \lambda_j \iota)^{m(j)} Q_j(\alpha) \prod_{i\neq j} (\alpha - \lambda_j \iota)^{m(j)} \mathbf{u} \\ &= Q_j(\alpha) \prod_{j=1}^{r} (\alpha - \lambda_j \iota)^{m(j)} \mathbf{u} \\ &= Q_j(\alpha) Q(\alpha) \mathbf{u} = Q_j(\alpha) \mathbf{0} = \mathbf{0}\end{aligned}$$

we have $\mathbf{u}_j \in U_j$. Thus

$$U = U_1 + U_2 + \ldots + U_r.$$

(ii) Next we observe that if $\mathbf{v} \in U_k$ with $k \neq j$ then

$$Q_j(\alpha) \prod_{i\neq j} (\alpha - \lambda_j \iota)^{m(j)} \mathbf{v} = \mathbf{0}$$

Thus if $\mathbf{v} \in U_k$ we have

$$\mathbf{v} = \sum_{j=1}^{r} Q_j(\alpha) \prod_{i\neq j} (\alpha - \lambda_i \iota)^{m(i)} \mathbf{v} = Q_k(\alpha) \prod_{i\neq k} (\alpha - \lambda_i \iota)^{m(i)} \mathbf{v}.$$

In particular $\prod_{i\neq k} (\alpha - \lambda_i \iota)^{m(i)} \mathbf{v} \neq \mathbf{0}$ if $\mathbf{v} \neq \mathbf{0}$.

Now if $v_j \in U_j$ and

$$\sum_{j=1}^{r} \mathbf{v}_j = \mathbf{0}$$

then

$$\prod_{i\neq k} (\alpha - \lambda_j \iota)^{m(i)} \mathbf{v}_k = \prod_{i\neq k} (\alpha - \lambda_j \iota)^{m(i)} \sum_{j=1}^{r} \mathbf{v}_j \mathbf{0}$$

so $\mathbf{v}_k = \mathbf{0}$ for each k.

Thus

$$U = U_1 \oplus U_2 \oplus \cdots \oplus U_r$$

(iii) We have

$$\mathbf{u} \in U_j \Rightarrow (\alpha - \lambda_j)^{m(j)}\mathbf{u} = \mathbf{0} \Rightarrow (\alpha - \lambda_j)^{m(j)}\alpha\mathbf{u} = \alpha(\alpha - \lambda_j)^{m(j)}\mathbf{u} = \mathbf{0} \Rightarrow \alpha\mathbf{u} \in U_j.$$

If $\mathbf{u} \in U$ then

$$\mathbf{u} = \sum_{j=1}^{r} \mathbf{u}_j$$

and

$$\alpha\mathbf{u} = \alpha\sum_{j=1}^{r} \mathbf{u}_j = \sum_{j=1}^{r} \alpha\mathbf{u}_j = \sum_{j=1}^{r} \alpha_j\mathbf{u}_j = \alpha = \alpha_1 \oplus \alpha_2 \oplus \cdots \oplus \alpha_r\mathbf{u}$$

so

$$\alpha = \alpha_1 \oplus \alpha_2 \oplus \cdots \oplus \alpha_r.$$

(iiv) By the definition of U_j, $(\alpha - \lambda_j\iota)^{m(j)}U_j = \mathbf{0}$ so the minimal polynomial for α_j must have the form. $(t - \lambda_j)^{p(j)}$. But If $\mathbf{u} \in U$ then $\mathbf{u} = \sum_{s=1}^{r} \mathbf{u}_s$ with $\mathbf{u}_s \in U_s$

$$\prod_{j=1}^{r}(\alpha - \lambda_j)^{p(j)}\mathbf{u} = \sum_{s=1}^{r}\prod_{j=1}^{r}(\alpha - \lambda_j)^{p(j)}\mathbf{u}_j = \sum_{s=1}^{r}\mathbf{0} = \mathbf{0}.$$

so the minimal polynomial must divide $\prod_{j=1}^{r}(\alpha - \lambda_j)^{p(j)}$ so $p(j) = m(j)$ for all J.

Exercise 12.3.14

Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$ be a basis for U with respect to which α has matrix A and $\mathbf{e}_{r+1}, \mathbf{e}_{r+2}, \ldots, \mathbf{e}_{r+s}$ be a basis for V with respect to which β has matrix B Then $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{r+s}$ is a basis for W with respect to which $\alpha \oplus \beta$ has matrix

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

Thus

$$\chi_{\alpha\oplus\beta}(t) = \det \begin{pmatrix} tI - A & 0 \\ 0 & tI - B \end{pmatrix} = \det(tI - A)\det(tI - B) = \chi_\alpha(t)\chi\beta(t).$$

If α has minimal polynomial m_α, β has minimal polynomial m_β and P is a polynomial

$$\begin{aligned} P(\alpha \oplus \beta) = 0 &\Leftrightarrow P(\alpha \oplus \beta)(\mathbf{u} + \mathbf{v}) = \mathbf{0} \forall \mathbf{u} \in U,\ \mathbf{v} \in V \\ &\Leftrightarrow P(\alpha)\mathbf{u} + P(\beta)\mathbf{v} = \mathbf{0} \forall \mathbf{u} \in U,\ \mathbf{v} \in V \\ &\Leftrightarrow P(\alpha)\mathbf{u} = P(\beta)\mathbf{v} = \mathbf{0} \forall \mathbf{u} \in U,\ \mathbf{v} \in V \\ &\Leftrightarrow m_\alpha \text{ and } m_\beta \text{ divide } P \end{aligned}$$

Thus the minimal polynomial of $\alpha \oplus \beta$ is the lowest common multiple of (ie the monic polynomial of lowest degree dividing) m_α and m_β.

Exercise 12.3.15

(i) If $(t-\lambda)$ is a factor of the characteristic polynomial then λ is an eigenvalue and $(t-\lambda)$ must be a factor of the minimal polynomial. Thus we cannot choose $S(t) = t+1$, $Q(t) = t$.

(ii) If $m = 1$ set $A = 0$. If $m = n$ let $A = (a_{ij})$ be the $n \times n$ matrix with $a_{i,i+1} = 1$ $a_{ij} = 0$ otherwise.

If $n > m > 1$ let $B = 0_{n-m}$ the $n - m + \times n - m$ zero matrix and let C be the $m \times m$ matrix with $c_{i,i+1} = 1$ $c_{ij} = 0$ otherwise. Take

$$A = \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$$

$A - \lambda I$ has characteristic equation $(t-\lambda)^n$ and minimal polynomial $(t-\lambda)^m$

(iii) We can write $P(t) = \prod i = 1^r (t - \lambda_i)^{m(i)}$ and $Q(t) = \prod i = 1^r (t - \lambda_i)^{m(i)}$ with the λ_j distinct and $m(j) \geq p(j) \geq 1$. By (ii) we can find a $m(j) \times m(j)$ matrix $B_j = A_j - \lambda I_{m(j)}$ with characteristic polynomial $(t-\lambda)^{m(j)}$ and minimal polynomial $(t-\lambda)^{p(j)}$. Take

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 & 0 & \\ 0 & A_2 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \\ 0 & 0 & 0 & \dots & A_{r-1} & 0 \\ 0 & 0 & 0 & \dots & 0 & A_r \end{pmatrix}$$

A has the required properties either by direct computation or using Exercise 12.3.14.

Exercise 12.4.3

Since $\alpha \mathbf{e}_j = \mathbf{e}_{j+1}$ for $1 \leq j \leq n-1$ and $\alpha \mathbf{e}_n = \mathbf{0}$ we have $A = (a_{ij})$ with $a_{i,i+1} = 1$ $a_{ij} = 0$ otherwise.

Exercise 12.4.4

If $\alpha^n \neq 0$, then $\mathbf{e}$, $\alpha\mathbf{e}$, $\ldots \alpha^n\mathbf{e}$ are linearly independent a and a space of dimension n would contain $n+1$ linearly independent vectors which is impossible.

Exercise 12.4.7

We have $\alpha^k \mathbf{e} \neq \mathbf{0}$ for only finitely many values of k.

Exercise 12.4.11

Observe that (d_{jj}) is a 1×1 Jordan matrix.

Exercise 12.4.12

(i) A is the matrix of a linear map $\alpha : U \to U$

$$\alpha = \alpha_1 \oplus \alpha_2 \dots \alpha_r$$

with $U = U_1 \oplus U_2 \dots U_r$ where α_j is a linear map on U_j with $\alpha_j = \lambda_j \iota + \beta_j$ with β_j a nilpotent linear map on U_j with $\beta_j^{k_j} = 0$, $\beta^{k_j - 1} \neq 0$ $\dim U_j = k_j$ Thus

$$(\lambda\iota - \alpha)^{-k}\mathbf{0} = \bigoplus_{j=1}^{r} (\lambda\iota - \alpha_j)^{-k}\mathbf{0}$$

and

$$\dim(\lambda\iota - \alpha)^{-k}\mathbf{0} = \sum_{j=1}^{r} \dim(\lambda\iota - \alpha_j)^{-k}\mathbf{0} = \sum_{\lambda=\lambda_j} \min\{k, k_j\}.$$

Thus

$$\{\mathbf{x} \in \mathbb{C}^n \,:\, (\lambda I - A)^k \mathbf{x} = \mathbf{0}\} = \sum_{\lambda=\lambda_j} \min\{k, k_j\}$$

(ii) By (i)

$$\sum_{\lambda=\lambda_j} \min\{k, k_j\} = \sum_{\lambda=\tilde{\lambda}_i} \min\{k, \tilde{k}_i\}$$

so $\tilde{r} = r$ and, possibly after renumbering, $\tilde{\lambda}_j = \lambda_j$ and $\tilde{k}_j = k_j$ for $1 \leq j \leq r$.

Exercise 12.4.13

Observe by looking at the effect on z^n that $T^{n+1} = 0$ but $T^n \neq 0$. Thus, since $\mathcal{P}_n$ has dimension $n+1$, T has Jordan form

$$J_{n+1}(0) = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

We observe that $e_r(z) = z^r$ is an eigenvector with eigenvalue r [$n \geq r \geq 0$]. Thus the Jordan normal form is the diagonal matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 2 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & n-1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & n \end{pmatrix}.$$

Exercise 12.4.14

(i) If λ is a root of χ_α then we can find an associated eigenvector $\mathbf{e}$. Since

$$\mathbf{e} \in \dim\{\mathbf{u} : (\alpha - \lambda\iota)(\mathbf{u}) = \mathbf{0}\}$$

we have $m_g(\lambda) \geq 1$.

Now take a basis $\mathbf{e}_1$, $\mathbf{e}_2$ ..., $\mathbf{e}_{m_g(\lambda)}$ for

$$\{\mathbf{u} : (\alpha - \lambda\iota)(\mathbf{u}) = \mathbf{0}\}$$

and extend it to a basis $\mathbf{e}_1$, $\mathbf{e}_2$ ..., $\mathbf{e}_n$ for U. With respect to this basis α has matrix

$$A = \begin{pmatrix} \lambda I_{m_g(\lambda)} & 0 \\ R & S \end{pmatrix}$$

with $I_{m_g(\lambda)}$ the $m_g(\lambda) \times (m_g(\lambda)$ identity matrix so

$$\chi_\alpha(t) = (t - \lambda)^{m_g(\lambda)} \chi_R(t)$$

and $m_a(\lambda) \geq m_g(\lambda)$.

(ii) Suppose $s \geq r \geq 1$ Let $B = (b_{ij})$ be the $r \times r$ matrix with $b_{i,i+1} = 1$ and $b_{ij} = 0$ and $A(r, s)$ the $s \times s$ matrix with

$$A(r, s) = \begin{pmatrix} \lambda B & 0 \\ 0 & 0 \end{pmatrix}.$$

If $\alpha_{r,s}$ is the linear map associated with $A(r, s)$ then, for this linear map,

$$m_g(\lambda) = \begin{cases} r & \text{if } \lambda = 0 \\ = 0 & \text{otherwise} \end{cases}, \quad m_a(\lambda) = \begin{cases} s & \text{if } \lambda = 0 \\ = 0 & \text{otherwise} \end{cases}$$

If we now set $\beta_{r,s} = \mu\iota - \alpha_{r,s}$ then for this linear map

$$m_g(\lambda) = \begin{cases} r & \text{if } \lambda = \mu \\ = 0 & \text{otherwise} \end{cases}, \quad m_a(\lambda) = \begin{cases} s & \text{if } \lambda = \mu \\ = 0 & \text{otherwise} \end{cases}$$

Now take a vector space

$$U = U_1 \oplus U_2 \oplus \ldots \oplus U_r$$

with $\dim U_k = n_a(\lambda_k)$. By the previous paragraph we can find $\gamma_k : U_k \to U_k$ linear such that for this linear map

$$m_g(\lambda) = \begin{cases} n_g(\lambda_k) If & \text{if } \lambda = \lambda_k \\ = 0 & \text{otherwise} \end{cases}, \quad n_a(\lambda) = \begin{cases} s & \text{if } \lambda = \lambda_k \\ = 0 & \text{otherwise} \end{cases}$$

If we set

$$\alpha = \gamma_1 \oplus \gamma_2 \oplus \ldots \oplus \gamma_r$$

then the required properties can be read off.

(iii) Suppose α has the associated Jordan form

$$A = \begin{pmatrix} J_{k_1}(\lambda_1) & & & & & \\ & J_{k_2}(\lambda_2) & & & & \\ & & J_{k_3}(\lambda_3) & & & \\ & & & \ddots & & \\ & & & & J_{k_{r-1}}(\lambda_{r-1}) & \\ & & & & & J_{k_r}(\lambda_r) \end{pmatrix}$$

(using the notation of Theorem 12.4.10).

We observe that the characteristic polynomial

$$\chi_\alpha(t) = \chi_A(t) = \prod_{j=1}^{r} \chi_{J_{k_j}}(t) = \prod_{j=1}^{r} (t - \lambda_j)^{k(j)}$$

so

$$m_a(\lambda) = \sum_{\lambda_j = \lambda} k(j).$$

Observe that the dimension of the space of solutions of

$$J_k \mathbf{x} = \lambda \mathbf{x}$$

is 1 if $\lambda = \lambda_k$ and zero otherwise we see that

$$m_a(\lambda) = \sum_{\lambda_j = \lambda} 1 = \operatorname{card}\{ : \lambda_j = \lambda\}$$

Exercise 12.5.1

Let U be a finite dimensional vector space over $\mathbf{C}$. If $\alpha \in \mathcal{L}(U,U)$ then the Jordan normal form theorem tells that we can write $U = \bigoplus j = 1^r U_j$ and find $\alpha_j \in \mathcal{L}(U_j, U_j)$ such that $\alpha = \bigoplus j = 1^r \alpha_j$ and U_j has a basis with respect to which α_j has an $n(j) \times n(j)$ matrix

$$A_j = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

Now

$$\det(t\iota - \alpha) = \prod_{j=1}^{r} \det(t\iota - \alpha_j) \prod_{j=1}^{r} (t - \lambda_j)^{n(j)}$$

and

$$\prod_{j=1}^{r} (\alpha - \lambda_j)^{n(j)} = \bigoplus j = 1^r \prod_{i=1}^{r} (\alpha_j - \lambda_i)^{n(j)} = 0$$

(since $(\alpha_j - \lambda_j)^{n(j)} = 0$) so $\chi_\alpha(\alpha) = 0$.

Lemma 12.2.1 tells us that there exists a non-zero polynomial P with $P(\alpha) = 0$. This is all we need to show that there is a minimal polynomial. The proof of the Jordan form theorem only uses the minimal polynomial.

Exercise 12.5.2

Consider the $s \times s$ matrix

$$B = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

If $\lambda \neq 0$, B is invertible so $\operatorname{rank} B^j = s$ for all $j \geq 0$. If $\lambda = 0$ then direct calculation shows that $\operatorname{rank} B^j = \max\{0, k-j\}$ for all $j \geq 0$.

If $\alpha = \prod_{k=1}^r \alpha_k$ then $\operatorname{rank} \alpha^j = \sum_{k=1}^r \operatorname{rank} \alpha_k^j$

Thus if the Jordan normal form of an $n \times n$ A contains m nilpotent blocks

$$K_p = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

with K_p an $n(p) \times n(p)$ matrix, then

$$rank A^j = \left(n - \sum_{p=1}^m n_p\right) + \sum_{p=1}^m \max\{0, n_p - j\}$$

. (ii) Observe that if we set

$$r(j) = \left(n - \sum_{p=1}^m n_p\right) + \sum_{p=1}^m \max\{0, n_p - j\}$$

then for some $m = \max_p n(p) =\leq n$ together with the condition

$$r_0 - r_1 \geq r_1 - r_2 \geq r_2 - r_3 \geq \dots \geq r_{m-1} - r_m.$$

Let $q(j) = (s_j - s_{j-1}) - (s_{j+1} - s_j)$. Let A be the $n \times n$ matrix with qj nilpotent Jordan blocks of size $j \times j$ and $s(m)$ 1×1 Jordan blocks (1). If α has matrix A with respect to some basis then α satisfies the required conditions.

Exercise 12.5.3

(i) We have

$$A_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\ A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\ A_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},\ A_5 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

(ii) For A_1 $\dot{x}_r = 0$ so $\mathbf{x}(t) = \mathbf{c}$ a constant.

For A_2 $\dot{x}_1 = x_2$, $\dot{x}_r = 0$ otherwise so $\mathbf{x}(t) = (c_2t + c_1, c_2, c_3, c_4)^T$ with c_j constants.

For A_3 $\dot{x}_1 = x_2$, $\dot{x}_2 = x_3$, $\dot{x}_r = 0$ otherwise so $\mathbf{x}(t) = (c_3t^2/2 + c_2t + c_1, c_2t + c_3, c_3, c_4)^T$ with c_j constants.

Similarly for A_4, $\mathbf{x}(t) = (c_4t^3/6 + c_3t^2/2 + c_2t + c_1, c_4t^2/2 + c_3t + c_2, c_4t + c_3, c_4)^T$ with c_j constants.

Similarly for A_5 $\mathbf{x}(t) = (c_2t + c_1, c_2, c_4t + c_3, c_4)^T$

(iii) If we write $\mathbf{x}(t) = e^{\lambda}\mathbf{y}(t)$ then

$$\mathbf{x}'(t) = A_j\mathbf{x}(t)$$

Thus the general solution of

$$\mathbf{y}'(t) = (\lambda I + A_j)\mathbf{y}(t)$$

is $\mathbf{y}(t) = e^{-\lambda t}\mathbf{x}(t)$ where $\mathbf{x}(t)$ is the general solution of

$$\mathbf{x}'(t) = A\mathbf{x}(t).$$

(iv) Suppose

$$B = \begin{pmatrix} B_1 & 0 & \dots & 0 & 0 \\ 0 & B_2 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & B_{r-1} & 0 \\ 0 & 0 & \dots & 0 & B_r \end{pmatrix}$$

with $B_j = \lambda_j I + C_j$ where C_j is an $s(j) \times s(j)$ nilpotent Jordan matrix. Then the general solution of

$$\mathbf{x}'(t) = B\mathbf{x}(t)$$

is $\mathbf{x} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_r)^T$ with

$$\mathbf{y}_j(t) = e^{-\lambda_j t}(P_j^{(s(j))}(t), P_j^{(s(j)-1)}(t), \dots, P_j(t))^T$$

with P_j any polynomial of degree $s(j) - 1$.

(iv) Find M invertible and B a matrix in Jordan form such that

$$MAM^{-1} = B$$

then writing $\mathbf{y} = M\mathbf{x}$ we have

$$\bigstar \qquad \dot{\mathbf{y}} = B\mathbf{y}.$$

The general solution of

$$\mathbf{x}'(t) = A\mathbf{x}(t).$$

is $\mathbf{x} = M^{-1}\mathbf{y}$ where $\mathbf{y}$ is the general solution of $\bigstar$.

Exercise 12.5.4

Observe that

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ & \vdots & & & \vdots & \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_{n-1} & -a_{n-2} & -a_{n-3} & \dots & -a_1 & -a_0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{pmatrix}$$

$$\Leftrightarrow \dot{x}_j = x_{j+1}\ [0 \leq j \leq n-2], -\sum_{j=0}^{n-2} a_j x_j = x_{n-1}$$

$$\Leftrightarrow x^{(n)} + \sum_{j=1}^{n-2} a_j x^{(n-j)} = 0$$

Exercise 12.5.5

Suppose $U = \bigoplus_{j=1}^{r} U_j$, U_j has dimension k_j $\beta_j : U_j \to U_j$ has $\beta_j^{k(j)-1} \neq 0$, $\beta_j^{k_j} = 0$, $\alpha_j = \lambda_j \iota + \beta_j$ and $\alpha = \bigoplus_{j=1}^{r} \alpha_j$.

Write $s(\lambda) = \max_{\lambda=\lambda_j} k_j$. If P is a polynomial

$$P(\alpha) = 0 \Leftrightarrow P(\alpha_j) = 0 \ \forall j \leftrightarrow (t-\lambda)^s(\lambda)|P(t).$$

Thus the characteristic polynomial equals the minimal polynomial if and only if

$$\prod_{j=1}^{r}(t-\lambda)^{k_j} = \prod_{s(\lambda)\neq 0}(t-\lambda)^{s(\lambda)}$$

that is to say, if and only if all the λ_j are distinct.

Exercise 12.5.6

Writing $\mathbf{e}_j$ for the vector with 1 in the jth place and zero everywhere else. By induction

$$A^j\mathbf{e}_n = \mathbf{e}_{n-j} + \mathbf{f}_{n-j}$$

where

$$\mathbf{f}_{n-j} \in \operatorname{span}\{\mathbf{e}_r : n \geq r \geq n-j+1\}.$$

Thus the $A^j\mathbf{e}_n$ with $n-1 \geq j \geq 0$ are linearly independent and

$$\left(\sum_{j=0}^{n-1} b_j A^j\right)\mathbf{e} = \mathbf{0} \Leftrightarrow b_j = 0 \ \forall 0 \leq j \leq n-1.$$

Thus the minimal polynomial has degree at least n. Since the minimal polynomial divides the characteristic polynomial and the characteristic polynomial has degree n it follows that the two polynomials are equal. By Exercise 12.5.5, it follows that that there is a Jordan form associated with A in which all the blocks $J_k(\lambda_k)$ have distinct λ_k.

Exercise 12.5.7

(i) If A is an $n \times n$ matrix consisting of a single Jordan block with associated eigenvalue λ and

$$\dot{\mathbf{x}} = A\mathbf{x}$$

then writing $\mathbf{y}(t) = e^{-\lambda t}\mathbf{x}(t)$ we have

$$\dot{\mathbf{y}}(t) = -\lambda e^{-\lambda t}\mathbf{x}(t) + e^{-\lambda t}\dot{\mathbf{x}}(t) = -\lambda e^{-\lambda t}\mathbf{x}(t) + e^{-\lambda t}A\mathbf{x}(t) = B\dot{\mathbf{y}}(t)$$

where B is a single Jordan block with associated eigenvalue 0. Thus

$$y^{(n)} = 0$$

and $y(t) = \sum_{j=0}^{n-1} c_j t^j$ for some $c_j \in \mathbb{R}$.

Conversely, if $y(t) = \sum_{j=0}^{n-1} c_j t^j$, for some $c_j \in \mathbb{R}$, then

$$\dot{\mathbf{y}}(t) = B\dot{\mathbf{y}}(t).$$

Thus

$$x(t) = e^{\lambda t}\sum_{j=0}^{n-1} c_j t^j$$

where the $c_j \in \mathbb{R}$ are arbitrary.

(ii) Suppose that A is associated with a Jordan form B with blocks the $n_k \times n_k$ Jordan block $J_k(\lambda_k)$ with all the $lambda_k$ distinct $[1 \leq k \leq r]$. Writing $A = MBM^{-1}$ with M non-singular we see that if $\mathbf{y}$ is a solution of $\dot{\mathbf{y}} = B\mathbf{y}$ if and only $\mathbf{x} = M\mathbf{y}$ is a solution of $\dot{\mathbf{x}} = A\mathbf{x}$. Thus the only possible solutions of ★ are

$$x(t) = \sum_{k=1}^{r} e^{\lambda_k t}P_k(t).$$

with the P_k general polynomials of degree $n_k - 1$.

To show that every x of this form is a solution observe that (by linearity) it suffices to show that

$$e^{\lambda_k t}P_k(t)$$

is a solution with the P_k a general polynomials of degree $n_k - 1$. By considering $e^{-\lambda_k t}x(t)$ we see that it suffices to show that if the characteristic equation of A has 0 as an p times repeated root the every polynomial of degree $r - 1$ satisfies ★. But if A has 0 as an p times repeated root then $a_0 = a_1 = \ldots = a_{r-1} = 0$ so the previous sentence is automatic.

Exercise 12.5.8

(i) We have

$$\binom{r}{k-1}+\binom{r}{k} = k\frac{r!}{k!(r+1-k)!}+(r+1-k)k\frac{r!}{k!(r+1-k)!} = (r+1)\frac{r!}{k!(r+1-k)!} = \binom{r+1}{k}$$

(ii) If $k=0$ our system becomes

$$u_r(0) = u_{r-1}(0)$$

with the general solution $u_r(0) = b_0$.

If $k=1$ we have $u_r(0) = b_0$ as before and

$$u_r(1) - u_{r-1}(1) = b_0$$

If we set

$$v_r(1) = u_r(1) - b_0 r$$

we obtain

$$v_r(1) - v_{r-1}(1) = 0$$

so as before $v_r(1) = b_1$ for some freely chosen b_1. Thus $u_r(1) = b_0 r + b_1$.

Suppose that the solution for a given $k \geq 0$ is

$$u_r(k) = \sum_{j=0}^{k} b_{k-j}\binom{r}{j}.$$

Then

$$u_r(k+1) - u_{r-1}(k+1) = \sum_{j=0}^{k} b_j\binom{r}{j},$$

and setting

$$v_r(k+1) = u_r(k+1) - \sum_{j=0}^{k} b_{k-j}\binom{r}{j+1},$$

we obtain

$$v_r(k+1) - v_{r-1}(k+1) = 0$$

so, as before, $v_r(k+1) = b_{k+1}$ for some freely chosen b_{k+1}. Thus

$$u_r(k+1) = \sum_{j=0}^{k+1} b_{k+1-j}\binom{r}{j}.$$

Thus, by induction,

$$u_r(k) = \sum_{j=0}^{k} b_{k-j}\binom{r}{j}.$$

(ii) If we set $u_r(k) = \lambda^{-r} v_r(k)$ then the $\mathbf{u}(k)$ satisfy the equations in (i). Thus

$$v_r(k) = \lambda^r \sum_{j=0}^{k} b_{k-j} \binom{r}{j}.$$

(iii) The equation

$$u_r + \sum_{j=0}^{n-1} a_j u_{j-n+r} = 0$$

may be rewritten as

$$\mathbf{u}_{t+1} + A\mathbf{u}_t = \mathbf{0}$$

where $\mathbf{u}_t$ is the column vector $(u_t, u_{t+1}, \ldots, u_{t+n-1})$ and

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ & \vdots & & & \vdots & \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$

By Exercise 12.5.6 we know that we can associate A with a Jordan form B in which in which the q blocks $J_k(\lambda_k)$ are of size $s_k \times s_k$ and have distinct λ_k.

Let $B = M^{-1}AM$ with M invertible.

By (ii) the general solution of

$$\mathbf{v}_{t+1} + B\mathbf{v}_t = \mathbf{0}$$

is

$$\mathbf{v}_t \sum_{k=1}^{q} \lambda_k^t \mathbf{e}_k \sum_{j=0}^{s_k} b_{k,s_k-j} \binom{r}{j}$$

so

$$u_t = \sum_{k=1}^{q} \lambda_k^t \mathbf{e}_k \sum_{j=0}^{s_k} c_{k,s_k-j} \binom{r}{j}.$$

We have shown that u_t must be of the form just given. Direct computation shows that any u_t of the form just given is a solution.

Exercise 12.5.9

All roots characteristic equation the same.

$$\begin{pmatrix}\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda\end{pmatrix}, \begin{pmatrix}\lambda & 1 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda\end{pmatrix}, \begin{pmatrix}\lambda & 1 & 0\\ 0 & \lambda & 1\\ 0 & 0 & \lambda\end{pmatrix}.$$

Characteristic equation has two distinct roots

$$\begin{pmatrix}\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \mu\end{pmatrix}, \begin{pmatrix}\lambda & 1 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \mu\end{pmatrix}$$

with $\lambda \neq \mu$. Characteristic equation has three distinct roots

$$\begin{pmatrix}\lambda & 0 & 0\\ 0 & \mu & 0\\ 0 & 0 & \nu\end{pmatrix}, \begin{pmatrix}\lambda & 1 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \mu\end{pmatrix}$$

with λ, μ, ν distinct.

Exercise 12.5.10

We have $A^4 - A^2 = 0$ so the minimal polynomial m must divide $t^4 - t^2 = t^2(t-1)(t+1)$

Since $(A - I)A = A^2 - A \neq 0$ the minimal polynomial cannot be 1, t, $t-1$ or $t(t-1)$

(i) $m(t) = t^2$. The characteristic polynomial must be t^5. The Jordan blocks must be of size 2×2 or less. There are *three* possibilities

$$0, \begin{pmatrix} J & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} J & 0 & 0 \\ 0 & J & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ with } J = \begin{pmatrix} J & 0 \\ 0 & 0 \end{pmatrix}$$

(ii) $m(t) = t + 1$. The Jordan normal form is $-I$ giving *one* possibility.

(iii) $m(t) = t^2 - 1$. The Jordan normal form is diagonal with r entries 1, $5 - r$ entries -1 $[1 \leq r \leq 4]$ and the characteristic equation is $(t-1)^r(t+1)^s$ There are *four* possibilities.

(iii) $m(t) = t(t+1)$. The Jordan normal form is diagonal with r entries -1, $5-r$ entries -0 $[1 \leq r \leq 4]$ and the characteristic equation is $(t+1)^r t^s$ There are *four* possibilities.

(iv) $m(t) = t^2(t-e)$ with $e = \pm 1$. The Jordan normal form either has one block J and then r diagonal entries e and $3 - r$ diagonal entries 0 $[1 \leq r \leq 3]$ giving a characteristic polynomial $t^{5-r}(t-e)^r$ or two blocks J and a diagonal entry e giving a characteristic polynomial $t^4(t-e)$ There are *eight* possibilities.

(v) $m(t) = t(t-1)(t+1)$. (iii) $m(t) = t^2 - 1$. The Jordan normal form is diagonal with r entries 1, s entries -1, u entries 0 $[r,\, s,\, u \geq 1,\ r + s + u = 5]$ and the characteristic equation is $(t-1)^r(t+1)^s t^u$ There are *six* possibilities.

(vi) $m(t) = t^2(t-1)(t+1)$. The Jordan normal form has one block J and r diagonal entries 1, s diagonal entries -1 $[r,\, s \geq 1,\ r + s \leq 3]$ and $3 - r - s$ diagonal entries 0. The characteristic equation is $(t-1)^r(t+1)^s t^{5-r-s}$. There are *three* possibilities.

There are 29 possible Jordan forms.

Exercise 12.5.11

$$\det(tI - M) = (t-1)\det\begin{pmatrix} t-1 & 0 & 0 \\ 1 & t-2 & 0 \\ 0 & 0 & t-2 \end{pmatrix} = (t-1)^2(t-2)^2$$

Looking at the eigenvalue 1 we have

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

implies

$$\begin{aligned} x + z &= x \\ y &= y \\ -y + 2z &= z \\ 2w &= w \end{aligned}$$

so

$$M(x, y, z, w)^T = (x, y, z, w)^T \Leftrightarrow (x, y, z, w) = (t, 0, 0, 0)$$

for general t.

Thus $\mathbf{e}_2 = (1, 0, 0, 0)^T$ is an eigenvector eigenvalue 1. If $(M-I)\mathbf{x} = \mathbf{e}$ then

$$\begin{aligned} z &= 1 \\ 0 &= 0 \\ -y + z &= 0 \\ w &= 0 \end{aligned}$$

Thus writing $\mathbf{e}_1 = (0, 1, 1, 0)^T$ we have $(M-I)\mathbf{e}_1 = \mathbf{e}_2 \neq \mathbf{0}$ $(M-I)^2\mathbf{e}_1 = \mathbf{0}$.

Looking at the eigenvalue 2 we have

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = 2\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$$

implies

$$\begin{aligned} x + z &= 2x \\ y &= 2y \\ -y + 2z &= 2z \\ 2w &= 2w \end{aligned}$$

so

$$M(x,y,z,w)^T = (x,y,z,w)^T \Leftrightarrow (x,y,z,w) = (t,t,0,s)$$

for general t and s. Thus the eigenspace has dimension 2 and basis $(1,1,0,0)^T$, $(0,0,0,1)^T$.

The characteristic polynomial is $(t-1)^2(t-2)^2$ minimal polynomial is $(t-1)^2(t-2)$. A Jordan form is given by

$$J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

An appropriate basis is $(0,1,1,0)^T$, $(1,0,0,0)^T$ $(1,1,0,0)^T$, $(0,0,0,1)^T$ and we can take M to be the matrix with these columns, that is to say

$$M = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Exercise 13.2.2

(i) The tables are

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

By inspection of the addition table $x + x = 0$ for all x.

(ii) $x = -x \Rightarrow (1 + 1)x = 0 \Rightarrow x = 0$.

Exercise 13.2.3

(i) If $c \neq 0$ then

$$0 = c^{-1}(cd) = (c^{-1}c)d = d$$

(ii) If $a^2 = b^2$ then $(a+b)(a-b) = a^2 - b^2 = 0$ so $a + b = 0$ and/or $b - a = 0$ i.e. $a = -b$ and or $a = b$.

(iii) By (ii) (and the fact that $1 + 1 \neq 0$, the equation $x^2 = c$ has no solutions or two solutions or $c = 0$ and we have $x = 0$. Thus every element y is a square root (just look at y^2) but every square has two distinct square roots or is 0. Thus k is odd, there are $(k-1)/2$ non zero squares and $k+1)/2$ squares in all.

(iv) If we work in $\mathbf{Z}_2$, then $0^2 = 0$ and $1^2 = 1$. Every element is a square. (Note $-1 = 1$)

Exercise 13.2.4

Suppose $x,\ y,\ z \in \mathbf{R}$ $\lambda,\ \mu \in \mathbb{Q}$.

(i) $(x\dot{+}y)\dot{+}z = (x+y)+z = x+(y+z) = (x\dot{+}y)\dot{+}z = x\dot{+}(y\dot{+}z)$.

(ii) $x\dot{+}y = x+y = y+x = y\dot{+}x$.

(iii) $x\dot{+}0 = x+0 = x$.

(iv) $\lambda\dot{\times}(x\dot{+}y) = \lambda(x+y) = \lambda x + \lambda y = \lambda\dot{\times}x\dot{+}\lambda\dot{\times}y$.

(v) $(\lambda+\mu)\dot{\times}x = (\lambda+\mu)x = \lambda x + \mu x = \lambda\dot{\times}x\dot{+}\mu\dot{\times}y$.

(vi) $(\lambda\mu)\dot{\times}x = (\lambda\mu)x = \lambda(\mu x) = \lambda\dot{\times}(\mu\dot{\times}x)$.

(vii) $1\dot{\times}x = 1x = x$ and $0\dot{\times}x = 0x = 0$.

Exercise 13.2.6

(i) If
$$MAM^{-1} = D \text{ with } D = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$
then
$$t^2 - a = \det(tI - A) = \det(tI - D) = (t - \lambda)(t - \mu)$$
so $t^2 = a$ has a solution.

Suppose $c^2 = a \neq 0$. The eigenvalues are $\pm c$ which are distinct so by our standard arguments A is diagonalisable.

(ii) Our standard argument shows that A is not diagonalisable.

(iii) We know that 2 has no square root in $\mathbf{Q}$ so by (i), no M exists.

(iv) Since some a do not have square roots the corresponding a are not diagonalisable.

(v) If $A = M^{-1}DM$ with D diagonal then $I = A^2 = M^{-1}D^2M$ so $D^2 = I$. Thus (since we work in $\mathbb{Z}_2$, $D = I$ so $A = I$ which is absurd.

If we set $\mathbf{e}_1 = (1, 1)^T$ and $\mathbf{e}_2 = (1, 0)^T$ we have a basis and $A\mathbf{e}_1 = \mathbf{e}_1$, $A\mathbf{e}_1 = \mathbf{e}_1 + \mathbf{e}_2$, Thus setting
$$M = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$
we have
$$MAM^1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Exercise 13.2.7

The tables are

$+$	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$(0,0)$	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$(1,0)$	$(1,0)$	$(0,0)$	$(1,1)$	$(0,1)$
$(0,1)$	$(0,1)$	$(1,1)$	$(0,0)$	$(1,0)$
$(1,1)$	$(1,1)$	$(0,1)$	$(1,0)$	$(0,0)$

$\times$	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$
$(1,0)$	$(0,0)$	$(1,0)$	$(1,1)$	$(0,1)$
$(0,1)$	$(0,0)$	$(0,1)$	$(1,1)$	$(1,0)$
$(1,1)$	$(1,0)$	$(1,1)$	$(1,0)$	$(0,1)$

We check the axioms for a field.

(i) $(a,b)+(c,d)=(a+c,b+d)=(c,d)+(a,b)$.

(ii) We have

$$\begin{aligned}((a,b)+(c,d))+(e,f)&=(a+c,b+d)+(e,f)\\&=\big((a+c)+e,(b+d)+f\big)\\&=\big(a+(c+e),b+(d+f)\big).\end{aligned}$$

(iii) $(a,b)+(0,0)=(a,b)$

(iv) $(a,b)+(a,b)=(0,0)$

(v) We have

$$\begin{aligned}(a,b)\times(c,d)&=(ac+bd,ad+bc+bd)\\&=(ca+db,da+cb+db)=(c,d)\times(a,b).\end{aligned}$$

(vi) We have

$$\begin{aligned}&\big((a,b)\times(c,d)\big)\times(e,f)=(ac+bd,ad+bc+bd)\times(e,f)\\&=\big((ac+bd)e+(ad+bc+bd)f,\\&\qquad\qquad(ac+bd)f+(ad+bc+bd)e+(ad+bc+bd)f\big)\\&=\big(a(ce+df)+b(de+cf+df),\\&\qquad\qquad a(de+cf+df)+b(ce+df)+b(de+cf+df)\big)\\&=(a,b)\times\big((c,d)\times(e,f)\big)\end{aligned}$$

(vii) $(a,b)\times(1,0)=(a,b)$.

(viii) By inspection of the multiplication table we see that if $(a,b)\neq(0,0)$, we can find (c,d) with $a\times(a,b)\times(c,d)=(1,0)$.

(ix) We have

$$\begin{aligned}(a,b)\times\big((c,d)+(e,f)\big)&=(a,b)\times(c+e,d+f)\\&=\big(a(c+e)+b(c+e),a(d+f)+b(c+e)+b(d+f)\big)\\&=(ac+bd,ad+bd+bc)+(ae+bf,af+bf+be)\\&=(a,b)\times(c,d)+(a,b)\times(e,f)\end{aligned}$$

(x) By inspection of the multiplication table, all the elements are squares.

Exercise 13.2.8

(i) Suppose $2 \neq 0$.

If $A = B + C$ with $B^T = B$ and $C^T = -C$ then

$$2B = B + C + B - C = (B + C) + (B + C)^T = A + A^T$$

so $B = 2^{-1}(A + A^T)$ and $C = A - B = 2^{-1}(A - A^T)$.

Conversely if $B = 2^{-1}(A + A^T)$ and $C = A - B = 2^{-1}(A - A^T)$ then

$$\begin{aligned} B^T &= 2^{-1}(A^T + A^{TT}) = B \\ C^T &= 2^{-1}(A^T - A^{TT}) = -C \\ B + C &= 2^{-1}(A + A^T) + 2^{-1}(A - A^T) = A \end{aligned}$$

(ii) Suppose $2 = 0$. Then $a + a = (1 + 1)a = 2a = 0a = 0$ so $a = -a$ and $(a_{ij}) = (-a_{ij})$ so $A = -A$. Thus

$$A = A^T \Leftrightarrow A = -A^T.$$

We work over $\mathbf{Z}_2$. If $A = B+C$ with $B^T = B$, $C^T = C$ then $A^T = A$. Thus

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

cannot be written as the sum of a symmetric and an antisymmetric matrix.

On the other hand the zero matrix 0 and the identity matrix I are symmetric so antisymmetric and

$$0 = 0 + 0 = I + I.$$

Exercise 13.2.9

(i) Observe that $0^2 + 0 = 0 + 0$ and $1^2 + 1 = 1 + 1 = 0$ so $x^2 = x = 0$ for all $x \in \mathbb{Z}_2$.

(ii) If $\mathbb{G}$ is a finite field there are only a finite number of distinct functions $f : \mathbb{G} \to \mathbb{G}$ but there are infinitely many polynomials. Thus there must exist distinct polynomials P and Q which give rise to the same function. Consider $P - Q$.

Exercise 13.3.1

If U is a vector space over $\mathbb{G}$ then $W \subseteq U$ is a subspace if and only if.

(i) $\mathbf{0} \in W$.

(ii) $\mathbf{x}, \mathbf{y} \in W \rightarrow \mathbf{x} = \mathbf{y} \in W$.

(iii) $\lambda \in \mathbb{G},\ \mathbf{x} \in W \Rightarrow \lambda\mathbf{x} \in W$.

If $\mathbb{G} = \mathbb{Z}_2$ then, if $\lambda \in \mathbb{G},\ \mathbf{x} \in W$ either $\lambda = 0$ and $\lambda\mathbf{x} = \mathbf{0}$ so, if condition (ii) is true $\lambda\mathbf{x} \in W$, or $\lambda = 1$ and $\lambda\mathbf{x} = \mathbf{x} \in W$.

Thus we can drop condition (iii) in this case.

Exercise 13.3.2

If U is infinite then all three statements are false.

If U is finite then (since U spans itself) it has a finite spanning set and so a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\ldots$, $\mathbf{e}_m$. The elements of U are the vectors

$$\sum_{j=1}^{m} x_j \mathbf{e}_j$$

with $x_j \in \mathbb{Z}_2$, each choice of the x_j giving a distinct vector so U has 2^m elements.

If U is isomorphic to $\mathbf{Z}_2^q$ they must have the same number of elements so $q = m$.

The map $\theta : \mathbf{Z}_2^m \to U$ given by

$$\theta(x_1, x_2, \ldots, x_m) = \sum_{j=1}^{m} x_j \mathbf{e}_j.$$

Thus the three given statements are equivalent.

Exercise 13.3.3

We have

$$\alpha(\lambda\mathbf{e}+\mu\mathbf{f}) = \sum_{j=1}^{n} p_j(\lambda e_j+\mu f_j) = \lambda\sum_{j=1}^{n} p_j e_j+\mu\sum_{j=1}^{n} p_j f_j = \lambda\alpha(\mathbf{e})+\mu\alpha(\mathbf{f})$$

so α is linear.

Let $\mathbf{e}(j)$ be the vector with 1 in the jth place and 0 elsewhere. If α is linear, then, setting $p_j = \alpha\mathbf{e}(j)$ we have

$$\alpha(\mathbf{e}) = \alpha\left(\sum_{j=1}^{n} e_j\mathbf{e}(j)\right) = \sum_{j=1}^{n} e_j\alpha\mathbf{e}(j) = \sum_{j=1}^{n} p_j e_j.$$

Exercise 13.3.4

Codeword $(0, 0, 0, 1, 1, 1, 1)$.

Error in fourth place gives $(0, 0, 0, 0, 1, 1, 1)$.

$z_1 = 0,\ z_2 = 0,\ z_3 = 1$ tells us error in 4th place. Recover $(0, 0, 0, 1, 1, 1, 1)$.

Exercise 13.3.5

The tape will be accepted if each line contains an even number of errors. Since the probability of errors is small (and the number of bits on each line is small) the probability of one error is much greater than the probability of an odd number of errors greater than 1. Thus

$$\begin{aligned}\Pr(\text{odd number of errors in one line}) &\approx \Pr(\text{exactly one error in one line})\\ &= 8 \times 10^{-4} \times (1-10^{-4})^7 \approx 8 \times 10^{-4}.\end{aligned}$$

Since the the probability λ of an odd number of errors in one line is very small but there are a large number N of lines, we may use the Poisson approximation to get

$$1 - \Pr(\text{odd number of errors in some line}) \approx e^{-\lambda N} \approx e^{-8} \approx 0.00034$$

and conclude that the probability of acceptance is less than .04%.

If we use the Hamming scheme, then, instead of having 7 freely chosen bits (plus a check bit) on each line, we only have 4 freely chosen bits (plus three check bits plus an unused bit) per line so we need approximately

$$\frac{1}{4} \times 7 \times 10^4 = 1.75 \times 10^4$$

lines.

If a line contains at most one error, it will be correctly decoded. A line will fail to be correctly decoded if it contains exactly two errors and it may fail to be correctly decoded if it contains more than two errors. Since the probability of errors is small (and the number of bits on each line is small), the probability of two errors is much greater than the probability of more than two. Thus

$$\begin{aligned}\Pr(\text{decoding failure for one line}) &\approx \Pr(\text{exactly two errors in one line})\\ &= \binom{7}{2} \times (10^{-4})^2 \times (1-10^{-4})^5 \approx 21 \times 10^{-8}.\end{aligned}$$

Since the the probability of a decoding error in one line is very small but there are a large number of lines, we may use the Poisson approximation to get

$$\begin{aligned}\Pr(\text{decoding error for some line}) &= 1 - \Pr(\text{no decoding error in any line})\\ &\approx 1 - e^{-21\times 10^{-8}\times 17500} \approx 1 - e^{-.003675} \approx .9963\end{aligned}$$

and conclude that the probability of a correct decode is greater than 99.6%.

Exercise 13.3.6

(i) Observe that

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

can be rewritten as

$$\begin{aligned} c_1 + c_3 + c_5 + c_7 &\equiv 0 \\ c_2 + c_3 + c_6 + c_7 &\equiv 0 \\ c_4 + c_5 + c_6 + c_7 &\equiv 0. \end{aligned}$$

(ii) Observe that $A\mathbf{e}_j^T$ is the j row of A and that the jth row of A is the binary digits of j written in reverse order.

By linearity $A(\mathbf{c} + \mathbf{e}_j)^T = A\mathbf{c}^T + A\mathbf{e}_j^T = A\mathbf{e}_j^T = (a_1, a_2, a_3)^T$ so the statement follows.

Exercise 13.3.7

$A\mathbf{e}_i^T \neq A\mathbf{e}_j^T$ for $i \neq j$ since otherwise an error in the jth place would have the same effect as an error in the ith place. Since

$$\mathbf{x} + \mathbf{y} = \mathbf{0} \Leftrightarrow \mathbf{x} = \mathbf{y}$$

we have

$$A(\mathbf{e}_i^T + \mathbf{e}_j^T) \neq \mathbf{0}^T$$

Thus by linearity

$$\begin{aligned} A(\mathbf{c} + \mathbf{e}_i)^T + \mathbf{e}_j)^T) &= A\mathbf{c} + A\mathbf{c}^T + A\mathbf{e}_j^T \\ &= A(\mathbf{e}_i^T + \mathbf{e}_j^T) \neq \mathbf{0} \end{aligned}$$

and $\mathbf{c} + \mathbf{e}_i)^T + \mathbf{e}_j$ is not a codeword.

(ii) $A\mathbf{e}_j \in \mathbb{Z}^3 \setminus \{\mathbf{0}\}$. Since the $A\mathbf{e}_j$ are distinct for distinct j and $\{j : 1 \leq j \leq 7\}$ has the same number of elements as $\mathbb{Z}^3 \setminus \{\mathbf{0}\}^T$ we have bijection and so since

$$A(\mathbf{e}_i^T + \mathbf{e}_j^T) \neq \mathbf{0}$$

there must be a k with

$$A(\mathbf{e}_i^T + A\mathbf{e}_j^T) = A(\mathbf{e}_k^T).$$

If $k = j$ we would have $A\mathbf{e}_i^T = \mathbf{0}^T$ which is impossible. Thus $j \neq k$ and similarly $j \neq i$.

If we have an error in the i and jth place the Hamming system will assume an error in the kth place.

Exercise 13.3.9

Observe that if $A = (a_{ij})$

$$\begin{aligned}
\mathbf{c} \in C &\Leftrightarrow A\mathbf{c}^T = \mathbf{0}^T \\
&\Leftrightarrow \sum_{j=1}^{n} a_{ij}c_j = 0 \text{ for } 1 \leq j \leq r \\
&\Leftrightarrow \alpha_j \mathbf{c} = 0 \text{ for } 1 \leq j \leq r
\end{aligned}$$

where $\alpha_j \in (\mathbf{Z}_2^n)'$ is given by

$$\alpha_j \mathbf{c} = \sum_{j=1}^{n} a_{ij}c_j.$$

Exercise 13.3.10

The null space of a linear map is a subspace.

Exercise 13.3.12

Let α be the linear map associated with A for the standard bases. Since A has rank r so does α and

$$\dim C = \dim \alpha^{-1}(0) = n - \operatorname{rank}(\alpha) = n - r$$

so C has 2^{n-r} elements.

Exercise 13.3.13

(i) Suppose C has basis $\mathbf{g}(j)$ $[1 \leq r \leq n]$. Consider an $r \times n$ matrix U with jth row $\mathbf{g}(j)$. We can use elementary elementary row operations so that (after reordering coordinates if necessary) we get $V = \begin{pmatrix} I & B \end{pmatrix}$. Since elementary row operations leave the space spanned by the rows of a matrix unchanged, C has basis the rows $\mathbf{e}(j)$ of V. This gives the required result.

(ii) The rank nullity theorem tells us that the Hamming code has dimension 4. If we do not interchange columns then considering the effect of choosing one of the c_3, c_5, c_6, c_7 to be 1 and the rest 0 we obtain 4 independent vectors.

$$(1,1,1,0,0,0,0),\ (1,0,0,1,1,0,0)$$
$$(0,1,0,1,0,1,0),\ (1,1,0,1,0,0,1)$$

which therefore form a basis.

(iii) α is linear and surjective so (since the spaces have the same dimension) an isomorphism.

$$\begin{aligned}
\big(\lambda\mathbf{x} + \mu\mathbf{y}, \beta(\lambda\mathbf{x} + \mu\mathbf{y})\big) &= \alpha(\lambda\mathbf{x} + \mu\mathbf{y}) \\
&= \lambda\alpha\mathbf{x} + \mu\alpha\mathbf{y} \\
&= \lambda(\mathbf{x}, \beta\mathbf{x}) + \mu(\mathbf{y}, \beta\mathbf{y}) \\
&= (\lambda\mathbf{x} + \mu\mathbf{y}, \lambda\beta\mathbf{x} + \mu\beta\mathbf{y})
\end{aligned}$$

so

$$\beta(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\beta\mathbf{x} + \mu\beta\mathbf{y}$$

and β is linear.

(iv) Neither is true. Take $n = 2r$ $\mathbf{b}(j) = \mathbf{0}$.

Exercise 13.3.14

The probability of exactly 2 errors in a line (ie codeword), in which case the correction is erroneous, is

$$\binom{7}{2}\left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right)^5 \approx 0.124$$

The probability of no such problem in 17 500 lines is less than $(9/10)^{17500}$ which is negligible.

Exercise 13.3.15

If $p = 1/2$ then, whatever the message sent, the received sequence is a series of independent random variables each taking the value 0 with probability $1/2$ and the value 1 with probability $1/2$.

If $p > 1/2$, then replacing each received 0 by 1 and each received 1 by 0 gives a system for the probability of an error in one bit is $1 - p$.

Exercise 13.3.16

By definition

$$A_1 = (1),\ A_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},\ A_3 = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 &) & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

and

$$A_4 = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

(ii) The 2^rth column has a 1 in the rth place with all other entries zero, so these columns are linearly independent and A has rank n.

(iii) Let us write $j = \sum_{i=1}^{n+1} a_{ij} 2^{i-1}$ with $a_{ij} \in \{0, 1\}$. Then, if $0 \leq j < 2^n$, $a_{ij} = a_{i(j+2^n)}$ for $1 \leq i \leq n-1$ and $a_{n,j} = 0$, $a_{n,(j+2^n)} = 1$. We have $a_{i2^{n-1}} = \delta_{in}$.

Exercise 13.3.17

Let $\mathbf{e}_k$ be a row vector of length $2^n - 1$ with 1 in the kth place an zero elsewhere. Write $k = \sum_{i=1}^{n-1} b_i 2^{i-1}$ with $b_i \in \{0, 1\}$. Then, if $\mathbf{c}$ is a code word,

$$A(\mathbf{c} + \mathbf{e}_k)^T = A\mathbf{e}_k^T = \mathbf{b},$$

so we can recover from a single mistake by the rules:-

(1) If $\mathbf{x}$ received and $A\mathbf{x}^T = \mathbf{0}^T$, assume message correct.

(2) If $\mathbf{x}$ received and $A\mathbf{x}^T = \mathbf{b}^T$, assume mistake in $\sum_{i=1}^{n} b_i 2^{i-1}$th place.

Since A has full rank (see previous question), the rank-nullity theorem tells us that C has 2^{n-r} elements.

Exercise 13.3.18

Since the error rate is very low we can use the Poisson approximation. (But exact calculation is easy using a calculator.) We have a rate $(5 \times 10^7) \times 10^{-7} = 10$ so the probability that no error occurs is roughly e^{-5} which is very small.

By the Poisson approximation (or directly by bounding terms) the probability of more than one error in a line is roughly

$$p = \frac{1}{2!}\lambda^2 e^{-\lambda}$$

with $\lambda = 63 \times 10^{-7}$. Thus

$$p \approx 2 \times 10^{-10}$$

There are roughly 10^5 lines so the probability of an error in the decoded message is (using the Poisson approximation) less than about $\exp(-2 \times 10^{-5})$ which is pretty small.

Exercise 14.1.3

$C([a,b])$ is a subset of the vector space $\mathbb{R}^{[a,b]}$ of functions $f : [a,b] \to \mathbb{R}$. We observe that $0 \in C([a,b])$ and that

$$f,\ g \in C([a,b]) \Rightarrow \lambda f + \mu g \in C([a,b])$$

so $C([a,b])$ is a subspace of $\mathbb{R}^{[a,b]}$ and so a vector space.

We have

$$\langle f, f \rangle = \int_a^b f(x)^2\, dx \geq 0$$

and since the integral of a positive continuous function is zero if and only if the function is zero

$$\langle f, f \rangle = 0 \Rightarrow \int_a^b f(x)^2\, dx = 0 \Rightarrow f \times f = 0 \Rightarrow f = 0.$$

Also

$$\langle f, g \rangle = \int_a^b f(x)g(x)\, dx = \int_a^b g(x)f(x)\, dx = \langle g, f \rangle,$$

$$\begin{aligned} \langle f, g + h \rangle &= \int_a^b f(x)(g(x) + h(x)\, dx \\ &= \int_a^b f(x)g(x) + f(x)h(x)\, dx \\ &= \langle f, g \rangle + \langle f, g \rangle \end{aligned}$$

and

$$\langle \lambda f, g \rangle = \int_a^b \lambda f(x)g(x)\, dx = \lambda \int_a^b f(x)g(x)\, dx = \lambda \langle f, g \rangle.$$

Exercise 14.1.4

(i) If $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$, we have equality. If λ is a real number then we have

$$\begin{aligned}
0 &\leq \|\lambda\mathbf{x} + \mathbf{y}\|^2 \\
&= \langle \lambda\mathbf{x} + \mathbf{y}, \lambda\mathbf{x} + \mathbf{y} \rangle \\
&= \lambda^2\|\mathbf{x}\|^2 + 2\lambda\langle \mathbf{x}, \mathbf{y}\rangle + \|\mathbf{y}\|^2 \\
&= \left(\lambda\|\mathbf{x}\| + \frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|}\right)^2 + \|\mathbf{y}\|^2 - \left(\frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|}\right)^2.
\end{aligned}$$

If we set

$$\lambda = -\frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|^2},$$

we obtain

$$0 \leq (\lambda\mathbf{x} + \mathbf{y}) \cdot (\lambda\mathbf{x} + \mathbf{y}) = \|\mathbf{y}\|^2 - \left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|}\right)^2.$$

Thus

$$\bigstar \qquad \|\mathbf{y}\|^2 - \left(\frac{\langle \mathbf{x}, \mathbf{y}\rangle}{\|\mathbf{x}\|}\right)^2 \geq 0$$

with equality only if

$$0 = \|\lambda\mathbf{x} + \mathbf{y}\|$$

so only if

$$\lambda\mathbf{x} + \mathbf{y} = \mathbf{0}.$$

Rearranging the terms in $\bigstar$ we obtain

$$\langle \mathbf{x}, \mathbf{y}\rangle^2 \leq \|\mathbf{x}\|^2\|\mathbf{y}\|^2$$

and so

$$|\langle \mathbf{x}, \mathbf{y}\rangle| \leq \|\mathbf{x}\|\|\mathbf{y}\|$$

with equality only if

$$\lambda\mathbf{x} + \mathbf{y} = \mathbf{0}.$$

We observe that, if λ', $\mu' \in \mathbb{R}$ are not both zero and $\lambda'\mathbf{x} = \mu'\mathbf{y}$, then $|\langle \mathbf{x}, \mathbf{y}\rangle| = \|\mathbf{x}\|\|\mathbf{y}\|$.

Thus $|\langle \mathbf{x}, \mathbf{y}\rangle| \leq \|\mathbf{x}\|\|\mathbf{y}$ with equality if and only if $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent.

(ii) By definition $\|\mathbf{x}\| \geq 0$. Further

$$\|\mathbf{x}\| \geq 0 \Leftrightarrow \langle \mathbf{x}, \mathbf{x}\rangle = 0 \Leftrightarrow \mathbf{x} = 0.$$

Since $\langle \lambda\mathbf{x}, \lambda\mathbf{x}\rangle = \lambda^2\langle \mathbf{x}, \mathbf{x}\rangle$, we have $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$.

Finally

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + 2\langle \mathbf{x}, \mathbf{y} + \|\mathbf{y}\|^2\\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\|\mathbf{y}\| + \|\mathbf{y}\|^2\\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2,\end{aligned}$$

so $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

(iii) If we use the inner product of Exercise 14.1.3, the Cauchy–Schwarz inequality takes the form

$$\left(\int_a^b f(t)g(t)\,dt\right)^2 \leq \int_a^b f(t)^2\,dt \int_a^b g(t)^2\,dt.$$

Exercise 14.1.5

(i) We have $p_0 = 1$, so the result is true for $n = 0$. If p_r is polynomial of degree r for $r \leq n-1$, then

$$\sum_{j=0}^{n-1} \frac{\langle q_n, p_j \rangle}{\|p_j\|^2}$$

is a polynomial of degree at most $n-1$, so

$$p_n = q_n - \sum_{j=0}^{n-1} \frac{\langle q_n, p_j \rangle}{\|p_j\|^2} p_j$$

is a monic polynomial of degree n.

$$p_1(x) = x - \frac{\langle q_1, p_0 \rangle}{\|p_0\|^2} p_0(x) = x.$$

$$\begin{aligned} p_1(x) &= x^2 - \sum_{j=0}^{1} \frac{\langle q_n, p_j \rangle}{\|p_j\|^2} p_j(x) \\ &= x^2 - \left(\int_{-1}^{1} t^2 \, dt \right) \left(\int_{-1}^{1} 1^2 \, dt \right) = x^2 - \tfrac{1}{3}. \end{aligned}$$

(ii) 0 is a polynomial of degree less than n. If P and Q are polynomials of degree at most n so is $\lambda P = \mu Q$. Writing $q_r(x) = x^r$, we have the trivial equality

$$\sum_{r=0}^{n} a_r x^r = \left(\sum_{r=0}^{n} a_r q_r \right)(x)$$

so the q_j with $0 \leq j \leq n$ span $\mathcal{P}_n$. Since a non-zero polynomial of degree at most n can have at most n zeros

$$\sum_{r=0}^{n} a_r q_r = 0 \Rightarrow a_r = 0 \ [0 \leq r \leq n]$$

so the q_r form a basis and $\mathcal{P}_n$ has dimension $n+1$.

The $p_j \|p_j\|^{-1}$ form a collection of $n+1$ orthonormal vectors so form an orthonormal basis.

(iii) If P is monic of degree $n+1$, then

$$P = p_{n+1} + Q$$

with Q a polynomial of degree at most n Now

$$\begin{aligned} P \perp \mathcal{P}_n &\Rightarrow (P - p_{n+1}) \perp \mathcal{P}_n \Rightarrow Q \perp \mathcal{P}_n \\ &\Rightarrow Q \perp Q \Rightarrow \langle Q, Q \rangle = 0 \\ &\Rightarrow Q = 0 \Rightarrow P = p_{n+1}. \end{aligned}$$

By definition p_{n+1} is monic and

$$p_{n+1} \perp \operatorname{span}\{p_0,\, p_1,\, \ldots,\, p_n\}$$

and so, since $\mathcal{P}_n \operatorname{span}\{p_0,\, p_1,\, \ldots,\, p_n\}$, we have $p_{n+1} \perp \mathcal{P}_n$.

Exercise 14.1.11

(i) Observe that

$$\sum_{j=1}^{n} |\lambda_j| \geq \sum_{j=1}^{n} \lambda_j = \int_{-1}^{1} 1\,dt = 2.$$

Similarly $\sum_{j=n+1}^{2n} |\lambda_j| \geq 2$.

(ii) We have

$$\sum_{j=1}^{2n} \mu_j P(t_j) = (\epsilon^{-1}+1)\sum_{j=1}^{n} \lambda_j P(t_j) - \epsilon^{-1}\sum_{j=n+1}^{2n} \lambda_j P(t_j) = \int_{-1}^{1} P(t)\,dt$$

for all polynomials P of degree n or less.

Let f be the simplest piecewise linear functional with

$$f(t_j = \begin{cases} \epsilon & \text{if } 1 \leq j \leq n \text{ and } \lambda_j \geq 0 \\ -\epsilon & \text{if } 1 \leq j \leq n \text{ and } \lambda_j < 0 \\ \epsilon & \text{if } n+1 \leq j \leq 2n \text{ and } \lambda_j < 0 \\ -\epsilon & \text{if } n+1 \leq j \leq 2n \text{ and } \lambda_j \geq 0 \end{cases}$$

with

$$\begin{aligned} f(-1) &= 0 && \text{if } -1 \notin \{t_j : 1 \leq j \leq 2n\} \\ f(1) &= 0 && \text{if } 1 \notin \{t_j : 1 \leq j \leq 2n\}. \end{aligned}$$

Then $|f(t)| \leq \epsilon$ for all $t \in [-1,1]$, but

$$\sum_{j=1}^{2n} \mu_j P(t_j) = \sum_{j=1}^{2n} |\mu_j| \geq 2(1+\epsilon^{-1}) + 2\epsilon^{-1}.$$

Exercise 14.1.12

(i) Observe that

$$\langle f, f \rangle = \int_a^b f(x)^2 r(x)\, dx \geq 0,$$

and, since $x \mapsto f(x)^2 r(x)$ is continuous and positive

$$\begin{aligned}\langle f, f \rangle = 0 &\Rightarrow \int_a^b f(x)^2 r(x)\, dx = 0\\ &\Rightarrow f(x)^2 r(x) = 0\ \forall x\\ &\Rightarrow f(x) = 0\ \forall x\end{aligned}$$

Automatically

$$\langle f, g \rangle = \int_a^b f(x)g(x)r(x)\, dx = \int_a^b g(x)f(x)r(x)\, dx = \langle g, f \rangle$$

and

$$\begin{aligned}\langle \lambda_1 f_1 + \lambda_2 f_2, g \rangle &= \int_a^b (\lambda_1 f_1(x) + \lambda_2 f_2(x))g(x)r(x)\, dx\\ &= \int_a^b \lambda_1 f_1(x)g(x) + \lambda_2 f_2 g(x)r(x)\, dx\\ &= \lambda_1 \langle f_1, g \rangle + \lambda_2 \langle f_2, g \rangle.\end{aligned}$$

(ii) This is just the Cauchy–Schwartz inequality

$$\langle f, f \rangle \langle g, g \rangle \geq \langle f, g \rangle^2.$$

(iii) We define P_n inductively. Set $P_0 = 1$. Suppose that P_j is a monic polynomial of degree j with the P_j orthogonal with respect to our inner product $[0 \leq j \leq n-1]$. Then, if $Q_n(x) = x^n$,

$$P_n = Q_n - \sum_{j=0}^{n-1} \frac{\langle Q_n, P_j \rangle}{\langle P_j, P_j \rangle}$$

is a monic polynomial of degree n orthogonal to all the P_j with $j \leq n-1$.

Since the subspace $\mathcal{P}_{n-1}$ of polynomials of degree at most $n-1$ has dimension n and the P_j are non-zero and mutually orthogonal the P_j with $j \leq n-1$ form a basis for $\mathcal{P}_{n-1}$ and $\langle P_n, Q \rangle = 0$ for all polynomials Q of degree $n-1$ or less.

(iii) Let x_1, x_2, $\ldots x_k$ be the real roots of P_n which have odd order and lie in (a, b). Setting $Q(x) = \prod_{j=1}^k (x - x_k)$, we see that $P_n Q$ is single signed so

$$\langle P_n, Q\rangle = \int_a^b P_n(x)Q(x)r(x)\,dx \neq 0$$

and $k \geq n$. Thus $k = n$ and P_n has n distinct real roots all lying in (a, b).

If we set $e_v(x) = \prod_{j\neq v, 1\leq j\leq n}(x - x_j)(x_v - x_j)^{-1}$ and

$$\lambda_v = \int_a^b e_v(x)r(x)\,dx$$

then (since $f - \sum_{v=1}^n f_(x_v)e_v$ is a polynomial of degree at most $n-1$ vanishing at n points) we have $f = \sum_{v=1}^n f(x_v)e_v$ and

$$\int_a^b f(x)r(x)\,dx = \sum_{v=1}^n \lambda_v f_(x_v)$$

for all f polynomials of degree $n-1$ or less.

If f is polynomials of degree $2n-1$ or less, then $f = gP_n + h$ with g and h polynomials of degree $n-1$ or less, so

$$\begin{aligned}
\int_a^b f(x)r(x)\,dx &= \int_a^b g(x)P_n r(x)\,dx + \int_a^b h(x)r(x)\,dx \\
&= 0 + \sum_{v=1}^n \lambda_v g_(x_v) \\
&= \sum_{v=1}^n \lambda_v g_(x_v) + \sum_{v=1}^n \lambda_v 0 \\
&= \sum_{v=1}^n \lambda_v g_(x_v) + \sum_{v=1}^n \lambda_v P_n(x_v)h(x_v) \\
&= \sum_{v=1}^n \lambda_v f(x_v).
\end{aligned}$$

Exercise 14.1.14

We need to show that $\langle P_n, qP_{n-1}\rangle > 0$.

Now qP_{n-1} and P_n are both monic polynomials of degree n, so $qP_{n-1} - P_n$ has degree at most $n-1$. Thus

$$\langle P_n, qP_{n-1} - P_n\rangle = 0$$

and

$$\langle P_n, qP_{n-1}\rangle = \langle P_n, P_n\rangle > 0.$$

Exercise 14.1.16

Observe that

$$\begin{aligned} 0 \leq \int_1^1 f_n(x)^2\,dx &= (2^{n+1}-1)\int_{-2^{-8n}}^{2^{-8n}} f_n(x)^2\,dx \\ &\leq 2^{n+1}2^{-8n+1}2^{2n} = 2^{-5n+2} \to 0 \end{aligned}$$

as $n \to \infty$ so $\|f_n\| \to 0$.

However, if u is an integer with $|u| < 2^m - 1$, then $f(u2^{-m}) = 2^n$ when $n \geq m$ and so $f_n(u2^{-m}) \to \infty$ as $n \to \infty$.

Exercise 14.2.1

If $\dim U = 0$, there is nothing to prove.

Otherwise, we construct the $\mathbf{e}_j$ inductively. Since $\dim U > 0$ we can find a $\mathbf{a} \neq \mathbf{0}$. Set $\mathbf{e}_1 = \|\mathbf{a}\|^{-1}\mathbf{a}$. If $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_r$ have been found as orthonormal vectors either $r = n$ and we have an orthonormal basis or we can find

$$\mathbf{b} \notin \operatorname{span}\{\mathbf{e}_1,\, \mathbf{e}_2,\, \ldots, \mathbf{e}_r\}$$

The Gram–Schmidt method now produces an $\mathbf{e}_{r+1}$ with $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_{r+1}$ orthonormal.

The process terminates in an orthonormal basis.

Since the $\mathbf{e}_j$ are linearly independent and span the map

$$\theta\left(\sum_{j=1}^n x_j\mathbf{e}_j\right) = (x_1, x_2, \ldots, x_n)^T$$

is a well defined bijection. Since

$$\begin{aligned}
\theta\left(\lambda\sum_{j=1}^n x_j\mathbf{e}_j + \mu\sum_{j=1}^n y_j\mathbf{e}_j\right) &= \theta\left(\sum_{j=1}^n (\lambda x_j + \mu y_j)\mathbf{e}_j\right)\\
&= (\lambda x_1 + \mu y_1, \lambda x_2 + \mu y_2, \ldots, \lambda x_n + \mu y_n)^T\\
&= \lambda(x_1, x_2, \ldots, x_n)^T + \mu(y_1, y_2, \ldots, y_n)^T\\
&= \lambda\theta\left(\lambda\sum_{j=1}^n x_j\mathbf{e}_j\right) + \mu\theta\left(\sum_{j=1}^n y_j\mathbf{e}_j\right)
\end{aligned}$$

θ is linear. Finally,

$$\theta\left(\sum_{j=1}^n x_j\mathbf{e}_j)\right)\cdot\theta\left(\sum_{j=1}^n y_j\mathbf{e}_j\right) = \sum_{j=1}^n x_jy_j = \left\langle \theta\left(\sum_{j=1}^n x_j\mathbf{e}_j\right), \theta\left(\sum_{j=1}^n y_j\mathbf{e}_j\right)\right\rangle.$$

Exercise 14.2.3

Choose an orthonormal basis

$$\mathbf{e}_1,\ \mathbf{e}_2,\ \ldots,\ \mathbf{e}_r$$

for V. Then by Bessel's inequality

$$\|\mathbf{a} - \sum_{j=1}^{r} \lambda_j \mathbf{e}_j\|$$

has a unique minimum when $\lambda_j = \langle \mathbf{a}, \mathbf{e}_j \rangle$.

If $\mathbf{x} \in V$

$$\begin{aligned}\langle \mathbf{a} - \mathbf{x}, \mathbf{v} \rangle = 0\ \forall \mathbf{v} \in V &\Leftrightarrow \langle \mathbf{a} - \mathbf{x}, \mathbf{e}_j \rangle = 0\ \forall 1 \leq j \leq r\\ &\Leftrightarrow \langle \mathbf{x}, \mathbf{e}_j \rangle = \langle \mathbf{a}, \mathbf{e}_j \rangle\ \forall 1 \leq j \leq r\\ &\Leftrightarrow \mathbf{x} = \sum_{j=1}^{r} \langle \mathbf{a}, \mathbf{e}_j \rangle \mathbf{e}_j\end{aligned}$$

so we are done.

Exercise 14.2.5★

Exercise 14.2.6

Observe that

$$\begin{aligned}\langle \mathbf{u}, \mathbf{a}\rangle &= \left\langle \sum_{j=1}^{n} \alpha(\mathbf{e}_j)\mathbf{e}_j, \mathbf{u} \right\rangle \\ &= \sum_{j=1}^{n} \alpha(\mathbf{e}_j)\langle \mathbf{e}_j, \mathbf{a}\rangle \\ &= \alpha\left(\sum_{j=1}^{n}\langle \mathbf{e}_j, \mathbf{a}\rangle \mathbf{e}_j\right) \\ &= \alpha\mathbf{u}.\end{aligned}$$

Exercise 14.2.7

We have a map $\alpha : U \to \mathbb{R}$ so we only need check linearity.

But, if $\mathbf{u}$, $\mathbf{v} \in U$ and λ, $\mu \in \mathbb{R}$ we have

$$\alpha(\lambda\mathbf{u} + \mu\mathbf{v}) = \langle \lambda\mathbf{u} + \mu\mathbf{v}, \mathbf{a}\rangle = \lambda\langle \mathbf{u}, \mathbf{a}\rangle + \mu\langle \mathbf{v}, \mathbf{a}\rangle = \lambda\alpha\mathbf{u} + \mu\alpha\mathbf{v},$$

so we are done.

EXERCISE 14.2.10

Observe that

$$\begin{aligned}\langle \mathbf{u}, \Phi(\lambda\alpha + \mu\beta)\mathbf{v}\rangle &= \langle(\lambda\alpha + \mu\beta)\mathbf{u}, \mathbf{v}\rangle\\ &= \langle\lambda\alpha(\mathbf{u}) + \mu\beta(\mathbf{u}), \mathbf{v}\rangle\\ &= \lambda\langle\alpha\mathbf{u}, \mathbf{v}\rangle + \mu\langle\beta\mathbf{u}, \mathbf{v}\rangle\\ &= \lambda\langle\mathbf{u}, \alpha^*\mathbf{v}\rangle + \mu\langle\mathbf{u}, \beta^*\mathbf{v}\rangle \quad = \langle\mathbf{u}, (\lambda\alpha^* + \mu\beta^*)\mathbf{v}\rangle\\ &= \langle\mathbf{u}, (\lambda\Phi\alpha + \mu\Phi\beta)\mathbf{v}\rangle\end{aligned}$$

for all $\mathbf{u} \in U$, so

$$\Phi(\lambda\alpha + \mu\beta)\mathbf{v} = (\lambda\Phi\alpha + \mu\Phi\beta)\mathbf{v}$$

for all $\mathbf{v} \in U$ so

$$\Phi(\lambda\alpha + \mu\beta) = \lambda\Phi\alpha + \mu\Phi\beta$$

for all α, $\beta \in \mathcal{L}(U, U)$ and λ, $\mu \in \mathbb{F}$, so Φ is linear.

Exercise 14.2.12

$\mathcal{D}$ is a subset of the inner product space $C([0,1])$. $0 \in \mathcal{D}$ and

$$\lambda,\ \mu \in \mathbb{R},\ \ f,\ g \in \mathcal{D} \Rightarrow \lambda f + \mu g \in \mathcal{D}.$$

Thus $\mathcal{D}$ is a subspace of the inner product space $C([0,1])$ and so an inner product space.

Further,

$$\alpha(\lambda f + \mu g) = \big(p(\lambda f + \mu g)'\big)' = \big(\lambda p f' + \mu g f'\big)' = \lambda \alpha f + \mu \alpha g.$$

so α is linear.

By integration by parts,

$$\begin{aligned}\langle \alpha f, g\rangle &= \int_0^1 (f'p)'(t) g(t)\, dt\\ &= \big[f'(t)p(t)g(t)\big]_0^1 - \int_0^1 f'(t)p(t)g'(t)\, dt\\ &= -\int_0^1 f'(t)p(t)g'(t)\, dt\\ &= \langle \alpha g, f\rangle = \langle f, \alpha g\rangle,\end{aligned}$$

so $\alpha^* = \alpha$.

If we write $(\beta f)(t) = tf(t)$ then β is a self-adjoint linear functional.

Exercise 14.3.2

We have
$$\langle \mathbf{z}, \lambda\mathbf{w}\rangle_P = \langle \lambda\mathbf{w}, \mathbf{z}\rangle_P^* = \big(\lambda^*\langle \mathbf{w}, \mathbf{z}\rangle_P\big)^* = \lambda\langle \mathbf{w}, \mathbf{z}\rangle_P^* = \lambda\langle \mathbf{z}, \mathbf{w}\rangle_P.$$

Let $\langle\ ,\ \rangle_M$ be defined as in the question. (i) Since $\langle \mathbf{z}, \mathbf{z}\rangle_P$ is real and positive, so is $\langle \mathbf{z}, \mathbf{z}\rangle_M = \langle \mathbf{z}, \mathbf{z}\rangle_P^*$.

(ii) $\langle \mathbf{z}, \mathbf{z}\rangle_M = 0 \Rightarrow \langle \mathbf{z}, \mathbf{z}\rangle_P = 0 \Rightarrow \mathbf{z} = \mathbf{0}$.

(iii) $\langle \lambda\mathbf{z}, \mathbf{w}\rangle_M = \langle \lambda\mathbf{z}, \mathbf{w}\rangle_P^* = \big(\lambda^*\langle \mathbf{z}, \mathbf{w}\rangle_P\big)^* = \lambda\langle \mathbf{z}, \mathbf{w}\rangle_P^* = \lambda\langle \mathbf{z}, \mathbf{w}\rangle_M$.

(iv) We have
$$\begin{aligned}\langle \mathbf{z}+\mathbf{u}, \mathbf{w}\rangle_M &= \langle \mathbf{z}+\mathbf{u}, \mathbf{w}\rangle_P^*\\ &= \big(\langle \mathbf{z}, \mathbf{w}\rangle_P + \langle \mathbf{u}, \mathbf{w}\rangle_P\big)^*\\ &= \langle \mathbf{z}, \mathbf{w}\rangle_P^* + \langle \mathbf{u}, \mathbf{w}\rangle_P^*\\ &= \langle \mathbf{z}, \mathbf{w}\rangle_M + \langle \mathbf{u}, \mathbf{w}\rangle_M\end{aligned}$$

(v) $\langle \mathbf{w}, \mathbf{z}\rangle_M = \langle \mathbf{w}, \mathbf{z}\rangle_P^* = \langle \mathbf{z}, \mathbf{w}\rangle_P^{**} = \langle \mathbf{z}, \mathbf{w}\rangle_M^*$.

Essentially same arguments show that if $\langle\ ,\ \rangle_M$ is a mathematician's inner product then
$$\langle \mathbf{z}, \mathbf{w}\rangle_P = \langle \mathbf{z}, \mathbf{w}\rangle_M^*$$
defines a physicist's inner product.

Exercise 14.3.3

$C([a,b])$ is a subset of the vector space $\mathbb{C}^{[a,b]}$ of functions $f : [a,b] \to \mathbb{C}$. We observe that $0 \in C([a,b])$ and that

$$f,\ g \in C([a,b]) \Rightarrow \lambda f + \mu g \in C([a,b])$$

so $C([a,b])$ is a subspace of $\mathbb{R}^{[a,b]}$ and so a vector space.

We have

$$\langle f, f\rangle = \int_a^b |f(x)|^2\,dx \geq 0$$

and since the integral of a positive continuous function is zero if and only if the function is zero

$$\langle f, f\rangle = 0 \Rightarrow \int_a^b |f(x)|^2\,dx = 0 \Rightarrow f \times f = 0 \Rightarrow f = 0.$$

Also

$$\begin{aligned}\langle f, g\rangle &= \int_a^b f(x)g(x)^*\,dx = \int_a^b (g(x)f(x)^*)^*\,dx\\ &= \left(\int_a^b g(x)f(x)^*\,dx\right)^* = \langle g, f\rangle,\end{aligned}$$

and

$$\begin{aligned}\langle f, g+h\rangle &= \int_a^b f(x)(g(x) + h(x)^*\,dx\\ &= \int_a^b f(x)g(x)^* + f(x)h(x)^*\,dx = \langle f, g\rangle + \langle f, g\rangle,\end{aligned}$$

whilst

$$\langle \lambda f, g\rangle = \int_a^b \lambda f(x)g(x)^*\,dx = \lambda\int_a^b f(x)g(x)^*\,dx = \lambda\langle f, g\rangle.$$

Exercise 14.3.4

(i) We have

$$\begin{aligned}\langle \mathbf{f}, \mathbf{e}_j \rangle &= \left\langle \sum_{r=1}^{n} a_r \mathbf{e}_r, \mathbf{e}_j \right\rangle \\ &= \sum_{r=1}^{n} a_r \langle \mathbf{e}_r, \mathbf{e}_j \rangle = \sum_{r=1}^{n} a_r \delta_{rj} = a_j\end{aligned}$$

Note that

$$\langle \mathbf{e}_j, \mathbf{f} \rangle = \langle \mathbf{f}, \mathbf{e}_j \rangle^* = a_j^*.$$

(ii) We have

$$\langle \mathbf{e}_r, \mathbf{e}_s \rangle = \int_0^1 \exp(2\pi irt) \exp(-2\pi ist)\, dt = \int_0^1 \exp\big(2\pi i(r-s)t\big)\, dt = \delta_{rs}.$$

Exercise 14.3.5

(i) Observe that

$$\langle \mathbf{v}, \mathbf{e}_r \rangle = \langle \mathbf{x}, \mathbf{e}_r \rangle - \sum_{j=1}^{k} \langle \mathbf{x}, \mathbf{e}_r \rangle \langle \mathbf{e}_j, \mathbf{e}_r \rangle = \langle \mathbf{x}, \mathbf{e}_r \rangle - \langle \mathbf{x}, \mathbf{e}_r \rangle = 0$$

for all $1 \leq r \leq k$.

(ii) If $\mathbf{v} = \mathbf{0}$, then

$$\mathbf{x} = \sum_{j=1}^{k} \langle \mathbf{x}, \mathbf{e}_j \rangle \mathbf{e}_j \in \operatorname{span}\{\mathbf{e}_1,\, \mathbf{e}_2,\, \ldots,\, \mathbf{e}_k\}.$$

If $\mathbf{v} \neq \mathbf{0}$, then $\|\mathbf{v}\| \neq 0$ and

$$\begin{aligned} \mathbf{x} &= \mathbf{v} + \sum_{j=1}^{k} \langle \mathbf{x}, \mathbf{e}_j \rangle \mathbf{e}_j \\ &= \|\mathbf{v}\| \mathbf{e}_{k+1} + \sum_{j=1}^{k} \langle \mathbf{x}, \mathbf{e}_j \rangle \mathbf{e}_j \in \operatorname{span}\{\mathbf{e}_1,\, \mathbf{e}_2,\, \ldots,\, \mathbf{e}_{k+1}\}. \end{aligned}$$

(iii) If $\mathbf{e}_1$, $\mathbf{e}_2$, $\ldots$, $\mathbf{e}_k$ do not form a basis for U, then we can find

$$\mathbf{x} \in U \setminus \operatorname{span}\{\mathbf{e}_1,\, \mathbf{e}_2,\, \ldots,\, \mathbf{e}_k\}.$$

Defining $\mathbf{v}$ as in part (i), we see that $\mathbf{v} \in U$ and so the vector $\mathbf{e}_{k+1}$ defined in (ii) lies in U. Thus we have found orthonormal vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\ldots$, $\mathbf{e}_{k+1}$ in U. If they form a basis for U we stop. If not, we repeat the process. Since no set of $n+1$ vectors in U can be orthonormal (because no set of $n+1$ vectors in U can be linearly independent), the process must terminate with an orthonormal basis for U of the required form.

Exercise 14.3.7

Let V have an orthonormal basis $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n-1}$. By using the Gram–Schmidt process, we can find $\mathbf{e}_n$ such that $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are orthonormal and so a basis for U. Setting $\mathbf{b} = \mathbf{e}_n$, we have the result.

Exercise 14.3.8

Uniqueness. If $\alpha\mathbf{u} = \langle \mathbf{u}, \mathbf{a}_1 \rangle = \langle \mathbf{u}, \mathbf{a}_2 \rangle$ for all $\mathbf{u} \in U$, then

$$\langle \mathbf{u}, \mathbf{a}_1 - \mathbf{a}_2 \rangle = 0$$

for all $\mathbf{u} \in U$ and, choosing $\mathbf{u} = \mathbf{a}_1 - \mathbf{a}_2$, we conclude, in the usual way, that $\mathbf{a}_1 - \mathbf{a}_2 = \mathbf{0}$.

Existence. If $\alpha = 0$, then we set $\mathbf{a} = \mathbf{0}$. If not, then α has rank 1 (since $\alpha(U) = \mathbb{R}$) and, by the rank-nullity theorem, α has nullity $n - 1$. In other words,

$$\alpha^{-1}(\mathbf{0}) = \{\mathbf{u} : \alpha\mathbf{u} = 0\}$$

has dimension $n - 1$. By Lemma 14.2.2, we can find a $\mathbf{b} \neq \mathbf{0}$ such that

$$\alpha^{-1}(\mathbf{0}) = \{\mathbf{x} \in U : \langle \mathbf{x}, \mathbf{b} \rangle = 0\},$$

that is to say,

$$\alpha(\mathbf{x}) = 0 \Leftrightarrow \langle \mathbf{x}, \mathbf{b} \rangle = 0$$

If we now set

$$\mathbf{a} = \|\mathbf{b}\|^{-2}\alpha(\mathbf{b})\mathbf{b},$$

we have

$$\alpha\mathbf{x} = 0 \Leftrightarrow \langle \mathbf{x}, \mathbf{a} \rangle = 0$$

and

$$\alpha\mathbf{a} = \|\mathbf{b}\|^{-2}\alpha(\mathbf{b})^2 = \langle \mathbf{a}, \mathbf{a} \rangle.$$

Now suppose that $\mathbf{u} \in U$. If we set

$$\mathbf{x} = \mathbf{u} - \frac{\alpha\mathbf{u}}{\alpha\mathbf{a}}\mathbf{a},$$

then $\alpha\mathbf{x} = 0$ so $\langle \mathbf{x}, \mathbf{a} \rangle = 0$, that is to say,

$$0 = \left\langle \mathbf{u} - \frac{\alpha\mathbf{u}}{\alpha\mathbf{a}}\mathbf{a}, \mathbf{a} \right\rangle = \langle \mathbf{u}, \mathbf{a} \rangle - \frac{\alpha\mathbf{a}}{\langle \mathbf{a}, \mathbf{a} \rangle}\alpha\mathbf{u} = \langle \mathbf{u}, \mathbf{a} \rangle$$

and we are done.

Exercise 14.3.9

Since

$$\langle \lambda\mathbf{u} + \mu\mathbf{v}, \mathbf{a}\rangle = \lambda\langle \mathbf{u}, \mathbf{a}\rangle + \mu\langle \mathbf{u}, \mathbf{a}\rangle$$

$\theta(\mathbf{a} \in U'$. The previous exercise tells us that θ is surjective.

$$\theta(\mathbf{a}) = \theta(\mathbf{b}) \Rightarrow \langle \mathbf{b} - \mathbf{a}, \mathbf{b} - \mathbf{a}\rangle \Rightarrow \mathbf{a} = \mathbf{b}$$

so θ is injective.

Finally,

$$\begin{aligned}
\theta(\lambda\mathbf{a} + \mu\mathbf{b})(\mathbf{u}) &= \langle \mathbf{u}, \lambda\mathbf{a} + \mu\mathbf{b}\rangle \\
&= \lambda^*\langle \mathbf{u}, \mathbf{a}\rangle + \mu^*\langle \mathbf{u}, \mathbf{b}\rangle \\
&= \lambda^*\theta(\mathbf{a})\mathbf{u} + \mu^*\theta(\mathbf{b})\mathbf{u} \\
&= \big(\lambda^*\theta(\mathbf{a}) + \mu^*\theta(\mathbf{b})\big)\mathbf{u}
\end{aligned}$$

for all $\mathbf{u} \in U$. Thus

$$\theta(\lambda\mathbf{a} + \mu\mathbf{b}) = \lambda^*\theta(\mathbf{a}) + \mu^*\theta(\mathbf{b})$$

as required.

Exercise 14.3.10★

Exercise 14.3.11

Observe that, if $\mathbf{v} \in U$, the map
$$\theta(\mathbf{u}) = \langle \alpha\mathbf{u}, \mathbf{v} \rangle$$
lies in U', so the Riesz representation theorem tells us that there is a unique $\alpha^*\mathbf{v} \in U$ such that
$$\langle \alpha\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \alpha^*\mathbf{v} \rangle$$
for all $\mathbf{v} \in U$.

Since
$$\begin{aligned}\langle \mathbf{u}, \alpha^*(\lambda\mathbf{v} + \mu\mathbf{w}) \rangle &= \langle \alpha\mathbf{u}, (\lambda\mathbf{v} + \mu\mathbf{w}) \rangle \\ &= \lambda^*\langle \alpha\mathbf{u}, \mathbf{v} \rangle + \mu^*\langle \alpha\mathbf{u}, \mathbf{w} \rangle \\ &= \lambda^*\langle \mathbf{u}, \alpha^*\mathbf{v} \rangle + \mu^*\langle \mathbf{u}, \alpha^*\mathbf{w} \rangle \\ &= \langle \mathbf{u}, \lambda\alpha^*\mathbf{v} + \mu\alpha^*\mathbf{w} \rangle\end{aligned}$$
for all $\mathbf{u} \in U$, we have
$$\alpha^*(\lambda\mathbf{v} + \mu\mathbf{w}) = \lambda\alpha^*\mathbf{v} + \mu\alpha^*\mathbf{w}$$
so $\alpha^* \in \mathcal{L}(U, U)$.

Again,
$$\begin{aligned}\langle \mathbf{u}, (\lambda\alpha + \mu\beta)^*\mathbf{v} \rangle &= \langle (\lambda\alpha + \mu\beta)\mathbf{u}, \mathbf{v} \rangle \\ &= \lambda\langle \alpha\mathbf{u}, \mathbf{v} \rangle + \mu\langle \beta\mathbf{u}, \mathbf{v} \rangle \\ &= \lambda\langle \mathbf{u}, \alpha^*\mathbf{v} \rangle + \mu\langle \mathbf{u}, \beta^*\mathbf{v} \rangle \\ &= \langle \mathbf{u}, \lambda^*\alpha^*\mathbf{v} + \mu^*\beta^*\mathbf{v} \rangle \\ &= \langle \mathbf{u}, (\lambda^*\alpha^* + \mu^*\beta^*)\mathbf{v} \rangle\end{aligned}$$
for all $\mathbf{u} \in U$, so
$$(\lambda\alpha + \mu\beta)^*\mathbf{v} = (\lambda^*\alpha^* + \mu^*\beta^*)\mathbf{v}$$
for all $\mathbf{v} \in U$, so
$$(\lambda\alpha + \mu\beta)^* = \lambda^*\alpha^* + \mu^*\beta^*$$
and he map $\Psi : \mathcal{L}(U, U) \to \mathcal{L}(U, U)$ given by $\Psi\alpha = \alpha^*$ satisfies
$$\Psi(\lambda\alpha + \mu\beta) = \lambda^*\Psi\alpha + \mu^*\Psi\beta.$$
Finally,
$$\begin{aligned}\langle \alpha^{**}\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \alpha^{**}\mathbf{u} \rangle &= \langle \alpha^*\mathbf{v}, \mathbf{u} \rangle^* \\ = \langle \mathbf{u}, \alpha^*\mathbf{v} \rangle &= \langle \alpha\mathbf{u}, \mathbf{v} \rangle\end{aligned}$$
for all $\mathbf{v} \in U$, so
$$\alpha^{**}\mathbf{u} = \mathbf{u}$$
for all $\mathbf{u} \in U$, so $\alpha^{**} = \alpha$ and Ψ is its own inverse, so bijective so an anti-isomorphism.

Exercise 14.3.12

If α has matrix $A = (a_{ij})$ and α^* has matrix $B = (b_{ij})$ with respect to the given basis, then

$$\begin{aligned}
b_{rj} &= \left\langle \sum_{i=1}^{n} b_{ij}\mathbf{e}_i, \mathbf{e}_r \right\rangle \\
&= \langle \alpha^*\mathbf{e}_j, \mathbf{e}_r \rangle = \langle \mathbf{e}_r \alpha^* \mathbf{e}_j \rangle^* = \langle \alpha\mathbf{e}_r, \mathbf{e}_j \rangle \\
&= \left\langle \sum_{k=1}^{n} a_{kr}\mathbf{e}_k, \mathbf{e}_j \right\rangle^* \\
&= a_{jr}^*,
\end{aligned}$$

so $B = A^*$.

Exercise 14.3.13

(i) Since V and $V^{\perp}$ are complementary,

$$\dim V + \dim V^{\perp} = \dim U.$$

For the same reason, $\dim V^{\perp} + \dim V^{\perp\perp} = \dim U$, so $\dim V^{\perp\perp} + \dim V$.

Since

$$\mathbf{x} \in V \Rightarrow \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle = 0 \ \forall \mathbf{y} \in V^{\perp}$$

we have $V \subseteq V^{\perp\perp}$ so $V = V^{\perp\perp}$.

(ii) The statement is equivalent to saying that, if $\mathbf{u} \in U$, there are unique $\mathbf{a} \in V$ and $\mathbf{b} \in V^{\perp}$ such that $\mathbf{u} = \mathbf{a} + \mathbf{b}$. and this is the same as saying that V and $V^{\perp}$ are complementary.

(iii) We have

$$\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2 = (\lambda_1 \pi \mathbf{u}_1 + \lambda_2 \pi \mathbf{u}_2) + (\lambda_1 (\iota - \pi) \mathbf{u}_1 + \lambda_2 (\iota - \pi) \mathbf{u}_2)$$

and $\lambda_1 \pi \mathbf{u}_1 + \lambda_2 \pi \mathbf{u}_2 \in V$, $\lambda_1 (\iota - \pi) \mathbf{u}_1 + \lambda_2 (\iota - \pi) \mathbf{u}_2 \in V^{\perp}$ so, by uniqueness,

$$\pi(\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2) = \lambda_1 \pi \mathbf{u}_1 + \lambda_2 \pi \mathbf{u}_2.$$

Thus π and so $\iota - \pi$ are linear.

Since $\pi \mathbf{u} \in V$ we have, by uniqueness $\pi^2 \mathbf{U} = \pi(\pi(\mathbf{u}) = \pi \mathbf{u}$ for all $\mathbf{u} \in U$ so $\pi = \pi^2$.

If $\mathbf{x}$, $\mathbf{y} \in U$, then

$$\begin{aligned}
\langle \pi \mathbf{x}, \mathbf{y} \rangle &= \langle \pi \mathbf{x}, \pi \mathbf{y} + (\iota - \pi) \mathbf{y} \rangle \\
&= \langle \pi \mathbf{x}, \pi \mathbf{y} \rangle + \langle \pi \mathbf{x}, (\iota - \pi) \mathbf{y} \rangle \\
&= \langle \pi \mathbf{x}, \pi \mathbf{y} \rangle \\
&= \langle \pi \mathbf{x}, \pi \mathbf{y} \rangle + \langle (\iota - \pi) \mathbf{x}, \pi \mathbf{y} \rangle \\
&= \langle \mathbf{x}, \pi \mathbf{y} \rangle
\end{aligned}$$

so $\pi = \pi^*$

Exercise 14.3.14

(i)⇒(ii)

Let $U = \alpha(V)$ and $W = (\iota - \alpha)(V)$. Since

$$\mathbf{v} = \alpha\mathbf{v} + (\iota - \alpha)\mathbf{v}$$

$U + W = V$. If $\mathbf{u} \in U$ then $\mathbf{u} = \alpha\mathbf{v}$ for some $\mathbf{v} \in V$ so

$$\alpha\mathbf{u} = \alpha^2\mathbf{v} = \alpha\mathbf{v} = \mathbf{u} = \iota_U\mathbf{u}.$$

If $\mathbf{w} \in W$ then $\mathbf{u} = (\iota - \alpha)\mathbf{v}$ for some $\mathbf{v} \in V$ so

$$\alpha\mathbf{w} = (\alpha - \alpha^2)\mathbf{v} = \mathbf{0}.$$

Finally, if $\mathbf{u} \in U$, $\mathbf{w} \in W$ the results of the previous paragraph give

$$\langle \mathbf{u}, \mathbf{w} \rangle = \langle \alpha\mathbf{u}, (\iota - \alpha)\mathbf{w} \rangle = \langle \mathbf{u}, \alpha(\iota - \alpha)\mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{0} \rangle = 0.$$

(ii)⇒(iii) Immediate.

(iii)⇒(iv) Take an orthonormal basis for U and an orthonormal basis for W. Together they form an orthonormal basis for V with the required properties.

(iv)⇒(i) Let A be the matrix of α with respect to the specified basis. Then $A^* = A$ and $A^2 = A$, so the results for α follow.

An orthogonal projection automatically obeys the condition $\alpha^2 = \alpha$.

Give $\mathbb{C}^2$ the standard inner product. Let

$$\alpha(z, w) = (\tfrac{1}{3}z + \tfrac{2}{3}w, \tfrac{1}{3}z + \tfrac{2}{3}w).$$

By inspection, α is linear and $\alpha^2 = \alpha$. By observing that the matrix A of α with respect to the standard basis is real, but not symmetric, we see that $A^* \neq A$ so $\alpha^* \neq \alpha$. Thus α is a projection but not an orthogonal projection.

If α and β are orthogonal projections and $\alpha\beta = \beta\alpha$, then

$$(\alpha\beta)^2 = (\alpha\beta)(\alpha\beta) = \alpha(\beta\alpha)\beta = \alpha(\alpha\beta)\beta = \alpha^2\beta^2 = \alpha\beta$$

and $(\alpha\beta)^* = (\beta\alpha)^* = \alpha^*\beta^* = \alpha\beta$.

Exercise 14.3.15

We have

$$\rho_W^2 = 4\pi_W^2 - 4\pi_w + \iota = \iota.$$

$$\begin{aligned}\|\rho_W \mathbf{x}\|^2 &= \langle(2\pi_W - \iota)\mathbf{x}, (2\pi_W - \iota)\mathbf{x}\rangle = \langle(2\pi_W - \iota)\mathbf{x}, (2\pi_W - \iota)\mathbf{x}\rangle \\ &= \langle\mathbf{x}, \mathbf{x}\rangle = \|\mathbf{x}\|^2\end{aligned}$$

so ρ_W is an isometry.

$$\begin{aligned}\rho_W \rho_{W^\perp} &= (2\pi_W - \iota)(2\pi_{W^\perp} - \iota) \\ &= 4\pi_W \pi_{W^\perp} - 2\pi_W - 2\pi_{W^\perp} + \iota = 0 - 2\iota + \iota = -\iota.\end{aligned}$$

Exercise 15.1.1

$d_1(A,B) = \max_{i,j} |a_{ij} - b_{ij}| = \|A - B\|_1$.

(i) $|a_{ij}| \geq 0 \ \forall i, j$ so $\|A\|_1 \geq 0$.

(ii) $\|A\|_1 = 0 \Rightarrow |a_{ij}| = 0 \ \forall i, j \Rightarrow a_{ij} = 0 \ \forall i, j \Rightarrow A = 0$.

(iii) $\|\lambda A\|_1 = \max_{i,j} |\lambda a_{ij}| = \max_{i,j} |\lambda||a_{ij}| = |\lambda| \max_{i,j} |a_{ij}| = |\lambda|\|A\|_1$.

(iv) $|a_{ij} - b_{ij}| \leq |a_{ij}| + |b_{ij}| \ \forall i, j$ so

$$\max_{i,j} |a_{ij} - b_{ij}| \leq \max_{i,j} |a_{ij}| + \max_{i,j} |b_{ij}|$$

and $\|A + B\|_1 \leq \|A\|_1 + \|B\|_1$.

$d_2(A,B) = \sum_{i,j} |a_{ij} - b_{ij}| = \|A - B\|_2$.

(i) $|a_{ij}| \geq 0 \ \forall i, j$ so

$$\|A\|_2 = \sum_{i,j} |a_{ij}| \geq 0.$$

(ii) $\|A\|_2 = 0 \Rightarrow |a_{ij}| = 0 \ \forall i, j \Rightarrow a_{ij} = 0 \ \forall i, j \Rightarrow A = 0$.

(iii) We have

$$\begin{aligned} \|\lambda A\|_2 &= \sum_{i,j} |\lambda a_{ij}| = \sum_{i,j} |\lambda||a_{ij}| \\ &= |\lambda| \sum_{i,j} |a_{ij}| = |\lambda|\|A\|_2 \end{aligned}$$

(iv) We have

$$\begin{aligned} \|A + B\|_2 &= \sum_{i,j} |a_{ij} + b_{ij}| \leq \sum_{i,j} |a_{ij}| + |b_{ij}| \\ &= \sum_{i,j} |a_{ij}| + \sum_{i,j} |b_{ij}| = \|A\|_2 + \|B\|_2. \end{aligned}$$

$d_3(A,B) = \left(\sum_{i,j} |a_{ij} - b_{ij}|^2\right)^{1/2} = \|A - B\|_3$.

$\| \ \|_3$ is just the standard norm for an inner product space of dimension n^2.

Exercise 15.1.3

Observe that, using the Cauchy–Schwarz inequality,

$$\begin{aligned}
\|\alpha\mathbf{x}\| &= \left(\sum_{i=1}^{n}\left|\sum_{j=1}^{n}a_{ij}x_j\right|^2\right)^{1/2} \leq \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\sum_{j=1}^{n}|x_j|^2\right)^{1/2}\\
&= \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\|\mathbf{x}\|^2\right)^{1/2} \leq \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}\|\mathbf{x}\|.
\end{aligned}$$

Exercise 15.1.7

Since $U = (\iota - \alpha)^{-1}(\mathbf{0})$ and $W = \alpha^{-1}(\mathbf{0})$ are orthogonal complements, we can write any $\mathbf{v} \in V$ in the form

$$\mathbf{v} = \mathbf{u} + \mathbf{w}$$

with $\mathbf{u} \in U$ and $\mathbf{w} \in W$. Since $\mathbf{u} \perp \mathbf{w}$,

$$\|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{w}\|^2$$

and

$$\|\pi\mathbf{v}\| = \|\mathbf{u}\| \leq \|\mathbf{v}$$

so $\|\pi\| \leq 1$.

Since $\pi \neq 0$, we can find $\mathbf{x}$ with $\mathbf{u} = \pi\mathbf{x} \neq \mathbf{0}$ so $\|\mathbf{u}\| \neq 0$, Since $\pi(\mathbf{u} = \mathbf{u}$, we have $\|\pi\| = 1$.

We use row vectors. Let $K > 0$. If $\alpha(x, y) = (x + Ky, 0)$, then α is linear and $\alpha^2(x, y) = (x + Ky, 0) = \alpha(x, y)$, so α is a projection. Since $\|(0, 1)\| = 1$ and $\|\alpha(0, 1)\| = K$, $\|\alpha\| \geq K$.

Exercise 15.1.9

If $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ is an orthonormal basis for U, then the map

$$\theta\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = (x_1, x_2, \ldots, x_n)^T$$

is a linear inner product preserving (and so norm preserving) isomorphism $\theta : U \to \mathbb{F}^n$ where $\mathbb{F}^n$ is equipped with the standard dot product.

Exercise 15.1.10

Observe that
$$\|I\mathbf{x}\| = \|\mathbf{x}\|$$
so $\|I\| = 1$.

(i) $\|I + (-I)\| = \|0\| = 0$, $\|I\| = \|-I\| = 1$.

(ii) $\|I + I)\| = \|2I\| = 2$, $\|I\| = 1$.

(iv) $\|II\| = \|I\| = 1$.

We now attack (iii). If
$$J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$
then
$$\|J(x,y)^T\| = \|(y,0)^T\| = |y| \leq \|(x,y)^T\|$$
so $\|J\|| \leq 1$. However $\|(1,0)^T\| = \|(0,1)^T\| = 1$ and $J(0,1)^T = (1,0)^T$ so $\|J\| = 1$ but
$$\|JJ\| = \|0\| = 0.$$

Exercise 15.1.11

$$\det(tI - A) = (t-1)^2 - \mu^2 = t^2 - 2t + \eta$$

so the eigenvalues are $1 \pm (1-\eta^2)^{1/2}$

The eigenvectors corresponding to $1 + (1-\eta^2)^{1/2}$ are given by

$$x + \mu y = (1 + (1-\eta^2)^{1/2})x$$
$$\mu x + y = (1 + (1-\eta^2)^{1/2})y$$

i.e $x = y$ (as we could spot directly). The eigenvectors corresponding to $1 - (1-\eta^2)^{1/2}$ are given by

$$x + \mu y = (1 - (1-\eta^2)^{1/2})x$$
$$\mu x + y = (1 - (1-\eta^2)^{1/2})y$$

i.e $x = -y$ (as we could spot directly).

Thus we have a basis of orthonormal eigenvectors $\mathbf{e}_1 = 2^{-1/2}(1,1)^T$ eigenvalue $\lambda_1 = 1 + (1-\eta^2)^{1/2}$ and $\mathbf{e}_2 = 2^{-1/2}(1,-1)^T$ eigenvalue $\lambda_2 = 1 - (1-\eta^2)^{1/2}$.

We have

$$\begin{aligned}\|T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2\| &= \|\lambda_1\mathbf{x}_1\mathbf{e}_1 - \lambda_2 x_2\mathbf{e}_2\| \\ &= \big((\lambda_1 x_1)^2 + (\lambda_2 x_2)\big)^{1/2} \\ &\leq \lambda_1(x_1^2 + x_2^2)^{1/2} \qquad \leq \lambda_1\|(x_1\mathbf{e}_1 + x_2\mathbf{e}_2\|\end{aligned}$$

so $\|A\| \leq \lambda_1$ (in fact, since $A\mathbf{e}_1 = \lambda_1\mathbf{e}_1$ we have $\|A\| = \lambda_1$) and small changes in $\mathbf{x}$ produce small changes in $\mathbf{y}$.

Observe that taking, $\mathbf{y} = \mathbf{e}_2$, we have $\mathbf{x} = \mathbf{A}^{-1} = \lambda_2^{-1}$ which is very large.

(ii) Observe that looking at A^{-1} we have a basis of orthonormal eigenvectors $\mathbf{e}_1 = 2^{-1/2}(1,1)^T$ eigenvalue $\lambda_1 = 1 + (1-\eta^2)^{-1/2}$ and $\mathbf{e}_2 = 2^{-1/2}(1,-1)^T$ eigenvalue $\lambda_2 = 1 - (1-\eta^2)^{-1/2}$. Thus arguing as in (i), $\|A^{-1}\| = \lambda_2^{-1}$.

Now $\det B = \det A \det A^{-1} = \det AA^{-1} = \det I = 1$ and

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

By looking at vectors of the form $(0,0,x,y)$ we see that $\|B\| \geq \|A^{-1}\|$ and by looking at vectors of the form $(x,y,0,0)$ we see that $\|B\| \geq \|A^{-1}\|$. Thus if η is very small, both $\|B\|$ and $\|B^{-1}\|$ are very large.

Exercise 15.1.12

(i) We have $c(A) = \|A\|\|A^{-1}| \geq \|AA^{-1}\| = \|I\| = 1$.

(ii) We have
$$c(\lambda A) = \|\lambda A\|\|(\lambda A)^{-1}\| = \|\lambda A\|\|\lambda^{-1}A^{-1}| = |\lambda||\lambda^{-1}|\|A\|\|A^{-1}\| = c(A).$$

EXERCISE 15.1.13

Let us use an orthonormal basis $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$. Observe that

$$n^{-2}\sum_{i,j}|a_{ij}| \leq n^{-2}\sum_{i,j}\max_{r,s}|a_{rs}| = \max_{r,s}|a_{rs}|.$$

and that

$$\begin{aligned}|a_{rs}| &\leq \left(\sum_{j=1}^{n}|a_{js}|^2\right)^{1/2}\\ &= \left\|\sum_{j=1}^{n}a_{js}\mathbf{e}_j\right\|\\ &= \|\alpha\mathbf{e}_s\| \leq \|\alpha\|\|\mathbf{e}_s\| = \|\alpha\|.\end{aligned}$$

If $\mathbf{x} = \sum_{j=1}^{n} x_j\mathbf{e}_j$ then $|x_j| = |\langle\mathbf{x}, \mathbf{e}_j\rangle| \leq \|\mathbf{x}\|$ so

$$\begin{aligned}\|\alpha\mathbf{x}\| &= \left\|\sum_{r=1}^{n}x_r\alpha(\mathbf{e}_r)\right\| \leq \sum_{r=1}^{n}|x_r|\|\alpha(\mathbf{e}_r)\|\\ &\leq \|\mathbf{x}\|\sum_{r=1}^{n}\|\alpha(\mathbf{e}_r)\| = \|\mathbf{x}\|\sum_{r=1}^{n}\left\|\sum_{j=1}^{n}a_{jr}\mathbf{e}_j\right\|\\ &\leq \|\mathbf{x}\|\sum_{r=1}^{n}\sum_{j=1}^{n}|a_{jr}|\|\mathbf{e}_j\| \leq \|\mathbf{x}\|\sum_{r=1}^{n}\sum_{j=1}^{n}|a_{jr}|,\end{aligned}$$

so

$$\|\alpha\| \leq \sum_{i,j}|a_{ij}|.$$

Finally,

$$\sum_{i,j}|a_{ij}| \leq \sum_{i,j}\max_{r,s}|a_{rs}| = n^2\max_{r,s}|a_{rs}|.$$

Since

$$\|\alpha\mathbf{x}\| \leq \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}$$

for all $\|\mathbf{x}\| \leq 1$, we have

$$\|\alpha\mathbf{y}\| \leq \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}\|\mathbf{y}\|$$

for all $\mathbf{y}$ and

$$\|\alpha\| \leq \left(\sum_{i,j}a_{ij}^2\right)^{1/2}.$$

Exercise 15.2.2

A triangular matrix A has characteristic polynomial

$$\chi_A(t) = \prod_{j=1}^{n}(t - a_{ii})$$

so, if A is real, its characteristic equation has all its roots real. Thus, if α is triangularisable, its characteristic equation has all its roots real. Consider the matrix

$$B = \begin{pmatrix} C & 0 \\ 0 & I \end{pmatrix}$$

with I the $n-2 \times n-2$ identity matrix and

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

B has characteristic polynomial $(t-1)^{n-2}(t^2+1)$ so not all its roots are real. If β is the linear map corresponding to B with respect to some basis then β is not triangularisable.

If $\dim V = 1$ a 1×1 matrix is triangular so all endomorphisms are triangularisable.

The fact that $A = QR$ with R upper triangular and Q orthogonal does not imply that there is an invertible matrix P and an upper triangular matrix T with $A = PAP^{-1}$.

Exercise 15.2.4

Let A be the matrix of an endomorphism α with respect to some orthonormal basis. We can find $\alpha(n) \in \mathcal{L}(U, U)$ with $\|\alpha(n) - \alpha\| \to 0$ as $n \to \infty$ such that the characteristic equations of the $\alpha(n)$ have no repeated roots. Let $A(n)$ be the matrix of $\alpha(n)$ with respect to the given basis. Then

$$\max_{i,j} |a_{ij}(n) - a_{ij}| \leq \|\alpha(n) - \alpha\| \to 0$$

as $n \to \infty$.

Exercise 15.3.2

We have
$$\det(tI - A) = (t - 1/2)(t - 1/4),$$
so A has eigenvalues $1/2$ and $1/4$

By induction on n,
$$A^n(0,1)^T = (2^{-n+1}K, 0)^T$$
for all $n \geq 1$. Thus, if $K = 2^N L$, we have $\|A^N(0,1)^T\| > L$.

Exercise 15.3.3

We consider column vectors. Observe that, if $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ and $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, then

$$\begin{aligned}\|\mathbf{z}\|^2 &= \sum_{j=1}^{n} |z_j|^2 \\ &= \sum_{j=1}^{n} (|x_j|^2 + |y_j|^2) \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\end{aligned}$$

Using the notation introduced in the first paragraph,

$$\begin{aligned}\|A^n\mathbf{z}\|^2 &= \|A^n(\mathbf{x} + i\mathbf{y})\|^2 \\ &= \|A^n\mathbf{x} + iA^n\mathbf{y})\|^2 \\ &= \|A^n\mathbf{x}\|^2 + \|A^n\mathbf{y})\|^2,\end{aligned}$$

so

$$\|A^n\mathbf{x}\|, \|A^n\mathbf{y})\| \to 0 \Rightarrow \|A^n\mathbf{z})\| \to 0.$$

The result follows.

Exercise 15.3.7

(i) We have

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{j+s} a_{ik}b_{kj} = \sum_{j+s+1 \le k \le i-r+1} a_{ik}b_{kj} = 0$$

if $i - r \le j + s + 1$.

(ii) If $D = (d_{ij})$ is diagonal, then $\|D\| = \max |d_{jj}| = \rho(D)$.

Exercise 15.3.8

The computations by which we obtained the result for r, $s \geq 0$ remain valid and the result remains true.

Exercise 15.3.11

Let A be an $n \times n$ matrix with non-zero diagonal entries and $\mathbf{b} \in \mathbb{F}^n$ a column vector. Suppose that the equation

$$A\mathbf{x} = \mathbf{b}$$

has the solution $\mathbf{x}^*$. Let us write D for the $n \times n$ diagonal matrix with diagonal entries the same as those of A and set $B = A - D$. If $\mathbf{x}_0 \in \mathbb{R}^n$ and

$$\mathbf{x}_{j+1} = D^{-1}(\mathbf{b} - B\mathbf{x}_j),$$

then

$$\|\mathbf{x}_j - \mathbf{x}^*\| \leq \|D^{-1}B\|^j$$

and $\|\mathbf{x}_j - \mathbf{x}^*\| \to 0$ whenever $\rho(D^{-1}B) < 1$. If $\rho(D^{-1}) = 1$. we can find an $\mathbf{x}_0$ such that $\|\mathbf{x}_j - \mathbf{x}^*\| \not\to 0$.

To prove this, observe that $-D^{-1}B\mathbf{x}^* = D^{-1}\mathbf{b} - \mathbf{x}^*$ and so

$$\mathbf{x}_{j+1} - \mathbf{x}^* = D^{-1}(\mathbf{b} - B\mathbf{x}_j) - \mathbf{x}^* = -D^{-1}B(\mathbf{x}_j - \mathbf{x}^*)$$

Thus

$$\|\mathbf{x}_n - \mathbf{x}^*\| \leq \|(D^{-1}B)^n\|\|\mathbf{x}_0 - \mathbf{x}^*\|$$

If $\rho(D^{-1}B) < 1$ $\|(D^{-1}B)^n\| \to 0$ the result follows.

If $\rho(D^{-1}B) \geq 1$ we can find an eigenvector $\mathbf{e}$ with eigenvalue having absolute value at least 1. If we set $\mathbf{x}_0 = \mathbf{x}^* + \mathbf{e}_0$ convergence fails.

To prove the last part, suppose, if possible, that we can find an eigenvector $\mathbf{y}$ of $D^{-1}B$ with eigenvalue λ such that $|\lambda| \geq 1$. Then $D^{-1}B\mathbf{y} = \lambda\mathbf{y}$ and so $B\mathbf{y} = \lambda D\mathbf{y}$. Thus

$$\sum_{j \neq i}^{n} a_{ij}y_j = \lambda a_{ii}y_i$$

and so

$$|a_{ii}||y_i| \leq \sum_{j \neq i} |a_{ij}||y_j|$$

for each i and (since $\mathbf{y} \neq \mathbf{0}$)

$$|a_{ii}||y_i| < \sum_{j \neq i} |a_{ij}||y_j|$$

for at least one value of i.

Summing over all i and interchanging the order of summation, we get

$$\sum_{i=1}^{n} |a_{ii}||y_i| \leq \sum_{i=1}^{n} \sum_{j \neq i} |a_{ij}||y_j| = \sum_{j=1}^{n} \sum_{i \neq j} |a_{ij}||y_j|$$
$$= \sum_{j=1}^{n} |y_j| \sum_{i \neq j} |a_{ij}| < \sum_{j=1}^{n} |y_j||a_{jj}| = \sum_{i=1}^{n} |a_{ii}||y_i|$$

which is absurd.

Thus all the eigenvalues of $D^{-1}B$ have absolute value less than 1 and we may apply the first part of the question.

Exercise 15.4.3

If α has matrix A with respect to given orthonormal basis, α^* has matrix A^*, $\alpha\alpha^*$ has matrix AA^* and $\alpha^*\alpha$ has matrix A^*A.

Thus $\alpha\alpha^* = \alpha^*\alpha \Leftrightarrow A^*A = AA^*$.

Exercise 15.4.6

If α is diagonalisable with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$, then U is the direct sum of the spaces

$$E_j = \{\mathbf{e} : \alpha\mathbf{e} = \lambda_j\mathbf{e}\}.$$

Moreover $E_j \perp E_k$, that is to say

$$\mathbf{e}_j \in E_j,\ \mathbf{e}_k \in E_k \Rightarrow \langle \mathbf{e}_j, \mathbf{e}_k\rangle = 0.$$

when $j \neq k$. It follows that if π_j is the (unique) orthogonal projection with $\pi_j U = E_j$ we have $\pi_j\pi_k = 0$ for all $j \neq k$ If $\mathbf{u} \in U$, we can write $\mathbf{u}$ uniquely as

$$\bigstar \qquad \mathbf{u} = \mathbf{e}_1 + \mathbf{e}_2 + \ldots + \mathbf{e}_m$$

with $\mathbf{e}_j \in E_j$. Since $\pi_j\pi_k = 0$ for all $j \neq k$ we have $\pi_j\mathbf{e}_k = \pi_j\pi_k\mathbf{e}_k = 0$. Thus, applying π_j to both sides of $\bigstar$, $\pi_j\mathbf{u} = \mathbf{e}_j$. Thus

$$\iota\mathbf{u} = \sum_{j=1}^{m} \pi_j\mathbf{u}$$

and

$$\alpha\mathbf{u} = \sum_{j=1}^{m} \alpha\mathbf{e}_j = \sum_{j=1}^{m} \lambda_j\mathbf{e}_j = \sum_{j=1}^{m} \lambda_j\pi_j\mathbf{u}.$$

for all $\mathbf{u} \in U$. It follows that

$$\iota = \pi_1 + \pi_2 + \ldots + \pi_m \text{ and } \alpha = \lambda_1\pi_1 + \lambda_2\pi_2 + \ldots + \lambda_m\pi_m.$$

Conversely, if the stated conditions hold and we write $E_j = \pi_j U$, we see that E_j is a subspace and $\pi_i E_j = \pi_i\pi_j E_j = 0$ for $i \neq j$. Since $iota = \pi_1 + \pi_2 + \ldots + \pi_m$, we have

$$\mathbf{u} = \iota\mathbf{u} = \sum_{j=1}^{m} \pi_j\mathbf{u} = \sum_{j=1}^{m} \mathbf{u}_j$$

with $\mathbf{u}_j = \pi_j\mathbf{u} \in E_j$. On the other hand, if $\mathbf{e}_j \in E_j$ and

$$\sum_{j=1}^{m} \mathbf{e}_j = \mathbf{0}$$

then applying π_i to both sides we get $\mathbf{e}_i = \mathbf{0}$ for all i. Thus

$$U = E_1 \oplus E_2 \oplus \ldots \oplus E_m.$$

Moreover, since $\pi_j\pi_k = 0$ we have $E_j \perp E_k$ for $k \neq j$.

Let $\mathcal{E}_j$ be an orthonormal basis for E_j. The set $\mathcal{E} = \bigcup_{j=1}^{m} \mathcal{E}_j$ is an orthonormal basis for U. Since α has a diagonal matrix with respect to this basis, α is normal.

α is self-adjoint if and only if there exist distinct $\lambda_j \in \mathbb{R}$ and othogonal projections π_j such that $\pi_k \pi_j = 0$ when $k \neq j$,

$$\iota = \pi_1 + \pi_2 + \ldots + \pi_m \, textand \, \alpha = \lambda_1 \pi_1 + \lambda_2 \pi_2 + \ldots + \lambda_m \pi_m.$$

Exercise 15.4.7

If α is unitary, then $\alpha = \alpha^{-1}$ so

$$\alpha\alpha^* = \iota = \alpha\alpha^*$$

and α is normal.

$$\begin{aligned}
\beta \text{ unitary} &\Leftrightarrow \beta\beta^* = \iota\\
&\Leftrightarrow \alpha^{-1}\alpha^*(\alpha^{-1}\alpha^*)^* = \iota\\
&\Leftrightarrow \alpha^{-1}\alpha^*\alpha^{**}(\alpha^*)^{-1} = \iota\\
&\Leftrightarrow \alpha^{-1}\alpha^*\alpha(\alpha^*)^{-1} = \iota\\
&\Leftrightarrow \alpha^*\alpha = \alpha\alpha^*\\
&\Leftrightarrow \alpha \text{ normal}
\end{aligned}$$

Exercise 15.4.9

(i) Choose an orthonormal basis such that α is represented by a diagonal matrix D with $d_{jj} = e^{i\theta_j}$ with θ_j real. Let $f(t)$ be represented by a diagonal matrix D with $d_{jj}(t) = e^{i\theta_j t}$. Then

$$\|f(t) - f(s)\| \leq \max_j |e^{i\theta_j t} - e^{i\theta_j s}| \leq \max |\theta_j||t - s|$$

so f is continuous, $f(t)$ is unitary, $f(0) = \iota$, $f(1) = \alpha$.

Recall that the unitary maps form a group. Let $\alpha = \beta^{-1}\gamma$. Since α is unitary we can find f as in the first paragraph. Set $g(t) = \beta f(t)$.

(ii) By considering matrix representations, we know that $\det : \mathcal{L}(U, U) \to \mathbb{R}$ is continuous. Thus, if f is as stated, $\det f(t)$ is a continuous function of t. But $\det f(1) = -1$, $\det f(0) = 1$ so, by the intermediate value theorem, we can find an $s \in [0, 1]$ such that $\det f(s) = 0$ and so $f(s)$ is not invertible.

EXERCISE 16.1.5

If U is finite dimensional and

$$\alpha(\mathbf{x}, \mathbf{y}) = 0 \text{ for all } \mathbf{x} \in U \Rightarrow \mathbf{y} = 0,$$

then $\theta_R : U \to U'$ is an isomorphism.

Proof. We observe that the stated condition tells us that θ_R is injective since

$$\begin{aligned}\theta_R(\mathbf{y}) = 0 &\Rightarrow \theta_R(\mathbf{y})(\mathbf{x}) = 0 \text{ for all } \mathbf{x} \in U\\ &\Rightarrow \alpha(\mathbf{x}, \mathbf{y}) = 0 \text{ for all } \mathbf{x} \in U\\ &\Rightarrow \mathbf{y} = 0.\end{aligned}$$

Since $\dim U = \dim U'$, it follows that θ_R is an isomorphism. $\square$

Exercise 16.1.9

Observe, that if $\mathbf{x} \neq 0$, then at least one of the following two things must occur:-

(A) $x_1 \neq 0$, so $\beta\big((x_1, x_2)^T, (0, 1)^T\big) = x_1 \neq 0$.

(B) $x_2 \neq 0$, so $\beta\big((x_1, x_2)^T, (1, 0)^T\big) = -x_2 \neq 0$.

Thus β is non-degenerate, but

$$\beta(\mathbf{x}, \mathbf{x}) = x_1 x_2 - x_2 x_1 = 0$$

for all $\mathbf{x}$.

We cannot find an α of the type required. If α is degenerate, then there exists an $\mathbf{x} \neq \mathbf{0}$ with $\alpha(\mathbf{x}, \mathbf{y}) = 0$ for all $\mathbf{y}$ so, in particular, $\alpha(\mathbf{x}, \mathbf{x}) = 0$.

Exercise 16.1.14

Observe that, with the notation of Definition 16.1.12,

$$\begin{aligned}
q(\mathbf{u}+\mathbf{v}) + q(\mathbf{u}-\mathbf{v}) &= \alpha(\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}) + \alpha(\mathbf{u}-\mathbf{v},\mathbf{u}-\mathbf{v})\\
&= \alpha(\mathbf{u},\mathbf{u}+\mathbf{v}) + \alpha(\mathbf{v},\mathbf{u}+\mathbf{v})\\
&\qquad + \alpha(\mathbf{u},\mathbf{u}-\mathbf{v}) + \alpha(\mathbf{v},\mathbf{u}-\mathbf{v})\\
&= \alpha(\mathbf{u},\mathbf{u}) + \alpha(\mathbf{u},\mathbf{v}) + \alpha(\mathbf{v},\mathbf{u}) + \alpha(\mathbf{v},\mathbf{v})\\
&\qquad + \alpha(\mathbf{u},\mathbf{u}) - \alpha(\mathbf{u},\mathbf{v}) - \alpha(\mathbf{v},\mathbf{u}) + \alpha(\mathbf{v},\mathbf{v})\\
&= 2\alpha(\mathbf{u},\mathbf{u}) + 2\alpha(\mathbf{u},\mathbf{u})\\
&= 2\big(q(\mathbf{u}) + q(\mathbf{v})\big).
\end{aligned}$$

Exercise 16.1.16

Observe that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_i b_{ij} x_j = \sum_{j=1}^{n}\sum_{i=1}^{n} x_j b_{ij} x_i,$$

so

$$\sum_{i=1}^{n}\sum_{j=1}^{n} x_i b_{ij} x_j = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i \frac{b_{ij} + b_{ji}}{2} x_j$$

and q is a quadratic form with associated symmetric matrix

$$A = \tfrac{1}{2}(A + A^T).$$

EXERCISE 16.1.17

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 + x_3^2 = 1\}$$

is a sphere.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 + x_3^2 = 0\}$$

is a point.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 + x_3^2 = -1\} = \emptyset.$$

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 - x_3^2 = 1\}$$

is a one sheeted hyperboloid of revolution obtained by revolving the hyperbola

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 - x_3^2 = 1, x_2 = 0\}$$

about its semi-minor axis Ox_3.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 - x_3^2 = 0\}$$

is a circular cone.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 - x_3^2 = -1\}$$

is a two sheeted hyperboloid of revolution obtained by revolving the hyperbola

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 - x_3^2 = -1, x_2 = 0\}$$

about its semi-major axis Ox_3. (The two sheeted hyperboloid has two connected components. The one sheeted hyperboloid has one.)

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 = 1\}$$

is a circular cylinder.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 = 0\}$$

is the straight line give along the axis Ox_3.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 + x_2^2 = -1\} = \emptyset.$$

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 - x_2^2 = 1\} \text{ and } \mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 - x_2^2 = -1\}$$

are hyperbolic cylinders.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 - x_2^2 = 0\}$$

is a pair of planes intersecting at right angles.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 = 1\}$$

is a pair of parallel planes.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 = 0\}$$

is a plane.

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, x_1^2 = -1\} = \emptyset.$$

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, 0 = 1\} = \{\mathbf{x} \in \mathbb{R}^3 \,:\, 0 = -1\} = \emptyset$$

and

$$\{\mathbf{x} \in \mathbb{R}^3 \,:\, 0 = 0\} = \mathbb{R}^3.$$

Exercise 16.1.23

By a rotation (which changes neither volume nor determinant) we may suppose that our ellipsoid is

$$\sum_{i=1}^{n} d_i y_i^2 \leq L.$$

By successive scale changes $w_i = d_i^{1/2} y_i$ which multiply volume by $d_i^{1/2}$ we can transform our ellipsoid to

$$\sum_{i=1}^{n} w_i^2 \leq L$$

Now make a scale change in every direction $u_i = L^{-1/2} w_i$, to obtain the unit sphere of volume V_n

Our original volume must have been $L^{n/2} \prod_{j=1}^{n} d_j^{-1/2} V_n$ i.e,

$$(\det A)^{-1/2} L^{n/2} V_n.$$

Exercise 16.1.25

We know that

$$f_{\mathbf{Y}}(\mathbf{y}) = K \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} d_j y_j^2\right)$$

is a density function and so

$$\int_{\mathbb{R}^n} f_{\mathbf{Y}}(\mathbf{y})\, dV(\mathbf{y}) = 1.$$

Thus

$$\begin{aligned}
1 &= K \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\sum_{j=1}^{n} d_j y_j^2\right)\, dy_1 dy_2 \dots dy_n\\
&= K \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} \prod_{j=1}^{n} \exp(-d_j y_j^2/2)\, dy_1 dy_2 \dots dy_n\\
&= K \prod_{j=1}^{n} \int_{-\infty}^{\infty} \exp(-d_j y_j^2/2)\, dy_j\\
&= K \prod_{j=1}^{n} \left(d_j^{-1/2} \int_{-\infty}^{\infty} \exp(-t_j^2/2)\, dt_j\right)\\
&= K(2\pi)^{n/2} \prod_{j=1}^{n} d_j^{-1/2}.
\end{aligned}$$

Thus

$$K = (2\pi)^{-n/2}\left(\prod_{j=1}^{n} d_j\right)^{1/2}.$$

Exercise 16.1.26

(i) Suppose α sesquilinear, β Hermitian, γ skew-Hermitian and
$$\alpha = \beta + \gamma.$$
Then
$$\alpha(\mathbf{u},\mathbf{v}) = \beta(\mathbf{u},\mathbf{v}) + \gamma(\mathbf{u},\mathbf{v})$$
and
$$\alpha(\mathbf{v},\mathbf{u})^* = \big(\beta(\mathbf{v},\mathbf{u}) + \gamma(\mathbf{v},\mathbf{u})\big)^* = \beta(\mathbf{u},\mathbf{v}) - \gamma(\mathbf{v},\mathbf{u}).$$
Thus
$$\begin{aligned}\beta(\mathbf{u},\mathbf{v}) &= \tfrac{1}{2}(\alpha(\mathbf{u},\mathbf{v}) + \alpha(\mathbf{v},\mathbf{u})^*)\\ \gamma(\mathbf{u},\mathbf{v}) &= \tfrac{1}{2}(\alpha(\mathbf{u},\mathbf{v}) - \alpha(\mathbf{v},\mathbf{u})^*)\end{aligned}$$

Conversely, if α is sesquilinear, the definitions
$$\begin{aligned}\beta(\mathbf{u},\mathbf{v}) &= \tfrac{1}{2}(\alpha(\mathbf{u},\mathbf{v}) + \alpha(\mathbf{v},\mathbf{u})^*)\\ \gamma(\mathbf{u},\mathbf{v}) &= \tfrac{1}{2}(\alpha(\mathbf{u},\mathbf{v}) - \alpha(\mathbf{v},\mathbf{u})^*)\end{aligned}$$
give sesquilinear forms (the point at issue is that if $(\mathbf{u},\mathbf{v}) \mapsto \alpha(\mathbf{u},\mathbf{v})$ is sesquilinear, so is the map $(\mathbf{u},\mathbf{v}) \mapsto \alpha(\mathbf{v},\mathbf{u})^*$) with
$$\begin{aligned}\beta(\mathbf{v},\mathbf{u}) &= \tfrac{1}{2}(\alpha(\mathbf{v},\mathbf{u}) + \alpha(\mathbf{u},\mathbf{v})^*) = \beta(\mathbf{u},\mathbf{v})^*\\ \gamma(\mathbf{v},\mathbf{u}) &= \tfrac{1}{2}(\alpha(\mathbf{v},\mathbf{u})^* - \alpha(\mathbf{v},\mathbf{u})) = -\gamma(\mathbf{v},\mathbf{u})^*\end{aligned}$$

(ii) Since $\alpha(\mathbf{u},\mathbf{v}) = \alpha(\mathbf{v},\mathbf{u})^*$ for all $\mathbf{u}$, $\mathbf{v}$, we have $\alpha(\mathbf{u},\mathbf{u}) = \alpha(\mathbf{u},\mathbf{u})^*$, so $\alpha(\mathbf{u},\mathbf{u})$ is real for all $\mathbf{u} \in U$.

(iii) Set $\beta = i\alpha$. Then

$$\begin{aligned}&\alpha \text{ skew-Hermitian} \Leftrightarrow \alpha(\mathbf{v},\mathbf{u}) = -\alpha(\mathbf{u},\mathbf{v})^* \ \forall \mathbf{v},\mathbf{u}\\ &\Leftrightarrow \beta(\mathbf{u},\mathbf{v})^* = \big(i\alpha(\mathbf{u},\mathbf{v})\big)^* = -i\alpha(\mathbf{u},\mathbf{v})^* = i\alpha(\mathbf{v},\mathbf{u}) = \beta(\mathbf{v},\mathbf{u}) \ \forall \mathbf{v},\mathbf{u}\\ &\Leftrightarrow \beta \text{ Hermitian.}\end{aligned}$$

(iv) Observe that
$$\begin{aligned}\alpha(\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}) &= \alpha(\mathbf{u},\mathbf{u}) + \alpha(\mathbf{u},\mathbf{v}) + \alpha(\mathbf{v},\mathbf{u}) + \alpha(\mathbf{u},\mathbf{u})\\ &= \alpha(\mathbf{u},\mathbf{u}) + \alpha(\mathbf{u},\mathbf{v}) + \alpha(\mathbf{u},\mathbf{v})^* + \alpha(\mathbf{v},\mathbf{v})\\ &= \alpha(\mathbf{u},\mathbf{u}) + 2\Re\alpha(\mathbf{u},\mathbf{v}) + \alpha(\mathbf{v},\mathbf{v}).\end{aligned}$$
Thus
$$\alpha(\mathbf{u}-\mathbf{v},\mathbf{u}-\mathbf{v}) = \alpha(\mathbf{u},\mathbf{u}) + 2\Re\alpha(\mathbf{u},\mathbf{v}) + \alpha(\mathbf{v},\mathbf{v})$$
and
$$\alpha(\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}) - \alpha(\mathbf{u}-\mathbf{v},\mathbf{u}-\mathbf{v}) = 4\Re\alpha(\mathbf{u},\mathbf{v}).$$
It follows that
$$\alpha(\mathbf{u}+i\mathbf{v},\mathbf{u}+i\mathbf{v}) - \alpha(\mathbf{u}-i\mathbf{v},\mathbf{u}-i\mathbf{v}) = 4\Re\alpha(\mathbf{u},i\mathbf{v}) = -4\Re\big(i\alpha(\mathbf{u},\mathbf{v}) = \Im(\mathbf{u},\mathbf{v}).$$
The result follows.

(v) Choose any orthonormal basis $\mathbf{f}_1$, $\mathbf{f}_2$, …, $\mathbf{f}_n$ for U. Define a matrix $A = (a_{pq})$ by the formula

$$a_{pq} = \alpha(f_p, f_q).$$

By sesquilinearity

$$\alpha\left(\sum_{p=1}^{n} z_p \mathbf{f}_p, \sum_{q=1}^{n} w_q \mathbf{f}_q\right) = \sum_{p=1}^{n}\sum_{q=1}^{n} z_p a_{pq} w_q^*$$

for all z_p, $w_q \in \mathbb{C}$. Since α is Hermitian

$$a_{qp} = \alpha(f_q, f_p) = \alpha(f_p, f_q)^* = a_{pq}^*$$

Thus A is Hermitian and we can find a unitary matrix M such that $M^*AM = D$ where D is a real diagonal matrix whose entries are the eigenvalues of A appearing with the appropriate multiplicities. If we set $\mathbf{e}_q = \sum_{p=1}^{n} m_{pq} f_p$ then (since M is unitary) the $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ are an orthonormal basis and by direct calculation

$$\alpha\left(\sum_{r=1}^{n} z_r \mathbf{e}_r, \sum_{s=1}^{n} w_s \mathbf{e}_s\right) = \sum_{t=1}^{n} d_t z_t w_t^*$$

for all z_r, $w_s \in \mathbb{C}$.

Exercise 16.1.27

Suppose that U is vector space over $\mathbb{C}$ and $\alpha : U \times U \to \mathbb{C}$ is a bilinear form. Let us set

$$\theta_R(\mathbf{w})\mathbf{z} = \alpha(\mathbf{z}, \mathbf{w})$$

for $\mathbf{z}$, $\mathbf{w} \in U$.

We observe that

$$\begin{aligned}\theta_R(\mathbf{w})(\lambda_1\mathbf{z}_1 + \lambda_2\mathbf{z}_2) &= \alpha(\lambda_1\mathbf{z}_1 + \lambda_2\mathbf{z}_2, \mathbf{w})\\ &= \lambda_1\alpha(\mathbf{z}_1, \mathbf{w}) + \lambda_2\alpha(\mathbf{z}_2, \mathbf{w})\\ &= \lambda_1\theta_R(\mathbf{w})\mathbf{z}_1 + \lambda_2\theta_R(\mathbf{w})\mathbf{z}_2\end{aligned}$$

for all λ_1, $\lambda_2 \in \mathbb{C}$ and $\mathbf{z}_1$, $\mathbf{z}_2 \in U$. Thus $\theta_R(\mathbf{w}) \in U'$ for all $\mathbf{w} \in U$.

Now

$$\begin{aligned}\theta_R(\lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2)\mathbf{z}) &= \alpha(\mathbf{z}, \lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2)\\ &= \lambda_1^*\alpha(\mathbf{z}, \mathbf{w}_1) + \lambda_2^*\alpha(\mathbf{z}, \mathbf{w}_2)\\ &= \lambda_1^*\theta_R(\mathbf{w}_1)\mathbf{z} + \lambda_2^*\theta_R(\mathbf{w}_2)\mathbf{z}\\ &= \big(\lambda_1^*\theta_R(\mathbf{w}_1) + \lambda_2^*\theta_R(\mathbf{w}_2)\big)\mathbf{z}\end{aligned}$$

for all $\mathbf{z} \in U$. Thus

$$\theta_R(\lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2) = \lambda_1^*\theta_R(\mathbf{w}_1) + \lambda_2^*\theta_R(\mathbf{w}_2)$$

and $\theta_R : U \to U'$ is an anti-linear map.

Suppose in addition we know that U is finite dimensional and that

$$\alpha(\mathbf{z}, \mathbf{w}) = 0 \text{ for all } \mathbf{z} \in U \Rightarrow \mathbf{w} = 0.$$

Then θ_R is injective since

$$\begin{aligned}\theta_R(\mathbf{w}) = 0 &\Rightarrow \theta_R(\mathbf{w})(\mathbf{z}) = 0 \text{ for all } \mathbf{z} \in U\\ &\Rightarrow \alpha(\mathbf{z}, \mathbf{w}) = 0 \text{ for all } \mathbf{w} \in U\\ &\Rightarrow \mathbf{w} = 0.\end{aligned}$$

Since $\dim U = \dim U'$, it follows that θ_R is an isomorphism.

Alternatively, observe that if $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ form a basis for U

$$\begin{aligned}\sum_{j=1}^n \lambda_j\theta_R\mathbf{e}_j = \mathbf{0} &\Rightarrow \theta_R\left(\sum_{j=1}^n \lambda_j^*\mathbf{e}_j\right) = \mathbf{0}\\ &\Rightarrow \sum_{j=1}^n \lambda_j^*\mathbf{e}_j = \mathbf{0}\\ &\Rightarrow \lambda_j^* = 0\ \forall j \Rightarrow \lambda_j = 0\ \forall j\end{aligned}$$

so $\theta_R\mathbf{e}_1$, $\theta_R\mathbf{e}_2$, …, $\theta_R\mathbf{e}_n$ are linearly independent so a basis for U' so span U'. Thus θ_R is surjective.

Exercise 16.2.4

Let $\mathbf{f}_1$, $\mathbf{f}_2$, $\ldots$, $\mathbf{f}_n$ be the basis associated with S. Define

$$\alpha\left(\sum_{i=1}^{n} x_i\mathbf{f}_i, \sum_{j=1}^{n} y_j\mathbf{f}_j\right) = \sum_{k=1}^{n} x_k y_k$$

for all x_i, $y_j \in \mathbb{R}$. Then, by inspection, α is a symmetric form on $\mathbb{R}^n$.

By Theorem 16.2.1 we can find a basis $\mathbf{e}_1$, $\mathbf{e}_2$, $\ldots$, $\mathbf{e}_n$ and positive integers p and m with $p+m \leq n$ such that

$$\alpha\left(\sum_{i=1}^{n} x_i\mathbf{e}_i, \sum_{j=1}^{n} y_j\mathbf{e}_j\right) = \sum_{k=1}^{p} x_k y_k - \sum_{k=p+1}^{p+m} x_k y_k$$

for all x_i, $y_j \in \mathbb{R}$. We have

$$q\left(\sum_{i=1}^{n} y_i\mathbf{e}_i\right) = \alpha\left(\sum_{i=1}^{n} y_i\mathbf{e}_i, \sum_{j=1}^{n} y_j\mathbf{e}_j\right) = \sum_{k=1}^{p} y_k^2 - \sum_{k=p+1}^{p+m} y_k^2$$

as required.

Exercise 16.2.8

A is two quadrants, B two lines at right angles.

$(1, 0)$, $(-1, 1) \in A$, but $(0, 1) = (1, 0) + (-1, 1) \notin A$. Thus A is not a subspace.

$(1, 0)$, $(-1, 1) \in B$, but $(0, 1) = (1, 0) + (-1, 1) \notin B$. Thus A is not a subspace.

Exercise 16.2.9

(i) Choose M orthogonal so that $M^T AM = D$ with D diagonal. Since matrix rank is unchanged by multiplication on the left or right by invertible matrices, D has the same matrix rank as A. But the signature rank of A and the matrix rank of D are both the number of non-zero entries of D. Thus the signature and matrix ranks are identical.

(ii) We continue with the notation of (i). The rank of q is the number of non-zero characteristic roots (multiple roots counted multiply). Now

$$\text{signature } q = (\text{no. strictly pos. entries } D) - (\text{no. strictly neg. entries } D),$$

so the signature of D is the number of strictly positive characteristic roots minus the number of strictly negative characteristic roots of A (multiple roots counted multiply).

EXERCISE 16.2.10

(i) Observe that

$$\gamma(P\mathbf{z}) = (P\mathbf{z}^T)A(P\mathbf{z}) = (\mathbf{z}^T P^T)A(P\mathbf{z}) = \mathbf{z}^T(P^T AP)\mathbf{z}.$$

(ii) We cannot use results on the diagonalisation of real symmetric matrices or Hermitian matrices, but we can use completing the square.

(α) Suppose that $A = (a_{ij})_{1\leq i,j\leq n}$ is a symmetric $n \times n$ matrix with $a_{11} \neq 0$, and we set $b_{ij} = a_{11}^{-1}(a_{11}a_{ij} - a_{1i}a_{1j})$. Then $B = (b_{ij})_{2\leq i,j\leq n}$ is a symmetric matrix. Further, if $\mathbf{z} \in \mathbb{C}^n$ and (after choosing one square root $a_{11}^{1/2}$ of a_{11}) we set

$$w_1 = a_{11}^{1/2}\left(z_1 + \sum_{j=2}^{n} a_{1j}a_{11}^{-1/2}z_j\right)$$

and $w_j = z_j$ otherwise, we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} z_i a_{ij} z_j = w_1^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} w_i b_{ij} w_j.$$

(β) Suppose that $A = (a_{ij})_{1\leq i,j\leq n}$ is a symmetric $n \times n$ matrix. Suppose further that $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ is a permutation (that is to say σ is a bijection) and we set $b_{ij} = a_{\sigma i\,\sigma j}$. Then $C = (c_{ij})_{1\leq i,j\leq n}$ is a symmetric matrix. Further, if $\mathbf{z} \in \mathbb{C}^n$ and we set $w_j = x_{\sigma(j)}$, we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} z_i a_{ij} z_j = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i c_{ij} w_j.$$

(γ) Suppose that $n \geq 2$, $A = (a_{ij})_{1\leq i,j\leq n}$ is a symmetric $n\times n$ matrix and $a_{11} = a_{22} = 0$, but $a_{12} \neq 0$. Then there exists a symmetric $n \times n$ matrix C with $c_{11} \neq 0$ such that, if $\mathbf{z} \in \mathbb{C}^n$ and we set $w_1 = (z_1 + z_2)/2$, $w_2 = (z_1 - z_2)/2$, $w_j = z_j$ for $j \geq 3$, we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} z_i a_{ij} z_j = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i c_{ij} w_j.$$

Combining (α), (β) and (γ) we have the following.

(δ) Suppose that $A = (a_{ij})_{1\leq i,j\leq n}$ is a non-zero symmetric $n \times n$ matrix. Then we can find an $n \times n$ invertible matrix $M = (m_{ij})_{1\leq i,j\leq n}$ and a symmetric $(n-1) \times (n-1)$ matrix $B = (b_{ij})_{2\leq i,j\leq n}$ such that, if $\mathbf{z} \in \mathbb{C}^n$ and we set $w_i = \sum_{j=1}^{n} m_{ij}z_j$, then

$$\sum_{i=1}^{n}\sum_{j=1}^{n} z_i a_{ij} z_j = w_1^2 + \sum_{i=2}^{n}\sum_{j=2}^{n} w_i b_{ij} w_j.$$

Repeated use of (δ) gives the desired result.

(iii) Observe that if P is invertible so is P^T, so (using matrix rank)

$$\operatorname{rank}(P^T AP) = \operatorname{rank} A.$$

Thus, if

$$\gamma(P\mathbf{z}) = \sum_{u=1}^{r} z_u^2,$$

we have $r = \operatorname{rank} A$.

(iv) Observe that $\gamma(i\mathbf{z}) = -\gamma(\mathbf{z})$.

(v) By (i) and (ii) we can choose a basis $\mathbf{e}_j$ for $\mathbb{C}^n$ so that

$$\gamma\left(\sum_{j=1}^{n} z_j \mathbf{e}_j\right) = \sum_{u=1}^{r} z_u^2.$$

Let s be the integer part of $r/2$. If F is the subspace spanned by the vectors $\mathbf{e}_{2k-1} - i\mathbf{e}_{2k}$ with $1 \leq k \leq s$ and the vectors $\mathbf{e}_l$ with $r+1 \leq l \leq n$. Then F has dimension at least m and any subspace E of F with dimension m will have desired properties.

(vi) Let $E_0 = \mathbb{C}^m$. Using (v) and proceeding inductively we can find subspaces E_j such that E_j is a subspace of E_{j-1}, $\dim E_j \geq [2^{-j}n]$ (using the standard integer part notation) and $\gamma_j|E_j = 0$ $[1 \leq j \leq k]$. We have $\dim E_k \geq 1$, so we may choose a non-zero $\mathbf{z} \in E_k$ to obtain the required result.

Exercise 16.2.12

If $a = 0$, we have Example 16.2.11. The rank is 3 and the signature -1.

If $a \neq 0$, then, setting $y_1 = (x_1 - 2^{-1}a^{-1}x_2 - 2^{-1}a^{-1})$, $y_2 = x_2$, $y_3 = x_3$, we have
$$x_1x_2+x_2x_3+x_3x_1+ax_1^2 = ay_1^2+4^{-1}a^{-1}y_2^2+4^{-1}a^{-1}y_2^2+(1-2^{-1}a^{-1})y_2y_3$$
Setting $u_1 = y_1$, $u_2 = y_2 + 2a(1 - 2^{-1}a^{-1})y_3$ $u_3 = y_3$, we have
$$x_1x_2 + x_2x_3 + x_3x_1 + ax_1^2 = au_1^2 + 4^{-1}a^{-1}u_2^2 - a(1 - 2^{-1}a^{-1})^2u_3.$$

If $a = 1/2$, we see that q has rank and signature 2.

If $a > 0$ and $a \neq 1/2$, q has rank 3 and signature 1.

If $a < 0$, q has rank 3 and signature -1.

Exercise 16.2.13

(i) We know that we can find a unitary matrix U such that $C = U^*AU$ is diagonal with real entries. Let D be the diagonal matrix with $d_{ii} = 1$ if $c_{ii} = 0$ and $d_{ii} = c_{ii}^{-1/2}$ (positive square root) otherwise. Setting $P = DU$ we see that P^*AP is diagonal with diagonal entries taking the values 1, -1 or 0.

(ii) We have

$$\begin{aligned}\left(\sum_{r=1}^{n}\sum_{s=1}^{n} z_r a_{rs} z_s^*\right)^* &= \sum_{r=1}^{n}\sum_{s=1}^{n} z_r^* a_{rs}^* z_s\\ &= \sum_{r=1}^{n}\sum_{s=1}^{n} z_r^* a_{sr} z_s\\ &= \sum_{r=1}^{n}\sum_{s=1}^{n} z_r a_{rs} z_s^*\end{aligned}$$

so $\sum_{r=1}^{n}\sum_{s=1}^{n} z_r a_{rs} z_s^*$ is real.

(iii) If $P_1^*AP_1$ has k entries 1, then if E is the subspace of $\mathbb{C}^n$ spanned by $P_1\mathbf{e}$ with $\mathbf{e}$ a column vector with ith entry 1 and all other entries 0 when the ith diagonal entry of $P_1^*AP_1$ is 1, we have

$$\dim E = k, \mathbf{z}^*A\mathbf{z} > 0\ \forall \mathbf{z} \in E \setminus \{\mathbf{0}\}.$$

If $P_2^*AP_2$ has k' entries 1 then if F is the subspace of $\mathbf{C}^n$ spanned by $P_2\mathbf{e}$ with $\mathbf{e}$ a column vector with ith entry 1 and all other entries 0 when the ith diagonal entry of $P_1^*AP_1$ is -1 or 0, we have

$$\dim F = n - k', \mathbf{z}^*A\mathbf{z} \leq 0\ \forall \mathbf{z} \in F.$$

Thus $E \cap F = \{\mathbf{0}\}$, $\dim E + \dim F \leq n$ and $n - k' + k \leq n$. It follows that $k \leq k'$. The same argument shows that $k' \leq k$ so $k = k'$.

In the same way (or by considering $-A$) we see that $P_1^*AP_1$ and $P_2^*AP_2$ have the same number of entries -1 and so since they have the same number of entries 1 and -1 must have the same number of entries 0 along the diagonal.

Exercise 16.3.2

(i) Let U be a vector space over $\mathbb{R}$. A quadratic form $q : U \to \mathbb{R}$ is said to be *negative semi-definite* if

$$q(\mathbf{u}) \leq 0 \text{ for all } \mathbf{u} \in U$$

and *strictly negative definite* if

$$q(\mathbf{u}) < 0 \text{ for all } \mathbf{u} \in U \text{ with } \mathbf{u} \neq \mathbf{0}.$$

By inspection $q : U \to \mathbb{R}$ is strictly negative definite (respectively negative semi-definite) if and only if $-q$ is strictly positive definite (respectively positive semi-definite).

(ii) If q is a quadratic form over a finite dimensional real vector space, then we can find a basis $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$ such that

$$q\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = \sum_{j=1}^{r} x_j^2 - \sum_{j=r+1}^{r+s} x_j^2.$$

Setting

$$q_1\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = \sum_{j=1}^{r} x_j^2 \text{ and } q_2\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = -\sum_{j=r+1}^{r+s} x_j^2$$

we see that $q = q_1 + q_2$ with q_1 a positive semi-definite quadratic form and q_2 a negative semi-definite quadratic form.

The observation

$$0 = 0 + 0 = x^2 - x^2$$

shows that the decomposition is not unique even for a space of dimension 1.

Exercise 16.3.3

We have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i(\mathbb{E}X_iX_j)c_j = \mathbb{E}\left(\sum_{j=1}^{n} c_jX_j\right)^2 \geq 0$$

with equality if and only if

$$\Pr\left(\sum_{j=1}^{n} c_jX_j = 0\right) = 1.$$

Exercise 16.3.4

We know that we can find a basis $\mathbf{e}_j$ such that
$$\alpha\left(\sum_{j=1}^{n} x_j\mathbf{e}_j\right) = \sum_{j=1}^{u} x_j^2 - \sum_{j=u+1}^{v} x_j^2.$$

If $v < n$, then $q(\mathbf{e}_n) = 0$ and q is not strictly positive definite. If $v \geq u+1$, $q(\mathbf{e}_{u+1}) = -1$ so q is not positive semi-definite.

By inspection, $u = v = n$ implies q strictly positive definite so q is strictly positive definite if and only if $n = u = V$ i.e. q has rank and signature n.

By inspection, $u = v$ implies q positive semi-definite so q is positive semi-definite if and only if $u = v$, i.e. q has rank and signature equal.

Finally q is negative semi-definite if and only if $-q$ is positive semi-definite, i.e. q has signature equal to minus its rank.

Exercise 16.3.5

The conditions

(i) $\alpha(\mathbf{u}, \mathbf{v}) = \alpha(\mathbf{v}, \mathbf{u})$

(ii) $\alpha(\lambda\mathbf{u}, \mathbf{v}) = \lambda\alpha(\mathbf{u}, \mathbf{v})$

(iii) $\alpha(\mathbf{u} + \mathbf{w}, \mathbf{v}) = \alpha(\mathbf{u}, \mathbf{v}) + \alpha(\mathbf{w}, \mathbf{v})$

(for all $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in U$ and $\lambda \in \mathbb{R}$) say that α is a quadratic form.

The additional conditions

(iv) $\alpha(\mathbf{u}, \mathbf{u}) \geq 0\ \forall \mathbf{u} \in U$

(v) $\alpha(\mathbf{u}, \mathbf{u}) = 0 \Rightarrow \mathbf{u} = \mathbf{0}$

say that α gives rise to a positive definite quadratic form.

Conditions (i) to (v) say that α is an inner product.

Exercise 16.3.6

By translation we may take $\mathbf{a} = \mathbf{0}$ and by subtracting a constant we may take $f(\mathbf{0}) = 0$.

Since f is smooth we know by the local Taylor's theorem that

$$f(\mathbf{h}) = \frac{\partial f}{\partial x_i}(\mathbf{0})h_i + \frac{\partial f}{\partial x_i \partial x_j}(\mathbf{0})h_i h_j + \epsilon(\mathbf{h})\|\mathbf{h}\|^2$$

with $\epsilon(\mathbf{h}) \to 0$ as $\|\mathbf{h}\| \to 0$. (We use the summation convention.)

(i) Suppose we have a minimum at $\mathbf{0}$. If $\mathbf{e}_i$ is the vector with 1 in the ith place and 0 elsewhere

$$\frac{\partial f}{\partial x_i}(\mathbf{0}) = \lim_{h_i \to 0+} \frac{f(h\mathbf{e}_i)}{h} \geq 0$$

and

$$\frac{\partial f}{\partial x_i}(\mathbf{0}) = \lim_{h_i \to 0-} \frac{f(h\mathbf{e}_i)}{h} \leq 0$$

so $\frac{\partial f}{\partial x_i} = 0$.

We now have

$$f(\mathbf{h}) = H_{ij}h_i h_j + \epsilon(\mathbf{h})\|\mathbf{h}\|^2$$

with H the Hessian. Let $\mathbf{u}$ be any non-zero vector We have

$$0 \leq f(h\mathbf{u}) = h^2\big(H_{ij}u_i u_j + \epsilon(h\mathbf{u})\|\mathbf{u}\|^2\big),$$

so

$$0 \leq H_{ij}u_i u_j + \epsilon(h\mathbf{u})\|\mathbf{u}\|^2$$

and, allowing $h \to 0$,

$$0 \leq H_{ij}u_i u_j.$$

Thus H is positive semi-definite.

(ii) We have

$$f(\mathbf{h}) = H_{ij}h_i h_j + \epsilon(\mathbf{h})\|\mathbf{h}\|^2$$

with H the Hessian. Since $\mathbf{u} \mapsto H\mathbf{u}$ is continuous and the surface S of the unit ball is compact, we know that H takes a minimum on S which is non-zero so there exists an $\eta > 0$ with

$$H(\mathbf{u}) \geq \eta$$

for all $\|\mathbf{u}\| = 1$.

Thus

$$f(\mathbf{h}) \geq \|\mathbf{h}\|^2\big((\eta - \epsilon(\mathbf{h})\big).$$

There exists a $\delta > 0$ such that $|\epsilon(\mathbf{h})| < \eta/2$ for all $\|\mathbf{h}\| < \delta$.

If $\|\mathbf{h}\| < \delta$, then $f(\mathbf{h}) \geq (\eta/2)\|\mathbf{h}\|^2$.

Thus f has a strict minimum at $\mathbf{0}$.

(ii) Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^4$. We have a strict minimum at 0, but the Hessian $H = \big(f''(0)\big) = (0)$ is not strictly positive definite.

Exercise 16.3.7

Since

$$\left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right) = (\cos x \sin y, \sin x \cos y)$$

the stationary points are

$$(x,y) = \big(n\pi, m\pi\big) \text{ and}(x,y) = \big((n+\tfrac{1}{2})\pi, (m+\tfrac{1}{2})\pi\big).$$

The Hessian

$$H = \begin{pmatrix} -\sin x \sin y & \cos x \cos y \\ \cos x \cos y & -\sin x \sin y \end{pmatrix}$$

so if $(x,y) = \big(n\pi, m\pi\big)$

$$H = \begin{pmatrix} 0 & (-1)^{n+m} \\ (-1)^{n+m} & 0 \end{pmatrix}$$

which is neither positive semi-definite nor negative semi-definite (note for example that $4xy = (x+y)^2 - (x-y)^2$) and we have a saddle, whilst if $(x,y) = \big((n+\frac{1}{2})\pi, (m+\frac{1}{2})\pi\big)$

$$H = \begin{pmatrix} (-1)^{n+m} & 0 \\ 0 & (-1)^{n+m} \end{pmatrix}$$

which is strictly positive definite if $n+m$ is even (so we have a minimum) and strictly negative definite if $n+m$ is even (so we have a maximum).

Exercise 16.3.11

We have
$$\tilde{A}^T = (\tilde{L}\tilde{L}^T)^T = \tilde{L}^{TT}\tilde{L}^T = \tilde{L}\tilde{L}^T = \tilde{A},$$
so $\tilde{A}$ is symmetric.

Observe that
$$\mathbf{x}^T\tilde{A}\mathbf{x} = \mathbf{x}^T\tilde{L}\tilde{L}^T\mathbf{x} = \|\tilde{L}^T\mathbf{x}\|^2 \geq 0$$
so $\tilde{A}$ is symmetric positive semi-definite. The case $L = 0$ shows that $\tilde{A}$ need not be strictly positive definite.

If L is lower triangular with non-zero diagonal entries, then $\det L \neq 0$, so L is non-singular and
$$\begin{aligned}
\mathbf{x}^T A\mathbf{x} = 0 &\Rightarrow \mathbf{x}^T LL^T\mathbf{x} = 0\\
&\Rightarrow \|L^T\mathbf{x}\|^2 = 0\\
&\Rightarrow \|L^T\mathbf{x}\| = 0\\
&\Rightarrow L^T\mathbf{x} = 0\\
&\Rightarrow \mathbf{x} = \mathbf{0},
\end{aligned}$$
so $A = LL^T$ is a symmetric strictly positive definite matrix.

Exercise 16.3.12

If, at the first stage $a_{11} < 0$, then A is not positive definite.

If $a_{11} > 0$, then either the Schur matrix B is positive definite, in which case, A is or the Schur matrix B is not positive definite, in which case, A is not. (Consider vectors whose first term is 0.)

If at each stage, the upper corner entry is strictly positive then our proof shows that A is positive definite. If at some stage it is not, we know that A is not positive definite.

Exercise 16.3.14

We need to look at $\mathbf{l}_1 = (2, -3, 1)^T$. Now

$$A - \mathbf{l}_1\mathbf{l}_1^T = \begin{pmatrix} 4 & -6 & 2 \\ -6 & 8 & -5 \\ 2 & -5 & 14 \end{pmatrix} - \begin{pmatrix} 4 & -6 & 2 \\ -6 & 9 & -3 \\ 2 & -3 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & -2 & 13 \end{pmatrix}.$$

The Schur matrix

$$B = \begin{pmatrix} -1 & -2 \\ -2 & 13 \end{pmatrix}$$

has strictly negative upper corner entry, so A is not positive semi-definite.

Exercise 16.3.15

(i) Look at the first formula of the proof of part (ii) of Theorem16.3.10.

(ii) Looking at the proof of of part (i) of Theorem16.3.10 we see that the computation of $\mathbf{l}$ involves roughly n operations and the computation of B involves roughly n^2 operations. Thus the reduction of our problem from $n \times n$ matrices to $(n-1) \times (n-1)$ matrices involves roughly $2n^2$ operations and the total number of operations required is roughly

$$2\sum_{r=1}^{n} r^2 \approx \frac{2}{3}n^3$$

operations.

[If the reader's definition of an operation agrees with mine, this argument shows that we can certainly take $K = 1$ (and, if we really needed, as close to $2/3$ as we wished) for n large.]

Exercise 16.3.16

If we take $\mathbf{l}_1 = (1, \frac{1}{2}, \frac{1}{3})^T$, then

$$H_3 - \mathbf{l}_1\mathbf{l}_1^T = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix} - \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{9} \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{12} & \frac{1}{12} \\ 0 & \frac{1}{12} & \frac{4}{45} \end{pmatrix}$$

If we take $\mathbf{l}_2 = \big((\frac{1}{12})^{1/2}, (\frac{1}{12})^{1/2}\big)^T$, then

$$\begin{pmatrix} \frac{1}{12} & \frac{1}{12} \\ \frac{1}{12} & \frac{4}{45} \end{pmatrix} - \mathbf{l}_2\mathbf{l}_2^T = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{180} \end{pmatrix}$$

Thus $H_3 = LL^T$ with

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2^{-1} & 12^{-1/2} & 0 \\ 3^{-1} & 12^{-1/2} & 180^{-1/2} \end{pmatrix}$$

Second part

If we take $\mathbf{l}_1 = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4})^T$, then

$$H_4 - \mathbf{l}_1\mathbf{l}_1^T = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \lambda \end{pmatrix} - \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{6} & \frac{1}{8} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{9} & \frac{1}{12} \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{12} & \frac{1}{16} \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{12} & \frac{1}{12} & \frac{3}{40} \\ 0 & \frac{1}{12} & \frac{4}{45} & \frac{1}{12} \\ 0 & \frac{3}{40} & \frac{1}{12} & \lambda - \frac{1}{16} \end{pmatrix}$$

If we take $\mathbf{l}_2 = ((12)^{-1/2}, (12)^{-1/2}, (12)^{1/2}\frac{3}{40})^T$ then

$$\begin{pmatrix} \frac{1}{12} & \frac{1}{12} & \frac{3}{40} \\ \frac{1}{12} & \frac{4}{45} & \frac{1}{12} \\ \frac{3}{40} & \frac{1}{12} & \lambda - \frac{1}{16} \end{pmatrix} - \mathbf{l}_2\mathbf{l}_2^T = \begin{pmatrix} \frac{1}{12} & \frac{1}{12} & \frac{3}{40} \\ \frac{1}{12} & \frac{4}{45} & \frac{1}{12} \\ \frac{3}{40} & \frac{1}{12} & \lambda - \frac{1}{16} \end{pmatrix} - \begin{pmatrix} \frac{1}{12} & \frac{1}{12} & \frac{3}{40} \\ \frac{1}{12} & \frac{1}{12} & \frac{3}{40} \\ \frac{3}{40} & \frac{3}{40} & \frac{27}{400} \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 180^{-1} & 120^{-1} \\ 0 & 120^{-1} & \lambda - \frac{13}{100} \end{pmatrix}$$

If we take $\mathbf{l}_3 = ((180)^{-1/2}, (180)^{1/2}120^{-1})^T)$

$$\begin{pmatrix} 180^{-1} & 120^{-1} \\ 120^{-1} & \lambda - \frac{13}{100} \end{pmatrix} - \mathbf{l}_3\mathbf{l}_3^T = \begin{pmatrix} 0 & 0 \\ 0 & \lambda - \frac{57}{400} \end{pmatrix}$$

Thus $\lambda_0 = 57/400$. We observe that

$$\lambda_0 - \frac{1}{7} = \frac{1}{2800}.$$

(The 'Hilbert matrices' H_n are 'only just positive definite' and 'only just invertible' and are often used to test computational methods.)

If $\lambda \geq \lambda_0$ then $H_4 = LL^T$ with

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2^{-1} & 12^{-1/2} & 0 & 0 \\ 3^{-1} & 12^{-1/2} & 180^{-1/2} & 0 \\ 4^{-1} & (12)^{1/2}\frac{3}{40} & (180)^{1/2}120^{-1} & (\lambda - \frac{57}{400})^{1/2} \end{pmatrix}$$

Exercise 16.3.17

(i) The result is true when $n = 1$, since then the only root is a_0 which is positive by hypothesis.

Suppose the result is true for $m \geq n \geq 1$ and P is a polynomial of the given form of degree $m + 1$. We observe that P' is a polynomial of the given form of degree m and so all its real zeros are strictly positive. Thus, if $m + 1$ is even, $P'(t) < 0$ for $t \leq 0$, so P is decreasing as t runs from $-\infty$ to 0. But $P(0) = a_0 > 0$ so $P(t) > 0$ for $t \leq 0$ and P has no non-positive real zeros. If $m + 1$ is odd, $P'(t) > 0$ for $t \leq 0$, so P is increasing as t runs from $-\infty$ to 0. But $P(0) = -a_0 < 0$, so $P(t) < 0$ for $t \leq 0$ and P has no non-positive real zeros.

The result follows by induction.

(ii) The only if part is immediate. To obtain the if part apply part (i) to

$$P(t) = \det(tI - A) = \prod_{j=1}^{n}(t - \lambda_j) = t^n + \sum_{j=1}^{n}(-1)^{n-j}a_j t^j,$$

observing that

$$a_{n-1} = \sum_{j=1}^{n}\lambda_j,\ a_{n-2} = \sum_{j\neq i}\lambda_i\lambda_j > 0,\ a_{n-3} = \sum_{i,j,k \text{ distinct}}\lambda_i\lambda_j\lambda_k.\ \ldots.$$

(iii) Let

$$A = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix}.$$

Then

$$\det(tI - A) = (t-1)^2 + 9 = t^2 - 2t + 8$$

is a polynomial with no real roots (since $(t-1)^2 + 9 > 0$ for all real t).

(iv) Observe that

$$\det(tI - A) = t^3 - b_2t^2 + b_1t - b_0$$

with

$$b_2 = a_{11} + a_{22} + a_{33}$$
$$b_1 = a_{11}a_{22} + a_{22}a_{33} + a_{33}a_{22} - a_{12}a_{21} - a_{23}a_{32} - a_{31}a_{13} \quad > 0$$
$$b_0 = \det A$$

and apply the results above.

Exercise 16.3.19

Suppose P is invertible with P^TAP and P^TBP diagonal. Since A is not semi-positive definite P^TAP is not and, since A is not semi-negative definite, P^TAP is not. Thus

$$P^TAP = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

with a and b non-zero of opposite sides. Observing that

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}$$

and

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}\begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} = \begin{pmatrix} \lambda^2 c & 0 \\ 0 & \mu^2 d \end{pmatrix}$$

we see that there is a non-singular matrix Q with

$$Q^T\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}Q = A$$

and Q^TCQ diagonal whenever C is diagonal.

Setting $M = PQ$ we see that M is invertible $M^TAM = A$ and M^TBM is diagonal.

Let

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

Since $M^TAM = A$ we have

$$\begin{aligned}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &= A = M^TAM \\ &= \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ -c & -d \end{pmatrix} \\ &= \begin{pmatrix} a^2 - c^2 & ab - cd \\ ba - cd & b^2 - d^2 \end{pmatrix}\end{aligned}$$

Thus $ab = cd$ and $d > 0$, $a > 0$

On the other hand

$$M^TBM = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} c & d \\ a & b \end{pmatrix} = \begin{pmatrix} 2ac & ad + bc \\ bc + ad & 2bd \end{pmatrix}$$

so if M^TBM is diagonal $bc = -ad$.

We now get

$$c^2d = abc = -a^2d$$

so, since $d \neq 0$, we have $c = a = 0$ which contradicts the condition $a^2 - c^2 = 1$.

(ii) Since $\det A$, $\det B \neq 0$, A and B have rank 2. A has signature 0 by inspection. B has signature 0 since $xy = \frac{1}{4}\big((x+y)^2 - (x-y)^2\big)$.

The transformation $\mathbf{x} \mapsto M\mathbf{x}$ leaves lines $x = y$ and $x = -y$ unchanged (or interchanges them) since these are the asymptotes to $x^2 - y^2 = K$ for all K. It therefore gives dilation along the directions $x = y$ $x = -y$ (and possibly interchanges these two axes) and will therefore not transform $xy = k$ in the desired manner.

(iii) C is strictly negative definite and the same argument which gave simultaneous diagonalisation when one form is strictly positive definite will work if one form is strictly negative definite.

Exercise 16.4.2

We have

$$\begin{aligned}
q\left(\sum_{i=1}^{n} x_i \mathbf{e}_i\right) &= \alpha\left(\sum_{i=1}^{n} x_i \mathbf{e}_i, \sum_{j=1}^{n} x_j \mathbf{e}_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i a_{ij} x_j \\
&= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i a_{ij} x_j + x_i a_{ji} x_j) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} 0 = 0.
\end{aligned}$$

Exercise 16.4.3

Observe that
$$(iA)^* = -iA^* = -iA^T = iA.$$
Thus iA is Hermitian and so has real eigenvalues. If $-\lambda$ is an eigenvalue of iA then $i\lambda = (-i)(-\lambda)$ is an eigenvalue of $A = (-i)(iA)$. Thus the eigenvalues of A are purely imaginary.

Observe that
$$(M^T BM)^T = M^T B^T M^{TT} = -M^T BM$$
so the eigenvalues of $M^T BM$ are purely imaginary. But $M^T BM$ is a real diagonal matrix whose diagonal entries are its eigenvalues which are thus real and so must be zero. Thus $M^T BM = 0$ and so $B = 0$.

Exercise 16.4.6

(i) Let $\mathbf{u}_j$ $[1 \leq j \leq n]$ form a basis for an n-dimensional complex vector space U. Let α be the antisymmetric form associated with A for this basis.

Let $\mathbf{e}_j$ $[1 \leq j \leq n]$ be the basis found in Theorem 16.4.5. If M is the appropriate basis change matrix $M^T AM$ has the required form.

(ii) Observe that $2m = \operatorname{rank} M^T AM = \operatorname{rank} A$.

Exercise 16.4.7

Observe that, by Theorem 16.4.5, if α is non-singular, $2m = n$, so n is even.

Exercise 16.4.8

(i) True. Since we can find a non-singular M such that M^TAM has m copies of
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
along the diagonal and all other entries 0
$$\operatorname{rank} A = \operatorname{rank} M^TAM = 2m.$$

(ii) False. Take
$$A = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}.$$

(iii) True. The non-real roots of a real polynomial occur in conjugate pairs. P_A is a real polynomial with roots the eigenvalues (over $\mathbb{C}$) and these are purely imaginary (see Exercise 16.4.3). Thus
$$P_A(t) = t^{n-2m}\prod_{j=1}^{m}(t+id_r)(t-id_r) = \prod_{j=1}^{m}(t^2+d_r^2)$$
with d_r real.

(iv) False. Take
$$A = \begin{pmatrix} 0 & 4 \\ -1 & 0 \end{pmatrix}.$$

(v) True. Let A be the $n \times n$ matrix with the m matrices
$$\begin{pmatrix} 0 & d_r \\ -d_r & 0 \end{pmatrix}$$
along the diagonal and all other entries 0.

Exercise 16.4.9

We observe that

$$A \text{ is skew-Hermitian} \Leftrightarrow A = iB \text{ where } B \text{ is Hermitian.}$$

(i) If A is skew-Hermitian, then, since iA is Hermitian, we can find unitary P such that $iP^*AP = P^*(iA)P$ is a real diagonal matrix D and so $P^*AP = -iD$ is diagonal with diagonal entries purely imaginary.

(ii) If A is skew-Hermitian, then, since iA is Hermitian, Exercise 16.2.13 tells us that we can an invertible matrix P such that P^*iAP such that P^*iAP is diagonal with diagonal entries taking the values 1, -1 or 0 and so P^*AP is diagonal with diagonal entries taking the values i, $-i$ or 0.

(iii) Suppose that A is a skew-Hermitian matrix and P_1, P_2 are invertible matrices such that $P_1^*AP_1$ and $P_2^*AP_2$ are diagonal with diagonal entries taking the values i, $-i$ or 0. Then the number of entries of each type is the same for both diagonal matrices.

This follows at once from Exercise 16.2.13 (iii) by considering iA.