Mark V. Lawson

Solutions to Algebra and Geometry

CHAPTER 2

Proofs

Exercises 2.3

1. (a) C was a knight. The argument runs as follows. If you ask a knight what he is he will say he is a knight. If you ask a knave what he is he is obliged to lie and so also say that he is a knight. Thus no one on this island can say they are a knave. This means that B is a knave. Hence C was correct in saying that B lies and so C was a knight.

   (b) A is a knave, B a knight and C a knave. The argument runs as follows. If all three were knaves then C would be telling the truth which contradicts the fact that he is a knave. Thus at least one of the three is a knight and at most two are knights. It follows that C is a knave. Suppose that there were exactly two knights. Then both A and B would be knights but they contradict each other. It follows that exactly one of them is a knight. Hence A is knave and B is a knight.

2. Sam drinks water and Mary owns the aardvark. The following table shows all the information you should have deduced.

	1	2	3	4	5
House	Yellow	Blue	Red	White	Green
Pet	Fox	Horse	Snails	Dog	Aardvark
Name	Sam	Tina	Sarah	Charles	Mary
Drink	Water	Tea	Milk	Orange juice	Coffee
Car	Bentley	Chevy	Oldsmobile	Lotus	Porsche

The starting point is the following table.

	1	2	3	4	5
House					
Pet					
Name					
Drink					
Car					

Using clues (h), (i) and (n), we can make the following entries in the table.

	1	2	3	4	5
House		Blue			
Pet					
Name	Sam				
Drink			Milk		
Car					

There are a number of different routes from here. I shall just give some examples of how you can reason. Clue (a) tells us that Sarah lives in the red house. Now she cannot live in the first house, because Sam lives there, and she cannot live in the second house because that is blue. We are therefore left with the following which summarizes all the possibilities so far.

	1	2	3	4	5
House	Yellow? White? Green?	Blue	Red?	Red?	Red?
Pet					
Name	Sam		Sarah?	Sarah?	Sarah?
Drink			Milk		
Car					

Clue (b) tells us that Charles owns the dog. It follows that Sam cannot own the dog. We are therefore left with the following possibilities.

	1	2	3	4	5
House	Yellow? White? Green?	Blue	Red?	Red?	Red?
Pet	Fox? Horse? Snails? Aardvark?				
Name	Sam		Sarah?	Sarah?	Sarah?
Drink			Milk		
Car					

Both Questions 1 and 2 demonstrate some of the logic needed in mathematics.

3. I shall prove that the sum of an odd number and an even number is odd. The other cases are proved similarly. Let m be even and n be odd. Then $m = 2r$ and $n = 2s+1$ for some integers r and s. Thus $m+n = 2r+2s+1 = 2(s+r)+1$. Thus $m+n$ is odd.

4. Draw a diagonal across the quadrilateral dividing the figure into two triangles. The sum of the interior angles of the figure is equal to the sum of the angles in the two triangles. This is 360°.

5. (a) Two applications of Pythagoras' theorem give $\sqrt{2^2+3^2+7^2} = \sqrt{62}$.

(b) Draw a diagonal of the square and then construct the square with side that diagonal. This has twice the area by Pythagoras' theorem.

(c) The area of such a triangle is $\frac{1}{2}xy$ but we are told that it equals $\frac{1}{4}z^2$. Hence $\frac{1}{2}xy = \frac{1}{4}z^2$. By Pythagoras' theorem $z^2 = x^2 + y^2$. Substituting this value for z^2 we get that $\frac{1}{2}xy = \frac{1}{4}(x^2 + y^2)$. Rearrange this to get $(x-y)^2 = 0$. Hence $x - y = 0$ and so $x = y$. It follows that the triangle is isosceles.

6. (a) Every natural number n can be written $n = 10q + r$ where $0 \leq r \leq 9$. It follows that $n^2 = 10(10q^2 + 2qr) + r^2$. Thus the last digit of n^2 is equal to the last digit of r^2. Direct calculation now shows that this can only be one of $0, 1, 4, 5, 6, 9$. The converse is not true because 14 is a natural number ending in 4 but is not a perfect square.

(b) There are two statements to prove (1) if n is even then its last digit is even and (2) if the last digit of n is even then n is even. We prove (1). We may write $n = 10q + r$. Thus $r = n - 10q$. But n is even and $10q$ is even so it follows that r is even. We prove (2). We may write $n = 10q + r$ where r is even. Thus $r = 2s$. It follows that $n = 10q + 2s$. It is now clear that n is even.

(c) Again there are two statements to prove (1) if n is divisible by 9 then the sum of the digits in n is divisible by 9 and (2) if the sum of the digits in n are divisible by 9 then n is divisible by 9. The proof is based on the following observation. Each power of 10 leaves remainder 1 when divided by 9. Thus $10 = 9 + 1$ and $100 = 99 + 1$ etc. It follows that n can be written as a multiple of 9 plus the sum of its digits. The proofs of (1) and (2) now follow easily.

7. The proof follows exactly the same script as that for showing $\sqrt{2}$ is irrational, except that you need to prove the following result: if 3 divides a^2 then 3 divides a. To prove this we need the following. We may write $a = 3q + r$ where $r = 0, 1, 2$. It follows that $a^2 = 3(3q^2 + 2r) + r$. Thus a^2 is divisible by 3 if and only if a is divisible by 3.

8. Show that the angles $A\hat{D}B = A\hat{C}B$ using Proposition III.21. Deduce that triangle AXD is similar to triangle ABC. Show that the angles $D\hat{C}A = D\hat{B}A$ using Proposition III.21. The angles $D\hat{A}C = B\hat{A}X$ from what we already know and how the angles are formed. Thus the triangles AXB and ACD are similar. The two equations follow immediately from the fact that we have two sets of similar triangles. The remainder of the proof is just routine algebra using these two equations where we eliminate e.

9. (a) Lines in the plane have equations of the form $y = tx + s$ for some t and s. Our lines passes through the point $(-1, 0)$ and so $t = s$. Thus it has the form $y = t(x + 1)$.

(b) To find the coordinates of the point P, we have to solve the equation $x^2 + t^2(x+1)^2 = 1$ for x. This yields $x = \frac{1-t^2}{1+t^2}$ since we are excluding $x =$

-1. Substitute this value into $1 = x^2 + y^2$ and we obtain the corresponding value $y = \frac{2t}{1+t^2}$.

(c) If t is rational then the corresponding point (x,y) is rational. If (r,s) is a rational point on the unit circle then there is a corresponding rational value of t that yields that point since $s = t(r+1)$.

(d) The rational points on the circle consist of the point $(-1,0)$ and all points of the form $\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$ where t is any rational number.

With a little more work, it is possible to use this result to derive a formula for all Pythagorean triples.

10. We use the trigonometric identity

$$\cos\frac{\theta}{2} = \sqrt{\frac{1+\cos\theta}{2}},$$

where $0 \leq \theta \leq \frac{\pi}{2}$, and the fact that $\cos 45^\circ = \frac{1}{\sqrt{2}}$. An isosceles triangle with side 1 and enclosed angle θ has base $\sqrt{2}\sqrt{1-\cos\theta}$. We may now calculate the lengths of the sides of a square (4-gon), octogon (8-gon), 16-gon, 32-gon etc as follows.

n	4	8	16	32
side	$\sqrt{2}$	$\sqrt{2-\sqrt{2}}$	$\sqrt{2-\sqrt{2+\sqrt{2}}}$	$\sqrt{2-\sqrt{2+\sqrt{2+\sqrt{2}}}}$

It is clear this pattern continues and it is now easy to derive our approximations to π. This result suggests, but does not prove, that π is a special kind of number. See Section 7.10.

11. This is called the *Collatz problem* or the $3x+1$ *problem*. Nobody has yet found a proof. It is therefore conceivable that there is a number where the process described in the question does not terminate. I included this question to show that unsolved problems are not limited to what you might regard as advanced mathematics.

CHAPTER 3

Foundations

Exercises 3.1

1. It is true that $A = B$ because the order in which the elements of a set are listed is immaterial. It is true that $B = C$ because repetitions of elements are ignored.

2. (a) $\{2,4,6,8,10\}$.
   (b) $\{1,3,5,7,9\}$.
   (c) $\{6,7,8,9,10\}$.
   (d) $\emptyset$.
   (e) $\{2,3,5,7\}$.
   (f) $\{1,2,3,4,7,8,9,10\}$.

3. (a) There are 2^2 subsets with $1+2+1$ elements.
   (b) There are 2^3 subsets with $1+3+3+1$ elements.
   (c) There are 2^4 subsets with $1+4+6+4+1$ elements.
   (d) If there are n elements in the set there appear to be 2^n subsets. The pattern of subsets appears to follow the entries of Pascal's triangle. We shall explain these patterns in Section 3.9.

4. $\{(A,a),(A,b),(B,a),(B,b),(C,a),(C,b)\}$.

5. mn. This result will be proved in Section 3.9

6. m^n. This result will be proved in Section 3.9.

7. (a) $S \cup (T \cap V) = \{4,5,7,8,10,20,23\}$.
   (b) $S \setminus (T \cap V) = \{4,7,8,23\}$.
   (c) $(S \cap T) \setminus V = \{7\}$.

8. We have that $A \cap B = \{d,e,f\}$ and $(A \cup B) \setminus (A \cap B) = \{a,b,c,g,h,k\}$. Thus $A \setminus ((A \cup B) \setminus (A \cap B)) = \{d,e,f\}$.

9. $\{(2,a),(2,b),(3,a),(3,b)\}$.

10. There are two statements to prove.

    (a) If $A = B$ then $A \subseteq B$ and $B \subseteq A$.

    (b) If $A \subseteq B$ and $B \subseteq A$ then $A = B$.

    Proof of (1). Assume that $A = B$. Let $a \in A$. Then since $A = B$, it follows that $a \in B$. We have proved that $A \subseteq B$. The proof of the reverse inclusion is similar.

    Proof of (2). We assume that $A \subseteq B$ and $B \subseteq A$. Suppose that $A \neq B$. Then there are two possibilities derived from the definition of equality of sets. Either there is an $a \in A$ such that $a \notin B$, or there is $b \in B$ such that $b \notin A$. But both of these two possibilities are ruled out by our assumptions. It follows that $A = B$.

11. The sets are $\{\emptyset\}$, $\{\emptyset,\{\emptyset\}\}$, $\{\emptyset,\{\emptyset\},\{\emptyset,\{\emptyset\}\}\}$. The cardinalities are 1, 2 and 3, respectively. This is how the natural numbers are defined in set theory.

Exercises 3.2

1. (a)

p	q	r	$p \wedge (q \wedge r)$	$(p \wedge q) \wedge r$
T	T	T	T	T
T	T	F	F	F
T	F	T	F	F
T	F	F	F	F
F	T	T	F	F
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

(b)

p	q	$p \wedge q$	$q \wedge p$
T	T	T	T
T	F	F	F
F	T	F	F
F	F	F	F

(c)

p	q	r	$p \vee (q \vee r)$	$(p \vee q) \vee r$
T	T	T	T	T
T	T	F	T	T
T	F	T	T	T
T	F	F	T	T
F	T	T	T	T
F	T	F	T	T
F	F	T	T	T
F	F	F	F	F

(d)

p	q	$p \vee q$	$q \vee p$
T	T	T	T
T	F	T	T
F	T	T	T
F	F	F	F

(e)

p	q	r	$p \vee (q \wedge r)$	$(p \vee q) \wedge p \vee r()$
T	T	T	T	T
T	T	F	T	T
T	F	T	T	T
T	F	F	T	T
F	T	T	T	T
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

(f)

p	q	r	$p \wedge \neg(q \vee r)$	$(p \wedge \neg q) \wedge (p \wedge \neg r)$
T	T	T	F	F
T	T	F	F	F
T	F	T	F	F
T	F	F	T	T
F	T	T	F	F
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

(g)

p	q	r	$p \wedge \neg(q \wedge r)$	$(p \wedge \neg q) \vee (p \wedge \neg r)$
T	T	T	F	F
T	T	F	T	T
T	F	T	T	T
T	F	F	T	T
F	T	T	F	F
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

(h)

p	$p \wedge p$
T	T
F	F

(i)

p	$p \vee p$
T	T
F	F

2. The arguments are identical to the one in Example 3.2.3. The proofs of properties (3) and (6) from Theorem 3.2.1 are immediate.

3. (a) This follows from the truth-table below.

p	q	r	$p \operatorname{xor} (q \operatorname{xor} r)$	$(p \operatorname{xor} q) \operatorname{xor} r$
T	T	T	T	T
T	T	F	F	F
T	F	T	F	F
T	F	F	T	T
F	T	T	F	F
F	T	F	T	T
F	F	T	T	T
F	F	F	F	F

(b) This follows from the truth-table below.

p	q	r	$p \wedge (q \operatorname{xor} r)$	$(p \wedge q) \operatorname{xor} (p \wedge r)$
T	T	T	F	F
T	T	F	T	T
T	F	T	T	T
T	F	F	F	F
F	T	T	F	F
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

4. (a)⇒(b). If $A \setminus B$ were non-empty, there would be an element $x \in A$ and $x \notin B$. But this contradicts the assumption that $A \subseteq B$.

   (b)⇒(c). Clearly, $B \subseteq A \cup B$. Let $x \in A \cup B$. If $x \in B$ there is nothing to prove. Suppose that $x \in A$ and $x \notin B$. This however contradicts our assumption. Thus $A \cup B \subseteq B$ and so $A \cup B = B$.

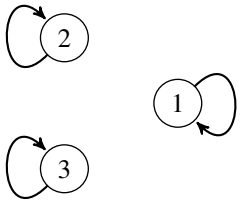
   (c)⇒(d). By assumption $B = A \cup B$. Then $A \cap B = A \cap (A \cup B) = A \cup (A \cap B) = A$ using the distributivity law and the fact that $A \cap B \subseteq A$.

   (d)⇒(a). We have that $A = A \cap B \subseteq B$ and so $A \subseteq B$, as required.

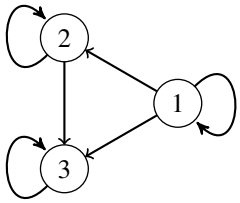
5. Let $x \in X \setminus B$. Then $x \in X$ and $x \notin B$. But $x \notin B$ implies that $x \notin A$. Thus $x \in X$ and $x \notin A$ and so by definition $x \in X \setminus A$.

Exercises 3.3

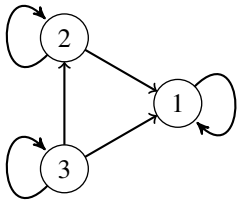
1. (a) Set of ordered pairs $\{(1,1),(2,2),(3,3)\}$. The digraph of $=$.



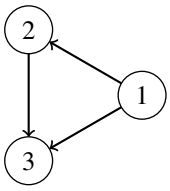
(b) Set of ordered pairs $\{(1,1),(2,2),(3,3),(1,2),(1,3),(2,3)\}$. The digraph of $\leq$.



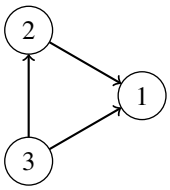
(c) Set of ordered pairs $\{(1,1),(2,2),(3,3),(2,1),(3,1),(3,2)\}$. The digraph of $\geq$.



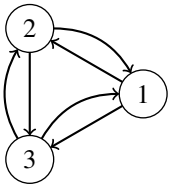
(d) Set of ordered pairs $\{(1,2),(1,3),(2,3)\}$. The digraph of $<$.



(e) Set of ordered pairs $\{(2,1),(3,1),(3,2)\}$. The digraph of $>$.



(f) Set of ordered pairs $\{(2,1),(3,1),(3,2),(1,2),(1,3),(2,3)\}$. The digraph of $\neq$.



2. (a) Symmetric, reflexive and transitive.

   (b) Antisymmetric, reflexive and transitive.

   (c) Antisymmetric, reflexive and transitive.

   (d) Transitive.

   (e) Transitive.

   (f) Symmetric.

3. (a) Show $\alpha \circ (\beta \circ \gamma) \subseteq (\alpha \circ \beta) \circ \gamma$ and $(\alpha \circ \beta) \circ \gamma \subseteq \alpha \circ (\beta \circ \gamma)$. I shall prove the first inclusion to demonstrate the idea which can be used throughout. Let $(x,y) \in \alpha \circ (\beta \circ \gamma)$. By definition there is an element u such that $(x,u) \in \alpha$ and $(u,y) \in \beta \circ \gamma$. By definition there is an element v such that $(u,v) \in \beta$ and $(v,y) \in \gamma$. Now reconstitute this information to get the desired result. From $(x,u) \in \alpha$ and $(u,v) \in \beta$ we obtain $(x,v) \in \alpha \circ \beta$. From $(x,v) \in \alpha \circ \beta$ and $(v,y) \in \gamma$ we obtain $(x,y) \in (\alpha \circ \beta) \circ \gamma$, as required.

   (b) Show that $\Delta \circ \alpha \subseteq \alpha$ and $\alpha \subseteq \Delta \circ \alpha$. The other equality is proved similarly.

   (c) Prove that $(\alpha \circ \beta)^{-1} \subseteq \beta^{-1} \circ \alpha^{-1}$ and $\beta^{-1} \circ \alpha^{-1} \subseteq (\alpha \circ \beta)^{-1}$.

Exercises 3.4

1. (a) Injective but not surjective since zero is omitted from the image.

   (b) Bijective.

   (c) Surjective but not injective since both -1 and 1 are mapped to the same element.

2. (a) $x \mapsto x+2$.

   (b) $x \mapsto x^2+1$.

(c) $x \mapsto x^2 + 2x + 1$.

(d) $x \mapsto 9x^2 + 12x + 3$.

(e) $x \mapsto 9x^2 + 12x + 3$.

3. (a) Disjoint cycles $(19)(28)(37)(46)(5)$ but 1-cycles are usually omitted. This is also a product of transpositions.

(b) Disjoint cycles $(124875)(36)$. Transpositions $(15)(17)(18)(14)(12)(36)$.

(c) Disjoint cycles (183426759). Transpositions $(19)(15)(17)(16)(12)(14)(13)(18)$.

4. (a) If $f(\emptyset)$ were non-empty it would have to contain an element $f(a)$ where $a \in \emptyset$, which is impossible.

(b) Let $f(a) \in f(A)$ where $a \in A$. Then $a \in B$ since $A \subseteq B$ and so $f(a) \in f(B)$, as required.

(c) Show that $f(A \cup B) \subseteq f(A) \cup f(B)$ and $f(A) \cup f(B) \subseteq f(A \cup B)$.

(d) Show that $f(A \cap B) \subseteq f(A) \cap f(B)$. Consider the function $f\colon \{1,2,3\} \to \{a,b\}$ given by $f(1) = f(3) = a$ and $f(2) = b$. Let $A = \{1,2\}$ and $B = \{2,3\}$. We have that $A \cap B = \{2\}$ and so $f(A \cap B) = \{b\}$. But $f(A) \cap f(B) = \{a,b\}$.

5. (a) If $f^{-1}(\emptyset)$ were non-empty it would have to contain an element a such that $f(a) \in \emptyset$, which is impossible.

(b) This follows from the fact that f is a function and so every element in the domain is mapped somewhere.

(c) Let $a' \in f^{-1}(A)$. Then $f(a') \in A$ and so $f(a') \in B$. It follows that $a' \in f^{-1}(B)$.

(d) Let $a' \in f^{-1}(A \cup B)$. Then $f(a') \in A$ or $f(a') \in B$. Thus $a' \in f^{-1}(A)$ or $a' \in f^{-1}(B)$. Hence one inclusion is proved. Let $a' \in f^{-1}(A \cup B)$. Then $f(a') \in A$ or $f(a') \in B$. It follows that $a' \in f^{-1}(A)$ or $a' \in f^{-1}(B)$ and the other inclusion is proved.

(e) Let $a' \in f^{-1}(A \cap B)$. Then $f(a') \in A \cap B$. Thus $f(a') \in A$ and $f(a') \in B$. Hence $a' \in f^{-1}(A)$ and $a' \in f^{-1}(B)$. Thus one inclusion is proved.

Now let $a' \in f^{-1}(A) \cap f^{-1}(B)$. Then $a' \in f^{-1}(A)$ and $a' \in f^{-1}(B)$. Thus $f(a') \in A$ and $f(a') \in B$. It follows that $a' \in f^{-1}(A \cap B)$ and the other inclusion is proved.

6. It's enough to prove that $f1_X = f$ since the other equality is similar. Both $f1_X$ and f have the same domain and codomain. And $(f1_X)(x) = f(1_X(x)) = f(x)$.

7. (a) Suppose that $(gf)(x) = (gf)(x')$. Then $g(f(x)) = g(f(x'))$. But g is injective and so $f(x) = f(x')$, and f is injective and so $x = x'$. It follows that gf is injective.

(b) Let $z \in Z$. Since g is surjective there exists $y \in Y$ such that $g(y) = z$. Since f is surjective there exists $x \in X$ such that $f(x) = y$. It follows that $(gf)(x) = z$ and so gf is surjective

8. We use the following property of a finite set X. If $Y \subseteq X$ and $|Y| = |X|$ then $X = Y$. Suppose that f is injective. Then the image of f contains the same number of elements as X. Thus it must actually be X. It follows that f is surjective. Suppose that f is surjective. For each $x \in X$ choose one x' such that $f(x') = x$. The set of all such elements is a subset of X with the same number of elements as X. Thus it must actually be X. It follows that there is a unique such x' for each x and so f is injective

9. (a) Let $x \in X$ be any element. Then $h(f(x)) = h(g(x))$. But h is injective and so $f(x) = g(x)$. But x was arbitrary and so $f = g$.

   (b) Suppose that h is not injective. Then there are $y, y' \in Y$ such that $h(y) = h(y')$. Let $\{1\}$ be any one element set. Defien $f(1) = y$ and $g(1) = y'$. Then $f \neq g$ but $hf = hg$ which is a contradiction. Thus h is injective,

   (c) Let $h\colon X \to Y$ and let $f, g\colon Y \to Z$. Then if $fh = gh$ then $f = g$.

   Suppose that h is surjective and that $fh = gh$. Let $y \in Y$ be arbitrary. Then $g(x) = y$ for some x. Thus $f(y) = g(y)$ for all $y \in Y$ and so $f = g$.

   Suppose that $fh = gh$ implies that $f = g$. We prove that h is surjective. Suppose not. Let $y \in Y$ be an element not in the image of h. Define $f\colon Y \to \{0,1\}$ by $f(y) = 0$ and f of everything else in Y maps to 1. Define $g\colon Y \to \{0,1\}$ by $g(y) = 1$ and g of everything else in Y maps to 1. Then $fh = gh$ but $f \neq g$.

Exercises 3.5

1. (a) This is not a partition because the element 7 is missing.

   (b) This is not a partition because the element 5 occurs in two different blocks.

   (c) This is a partition.

2. (a) Horizontal lines in the plane.

   (b) Vertical lines in the plane.

   (c) All lines parallel to $y = x$.

   (d) All lines parallel to $y = 2x$.

3. There are 15 partitions altogether. There are two extreme cases $\{\{1\},\{2\},\{3\},\{4\}\}$ and $\{1,2,3,4\}$.
   Then $\{\{1\},\{2,3,4\}\}, \{\{2\},\{1,3,4\}\}, \{\{3\},\{1,2,4\}\}, \{\{4\},\{1,2,3\}\}$.
   Then $\{\{1,2\},\{3,4\}\}, \{\{1,3\},\{2,4\}\}, \{\{1,4\},\{2,3\}\}$.
   Then $\{\{1\},\{2\},\{3,4\}\}, \{\{1\},\{3\},\{2,4\}\}, \{\{1\},\{4\},\{2,3\}\}, \{\{2\},\{3\},\{1,4\}\}$.
   $\{\{3\},\{4\},\{1,2\}\}$. $\{\{2\},\{4\},\{1,3\}\}$.

4. Let P be a partition on the set X consisting of blocks A_i. Define $x \sigma_P y$ on X if and only if $x, y \in A_i$ for some block A_i. Reflexive. For each $x \in X$ we have that $x \in A_i$ for some block. Thus $x \sigma_P x$. Symmetric. Suppose that $x \sigma_P y$.

Then $x,y \in A_i$ for some block. But then $y,x \in A_i$ for some block. Thus $y \sigma_P x$. Transitive. Let $x \sigma_P y$ and $y \sigma_P z$. Then $x,y \in A_i$ and $y,z \in A_j$. Then $A_i = A_j$ and so $x \sigma_P z$.

5. Let $f: X \to Y$ be a function. Define $\sim$ on X by $x \sim x'$ if and only if $f(x) = f(x')$. We prove that $\sim$ is an equivalence relation. For each $x \in X$ we have that $f(x) = f(x)$. Thus $\sim$ is reflexive. Let $x \sim y$. Then by definition $f(x) = f(y)$. Clearly $f(y) = f(x)$ and so $y \sim x$. It follows that $\sim$ is symmetric. Finally suppose that $x \sim y$ and $y \sim z$. Then $f(x) = f(y)$ and $f(y) = f(z)$. It follows that $f(x) = f(z)$ and so $x \sim y$. It follows that $\sim$ is transitive. We have therefore proved that $\sim$ is an equivalence relation.

Exercises 3.6

1. (a) We have that $x \preceq x$ and so $x \equiv x$. Symmetry is built into the definition. Transitivity is immediate.

   (b) Since $x \preceq x$ we have that $[x] \leq [x]$. Suppose that $[x] \leq [y]$ and $[y] \leq [x]$. Then $x' \preceq y'$ and $y'' \preceq x''$ for some $x', x'' \in [x]$ and $y', y'' \in [y]$. Now $x \equiv x' \preceq y' \equiv y$ thus $x \preceq y$ and $y \equiv y'' \preceq x'' \equiv x$ thus $y \preceq x$. It follows that $x \equiv y$ and so $[x] = [y]$. Suppose that $[x] \leq [y]$ and $[y] \leq [z]$. Then $x' \preceq y'$ and $y'' \preceq z'$. Thus $x' \preceq y' \equiv y'' \preceq z'$ and so $x' \preceq z'$ from which we get that $[x] \leq [z]$.

2. (a) $(a,a) \leq (a,a)$ since $a = a$ and $a \leq a$. Suppose that $(a,b) \leq (c,d)$ and $(c,d) \leq (a,b)$. Either $b < d$ or $b = d$ and $a \leq c$. Either $d < b$ or $d = b$ and $c \leq a$. The case $b < d$ is incompatible with what follows so $b = d$ amd $a \leq c$. But then $c \leq a$ and so $a = c$. Antisymmetry is therefore proved.

   Suppose that $(a,b) \leq (c,d)$ and $(c,d) \leq (e,f)$. Then either $b < d$ or $b = d$ and $a \leq c$. And either $d < f$ or $d = f$ and $c \leq e$. We just track through the possibilties.

   $b < d$ and $d < f$ so that $b < f$.

   $b < d$ and $d = f$ and $c \leq e$. Then $b < f$.

   $b = d$ and $a \leq c$. $d < f$ and so $b < f$.

   $b = d$ and $a \leq c$. $d = f$ and $c \leq e$. Then $b = f$ and $a \leq e$.

   (b) Let (a,b) and (c,d) be any elements. If $b < d$ then $(a,b) \leq (c,d)$. If $b = d$ then if $a \leq c$ we have that $(a,b) \leq (c,d)$ and if $c \leq a$ then $(c,d) \leq (a,b)$. If $d < b$ then $(c,d) \leq (a,b)$. We have therefore proved that this is a linear order.

3. Let $\leq$ be a partial order on X. Define $<$ by $a < b$ if and only if $a \leq b$ and $a \neq b$. It is easy to check that $<$ is a strict order. Let $\prec$ be a strict order on X. Define $\preceq$ by $a \preceq b$ if and only if $a \prec b$ or $a = b$. It is easy to check that $\preceq$ is a partial order on X.

Exercises 3.7

1. A concrete example is $(\mathbb{N}, <)$. Let X be a set on which such a ρ is defined. Let $x_1 \in X$ be any element. Then there exists $x_2 \in X$ such that $x_1 \rho x_2$. We cannot have $x_1 = x_2$ because then we would have $x_1 \rho x_1$ which is not allowed. Given x_2 there is x_3 such that $x_2 \rho x_3$. We cannot have $x_3 = x_2$. Could we have $x_3 = x_1$? The answer's no because $x_1 \rho x_3$. Continuing in this way we get an infinite subset $x_1, x_2, x_3 \ldots$ of X and so X is infinite.

2. $S \Rightarrow T$ is a theorem. But $T \Rightarrow S$ is not a theorem. Here is a counter-example. Let $A(x)$ be interpreted as 'x is an even natural number' and $B(x)$ is interpreted as 'x is an odd natural number'. Then $(\exists x)A(x) \wedge (\exists x)B(x)$ is true since $A(2)$ is true and $A(1)$ is true. But $(\exists x)(A(x) \wedge B(x))$ is false since there is no natural number that is both odd and even.

Exercises 3.8

1. Base step: check that the strict inequality holds when $n = 3$. We have that $3^2 = 9$ and $2 \times 3 + 1 = 7$ and $9 > 7$. Thus the base step holds. (IH): assume that $k^2 > 2k + 1$. We now have to prove that $(k+1)^2 > 2(k+1) + 1$ using (IH). We have that $(k+1)^2 = k^2 + 2k + 1$. By (IH), we have that $k^2 > 2k + 1$. Hence

$$(k+1)^2 > 2k + 1 + 2k + 1 = 2k + 2 + 2k > 2k + 2 + 1 = 2(k+1) + 1$$

since $2k > 1$. By the induction principle, we deduce that the inequality holds for all $n \geq 3$.

2. Base step: check that strict inequality holds when $n = 5$. We have that $2^5 = 32$ and $5^2 = 25$ and $32 > 25$. Thus the base step holds. (IH): assume that $2^k > k^2$. We now prove that $2^{k+1} > (k+1)^2$ using (IH). We have that

$$2^{k+1} = 22^k > 2k^2$$

using (IH). But $2k^2 = k^2 + k^2 > k^2 + 2k + 1$ by Question 1. Thus $2^{k+1} > (k+1)^2$. By the induction principle, we deduce that the inequality holds for all $n \geq 5$.

3. Base step: 6 is divisible by 3 and so this case holds. (IH) assume that $4^k + 2$ is divisible by 3. We prove that $4^{k+1} + 2$ is divisible by 3. We have that $4^{k+1} + 2 = 44^k + 2 = 3 \cdot 4^k + (4^k + 2)$ which is clearly divisible by 3.

4. We have that $1 + 2 + \ldots + k + (k+1) = \frac{1}{2}k(k+1) + (k+1)$. This is equal to $\frac{(k+1)(k+2)}{2}$.

5. We have that $2 + 4 + \ldots + 2k + 2(k+1) = k(k+1) + 2(k+1)$. This is equal to $(k+1)(k+2)$.

6. We have that $1^3 + 2^3 + \ldots + k^3 + (k+1)^3 = \left(\frac{k(k+1)}{2}\right)^2 + (k+1)^3$. This is equal to $\left(\frac{(k+1)(k+2)}{2}\right)^2$.

Exercises 3.9

1. (a) Let A be the set of starters, B the set of main courses, and C the set of drinks. Then an element of the set $A \times B \times C$ consists of a starter, followed by a main course, followed by a drink. Thus the set $A \times B \times C$ is the set of all possible such dinners. The cardinality of this set is $2 \cdot 3 \cdot 4$. Thus there are 24 such dinners.

   (b) The argument is the same as for (1) above. The number of possible dates is therefore $31 \cdot 12 \cdot 3000 = 1,116,000$.

   (c) This is a question about permutations. The answer is 10!.

   (d) This is the same as the previous answer.

   (e) This is a question about 3-permutations. The answer is $8 \cdot 7 \cdot 6 = 336$.

   (f) This is a question about combinations. The answer is $\binom{52}{13}$.

   (g) This is a question about combinations. The answer is $\binom{10}{4}$.

   (h) Order matters but repetition is not allowed and so this is a question about 3-permutations. The answer is is $9 \times 8 \times 7 = 504$ ways. I should add that order is implicit in stating the rôles: chairman, secretary and treasurer.

   (i) Since order is not important and repetitions are not allowed, this is a question about combinations and so the solution is

   $$\binom{49}{6} = 13,983,816.$$

   (j) We are just counting sequences and so the solution is $5^4 = 625$.

2. (a) n^n.

   (b) $n!$.

3. Think of a novel as one long string of symbols. This string has length

   $$250 \times 45 \times 60 = 675,000.$$

   But each symbol can be one of 100 possibilities and so the number of possible novels is

   $$100^{675,000}.$$

   It's more convenient to write this as a power of 10 and so we get

   $$10^{1,350,000}$$

   possible novels. For comparison purposes, the number of atoms in the universe is estimated to be 10^{80}.

4. (a) This is an important result. The partition counting principle tells us how to calculate the number of elements in the disjoint union of two sets. This result tells us how to calculate the number of elements in an *arbitrary* union. The result should make sense intuitively (why?). Observe that

$$A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A)$$

where the union on the RHS is disjoint. Thus

$$|A \cup B| = |A \setminus B| + |A \cap B| + |B \setminus A| .$$

Now $A = (A \setminus B) \cup (A \cap B)$, which is a disjoint union, and so

$$|A| = |A \setminus B| + |A \cap B| .$$

Similarly

$$|B| = |B \setminus A| + |A \cap B| .$$

Thus

$$|A \setminus B| = |A| - |A \cap B|$$

and

$$|B \setminus A| = |B| - |A \cap B| .$$

Substituting this in the our first expression for $|A \cup B|$ gives the result.

(b) To prove the second claim, we calculate as follows. We use the properties of the Boolean set operations. First,

$$|(A \cup B) \cup C| = |A \cup B| + |C| - |(A \cup B) \cap C| .$$

We can appeal to our result above to get

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |(A \cup B) \cap C| .$$

Finally, we deal the last term by writing $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$. We therefore get

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C| .$$

5. Denote by M the set of Montagues and by C the set of Capulates. A commitee is an ordered pair (A, B) where $A \subseteq M$ and $|A| = 3$ and $B \subseteq C$ where $|B| = 2$. The number of such subsets A is $\binom{7}{3}$ and the number of such subsets B is $\binom{5}{2}$. Thus the total number of commitees is $\binom{7}{3}\binom{5}{2}$.

6. We are interested in the number of r-element subsets of an $n+1$-element set. Pick an element of the $n+1$-element set and fix it. An r-element subset either contains this element or does not. We first count the number of r-element subsets that contain this element. We have to choose $r-1$ elements from n elements. Now we count the number of r-element subsets that do not contain this element. We have to choose r-elements from n. The result now follows by the partition counting principle.

Exercises 3.10

1. Let $f\colon A \to C$ and $g\colon B \to D$ be bijections. Define $k\colon A \cup B \to C \cup D$ by $k(x) = f(x)$ if $x \in A$ and $k(x) = g(x)$ if $x \in B$. Then k is a bijection.

2. Let $f\colon A \to C$ and $g\colon B \to D$ be bijections. Define $k\colon A \times C \to B \times D$ by $k(a,c) = (f(a), g(c))$. Then k is a bijection.

3. Let $f\colon X \to X'$ and $g\colon Y \to Y'$ be bijections. Define $k\colon X \sqcup Y \to X' \sqcup Y'$ by $k(x,0) = (f(x),0)$ and $k(y,1) = (g(y),1)$. Then k is a bijection.

4. We use the terms associative and commutative from Section 4.1. Observe that $A \sqcup B \cong B \sqcup A$ and $A \times B \cong B \times A$. Thus addition and multiplication of cardinals is commutative. This means that we need only check the calculations on and above the main diagonals in both cases. We also have that $(A \sqcup B) \sqcup C \cong A \sqcup (B \sqcup C)$ and $(A \times B) \times C \cong A \times (B \times C)$. Thus the operations are associative. Observe that $\emptyset \times A = \emptyset$. To prove $\mathbb{N} \times \mathbb{R} \cong \mathbb{R}$ it is enough to prove that $\mathbb{N} \times (0,1) \cong \mathbb{R}$. For finite cardinals it is true that $a^2 = a$ implies that $a = 0$ or $a = 1$. But it is not true for infinite cardinals since $\aleph_0^2 = \aleph_0$ and $c^2 = c$.

CHAPTER 4

Algebra redux

Exercises 4.1

1. Let $(a,b),(c,d),(e,f)$ be any three elements. Then $[(a,b)(c,d)](e,f) = (a,d)(e,f) = (a,f)$, and $(a,b)[(c,d)(e,f)] = (a,b)(c,f) = (a,f)$. Thus the operation is associative.

2. Let x,y,x be any three elements. Then $(xy)z = xz = x$ and $x(yz) = xy = x$. Thus the operation is associative. Clearly, every element is idempotent. Suppose that e is an identity and x is any element. Then $xe = x$ and $xe = e$. Thus $x = e$. It follows that the only case there is an identity is when the set X contains exactly one element.

3. Reflexivity. Every element is idempotent and so $e = e^2 = ee$. This means that $e \leq e$. Antisymmetry. Suppose that $e \leq f$ and $f \leq e$. By definition $e = ef$ and $f = fe$. But by commutativity we have that $e = f$. Transitivity. Let $e \leq f$ and $f \leq g$. Then by definition $e = ef$ and $f = fg$. Thus $e = e(fg) = (ef)g$ by associativity. But $ef = e$ thus $e = eg$. By definition $e \leq g$.

4. Let x and y be any two element. Then $(xy)^2 = 1$. Thus $xyxy = 1$ and so $xy = y^{-1}x^{-1}$. But $x^2 = 1$ implies that $x = x^{-1}$. It follows that $xy = yx$.

5. (a)⇒(b). Since $a \cdot a^2 = a^2 \cdot a$ we have that $a = a^2$. Now using this result $a \cdot aba = aba \cdot a$. It follows that $a = aba$.

   (b)⇒(c). We have that $a^3 = a$. Thus $a^4 = a^2$. It follows that $a = a \cdot a^2 \cdot a = a^4 = a^2$. Now $abc = ab(cac) = a(bc)ac = aac = ac$.

   (c)⇒(a). Suppose that $ab = ba$. Then $aba = baa$. Thus $a^2 = ba$ and so $a = ba$. But $ba = bab = b^2 = b$. Thus $a = b$.

6. (a) There are at most 6 elements: a,b,ab,ba,aba,bab. As a starting point, observe that any product of length four of as and bs must equal a product of three or less as or bs. Thus, for example, $abaa = aba$ and $abab = ab$. This gives us an upper bound for the number of elements. Now systematically go through all possibilities to show that there are at most 6 elements.

(b) There are at most 159 elements. This is similar to part (a) but requires more work. As a starting point, observe that any product of length six of as, bs and cs must equal a product of five or less as, bs and cs.

Exercises 4.2

1. (a) Start on the lefthand side. Expand brackets using the distributivity laws to get $aa+ba+ab+bb$. Commutativity tells us that $ba=ab$. Now simplify using standard abbreviations to get $a^2+2ab+b^2$.

   (b) Start on the lefthand side. Write $(a+b)^3=(a+b)(a+b)^2$. Use the distributivity laws applied to part (a) above to get $aa^2+a2ab+ab^2+ba^2+b2ab+bb^2$. Using commutativity and abbreviations this is just $a^3+2a^2b+ab^2+ba^2+2ab^2+b^3$. Now use commutativity and abbreviations to get $a^3+3a^2b+3ab^2+b^3$.

   (c) Start on the righthand side. Expand brackets using the distributivity laws to get $aa+ba-ab-bb$. Using commutativity the middle terms cancel, and using abbreviations we get a^2-b^2.

   (d) Expand the lefthand side using the distributivity laws to get $a^2c^2+a^2d^2+b^2c^2+b^2d^2$. Now expand the righthand side using the distributivity laws and commutativity to get $a^2c^2-2abcd+b^2d^2+a^2d^2+2abcd+b^2c^2$. By commutativity two terms cancel. We have shown that the lefthand side is equal to the righthand side.

2. (a) 8.

   (b) $\sqrt[3]{2}$.

   (c) $\frac{1}{16}$.

   (d) $\frac{1}{(\sqrt{2})^3}=\frac{1}{\sqrt{8}}$.

3. (a) This is the sum of the second column and so is $2+6+10=18$.

   (b) This is the sum of the third row and so is $9+10+11+12=42$.

   (c) This is the sum of the squares of all entries evaluated a row at a time and so is $(1^2+2^2+3^2+4^2)+(5^2+6^2+7^2+8^2)+(9^2+10^2+11^2+12^2)$ which is $30+174+446=650$.

4. Observe that

$$b(a_1+a_2+a_3+\ldots+a_n+a_{n+1})=b((a_1+a_2+a_3+\ldots+a_n)+a_{n+1})$$

   which is equal to

$$b(a_1+a_2+a_3+\ldots+a_n)+ba_{n+1}$$

   by the distributivity law. By (IH), this is equal to $ba_1+ba_2+ba_3+\ldots+ba_n+ba_{n+1}$.

5. If $a=0$ then the result is not true since $0\times 1=0\times 2$ and $1\neq 2$. Assume that $a\neq 0$. Then the result is true because from $ab=ac$ we get that $a^{-1}(ab)=a^{-1}(ac)$. By associativity, $a^{-1}(ab)=(a^{-1}a)b$ and $a^{-1}(ac)=(a^{-1}a)c$. But $a^{-1}a=1$. Thus $1b=1c$ and so $b=c$.

6. (a) Observe that

$$(a+b\sqrt{2})+(c+d\sqrt{2})=(a+c)+(b+d)\sqrt{2}$$

and

$$(a+b\sqrt{2})(c+d\sqrt{2})=(ac+2bd)+(ad+bc)\sqrt{2}.$$

This prove closure under addition and multiplication.

(b) Observe that

$$\frac{1}{a+b\sqrt{2}}=\frac{a-b\sqrt{2}}{a^2-2b^2}$$

which is obtained from $(a+b\sqrt{2})(a-b\sqrt{2})=a^2-2b^2$. If at least one of a and b is non-zero then this is non-zero, because if it were equal to zero we would be able to show that $\sqrt{2}$ is rational. Observe also that $0,1\in F$ and that if $a+b\sqrt{2}\in F$ then $-a-b\sqrt{2}\in F$. The fact that the axioms for a field hold for F is now immediate from the fact that they hold for $\mathbb{R}$.

Exercises 4.3

1. (a) 16 two real roots.
   (b) 0 repeated root.
   (c) -16 no real roots.

2. (a) Complete the square $x^2+10x+16=(x+5)^2-25+16=(x+5)^2-9$. Thus $x=-2,-8$.
   (b) Complete the square $x^2+4x+2=(x+2)^2-4+2=(x+2)^2-2$. Thus $x=-2\pm\sqrt{2}$.
   (c) Complete the square $2x^2-x-7=2[x^2-\frac{1}{2}x-\frac{7}{2}]=2[(x-\frac{1}{4})^2-\frac{1}{16}-\frac{7}{2}]=2[(x-\frac{1}{4})^2-\frac{57}{16}]$. Thus $x=\frac{1\pm\sqrt{57}}{4}$.

3. We are given that $x+y=a$ and $xy=b$. Suppose that $b\neq 0$. Then $x,y\neq 0$. Put $y=\frac{b}{x}$. This leads to the quadratic $x^2-ax+b=0$. Solving this yields $x=\frac{1}{2}\left(a+\sqrt{a^2-4b}\right)$ and $y=\frac{1}{2}\left(a-\sqrt{a^2-4b}\right)$, where we note that it doesn't matter which value is assigned to x as long as the corresponding value is assigned to y. Suppose that $b=0$. Then without loss of generality, we may assume that $x=0$. Then $y=a$.

4. We have that $2x_1=-b+\sqrt{D}$ and $2x_2=-b-\sqrt{D}$. Thus $2x_1-2x_2=\sqrt{D}+\sqrt{D}$. It follows that $x_1-x_2=\sqrt{D}$ and so $\Delta=(x_1-x_2)^2$.

5. (a) Suppose that $b \neq 0$. Then $\sqrt{c} = \frac{a-b^2-c}{2b}$ which contradicts the assumption that $\sqrt{c}$ is irrational.

   (b) We use part (a). We have that $\sqrt{d} = (a-c) + \sqrt{b}$. Thus $a = c$ and so $\sqrt{b} = \sqrt{d}$.

   (c) Put $\left(\sqrt{x} + \sqrt{y}\right)^2 = a + \sqrt{b}$ and solve for x and y. We get first $x + y = a$ and $2\sqrt{xy} = \sqrt{b}$ by part (b). Thus $x = \frac{1}{2}\left(a + \sqrt{a^2 - b}\right)$ and $y = \frac{1}{2}\left(a - \sqrt{a^2 - b}\right)$ as one solution with the other being -1 times these. Check that

$$\left(\tfrac{1}{\sqrt{2}}\sqrt{a + \sqrt{a^2 - b}} + \tfrac{1}{\sqrt{2}}\sqrt{a - \sqrt{a^2 - b}}\right)^2 = a + \sqrt{b}.$$

Exercises 4.4

1. $$\sum_{i=0}^{8} \binom{8}{i} x^i.$$

2. $$\sum_{i=0}^{8} \binom{8}{i} (-1)^i x^{8-i}.$$

3. The coefficient is $\binom{10}{2}$.

4. The coefficient is $\binom{6}{3} \cdot 3^3 \cdot 4^3$.

5. The binomial expansion is

$$\sum_{i=0}^{9} \binom{9}{i} \left(3x^2\right)^{9-i} \left(-\frac{1}{2x}\right)^i.$$

   Expanding each term carefully we get

$$\sum_{i=0}^{9} \binom{9}{i} \cdot (-1)^i \cdot 3^{9-i} \cdot 2^{-i} \cdot x^{18-2i} \cdot x^{-i}.$$

   This is equal to

$$\sum_{i=0}^{9} \binom{9}{i} \cdot (-1)^i \cdot 3^{9-i} \cdot 2^{-i} \cdot x^{18-3i}.$$

   We need to calculate the coefficient of x^3. This means we need to calculate the coefficient where $i = 5$. This is $-\binom{9}{5} \cdot 3^4 \cdot 2^{-5}$.

   We need to calculate the value of the constant term. This means we need to calculate the term where $i = 6$. This is $\binom{9}{6} \cdot 3^3 \cdot 2^{-6}$.

6. (a) Put $x = y = 1$ in the binomial theorem.
   (b) Put $x = 1$ and $y = -1$ in the binomial theorem.
   (c) Put $x = 1$ and $y = \frac{1}{2}$ in the binomial theorem.

7. We can prove that
$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}$$
by calculating the LHS and the RHS separately and then showing that they are equal. But a more conceptual proof was developed in Question 6 of Exercises 3.9.

8. We prove the binomial theorem by induction. We have that
$$(x+y)^{n+1} = (x+y)(x+y)^n.$$
By (IH), this is equal to
$$(x+y)\left(\sum_{i=0}^{n}\binom{n}{i}x^i y^{n-i}\right).$$
Using the distributivity law we get
$$\sum_{j=0}^{n}\binom{n}{j}x^{j+1}y^{n-j} + \sum_{i=0}^{n}\binom{n}{i}x^i y^{n-i+1}.$$
The first summand above can be written
$$\sum_{i=1}^{n+1}\binom{n}{i-1}x^i y^{n-i+1}.$$
We therefore have
$$y^{n+1} + \sum_{i=1}^{n}\left(\binom{n}{i-1} + \binom{n}{i}\right)x^i y^{(n+1)-i}.$$
The result now follows by the previous question.

9. (a) Expand $(x+y)^{2n}$ using the binomial theorem to get
$$\sum_{k=0}^{2n}\binom{2n}{k}x^k y^{2n-k}.$$
The term in $x^n y^n$ is just
$$\binom{2n}{n}x^n y^n.$$
But also
$$(x+y)^{2n} = (x+y)^n(x+y)^n.$$

Thus $(x+y)^{2n}$ is also equal to the square of the binomial expansion of $(x+y)^n$. This is equal to

$$\left(\sum_{i=0}^{n}\binom{n}{i}x^i y^{n-i}\right)\left(\sum_{j=0}^{n}\binom{n}{j}x^j y^{n-j}\right)$$

which is just

$$\sum_{i=0}^{n}\sum_{j=0}^{n}\binom{n}{i}\binom{n}{j}x^{i+j}y^{2n-(i+j)}.$$

The term in $x^n y^n$ is just the sum of those terms in which $i+j=n$. This is equal to $x^n y^n$ times

$$\binom{n}{0}\binom{n}{n}+\binom{n}{1}\binom{n}{n-1}+\ldots+\binom{n}{n}\binom{n}{0}.$$

Equating the coefficients of the same terms in $x^n y^n$ in both expressions yields the result.

(b) The total number of such commitees is $\binom{2n}{n}$. On the other hand a committee with i Montagues and $n-i$ Capulets can be chosen in $\binom{n}{i}\binom{n}{n-i}=\binom{n}{i}^2$ ways.

Exercises 4.5

1. (2) $a=a+0=a+(a\bar{a})=(a+a)\cdot(a+\bar{a})=(a+a)1=a+a$.

   (4) $1+a=(a+\bar{a})+a=\bar{a}+a+a=\bar{a}+a=1$.

   (6) $a+\bar{a}b=(a+ab)+\bar{a}b=a+(ab+\bar{a}b)=a+(a+\bar{a})b=a+1b=a+b$.

2. (a) We are given that $a+b=1$ and $a\cdot b=0$. We prove that $b=\bar{a}$. We have that $a+\bar{a}=1$. Thus $a+b=a+\bar{a}$. Multiply both sides by b to get $b(a+b)=b(a+\bar{a})$. Thus $ba+b\bar{a}=ba+b^2$. We are given that $ab=0$. Thus $b\bar{a}=b$. By assumption, $a+b=1$. Thus $\bar{a}(a+b)=\bar{a}$. But $a\bar{a}=0$. Thus $\bar{a}b=\bar{a}$. We have therefore proved that $b=\bar{a}$, as required.

   (b) We use the fact that complements are unique proved in part (a). We have that $\bar{a}+\bar{\bar{a}}=1$ and $\bar{a}\bar{\bar{a}}=0$. We deduce that $a=\bar{\bar{a}}$.

   (c) We use uniqueness of complements.
   Show that $(a+b)+\bar{a}\bar{b}=1$ and $\bar{a}\bar{b}(a+b)=0$.
   The proof of the other case is similar.

3. (a) $\bar{x}\bar{y}z+x\bar{y}\bar{z}+x\bar{y}z=(x+z)\bar{y}$.

   (b) $xy+x\bar{y}+\bar{x}y=x+y$.

   (c) $x+yz+\bar{x}y+\bar{y}xz=x+y$.

4. Observe that $1 * b = 1 \cdot \bar{b} = \bar{b}$. Thus complementation can be implemented. Also $a * \bar{b} = a \cdot b$. Thus multiplication can be implemented. Finally $\overline{\bar{a} * b} = \overline{\bar{a} \cdot \bar{b}} = a + b$. Thus addition can be implemented. The operation $*$ models an electronic switch, in fact a transistor because if $b = 0$ then $a * b = a$ whereas if $b = 1$ then $a * b = 0$.

Exercises 4.6

1. Let a and b both be top elements. Then $a \leq b$ and $b \leq a$ and so $a = b$. The argument for the uniqueness of bottom elements is essentially the same.

2. Suppose that $\frac{a}{b} < \frac{c}{d}$. We may assume without loss of generality, that $b, d > 0$ and, by assumption, $ad < bc$. It is now easy to check that

$$\frac{ad + bc}{2bd}$$

does the trick.

3. (a) We are given that $a \leq b$. Add $-b$ to both sides to get $-b + a \leq 0$. Now add $-a$ to both sides to get $-b \leq -a$. There are, of course, a number of intermediate steps which I have omitted.

   (b) There are two cases to consider. Suppose first that $0 < a$. Then by (O6), we have that $0 < a^2$ where I have omitted certain intermediate steps. Suppose now that $a < 0$. Then by part (a), we have that $0 < -a$. It follows that $0 < (-a)^2$. We now use the result that $-1 \cdot b = -b$ which I leave you to prove. We now get that $0 < a^2$ using the result we proved earlier that $(-1)^2 = 1$.

   (c) Since $1 = 1^2$ we have by part (b) that $0 < 1$. I will prove first that if $0 < a$ then $0 < a^{-1}$. In fact, if $a^{-1} \leq 0$ then $aa^{-1} \leq 0$ and so $1 \leq 0$ which is a contradiction. It follows that from $a < b$ we get that $a^{-1}a < a^{-1}b$ and so $1 < a^{-1}b$. Thus in a similar way $1b^{-1} < a^{-1}bb^{-1}$ and so $b^{-1} < a^{-1}$.

CHAPTER 5

Number theory

Exercises 5.1

1. (a) The quotient is 5 and the remainder is 0.
   (b) The quotient is 4 and the remainder is 4.
   (c) The quotient is 30 and the remainder is 4.

2. (a) 1103_5.
   (b) 109_{12}.
   (c) 99_{16}.

3. (a) 3499.
   (b) 19006.
   (c) 386556.

4. (a) The required fraction is $\frac{534-5}{1000-10} = \frac{529}{990}$ which is also in its lowest terms.
   (b) The required fraction is $\frac{2106-2}{10{,}000-10} = \frac{2104}{9990}$. The fraction in its lowest terms is $\frac{1052}{4995}$.
   (c) The required fraction is $\frac{76923}{999{,}999}$. The fraction in its lowest terms is $\frac{1}{13}$. An algorithm for writing a fraction in its lowest terms will be described in the next section.

5. (a) This follows from the fact that $a = 1 \times a$.
   (b) We are given that $b = am$ and $a = bn$ for some integers m and n. Thus $a = bn = amn$. Cancelling a we get that $mn = 1$. Thus either $m = n = 1$ or $m = n = -1$.
   (c) We are given that $b = am$ and $c = bn$. Thus $c = bn = amn$. But this means that $a \mid c$, as required.
   (d) We are given that $b = am$ and $c = an$. Thus $b + c = am + an = a(m+n)$. It follows that $a \mid (b+c)$ as required.

Exercises 5.2

1. (a) 5.

   (b) 9.

   (c) 7.

2. (a) We have that $267 = 2 \cdot 112 + 43$, $112 = 2 \cdot 43 + 26$, $43 = 1 \cdot 26 + 17$, $26 = 1 \cdot 17 + 9$, $17 = 1 \cdot 9 + 8$, $9 = 1 \cdot 8 + 1$. Thus $\gcd(112, 267) = 1$. We now run these calculations backwards to get

   $$1 = 31 \cdot 112 - 13 \cdot 267.$$

   (b) We have that $1870 = 7 \cdot 242 + 176$, $242 = 1 \cdot 176 + 66$, $176 = 2 \cdot 66 + 44$, $66 = 1 \cdot 44 + 22$. Thus $\gcd(242, 1870) = 22$. We now run these calculations backwards to get

   $$22 = 31 \cdot 242 - 4 \cdot 1870.$$

3. (a) $\gcd(10, 15) = 5$ and 5 does not divide 7. Thus this equation has no integer solutions.

   (b) $\gcd(5, 7) = 1$ thus this equation has infinitely many solutions. We have that $5(-4) + 7(3) = 1$. Thus a particular solution is $(-4, 3)$. The general solution is therefore $(-4 - 7t, 3 + 5t)$ where $t \in \mathbb{Z}$.

   (c) $\gcd(242, 1870) = 22$. Since 22 divides 66 this equation has infinitely many solutions. We have that $22 = 31 \cdot 242 - 4 \cdot 1870$. Thus $66 = 93 \cdot 242 - 12 \cdot 1870$. It follows that a particular solution is $(93, -12)$. Thus the general solution is $(93 - 85t, -12 + 11t)$ where $t \in \mathbb{Z}$.

4. (a) Draw up a table with three columns and a number of rows. The first row is labelled 0,1,2. the second is labelled 3,4,5 and so forth. Circle those numbers that can be written in the form $3x + 5y$ where $x, y \in \mathbb{N}$. We say that a row is complete if all numbers in that row are circled. Observe that $9 = 3 \cdot 3$ and $10 = 2 \cdot 5$ and $11 = 2 \cdot 3 + 5$. Since this row is complete all subsequent rows are complete. Now $8 = 3 + 5$ but 7 cannot be written in the given way. It follows that 7 is the largest value that cannot be made.

   (b) We claim that the Frobenius number is $ab - a - b$. Suppose that $n = ax + by$ where $x, y \in \mathbb{Z}$. By the remainder theorem, we can write $x = qb + r$ where $0 \leq r < b$. Hence $n = ar + b(aq + y)$. Suppose that $n = ar + by$ and $n = ar' + by'$ where $0 \leq r, r' < b$. Then $a(r - r') = b(y - y')$. The element b divides the RHS so must divide $r - r'$ but this can only occur if $r = r'$. It then follows that $y = y'$.

   Thus every integer n can be uniquely written in the form $n = ax + by$ where we can choose $0 \leq x < b$. We want to find the largest natural number n so that if $n = ax + by$ then $y < 0$. It follows that such an n can be written $n = a(b - 1) + b(-1) = ab - a - b$.

5. Let d be any natural number that divides a, b and c. Since d divides a and b it must divide $\gcd(a,b)$. But since d divides $\gcd(a,b)$ and c it must divide $\gcd(\gcd(a,b),c)$. Thus the lefthand side divides the righthand side. Now let $h = \gcd(\gcd(a,b),c)$. Then h divides $\gcd(a,b)$ and c. But if h divides $\gcd(a,b)$ then it must divide a and b. It follows that the righthand side divides the lefthand side. Hence the two sides must equal each other being natural numbers that are mutually divisible.

   The proof of the other claim follows by symmetry.

   Define a binary operation on the set $\mathbb{N} \setminus \{0\}$ by $a * b = \gcd(a,b)$. We have proved that this operation is associative.

   By generalized associativity, we have that

$$\gcd(910,780,286,195) = ((910 * 780) * 286) * 195.$$

   Thus

$$((910 * 780) * 286) * 195 = (130 * 286) * 195 = 26 * 195 = 13.$$

6. Let u and v be two non-zero natural numbers. First observe that if $u = v$ then $u \circ u = u$ and $\gcd(u,u) = u$ and there is nothing to prove. Without loss of generality, assume that $u > v$. But then

$$v \circ u = v \circ (v + (v - u)) = v \circ (u - v)$$

   where $u - v < u$. Observe that $\gcd(u,v) = \gcd(v,u-v)$. This process can be repeated using commutativity to swop sides if necessary. It will terminate when the two numbers are equal and then we use our first observation to deduce the result.

7. By construction $d = au + bv$ for some $u, v \in \mathbb{Z}$. Thus any common divisor of a and b divides d. We claim that d divides both a and b. Suppose not. Then $a = qd + r$ where $0 < r < d$. But $r = a - qd \in I$ and $r < d$. This is a contradiction. Thus $d \mid a$ and similarly $d \mid b$. It follows that d is the greatest common divisor of a and b.

Exercises 5.3

1. The primes less than 100 are

$$2,3,5,7,11,13,17,19,23,29,31,37,41,$$

$$43,47,53,59,61,67,71,73,79,83,89,97.$$

2. (a) $\sqrt{131} = 11 \cdot \ldots$. We therefore try the primes 2, 3, 5, 7, and 11. None of them works and so 131 is a prime.

(b) $\sqrt{689} = 26 \cdot \ldots$. We therefore try the primes 2, 3, 5, 7, 11, 13, 17, 19, and 23. We find that $13 \mid 689$ and $689 = 13 \cdot 53$. But 53 is a prime and so this is the prime factorization of our number.

(c) $\sqrt{5491} = 74 \cdot \ldots$. We therefore try the primes 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, and 73. We find that $17 \mid 5491$ and $5491 = 17 \cdot 323$. Now $\sqrt{323} = 17 \cdot \ldots$. We therefore try the primes 2, 3, 5, 7, 11, 13, and 17. We find that $323 = 17 \cdot 19$. But 19 is a prime. Thus $5491 = 17^2 \cdot 19$ is the prime factorization of our number.

3. (a) $\gcd(22,121) = 11$. Thus $\operatorname{lcm}(22,121) = \frac{22 \cdot 121}{11} = 242$.

(b) $\gcd(48,72) = 24$. Thus $\operatorname{lcm}(48,72) = \frac{48 \cdot 72}{24} = 144$.

(c) $\gcd(25,116) = 1$. Thus $\operatorname{lcm}(25,116) = \frac{25 \cdot 116}{1} = 2,900$.

4. The greatest common divisor is therefore $2^2 \cdot 5^5 \cdot 11^2$ and the least common multiple is therefore $2^4 \cdot 3 \cdot 5^6 \cdot 11^4$.

5. Since $\sqrt{ab} = \sqrt{a}\sqrt{b}$ it is enough to prove the result for powers of primes. Let $a = p^n$ where $n \geq 1$. Suppose that $n = 2m$ then $\sqrt{p^n} = p^m$. Suppose that $n = 2m+1$ then $\sqrt{p^n} = p^m\sqrt{p}$.

(a) $\sqrt{2}\sqrt{5}$.

(b) $\sqrt{2}\sqrt{3}\sqrt{7}$.

(c) $3\sqrt{2}\sqrt{3}$.

6. (a) This is proved by multiplying out the righthand side.

(b) This follows from part (a) by replacing y by $-y$.

(c) This follows from part (a).

(d) This follows from part (c).

(e) This follows by repeated application of part (b). We have that $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$ and $F_4 = 65,537$ are all primes. On the other hand $F_5 = 4,294,967,297 = 641 \cdot 6,700,417$, a result proved by Euler.

7. (a) Suppose that $n = u^2 - v^2$. Then $n = (u+v)(u-v)$. Put $a = u+v$ and $b = u-v$. Then $n = ab$ and $a \geq b > 0$. Suppose that $n = ab$ where $a \geq b > 0$. Put $u = \frac{1}{2}(a+b)$ and $v = \frac{1}{2}(a-b)$. Then $n = u^2 - v^2$. It is easy to check that these two procedures are mutually inverse and so they define a bijection.

(b) It factors as $(45,041 + 1,020)(45,041 - 1,020) = 46,061 \cdot 44,021$.

(c) If the factors a and b are close together then they are each roughly equal to the $\sqrt{200,819}$ which is approximately 448. Now $449^2 - 200819 = 782$ which is not a perfect square. But $450^2 - 200819 = 1681 = 41^2$ which is a perfect square. Thus $200819 = 450^2 - 41^2$ and we can factorize our number.

8. (a) Let p be an odd prime. By the remainder theorem $p = 4q + r$ where $0 \leq r < 4$. Neither $r = 0$ nor $r = 2$ is possible. Thus $r = 1$ or $r = 3$.

   (b) Let the first m consecutive primes of the form $4n+3$ be $p_1, \ldots, p_m$ where $p_1 = 3$. Put $N = 4p_2 \ldots p_m + 3$. Observe that we exclude $p_1 = 3$ so that N is not divisible by 3. If N is a prime then it is bigger than p_m and of the requisite form. If N is not prime then it has a prime factor. Not all prime factors of N can be of the form $4n+1$ because the product of numbers of the form $4a+1$ and $4b+1$ is also of the form $4c+1$. Thus N has a prime factor of the form $4n+3$ or 2. We can rule out the latter possibility since N is odd. Thus N has a prime factor of the form $4n+3$ which cannot be any of $p_1, \ldots, p_m$. Thus N has a prime factor of the requisite form greater than p_m.

9. By definition $a^2 + b^2 = c^2$. Suppose that a and b are both even. Then a^2 and b^2 are both even and so c^2 is even and so c is even. This contradicts the assumption that the triple is primitive. Suppose that a and b are both odd. Then a^2 and b^2 are both odd and so c^2 is even. Thus c is even. It follows that c^2 is divisible by 4. But the LHS leaves the remainder 2 when divided by 4 which is impossible. Thus exactly one of a and b is even.

   It is easy to show that $(p^2 - q^2, 2pq, p^2 + q^2)$ is a Pythgorean triple. We show that it is primitive. Suppose that n is a prime that divides all three of these numbers. Since exactly one of p and q is even, it follows that $p^2 - q^2$ is odd. Thus n cannot be 2. Now n divides pq and n divides $p^2 + q^2$. It follows that n divides $p + q$. But n divides pq and so n divides p or n divides q because p and q are coprime. Together this implies that n divides both p and q which is a contradiction. It follows that this really is a primitive Pythagorean triple.

   Let (a,b,c) be a primitive Pythogorean triple where b is even. Then $\left(\frac{a}{c}\right)^2 + \left(\frac{b}{c}\right)^2 = 1$. It follows that $(\frac{a}{c}, \frac{b}{c})$ is a rational point on the unit circle. Thus by Question 9 of Exercises 2.3, we have that

$$\frac{a}{c} = \frac{1-t^2}{1+t^2} \text{ and } \frac{b}{c} = \frac{2t}{1+t^2}$$

   for some rational t. We can write $t = \frac{q}{p}$ where $p > q > 0$ and p and q are coprime. Thus

$$\frac{a}{c} = \frac{p^2-q^2}{p^2+q^2} \text{ and } \frac{b}{c} = \frac{2pq}{p^2+q^2}.$$

   I claim that it is not possible for both p and q to be odd. Suppose they were. Then $p = 2m+1$ and $q = 2n+1$ for some m and n. From

$$\frac{a}{c} = \frac{p^2-q^2}{p^2+q^2}$$

   we deduce that a is even which is a contradiction.

We prove that $q^2 - p^2$ and $p^2 + q^2$ are coprime. Let r be a prime that divides them both. Then $r = 2$ or r divides q^2. Suppose the latter. Then r divides q and so r divides p^2 and so r divides p which is a contradiction. It follows that $r = 2$. The only way this can occur is if both p and q are odd which we have ruled out above.

We prove that $2pq$ and $p^2 + q^2$ are coprime. Let r be a prime that divides both. We can rule out $r = 2$ as above. If r divides p then r divides p^2 and so r divides q^2 and so r divides q. This is a contradiction. A similar argument shows that r cannot divide q.

We can now deduce that $a = (p^2 - q^2)s$ and $b = 2pqs$ and $c = (p^2 + q^2)s$ for some natural number s. But by assumption, the triple is primitive and so $s = 1$.

Exercises 5.4

1. Observe that $10 \equiv 1 \pmod 3$. Thus $10^r \equiv 1 \pmod 3$ for all $r \geq 1$. Let n have decimal representation $a_m \ldots a_0$. Then $n = a_m 10^m + \ldots + a_0$. Thus $n \equiv a_m + \ldots + a_0 \pmod 3$ The result now follows.

2. Let $n = x^2 + y^2$. We have that $x \equiv 0, 1, 2, 3 \pmod 4$. Thus $x^2 \equiv 0, 1 \pmod 4$. Hence $x^2 + y^2 \equiv 0, 1, 2 \pmod 4$.

3. The values are tabulated below.

n	$\phi(n)$
1	1
2	1
3	2
4	2
5	4
6	2
7	6
8	4
9	6
10	4
11	10
12	4

4. (a)

	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

(b)

	1	3	7	9
1	1	3	7	9
3	3	9	1	7
7	7	1	9	3
9	9	7	3	1

(c)

	1	5	7	11
1	1	5	7	11
5	5	1	11	7
7	7	11	1	5
11	11	7	5	1

5. (a) The following is a table of squares and so of square roots modulo 13.

n	1	2	3	4	5	6	7	8	9	10	11	12
n^2	1	4	9	3	12	10	10	12	3	9	4	1

(b) Working in $\mathbb{Z}_{13}$ we have that $(x+7a)^2 = x^2+ax+10a^2 = -b+10a^2 = 12b+10a^2 = 10(a^2+9b)$.

(a) $\Delta = 0$. Thus $(x+1)^2 = 0$ and so we get one root repeated $x = 12$.

(b) $\Delta = 1$. Thus $(x+8)^2 = 10$. It follows that $x+8 = 6$ or $x+8 = 7$. Hence the roots are $x = 11$ and $x = 12$.

(c) $\Delta = 5$ which has no square roots Thus the quadratic has no roots.

CHAPTER 6

Complex numbers

Exercises 6.1

1. (a) $6+4i$.
   (b) $5+14i$.
   (c) $28+96i$.
   (d) $\frac{11}{17}+\frac{10}{17}i$.
   (e) $\frac{3}{2}-\frac{5}{2}i$.
   (f) $\frac{-31}{200}+\frac{367}{200}i$.
2. (a) $\pm\frac{1}{\sqrt{2}}(1-i)$.
   (b) $\pm(\sqrt{2}+\sqrt{3}i)$.
   (c) $\pm(6-7i)$.
3. (a) $\frac{1}{2}(-1\pm\sqrt{3}i)$.
   (b) $\frac{1}{4}(3\pm\sqrt{7}i)$.
   (c) The roots are $1+2i$ and $1+i$.
4. Let $a+bi$ be a Gaussian integer. Squaring it, we get $(a^2-b^2)+2abi$. Observe now that $(a^2-b^2)^2+(2ab)^2=(a^2+b^2)^2$. Thus

$$\left(a^2-b^2, 2ab, a^2+b^2\right)$$

   is a Pythagorean triple.

Exercises 6.2

1. The set S^1 consists of all complex numbers with modulus 1. Let $u,v\in S^1$. Then $|uv|=|u||v|$ and so uv has modulus 1. Let $u\in S^1$. Then $u=\cos\theta+i\sin\theta$ for some θ. In this case, $u^{-1}=\cos\theta-i\sin\theta$. Thus $\left|u^{-1}\right|=1$.

2. By De Moivre's theorem

$$(\cos x + i\sin x)^5 = \cos 5x + i\sin 5x.$$

Expand the LHS using the binomial theorem to get

$$(\cos x)^5 + 5(\cos x)^4(i\sin x) + 10(\cos x)^3(i\sin x)^2 + 10(\cos x)^2(i\sin x)^3 + 5(\cos x)(i\sin x)^4 + (i\sin x)^5.$$

Simplifying and equating real and complex parts we get

$$\cos 5x = \cos^5 x - 10\cos^3 \sin^2 x + 5\cos x \sin^4 x$$

and

$$\sin 5x = 5\cos^4 x \sin x - 10\cos^2 x \sin^3 x + \sin^5 x.$$

3. This is an induction argument. We have that

$$(\cos\theta + i\sin\theta)^{n+1} = (\cos n\theta + i\sin n\theta)(\cos\theta + i\sin\theta)$$

by the induction hypothesis and this equals $\cos(n+1)\theta + i\sin(n+1)\theta$ from multiplication of complex numbers in polar form.

4. Let $z = a + bi$.

   (a) This follows from $z + \bar{z} = a + bi + a - bi = 2a$.

   (b) The real part of z is a whereas $|z| = \sqrt{a^2+b^2}$. Clearly $a \leq \sqrt{a^2+b^2}$.

   (c) Both equal $\sqrt{a^2+b^2}$.

   (d) We have that

$$|u+v|^2 = (u+v)\overline{(u+v)} = |u|^2 + |v|^2 + u\bar{v} + v\bar{u}.$$

   By part (a) above, we have that $u\bar{v} + v\bar{u} = 2\mathrm{Re}(u\bar{v})$. By part (b) above, we have that $\mathrm{Re}(u\bar{v}) \leq |u\bar{v}|$. By part (c) above, we have that $|u\bar{v}| = |uv|$. Thus

$$|u|^2 + |v|^2 + u\bar{v} + v\bar{u} \leq |u|^2 + |v|^2 + 2|uv| = (|u| + |v|)^2$$

   from which the result follows.

Exercises 6.3

1. (a) We have that $e^{ix} = \cos x + i\sin x$ and $e^{-ix} = \cos(-x) + i\sin(-x) = \cos x - i\sin x$. Thus $e^{ix} - e^{-ix} = 2i\sin x$. It follows that $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$

   (b) Using the calculations of (a) above, we have that $e^{ix} + e^{-ix} = 2\cos x$ and so we get the result.

We have that $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$. Taking fourth powers of both sides we get that

$$\cos^4 x = \frac{1}{16}\left(e^{4ix} + 4e^{2ix} + 6 + 4e^{-2ix} + e^{-4ix}\right).$$

This simplifies to

$$\frac{1}{8}(\cos 4x + 4\cos 2x + 3),$$

as required.

2. Put $z = i^i$. We extend, without worrying about the legitimacy of doing so, the results on powers of numbers described in Section 4.2. We therefore interpret i^i to mean $\exp(i\ln(i))$. Now $\exp(i(\frac{\pi}{2} + 2\pi k)) = i$ where k is any integer. It follows that $\ln(i) = i(\frac{\pi}{2} + 2\pi k)$. Thus $\ln(z) = -(\frac{\pi}{2} + 2\pi k)$. Hence

$$i^i = \exp(-\left(\tfrac{\pi}{2} + 2\pi k\right)).$$

All of these values are real.

Exercises 6.4

1. I shall prove that the multiplication is associative, commutative, that there is a multiplicative identity, and that every non-zero element has an inverse. For associativity, let $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$ and $\mathbf{c} = (c_1, c_2)$. Then $\mathbf{a}(\mathbf{bc}) = (\mathbf{ab})\mathbf{c}$ and both are equal to

$$(a_1b_1c_1 - a_1b_2c_2 - a_2b_1c_2 - a_2b_2c_1, a_1b_1c_2 + a_1b_2c_1 + a_2b_1c_1 - a_2b_2c_2).$$

Commutativity is easy to check as is the fact that $\mathbf{a1} = \mathbf{a}$. It is also easy to check that when $\mathbf{a} \neq \mathbf{0}$ we have that $\mathbf{a}\mathbf{a}^{-1} = \mathbf{1}$.

2. The two systems are similar in all respects except two. The multiplicative identity is $(1,1)$. But there are many non-zero element that do not have multiplicative inverses. In fact all the elements of the form $(a,0)$ and of the form $(0,a)$ where $a \neq 0$ are non-zero but don't have inverses. In addition, it is possible for non-zero elements to multiply together and get zero. For example $(0,a)(b,a) = (0,0)$ where $a, b \neq 0$.

CHAPTER 7

Polynomials

Exercises 7.2

1. (a) The quotient is $2x^2 - 3x$ and the remainder is 1.
   (b) The quotient is $x^2 + 2x - 3$ and the remainder is -7.
   (c) The quotient is $x^2 - 3x + 8$ and the remainder is $-27x + 7$.

Exercises 7.4

1. (a) We are given that 4 is a root and so we know that $x-4$ is a factor. Dividing out we get $3x^2 - 8x + 4$. This is a quadratic and so we can find its roots by means of completing the square. We get 2 and $\frac{2}{3}$. Thus the roots are $4, 2, \frac{2}{3}$.
   (b) We are given that -1 and -2 are roots and so $(x+1)(x+2)$ is a factor. Dividing out we get $x^2 - x + 1$. The roots of this quadratic are $\frac{1}{2}(1 \pm i\sqrt{3})$. Thus the roots are $-1, -2, \frac{1}{2}(1 \pm i\sqrt{3})$.

2. The required cubic is $(x-2)(x+3)(x-4) = x^3 - 3x^2 - 10x + 24$.

3. The required quartic is $(x-i)(x+i)(x-1-i)(x-1+i) = x^4 - 2x^3 + 3x^2 - 2x + 2$.

4. By assumption $x^3 + ax^2 + bx + c = (x - x_1)(x - x_2)(x - x_3)$. Multiplying out the RHS we get

$$x^3 - (x_1 + x_2 + x_3)x^2 + (x_1x_2 + x_1x_3 + x_2x_3)x - x_1x_2x_3.$$

   Now equate with the coefficients of the LHS to get

$$a = -(x_1 + x_2 + x_3),\ b = x_1x_2 + x_1x_3 + x_2x_3,\ c = -x_1x_2x_3.$$

5. The polynomial in question has real coefficients and so the complex roots come in complex conjugate pairs. It follows therefore that $3 - i\sqrt{2}$ is also a root. Thus $(x - 3 - i\sqrt{2})(x - 3 + i\sqrt{2}) = x^2 - 6x + 11$ is a factor. Dividing out we get $x^2 + 7x + 6$. This factorizes as $(x+1)(x+6)$ and so its roots are -1 and -6. Thus the roots are $-1, -6, 3 + i\sqrt{2}, 3 - i\sqrt{2}$.

6. The polynomial in question has real roots and so $1+i\sqrt{5}$ is another root. Thus $(x-1-i\sqrt{5})(x-1+i\sqrt{5})$ is a factor. Dividing out by x^2-2x+6 we get x^2-2. This factorizes as $(x-\sqrt{2})(x+\sqrt{2})$. Thus the roots are $1+i\sqrt{5}, 1-i\sqrt{5}, \sqrt{2}, -\sqrt{2}$.

7. (a) -1 is a root and so $x+1$ is a factor. We can write the polynomial as the product $(x+1)(x^2+1)$. The roots are therefore $-1, i, -i$.

   (b) -2 is a root and so $x+2$ is a factor. We can therefore write the polynomial as $(x+2)(x^2-3x+3)$. The roots are therefore $-2, \frac{1}{2}(3+i\sqrt{3}), \frac{1}{2}(3-i\sqrt{3})$.

   (c) 1 is a root and so we get a first factorization of our polynomial as $(x-1)(x^3+5x+6)$. -1 is a root of x^3+5x+6. We may therefore factorize $x^3+5x+6=(x+1)(x^2-x+6)$. The quadratic has the roots $\frac{1}{2}(1\pm i\sqrt{23})$. The roots are therefore $1, -1, \frac{1}{2}(1+i\sqrt{23}), \frac{1}{2}(1-i\sqrt{23})$.

8. (a) Show that 1 is a root and then divide by $x-1$ to get the required factorization $(x-1)(x^2+x+1)$. Observe that x^2+x+1 has complex roots and so cannot be factorized further in terms of real polynomials.

   (b) This is a difference of two squares and so a first factorization is $(x^2+1)(x^2-1)$ and thus the required factorization is $(x-1)(x+1)(x^2+1)$. Observe that x^2+1 has complex roots and so cannot be factorized further in terms of real polynomials.

   (c) Put $y=x^2$ and Solve $y^2+1=0$. The solutions are $\pm i$. Thus $x^2=i$ or $x^2=-i$. Taking square roots, yields $x=\frac{1}{\sqrt{2}}(1+i), \frac{-1}{\sqrt{2}}(1+i), \frac{1}{\sqrt{2}}(-1+i), \frac{1}{\sqrt{2}}(1-i)$. Now we collect together complex conjugate pairs, to get $(x-\frac{1}{\sqrt{2}}(1+i))(x-\frac{1}{\sqrt{2}}(1-i))=x^2-\sqrt{2}x+1$, and $x^2+\sqrt{2}x+1$. Thus $x^4+1=(x^2-\sqrt{2}x+1)(x^2+\sqrt{2}x+1)$. [A student made a nice observation that leads to a much quicker solution to this question. Observe that $x^4+1=(x^2+1)^2-2x^2$. How does this help?]

9. Here is the case where $r=3$.

   - $p_1(x)=\frac{(x-\lambda_2)(x-\lambda_3)}{(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)}$.
   - $p_2(x)=\frac{(x-\lambda_1)(x-\lambda_3)}{(\lambda_2-\lambda_1)(\lambda_2-\lambda_3)}$.
   - $p_3(x)=\frac{(x-\lambda_1)(x-\lambda_2)}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}$.

   The general case is now easy to write down.

Exercises 7.5

1. $1, i, -1, -i$.

2. Let $\omega=\frac{1}{2}\left(1+i\sqrt{3}\right)$. Then the roots are $1, \omega, \omega^2, \omega^3, \omega^4, \omega^5$.

3. Let $\omega = \frac{1}{\sqrt{2}}(1+i)$. Then the roots are $1, \omega, \omega^2, \omega^3, \omega^4, \omega^5, \omega^6, \omega^7$.

4. This question shows the sorts of insights that are needed to calculate explicit radical expressions for nth roots.

5. (a) The cube roots are
   - $2(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}) = 2i$.
   - $2(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}) = -\sqrt{3} - i$.
   - $2(\cos\frac{11\pi}{6} + i\sin\frac{11\pi}{6}) = \sqrt{3} - i$.

   (b) The fourth roots are
   - $\sqrt[4]{2}\left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right) = i\sqrt[4]{2}$.
   - $\sqrt[4]{2}\,(\cos\pi + i\sin\pi) = -\sqrt[4]{2}$.
   - $\sqrt[4]{2}\left(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}\right) = -i\sqrt[4]{2}$.
   - $\sqrt[4]{2}\,(\cos 2\pi + i\sin 2\pi) = \sqrt[4]{2}$.

   (c) Observe that $1+i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right)$. The sixth roots are
   - $\sqrt[12]{2}\left(\cos\frac{\pi}{24} + i\sin\frac{\pi}{24}\right)$.
   - $\sqrt[12]{2}\left(\cos\frac{9\pi}{24} + i\sin\frac{9\pi}{24}\right)$.
   - $\sqrt[12]{2}\left(\cos\frac{17\pi}{24} + i\sin\frac{17\pi}{24}\right)$.
   - $\sqrt[12]{2}\left(\cos\frac{25\pi}{24} + i\sin\frac{25\pi}{24}\right)$.
   - $\sqrt[12]{2}\left(\cos\frac{33\pi}{24} + i\sin\frac{33\pi}{24}\right)$.
   - $\sqrt[12]{2}\left(\cos\frac{41\pi}{24} + i\sin\frac{41\pi}{24}\right)$.

Exercises 7.6

1. (a) $x^2 + 2x + 3$.
   (b) 1.
   (c) $x + 2$.

Exercises 7.8

1. (a) $\frac{1}{x+1} + \frac{2}{x+2}$.
   (b) $\frac{3}{x+2} + \frac{x}{x^2+1}$.
   (c) $\frac{x+1}{x^2+x+1} + \frac{x}{(x^2+x+1)^2}$.
   (d) $\frac{\frac{\sqrt{2}}{4}x + \frac{1}{2}}{x^2+\sqrt{2}x+1} + \frac{\frac{1}{2} - \frac{\sqrt{2}}{4}x}{x^2-\sqrt{2}x+1}$.

Exercises 7.9

1. Calculate $(x-u-v)(x-u\omega-v\omega^2)(x-u\omega^2-v\omega)$.

2. Calculate $(x-u^3)(x-v^3)$.

3. We use the substitution $x = y - 1$ to get the reduced cubic $x^3 - 6x - 9$ where y has been relabelled x. We now solve $-3uv = -6$ and $-u^3 - v^3 = -9$. Thus u^3 and v^3 are the roots of $t^2 - 9t + 8 = 0$. Thus $t = 1$ or $t = 8$. Put $1 = u^3$. One cube root of unity is $u = 1$ but $uv = 2$ thus $v = 2$. The roots of the reduced cubic are therefore 3, $\omega + 2\omega^2$ and $\omega^2 + 2\omega$. Thus the roots of the original equation are 2, $\omega + 2\omega^2 + 1$ and $\omega^2 + 2\omega + 1$. That is 2 and $\frac{1}{2}(-5 \pm i\sqrt{3})$.

4. The substitution that works is $x = y - \frac{a_3}{4}$.

5. (a) This is simply $x^4 - (x+1)^2 = \big(x^2 - (x+1)\big)\big(x^2 + (x+1)\big)$.

   (b) This becomes $(x^2 + 2)^2 = 3$.

6. The even permutations are ι, (123) and (132). The odd permutations are (12), (23) and (13).

7. (a) Multiply out the RHS and show that it simplifies to the LHS.

   (b) Multiply out the RHS and show that it simplifies to the LHS.

   (c) The polynomial is symmetric and so must be equal to the polynomial obtained by interchanging x and y.

   (d) This follows from (c) above.

   (e) Write the polynomial in the form derived in (d) above. Now repeatedly apply the reduction rules derived in (a) and (b). This will result in a polynomial in the variables $\sigma_1 = x + y$ and $\sigma_2 = xy$.

   (f) We first rewrite the polynomial as

   $$(x^4 + y^4) + 4(x^3y + xy^3) + 6x^2y^2$$

   using (d). Observe that $x^2y^2 = (xy)^2 = \sigma_2^2$. Thus we obtain

   $$(x^4 + y^4) + 4(x^3y + xy^3) + 6\sigma_2^2.$$

   By (b), we have that $x^3y + xy^3 = xy(x^2 + y^2)$. But $x^2 + y^2 = (x+y)^2 - 2xy$. Thus $x^3y + xy^3 = \sigma_2\sigma_1^2 - 2\sigma_2^2$. Thus we obtain

   $$(x^4 + y^4) + 4(\sigma_2\sigma_1^2 - 2\sigma_2^2) + 6\sigma_2^2 = (x^4 + y^4) + 4\sigma_2\sigma_1^2 - 2\sigma_2^2.$$

   By (a), we have that $x^4 + y^4 = (x+y)(x^3 + y^3) - xy(x^2 + y^2)$. Likewise $x^2 + y^2 = (x+y)^2 - 2xy$ and $x^3 + y^3 = (x+y)(x^2 + y^2) - xy(x+y)$. It follows that

   $$x^4 + y^4 = \sigma_1^4 - 4\sigma_1^2\sigma_2 + 2\sigma_2^2.$$

   Thus our polynomial has the form σ_1^4 as expected.

Exercises 7.11

1. A transversal is given by polynomials of the form $a + bx$ where $a, b \in \mathbb{R}$. Multiplication is given by $(a + bx)(c + dx) = ac + (ad + bc)x$. This does not form a field since x^2 is clearly a reducible polynomial over $\mathbb{R}$.

2. There are 8 element and they form a field since x^3+x+1 is irreducible over $\mathbb{Z}_2$. The elements are

$$0,1,x,x^2,x+1,x^2+x,x^2+x+1,x^2+1.$$

The construction of the Cayley table can be considerably eased by observing that $x^3 \equiv x+1$, $x^4 \equiv x^2+x$, $x^5 \equiv x^2+x+1$, $x^6 \equiv x^2+1$ and $x^7 = 1$.

3. There are 9 elements and they form a field since x^2+1 is irreducible over $\mathbb{Z}_3$. The elements are

$$0,1,2,x,x+1,x+2,2x,2x+1,2x+2.$$

As above, the construction of the Cayley table can be considerably eased by observing that

- $(x+1)^2 \equiv 2x$.
- $(x+1)^3 \equiv 2x+1$.
- $(x+1)^4 \equiv 2$.
- $(x+1)^5 \equiv 2x+2$.
- $(x+1)^6 \equiv x$.
- $(x+1)^7 \equiv x+2$.
- $(x+1)^8 \equiv 1$.

CHAPTER 8

Matrices

Exercises 8.1

1. (a)
$$-3B = -3\begin{pmatrix} 1 & 4 \\ -1 & 1 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -3 & -12 \\ 3 & -3 \\ 0 & -9 \end{pmatrix}.$$

(b)
$$A+B = \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 4 \\ -1 & 1 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1+1 & 2+4 \\ 1-1 & 0+1 \\ -1+0 & 1+3 \end{pmatrix} = \begin{pmatrix} 2 & 6 \\ 0 & 1 \\ -1 & 4 \end{pmatrix}.$$

(c)
$$A-B = \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 4 \\ -1 & 1 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1-1 & 2-4 \\ 1+1 & 0-1 \\ -1-0 & 1-3 \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 2 & -1 \\ -1 & -2 \end{pmatrix}.$$

2. (a)
$$AB = \begin{pmatrix} 0+8+6 & 0+0+4 & 0-16+0 \\ -1+2+9 & 3+0+6 & -5-4+0 \\ 2+0+6 & -6+0+4 & 10+0+0 \end{pmatrix} = \begin{pmatrix} 14 & 4 & -16 \\ 10 & 9 & -9 \\ 8 & -2 & 10 \end{pmatrix}.$$

(b)
$$BA = \begin{pmatrix} 0+3+10 & 4-3+0 & 2-9+10 \\ 0+0-8 & 8+0+0 & 4+0-8 \\ 0-2+0 & 12+2+0 & 6+6+0 \end{pmatrix} = \begin{pmatrix} 13 & 1 & 3 \\ -8 & 8 & -4 \\ -2 & 14 & 12 \end{pmatrix}.$$

3. (a)
$$BA = \begin{pmatrix} 0 & 1 \\ -1 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0+0 & 0-1 \\ -3+0 & -1-1 \\ 9+0 & 3-1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -3 & -2 \\ 9 & 2 \end{pmatrix}.$$

(b)

$$AA = \begin{pmatrix} 3 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 9+0 & 3-1 \\ 0+0 & 0+1 \end{pmatrix} = \begin{pmatrix} 9 & 2 \\ 0 & 1 \end{pmatrix}.$$

(c)

$$CB = \begin{pmatrix} 1 & 0 & 3 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 0+0+9 & 1+0+3 \\ 0-1+3 & -1+1+1 \end{pmatrix} = \begin{pmatrix} 9 & 4 \\ 2 & 1 \end{pmatrix}.$$

(d)

$$AC = \begin{pmatrix} 3 & 1 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 3 \\ -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3-1 & 0+1 & 9+1 \\ 0+1 & 0-1 & 0-1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 10 \\ 1 & -1 & -1 \end{pmatrix}.$$

(e)

$$BC = \begin{pmatrix} 0 & 1 \\ -1 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 3 \\ -1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0-1 & 0+1 & 0+1 \\ -1-1 & 0+1 & -3+1 \\ 3-1 & 0+1 & 9+1 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 1 \\ -2 & 1 & -2 \\ 2 & 1 & 10 \end{pmatrix}.$$

(f)

$$C^T A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 3+0 & 1+1 \\ 0+0 & 0-1 \\ 9+0 & 3-1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 0 & -1 \\ 9 & 2 \end{pmatrix}.$$

4.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \\ 4 & 8 & 12 \end{pmatrix}.$$

5.

$$(AB)C = \left(\begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ -2 & 1 \end{pmatrix}\right)\begin{pmatrix} -1 & 2 & 3 \\ 4 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ -3 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} -1 & 2 & 3 \\ 4 & 0 & 1 \end{pmatrix}$$

which gives

$$(AB)C = \begin{pmatrix} 0 & 8 & 13 \\ 3 & -6 & -9 \\ 12 & 0 & 3 \end{pmatrix}.$$

The same matrix arises from the calculation $A(BC)$.

6.

$$\begin{pmatrix} 7i-3 & 8i \\ 4i & 9i \end{pmatrix}.$$

7.
$$\begin{pmatrix} ad & 0 & 0 \\ 0 & be & 0 \\ 0 & 0 & cf \end{pmatrix}.$$

8. (a)
$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

(b)
$$\begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix}.$$

(c)
$$\begin{pmatrix} b & a & c \\ e & d & f \\ h & g & i \end{pmatrix}.$$

9. (a)
$$A^T = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \end{pmatrix}.$$

(b)
$$B^T = \begin{pmatrix} 1 & 2 & 3 \\ -3 & 0 & 2 \\ 5 & -4 & 0 \end{pmatrix}.$$

(c)
$$C^T = \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}.$$

10. I give just the top lefthand side of the table.

	I	X	Y	Z
I	I	X	Y	Z
X	X	$-I$	Z	$-Y$
Y	Y	$-Z$	$-I$	X
Z	Z	Y	$-X$	$-I$

In Chapter 9, I shall explain the connection between these matrices and Hamilton's quaternions.

11. (a) When we calculate $\mathbf{c} = H\mathbf{r}^T$ we are recalculating the check digits c_1, c_2, c_3 and we get the values c_1', c_2', c_3'. In fact $\mathbf{c}_1 = c_3'$, $\mathbf{c}_2 = c_2'$ and $\mathbf{c}_3 = c_1'$ If they are all zero, then using the Venn diagram, you can see that no errors have occurred. If $\mathbf{c}$ matches the i column of H then the pattern of zeros and ones applied to the Venn diagram enables you to locate the bit that must be incorrect. The columns of H are arranged so that the ith column is the binary representation of the number i.

(b) To correct an error it is enough to say which bit is in error and this can be done using binary. Although it is the 4 bits of the message that interest us when we add check bits these too could be in error. To describe the error in 4 bits we need at least two bits in addition. But then we have 6 bits and to describe an error in 6 positions we need an extra bit. Thus we have 4 information bits and 3 error-correction bits. The error can occur in any one of the 7 bits and 3 bits can describe those 7 positions.

Exercises 8.2

1. $$\left(\begin{pmatrix} 2 & 0 \\ 7 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\right) + \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$$

is equal to

$$\left(\begin{pmatrix} 3 & 1 \\ 8 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\right) + \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$$

is equal to

$$\begin{pmatrix} 3 & 2 \\ 9 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$$

is equal to

$$\begin{pmatrix} 5 & 4 \\ 12 & 3 \end{pmatrix}$$

where we use the associative law of matrix addition throughout.

2. $$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \left(\begin{pmatrix} 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -4 \end{pmatrix} \right) \begin{pmatrix} 3 & 1 & 5 \end{pmatrix}$$

is equal to

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \left(\begin{pmatrix} -3 \end{pmatrix} \begin{pmatrix} 3 & 1 & 5 \end{pmatrix} \right)$$

is equal to

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} -9 & -3 & -15 \end{pmatrix}$$

is equal to

$$\begin{pmatrix} -9 & -3 & -15 \\ -18 & -6 & -30 \\ -27 & -9 & -45 \end{pmatrix}$$

where we use the associative law of multiplication throughout.

3. $A^2 = \begin{pmatrix} 0 & -3 \\ 3 & 3 \end{pmatrix}, A^3 = \begin{pmatrix} -3 & -6 \\ 6 & 3 \end{pmatrix}, A^4 = \begin{pmatrix} -9 & -9 \\ 9 & 0 \end{pmatrix}.$

4. $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 5 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 8 \\ 5 \end{pmatrix}.$

   Each pair of entries consists of successive terms of the Fibonacci sequence.

5. $$A^2 = \begin{pmatrix} \cos(\theta+\phi) & \sin(\theta+\phi) \\ -\sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}$$

   using the addition formulae for sines and cosines.

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

7. We have already seen that $(A+B)^2 \neq A^2 + 2AB + B^2$ in general unless you know that $AB = BA$. More generally, you cannot use the binomial theorem to expand expressions such as $(A+B)^n$ unless you know that A and B commute.

8. (a) The adjacency matrix A is given by

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

   (b) To determine the number of paths of length 2 we calculate A^2 and get

$$\begin{pmatrix} 2 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 \end{pmatrix}.$$

   The number of paths we require is $(A^2)_{12} = 2$.

   (c) To determine the number of paths of length 3 we calculate A^3 and get

$$\begin{pmatrix} 0 & 0 & 4 & 4 \\ 0 & 0 & 4 & 4 \\ 4 & 4 & 0 & 0 \\ 4 & 4 & 0 & 0 \end{pmatrix}.$$

   The number of paths we require is $(A^3)_{14} = 4$.

9. The zero matrix.

10. (a) Let A be an $m \times n$ matrix. Then A^T is an $n \times m$ matrix. But by assumption, $A^T = A$ and so $m = n$ and A is square.

(b) If A is $m \times n$ then A^T is $n \times m$ and so the product AA^T is defined. We now calculate $(AA^T)^T = (A^T)^T A^T = AA^T$. Thus AA^T is symmetric, as claimed.

(c) This is an 'if and only if' statement and so there are two things to prove.

Suppose first that $AB = BA$. I prove that AB is symmetric. We calculate $(AB)^T = B^T A^T = BA$, since B and A are symmetric. We now use the fact that A and B commute to get that $(AB)^T = AB$. Thus AB is symmetric.

Now suppose that AB is symmetric. We need to prove that A and B commute. By assumption, $(AB)^T = AB$. On the other hand, $(AB)^T = B^T A^T = BA$. Thus $AB = BA$, as required.

11. (a) If $A = (a_{ij})$ is skew-symmetric then $(A^T)_{ij} = -a_{ij}$, i.e. $a_{ji} = -a_{ij}$, for all i, j. In particular, putting $i = j$ we get $a_{ii} = -a_{ii}$, so $a_{ii} = 0$ for all i, i.e. the diagonal elements are zero.

(b)
$$(B + B^T)^T = B^T + (B^T)^T = B^T + B = B + B^T,$$
so $B + B^T$ is symmetric.
$$(B - B^T)^T = B^T - (B^T)^T = B^T - B = -(B - B^T),$$
so $B - B^T$ is skew-symmetric.

(c) Hence $B = \frac{1}{2}(B + B^T) + \frac{1}{2}(B - B^T)$ is the sum of a symmetric matrix and a skew-symmetric matrix.

12. Here is the proof for the case $n = 2$. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. By assumption, $AB = BA$ for all choices of 2×2 matrices B.

Choosing $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ gives us that $b = c = 0$.

Choosing $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ gives us that $a = b$, hence result.

This proof can now be generalized to square matrices of arbitrary size.

Exercises 8.3

1. (a) This is a consistent system of equations with infinitely many solutions. The solution set is
$$\left\{ \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

(b) This is an inconsistent system that has no solutions.

(c) This is a consistent system with a unique solution

$$\begin{pmatrix} -\frac{3}{5} \\ \frac{14}{5} \\ -\frac{7}{5} \end{pmatrix}.$$

(d) This is a consistent system with a unique solution

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

(e) This is a consistent system with a unique solution

$$\begin{pmatrix} -6 \\ 5 \\ -1 \end{pmatrix}.$$

(f) This is a consistent system with infinitely many solutions The solution set is

$$\left\{ \begin{pmatrix} -3 \\ 2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

2. This is a consistent system with infinitely many solutions. The solution set is

$$\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}.$$

3. The important feature of each of the three elementary row operations is that they are reversible. Clearly interchanging the rows will not change the solution set nor will multiplying any row by a non-zero scalar. In the third elementary row operation, we change the jth equation to get a new jth equation. But if we carry out the elementary row operation $R_j \leftarrow R_j - \lambda R_i$ then we get the original jth equation back again.

Exercises 8.4

1. (a) 5.
   (b) 0.
   (c) 5.
   (d) 2.
   (e) -1200.

(f) 33.

(g) 4.

(h) 0.

2. We have that $(1-x)(3-x)-8=0$. Thus $x^2-4x-5=0$. Hence $(x+1)(x-5)=0$. It follows that $x=-1$ or $x=5$.

3. x.

4. Suppose first that $ad=bc\neq 0$. The all of $a,b.c,d$ are non-zero. Let

$$\lambda=\frac{a}{b}=\frac{c}{d}.$$

Then

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

and so one column of the matrix is a scalar multiple of the other. Suppose now that $ad=bc=0$. Then either $a=0$ or $d=0$, and $b=0$ or $c=0$. Therefore there are nine possible outcomes which lead to nine matrices. In every case, one column is a scalar multiple of the other. We also have to prove the converse: namely, that if one column is a scalar multiple of the other then the determinant of the matrix is zero. There are two cases to consider and they are proved by direct computation.

Exercises 8.5

1. (a) $\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$.

(b) $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$.

(c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$.

(d) $\begin{pmatrix} \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ 1 & -1 & -1 \\ -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{pmatrix}$.

(e) $\begin{pmatrix} 6 & -2 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$.

(f) $\begin{pmatrix} \frac{2}{9} & -\frac{2}{9} & \frac{1}{9} \\ \frac{2}{9} & \frac{1}{9} & -\frac{2}{9} \\ \frac{1}{9} & \frac{2}{9} & \frac{2}{9} \end{pmatrix}$.

2. The result follows from the fact that elementary row operations are reversible. More formally, it follows from the following observation. Let ε be any one of the elementary row operations and write $\varepsilon(A)$ to mean the matrix that results by applying ε to A. Define $E = \varepsilon(I)$. Then $\varepsilon(A) = EA$ and E is an invertible matrix.

3. (a) By definition, $AA^{-1} = I$. Thus by properties of determinants

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

It follows that $\det(A) \neq 0$ if and only if $\det(A^{-1}) \neq 0$, and that in this case $\det(A^{-1}) = \det(A)^{-1}$.

(b) By properties of deterninants, $\det(A) = \det(A^T)$. Thus $\det(A) \neq 0$ if and only if $\det(A^T) \neq 0$. Now $AA^{-1} = I = A^{-1}A$ and so $(A^{-1})^T A^T = I = A^T(A^{-1})^T$ from properties of the transpose. But these equations say that the inverse of A^T is $(A^{-1})^T$ and we have proved the result.

4. Observe that if such a matrix B exists then by properties of determinants $\det(A) \neq 0$. Thus A is invertible and, in fact, $B = A^{-1}$.

5. (a) The inverse matrix is

$$\begin{pmatrix} 10 & 11 \\ 11 & 10 \end{pmatrix}.$$

(b) The inverse matrix is

$$\begin{pmatrix} 3 & 8 \\ 22 & 1 \end{pmatrix}.$$

The decoded message is ULTRAZ We have used the Z at the end to pad the message out to have even length.

6. Injectivity: suppose that $f(\mathbf{x}) = f(\mathbf{x}')$. Then $A\mathbf{x} = A\mathbf{x}'$. Multiplying both sides of the equation by A^{-1} we get that $\mathbf{x} = \mathbf{x}'$. Surjectivity: let $\mathbf{y} \in \mathbb{R}^n$. Define $\mathbf{x} = A^{-1}\mathbf{y}$. Then $f(\mathbf{x}) = \mathbf{y}$.

7. (a) The determinant of the matrix A assembled from these column vectors is zero. This means that the vectors are linearly dependent. To find out how, we need to find a solution to $A\mathbf{x} = \mathbf{0}$. One that works is $(2,-1,3)$.

(b) The determinant is zero again and so the vectors are linearly dependent. A solution to the associated homogenous equations is $(2,5,-1)$.

(c) The determinant is 99 and so non-zero. This means the vectors are linearly independent.

8. This follows by Theorem 8.3.7 since we have three equations but four unknowns.

Exercises 8.6

1. (a)
$$C^{-1} = \begin{pmatrix} 3 & -2 & -2 \\ -1 & 1 & 1 \\ 2 & -1 & -2 \end{pmatrix}.$$

(b)
$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

2. (a) The characteristic polynomial is $x^2 - 5x + 4$. The eigenvalues are 1 and 4.

   (b) The characteristic polynomial is $-x^3 + 6x^2 - 3x - 10$. The eigenvalues are $2, -1$ and 5.

3. (a) Characteristic polynomial $-x^3 + 6x^2 - 11x + 6$. Eigenvalues are $1, 2, 3$. Corresponding eigenvectors are, respectively,
$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 6 \\ 1 \\ 3 \end{pmatrix}.$$
The matrix with these vectors as columns is invertible and so the matrix is diagonalizable.

   (b) Characteristic polynomial $-x^3 + 5x^2 - 8x + 4$. Eigenvalues are $1, 2, 2$. Corresponding eigenvectors are, respectively,
$$\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$
The matrix with these vectors as columns is invertible and so the matrix is diagonalizable.

   (c) Characteristic polynomial $-x^3 + 5x^2 - 8x + 4$. Eigenvalues are $1, 2, 2$. Corresponding eigenvectors are, respectively,
$$\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$
There are not enough linearly independent eigenvectors and so the matrix is not diagonalizable.

4. We use properties of invertible matrices. By definition $A \equiv B$ if and only if $B = P^{-1}AP$ for some invertible matrix P. Reflexivity: $A = IAI$. Symmetry: from $B = P^{-1}AP$ we get that $A = PBP^{-1}$. Transitivity: suppose that $A \equiv B$ and $B \equiv C$. Then $B = P^{-1}AP$ and $C = Q^{-1}BQ$. Thus $C = Q^{-1}P^{-1}APQ$ and $(PQ)^{-1} = Q^{-1}P^{-1}$.

5. (a)
$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

(b) The characteristic polynomial is $x^2 - x - 1$. The eigenvalues are
$$\phi = \frac{1}{2}\left(1+\sqrt{5}\right) \text{ and } \overline{\phi} = \frac{1}{2}\left(1-\sqrt{5}\right).$$
Observe that $\phi - \overline{\phi} = \sqrt{5}$.

(c) Eigenvectors are, respectively,
$$\begin{pmatrix} -1 \\ \overline{\phi} \end{pmatrix} \text{ and } \begin{pmatrix} -1 \\ \phi \end{pmatrix}$$
Put
$$P = \begin{pmatrix} -1 & -1 \\ \overline{\phi} & \phi \end{pmatrix}$$
Then
$$P^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\phi & -1 \\ \overline{\phi} & 1 \end{pmatrix}$$
It follows that the matrix A is diagonalizable. In fact
$$P^{-1}AP = \begin{pmatrix} \phi & 0 \\ 0 & \overline{\phi} \end{pmatrix}$$

(d) We have that
$$A = P \begin{pmatrix} \phi & 0 \\ 0 & \overline{\phi} \end{pmatrix} P^{-1}.$$
Thus
$$A^n = P \begin{pmatrix} \phi^n & 0 \\ 0 & \overline{\phi}^2 \end{pmatrix} P^{-1}.$$
It follows from this that the number of paths from vertex 1 to itself is given by the formula
$$\frac{1}{\sqrt{5}}\left(\phi^{n+1} - \bar{\phi}^{n+1}\right).$$
Compare this with the results of Section 5.5.

6. $-x^3 + a_2x^2 + a_1x + a_0$. Every real cubic polynomial where the coefficient of x^3 is -1 occurs as the characteristic polynomial of some real matrix.

7. Prove by induction that $A^n\mathbf{v} = \lambda^n\mathbf{v}$. To calculate the effect of $f(A)$ on $\mathbf{v}$ we simply calculate the effect of terms such as aA^n on $\mathbf{v}$, which is just $a\lambda^n\mathbf{v}$ and add the results up.

8. (a) The proofs that the set is closed under addition and subtraction are almost immediate. We prove closure under multiplication. Observe that

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c & -d \\ d & c \end{pmatrix} = \begin{pmatrix} ac-bd & -(ad+bc) \\ ad+bc & ac-bd \end{pmatrix}.$$

(b) Calculate

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

and verify that you obtain the same matrix as in part (a).

(c) The determinant of

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

is a^2+b^2 which is zero if and only if $a=0=b$ if and only if the matrix is the zero matrix. It follows that the non-zero elements of the set are always invertible. We have that

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}^{-1} = \frac{1}{a^2+b^2}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

which is again an element of the set.

(d) This follows from the Section 8.2 and what we proved above.

(e) Straightforward.

Exercises 8.7

1. (a) $(-13)267+(31)112=1$.

(b) $(-4)1870+(31)242=22$.

(c) $(11)1079+(-16)741=13$.

CHAPTER 9

Vectors

Exercises 9.1

1. (a) $BD = \mathbf{c} - \mathbf{a}$.
   (b) $AE = \mathbf{a} + \mathbf{c}$.
   (c) $DE = \mathbf{a}$.
   (d) $CF = \mathbf{c}$.
   (e) $AC = \mathbf{a} + \mathbf{b}$.
   (f) $BF = \mathbf{b} + \mathbf{c}$.

2. If the quadrilateral is a parallelogram then $\mathbf{a} = -\mathbf{c}$. Conversely, suppose that $\mathbf{a} + \mathbf{c} = \mathbf{0}$. Then because $\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} = \mathbf{0}$ we deduce that $\mathbf{b} + \mathbf{d} = \mathbf{0}$ and so the shape is a parallelogram.

3. (a) $EA = -(\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d})$.
   (b) $DA = -(\mathbf{a} + \mathbf{b} + \mathbf{c})$.
   (c) $DB = -(\mathbf{b} + \mathbf{c})$.
   (d) $CA = -(\mathbf{a} + \mathbf{b})$.
   (e) $EC = -(\mathbf{c} + \mathbf{d})$.
   (f) $BE = \mathbf{b} + \mathbf{c} + \mathbf{d}$.

4. The remaining sides are: $\mathbf{b} - \mathbf{a}$, $-\mathbf{a}$, $-\mathbf{b}$, $\mathbf{a} - \mathbf{b}$. This was obtained by dividing up the hexagon into equilateral triangles and then observing which lines were parallel to each other.

5. Calculate
$$(\|\mathbf{a}\|\,\mathbf{b} + \|\mathbf{b}\|\,\mathbf{a}) \cdot (\|\mathbf{a}\|\,\mathbf{b} - \|\mathbf{b}\|\,\mathbf{a})$$
using distributivity and the fact that $\|\mathbf{a}\|^2 = \mathbf{a} \cdot \mathbf{a}$. The answer is zero, and so the vectors are orthogonal.

6. Calculate
$$\left(\mathbf{b} - \frac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a}\right)\cdot\mathbf{a}$$
using the distributive law, the fact that $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}$, and the fact that $\|\mathbf{a}\|^2 = \mathbf{a}\cdot\mathbf{a}$.

7. (a) $\mathbf{0}$.
   (b) First we expand using distributivity
$$(\mathbf{u}+\mathbf{v})\times(\mathbf{u}-\mathbf{v}) = \mathbf{u}\times\mathbf{u} + \mathbf{v}\times\mathbf{u} - \mathbf{u}\times\mathbf{v} + \mathbf{v}\times\mathbf{v}.$$
   But $\mathbf{u}\times\mathbf{u} = \mathbf{0} = \mathbf{v}\times\mathbf{v}$ and $\mathbf{u}\times\mathbf{v} = -\mathbf{v}\times\mathbf{u}$. Thus the answer is $2\mathbf{v}\times\mathbf{u}$.

8. We calculate
$$\mathbf{a}\cdot(2\mathbf{b}-\mathbf{a}) = 2\mathbf{a}\cdot\mathbf{b} - \mathbf{a}\cdot\mathbf{a}$$
but $\mathbf{a}\cdot\mathbf{a} = 1$ and $\mathbf{a}\cdot\mathbf{b} = \cos\frac{\pi}{3} = \frac{1}{2}$. The result now follows.

9. The lefthand side is $(\mathbf{u}-\mathbf{v})^2 + (\mathbf{u}+\mathbf{v})^2$ which expands to
$$\mathbf{u}^2 - 2\mathbf{u}\cdot\mathbf{v} + \mathbf{v}^2 + \mathbf{u}^2 + 2\mathbf{u}\cdot\mathbf{v} + \mathbf{v}^2$$
which is just
$$2(\mathbf{u}^2 + \mathbf{v}^2)$$
as required. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors lying along two adjacent sides of the parallelogram. The diagonals are $\mathbf{u}+\mathbf{v}$ and $\mathbf{u}-\mathbf{v}$. The result now follows.

Exercises 9.2

1. (a) $\sqrt{6}$ and $\sqrt{14}$.
   (b) $3\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$.
   (c) $\mathbf{i} - \mathbf{j} - 2\mathbf{k}$.
   (d) 7.
   (e) 40°.
   (f) $\mathbf{i} - 5\mathbf{j} + 3\mathbf{k}$.
   (g) $\frac{1}{\sqrt{35}}(\mathbf{i} - 5\mathbf{j} + 3\mathbf{k})$.

2. $(\mathbf{i}\times\mathbf{i})\times\mathbf{k} = \mathbf{0}$, whereas $\mathbf{i}\times(\mathbf{i}\times\mathbf{k}) = -\mathbf{k}$. Thus the vector product is not associative.

3. 0 in both cases. This is a useful check when calculating vector products.

4. A diagonal is $\mathbf{i}+\mathbf{j}+\mathbf{k}$. The cosine of the angle between this vector and $\mathbf{i}$, one of the edges, is $\frac{1}{\sqrt{3}}$. Thus the angle is between 54 and 55 degrees.

5. The number $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$ is the determinant

$$\begin{vmatrix} 3 & -2 & -5 \\ 1 & 4 & -4 \\ 0 & 3 & 2 \end{vmatrix}$$

which is just 49.

Exercises 9.4

1. (a) We begin by finding the parametric equation of the line through the two given points. Let $\mathbf{r}$ be the position vector of a point on the line. Then the vectors

$$\mathbf{r}-(\mathbf{i}-\mathbf{j}+2\mathbf{k})$$

and

$$(2\mathbf{i}+3\mathbf{j}+4\mathbf{k})-(\mathbf{i}-\mathbf{j}+2\mathbf{k})$$

are parallel. Thus there is a scalar λ such that

$$\mathbf{r}=\mathbf{i}-\mathbf{j}+2\mathbf{k}+\lambda(\mathbf{i}+4\mathbf{j}+2\mathbf{k}).$$

This is the parametric equation of the line.

To obtain the non-parametric equation, we first equate components in the above equation and get

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

Now we eliminate the parameter s to get the non-parametric equations

$$x-1=\frac{y+1}{4},\quad \frac{y+1}{4}=\frac{z-2}{2}.$$

(b) We begin by finding the parametric equation of the plane through the three given points. First we must find two vectors that are parallel to the plane. Let

$$\mathbf{a}=(\mathbf{i}+2\mathbf{j}-\mathbf{k})-(\mathbf{i}+3\mathbf{k})=2\mathbf{j}-4\mathbf{k}$$

and

$$\mathbf{b}=(3\mathbf{i}-\mathbf{j}-2\mathbf{k})-(\mathbf{i}+3\mathbf{k})=2\mathbf{i}-\mathbf{j}-5\mathbf{k}.$$

Thus if $\mathbf{r}$ is the position vector of a point in the given plane then we have that

$$\mathbf{r}-(\mathbf{i}+3\mathbf{k})=\lambda\mathbf{a}+\mu\mathbf{b}$$

for some parameters λ and μ. Hence the parametric equation of the plane is

$$\mathbf{r}=\mathbf{i}+3\mathbf{k}+\lambda(2\mathbf{j}-4\mathbf{k})+\mu(2\mathbf{i}-\mathbf{j}-5\mathbf{k}).$$

To find the non-parametric equation of the given plane, we need to find a vector normal to the plane. The vector $\mathbf{a}\times\mathbf{b}$ will do the trick. This is

equal to $-14\mathbf{i}-8\mathbf{j}-4\mathbf{k}$. Thus if $\mathbf{r}$ is the position vector of a point in the given plane we have that

$$(\mathbf{r}-(\mathbf{i}+3\mathbf{k}))\cdot(14\mathbf{i}+8\mathbf{j}+4\mathbf{k})=0.$$

This simplifies to

$$7x+4y+2z=13.$$

(c) $x+y-z=3$.

2. (a) $\mathbf{r}=4\mathbf{i}+5\mathbf{j}+\mathbf{k}+\lambda(\mathbf{i}+\mathbf{j}+\mathbf{k})$.
   (b) $\mathbf{r}=5\mathbf{i}-4\mathbf{j}+\mu(2\mathbf{i}-3\mathbf{j}+\mathbf{k})$.
   (c) $\mathbf{i}+2\mathbf{j}-2\mathbf{k}$.

3. $\frac{1}{5}(7\mathbf{i}-8\mathbf{j})+\frac{\lambda}{5}(2\mathbf{i}-8\mathbf{j}+5\mathbf{k})$. But this could equally well be written $\frac{1}{5}(7\mathbf{i}-8\mathbf{j})+\mu(2\mathbf{i}-8\mathbf{j}+5\mathbf{k})$ where $\mu\in\mathbb{R}$.

4. Let θ be the angle that $\mathbf{q}-\mathbf{p}$ makes with the direction determined by $\mathbf{d}$. Then the required distance is $||\mathbf{q}-\mathbf{p}||\sin\theta$. This quickly leads to the result.

5. Let θ be the angle the vector $\mathbf{q}-\mathbf{p}$ makes with the normal $\mathbf{n}$. Then the required distance is $||\mathbf{q}-\mathbf{p}||\cos\theta$. This quickly leads to the result.

6. (a) The desired equation is $(x-1)^2+(y-1)^2+(z-1)^2=4$ (which can be multiplied out).
   (b) Write first in the form $(x-1)^2-1+(y-2)^2-4+(z-3)^2-9-2=0$ by completing the square. This gives us $(x-1)^2+(y-2)^2+(z-3)^2=16$. Thus the centre is $\mathbf{i}+2\mathbf{j}+3\mathbf{k}$ and the radius is 4.

7. Suppose that $\lambda\mathbf{u}+\mu\mathbf{v}+\nu\mathbf{w}=\mathbf{0}$. Take the inner product on both sides with $\mathbf{u}$. This yields $\lambda\mathbf{u}^2=0$. By assumption $\mathbf{u}^2\neq 0$ and so $\lambda=0$. This process may be repeated with $\mathbf{v}$ and $\mathbf{w}$ in turn giving $\mu=0$ and $\nu=0$, respectively. This implies that the original vectors are linearly independent.

8. $\mathbf{0}$.

9. The vector $\mathbf{v}\times\mathbf{w}$ is orthogonal to $\mathbf{v}$ and $\mathbf{w}$. The vector $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})$ is orthogonal to $\mathbf{v}\times\mathbf{w}$. Thus $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})$ is parallel to the plane determined by $\mathbf{v}$ and $\mathbf{w}$. It follows that $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})=\lambda\mathbf{v}+\mu\mathbf{w}$. Taking the inner product with $\mathbf{u}$ on both sides we get $\lambda(\mathbf{u}\cdot\mathbf{v})+\mu(\mathbf{u}\cdot\mathbf{w})=0$. The cases where $\mathbf{u}\cdot\mathbf{v}=0$ or $\mathbf{u}\cdot\mathbf{w}=0$ can be dealt with separately. Thus $\lambda=\gamma(\mathbf{u}\cdot\mathbf{w})$ and $\mu=-\gamma(\mathbf{u}\cdot\mathbf{v})$ for some γ. It follows that $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})=\gamma(\mathbf{u}\cdot\mathbf{w})\mathbf{v}-\gamma(\mathbf{u}\cdot\mathbf{v})\mathbf{w}$. To show that $\gamma=1$, calculate the $\mathbf{i}$ coordinates on both sides. They should both be $u_2v_1w_2+u_3v_1w_3-u_3v_3w_1-u_2v_2w_1$.

10. We need to compute $[\mathbf{c}\times\mathbf{d},\mathbf{a},\mathbf{b}]$. This equals $[\mathbf{a},\mathbf{b},\mathbf{c}\times\mathbf{d}]$. This is equal to $\mathbf{a}\cdot(\mathbf{b}\times(\mathbf{c}\times\mathbf{d}))$. We now use the previous result and tidy up.

Exercises 9.5

1. The characteristic polynomial is $\chi_A(x) = x^2 - 2x\cos\theta + 1$. Its discriminant is $4(\cos^2\theta - 1)$. For all angles except $\theta = 0, \pi$ the discriminant is negative and so there are no real eigenvalues. When $\theta = 0$ the matrix A is just the identity matrix. The eigenvalues are 1 (twice) and every non-zero vector is an eigenvector. When $\theta = \pi$ the matrix A is $-I$ and represents a rotation by π about the origin. The eigenvalues are -1 (twice) and every non-zero vector is an eigenvector. In all other cases, there are no real eigenvalues and so no eigenvectors. This makes sense geometrically since in this case the matrix represents a rotation about the origin.

2. The vector **u** is an eigenvector belonging to 1. This implies that the scalar multiplies of **u** define the mirror line which is therefore at an angle of $\frac{\theta}{2}$ to the x-axis. The vector **v** is the eigenvector belonging to -1. It is orthogonal to **u**.

3. (a) The first matrix arises by interchanging the two rows. The second by multiplying the first row by λ and the third by multiplying the second row by λ. The fourth arises by adding λ times the second row to the first, and the fifth by adding λ times the first to the second.

   (b) To prove invertibility simply calculate determinants in each case. It is easy to check that the inverse of each elementary matrix is also an elementary matrix. The first matrix is a reflection in the line $y = x$. The second stretches the x-coordinate by λ and the third the y-coordinate by λ. The fourth matrix is a *shear* parallel to the x-axis. The fifth matrix is a shear parallel to the y-axis. What these last two operations are is best seen by considering their effects on the unit square.

   (c) The matrix A is any $2 \times m$ matrix. The result follows by direct verification.

   (d) This is a consequence of the method described in Section 8.5 for calculating the inverse of a matrix by using elementary row operations. Let A be a 2×2 invertible matrix. Let $(\varepsilon_n \dots \varepsilon_1)A = I$ where the ε_i are elementary row operations. Then $E_n \dots E_1 A = I$. It follows that $A = E_1^{-1} \dots E_n^{-1}$. But the inverse of an elementary matrix is an elementary matrix and so we have written our invertible matrix as a product of elementary matrices.

4. The matrix is singular and so the columns are linearly dependent. There are various special cases. All entries zero, the first column only zero, the second column only zero, and neither of the two columns zero. These are therefore four possibilities.

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix}, \quad C = \begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix}, \quad D = \begin{pmatrix} a & \lambda a \\ c & \lambda c \end{pmatrix}.$$

   The matrix A collapses everything onto the origin. The matrix B collapses

everything onto the line determined by the vector $\begin{pmatrix} b \\ d \end{pmatrix}$. The matrix C collapses everything onto the line determined by the vector $\begin{pmatrix} a \\ c \end{pmatrix}$. The matrix D collapses everything onto the line determined by the vector $\begin{pmatrix} a \\ c \end{pmatrix}$. Thus everything is either collapsed to a point or to a line.

Exercises 9.7

1. (a) Let $n = x^2 + y^2 + z^2$. A square is congruent to one of $0, 1, 4$ modulo 8. By looking at cases, a sum of three squares cannot be congruent to 7 modulo 8.
   (b) We have that $3 = 1^2 + 1^2 + 1^2$ and $21 = 4^2 + 2^2 + 1^2$. But $63 \equiv 7 \pmod 8$ and so by part (1) cannot be written as a sum of three squares.

2. (a) $\mathbf{u}\mathbf{u}^* = a^2 + b^2 + c^2 + d^2$.
   (b) We may therefore define
   $$\mathbf{u}^{-1} = \frac{1}{a^2 + b^2 + c^2 + d^2}\mathbf{u}^*.$$

3. (a) The product $\mathbf{uv}$ is given by
   $$\begin{aligned}(\lambda\mu - u_1v_1 - u_2v_2 - u_3v_3) + \\ (\lambda v_1 + u_1\mu + u_2v_3 - u_3v_2)\mathbf{i} + \\ (\lambda v_2 + u_2\mu + u_3v_1 - u_1v_3)\mathbf{j} + \\ (\lambda v_3 + u_3\mu + u_1v_2 - u_2v_1)\mathbf{k}\end{aligned}$$
   You can now verify (!) that
   $$\begin{aligned}(\lambda^2 + u_1^2 + u_2^2 + u_3^2)(\mu^2 + v_1^2 + v_2^2 + v_3^2) = \\ (\lambda\mu - u_1v_1 - u_2v_2 - u_3v_3)^2 + \\ (\lambda v_1 + u_1\mu + u_2v_3 - u_3v_2)^2 + \\ (\lambda v_2 + u_2\mu + u_3v_1 - u_1v_3)^2 + \\ (\lambda v_3 + u_3\mu + u_1v_2 - u_2v_1)^2\end{aligned}$$
   (b) The equation above speaks for itself.

4. The product $(\lambda + iu_1, u_2 + iu_3)(\mu + iv_1, v_2 + iv_3)$ has first component equal to
   $$\lambda\mu - u_1v_1 - u_2v_2 - u_3v_3 + i\left(\lambda v_1 + \mu u_1 + \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}\right)$$
   and second component equal to
   $$\lambda v_2 + \mu u_2 + \begin{vmatrix} u_3 & u_1 \\ v_3 & v_1 \end{vmatrix} + i\left(\lambda v_3 + \mu u_3 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}\right).$$

5.

CHAPTER 10

Principal axes theorem

Exercises 10.1

1. Begin with $\frac{1}{2}\left(\mathbf{u}^2+\mathbf{v}^2-(\mathbf{u}-\mathbf{v})^2\right)$ and observe that $(\mathbf{u}-\mathbf{v})^2=\mathbf{u}^2-2\mathbf{u}\cdot\mathbf{v}+\mathbf{v}^2$. The result now follows.

2. (1) We are given that $A^TA=I$. We take determinants of both sides and using the fact that the determinant of a product is a product of determinants we get $\det(A^T)\det(A)=\det(I)=1$. But $\det(A^T)=\det(A)$. Thus $\det(A)^2=1$. It follows that $\det(A)=\pm 1$.

   2 By part (1) we know that A is invertible. Multiply the equation $A^TA=I$ on the right by A^{-1}. We get $A^T=A^{-1}$, as required.

   (3) Let A be orthogonal. To prove that A^T is orthogonal we calculate $(A^T)^TA^T=AA^T=I$ since $A^{-1}=A^T$.

   (4) Let A and B be orthogonal. We prove that AB is orthogonal by calculating $(AB)^TAB$. We have that $(AB)^TAB=B^TA^TAB$ using the fact that the transpose of a product is the reverse product of the transposes. But by assumption $A^TA=I$ and $B^TB=I$ it follows that $(AB)^TAB=I$. We have therefore proved that AB is also an orthogonal matrix.

   (5) Let A be symmetric and B orthogonal. We prove that B^TAB is symmetric. We calculate $(B^TAB)^T=B^TA^T(B^T)^T=B^TAB$ using the assumption that A is symmetric and the fact that taking the transpose twice leaves us where we started.

3. (a) We prove that $\mathrm{tr}(AB)=\mathrm{tr}(BA)$. Let $A=(a_{ij})$ and $B=(b_{ij})$. Then $\mathrm{tr}(AB)=\sum_{i=1}^n\sum_{j=1}^n a_{ij}b_{ji}$ and $\mathrm{tr}(BA)=\sum_{i=1}^n\sum_{j=1}^n b_{ij}a_{ji}$. By commutativity of multiplication of real numbers these two sums are equal.

   (b) There are no solutions. If X could be solved then using properties of the trace, the trace of the lefthand side would be zero, whereas the trace of the righthand side would be n.

4. Let A and B be orthogonal matrices with determinant 1. We have already

proved that AB is orthogonal, and $\det(AB) = \det(A)\det(B) = 1$. Thus the set $SO(3)$ is closed under multiplication. Let A be an orthogonal matrix with determinant 1. Its inverse is A^T. We have that

5. We have to prove that $\|A\mathbf{u} - A\mathbf{v}\| = \|\mathbf{u} - \mathbf{v}\|$. This follows by calculating $(A(\mathbf{u} - \mathbf{v}))^T(\mathbf{u} - \mathbf{v})$ using the fact that A is an orthogonal matrix.

6. We are given that $1 + 2\cos\phi \in \mathbb{Z}$. Thus $\cos\phi = \frac{n}{2}$. It follows that $n = 0, 1, -1, 2, -2$.

7. The result is straightforward in the 2×2 case, and using that result the calculation in the 3×3 case can be simplified. For the general case, let $n \geq 4$. Let A be an $n \times n$ matrix, and let B be the $(n-1) \times (n-1)$ matrix that results from A by crossing out the first row and first column. Then the result follows from the observation that $\det(1 + \varepsilon A) \approx (1 + \varepsilon a_{11})\det(B)$. This is proved using the expansion of $\det(A)$ along the top row. The first term in this expansion is $(1 + \varepsilon a_{11})\det(B)$. All the other terms have the form $(-1)^{1+j}\varepsilon a_{1j}$, where $j \neq 1$, times the $(n-1) \times (n-1)$ determinant associated with a_{1j}. In particular, they all have a factor of ε. But the $(n-1) \times (n-1)$ determinant associated with a_{1j} will also contain a factor of ε. It follows that the remaining terms in the expansion of the determinant along the top row will all contain a factor ε^2, and so they can be neglected.

Exercises 10.2

1. Calculate
$$\mathbf{u}' \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{v} - \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}}\right)(\mathbf{v} \cdot \mathbf{v}) = 0,$$
as required.

2. (a) The characteristic polynomial is $-x^3 + 36x^2 - 81x - 4374$. There are three distinct eigenvalues $27, 18, -9$. Eigenvectors associated with each of these eigenvalues, respectively, are
$$\begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}.$$
and so the corresponding orthogonal matrix is
$$\frac{1}{3}\begin{pmatrix} 2 & 1 & 2 \\ -1 & -2 & 2 \\ -2 & 2 & 1 \end{pmatrix}$$

   (b) The characteristic polynomial is $-x^3 + 27x^2 + 54$. There are three eigenvalues $6, -3, -3$, but this time we have a repeated egienvalue which will

cause us extra work. Eigenvectors associated with each of these eigenvalues, respectively, are

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

But the last two eigenvectors, which are associated with the same eigenvalue, are not orthogonal. We therefore use the Gram-Schmidt process: I shall keep the first of the two eigenvectors belonging to -3 and change the third, but you could equally well carry out the calculations the other way around. You will get different, but equally valid, answers. In this way, we get

$$\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} -4 \\ 5 \\ -2 \end{pmatrix}.$$

which are orthogonal. The corresponding orthogonal matrix is

$$\begin{pmatrix} \frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{\sqrt{45}} \\ \frac{2}{3} & 0 & \frac{2}{\sqrt{45}} \\ \frac{1}{3} & -\frac{2}{\sqrt{5}} & -\frac{2}{\sqrt{45}} \end{pmatrix}$$

Exercises 10.3

1. The result follows from the observation that $\mathrm{tr}(P^{-1}AP) = \mathrm{tr}((P^{-1}A)P) = \mathrm{tr}(P(P^{-1}A)) = \mathrm{tr}(A)$.

2. By orthogonal diagonalization we may assume that all cross-terms in the quadric have been removed. The equation therefore has the following form.

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + ax + by + cz + d = 0.$$

There are two cases to consider.

(a) $\lambda_1\lambda_2\lambda_3 \neq 0$.

(b) $\lambda_1\lambda_2\lambda_3 = 0$.

Case (a). We combine terms in x, those in y and those in z, and complete the squares. We then relabel variables and constants and end up with an equation of the form

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + d = 0.$$

Subsequent analysis leads to the canonical forms (1)–(6) in the table. This involves interchanging variables in some cases.

Case (b). By relabelling variables, this boils down to two cases.

(a) $\lambda_1 \neq 0$ and $\lambda_2 = \lambda_3 = 0$.

(b) $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$ and $\lambda_3 = 0$.

Subcase (a). The equation has the form

$$\lambda_1 x^2 + ax + by + cz + d = 0.$$

By completing the square for the expression $\lambda_1 x^2 + ax$ and relabelling, we may assume that the equation has the form

$$\lambda_1 x^2 + by + cz + d = 0.$$

If $b = c = 0$ then we obtain the canonical forms (15)–(16). Suppose that $b \neq 0$ and $c = 0$. Thus the equation has the form

$$\lambda_1 x^2 + by + d = 0.$$

This may be written

$$\lambda_1 x^2 + b(y + \tfrac{d}{b}) = 0.$$

We now change variable by translation and relabel to get

$$\lambda_1 x^2 + by = 0.$$

This leads to canonical form (14).

Subcase (b). The equation has the form

$$\lambda_1 x^2 + \lambda_2 y^2 + ax + by + cz + d = 0.$$

By completing the square for the terms involving x and y, respectively, and relabelling, we may assume the equation has the form

$$\lambda_1 x^2 + \lambda_2 y^2 + cz + d = 0.$$

There are now two cases. If $c = 0$, we get the canonical forms (9)–(13). If $c \neq 0$, we get the canonical forms (7) and (8).

3. (a) Matrix form of equation is

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -2 & 8 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (36).$$

The characteristic polynomial of the symmetric matrix is $x^2 - 13x + 36$. The eigenvalues are 9 and 4. Corresponding eigenvectors are

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix} \text{ and } \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

By changing variables and relabelling we get the equation $9x^2+4y^2=36$. The canonical form is therefore

$$\frac{x^2}{9}+\frac{y^2}{4}=1$$

which is an ellipse.

(b) Matrix form of the equation is

$$\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}+\begin{pmatrix} 14 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=(-3).$$

The characteristic polynomial of the symmetric marix is $x^2-6x-16$. The eigenvalues are -2 and 8 and the corresponding eigenvectors are

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Carrying out a change of variables and then relabelling we obtain

$$-2x^2+8y^2+\tfrac{12}{\sqrt{2}}x+\tfrac{16}{\sqrt{2}}y+3=0.$$

[Don't forget to normalize your eigenvectors to construct the orthogonal matrix]. Completing the squares we get

$$-2(x-\tfrac{3}{\sqrt{2}})^2+8(y+\tfrac{1}{\sqrt{2}})^2+8=0.$$

Relabelling variables we get

$$-2x^2+8y^2+8=0$$

or

$$\frac{x^2}{4}-y^2=1.$$

This is a hyperbola.

(c) Matrix form of equation is

$$\begin{pmatrix} x & y & z \end{pmatrix}\begin{pmatrix} 2 & 1 & -3 \\ 1 & 2 & -3 \\ -3 & -3 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=(0).$$

The eigenvalues are 9, 1, 0. Thus by changing variables and relabelling we obtain $9x^2+y^2=0$. This is canonical form (13) and is a pair of intersecting planes.