Chapter 2

Fields and vector spaces

2.1 Fields

1. Let F be a field.

   (a) We wish to prove that $-1 \neq 0$. Let us argue by contradiction, and assume that $-1 = 0$. Then $\alpha + (-1) = \alpha$ for all $\alpha \in F$; in particular, $1 + (-1) = 1$. But $1 + (-1) = 0$ by definition of -1, and therefore $1 = 0$ must hold. This contradicts the definition of a field (which states that 1 and 0 are distinct elements), and hence $-1 = 0$ cannot hold in a field.

   (b) It need not be the case that $-1 \neq 1$; in fact, in $\mathbf{Z}_2$, $1 + 1 = 0$, and therefore $-1 = 1$.

2. Let F be a field. We wish to show that the multiplicative identity 1 is unique. Let us suppose that $\gamma \in F$ satisfies $\alpha\gamma = \alpha$ for all $\alpha \in F$. We then have $1 = 1 \cdot \gamma$ (since γ is a multiplicative identity), and also $\gamma = 1 \cdot \gamma$ (since 1 is a multiplicative identity). This implies that $\gamma = 1$, and hence the multiplicative identity is unique.

3. Let F be a field and let $\alpha \in F$ be nonzero. We wish to show that the multiplicative inverse of α is unique. Suppose $\beta \in F$ satisfies $\alpha\beta = 1$. Then, multiplying both sides of the equation by α^{-1}, we obtain $\alpha^{-1}(\alpha\beta) = \alpha^{-1} \cdot 1$, or $(\alpha^{-1}\alpha)\beta = \alpha^{-1}$, or $1 \cdot \beta = \alpha^{-1}$. It follows that $\beta = \alpha^{-1}$, and thus α has a unique multiplicative inverse.

4. Let F be a field, and suppose $\alpha, \beta, \gamma \in F$.

   (a) We have $(-\alpha) + \alpha = 0$; since the additive inverse of $-\alpha$ is unique, this implies that $\alpha = -(-\alpha)$.

   (b) Using the associate and commutative properties of addition, we can rewrite $(\alpha + \beta) + (-\alpha + (-\beta))$ as $(\alpha + (-\alpha)) + (\beta + (-\beta)) = 0 + 0 = 0$. Therefore, $-(\alpha + \beta) = -\alpha + (-\beta)$.

   (c) As in the last part, we can use commutativity and associativity to show that $(\alpha - \beta) + (-\alpha + \beta) = 0$.

   (d) We have $\alpha\beta + \alpha(-\beta) = \alpha(\beta + (-\beta)) = \alpha \cdot 0 = 0$, and this proves that $-(\alpha\beta) = \alpha(-\beta)$.

   (e) This follows from the previous result and the commutative property of multiplication.

   (f) Applying the first, fourth, and fifth results, we have $(-\alpha)(-\beta) = -(\alpha(-\beta)) = -(-(\alpha\beta)) = \alpha\beta$.

   (g) Assume $\alpha \neq 0$. Then $\alpha(\alpha^{-1} + (-\alpha)^{-1}) = \alpha\alpha^{-1} + \alpha(-\alpha)^{-1} = 1 - (-\alpha)(-\alpha)^{-1} = 1 - 1 = 0$. Since $\alpha \neq 0$, $\alpha(\alpha^{-1} + (-\alpha)^{-1}) = 0$ implies that $\alpha^{-1} + (-\alpha)^{-1} = 0$, which in turn implies that $(-\alpha)^{-1} = -(\alpha^{-1})$.

   (h) Using the associative and commutative properties of multiplication, we can rewrite $(\alpha\beta)(\alpha^{-1}\beta^{-1})$ as $(\alpha\alpha^{-1})(\beta\beta^{-1}) = 1 \cdot 1 = 1$. This shows that $\alpha^{-1}\beta^{-1} = (\alpha\beta)^{-1}$.

   (i) Using the definition of subtraction, the distributive property, and the fourth property above, we have $\alpha(\beta - \gamma) = \alpha(\beta + (-\gamma)) = \alpha\beta + \alpha(-\gamma) = \alpha\beta + (-(\alpha\gamma)) = \alpha\beta - \alpha\gamma$.

   (j) This is proved in the same way as the previous result.

5. (a) By definition, $i^2 = (0 + 1 \cdot i)(0 + 1 \cdot i) = (0 \cdot 0 - 1 \cdot 1) + (0 \cdot 1 + 1 \cdot 0)i = -1 + 0 \cdot i = -1$.

   (b) We will prove only the existence of multiplicative inverses, the other properties of a field being straightforward (although possibly tedious) to verify. Let $a + bi$ be a nonzero complex number (which means that $a \neq 0$ or $b \neq 0$). We must prove that there exists a complex number $c + di$ such that $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$ equals $1 = 1 + 0 \cdot i$. This is equivalent to the pair of equations $ac - bd = 1$, $ad + bc = 0$. If $a \neq 0$, then the second equation yields $d = -bc/a$; substituting into the first equations yields $ac + b^2c/a = 1$, or $c = a/(a^2 + b^2)$. Then $d = -bc/a = -b/(a^2 + b^2)$. This solution is well-defined even if $a = 0$ (since in that case $b \neq 0$), and it can be directly verified that it satisfies $(a + bi)(c + di) = 1$ in that case also. Thus each nonzero complex number has a multiplicative inverse.

6. Let F be a field, and let $\alpha, \beta, \gamma \in F$ with $\gamma \neq 0$. Suppose $\alpha\gamma = \beta\gamma$. Then, multiplying both sides by γ^{-1}, we obtain $(\alpha\gamma)\gamma^{-1} = (\beta\gamma)\gamma^{-1}$, or $\alpha(\gamma\gamma^{-1}) = \beta(\gamma\gamma^{-1})$, which then yields $\alpha \cdot 1 = \beta \cdot 1$, or $\alpha = \beta$.

7. Let F be a field and let α, β be elements of F. We wish to show that the equation $\alpha + x = \beta$ has a unique solution. The proof has two parts. First, if x satisfies $\alpha + x = \beta$, then adding $-\alpha$ to both sides shows that x must equal $-\alpha + \beta = \beta - \alpha$. This shows that the equation has at most one solution. On the other hand, $x = -\alpha + \beta$ is a solution since $\alpha + (-\alpha + \beta) = (\alpha - \alpha) + \beta = 0 + \beta = \beta$. Therefore, $\alpha + x = \beta$ has a unique solution, namely, $x = -\alpha + \beta$.

8. Let F be a field, and let $\alpha, \beta \in F$. We wish to determine if the equation $\alpha x = \beta$ always has a unique solution. The answer is no; in fact, there are three possible cases. First, if $\alpha = 0$ and $\beta = 0$, then $\alpha x = \beta$ is satisfied by every element of F, and there are multiple solutions in this case. Second, if $\alpha = 0$, $\beta \neq 0$, then $\alpha x = \beta$ has no solution (since $\alpha x = 0$ for all $x \in F$ in this case). Third, if $\alpha \neq 0$, then $\alpha x = \beta$ has the unique solution $x = \alpha^{-1}\beta$. (Existence and uniqueness is proved in this case as in the previous exercise.)

9. Let F be a field.

   Let $\alpha, \beta, \gamma, \delta \in F$, with $\beta, \delta \neq 0$. We wish to show that

$$\frac{\alpha}{\beta} + \frac{\gamma}{\delta} = \frac{\alpha\delta + \beta\gamma}{\beta\delta}, \quad \frac{\alpha}{\beta} \cdot \frac{\gamma}{\delta} = \frac{\alpha\gamma}{\beta\delta}.$$

   Using the definition of division, the commutative and associative properties of multiplication, and finally the distributive property, we obtain

$$\begin{aligned}
\frac{\alpha}{\beta} + \frac{\gamma}{\delta} &= \alpha\beta^{-1} + \gamma\delta^{-1} = (\alpha \cdot 1)\beta^{-1} + (\gamma \cdot 1)\delta^{-1} \\
&= (\alpha(\delta\delta^{-1}))\beta^{-1} + (\gamma(\beta\beta^{-1}))\delta^{-1} = ((\alpha\delta)\delta^{-1})\beta^{-1} + ((\gamma\beta)\beta^{-1})\delta^{-1} \\
&= (\alpha\delta)(\delta^{-1}\beta^{-1}) + (\gamma\beta)(\beta^{-1}\delta^{-1}) = (\alpha\delta)(\delta\beta)^{-1}) + (\gamma\beta)(\beta\delta)^{-1} \\
&= (\alpha\delta)(\beta\delta)^{-1}) + (\gamma\beta)(\beta\delta)^{-1} = ((\alpha\delta) + (\gamma\beta))(\beta\delta)^{-1} \\
&= \frac{\alpha\delta + \beta\gamma}{\beta\delta}.
\end{aligned}$$

   Similarly,

$$\begin{aligned}
\frac{\alpha}{\beta} \cdot \frac{\gamma}{\delta} &= (\alpha\beta^{-1})(\gamma\delta^{-1}) = ((\alpha\beta^{-1})\gamma)\delta^{-1} = (\alpha(\beta^{-1}\gamma))\delta^{-1} = (\alpha(\gamma\beta^{-1}))\delta^{-1} \\
&= ((\alpha\gamma)\beta^{-1})\delta^{-1} = (\alpha\gamma)(\beta^{-1}\delta^{-1}) = (\alpha\gamma)(\beta\delta)^{-1} = \frac{\alpha\gamma}{\beta\delta}.
\end{aligned}$$

   Now assuming that $\beta, \gamma, \delta \neq 0$, we wish to show that

$$\frac{\alpha/\beta}{\gamma/\delta} = \frac{\alpha\delta}{\beta\gamma}.$$

Using the fact that $(\delta^{-1})^{-1} = \delta$, we have

$$\begin{aligned}\frac{\alpha/\beta}{\gamma/\delta} &= (\alpha\beta^{-1})(\gamma\delta^{-1})^{-1} = (\alpha\beta^{-1})(\gamma^{-1}\delta) = ((\alpha\beta^{-1})\gamma^{-1})\delta \\ &= (\alpha(\beta^{-1}\gamma^{-1}))\delta = (\alpha(\beta\gamma)^{-1})\delta = \alpha((\beta\gamma)^{-1}\delta) \\ &= \alpha(\delta(\beta\gamma)^{-1}) = (\alpha\delta)(\beta\gamma)^{-1} = \frac{\alpha\delta}{\beta\gamma}.\end{aligned}$$

10. Let F be a field, and let $\alpha \in F$ be given. We wish to prove that, for any $\beta_1, \ldots, \beta_n \in F$, $\alpha(\beta_1 + \cdots + \beta_n) = \alpha\beta_1 + \cdots + \alpha\beta_n$. We argue by induction on n. For $n = 1$, the result is simply $\alpha\beta_1 = \alpha\beta_1$. Suppose that for some $n \geq 2$, $\alpha(\beta_1 + \cdots + \beta_{n-1}) = \alpha\beta_1 + \cdots + \alpha\beta_{n-1}$ for any $\beta_1, \ldots, \beta_{n-1} \in F$. Let $\beta_1, \ldots, \beta_{n-1}, \beta_n \in F$. Then

$$\begin{aligned}&\alpha(\beta_1 + \cdots + \beta_n) = \alpha((\beta_1 + \cdots + \beta_{n-1}) + \beta_n) \\ =&\alpha(\beta_1 + \cdots + \beta_{n-1}) + \alpha\beta_n = \alpha\beta_1 + \cdots + \alpha\beta_{n-1} + \alpha\beta_n.\end{aligned}$$

(In the last step, we applied the induction hypothesis, and in the step preceding that, the distributive property of addition of multiplication.) This shows that $\alpha(\beta_1 + \cdots + \beta_n) = \alpha\beta_1 + \cdots + \alpha\beta_n$, and the general result now follows by induction.

11. (a) The space $\mathbf{Z}$ is not a field because multiplicative inverses do not exist in general. For example, $2 \neq 0$, yet there exists no $n \in \mathbf{Z}$ such that $2n = 1$.

    (b) The space $\mathbf{Q}$ of rational number is a field. Assuming the usual definitions for addition and multiplication, all of the defining properties of a field are straightforward to verify.

    (c) The space of positive real numbers is not a field because there is no additive identity. For any $z \in (0, \infty)$, $x + z > x$ for all $x \in (0, \infty)$.

12. Let $F = \{(\alpha, \beta) : \alpha, \beta \in \mathbf{R}\}$, and define addition and multiplication on F by $(\alpha, \beta)+(\gamma, \delta) = (\alpha+\gamma, \beta+\delta)$, $(\alpha, \beta) \cdot (\gamma, \delta) = (\alpha\gamma, \beta\delta)$. With these definitions, F is not a field because multiplicative inverses do not exists. It is straightforward to verify that $(0, 0)$ is an additive inverse and $(1, 1)$ is a multiplicative inverse. Then $(1, 0) \neq (0, 0)$, yet $(1, 0) \cdot (\alpha, \beta) = (\alpha, 0) \neq (1, 1)$ for all $(\alpha, \beta) \in F$. Since F contains a nonzero element with no multiplicative inverse, F is not a field.

13. Let $F = (0, \infty)$, and define addition and multiplication on F by $x \oplus y = xy$, $x \odot y = x^{\ln y}$. We wish to show that F is a field. Commutativity and associativity of addition follow immediately from these properties for ordinary multiplication of real numbers. Obviously 1 is an additive inverse, and the additive inverse of $x \in F$ is its reciprocal $1/x$. The properties of multiplication are less obvious, but note that $x \odot y = e^{\ln(y \odot x)} = e^{\ln(y)\ln(x)}$, and this formula makes both commutativity and associativity easy to verify. We also see that e is a multiplicative identity: $x \odot e = x^{\ln(e)} = x^1 = x$ for all $x \in F$. For any $x \in F$, $x \neq 1$, $y = e^{1/\ln(x)}$ is a multiplicative inverse. Finally, for any $x, y, z \in F$, $x \odot (y \oplus z) = x \odot (yz) = x^{\ln(yz)} = x^{\ln(y)+\ln(z)} = x^{\ln(y)}x^{\ln(z)} = (x \odot y)(x \odot z) = (x \odot y) \oplus (x \odot z)$. Thus the distributive property holds.

14. Suppose F is a set on which are defined two operations, addition and multiplication, such that all the properties of a field are satisfied except that addition is not assumed to be commutative. We wish to show that, in fact, addition must be commutative, and therefore F must be a field. We first note that it is possible to prove that $0 \cdot \gamma = 0$, $-1 \cdot \gamma = -\gamma$, and$-(-\gamma) = \gamma$ for all $\gamma \in F$ without invoking commutativity of addition. Moreover, for all $\alpha, \beta \in F$, $-\beta + (-\alpha) = -(\alpha + \beta)$ since $(\alpha + \beta) + (-\beta + (-\alpha)) = ((\alpha + \beta) + (-\beta)) + (-\alpha) = (\alpha + (\beta + (-\beta))) + (-\alpha) = (\alpha + 0) + (-\alpha) = \alpha + (-\alpha) = 0$. We therefore conclude that $-1 \cdot (\alpha + \beta) = -\beta + (-\alpha)$ for all $\alpha, \beta \in F$. But, by the distributive property, $-1 \cdot (\alpha + \beta) = -1 \cdot \alpha + (-1) \cdot \beta = -\alpha + (-\beta)$, and therefore $-\alpha + (-\beta) = -\beta + (-\alpha)$ for all $\alpha, \beta \in F$. Applying this property to $-\alpha$, $-\beta$ in place of α, β, respectively, yields $\alpha + \beta = \beta + \alpha$ for all $\alpha, \beta \in F$, which is what we wanted to prove.

15. (a) In $\mathbf{Z}_2 = \{0, 1\}$, we have $0 + 0 = 0$, $1 + 0 = 0 + 1 = 1$, and $1 + 1 = 0$. This shows that $-0 = 0$ (as always) and $-1 = 1$. Also, $0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0$, $1 \cdot 1 = 1$, and $1^{-1} = 1$ (as in any field).

(b) The addition and multiplication tables for $\mathbf{Z}_3 = \{0, 1, 2\}$ are

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

We see that $-0 = 0$, $-1 = 2$, $-2 = 1$, $1^{-1} = 1$, $2^{-1} = 2$.
The addition and multiplication tables for $\mathbf{Z}_5 = \{0, 1, 2, 3, 4\}$ are

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$+$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

We see that $-0 = 0$, $-1 = 4$, $-2 = 3$, $-3 = 2$, $-4 = 1$, $1^{-1} = 1$, $2^{-1} = 3$, $3^{-1} = 2$, $4^{-1} = 4$.

16. Let p be a positive integer that is prime. We wish to show that $\mathbf{Z}_p$ is a field. The commutativity of addition and multiplication in $\mathbf{Z}_p$ follow immediately from the commutativity of addition and multiplication of integers, and similarly for associativity and the distributive property. Obviously 0 is an additive identity and 1 is a multiplicative identity. Also, $1 \neq 0$. For any $\alpha \in \mathbf{Z}_p$, the integer $p - \alpha$, regarded as an element of $\mathbf{Z}_p$, satisfies $\alpha + (p - \alpha) = 0$ in $\mathbf{Z}_p$; therefore every element of $\mathbf{Z}_p$ has an additive inverse. It remains only to prove that every nonzero element of $\mathbf{Z}_p$ has a multiplicative inverse. Suppose $\alpha \in \mathbf{Z}_p$, $\alpha \neq 0$. Since $\mathbf{Z}_p$ has only finitely many elements, $\alpha, \alpha^2, \alpha^3, \ldots$ cannot all be distinct; there must exist positive integers k, ℓ, with $k > \ell$, such that $\alpha^k = \alpha^\ell$ in $\mathbf{Z}_p$. This means that the integers α^k, α^ℓ satisfy $\alpha^k = \lambda^\ell + np$ for some positive integer n, which in turn yields $\alpha^k - \alpha^\ell = np$ or $\alpha^\ell(\alpha^{k-\ell} - 1) = np$. Now, a basic theorem from number theory states that if a prime p divides a product of integers, then it must divide one of the integers in the product. In this case, p must divide α or $\alpha^{k-\ell} - 1$. Since $0 < \alpha < p$, p does not divide α, and therefore p divides $\alpha^{k-\ell} - 1$. Therefore, $\alpha^{k-\ell} - 1 = sp$, where s is a positive integer; this is equivalent to $\alpha^{k-\ell} = 1$ in $\mathbf{Z}_p$. Finally, this means that $\alpha\alpha^{k-\ell-1} = 1$ in $\mathbf{Z}_p$, and therefore α has a multiplicative inverse, namely, $\alpha^{k-\ell-1}$. This completes the proof that $\mathbf{Z}_p$ is a field.

17. Let p be a positive integer that is not prime. It is easy to see that 1 is a multiplicative identity in $\mathbf{Z}_p$. Since p is not prime, there exist integers m, n satisfying $1 < m, n < p$ and $mn = p$. But then, if m and n are regarded as elements of $\mathbf{Z}_p$, $m, n \neq 0$ and $mn = 0$, which is impossible in a field. Therefore, $\mathbf{Z}_p$ is not a field when p is not prime.

18. Let F be a finite field.

(a) Consider the elements $1, 1+1, 1+1+1, \ldots$ in F. Since F contains only finitely many elements, there must exist two terms in this sequence that are equal, say $1+1+\cdots+1$ (ℓ terms) and $1+1+\cdots+1$ (k terms), where $k > \ell$. We can then add -1 to both sides ℓ times to show that $1+1+\cdots+1$ ($k-\ell$ terms) equals 0 in F. Since at least one of the sequence $1, 1+1, 1+1+1, \ldots$ equals 0, we can define n to be the smallest integer greater than 1 such that $1 + 1 + \cdots + 1 = 0$ (n terms). We call n the characteristic of the field.

(b) Given that the characteristic of F is n, for any $\alpha \in F$, we have $\alpha + \alpha + \cdots + \alpha = \alpha(1+1+\cdots+1) = \alpha \cdot 0 = 0$ if the sum has n terms.

(c) We now wish to show that the characteristic n is prime. Suppose, by way of contradiction, that $n = k\ell$, where $1 < k, \ell < n$. Define $\alpha = 1 + 1 + \cdots + 1$ (k terms) and $\beta = 1 + 1 + \cdots + 1$ (ℓ terms). Then $\alpha\beta = 1 + 1 + \cdots + 1$ (n terms), so that $\alpha\beta = 0$. But this implies that $\alpha = 0$ or $\beta = 0$, which contradicts the definition of the characteristic n. This contradiction shows that n must be prime.

19. We are given that $\mathbf{H}$ represents the space of quaternions and the definitions of addition and multiplication in $\mathbf{H}$. The first two parts of the exercise are purely computational.

    (a) $i^2 = j^2 = k^2 = -1$, $ij = k$, $ik = -j$, $jk = i$, $ji = -k$, $ki = j$, $kj = -i$, $ijk = -1$.

    (b) $x\overline{x} = \overline{x}x = x_1^2 + x_2^2 + x_3^2 + x_4^2$.

    (c) The additive identity in $\mathbf{H}$ is $0 = 0 + 0i + 0j + 0k$. The additive inverse of $x = x_1 + x_2 i + x_3 j + x_4 k$ is $-x = -x_1 - x_2 i - x_3 j - x_4 k$.

    (d) The calculations above show that multiplication is not commutative; for instance, $ij = k$, $ji = -k$.

    (e) It is easy to verify that $1 = 1 + 0i + 0j + 0k$ is a multiplicative identity for $\mathbf{H}$.

    (f) If $x \in \mathbf{H}$ is nonzero, then $x\overline{x} = x_1^2 + x_2^2 + x_3^2 + x_4^2$ is also nonzero. It follows that

    $$\frac{\overline{x}}{x\overline{x}} = \frac{x_1}{x\overline{x}} - \frac{x_2}{x\overline{x}} i - \frac{x_3}{x\overline{x}} j - \frac{x_4}{x\overline{x}} k$$

    is a multiplicative inverse for x:

    $$x\frac{\overline{x}}{x\overline{x}} = \frac{x\overline{x}}{x\overline{x}} = 1.$$

20. In order to solve the first two parts of this problem, it is convenient to prove the following result. Suppose F is a field under operations $+, \cdot$, G is a nonempty subset of F, and G is a field under operations $\oplus, \odot$. Moreover, suppose that for all $x, y \in G$, $x + y = x \oplus y$ and $x \cdot y = x \odot y$. Then G is a subfield of F. We already know (since the operations on F reduce to the operations on G when the operands belong to G) that G is closed under $+$ and $\cdot$. We have to prove that the additive and multiplicative identities of F belong to G, which we will do by showing that $0_F = 0_G$ (that is, the additive identity of F equals the additive identity of G under $\oplus$) and $1_F = 1_G$ (which has the analogous meaning). To prove the first, notice that $0_G + 0_G = 0_G \oplus 0_G$ since $0_G \in G$, and therefore $0_G + 0_G = 0_G$. Adding the additive inverse (in F) -0_G to both sides of this equation yields $0_G = 0_F$. A similar proof shows that $1_F = 1_G$. Thus $0_F, 1_F \in G$. We next show that if $x \in G$ and $-x$ denotes the additive inverse of x in F, then $-x \in G$. We write $\ominus x$ for the additive inverse of x in G. We have $x \oplus (\ominus x) = 0$, which implies that $x + (\ominus x) = 0$. But then, adding $-x$ to both sides, we obtain $\ominus x = -x$, and therefore $-x \in G$. Similarly, if $x \in G$, $x \neq 0$, and x^{-1} denotes the multiplicative inverse of x in F, then $x^{-1} \in G$. This completes the proof.

    (a) We wish to show that $\mathbf{R}$ is a subfield of $\mathbf{C}$. It suffices to prove that addition and multiplication in $\mathbf{C}$ reduce to the usual addition and multiplication in $\mathbf{R}$ when the operands are real numbers. If $x, y \in \mathbf{R} \subset \mathbf{C}$, then $(x + 0i) + (y + 0i) = (x + y) + (0 + 0)i = (x + y) + 0i = x + y$. Similarly, $(x + 0i)(y + 0i) = (xy - 0 \cdot 0) + (x \cdot 0 + 0 \cdot y)i = xy + 0i = xy$. The the operations on $\mathbf{C}$ reduce to the operations on $\mathbf{R}$ when the operands are elements of $\mathbf{R}$, and therefore $\mathbf{R}$ is a subfield of $\mathbf{C}$.

    (b) We now wish to show that $\mathbf{C}$ is a subfield of $\mathbf{H}$ by showing that the operations of $\mathbf{H}$ reduce to the operations on $\mathbf{C}$ when the operands belong to $\mathbf{C}$. Let $x = x_1 + x_2 i, y = y_1 + y_2 i$ belong to $\mathbf{C}$, so that $x = x_1 + x_2 i + 0j + 0k, y = y_1 + y_2 i + 0j + 0k$ can be regarded as elements of $\mathbf{H}$. By definition,

    $$\begin{aligned} x + y &= (x_1 + x_2 i + 0j + 0k) + (y_1 + y_2 i + 0j + 0k) \\ &= (x_1 + y_1) + (x_2 + y_2)i + (0 + 0)j + (0 + 0)k \\ &= (x_1 + y_1) + (x_2 + y_2)i, \\ xy &= (x_1 + x_2 i + 0j + 0k)(y_1 + y_2 i + 0j + 0k) \\ &= (x_1 y_1 - x_2 y_2 - 0 \cdot 0 - 0 \cdot 0) + \\ &\quad (x_1 y_2 + x_2 y_1 + 0 \cdot 0 - 0 \cdot 0)i + \\ &\quad (x_1 \cdot 0 - x_2 \cdot 0 + 0 \cdot y_1 + 0 \cdot y_2)j + \\ &\quad (x_1 \cdot 0 + x_2 \cdot 0 - 0 \cdot y_2 + 0 \cdot y_1)k \\ &= (x_1 y_1 - x_2 y_2) + (x_1 y_2 + x_2 y_1)i. \end{aligned}$$

    Thus both operations on $\mathbf{H}$ reduce to the usual operations on $\mathbf{C}$, which shows that $\mathbf{C}$ is a subfield of $\mathbf{H}$.

(c) Consider the subset $S = \{a + bi + cj \ : \ a, b, c \in \mathbf{R}\}$ of $\mathbf{H}$. We wish to determine whether S is a subsfield of $\mathbf{H}$. In fact, S is not a subfield because it is not closed under multiplication. For example, $i, j \in S$, but $ij = k \notin S$.

2.2 Vector spaces

1. Let F be a field, and let $V = \{0\}$, with addition and scalar multiplication on V defined by $0 + 0 = 0$, $\alpha \cdot 0 = 0$ for all $\alpha \in F$. We wish to prove that V is a vector space over F. This is a straightforward verification of the defining properties. The commutative property of addition is vacuous, since V contains a single element. We have $(0 + 0) + 0 = 0 + 0 = 0 + (0 + 0)$, so the associative property holds. The definition $0 + 0 = 0$ shows both that 0 is an additive identity and that 0 is the additive inverse of 0, the only vector in V. Next, for all $\alpha, \beta \in F$, we have $\alpha(\beta \cdot 0) = \alpha \cdot 0 = 0 = (\alpha\beta) \cdot 0$, so the associative property of scalar multiplication is satisfied. Also, $\alpha(0 + 0) = \alpha \cdot 0 = 0 = 0 + 0 = \alpha \cdot 0 + \alpha \cdot 0$ and $(\alpha + \beta) \cdot 0 = 0 = \alpha \cdot 0 + \beta \cdot 0$, so both distributive properties hold. Finally, $1 \cdot 0 = 0$ by definition, so the final property of a vector space holds. Thus V is a vector space over F.

2. Let F be an infinite field, and let V be a nontrivial vector space over F. We wish to show that V contains infinitely many vectors. By definition, V contains a nonzero vector u. It suffices to show that, for all $\alpha, \beta \in F$, $\alpha \neq \beta$ implies $\alpha u \neq \beta u$, since then V contains the infinite subset $\{\alpha u \ : \ \alpha \in F\}$. Suppose $\alpha, \beta \in F$, $\alpha \neq \beta$. Then $\alpha u = \beta u$ if and only if $\alpha u - \beta u = 0$, that is, if and only if $(\alpha - \beta)u = 0$. Since $u \neq 0$ by assumption, this implies that $\alpha - \beta = 0$ by Theorem 5. Thus $\alpha u = \beta u$ implies $\alpha = \beta$. which completes the proof.

3. Let V be a vector space over a field F.

   (a) Suppose $z \in V$ is an additive identity. Then $z + 0 = z$ (since 0 is an additive identity) and $0 + z = 0$ (since z is an additive identity). Then $z = z + 0 = 0 + z = 0$, which shows that 0 is the only additive identity in V.

   (b) Let $u \in V$. If $u + v = 0$, then $-u + (u + v) = -u + 0$, which implies that $(-u + u) + v = -u$, or $0 + v = -u$, or finally $v = -u$. Thus the additive inverse $-u$ of u is unique.

   (c) Suppose $u, v \in V$. Then $(u+v)+(-u+(-v)) = ((u+v)+(-u))+(-v) = (u+(v+(-u)))+(-v) = (u+(-u+v))+(-v) = ((u+(-u))+v)+(-v) = (0+v)+(-v) = v+(-v) = 0$. By the preceding result, this shows that $-u + (-v) = -(u + v)$.

   (d) Suppose $u, v, w \in V$ and $u + v = u + w$. Then $-u + (u + v) = -u + (u + w)$, which implies that $(-u + u) + v = (-u + u) + w$, or $0 + v = 0 + w$, or finally $v = w$.

   (e) Suppose $\alpha \in F$ and 0 is the zero vector in V. Then $\alpha 0 + \alpha 0 = \alpha(0 + 0) = \alpha 0$; adding $-(\alpha 0)$ to both sides yields $\alpha 0 = 0$, as desired.

   (f) Suppose $\alpha \in F$, $u \in V$, and $\alpha u = 0$. If $\alpha \neq 0$, then α^{-1} exists and $\alpha^{-1}(\alpha u) = \alpha^{-1} \cdot 0$, which implies $(\alpha^{-1}\alpha)u = 0$ (applying the last result), which in turn yields $1 \cdot u = 0$ or finally $u = 0$. Therefore, $\alpha u = 0$ implies that $\alpha = 0$ or $u = 0$.

   (g) Suppose $u \in V$. Then $0 \cdot u + 0 \cdot u = (0 + 0) \cdot u = 0 \cdot u$. Adding $-(0 \cdot u)$ to both sides yields $0 \cdot u = 0$. We then have $u + (-1)u = 1 \cdot u + (-1)u = (1 + (-1))u = 0 \cdot u = 0$, which shows that $(-1)u = -u$.

4. We are to prove that if F is a field, then F^n is a vector space over F. This is a straightforward verification of the defining properties of a vector space, which follow in this case from the analogous properties of the field F. The details are omitted.

5. We are to prove that $F[a, b]$ (the space of all functions $f : [a, b] \to \mathbf{R}$) is a vector space over $\mathbf{R}$. Like the last exercise, this straightforward verification is omitted.

6. (a) Let p be a prime and n a positive integer. Since each of the n components of $x \in \mathbf{Z}_p^n$ can take on any of the p values $0, 1, \ldots, p - 1$, there are p^n distinct vectors in $\mathbf{Z}_p^n$.

(b) The elements of $\mathbf{Z}_2^2$ are $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$. We have $(0,0)+(0,0)=(0,0)$, $(0,0)+(0,1)=(0,1)$, $(0,0)+(1,0)=(1,0)$, $(0,0)+(1,1)=(1,1)$, $(0,1)+(0,1)=(0,0)$, $(0,1)+(1,0)=(1,1)$, $(0,1)+(1,1)=(1,0)$, $(1,0)+(1,0)=(0,0)$, $(1,0)+(1,1)=(0,1)$, $(1,1)+(1,1)=(0,0)$.

7. (a) The elements of $\mathcal{P}_1(\mathbf{Z}_2)$ are the polynomials 0, 1, x, $1+x$, which define distinct functions on $\mathbf{Z}_2$. We have $0+0=0$, $0+1=1$, $0+x=x$, $0+(1+x)=1+x$, $1+1=0$, $1+x=1+x$, $1+(1+x)=x$, $x+x=(1+1)x=0x=0$, $x+(1+x)=1+(x+x)=1$, $(1+x)+(1+x)=(1+1)+(x+x)=0+0=0$.

(b) Nominally, the elements of $\mathcal{P}_2(\mathbf{Z}_2)$ are $0, 1, x, 1+x, x^2, 1+x^2, x+x^2, 1+x+x^2$. However, since these elements are interpreted as functions mapping $\mathbf{Z}_2$ into $\mathbf{Z}_2$, it turns out that the last four functions equal the first four. In particular, $x^2=x$ (as functions), since $0^2=0$ and $1^2=1$. Then $1+x^2=1+x$, $x+x^2=x+x=0$, and $1+x+x^2=1+0=1$. Thus we see that the function spaces $\mathcal{P}_2(\mathbf{Z}_2)$ and $\mathcal{P}_1(\mathbf{Z}_2)$ are the same.

(c) Let V be the vector space consisting of all functions from $\mathbf{Z}_2$ into $\mathbf{Z}_2$. To specify $f\in V$ means to specify the two values $f(0)$ and $f(1)$. There are exactly four ways to do this: $f(0)=0$, $f(1)=0$ (so $f(x)=0$); $f(0)=1$, $f(1)=1$ (so $f(x)=1$); $f(0)=0$, $f(1)=1$ (so $f(x)=x$); and $f(0)=1$, $f(1)=0$ (so $f(x)=1+x$). Thus we see that $V=\mathcal{P}_1(\mathbf{Z}_2)$.

8. Let $V=(0,\infty)$, with addition $\oplus$ and scalar multiplication $\odot$ defined by $u\oplus v=uv$ for all $u,v\in V$ and $\alpha\odot u=u^\alpha$ for all $\alpha\in\mathbf{R}$ and all $u\in V$. We will prove that V is a vector space over $\mathbf{R}$. First of all, $\oplus$ is commutative and associative (because multiplication of real numbers has these properties). For all $u\in V$, $u\oplus 1=u\cdot 1=u$, so there is an additive identity. Also, if $u\in V$, then $1/u\in V$ satisfies $u\oplus(1/u)=u(1/u)=1$, so each vector has an additive inverse. Next, if $\alpha,\beta\in\mathbf{R}$ and $u,v\in V$, then $\alpha\odot(\beta\odot u)=\alpha\odot(u^\beta)=(u^\beta)^\alpha=u^{\alpha\beta}=(\alpha\beta)\odot u$, so the associative property of scalar multiplication holds. Also, $\alpha\odot(u\oplus v)=\alpha\odot(uv)=(uv)^\alpha=u^\alpha v^\alpha=(\alpha\odot u)\oplus(\alpha\odot v)$ and $(\alpha+\beta)\odot u=u^{\alpha+\beta}=u^\alpha u^\beta=(\alpha\odot u)\oplus(\beta\odot u)$. Thus both distributive properties hold. Finally, $1\odot u=u^1=u$. This completes the proof that V is a vector space over $\mathbf{R}$.

9. Let $V=\mathbf{R}^2$ with the usual scalar multiplication and the following nonstandard vector addition: $u\oplus v=(u_1+v_1,u_2+v_2+1)$ for all $u,v\in\mathbf{R}^2$. It is easy to check that commutativity and associativity of $\oplus$ hold, that $(0,-1)$ is an additive identity, and that each $u=(u_1,u_2)$ has an additive inverse, namely, $(-u_1,-u_2-2)$. Also, $\alpha(\beta u)=(\alpha\beta)u$ for all $u\in V$, $\alpha,\beta\in\mathbf{R}$ (since scalar multiplication is defined in the standard way). However, if $\alpha\in\mathbf{R}$, then $\alpha(u+v)=\alpha(u_1+v_1,u_2+v_2+1)=(\alpha u_1+\alpha v_1,\alpha u_2+\alpha v_2+\alpha)$, while $\alpha u+\alpha v=(\alpha u_1,\alpha u_2)+(\alpha v_1,\alpha v_2)=(\alpha u_1+\alpha v_1,\alpha u_2+\alpha v_2+1)$, and these are unequal if $\alpha\neq 1$. Thus the first distributive property fails to hold, and V is not a vector space over $\mathbf{R}$. (In fact, the second distributive property also fails.)

10. Let $V=\mathbf{R}^2$ with the usual scalar multiplication, and with addition defined by $u\oplus v=(\alpha_1u_1+\beta_1v_1,\alpha_2u_2+\beta_2v_2)$, where $\alpha_1,\alpha_2,\beta_1,\beta_2\in\mathbf{R}$ are fixed. We wish to determine what values of $\alpha_1,\alpha_2,\beta_1,\beta_2$ will make V a vector space over $\mathbf{R}$. We first note that $\oplus$ is commutative if and only if $\alpha_1=\beta_1$, $\alpha_2=\beta_2$. We therefore redefine $u\oplus v$ as $(\alpha_1u_1+\alpha_1v_1,\alpha_2u_2+\alpha_2v_2)=(\alpha_1(u_1+v_1),\alpha_2(u_2+v_2))$. Next, we have $(u\oplus v)\oplus w=(\alpha_1^2u_1+\alpha_1^2v_1+\alpha_1w_1,\alpha_2^2u_2+\alpha_2^2v_2+\alpha_2w_2)$, $u\oplus(v\oplus w)=(\alpha_1u_1+\alpha_1^2v_1+\alpha_1^2w_1,\alpha_2u_2+\alpha_2^2v_2+\alpha_2^2w_2)$. From this, it is easy to show that $(u\oplus v)\oplus w=u\oplus(v\oplus w)$ for all $u,v,w\in\mathbf{R}$ if and only if $\alpha_1^2=\alpha_1$ and $\alpha_2^2=\alpha_2$, that is, if and only if $\alpha_1=0$ or $\alpha_1=1$, and similarly for α_2. However, if $\alpha_1=0$ or $\alpha_2=0$, then no additive identity can exist. For suppose $\alpha_1=0$. Then $u\oplus v=(0,\alpha_2(u_2+v_2))$ for all $u,v\in\mathcal{V}$, and no $z\in V$ can satisfy $u\oplus z=u$ if $u_1\neq 0$. Similarly, if $\alpha_2=0$, then no additive identity can exist. Therefore, if V is to be a field over $\mathbf{R}^2$, then we must have $\alpha_1=\beta_1=\alpha_2=\beta_2=1$, and V reduces to $\mathbf{R}^2$ under the usual vector space operations.

11. Suppose V is the set of all polynomials (over $\mathbf{R}$) of degree exactly two, together with the zero polynomial. Addition and scalar multiplication are defined on V in the usual fashion. Then V is *not* a vector space over $\mathbf{R}$ because it is not closed under addition. For example, $1+x+x^2\in V$, $1+x-x^2\in V$, but $(1+x+x^2)+(1+x-x^2)=2+2x\notin V$.

12. (a) We wish to find a function lying in $C(0,1)$ but not in $C[0,1]$. A suitable function with a discontinuity at one of the endpoints provides an example. For example, $f(x)=1/x$ satisfies $f\in C(0,1)$ and

$f \notin C[0,1]$, as does $f(x) = 1/(1-x)$ or $f(x) = 1/(x-x^2)$. A different type of example is provided by $f(x) = \sin(1/x)$.

(b) The function $f(x) = |x|$ belongs to $C[-1,1]$ but not to $C^1[-1,1]$.

13. Let V be the space of all infinite sequences of real numbers, and define $\{x_n\} + \{y_n\} = \{x_n + y_n\}$, $\alpha\{x_n\} = \{\alpha x_n\}$. The proof that V is a vector space is a straightforward verification of the defining properties, no different than for $\mathbf{R}^n$, and will not be given here.

14. Let V be the set of all piecewise continuous functions $f : [a,b] \to \mathbf{R}$, with addition and scalar multiplication defined as usual for functions. We wish to show that V is a vector space over $\mathbf{R}$. Most of the properties of a vector space are automatically satisfied by V because it is a subset of the space of all real-valued functions on $[a,b]$, which is known to be a vector space. Specifically, commutativity and associativity of addition, the associative property of scalar multiplication, the two distributive laws, and the fact that $1 \cdot u = u$ for all $u \in V$ are all obviously satisfied. Moreover, the 0 function is continuous and hence by definition piecewise continuous, and therefore $0 \in V$. It remains only to show that V is closed under addition and scalar multiplication (then, since $-u = -1 \cdot u$ for any function u, each function $u \in V$ must have an additive inverse in V). Let $u \in V$, $\alpha \in \mathbf{R}$, and suppose u has points of discontinuity $x_1 < x_2 < \cdots < x_{k-1}$, where $x_1 > x_0 = a$ and $x_{k-1} < x_k = b$. Then u is continuous on each interval (x_{i-1}, x_i), $i = 1, 2, \ldots, k$, and therefore, by a simple theorem of calculus (any multiple of a continuous function is continuous), αu is also continuous on each (x_{i-1}, x_i). The one-sided limits of αu at $x_0, x_1, \ldots, x_k$ exist since, for example,

$$\lim_{x \to x_i^+} \alpha u(x) = \alpha \lim_{x \to x_i^+} u(x)$$

(and similarly for left-hand limits). Therefore, αu is piecewise continuous and therefore $\alpha u \in V$. Now suppose u, v belong to V. Let $\{x_1, x_2, \ldots, x_{\ell-1}\}$ be the union of the sets of points of discontinuity of u and of v, ordered so that $a = x_0 < x_1 < \cdots < x_{\ell-1} = x_\ell = b$. Then, since both u and v are continuous at all other points in (a,b), $u + v$ is continuous on every interval (x_{i-1}, x_i). Also, at each x_i, either $\lim_{x \to x_i} u(x)$ exists (if u is continuous at x_i, that is, if x_i is a point of discontinuity only for v), or the one-sided limits $\lim_{x \to x_i^+} u(x)$ and $\lim_{x \to x_i^-} u(x)$ both exist. In the first case, the two one-sided limits exist (and are equal), so in any case the two one-sided limits exist. The same is true for v. Thus, for each x_i, $i = 0, 1, \ldots, \ell - 1$,

$$\lim_{x \to x_i^+} (u(x) + v(x)) = \lim_{x \to x_i^+} u(x) + \lim_{x \to x_i^+} v(x),$$

and similarly for the left-hand limits at $x_1, x_2, \ldots, x_\ell$. This shows that $u + v$ is piecewise continuous, and therefore belongs to V. This completes the proof.

15. Suppose U and V are vector spaces over a field F, and define addition and scalar multiplication on $U \times V$ by $(u,v) + (w,z) = (u+w, v+z)$, $\alpha(u,v) = (\alpha u, \alpha v)$. We wish to prove that $U \times V$ is a vector space over F. In fact, the verifications of all the defining properties of a vector space are straightforward. For instance, $(u,v) + (w,z) = (u+w, v+z) = (w+u, z+v) = (w,z) + (u,v)$ (using the commutativity of addition in U and V), and therefore addition in $U \times V$ is commutative. Note that the additive identity in $U \times V$ is $(0,0)$, where the first 0 is the zero vector in U and the second is the zero vector in V. We will not verify the remaining properties here.

2.3 Subspaces

1. Let V be a vector space over F.

(a) Let $S = \{0\}$. Then $0 \in S$, S is closed under addition since $0 + 0 = 0 \in S$, and S is closed under scalar multiplication since $\alpha \cdot 0 = 0 \in S$ for all $\alpha \in F$. Thus S is a subspace of V.

(b) The entire space V is a subspace of V since $0 \in V$ and V is closed under addition and scalar multiplication by definition.

2. Suppose we adopt an alternate definition of subspace, in which "$0 \in S$" is replaced with "S is nonempty." We wish to show that the alternate definition is equivalent to the original definition. If S is a subspace according to the original definition, then $0 \in S$, and therefore S is nonempty. Hence S is a subspace according to the alternate definition. Conversely, suppose S satisfies the alternate definition. Then S is nonempty, so there exists $x \in S$. Since S is closed under scalar multiplication and addition, it follows that $-x = -1 \cdot x \in S$, and hence $0 = -x + x \in S$. Therefore, S satisfies the original definition of subspace.

3. Let V be a vector space over $\mathbf{R}$, and let $v \in V$ be nonzero. We wish to prove that $S = \{0, v\}$ is not a subspace of V. If S were a subspace, then $2v$ would lie in S. But $2v \neq 0$ by Theorem 5, and $2v \neq v$ (since otherwise adding $-v$ to both sides would imply that $v = 0$). Hence $2v \notin S$, and therefore S is not a subspace of V.

4. We wish to determine which of the given subsets are subspaces of $\mathbf{Z}_2^3$. Notice that since $\mathbf{Z}_2 = \{0, 1\}$, if S contains the zero vector, then it is automatically closed under scalar multiplication. Therefore, we need only check whether the given subset contains $(0, 0, 0)$ and is closed under addition.

   (a) $S = \{(0,0,0), (1,0,0)\}$. This set contains $(0,0,0)$ and is closed under addition since $(0,0,0) + (0,0,0) = (0,0,0)$, $(0,0,0) + (1,0,0) = (1,0,0)$, and $(1,0,0) + (1,0,0) = (0,0,0)$. Thus S is a subspace.

   (b) $S = \{(0,0,0), (0,1,0), (1,0,1), (1,1,1)\}$. This set contains $(0,0,0)$ and it can be verified that it is closed under addition (for instance, $(0,1,0) + (1,0,1) = (1,1,1)$, $(1,1,1) + (0,1,0) = (1,0,1)$, etc.). Thus S is a subspace.

   (c) $S = \{(0,0,0), (1,0,0), (0,1,0), (1,1,1)\}$ is not a subspace because it is not closed under addition: $(1,0,0) + (0,1,0) = (1,1,0) \notin S$.

5. Suppose S is a subset of $\mathbf{Z}_2^n$. We wish to show that S is a subspace of $\mathbf{Z}_2^n$ if and only if $0 \in S$ and S is closed under addition. Of course, the "only if" direction is trivial. The other direction follows as in the preceding exercise: If $0 \in S$, then S is automatically closed under scalar multiplication, since 0 and 1 are the only elements of the field $\mathbf{Z}_2$, and $0 \cdot v = 0$ for all $v \in S$, $1 \cdot v = v$ for all $v \in S$.

6. Define $S = \{x \in \mathbf{R}^2 : x_1 \geq 0, x_2 \geq 0\}$. Then S is not a subspace of $\mathbf{R}^2$, since it is not closed under scalar multiplication. For instance, $(1,1) \in S$ but $-1 \cdot (1,1) = (-1,-1) \notin S$.

7. Define $S = \{x \in \mathbf{R}^2 : ax_1 + bx_2 = 0\}$, where $a, b \in \mathbf{R}$ are constants. We will show that S is subspace of $\mathbf{R}^2$. First, $(0,0) \in S$, since $a \cdot 0 + b \cdot 0 = 0$. Next, suppose $x \in S$ and $\alpha \in \mathbf{R}$. Then $ax_1 + bx_2 = 0$, and therefore $a(\alpha x_1) + b(\alpha x_2) = \alpha(ax_1 + bx_2) = \alpha \cdot 0 = 0$. This shows that $\alpha x \in S$, and therefore S is closed under scalar multiplication. Finally, suppose $x, y \in S$, so that $ax_1 + bx_2 = 0$ and $ay_1 + by_2 = 0$. Then $a(x_1 + y_1) + b(x_2 + y_2) = (ax_1 + bx_2) + (ay_1 + by_2) = 0 + 0 = 0$, which shows that $x + y \in S$, and therefore that S is closed under addition. This completes the proof.

8. (a) The set $A = \{x \in \mathbf{R}^2 : x_1 = 0 \text{ or } x_2 = 0\}$ is closed under scalar multiplication but not addition. Closure under scalar multiplication holds since if $x_1 = 0$, then $(\alpha x)_1 = \alpha x_1 = \alpha \cdot 0 = 0$, and similarly for the second component. The set is not closed under addition; for instance, $(1,0), (0,1) \in A$, but $(1,0) + (0,1) = (1,1) \notin A$.

   (b) The set $Q = \{x \in \mathbf{R}^2 : x_1 \geq 0, x_2 \geq 0\}$ is closed under addition but not scalar multiplication. Since $(1,1) \in Q$ but $-1 \cdot (1,1) = (-1,-1) \notin Q$, we see that Q is not closed under scalar multiplication. On the other hand, if $x, y \in Q$, so that $x_1, x_2, y_1, y_2 \geq 0$, we see that $(x+y)_1 = x_1 + y_1 \geq 0 + 0 = 0$ and $(x+y)_2 = x_2 + y_2 \geq 0 + 0 = 0$. This shows that $x + y \in Q$, and therefore Q is closed under addition.

9. Let V be a vector space over a field F, let $u \in V$, and define $S = \{\alpha u : \alpha \in F\}$. We will show that S is a subspace of V. First, $0 \in V$ because $0 = 0 \cdot u$. Next, suppose $x \in S$ and $\beta \in F$. Since $x \in S$, there exists $\alpha \in F$ such that $x = \alpha u$. Therefore, $\beta x = \beta(\alpha u) = (\beta\alpha)u$ (using the associative property of scalar multiplication, which shows that βx belongs to S. Thus S is closed under scalar multiplication. Finally, suppose $x, y \in S$; then there exist $\alpha, \beta \in F$ such that $x = \alpha u$, $y = \beta u$, and $x + y = \alpha u + \beta u = (\alpha + \beta)u$

by the second distributive property. Therefore S is closed under addition, and we have shown that S is a subspace.

10. Let $\mathbf{R}$ be regarded as a vector space over $\mathbf{R}$. We wish to prove that $\mathbf{R}$ has no proper subspaces. It suffices to prove that if S is a nontrivial subspace of $\mathbf{R}$, then $S = \mathbf{R}$. So suppose S is a nontrivial subspace, which means that there exists $x \neq 0$ belonging to S. But then, given any $y \in \mathbf{R}$, $y = (yx^{-1})x$ belongs to S because S is closed under scalar multiplication. Thus $\mathbf{R} \subset S$, and hence $S = \mathbf{R}$.

11. We wish to describe all proper subspaces of $\mathbf{R}^2$. We claim that every proper subspace of $\mathbf{R}^2$ has the form $\{\alpha x : \alpha \in \mathbf{R}, \alpha \neq 0\}$, where $x \in \mathbf{R}^2$ is nonzero (geometrically, such a set is a line through the origin). To prove, this, let us suppose S is a proper subspace of $\mathbf{R}^2$. Then there exists $x \in S$, $x \neq 0$. Since S is closed under scalar multiplication, every vector of the form αx, $\alpha \in \mathbf{R}$, must belong to S. Therefore, S contains the set $\{\alpha x : \alpha \in \mathbf{R}, \alpha \neq 0\}$. Let us suppose that there exists $y \in S$ such that y cannot be written as $y = \alpha x$ for some $\alpha \in \mathbf{R}$. In this case, we argue that every $z \in \mathbf{R}^2$ belongs to S, and hence S is not a proper subspace of $\mathbf{R}^2$. To justify this conclusion, we first note that, since y is not a multiple of x, $x_1y_2 - x_2y_1 \neq 0$. Let $z \in \mathbf{R}^2$ be given and consider the equation $\alpha x + \beta y = z$. It can be verified directly that $\alpha = (y_2z_1 - y_1z_2)/(x_1y_2 - x_2y_1)$, $\beta = (x_1z_2 - x_2z_1)/(x_1y_2 - x_2y_1)$ satisfy this equation, from which it follows that $z \in S$ (since S is closed under addition and scalar multiplication). Therefore, if S contains any vector not lying in $\{\alpha x : \alpha \in \mathbf{R}, \alpha \neq 0\}$, then S consists of all of $\mathbf{R}^2$, and S is not a proper subspace of $\mathbf{R}^2$.

12. We wish to find a proper subspace of $\mathbf{R}^3$ that is not a plane. One such subspace is the x_1-axis: $S = \{x \in \mathbf{R}^3 : x_2 = x_3 = 0\}$. It is easy to verify that S is a subspace of $\mathbf{R}^3$, and geometrically, S is a line.

    More generally, using the results of Exercise 10, we can show that $\{\alpha x : \alpha \in \mathbf{R}\}$, where $x \neq 0$ is a given vector, is a proper subspace of $\mathbf{R}^3$. Such a subspace represents a line through the origin.

13. Consider the subset $\mathbf{R}^n$ of $\mathbf{C}^n$. Although $\mathbf{R}^n$ contains the zero vector and is closed under addition, it is not closed under scalar multiplication, and hence is not a subspace of $\mathbf{C}^n$. Here the scalars are complex numbers (since $\mathbf{C}^n$ is a vector space over $\mathbf{C}$), and, for example, $(1, 0, \ldots, 0) \in \mathbf{R}^n$, $i \in \mathbf{C}$, and $i(1, 0, \ldots, 0) = (i, 0, \ldots, 0)$ does not belong to $\mathbf{R}^n$.

14. Let $S = \{u \in C[a, b] : u(a) = u(b) = 0\}$. Then S is a subspace of $C[a, b]$. The zero function clearly belongs to S. Suppose $u \in S$ and $\alpha \in \mathbf{R}$. Then $(\alpha u)(a) = \alpha u(a) = \alpha \cdot 0 = 0$, and similarly $(\alpha u)(b) = 0$. It follows that $\alpha u \in S$, and S is closed under scalar multiplication. If $u, v \in S$, then $(u + v)(a) = u(a) + v(a) = 0 + 0 = 0$, and similarly $(u + v)(b) = 0$. Therefore S is closed under addition, and we have shown that S is a subspace of $C[a, b]$.

15. Let $S = \{u \in C[a, b] : u(a) = 1\}$. Then S is not a subspace of $C[a, b]$ because the zero function does not belong to S.

16. Let $S = \left\{u \in C[a, b] : \int_a^b u(x)\,dx = 0\right\}$. We will show that S is a subspace of $C[a, b]$. First, since the integral of the zero function is zero, we see that the zero function belongs to S. Next, suppose $u \in S$ and $\alpha \in \mathbf{R}$. Then $\int_a^b (\alpha u)(x)\,dx = \int_a^b \alpha u(x)\,dx = \alpha \int_a^b u(x)\,dx = \alpha \cdot 0 = 0$, and therefore $\alpha u \in S$. Finally, suppose $u, v \in S$. Then $\int_a^b (u+v)(x)\,dx = \int_a^b (u(x) + v(x))\,dx = \int_a^b u(x)\,dx + \int_a^b v(x)\,dx = 0 + 0 = 0$. This shows that $u + v \in S$, and we have proved that S is a subspace of $C[a, b]$.

17. Let V be the vector space of all (infinite) sequences of real numbers.

    (a) Define $Z = \{\{x_n\} \in V : \lim_{n\to\infty} x_n = 0\}$. Clearly the zero seqence converges to zero, and hence belongs to Z. If $\{x_n\} \in Z$ and $\alpha \in \mathbf{R}$, then $\lim_{n\to\infty} \alpha x_n = \alpha \lim_{n\to\infty} x_n = \alpha \cdot 0 = 0$, which implies that $\alpha\{x_n\} = \{\alpha x_n\}$ belongs to Z, and therefore Z is closed under scalar multiplication. Now suppose $\{x_n\}$, $\{y_n\}$ both belong to Z. Then $\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n = 0 + 0 = 0$. Therefore $\{x_n\} + \{y_n\} = \{x_n + y_n\}$ belongs to Z, Z is closed under addition, and we have shown that Z is a subspace of V.

(b) Define $S = \{\{x_n\} \in V : \sum_{n=1}^{\infty} x_n < \infty\}$. From calculus, we know that if $\sum_{n=1}^{\infty} x_n$ converges, then so does $\sum_{n=1}^{\infty} \alpha x_n = \alpha \sum_{n=1}^{\infty} x_n$ for any $\alpha \in \mathbf{R}$. Similarly, if $\sum_{n=1}^{\infty} x_n$, $\sum_{n=1}^{\infty} y_n$ converge, then so does $\sum_{n=1}^{\infty}(x_n + y_n) = \sum_{n=1}^{\infty} x_n + \sum_{n=1}^{\infty} y_n$. Using these facts, it is straightforward to show that S is closed under addition and scalar multiplication. Obviously the zero sequence belongs to S.

(c) Define $L = \{\{x_n\} \in V : \sum_{n=1}^{\infty} x_n^2 < \infty\}$. Here is it obvious that the zero sequence belongs to L and that L is closed under scalar multiplication. To prove that L is closed under addition, notice that, for any $x, y \in \mathbf{R}$, $(x-y)^2 \geq 0$ and $(x+y)^2 \geq 0$ together imply that $|xy| \leq (x^2+y^2)/2$. It follows that $(x_n + y_n)^2 = x_n^2 + 2x_n y_n + y_n^2 \leq 2(x_n^2 + y_n^2)$. It follows that if $\sum_{n=1}^{\infty} x_n^2$ and $\sum_{n=1}^{\infty} y_n^2$ both converge, and so does $\sum_{n=1}^{\infty}(x_n+y_n)^2$, with $\sum_{n=1}^{\infty}(x_n+y_n)^2 \leq 2\sum_{n=1}^{\infty} x_n^2 + 2\sum_{n=1}^{\infty} y_n^2$. From this we see that L is closed under addition, and thus L is a subspace.

By a common theorem of calculus, we know that if $\sum_{n=1}^{\infty} x_n$ converges, then $\lim_{n\to\infty} x_n = 0$, and the same is true if $\sum_{n=1}^{\infty} x_n^2$ converges. Therefore, S and L are subspaces of Z. However, the converse of this result is not true (if the sequence converges to zero, this does not imply that the corresponding series converges). Therefore, S and L are proper subspaces of Z. We know that L is not a subspace of S; for instance $\sum_{n=1}^{\infty}(1/n^2)$ converges, but $\sum_{n=1}^{\infty}(1/n)$ does not, which shows that $\{1/n\}$ belongs to L but not to S. Also, S is not a subspace of L, since $\{(-1)^n/\sqrt{n}\}$ belongs to S (by the alternating series test) but not to L.

18. Let V be a vector space over a field F, and let X and Y be subspaces of V.

(a) We will show that $X \cap Y$ is also a subspace of V. First of all, since $0 \in X$ and $0 \in Y$, it follows that $0 \in X \cap Y$. Next, suppose $x \in X \cap Y$ and $\alpha \in F$. Then, by definition of intersection, $x \in X$ and $x \in Y$. Since X and Y are subspaces, both are closed under scalar multiplication and therefore $\alpha x \in X$ and $\alpha x \in Y$, from which it follows that $\alpha \in X \cap Y$. Thus $X \cap Y$ is closed under scalar multiplication. Finally, suppose $x, y \in X \cap Y$. Then $x, y \in X$ and $x, y \in Y$. Since X and Y are closed under addition, we have $x + y \in X$ and $x + y \in Y$, from which we see that $x + y \in X \cap Y$. Therefore, $X \cap Y$ is closed under addition, and we have proved that $X \cap Y$ is a subspace of V.

(b) It is not necessarily that case that $X \cup Y$ is a subspace of V. For instance, let $V = \mathbf{R}^2$, and define $X = \{x \in \mathbf{R}^2 : x_2 = 0\}$, $Y = \{x \in \mathbf{R}^2 : x_1 = 0\}$. Thus $X \cup Y$ is not closed under addition, and hence is not a subspace of $\mathbf{R}^2$. For instance, $(1,0) \in X \subset X \cup Y$ and $(0,1) \in Y \subset X \cup Y$; however, $(1,0) + (0,1) = (1,1) \notin X \cup Y$.

19. Let V be a vector space over a field F, and let S be a nonempty subset of V. Define T to be the intersection of all subspaces of V that contain S.

(a) We wish to show that T is a subspace of V. First, 0 belongs to every subspace of V that contains S, and therefore 0 belongs to the intersection T. Next, suppose $x \in T$ and $\alpha \in F$. Then x belongs to every subspace of V containing S. Since each of these subspaces is closed under scalar multiplication, it follows that αx also belongs to each subspace, and therefore $\alpha x \in T$. Therefore, T is closed under scalar multiplication. Finally, suppose $x, y \in T$. Then both x and y belong to every subspace of V containing S. Since each subspace is closed under addition, it follows that $x + y$ belongs to every subspace of V containing S. Therefore $x + y \in T$, T is closed under addition, and we have shown that T is a subspace.

(b) Now suppose U is any subspace of V containing S. Then U is one of the sets whose intersection defines T, and therefore every element of T belongs to U by definition of intersection. It follows that $T \subset U$. This means that T is the smallest subspace of V containing S.

20. Let V be a vector space over a field F, and let S, T be subspaces of V. Define $S+T = \{s+t : s \in S, t \in T\}$. We wish to show that $S + T$ is a subspace of V. First of all, $0 \in S$ and $0 \in T$ because S and T are subspaces. Therefore, $0 = 0+0 \in S+T$. Next, suppose $x \in S+T$ and $\alpha \in F$. Then, by definition of $S+T$, there exist $s \in S$, $t \in T$ such that $x = s + t$. Since S and T are subspaces, they are closed under scalar multiplication, and therefore $\alpha s \in S$ and $\alpha t \in T$. It follows that $\alpha x = \alpha(s+t) = \alpha s + \alpha t \in S + T$. Thus $S + T$ is closed under scalar multiplication. Finally, suppose $x, y \in S + T$. Then there exist $s_1, s_2 \in S$,

$t_1, t_2 \in T$ such that $x = s_1 + t_1$, $y = s_2 + t_2$. Since S and T are closed under addition, we see that $s_1 + s_2 \in S$, $t_1 + t_2 \in T$, and therefore $x + y = (s_1 + t_1) + (s_2 + t_2) = (s_1 + s_2) + (t_1 + t_2) \in S + T$. It follows that $S + T$ is closed under addition, and we have shown that $S + T$ is a field.

2.4 Linear combinations and spanning sets

1. Write $u_1 = (-1, -2, 4, -2)$, $u_2 = (0, 1, -5, 4)$.

   (a) With $v = (-1, 0, -6, 6)$, the equation $\alpha_1 u_1 + \alpha_2 u_2 = v$ has a (unique) solution: $\alpha_1 = 1$, $\alpha_2 = 2$. This shows that $v \in \text{sp}\{u_1, u_2\}$.

   (b) With $v = (1, 1, 1, 1)$, the equation $\alpha_1 u_1 + \alpha_2 u_2 = v$ has no solution, and therefore $v \notin \text{sp}\{u_1, u_2\}$.

2. Let $S = \text{sp}\{e^x, e^{-x}\} \subset C[0, 1]$.

   (a) The function $f(x) = \cosh(x)$ belongs to S because $\cosh(x) = (1/2)e^x + (1/2)e^{-x}$.

   (b) The function $f(x) = 1$ does not belong to S because there are no scalars α_1, α_2 satisfying $\alpha_1 e^x + \alpha_2 e^{-x} = 1$ for all $x \in [0, 1]$. To prove this, note that any solution α_1, α_2 would have to satisfy the equations that result from substituting any three values of x from the interval $[0, 1]$. For instance, if we choose $x = 0$, $x = 1/2$, $x = 1$, then we obtain the equations

   $$\begin{aligned} \alpha_1 + \alpha_2 &= 1, \\ \alpha_1 e^{1/2} + \alpha_2 e^{-1/2} &= 1, \\ \alpha_1 e + \alpha_2 e^{-1} &= 1. \end{aligned}$$

   A direct calculation shows that this system is inconsistent. Therefore no solution α_1, α_2 exists, and $f \notin S$.

3. Let $S = \text{sp}\{1 + 2x + 3x^2, x - x^2\} \subset \mathcal{P}_2$.

   (a) There is a (unique) solution $\alpha_1 = 2$, $\alpha_2 = 1$ to $\alpha_1(1 + 2x + 3x^2) + \alpha_2(x - x^2) = 2 + 5x + 5x^2$. Therefore, $2 + 5x + 5x^2 \in S$.

   (b) There is no solution α_1, α_2 to $\alpha_1(1 + 2x + 3x^2) + \alpha_2(x - x^2) = 1 - x + x^2$. Therefore, $1 - x + x^2 \notin S$.

4. Let $u_1 = (1 + i, i, 2)$, $u_2 = (1, 2i, 2 - i)$, and define $S = \text{sp}\{u_1, u_2\} \subset \mathbf{C}^3$. The vector $v = (2 + 3i, -2 + 2i, 5 + 2i)$ belongs to S because $2u_1 + iu_2 = v$.

5. Let $S = \text{sp}\{(1, 2, 0, 1), (2, 0, 1, 2)\} \subset \mathbf{Z}_3^4$.

   (a) The vector $(1, 1, 1, 1)$ belongs to S because $2(1, 2, 0, 1) + (2, 0, 1, 2) = (1, 1, 1, 1)$.

   (b) The vector $(1, 0, 1, 1)$ does not belong to S because $\alpha_1(1, 2, 0, 1) + \alpha_2(2, 0, 1, 2) = (1, 0, 1, 1)$ has no solution.

6. Let $S = \text{sp}\{1 + x, x + x^2, 2 + x + x^2\} \subset \mathcal{P}_3(\mathbf{Z}_3)$.

   (a) If $p(x) = 1 + x + x^2$, then $0(1 + x) + 2(x + x^2) + 2(2 + x + x^2) = p(x)$, and therefore $p \in S$.

   (b) Let $q(x) = x^3$. Recalling that $\mathcal{P}_3(\mathbf{Z}_3)$ is a space of polynomials functions, we notice that $q(0) = 0$, $q(1) = 1$, $q(2) = 2$, which means that $q(x) = x$ for all $x \in \mathbf{Z}_3$. We have $1(1 + x) + 2(x + x^2) + 1(2 + x + x^2) = x = q(x)$, and therefore $q \in S$.

7. Let $u = (1, 1, -1)$, $v = (1, 0, 2)$ be vectors in $\mathbf{R}^3$. We wish to show that $S = \text{sp}\{u, v\}$ is a plane in $\mathbf{R}^3$. First note that if $S = \{x \in \mathbf{R}^3 : ax_1 + bx_2 + cx_3 = 0\}$, then (taking $x = u$, $x = v$) we see that a, b, c must satisfy $a + b - c = 0$, $a + 2c = 0$. One solution is $a = 2$, $b = -3$, $c = -1$. We will now prove that $S = \{x \in \mathbf{R}^3 : 2x_1 - 3x_2 - x_3 = 0\}$. First, suppose $x \in S$. Then there exist $\alpha, \beta \in \mathbf{R}$ such that $x = \alpha u + \beta v = \alpha(1, 1, -1) + \beta(1, 0, 2) = (\alpha + \beta, \alpha, -\alpha + 2\beta)$, and $2x_1 - 3x_2 - x_3 = 2(\alpha + \beta) - 3\alpha - (-\alpha + 2\beta) =$

$2\alpha + 2\beta - 3\alpha + \alpha - 2\beta = 0$. Therefore, $x \in \{x \in \mathbf{R}^3 : 2x_1 - 3x_2 - x_3 = 0\}$. Conversely, suppose $x \in \{x \in \mathbf{R}^3 : 2x_1 - 3x_2 - x_3 = 0\}$. If we solve the equation $\alpha u + \beta v = x$, we see that it has the solution $\alpha = x_2$, $\beta = x_1 - x_2$, and therefore $x \in S$. (Notice that $x_2(1,1,-1) + (x_1 - x_2)(1,0,2) = (x_1, x_2, 2x_1 - 3x_2)$, and the assumption $2x_1 - 3x_2 - x_3 = 0$ implies that $2x_1 - 3x_2 = x_3$.) This completes the proof.

8. The previous exercise does not hold true for every choice of $u, v \in \mathbf{R}^3$. For instance, if $u = (1,1,1)$, $v = (2,2,2)$, then $S = \text{sp}\{u, v\}$ is not a plane; in fact, S is easily seen to be the line passing through $(0,0,0)$ and $(1,1,1)$.

9. Let $v_1 = (1,-2,1,2)$, $v_2 = (-1,1,2,1)$, and $v_3 = (-7,9,8,1)$ be vectors in $\mathbf{R}^4$, and let $S = \text{sp}\{v_1, v_2, v_3\}$. Suppose $x \in S$, say $x = \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 = (\beta_1 - \beta_2 - 7\beta_3, -2\beta_1 + \beta_2 + 9\beta_3, \beta_1 + 2\beta_2 + 8\beta_3, 2\beta_1 + \beta_2 + \beta_3)$. The equation $\alpha_1 v_1 + \alpha_2 v_2 = (\beta_1 - \beta_2 - 7\beta_3, -2\beta_1 + \beta_2 + 9\beta_3, \beta_1 + 2\beta_2 + 8\beta_3, 2\beta_1 + \beta_2 + \beta_3)$ has a unique solution, namely, $\alpha_1 = \beta_1 - 2\beta_3$, $\alpha_2 = \beta_2 + 5\beta_3$. This shows that x is a linear combination of v_1, v_2 alone.

   Alternate solution: We can solve $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$ to obtain $\alpha_1 = -2$, $\alpha_2 = 5$, $\alpha_3 = -1$, which means that $-2v_1 + 5v_2 - v_3 = 0$ or $v_3 = -2v_1 + 5v_2$. Now suppose $x \in S$, say $x = \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3$. It follows that $x = \beta_1 v_1 + \beta_2 v_2 + \beta_3(-2v_1 + 5v_2) = (\beta_1 - 2\beta_3)v_1 + (\beta_2 + 5\beta_3)v_2$, and therefore x can be written as a linear combination of v_1 and v_2 alone.

10. Let $u_1 = (1,1,1)$, $u_2 = (1,-1,1)$, $u_3 = (1,0,1)$, and define $S_1 = \text{sp}\{u_1, u_2\}$, $S_2 = \text{sp}\{u_1, u_2, u_3\}$. We wish to prove that $S_1 = S_2$. We first note that if $x \in S_1$, then there exists scalars α_1, α_2 such that $x = \alpha_1 u_1 + \alpha_2 u_2$. But then x can be written as $x = \alpha_1 u_1 + \alpha_2 u_2 + 0 \cdot u_3$, which shows that x is a linear combination of u_1, u_2, u_3, and hence $x \in S_2$. Conversely, suppose that $x \in S_2$, say $x = \beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3 = (\beta_1 + \beta_2 + \beta_3, \beta_1 - \beta_2, \beta_1 + \beta_2 + \beta_3)$. We wish to show that x can be written as a linear combination of u_1, u_2 alone, that is, that there exist scalars α_1, α_2 such that $\alpha_1 u_1 + \alpha_2 u_2 = x$. A direct calculation shows that this equation has a unique solution, namely, $\alpha_1 = \beta_1 + \beta_3/2$, $\alpha_2 = \beta_2 + \beta_3/2$. This shows that $x \in \text{sp}\{u_1, u_2\} = S_1$, and the proof is complete. (The second part of the proof can be done as in the previous solution, by first showing that $u_3 = (1/2)u_1 + (1/2)u_2$.)

11. Let $S = \text{sp}\{(-1,-3,3), (-1,-4,3), (-1,-1,4)\} \subset \mathbf{R}^3$. We wish to determine if $S = \mathbf{R}^3$ or if S is a proper subspace of $\mathbf{R}^3$. Given an arbitrary $x \in \mathbf{R}^3$, we solve $\alpha_1(-1,-1,3) + \alpha_2(-1,-4,3) + \alpha_3(-1,-1,4) = (x_1, x_2, x_3)$ and find that there is a unique solution, namely, $\alpha_1 = -13x_1 + x_2 - 3x_3$, $\alpha_2 = 9x_1 - x_2 + 2x_3$, $\alpha_3 = 3x_1 + x_3$. This shows that every $x \in \mathbf{R}^3$ lies in S, and therefore $S = \mathbf{R}^3$.

12. Let $S = \text{sp}\{(-1,-5,1), (3,14,-4), (1,4,-2)\}$. Given an arbitary $x \in \mathbf{R}^3$, if we try to solve $\alpha_1(-1,-5,1) + \alpha_2(3,14,-4) + \alpha_3(1,4,-2) = (x_1, x_2, x_3)$, we find that there is a solution if and only if $6x_1 - x_2 + x_3 = 0$. Since not all $x \in \mathbf{R}^3$ satisfy this condition, S is a proper subspace of $\mathbf{R}^3$.

13. Let $S = \text{sp}\{1 - x, 2 - 2x + x^2, 1 - 3x^2\} \subset \mathcal{P}_2$. We wish to determine if S is a proper subspace of $\mathcal{P}_2$. Given any $p \in \mathcal{P}_2$, say $p(x) = c_0 + c_1 x + c_2 x^2$, we try to solve $\alpha_1(1 - x) + \alpha_2(2 - 2x + x^2) + \alpha_3(1 - 3x^2) = c_0 + c_1 x + c_2 x^2$. We find that there is a unique solution, $\alpha_1 = -6c_0 - 7c_1 - 2c_2$, $\alpha_2 = 3c_0 + 3c_1 + c_2$, $\alpha_3 = c_0 + c_1$. Therefore, each $p \in \mathcal{P}_2$ belongs to S, and therefore $S = \mathcal{P}_2$.

14. Suppose V is a vector space over a field F and S is a subspace of V. We wish to prove that $u_1, \ldots, u_k \in S$, $\alpha_1, \ldots, \alpha_k \in F$ imply that $\alpha_1 u_1 + \cdots + \alpha_k u_k \in S$. We argue by induction on k. For $k = 1$, we have that $\alpha_1 u_1 \in S$ because S is a subspace and therefore closed under scalar multiplication. Now suppose that, for some $k \geq 2$, $\alpha_1 u_1 + \cdots + \alpha_{k-1} u_{k-1} \in S$ for any $u_1, \ldots, u_{k-1} \in S$, $\alpha_1, \ldots, \alpha_{k-1} \in F$. Let $u_1, \ldots, u_k \in S$, $\alpha_1, \ldots, \alpha_k \in F$ be arbitrary. Then $\alpha_1 u_1 + \ldots + \alpha_k u_k = (\alpha_1 u_1 + \cdots + \alpha_{k-1} u_{k-1}) + \alpha_k u_k$. By the induction hypothesis, $\alpha_1 u_1 + \cdots + \alpha_{k-1} u_{k-1} \in S$, and $\alpha_k u_k \in S$ because S is closed under scalar multiplication. But then $\alpha_1 u_1 + \ldots + \alpha_k u_k = (\alpha_1 u_1 + \cdots + \alpha_{k-1} u_{k-1}) + \alpha_k u_k \in S$ because S is closed under addition. Therefore, by induction, the result holds for all $k \geq 1$, and the proof is complete.

15. Let V be a vector space over a field F, and let $u \in V$, $u \neq 0$, $\alpha \in F$. We wish to prove that $\text{sp}\{u\} = \text{sp}\{u, \alpha u\}$. First, if $x \in \text{sp}\{u\}$, then $x = \beta u$ for some $\beta \in F$, in which case we can write $x = \beta u + 0(\alpha u)$, which shows that x also belongs to $\text{sp}\{u, \alpha u\}$. Conversely, if $x \in \text{sp}\{u, \alpha u\}$, then there exist scalars $\beta, \gamma \in F$ such that $x = \beta u + \gamma(\alpha u)$. But then $x = (\beta + \gamma\alpha)u$, and therefore $x \in \text{sp}\{u\}$. Thus $\text{sp}\{u\} = \text{sp}\{u, \alpha u\}$.

16. Let V be a vector space over a field F, and suppose $x, u_1, \ldots, u_k, v_1, \ldots, v_\ell$ are vectors in V. Assume $x \in \text{sp}\{u_1, \ldots, u_k\}$ and $u_j \in \text{sp}\{v_1, \ldots, v_\ell\}$ for $j = 1, \ldots, k$. We wish to show that $x \in \text{sp}\{v_1, \ldots, v_\ell\}$. Since $u_j \in \text{sp}\{v_1, \ldots, v_\ell\}$, there exist scalars $\beta_{j,1}, \ldots, \beta_{j,\ell}$ such that $u_j = \beta_{j,1}v_1 + \cdots + \beta_{j,\ell}v_\ell$. This is true for each u_j, $j = 1, \ldots, k$. Also, $x \in \text{sp}\{u_1, \ldots, u_k\}$, so there exist $\alpha_1, \ldots, \alpha_k \in F$ such that $x = \alpha_1 u_1 + \cdots + \alpha_k u_k$. It follows that

$$\begin{aligned} x &= \alpha_1(\beta_{1,1}v_1 + \cdots + \beta_{1,\ell}v_\ell) + \alpha_2(\beta_{2,1}v_1 + \cdots + \beta_{2,\ell}v_\ell) + \cdots + \alpha_k(\beta_{k,1}v_1 + \cdots + \beta_{k,\ell}v_\ell) \\ &= \alpha_1\beta_{1,1}v_1 + \cdots + \alpha_1\beta_{1,\ell}v_\ell + \alpha_2\beta_{2,1}v_1 + \cdots + \alpha_2\beta_{2,\ell}v_\ell + \cdots + \alpha_k\beta_{k,1}v_1 + \cdots + \alpha_k\beta_{k,\ell}v_\ell \\ &= (\alpha_1\beta_{1,1} + \alpha_2\beta_{2,1} + \cdots + \alpha_k\beta_{k,1})v_1 + (\alpha_1\beta_{1,2} + \alpha_2\beta_{2,2} + \cdots + \alpha_k\beta_{k,2})v_2 + \cdots + \\ &\quad (\alpha_1\beta_{1,\ell} + \alpha_2\beta_{2,\ell} + \cdots + \alpha_k\beta_{k,\ell})v_\ell. \end{aligned}$$

This shows that $x \in \text{sp}\{v_1, \ldots, v_\ell\}$.

17. (a) Let V be a vector space over $\mathbf{R}$, and let u, v be any two vectors in V. We wish to prove that $\text{sp}\{u, v\} = \text{sp}\{u+v, u-v\}$. We first suppose that $x \in \text{sp}\{u+v, u-v\}$, say $x = \alpha(u+v) + \beta(u-v)$. Then $x = \alpha u + \alpha v + \beta u - \beta v = (\alpha + \beta)u + (\alpha - \beta)v$, which shows that $x \in \text{sp}\{u, v\}$. Conversely, suppose that $x \in \text{sp}\{u, v\}$, say $x = \alpha u + \beta v$. We notice that $u = (1/2)(u+v) + (1/2)(u-v)$, and $v = (1/2)(u+v) - (1/2)(u-v)$. Therefore, $x = \alpha((1/2)(u+v) + (1/2)(u-v)) + \beta((1/2)(u+v) - (1/2)(u-v)) = (\alpha/2 + \beta/2)(u+v) + (\alpha/2 - \beta/2)(u-v)$, which shows that $x \in \text{sp}\{u+v, u-v\}$. Therefore, $x \in \text{sp}\{u, v\}$ if and only if $x \in \text{sp}\{u+v, u-v\}$, and hence the two subspaces are equal.

    (b) The result just proved does not necessarily hold if V is a vector space over an arbitrary field F. More specifically, the first part of the proof is always valid, and therefore $\text{sp}\{u+v, u-v\} \subset \text{sp}\{u, v\}$ always holds. However, it is not always possible to write u and v in terms of $u+v$ and $u-v$, and therefore $\text{sp}\{u, v\} \subset \text{sp}\{u+v, u-v\}$ need not hold. For example, if $F = \mathbf{Z}_2$, $V = \mathbf{Z}_2^2$, $u = (1, 0)$, $v = (0, 1)$, then we have $u + v = (1, 1)$ and $u - v = (1, 1)$ (since $-1 = 1$ in $\mathbf{Z}_2$). It follows that $\text{sp}\{u+v, u-v\} = \{(0, 0), (1, 1)\}$, and hence $u, v \notin \text{sp}\{u+v, u-v\}$, which in turn means that $\text{sp}\{u, v\} \not\subset \text{sp}\{u+v, u-v\}$.

2.5 Linear independence

1. Let V be a vector space over a field F, and let $u_1, u_2 \in V$. We wish to prove that $\{u_1, u_2\}$ is linearly dependent if and only if one of these vectors is a multiple of the other. Suppose first that $\{u, v\}$ is linearly dependent. Then there exist scalars α_1, α_2, not both zero, such that $\alpha_1 u_1 + \alpha_2 u_2 = 0$. Suppose $\alpha_1 \neq 0$; then α_1^{-1} exists, and we have $\alpha_1 u_1 + \alpha_2 u_2 = 0 \Rightarrow \alpha_1 u_1 = -\alpha_2 u_2 \Rightarrow u_1 = -\alpha_1^{-1}\alpha_2 u_2$. Therefore, in this case, u_1 is a multiple of u_2. Similarly, if $\alpha_2 \neq 0$, we can show that u_2 is a multiple of u_1. Conversely, suppose one of u_1, u_2 is a multiple of the other, say $u_1 = \alpha u_2$. We can then write $u_1 - \alpha u_2 = 0$, or $1 \cdot u_1 + (-\alpha)u_2 = 0$, which, since $1 \neq 0$, shows that $\{u_1, u_2\}$ is linearly dependent. A similar proof shows that if u_2 is a multiple of u_1, then $\{u_1, u_2\}$ is linearly dependent. This completes the proof.

2. Let V be a vector space over a field F, and suppose $v \in V$. We wish to prove that $\{v\}$ is linearly independent if and only if $v \neq 0$. First, if $v \neq 0$, then $\alpha v = 0$ implies that $\alpha = 0$ by Theorem 5. It follows that $\{v\}$ is linearly independent if $v \neq 0$. On the other hand, if $v = 0$, then $1 \cdot v = 0$, which shows that $\{v\}$ is linearly dependent (there is a nontrivial solution to $\alpha v = 0$). Thus $\{v\}$ is linearly independent if and only if $v \neq 0$.

3. Let V be a vector space over a field F, and let $u_1, \ldots, u_n \in V$. Suppose $u_i = 0$ for some i, $1 \leq i \leq n$, and define scalars $\alpha_1, \ldots, \alpha_n \in F$ by $\alpha_k = 0$ if $k \neq i$, $\alpha_i = 1$. Then $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0 \cdot u_1 + \cdots + 0 \cdot u_{i-1} + 1 \cdot 0 + 0 \cdot u_{i+1} + \cdots + 0 \cdot u_n = 0$, and hence there is a nontrivial solution to $\alpha_1 u_1 + \cdots + \alpha_n u_n = 0$. This shows that $\{u_1, \ldots, u_n\}$ is linearly dependent.

4. Let V be a vector space over a field F, let $\{u_1, \ldots, u_k\}$ be a linearly independent subspace of V, and assume $v \in V$, $v \notin \text{sp}\{u_1, \ldots, u_k\}$. We wish to show that $\{u_1, \ldots, u_k, v\}$ is also linearly independent. We argue by contradiction and assume that $\{u_1, \ldots, u_k, v\}$ is linearly dependent. Then there exist scalars

$\alpha_1, \ldots, \alpha_k, \beta$, not all zero, such that $\alpha_1 u_1 + \cdots + \alpha_k u_k + \beta v = 0$. We now consider two cases. First, if $\beta = 0$, then not all of $\alpha_1, \ldots, \alpha_k$ are zero, and we see that $\alpha_1 u_1 + \cdots + \alpha_k u_k = \alpha_1 u_1 + \cdots + \alpha_k u_k + 0 \cdot v = 0$. This contradicts the fact that $\{u_1, \ldots, u_k\}$ is linearly independent. Second, if $\beta \neq 0$, then we can solve $\alpha_1 u_1 + \cdots + \alpha_k u_k + \beta v = 0$ to obtain $v = -\beta^{-1}\alpha_1 u_1 - \cdots - \beta^{-1}\alpha_k u_k$, which contradicts that $v \notin \text{sp}\{u_1, \ldots, u_k\}$. Thus, in either case, we obtain a contradiction, and the proof is complete.

5. We wish to determine whether each of the following sets is linearly independent or not.

   (a) The set $\{(1, 2), (1, -1)\} \subset \mathbf{R}^2$ is linearly independent by Exercise 1, since neither vector is a multiple of the other.

   (b) The set $\{(-1, -1, 4), (-4, -4, 17), (1, 1, -3)\}$ is linearly dependent. Solving

   $$\alpha_1(-1, -1, 4) + \alpha_2(-4, -4, 17) + \alpha_3(1, 1, -3) = 0$$

   shows that $\alpha_1 = 5, \alpha_2 = -1, \alpha_3 = 1$ is a nontrivial solution.

   (c) The set $\{(-1, 3, -2), (3, -10, 7), (-1, 3, -1)\}$ is linearly independent. Solving

   $$\alpha_1(-1, 3, -2) + \alpha_2(3, -10, 7) + \alpha_3(-1, 3, -1) = 0$$

   shows that the only solution is $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

6. We wish to determine whether each of the following sets of polynomials is linearly independent or not.

   (a) The set $\{1 - x^2, x + x^2, 3 + 3x - 4x^2\} \subset \mathcal{P}_2$ is linearly independent since the only solution to $\alpha_1(1 - x^2) + \alpha_2(x + x^2) + \alpha_3(3 + 3x - 4x^2) = 0$ is $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

   (b) The set $\{1 + x^2, 4 + 3x^2 + 3x^3, 3 - x + 10x^3, 1 + 7x^2 - 18x^3\} \subset \mathcal{P}_3$ is linearly dependent. Solving $\alpha_1(1 + x^2) + \alpha_2(4 + 3x^2 + 3x^3) + \alpha_3(3 - x + 10x^3) + \alpha_4(1 + 7x^2 - 18x^3) = 0$ yields a nontrivial solution $\alpha_1 = -25$, $\alpha_2 = 6$, $\alpha_3 = 0$, $\alpha_4 = 1$.

7. The set $\{e^x, e^{-x}, \cosh(x)\} \subset C[0, 1]$ is linearly dependent since $(1/2)e^x + (1/2)e^{-x} - \cosh(x) = 0$ for all $x \in [0, 1]$.

8. The subset $\{(0, 1, 2), (1, 2, 0), (2, 0, 1)\}$ of $\mathbf{Z}_3^3$ is linearly dependent because $1 \cdot (0, 1, 2) + 1 \cdot (1, 2, 0) + 1 \cdot (2, 0, 1) = (0, 0, 0)$.

9. We wish to show that $\{1, x, x^2\}$ is linearly dependent in $\mathcal{P}_2(\mathbf{Z}_2)$. The equation $\alpha_1 \cdot 1 + \alpha_2 x + \alpha_3 x^2 = 0$ has the nontrivial solution $\alpha_1 = 0, \alpha_2 = 1, \alpha_3 = 1$. To verify this, we must simply verify that $x + x^2$ is the zero function in $\mathcal{P}_2(\mathbf{Z}_2)$. Substituting $x = 0$, we obtain $0 + 0^2 = 0 + 0 = 0$, and with $x = 1$, we obtain $1 + 1^2 = 1 + 1 = 0$.

10. The set $\{(i, 1, 2i), (1, 1+i, i), (1, 3+5i, -4+3i)\} \subset \mathbf{C}^3$ is linearly dependent, because $\alpha_1(i, 1, 2i) + \alpha_2(1, 1+i, i) + \alpha_3(1, 3+5i, -4+3i) = (0, 0, 0)$ has the nontrivial solution $\alpha_1 = -2i, \alpha_2 = -3, \alpha_3 = 1$.

11. We have already seen that $\{(3, 2, 2, 3), (3, 2, 1, 2), (3, 2, 0, 1)\} \subset \mathbf{R}^4$ is linearly dependent, because $(3, 2, 2, 3) - 2(3, 2, 1, 2) + (3, 2, 0, 1) = (0, 0, 0, 0)$.

    (a) We can solve this equation for any one of the vectors in terms of the other two; for instance, $(3, 2, 2, 3) = 2(3, 2, 1, 2) - (3, 2, 0, 1)$.

    (b) We can show that $(-3, -2, 2, 1) \in \text{sp}\{(3, 2, 2, 3), (3, 2, 1, 2), (3, 2, 0, 1)\}$ by solving $\alpha_1(3, 2, 2, 3) + \alpha_2(3, 2, 1, 2) + \alpha_3(3, 2, 0, 1) = (-3, -2, 2, 1)$. One solution is $(-3, -2, 2, 1) = 3(3, 2, 2, 3) - 4(3, 2, 1, 2)$. Substituting $(3, 2, 2, 3) = 2(3, 2, 1, 2) - (3, 2, 0, 1)$, we obtain another solution: $(-3, -2, 2, 1) = 2(3, 2, 1, 2) - 3(3, 2, 0, 1)$.

12. We wish to show that $\{(-1,1,3),(1,-1,-2),(-3,3,13)\} \subset \mathbf{R}^3$ is linearly dependent by writing one of the vectors as a linear combination of the others. We will try to solve for the third vector in terms of the other two. (There is an element of trial and error involved here: Even if the three vectors form a linearly independent set, there is no guarantee that this will work; it could be, for instance, that the first two vectors form a linearly dependent set and the third vector does not lie in the span of the first two.) Solving $\alpha_1(-1,1,3)+\alpha_2(1,-1,-2) = (-3,3,13)$ yields a unique solution: $(-3,3,13) = 7(-1,1,3)+4(1,-1,-2)$. This shows that the set is linearly dependent.

    Alternate solution: We begin by solving $\alpha_1(-1,1,3) + \alpha_2(1,-1,-2) + \alpha_3(-3,3,13) = (0,0,0)$ to obtain $7(-1,1,3)+4(1,-1,-2)-(-3,3,13) = (0,0,0)$. We can easily solve this for the third vector: $(-3,3,13) = 7(-1,1,3) + 4(1,-1,-2)$.

13. We wish to show that $\{p_1,p_2,p_3\}$, where $p_1(x) = 1-x^2$, $p_2(x) = 1+x-6x^2$, $p_3(x) = 3-2x^2$, is linearly independent and spans $\mathcal{P}_2$. We first verify that the set is linearly independent by solving $\alpha_1(1-x^2)+\alpha_2(1+x-6x^2)+\alpha_3(3-2x^2) = 0$. This equation is equivalent to the system $\alpha_1+\alpha_2+3\alpha_2 = 0$, $\alpha_2 = 0$, $-\alpha_1-6\alpha_2-2\alpha_3 = 0$, and a direct calculation shows that the only solution is $\alpha_1 = \alpha_2 = \alpha_3 = 0$. To show that the set spans $\mathcal{P}_2$, we take an arbitrary $p \in \mathcal{P}_2$, say $p(x) = c_0 + c_1x + c_2x^2$, and solve $\alpha_1(1-x^2)+\alpha_2(1+x-6x^2)+\alpha_3(3-2x^2) = c_0+c_1x+c_2x^2$. This is equivalent to the system $\alpha_1+\alpha_2+3\alpha_2 = c_0$, $\alpha_2 = c_1$, $-\alpha_1-6\alpha_2-2\alpha_3 = c_2$. There is a unique solution: $\alpha_1 = -2c_0-16c_1-3c_2$, $\alpha_2 = c_1$, $\alpha_3 = c_0+5c_1+c_2$. This shows that $p \in \text{sp}\{p_1,p_2,p_3\}$, and, since p was arbitrary, that $\{p_1,p_2,p_3\}$ spans all of $\mathcal{P}_2$.

14. Let V be a vector space over a field F and let $\{u_1,\ldots,u_k\}$ be a linearly independent subset of V. Suppose $u,v \in V$, $\{u,v\}$ is linearly independent, and $u,v \notin \text{sp}\{u_1,\ldots,u_k\}$. We wish to determine whether $\{u_1,\ldots,u_k,u,v\}$ is necessarily linearly independent. In fact, this set need not be linearly independent. For example, take $V = \mathbf{R}^4$, $F = \mathbf{R}$, $k = 3$, and $u_1 = (1,0,0,0)$, $u_2 = (0,1,0,0)$, $u_3 = (0,0,1,0)$. With $u = (0,0,0,1)$, $v = (1,1,1,1)$, we see immediately that $\{u,v\}$ is linearly independent (neither vector is a multiple of the other), and that neither u nor v belongs to $\text{sp}\{u_1,u_2,u_3\}$. Nevertheless, $\{u_1,u_2,u_3,u,v\}$ is linear dependent because $v = u_1+u_2+u_3+u$.

15. Let V be a vector space over a field F, and suppose S and T are subspaces of V satisfying $S \cap T = \{0\}$. Suppose $\{s_1,\ldots,s_k\} \subset S$ and $\{t_1,\ldots,t_\ell\} \subset T$ are both linearly independent sets. We wish to prove that $\{s_1,\ldots,s_k,t_1,\ldots,t_\ell\}$ is linearly independent. Suppose scalars $\alpha_1,\ldots,\alpha_k$, $\beta_1,\ldots,\beta_\ell$ satisfy $\alpha_1s_1+\cdots+\alpha_ks_k+\beta_1t_1+\cdots+\beta_\ell t_\ell = 0$. We can rearrange this equation to read $\alpha_1s_1+\cdots+\alpha_ks_k = -\beta_1t_1-\cdots-\beta_\ell t_\ell$. If v is the vector represented by these two expressions, then $v \in S$ (since v is a linear combination of $s_1,\ldots,s_k$) and $v \in T$ (since v is a linear combination of $t_1,\ldots,t_\ell$). But the only vector in $S \cap T$ is the zero vector, and hence $\alpha_1s_1+\cdots+\alpha_ks_k = 0$, $-\beta_1t_1-\cdots-\beta_\ell t_\ell = 0$. The first equation implies that $\alpha_1 = \cdots = \alpha_k = 0$ (since $\{s_1,\ldots,s_k\}$ is linearly independent), while the second equation implies that $\beta_1 = \cdots = \beta_\ell = 0$ (since $\{t_1,\ldots,t_\ell\}$ is linearly independent). Therefore, $\alpha_1s_1+\cdots+\alpha_ks_k+\beta_1t_1+\cdots+\beta_\ell t_\ell = 0$ implies that all the scalars are zero, and hence $\{s_1,\ldots,s_k,t_1,\ldots,t_\ell\}$ is linearly independent.

16. Let V be a vector space over a field F, and let $\{u_1,\ldots,u_k\}$ and $\{v_1,\ldots,v_\ell\}$ be two linearly independent subsets of V. We wish to find a condition that implies that $\{u_1,\ldots,u_k,v_1,\ldots,v_\ell\}$ is linearly independent. By the previous exercise, a sufficient condition for $\{u_1,\ldots,u_k,v_1,\ldots,v_\ell\}$ to be linearly independent is that $S = \text{sp}\{u_1,\ldots,u_k\}$, $T = \text{sp}\{v_1,\ldots,v_\ell\}$ satisfy $S \cap T = \{0\}$. We will prove that this condition is also necessary. Suppose $\{u_1,\ldots,u_k\}$ and $\{v_1,\ldots,v_\ell\}$ are linearly independent subsets of V, and that $\{u_1,\ldots,u_k,v_1,\ldots,v_\ell\}$ is also linearly independent. Define S and T as above. If $x \in S \cap T$, then there exist scalars $\alpha_1,\ldots,\alpha_k \in F$ such that $x = \alpha_1u_1+\cdots+\alpha_ku_k$ (since $x \in S$), and also scalars $\beta_1,\ldots,\beta_\ell \in F$ such that $x = \beta_1v_1+\cdots+\beta_\ell v_\ell$ (since $x \in T$). But then $\alpha_1u_1+\cdots+\alpha_ku_k = \beta_1v_1+\cdots+\beta_\ell v_\ell$, which implies that $\alpha_1u_1+\cdots+\alpha_ku_k-\beta_1v_1-\cdots-\beta_\ell v_\ell = 0$. Since $\{u_1,\ldots,u_k,v_1,\ldots,v_\ell\}$ is linearly independent by assumption, this implies that $\alpha_1 = \cdots = \alpha_k = \beta_1 = \cdots = \beta_\ell = 0$, which in turn shows that $x = 0$. Therefore $S \cap T = \{0\}$, and the proof is complete.

17. (a) Let V be a vector space over $\mathbf{R}$, and suppose $\{x,y,z\}$ is a linearly independent subset of V. We wish to show that $\{x+y,y+z,x+z\}$ is also linearly independent. Let $\alpha_1,\alpha_2,\alpha_3 \in \mathbf{R}$ satisfy $\alpha_1(x+y)+$

$\alpha_2(y+z)+\alpha_3(x+z)=0$. This equation is equivalent to $(\alpha_1+\alpha_3)x+(\alpha_1+\alpha_2)y+(\alpha_2+\alpha_3)z=0$. Since $\{x,y,z\}$ is linearly independent, it follows that $\alpha_1+\alpha_3=\alpha_1+\alpha_2=\alpha_2+\alpha_3=0$. This system can be solved directly to show that $\alpha_1=\alpha_2=\alpha_3=0$, which proves that $\{x+y,y+z,x+z\}$ is linearly independent.

(b) We now show, by example, that the previous result is not necessarily true if V is a vector space over some field $F\neq\mathbf{R}$. Let $V=\mathbf{Z}_2^3$, and define $x=(1,0,0)$, $y=(0,1,0)$, and $z=(0,0,1)$. Obviously $\{x,y,z\}$ is linearly independent. On the other hand, we have $(x+y)+(y+z)+(x+z)=(1,1,0)+(0,1,1)+(1,0,1)=(1+0+1,1+1+0,0+1+1)=(0,0,0)$, which shows that $\{x+y,y+z,x+z\}$ is linearly dependent.

18. Let U and V be vector spaces over a field F, and define $W=U\times V$. Suppose $\{u_1,\ldots,u_k\}\subset U$ and $\{v_1,\ldots,v_\ell\}\subset V$ are linearly independent. We wish to show that $\{(u_1,0),\ldots,(u_k,0),(0,v_1),\ldots,(0,v_\ell)\}$ is also linearly independent. Suppose $\alpha_1,\ldots,\alpha_k,\beta_1,\ldots,\beta_\ell\in F$ satisfy $\alpha_1(u_1,0)+\cdots+\alpha_k(u_k,0)+\beta_1(0,v_1)+\cdots+\beta_\ell(0,v_\ell)=(0,0)$. This reduces to $(\alpha_1u_1+\cdots+\alpha_ku_k,\beta_1v_1+\cdots+\beta_\ell v_\ell)=(0,0)$, which holds if and only if $\alpha_1u_1+\cdots+\alpha_ku_k=0$ and $\beta_1v_1+\cdots+\beta_\ell v_\ell=0$. Since $\{u_1,\ldots,u_k\}$ is linearly independent, the first equation implies that $\alpha_1=\cdots=\alpha_k=0$, and, since $\{v_1,\ldots,v_\ell\}$ is linearly independent, the second implies that $\beta_1=\cdots=\beta_\ell=0$. Since all the scalars are necessarily zero, we see that $\{(u_1,0),\ldots,(u_k,0),(0,v_1),\ldots,(0,v_\ell)\}$ is linearly independent.

19. Let V be a vector space over a field F, and let $u_1,u_2,\ldots,u_n$ be vectors in V. Suppose a nonempty subset of $\{u_1,u_2,\ldots,u_n\}$, say $\{u_{i_1},\ldots,u_{i_k}\}$, is linearly dependent. (Here $1\leq k<n$ and $i_1,\ldots,i_k$ are distinct integers each satisfying $1\leq i_j\leq n$.) We wish to prove that $\{u_1,u_2,\ldots,u_n\}$ itself is linearly dependent. By assumption, there exist scalars $\alpha_{i_1},\ldots,\alpha_{i_k}\in F$, not all zero, such that $\alpha_{i_1}u_{i_1}+\cdots+\alpha_{i_k}u_{i_k}=0$. For each $i\in\{1,\ldots,n\}\setminus\{i_1,\ldots,i_k\}$, define $\alpha_i=0$. Then we have $\alpha_1u_1+\cdots+\alpha_nu_n=0+\alpha_{i_1}u_{i_1}+\cdots+\alpha_{i_k}u_{i_k}=0$, and not all of $\alpha_1,\ldots,\alpha_n$ are zero since at least one α_{i_j} is nonzero. This shows that $\{u_1,\ldots,u_n\}$ is linearly dependent.

20. Let V be a vector space over a field F, and suppose $\{u_1,u_2,\ldots,u_n\}$ is a linearly independent subset of V. We wish to prove that every nonempty subset of $\{u_1,u_2,\ldots,u_n\}$ is also linearly independent. The result to be proved is simply the contrapositive of the statement in the previous exercise, and therefore holds by the previous proof.

21. Let V be a vector space over a field F, and suppose $\{u_1,u_2,\ldots,u_n\}$ is linearly dependent. We wish to prove that, given any i, $1\leq i\leq n$, either u_i is a linear combination of $u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_n$ or these vectors form a linearly dependent set. By assumption, there exist scalars $\alpha_1,\ldots,\alpha_n\in F$, not all zero, such that $\alpha_1u_1+\cdots+\alpha_iu_i+\cdots+\alpha_nu_n=0$. We now consider two cases. If $\alpha_i\neq0$, the we can solve the latter equation for u_i to obtain $u_i=-\alpha_i^{-1}\alpha_1u_1-\cdots-\alpha_i^{-1}\alpha_{i-1}u_{i-1}-\alpha_i^{-1}\alpha_{i+1}u_{i+1}-\cdots-\alpha_i^{-1}\alpha_nu_n$. In this case, u_i is a linear combination of the remaining vectors. The second case is that $\alpha_i=0$, in which case at least one of $\alpha_1,\ldots,\alpha_{i-1},\alpha_{i+1},\ldots,\alpha_n$ is nonzero, and we have $\alpha_1u_1+\cdots+\alpha_{i-1}u_{i-1}+\alpha_{i+1}u_{i+1}+\cdots+\alpha_nu_n=0$. This shows that $\{u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_n\}$ is linearly dependent.

2.6 Basis and dimension

1. Suppose $\{v_1,v_2,\ldots,v_n\}$ is a basis for a vector space V.

(a) We wish to show that if any v_j is removed from the basis, the resulting set of $n-1$ vectors does not span V and hence is not a basis. This follows from Theorem 24: Since $\{v_1,v_2,\ldots,v_n\}$ is linearly independent, no v_j, $j=1,2,\ldots,n$, can be written as a linear combination of the remaining vectors. Therefore,

$$v_j\notin\operatorname{sp}\{v_1,\ldots,v_{j-1},v_{j+1},\ldots,v_n\},$$

which proves the desired result.

(b) Now we wish to show that if any vector $u\in V$, $u\notin\{v_1,v_2,\ldots,v_n\}$, is added to the basis, the resulting set of $n+1$ vectors is linearly dependent. This is immediate from Theorem 34: Since

the dimension of V is n, every set containing more than n vectors is linearly dependent. Since $\{v_1, v_2, \ldots, v_n, u\}$ contains $n+1$ vectors, it must be linearly dependent.

2. Consider the following vectors in $\mathbf{R}^3$: $v_1 = (-1, 4, -2)$, $v_2 = (5, -20, 9)$, $v_3 = (2, -7, 6)$. We wish to determine if $\{v_1, v_2, v_3\}$ is a basis for $\mathbf{R}^3$. If we solve $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = x$ for an arbitrary $x \in \mathbf{R}^3$, we find a unique solution: $\alpha_1 = 57x_1 + 12x_2 - 5x_3$, $\alpha_2 = 10x_1 + 2x_2 - x_3$ $\alpha_3 = 4x_1 + x_2$. By Theorem 28, this implies that $\{v_1, v_2, v_3\}$ is a basis for $\mathbf{R}^3$.

3. We now repeat the previous exercise for the vectors $v_1 = (-1, 3, -1)$, $v_2 = (1, -2, -2)$, $v_3 = (-1, 7, -13)$. If we try to solve $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = x$ for an arbitrary $x \in \mathbf{R}^3$, we find that this equation is equivalent to the following system:

$$\begin{aligned} -\alpha_1 + \alpha_2 - \alpha_3 &= x_1 \\ \alpha_2 + 4\alpha_3 &= 3x_1 + x_2 \\ 0 &= 8x_1 + 3x_2 + x_3. \end{aligned}$$

Since this system is inconsistent for most $x \in \mathbf{R}^3$ (the system is consistent only if x happens to satisfy $8x_1 + 3x_2 + x_3 = 0$), $\{v_1, v_2, v_3\}$ does not span $\mathbf{R}^3$ and therefore is not a basis.

4. Let $S = \text{sp}\{e^x, e^{-x}\}$ be regarded as a subspace of $C(\mathbf{R})$. We will show that $\{e^x, e^{-x}\}$, $\{\cosh(x), \sinh(x)\}$ are two different bases for S. First, to verify that $\{e^x, e^{-x}\}$ is a basis, we merely need to verify that it is linearly independent (since it spans S by definition). This can be done as follows: If $c_1 e^x + c_2 e^{-x} = 0$, where 0 represents the zero function, then the equation must hold for all values of $x \in \mathbf{R}$. So choose $x = 0$ and $x = \ln 2$; then c_1 and c_2 must satisfy

$$\begin{aligned} c_1 + c_2 &= 0, \\ 2c_1 + \frac{1}{2}c_2 &= 0. \end{aligned}$$

It is straightforward to show that the only solution of this system is $c_1 = c_2 = 0$, and hence $\{e^x, e^{-x}\}$ is linearly independent.

Next, since

$$\cosh(x) = \frac{1}{2}e^x + \frac{1}{2}e^{-x}, \ \cosh(x) = \frac{1}{2}e^x - \frac{1}{2}e^{-x},$$

we see that $\cosh(x), \sinh(x) \in S = \text{sp}\{e^x, e^{-x}\}$. We can verify that $\{\cosh(x), \sinh(x)\}$ is linearly independent directly: If $c_1 \cosh(x) + c_2 \sinh(x) = 0$, then, substituting $x = 0$ and $x = \ln 2$, we obtain the system

$$1 \cdot c_1 + 0 \cdot c_2 = 0, \ \frac{5}{4}c_1 + \frac{3}{4}c_2 = 0,$$

and the only solution is $c_1 = c_2 = 0$. Thus $\{\cosh(x), \sinh(x)\}$ is linearly independent. Finally, let f be any function in S. Then, by definition, f can be written as $f(x) = \alpha_1 e^x + \alpha_2 e^{-x}$ for some $\alpha_1, \alpha_2 \in \mathbf{R}$. We must show that $f \in \text{sp}\{\cosh(x), \sinh(x)\}$, that is, that there exist $c_1, c_2 \in \mathbf{R}$ such that

$$c_1 \cosh(x) + c_2 \sinh(x) = f(x).$$

This equation can be manipulated as follows:

$$\begin{aligned} & c_1 \cosh(x) + c_2 \sinh(x) = f(x) \\ \Leftrightarrow\ & c_1 \left(\frac{1}{2}e^x + \frac{1}{2}e^{-x}\right) + c_2 \left(\frac{1}{2}e^x - \frac{1}{2}e^{-x}\right) = \alpha_1 e^x + \alpha_2 e^{-x} \\ \Leftrightarrow\ & \left(\frac{1}{2}c_1 + \frac{1}{2}c_2\right) e^x + \left(\frac{1}{2}c_1 - \frac{1}{2}c_2\right) e^{-x} = \alpha_1 e^x + \alpha_2 e^{-x}. \end{aligned}$$

Since $\{\{e^x, e^{-x}\}$ is linearly independent, Theorem 26 implies that the last equation can hold only if

$$\frac{1}{2}c_1 + \frac{1}{2}c_2 = \alpha_1,\ \frac{1}{2}c_1 - \frac{1}{2}c_2\alpha_2.$$

This last system has a unique solution:

$$c_1 = \alpha_1 + \alpha_2,\ c_2 = \alpha_1 - \alpha_2.$$

This shows that $f \in \text{sp}\{\cosh(x), \sinh(x)\}$, and the proof is complete.

5. Let $p_1(x) = 1 - 4x + 4x^2$, $p_2(x) = x + x^2$, $p_3(x) = -2 + 11x - 6x^2$. We will determine whether $\{p_1, p_2, p_3\}$ is a basis for $\mathcal{P}_2$ or not by solving $\alpha_1 p_1(x) + \alpha_2 p_2(x) + \alpha_3 p_3(x) = c_0 + c_1 x + c_2 x^2$, where $c_0 + c_1 x + c_2 x^2$ is an arbitrary element of $\mathcal{P}_2$. A direct calculation shows that there is a unique solution for α_1, α_2, α_3:

$$\alpha_1 = 17c_0 + 2c_1 - 2c_2,\ \alpha_2 = 3c_2 - 2c_1 - 20c_0,\ \alpha_3 = 8c_0 + c_1 - c_2.$$

By Theorem 28, it follows that $\{p_1, p_2, p_3\}$ is a basis for $\mathcal{P}_2$.

6. Let $p_1(x) = 1 - x^2$, $p_2(x) = 2 + x$, $p_3(x) = x + 2x^2$. We will determine whether $\{p_1, p_2, p_3\}$ a basis for $\mathcal{P}_2$ by trying to solve

$$\alpha_1 p_1(x) + \alpha_2 p_2(x) + \alpha_3 p_3(x) = c_0 + c_1 x + c_2 x^2,$$

where c_0, c_1, c_2 are arbitrary real numbers. This equation is equivalent to the system

$$\begin{aligned} \alpha_1 + 2\alpha_2 &= c_0, \\ \alpha_2 + \alpha_3 &= c_1, \\ -\alpha_1 + 2\alpha_3 &= c_2. \end{aligned}$$

Solving this system by elimination leads to

$$\begin{aligned} \alpha_1 + 2\alpha_2 &= c_0, \\ \alpha_2 + \alpha_3 &= c_1, \\ 0 &= c_0 - 2c_1 + c_2, \end{aligned}$$

which is inconsistent for most values of c_0, c_1, c_2. Therefore, $\{p_1, p_2, p_3\}$ does not span $\mathcal{P}_2$ and hence is not a basis for $\mathcal{P}_2$.

7. Consider the subspace $S = \text{sp}\{p_1, p_2, p_3, p_4, p_5\}$ of $\mathcal{P}_3$, where

$$\begin{aligned} &p_1(x) = -1 + 4x - x^2 + 3x^3,\ p_2(x) = 2 - 8x + 2x^2 - 5x^3, \\ &p_3(x) = 3 - 11x + 3x^2 - 8x^3,\ p_4(x) = -2 + 8x - 2x^2 - 3x^3, \\ &p_5(x) = 2 - 8x + 2x^2 + 3x^3. \end{aligned}$$

   (a) The set $\{p_1, p_2, p_3, p_4, p_5\}$ is linearly dependent (by Theorem 34) because it contains five elements and the dimension of $\mathcal{P}_3$ is only four.

   (b) As illustrated in Example 39, we begin by solving

$$\alpha_1 p_1(x) + \alpha_2 p_2(x) + \alpha_3 p_3(x) + \alpha_4 p_4(x) + \alpha_5 p_5(x) = 0;$$

   this is equivalent to the system

$$\begin{aligned} -\alpha_1 + 2\alpha_2 + 3\alpha_3 - 2\alpha_4 + 2\alpha_5 &= 0, \\ 4\alpha_1 - 8\alpha_2 - 11\alpha_3 + 8\alpha_4 - 8\alpha_5 &= 0, \\ -\alpha_1 + 2\alpha_2 + 3\alpha_3 - 2\alpha_4 + 2\alpha_5 &= 0, \\ 3\alpha_1 - 5\alpha_2 - 8\alpha_3 - 3\alpha_4 + 3\alpha_5 &= 0, \end{aligned}$$

which reduces to

$$\begin{aligned}\alpha_1 &= 16\alpha_4 - 16\alpha_5,\\ \alpha_2 &= 9\alpha_4 - 9\alpha_5,\\ \alpha_3 &= 0.\end{aligned}$$

Since there are nontrivial solutions, $\{p_1, p_2, p_3, p_4, p_5\}$ is linearly dependent (which we already knew), but we can deduce more than that. By taking $\alpha_4 = 1$, $\alpha_5 = 0$, we see that $\alpha_1 = 16$, $\alpha_2 = 9$, $\alpha_3 = 0$, $\alpha_4 = 1$, $\alpha_5 = 0$ is one solution, which means that

$$16p_1(x) + 9p_2(x) + p_4(x) = 0 \;\Rightarrow\; p_4(x) = -16p_1(x) - 9p_2(x).$$

This shows that $p_4 \in \text{sp}\{p_1, p_2\} \subset \text{sp}\{p_1, p_2, p_3\}$. Similarly, taking $\alpha_4 = 0$, $\alpha_5 = 0$, we find that

$$-16p_1(x) - 9p_2(x) + p_5(x) = 0 \;\Rightarrow\; p_5(x) = 16p_1(x) + 9p_2(x),$$

and hence $p_5 \in \text{sp}\{p_1, p_2\} \subset \text{sp}\{p_1, p_2, p_3\}$. It follows from Lemma 19 that $\text{sp}\{p_1, p_2, p_3, p_4, p_5\} = \text{sp}\{p_1, p_2, p_3\}$. Our calculations above show that $\{p_1, p_2, p_3\}$ is linearly independent (if $\alpha_4 = \alpha_5 = 0$, then also $\alpha_1 = \alpha_2 = \alpha_3 = 0$). Therefore, $\{p_1, p_2, p_3\}$ is a linearly independent spanning set of S and hence a basis for S.

8. We wish to find a basis for $\text{sp}\{(1,2,1),(0,1,1),(1,1,0)\} \subset \mathbf{R}^3$. We will name the vectors v_1, v_2, v_3, respectively, and begin by testing the linear independence of $\{v_1, v_2, v_3\}$. The equation $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$ is equivalent to

$$\begin{aligned}\alpha_1 + \alpha_3 &= 0,\\ 2\alpha_2 + \alpha_2 + \alpha_3 &= 0,\\ \alpha_1 + \alpha_2 &= 0,\end{aligned}$$

which reduces to

$$\alpha_1 = -\alpha_3,\ \alpha_2 = \alpha_3.$$

One solution is $\alpha_1 = -1$, $\alpha_2 = 1$, $\alpha_3 = 1$, which shows that $-v_1 + v_2 + v_3 = 0$, or $v_3 = v_1 - v_2$. This in turns shows that $\text{sp}\{v_1, v_2, v_3\} = \text{sp}\{v_1, v_2\}$ (by Lemma 19). Clearly $\{v_1, v_2\}$ is linearly independent (since neither vector is a multiple of the other), and hence $\{v_1, v_2\}$ is a basis for $\text{sp}\{v_1, v_2, v_3\}$.

9. We wish to find a basis for $S = \text{sp}\{(1,2,1,2,1),(1,1,2,2,1),(0,1,2,0,2)\}$ in $\mathbf{Z}_3^5$. The equation

$$\alpha_1(1,2,1,2,1) + \alpha_2(1,1,2,2,1) + \alpha_3(0,1,2,0,2) = (0,0,0,0,0)$$

is equivalent to the system

$$\begin{aligned}\alpha_1 + \alpha_2 &= 0,\\ 2\alpha_1 + \alpha_2 + \alpha_3 &= 0,\\ \alpha_1 + 2\alpha_2 + 2\alpha_3 &= 0,\\ 2\alpha_1 + 2\alpha_2 &= 0,\\ \alpha_1 + \alpha_2 + 2\alpha_3 &= 0.\end{aligned}$$

Reducing this system by Gaussian elimination (in modulo 3 arithmetic), we obtain

$$\alpha_1 = \alpha_2 = \alpha_3 = 0,$$

which shows that the given vectors form a linearly independent set and therefore a basis for S.

10. We will show that $\{1+x+x^2, 1-x+x^2, 1+x+2x^2\}$ is a basis for $\mathcal{P}_2(\mathbf{Z}_3)$ by showing that there is a unique solution to

$$\alpha_1\left(1+x+x^2\right)+\alpha_2\left(1-x+x^2\right)+\alpha_3\left(1+x+2x^2\right)=c_0+c_1x+c_2x^2.$$

We first note that $1-x+x^2=1+2x+x^2$ in $\mathcal{P}_2(\mathbf{Z}_3)$, so we can write our equation as

$$\alpha_1\left(1+x+x^2\right)+\alpha_2\left(1+2x+x^2\right)+\alpha_3\left(1+x+2x^2\right)=c_0+c_1x+c_2x^2.$$

We rearrange the previous equation in the form

$$(\alpha_1+\alpha_2+\alpha_3)+(\alpha_1+2\alpha_2+\alpha_3)x+(\alpha_1+\alpha_2+2\alpha_3)x^2=c_0+c_1x+c_2x^2.$$

Since the polynomials involved are of degree 2 and the field $\mathbf{Z}_3$ contains 3 elements, this last equation is is equivalent to the system

$$\begin{aligned}\alpha_1+\alpha_2+\alpha_3&=c_0,\\ \alpha_1+2\alpha_2+\alpha_3&=c_1,\\ \alpha_1+\alpha_2+2\alpha_3&=c_2\end{aligned}$$

(cf. the discussion on page 45 of the text). Applying Gaussian elimination (modulo 3) shows that there is a unique solution:

$$\begin{aligned}\alpha_1&=c_0+2c_1+c_2,\\ \alpha_2&=2c_0+c_1,\\ \alpha_3&=2c_0+c_2.\end{aligned}$$

This in turn proves (by Theorem 28) that the given polynomials form a basis for $\mathcal{P}_2(\mathbf{Z}_3)$.

11. Suppose F is a finite field with q distinct elements.

    (a) Assume $n \leq q-1$. We wish to show that $\{1, x, x^2, \ldots, x^n\}$ is a linearly independent subset of $\mathcal{P}_n(F)$. (Since $\{1, x, x^2, \ldots, x^n\}$ clearly spans $\mathcal{P}_n(F)$, this will show that it is a basis for $\mathcal{P}_n(F)$, and hence that $\dim(\mathcal{P}_n(F)) = n+1$ in the case that $n \leq q-1$.) The desired conclusion follows from the discussion on page 45 of the text. If $c_0 \cdot 1 + c_1x + \cdots + c_nx^n = 0$ (where 0 is the zero function), then every element of F is a root of $c_0 \cdot 1 + c_1x + \cdots + c_nx^n$. Since F contains more than n elements and a nonzero polynomial of degree n can have at most n distinct roots, this is impossible unless $c_0 = c_1 = \ldots = c_n = 0$. Thus $\{1, x, \ldots, x^n\}$ is linearly independent.

    (b) Now suppose that $n \geq q$. The reasoning above shows that $\{1, x, x^2, \ldots, x^{q-1}\}$ is linearly independent in $\mathcal{P}_n(F)$ ($c_0 \cdot 1 + c_1x + \cdots + c_{q-1}x^{q-1}$ has at most $q-1$ distinct roots, and F contains more than $q-1$ elements, etc.). This implies that $\dim(\mathcal{P}_n(F)) \geq q$ in the case $n \geq q$.

12. Suppose V is a vector space over a field F, and S, T are two n-dimensional subspaces of V. We wish to prove that if $S \subset T$, then in fact $S = T$. Let $\{s_1, s_2, \ldots, s_n\}$ be a basis for S. Since $S \subset T$, this implies that $\{s_1, s_2, \ldots, s_n\}$ is a linearly independent subset of T. We will now show that $\{s_1, s_2, \ldots, s_n\}$ also spans T. Let $t \in T$ be arbitrary. Since T has dimension n, the set $\{s_1, s_2, \ldots, s_n, t\}$ is linearly dependent by Theorem 34. But then, by Lemma 33, t must be a linear combination of $s_1, s_2, \ldots, s_n$ (since no s_k is a linear combination of $s_1, s_2, \ldots, s_{k-1}$). This shows that $t \in \mathrm{sp}\{s_1, s_2, \ldots, s_n\}$, and hence we have shown that $\{s_1, s_2, \ldots, s_n\}$ is a basis for T. But then

$$T = \mathrm{sp}\{s_1, s_2, \ldots, s_n\} = S,$$

as desired.

13. Suppose V is a vector space over a field F, and S, T are two finite-dimensional subspaces of V with $S \subset T$. We are asked to prove that $\dim(S) \leq \dim(T)$. Let $\{s_1, s_2, \ldots, s_n\}$ be a basis for S. Since $S \subset T$, it follows that $\{s_1, s_2, \ldots, s_n\}$ is a linearly independent subset of T, and hence, by Theorem 34, any basis for T must have at least n vectors. It follows that $\dim(T) \geq n = \dim(S)$.

14. **(Note: This exercise belongs in Section 2.7 since the most natural solution uses Theorem 43.)** Let V be a vector space over a field F, and let S and T be finite-dimensional subspaces of V. We wish to prove that

$$\dim(S+T) = \dim(S) + \dim(T) - \dim(S \cap T).$$

We know from Exercise 2.3.19 that $S \cap T$ is a subspace of V, and since it is a subset of S, $\dim(S \cap T) \leq \dim(S)$. Since S is finite-dimensional by assumption, it follows that $S \cap T$ is also finite-dimensional, and therefore either $S \cap T = \{0\}$ or $S \cap T$ has a basis.

Suppose first that $S \cap T = \{0\}$, so that $\dim(S \cap T) = 0$. Let $\{s_1, s_2, \ldots, s_m\}$ be a basis for S and $\{t_1, t_2, \ldots, t_n\}$ be a basis for T. We will show that $\{s_1, \ldots, s_m, t_1, \ldots, t_n\}$ is a basis for $S+T$, from which it follows that

$$\dim(S+T) = m + n = m + n - 0 = \dim(S) + \dim(T) - \dim(S \cap T).$$

The set $\{s_1, \ldots, s_m, t_1, \ldots, t_n\}$ is linearly independent by Exercise 2.5.15. Given any $v \in S+T$, there exist $s \in S$, $t \in T$ such that $v = s + t$. But since $s \in S$, there exist scalars $\alpha_1, \ldots, \alpha_m \in F$ such that $s = \alpha_1 s_1 + \cdots + \alpha_m s_m$. Similarly, since $t \in T$, there exist $\beta_1, \ldots, \beta_n \in F$ such that $t = \beta_1 t_1 + \cdots + \beta_n t_n$. But then

$$v = s + t = \alpha_1 s_1 + \cdots + \alpha_m s_m + \beta_1 t_1 + \cdots + \beta_n t_n,$$

which shows that $v \in \text{sp}\{s_1, \ldots, s_m, t_1, \ldots, t_n\}$. Thus we have shown that $\{s_1, \ldots, s_m, t_1, \ldots, t_n\}$ is a basis for $S+T$, which completes the proof in the case that $S \cap T = \{0\}$.

Now suppose $S \cap T$ is nontrivial, with basis $\{v_1, \ldots, v_k\}$. Since $S \cap T$ is a subset of S, $\{v_1, \ldots, v_k\}$ is a linearly independent subset of S and hence, by Theorem 43, can be extended to a basis $\{v_1, \ldots, v_k, s_1, \ldots, s_p\}$ of S. Similarly, $\{v_1, \ldots, v_k\}$ can be extended to a basis $\{v_1, \ldots, v_k, t_1, \ldots, t_q\}$ of T. We will show that $\{v_1, \ldots, v_k, s_1, \ldots, s_p, t_1, \ldots, t_q\}$ is a basis of $S+T$. Then we will have

$$\dim(S) = k + p, \ \dim(T) = k + q, \ \dim(S \cap T) = k$$

and

$$\dim(S+T) = k + p + q = (k+p) + (k+q) - k = \dim(S) + \dim(T) - \dim(S \cap T),$$

as desired. First, suppose $v \in S+T$. Then, by definition of $S+T$, there exist $s \in S$ and $t \in T$ such that $v = s + t$. Since $\{v_1, \ldots, v_k, s_1, \ldots, s_p\}$ is a basis for S, there exist scalars $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_p \in F$ such that

$$s = \alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 s_1 + \cdots + \beta_p s_p.$$

Similarly, there exist $\gamma_1, \ldots, \gamma_k, \delta_1, \ldots, \delta_q \in F$ such that

$$t = \gamma_1 v_1 + \cdots + \gamma_k v_k + \delta_1 s_1 + \cdots + \delta_q s_q.$$

But then

$$\begin{aligned} v = s + t &= \alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 s_1 + \cdots + \beta_p s_p + \\ &\quad \gamma_1 v_1 + \cdots + \gamma_k v_k + \delta_1 s_1 + \cdots + \delta_q s_q \\ &= (\alpha_1 + \gamma_1) v_1 + \cdots + (\alpha_k + \gamma_k) v_k + \beta_1 s_1 + \cdots + \beta_p s_p + \\ &\quad \delta_1 t_1 + \cdots + \delta_q t_q \\ &\in \text{sp}\{v_1, \ldots, v_k, s_1, \ldots, s_p, t_1, \ldots, t_q\}. \end{aligned}$$

This shows that $\{v_1, \ldots, v_k, s_1, \ldots, s_p, t_1, \ldots, t_q\}$ spans $S+T$. Now suppose $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_p, \gamma_1, \ldots, \gamma_q \in F$ satisfy

$$\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 s_1 + \cdots + \beta_p s_p + \gamma_1 t_1 + \cdots + \gamma_q t_q = 0.$$

This implies that

$$\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 s_1 + \cdots + \beta_p s_p = -\gamma_1 t_1 - \cdots - \gamma_q t_q.$$

The vector on the left belongs to S, while the vector on the right belongs to T; hence both vectors (which are really the same) belong to $S \cap T$. But then $-\gamma_1 t_1 - \cdots - \gamma_q t_q$ can be written in terms of the basis $\{v_1, \ldots, v_k\}$ of $S \cap T$, say

$$-\gamma_1 t_1 - \cdots - \gamma_q t_q = \delta_1 v_1 + \cdots + \delta_k v_k.$$

But this gives two representations of the vector $-\gamma_1 t_1 - \cdots - \gamma_q t_q \in T$ in terms of the basis

$$\{v_1, \ldots, v_k, t_1, \ldots, t_q\}.$$

Since each vector in T must be uniquely represented as a linear combination of the basis vectors, this is possible only if $\gamma_1 = \cdots = \gamma_q = \delta_1 = \cdots = \delta_k = 0$. But then

$$\alpha_1 v_1 + \cdots + \alpha_k v_k + \beta_1 s_1 + \cdots + \beta_p s_p = 0,$$

and the linear independence of $\{v_1, \ldots, v_k, \sigma_1, \ldots, s_p\}$ implies that $\alpha_1 = \cdots = \alpha_k = \beta_1 = \cdots = \beta_p = 0$. We have thus shown that

$$\{v_1, \ldots, v_k, s_1, \ldots, s_p, t_1, \ldots, t_q\}$$

is linearly independent, which completes the proof.

15. Let V be a vector space over a field F, and let S and T be finite-dimensional subspaces of V. Consider the four subspaces

$$X_1 = S,\ X_2 = T,\ X_3 = S + T,\ X_4 = S \cap T.$$

For every choice of i, j with

$$1 \leq i < j \leq 4,$$

we wish to determine if $\dim(X_i) \leq \dim(X_j)$ or $\dim(X_i) \geq \dim(X_j)$ (or neither) must hold. First of all, since S and T are arbitrary subspaces, it is obvious that there need be no particular relationship between the dimensions of S and T. However, $S \subset S + T$ since each $s \in S$ can be written as $s = s + 0 \in S + T$ ($0 \in T$ because every subspace contains the zero vector). Therefore, by Exercise 13, $\dim(S) \leq \dim(S+T)$. By the same reasoning, $T \subset S + T$ and hence $\dim(T) \leq \dim(S + T)$. Next, $S \cap T \subset S$, $S \cap T \subset T$, and hence $\dim(S \cap T) \leq \dim(S)$, $\dim(S \cap T) \leq \dim(T)$. Finally, we have $S \cap T \subset S \subset S + T$, and hence $\dim(S \cap T) \leq \dim(S + T)$.

16. Let V be a vector space over a field F, and suppose S and T are subspaces of V satisfying $S \cap T = \{0\}$. Suppose $\{s_1, s_2, \ldots, s_k\} \subset S$ and $\{t_1, t_2, \ldots, t_\ell\} \subset T$ are bases for S and T, respectively. We wish to prove that

$$\{s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_\ell\}$$

is a basis for $S + T$. This was done in the course of proving the result in Exercise 14.

17. Let U and V be vector spaces over a field F, and let $\{u_1, \ldots, u_n\}$ and $\{v_1, \ldots, v_m\}$ be bases for U and V, respectively. We are asked to prove that

$$\{(u_1, 0), \ldots, (u_n, 0), (0, v_1), \ldots, (0, v_m)\}$$

is a basis for $U \times V$. First, let (u, v) be an arbitrary vector in $U \times V$. Then $u \in U$ and there exist $\alpha_1, \ldots, \alpha_n \in F$ such that $u = \alpha_1 u_1 + \cdots + \alpha_n u_n$. Similarly, $v \in V$ and there exist $\beta_1, \ldots, \beta_m \in F$ such that $v = \beta_1 v_1 + \cdots + \beta_m v_m$. It follows that

$$\begin{aligned}(u, v) &= (u, 0) + (0, v) \\ &= (\alpha_1 u_1 + \cdots + \alpha_n u_n, 0) + (0, \beta_1 v_1 + \cdots + \beta_m v_m) \\ &= \alpha_1 (u_1, 0) + \cdots \alpha_n (u_n, 0) + \beta_1 (0, v_1) + \cdots + \beta_m (0, v_m).\end{aligned}$$

This shows that $\{(u_1, 0), \ldots, (u_n, 0), (0, v_1), \ldots, (0, v_m)\}$ spans $U \times V$. Next, suppose $\alpha_1, \ldots, \alpha_n \in F$, $\beta_1, \ldots, \beta_m \in F$ satisfy

$$\alpha_1 (u_1, 0) + \cdots \alpha_n (u_n, 0) + \beta_1 (0, v_1) + \cdots + \beta_m (0, v_m) = 0.$$

Since the zero vector $\in U \times V$ is $(0,0)$, this yields

$$(\alpha_1 u_1 + \cdots + \alpha_n u_n, 0) + (0, \beta_1 v_1 + \cdots + \beta_m v_m) = (0,0),$$

which is equivalent to

$$\alpha_1 u_1 + \cdots + \alpha_n u_n = 0,\ \beta_1 v_1 + \cdots + \beta_m v_m = 0.$$

Since both $\{u_1, \ldots, u_n\}$ and $\{v_1, \ldots, v_m\}$ are linearly independent, it follows that $\alpha_1 = \cdots = \alpha_n = \beta_1 = \cdots = \beta_m = 0$. This shows that $\{(u_1, 0), \ldots, (u_n, 0), (0, v_1), \ldots, (0, v_m)\}$ is linearly independent, and the proof is complete.

18. We will prove that the number of elements in a finite field must be p^n, where p is a prime number and n is a positive integer.

Let F be a finite field.

(a) Let p be the characteristic of F. Then

$$0, 1, 1+1, 1+1+1, \ldots, 1+1+\cdots+1$$

($p-1$ terms in the last sum) are distinct elements of F, while $1+1+\cdots+1$ (p terms) is 0. We will write $2 = 1+1$, $3 = 1+1+1$, and so forth, thus labeling p distinct elements of F, namely, $0, 1, \ldots, p-1$. We can then show that $\{0, 1, 2, \ldots, p-1\} \subset F$ is a subfield of F isomorphic to $\mathbf{Z}_p$. Writing out a formal proof is difficult, because the symbols $0, 1, 2, \ldots, p-1$ have now three different meanings (they are elements of $\mathbf{Z}$, elements of $\mathbf{Z}_p$, and now elements of F). For the purposes of this proof, we will temporarily write $0_F, 1_F, \ldots, (p-1)_F$ for the elements of F, $0_{\mathbf{Z}_p}, 1_{\mathbf{Z}_p}, \ldots, (p-1)_{\mathbf{Z}_p}$ for the elements of $\mathbf{Z}_p$, and $0, 1, \ldots, p-1$ for the elements of $\mathbf{Z}$. Let us define $G = \{0_F, 1_F, \ldots, (p-1)_F\}$ and $\phi : G \to \mathbf{Z}_p$ by $\phi(k_F) = k_{\mathbf{Z}_p}$. If $k + \ell < p$, then $k_F + \ell_F$ is the sum of $k+\ell$ copies of 1_F, which is $(k+\ell)_F$ by definition. Similarly, $k_{\mathbf{Z}_p} + \ell_{\mathbf{Z}_p} = (k+\ell)_{\mathbf{Z}_p}$ by definition of addition in $\mathbf{Z}_p$. Therefore,

$$\phi(k_F + \ell_F) = \phi((k+\ell)_F) = (k+\ell)_{\mathbf{Z}_p} = k_{\mathbf{Z}_p} + \ell_{\mathbf{Z}_p} = \phi(k_F) + \phi(\ell_F).$$

On the other hand, if $0 \le k, \ell \le p-1$ and $k + \ell \ge p$, then $k_F + \ell_F$ is the sum of $k+\ell$ copies of 1_F, which can be written (by the associative property of addition) as the sum of p copies of 1_F plus the sum of $k+\ell-p$ copies of 1_F. This reduces to $0_F + (k+\ell-p)_F = (k+\ell-p)_F$. Similarly, by the definition of addition in $\mathbf{Z}_p$, $k_{\mathbf{Z}_p} + \ell_{\mathbf{Z}_p} = (k+\ell-p)_{\mathbf{Z}_p}$, and therefore, in this case also, we see

$$\phi(k_F + \ell_F) = \phi((k+\ell-p)_F) = (k+\ell-p)_{\mathbf{Z}_p} = k_{\mathbf{Z}_p} + \ell_{\mathbf{Z}_p} = \phi(k_F) + \phi(\ell_F).$$

Therefore, ϕ preserves addition.

Now, by the distributive law, $k_F \ell_F$ can be written as the sum of $k\ell$ copies of 1_F (a careful proof of this would require induction). If $k\ell = qp + r$, where $q \ge 0$ and $0 \le r \le p-1$, then, by the associative property of addition, we can write $k_F \ell_F$ as the sum of $q+1$ sums, the first q of them consisting of p copies of 1_F (and thus each equalling 0_F) and the last consisting of r copies of 1_F. It follows that $k_F \ell_F = r_F$. By definition of multiplication in $\mathbf{Z}_p$, we simiarly have $k_{\mathbf{Z}_p} \ell_{\mathbf{Z}_p} = r_{\mathbf{Z}_p}$, and hence

$$\phi(k_F \ell_F) = \phi(r_F) = r_{\mathbf{Z}_p} = k_{\mathbf{Z}_p} \ell_{\mathbf{Z}_p} = \phi(k_F)\phi(\ell_F).$$

Therefore, ϕ also preserves multiplication, and we have shown that G and $\mathbf{Z}_p$ are isomorphic as fields.

(b) We now drop the subscripts and identify $\mathbf{Z}_p$ with the subfield G of F. We wish to show that that F is a vector space over $\mathbf{Z}_p$. We already know that addition in F is commutative, associative, and has an identity, and that each element of F has an additive inverse in F. The associative property of scalar multiplication and the two distributive properties of scalar multiplication reduce to the associative property of multiplication and the distributive property in F. Finally, $1 \cdot u = u$ for all $u \in F$ since 1 is nothing more than the multiplicative identity in F. This verifies that F is a vector field over $\mathbf{Z}_p$.

(c) Since F has only a finite number of elements, it must be a finite-dimensional vector space over $\mathbf{Z}_p$. Let the dimension be n, and let $\{f_1, \ldots, f_n\}$ be a basis for F. Then every element of F can be written uniquely in the form

$$\alpha_1 f_1 + \alpha_2 f_2 + \cdots + \alpha_n f_n, \tag{2.1}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbf{Z}_p$. Conversely, for each choice of $\alpha_1, \alpha_2, \ldots, \alpha_n$ in $\mathbf{Z}_p$, (2.1) defines an element of F. Therefore, the number of elements of F is precisely the number of different ways to choose $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbf{Z}_p$, which is p^n.

2.7 Properties of bases

1. Consider the following vectors in $\mathbf{R}^3$: $v_1 = (1, 5, 4)$, $v_2 = (1, 5, 3)$, $v_3 = (17, 85, 56)$, $v_4 = (1, 5, 2)$, $v_5 = (3, 16, 13)$.

   (a) We wish to show that $\{v_1, v_2, v_3, v_4, v_5\}$ spans $\mathbf{R}^3$. Given an arbitrary $x \in \mathbf{R}^3$, the equation

   $$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4 + \alpha_5 v_5 = x$$

   is equivalent to the system

   $$\begin{aligned} \alpha_1 + \alpha_2 + 17\alpha_3 + \alpha_4 + 3\alpha_5 &= x_1, \\ 5\alpha_1 + 5\alpha_2 + 85\alpha_3 + 5\alpha_4 + 16\alpha_5 &= x_2, \\ 4\alpha_1 + 3\alpha_2 + 56\alpha_3 + 2\alpha_4 + 13\alpha_5 &= x_3. \end{aligned}$$

   Applying Gaussian elimination, this system reduces to

   $$\begin{aligned} \alpha_1 &= 17x_1 - 4x_2 + x_3 - 5\alpha_3 + \alpha_4, \\ \alpha_2 &= x_2 - x_1 - x_3 - 12\alpha_3 - 2\alpha_5, \\ \alpha_5 &= x_2 - 5x_1. \end{aligned}$$

   This shows that there are solutions regardless of the value of x; that is, each $x \in \mathbf{R}^3$ can be written as a linear combination of v_1, v_2, v_3, v_4, v_5. Therefore, $\{v_1, v_2, v_3, v_4, v_5\}$ spans $\mathbf{R}^3$.

   (b) Now we wish to find a subset of $\{v_1, v_2, v_3, v_4, v_5\}$ that is a basis for $\mathbf{R}^3$. According to the calculations given above, each $x \in \mathbf{R}^3$ can be written as a linear combination of $\{v_1, v_2, v_5\}$ (just take $\alpha_3 = \alpha_4 = 0$ in the system solved above). Since $\dim(\mathbf{R}^3) = 3$, any three vectors spanning $\mathbf{R}^3$ form a basis for $\mathbf{R}^3$ (by Theorem 45). Hence $\{v_1, v_2, v_5\}$ is a basis for $\mathbf{R}^3$.

2. Consider the following vectors in $\mathbf{R}^4$:

   $$u_1 = (1, 3, 5, -1),\ u_2 = (1, 4, 9, 0),\ u_3 = (4, 9, 7, -5).$$

   (a) We wish to show that $\{u_1, u_2, u_3\}$ is linearly independent, which we do by solving $\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 = 0$. This is equivalent to the system

   $$\begin{aligned} \alpha_1 + \alpha_2 + 4\alpha_3 &= 0, \\ 3\alpha_1 + 4\alpha_2 + 9\alpha_3 &= 0, \\ 5\alpha_1 + 9\alpha_2 + 7\alpha_3 &= 0, \\ -\alpha_1 - 5\alpha_3 &= 0. \end{aligned}$$

   Applying Gaussian elimination, we find that the only solution is $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Thus $\{u_1, u_2, u_3\}$ is linearly independent.

   (b) Since $\mathrm{sp}\{u_1, u_2, u_3\}$ is a three-dimensional subspace of $\mathbf{R}^4$, and hence a very small part of $\mathbf{R}^4$, almost every vector in $\mathbf{R}^4$ does not belong to $\mathrm{sp}\{u_1, u_2, u_3\}$ and hence would be a valid fourth vector for a basis. We choose $u_4 = (0, 0, 0, 1)$, and test whether $\{u_1, u_2, u_3, u_4\}$ is linearly independent. A direct calculation shows that $\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_4 u_4 = 0$ has only the trivial solution, and hence (by Theorem 45) $\{u_1, u_2, u_3, u_4\}$ is a basis for $\mathbf{R}^4$.

3. Let $p_1(x) = 2 - 5x$, $p_2(x) = 2 - 5x + 4x^2$.

   (a) Obviously $\{p_1, p_2\}$ is linearly independent, because neither polynomial is a multiple of the other.

   (b) Now we wish to find a polynomial $p_3 \in \mathcal{P}_2$ such that $\{p_1, p_2, p_3\}$ is a basis for $\mathcal{P}_2$. Since $\text{sp}\{p_1, p_2\}$ is a two-dimensional subspace of the three-dimensional space $\mathcal{P}_2$, almost any polynomial will do; we choose $p_3(x) = 1$. We then test for linear independence by solving $c_1p_1(x) + c_2p_2(x) + c_3p_3(x) = 0$. This equation is equivalent to

$$(2c_1 + 2c_2 + c_3) + (-5c_1 - 5c_2)x + 4c_2x^2 = 0,$$

   which in turn is equivalent to the system

$$\begin{aligned} 2c_1 + 2c_2 + c_3 &= 0, \\ -5c_1 - 5c_2 &= 0, \\ 4c_2 &= 0. \end{aligned}$$

   A direct calculation shows that the only solution is $c_1 = c_2 = c_3 = 0$, and hence $\{p_1, p_2, p_3\}$ is linearly independent. It follows from Theorem 45 that $\{p_1, p_2, p_3\}$ is a basis for $\mathbf{R}^3$.

4. Define $p_1, p_2, p_3, p_4, p_5 \in \mathcal{P}_2$ by

$$\begin{aligned} &p_1(x) = x,\ p_2(x) = 1 + x,\ p_3(x) = 3 + 5x, \\ &p_4(x) = 5 + 8x,\ p_5(x) = 3 + x - x^2. \end{aligned}$$

   (a) We first show that $\{p_1, p_2, p_3, p_4, p_5\}$ spans $\mathcal{P}_2$. Given an arbitrary $q(x) = a_0 + a_1x + a_2x^2$ in $\mathcal{P}_2$, the equation $c_1p_1(x) + c_2p_2(x) + c_3p_3(x) + c_4p_4(x) + c_5p_5(x) = q(x)$ is equivalent to

$$(c_2 + 3c_3 + 5c_4 + 3c_5) + (c_1 + c_2 + 5c_3 + 8c_4 + c_5)x - c_5x^2 = a_0 + a_1x + a_2x^2,$$

   and hence to the system

$$\begin{aligned} c_2 + 3c_3 + 5c_4 + 3c_5 &= a_0, \\ c_1 + c_2 + 5c_3 + 8c_4 + c_5 &= a_1, \\ -c_5 &= a_2. \end{aligned}$$

   Applying Gaussian elimination, we obtain the reduced system

$$\begin{aligned} c_1 &= a_1 - a_0 - 2a_2 - 2c_3 - 3c_4, \\ c_2 &= a_0 + 3a_2 - 3c_3 - 5c_4, \\ c_5 &= -a_2. \end{aligned}$$

   We see that, regardless of the values of a_0, a_1, a_2, there is a solution, and hence $\{p_1, p_2, p_3, p_4, p_5\}$ spans $\mathcal{P}_2$.

   (b) We now find a subset of $\{p_1, p_2, p_3, p_4, p_5\}$ that forms a basis for $\mathcal{P}_2$. From the above calculation, we see that every $q \in \mathcal{P}_2$ can be written as a linear combination of p_1, p_2, p_5. Since $\dim(\mathcal{P}_2) = 3$, Theorem 45 implies that $\{p_1, p_2, p_5\}$ is a basis for $\mathcal{P}_2$.

5. Let $u_1 = (1, 4, 0, -5, 1)$, $u_2 = (1, 3, 0, -4, 0)$, $u_3 = (0, 4, 1, 1, 4)$ be vectors in $\mathbf{R}^5$.

   (a) To show that $\{u_1, u_2, u_3\}$ is linearly independent, we solve the equation $\alpha_1u_1 + \alpha_2u_2 + \alpha_3u_3 = 0$, which is equivalent to the system

$$\begin{aligned} \alpha_1 + \alpha_2 &= 0, \\ 4\alpha_1 + 3\alpha_2 + 4\alpha_3 &= 0, \\ \alpha_3 &= 0, \\ -5\alpha_1 - 4\alpha_2 + \alpha_3 &= 0, \\ \alpha_1 + 4\alpha_3 &= 0. \end{aligned}$$

   A direct calculation shows that this system has only the trivial solution.

(b) To extend $\{u_1, u_2, u_3\}$ to a basis for $\mathbf{R}^5$, we need two more vectors. We will try $u_4 = (0, 0, 0, 1, 0)$ and $u_5 = (0, 0, 0, 0, 1)$. We solve $\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_4 u_4 + \alpha_5 u_5 = 0$ and find that the only solution is the trivial one. This implies that $\{u_1, u_2, u_3, u_4, u_5\}$ is linearly independent and hence, by Theorem 45, a basis for $\mathbf{R}^5$.

6. Consider the following vectors in $\mathbf{R}^5$:

$$u_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \; u_2 = \begin{bmatrix} -1 \\ 3 \\ 2 \\ 1 \\ -1 \end{bmatrix}, \; u_3 = \begin{bmatrix} 1 \\ 7 \\ 2 \\ 3 \\ 1 \end{bmatrix},$$
$$u_4 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \; u_5 = \begin{bmatrix} 2 \\ 10 \\ 3 \\ 6 \\ 2 \end{bmatrix}.$$

Let $S = \text{sp}\{u_1, u_2, u_3, u_4, u_5\}$. We wish to find a subset of $\{u_1, u_2, u_3, u_4, u_5\}$ that is a basis for S. We let x be an arbitrary vector in $\mathbf{R}^5$ and solve $\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + \alpha_4 u_4 + \alpha_5 u_5 = x$. This equation is equivalent to the system

$$\begin{aligned} \alpha_1 - \alpha_2 + \alpha_3 + \alpha_4 + 2\alpha_5 &= x_1, \\ 2\alpha_1 + 3\alpha_2 + 7\alpha_3 - 2\alpha_4 + 10\alpha_5 &= x_2, \\ 2\alpha_2 + 2\alpha_3 - \alpha_4 + 3\alpha_5 &= x_3, \\ \alpha_1 + \alpha_2 + 3\alpha_3 + \alpha_4 + 6\alpha_5 &= x_4, \\ \alpha_1 - \alpha_2 + \alpha_3 + \alpha_4 + 2\alpha_5 &= x_5. \end{aligned}$$

Applying Gaussian elimination, this system is equivalent to

$$\begin{aligned} \alpha_1 &= \frac{3}{2}x_1 + x_3 - \frac{1}{2}x_4 - 2\alpha_3 - 3\alpha_5, \\ \alpha_2 &= -\frac{1}{2}x_1 + \frac{1}{2}x_4 - \alpha_3 - 2\alpha_5, \\ \alpha_4 &= -x_1 - x_3 + x_4 - \alpha_5, \\ 0 &= -\frac{7}{2}x_1 + x_2 - 4x_3 + \frac{3}{2}x_4, \\ 0 &= -x_1 + x_5. \end{aligned}$$

We see first of all that S is a proper subspace of $\mathbf{R}^5$, since $x \notin S$ unless

$$-\frac{7}{2}x_1 + x_2 - 4x_3 + \frac{3}{2}x_4 = 0, \; -x_1 + x_5 = 0.$$

We also see that any $x \in S$ can be represented as a linear combination of u_1, u_2, u_4 by taking $\alpha_3 = \alpha_5 = 0$ in the above equations. Finally, it can be verified directly that $\{u_1, u_2, u_4\}$ is linearly independent and hence a basis for S.

7. Consider the following polynomials in $\mathcal{P}_3$:

$$\begin{aligned} &p_1(x) = 1 - 4x + x^2 + x^3, \; p_2(x) = 3 - 11x + x^2 + 4x^3, \\ &p_3(x) = -x + 2x^2 - x^3, \; p_4(x) = -x^2 + 2x^3, \\ &p_5(x) = 5 - 18x + 2x^2 + 5x^3. \end{aligned}$$

We wish to determine the dimension of $S = \text{sp}\{p_1, p_2, p_3, p_4, p_5\}$. We solve $c_1p_1(x) + c_2p_2(x) + c_3p_3(x) + c_4p_4(x) + c_5p_5(x) = 0$ to determine the linear dependence relationships among these vectors (notice that we already know that $\{p_1, p_2, p_3, p_4, p_5\}$ is linearly dependent because the dimension of $\mathcal{P}_3$ is only 4). This equation is equivalent to the system

$$\begin{aligned} c_1 + 3c_2 + 5c_5 &= 0, \\ -4c_1 - 11c_2 - c_3 - 18c_5 &= 0, \\ c_1 + c_2 + 2c_3 - c_4 + 2c_5 &= 0, \\ 2c_1 + 4c_2 - c_3 + 2c_4 + 5c_5 &= 0. \end{aligned}$$

Applying Gaussian elimation yields

$$c_1 = 0,\ c_2 = -\frac{5}{3}c_5,\ c_3 = \frac{1}{3}c_5,\ c_4 = c_5.$$

Choosing $c_5 = 1$, we see that

$$-\frac{5}{3}p_2(x) + \frac{1}{3}p_3(x) + p_4(x) + p_5(x) = 0.$$

We can solve this equation for $p_5(x)$, which shows that $p_5 \in \text{sp}\{p_2, p_3, p_4\} \subset \text{sp}\{p_1, p_2, p_3, p_4\}$. Therefore, $S = \text{sp}\{p_1, p_2, p_3, p_4\}$. The above calculations also show that the only solution of $c_1p_1(x) + c_2p_2(x) + c_3p_3(x) + c_4p_4(x) + c_5p_5(x) = 0$ with $c_5 = 0$ is the trivial solution, that is, the only solution of $c_1p_1(x) + c_2p_2(x) + c_3p_3(x) + c_4p_4(x) = 0$ is the trivial solution. Thus $\{p_1, p_2, p_3, p_4\}$ is linearly independent and hence a basis for S. (This also shows that S is four-dimensional and hence equals all of $\mathcal{P}_3$.)

8. Let $S = \text{sp}\{v_1, v_2, v_3, v_4\} \subset \mathbf{C}^3$, where

$$\begin{aligned} v_1 &= (1 - i, 3 + i, 1 + i),\ v_2 = (1, 1 - i, 3), \\ v_3 &= (i, -2 - 2i, 2 - i),\ v_4 = (2 - i, 7 + 3i, 2 + 5i). \end{aligned}$$

We will find a basis for S. We begin by solving $\alpha_1v_1 + \alpha_2v_2 + \alpha_3v_3 + \alpha_4v_4 = 0$, which is equivalent to the system

$$\begin{aligned} (1 - i)\alpha_1 + \alpha_2 + i\alpha_3 + (2 - i)\alpha_4 &= 0, \\ (3 + i)\alpha_1 + (1 - i)\alpha_2 + (-2 - 2i)\alpha_3 + (7 + 3i)\alpha_4 &= 0, \\ (1 + i)\alpha_1 + 3\alpha_2 + (2 - i)\alpha_3 + (2 + 5i)\alpha_4 &= 0. \end{aligned}$$

Applying Gaussian elimination yields

$$\alpha_1 = \alpha_3 - 2\alpha_4,\ \alpha_2 = -\alpha_3 - i\alpha_4.$$

Taking $\alpha_3 = 1$, $\alpha_4 = 0$, we see that v_3 can be written as a linear combination of v_1, v_2, and taking $\alpha_3 = 0$, $\alpha_4 = 1$, we see that v_4 can be written as a linear combination of v_1, v_2. Thus both v_3 and v_4 belong to $\text{sp}\{v_1, v_2\}$, which shows that $S = \text{sp}\{v_1, v_2\}$. Since neither v_1 nor v_2 is a multiple of the other, we see that $\{v_1, v_2\}$ is linearly independent and hence is a basis for S.

9. Consider the vectors $u_1 = (3, 1, 0, 4)$ and $u_2 = (1, 1, 1, 4)$ in $\mathbf{Z}_5^4$.

   (a) It is obvious that $\{u_1, u_2\}$ is linearly independent, since neither vector is a multiple of the other.

   (b) To extend $\{u_1, u_2\}$ to a basis for $\mathbf{Z}_5^4$, we must find vectors u_3, u_4 such that $\{u_1, u_2, u_3, u_4\}$ is linearly independent. We try $u_3 = (0, 0, 1, 0)$ and $u_4 = (0, 0, 0, 1)$. A direct calculation then shows that $\alpha_1u_1 + \alpha_2u_2 + \alpha_3u_3 + \alpha_4u_4 = 0$ has only the trivial solution. Therefore $\{u_1, u_2, u_3, u_4\}$ is linearly independent and hence, since $\dim(\mathbf{Z}_5^4) = 4$, it is a basis for $\mathbf{Z}_5^4$.

10. Let $S = \text{sp}\{v_1, v_2, v_3\} \subset \mathbf{Z}_3^3$, where

$$v_1 = (1, 2, 1),\ v_2 = (2, 1, 2),\ v_3 = (1, 0, 1).$$

We wish to find a subset of $\{v_1, v_2, v_3\}$ that is a basis for S. As usual, we proceed by solving $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$ to find the linear dependence relationships (if any). This equation is equivalent to the system

$$\begin{aligned} \alpha_1 + 2\alpha_2 + \alpha_3 &= 0, \\ 2\alpha_1 + \alpha_2 &= 0, \\ \alpha_1 + 2\alpha_2 + \alpha_3 &= 0, \end{aligned}$$

which reduces, by Gaussian elimination, to

$$\alpha_1 = \alpha_2,\ \alpha_3 = 0.$$

It follows that $v_1 + v_2 = 0$, or $v_2 = 2v_1$ (notice that $-1 = 2$ in $\mathbf{Z}_3$). Therefore $\text{sp}\{v_1, v_3\} = \text{sp}\{v_1, v_2, v_3\} = S$. It is obvious that $\{v_1, v_3\}$ is linearly independent (since neither vector is a multiple of the other), and therefore $\{v_1, v_3\}$ is a basis for S.

11. Let F be a field. We will show how to produce different bases for a nontrivial, finite-dimensional vector space over V.

(a) Let V be a 1-dimensional vector space over F, and let $\{u_1\}$ be a basis for V. Then $\{\alpha u_1\}$ is a basis for V for any $\alpha \neq 0$. To prove this, we first note that u_1 is nonzero since $\{u_1\}$ is linearly independent by assumption; therefore $\alpha_1(\alpha u_1) = 0$ implies that $(\alpha_1\alpha)u_1 = 0$ and hence (by Theorem 5, part 6) that $\alpha_1\alpha = 0$. Since $\alpha \neq 0$, this in turn yields $\alpha_1 = 0$, and hence $\{\alpha u_1\}$ is linearly independent. Now suppose $v \in V$. Since $\{u_1\}$ is a basis for V, there exists $\beta \in F$ such that $\beta u_1 = v$. But then $\left(\beta\alpha^{-1}\right)(\alpha u_1) = v$, which shows that $\{\alpha u_1\}$ spans V. Thus $\{\alpha u_1\}$ is a basis for V.

(b) Now let V be a 2-dimensional vector space over F, and let $\{u_1, u_2\}$ be a basis for V. We wish to prove that $\{\alpha u_1, \beta u_1 + \gamma u_2\}$ is a basis for V for any $\alpha \neq 0$, $\gamma \neq 0$. First, suppose $\alpha_1(\alpha u_1) + \alpha_2(\beta u_1 + \gamma u_2) = 0$. We can rewrite this equation as $(\alpha\alpha_1 + \beta\alpha_2)u_1 + (\gamma\alpha_2)u_2 = 0$ and, since $\{u_1, u_2\}$ is linearly independent, this implies that

$$\begin{aligned} \alpha\alpha_1 + \beta\alpha_2 &= 0, \\ \gamma\alpha_2 &= 0. \end{aligned}$$

Since $\gamma \neq 0$ by assumption, the second equation implies that $\alpha_2 = 0$. Then the first equation simplifies to $\alpha\alpha_1 = 0$, which implies (since $\alpha \neq 0$) that $\alpha_1 = 0$. This shows that $\{\alpha u_1, \beta u_1 + \gamma u_2\}$ is linearly independent.

Now suppose $v \in V$. Since $\{u_1, u_2\}$ is a basis for V, there exist $\beta_1, \beta_2 \in F$ such that $v = \beta_1 u_1 + \beta_2 u_2$. We wish to find $\alpha_1, \alpha_2 \in F$ such that $\alpha_1(\alpha u_1) + \alpha_2(\beta u_1 + \gamma u_2) = v$. We can rearrange this last equation to read $(\alpha\alpha_1 + \beta\alpha_2)u_1 + (\gamma\alpha_2) = v$, which then yields $(\alpha\alpha_1 + \beta\alpha_2)u_1 + (\gamma\alpha_2) = \beta_1 u_1 + \beta_2 u_2$. Since v has a unique representation as a linear combination of the basis vectors u_1, u_2, it follows that

$$\begin{aligned} \alpha\alpha_1 + \beta\alpha_2 &= \beta_1, \\ \gamma\alpha_2 &= \beta_2. \end{aligned}$$

This system can be solved to yield a unique solution:

$$\alpha_1 = \alpha^{-1}(\beta_1 - \beta\gamma^{-1}\beta_2),\ \alpha_2 = \gamma^{-1}\beta_2.$$

This shows that $v \in \text{sp}\{\alpha u_1, \beta u_1 + \gamma u_2\}$, and hence that $\{\alpha u_1, \beta u_1 + \gamma u_2\}$ spans V. Therefore, $\{\alpha u_1, \beta u_1 + \gamma u_2\}$ is a basis for V.

(c) Let V be a vector space over F with basis $\{u_1, \ldots, u_n\}$. We wish to generalize the previous parts of this exercise to show how to produce a collection of different bases for V. We choose any scalars

$$\alpha_{ij},\ i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, i,$$

with $\alpha_{ii} \neq 0$ for all $i = 1, 2, \ldots, n$. Then the set

$$\{\alpha_{11}u_1, \alpha_{21}u_1 + \alpha_{22}u_2, \ldots, \alpha_{n1}u_1 + \alpha_{n2}u_2 + \cdots + \alpha_{nn}u_n\} \tag{2.2}$$

is a basis for V. We can prove this by induction on n. We have already done the case $n = 1$ (and also $n = 2$). Let us assume that the construction leads to a basis if the dimension of the vector space is $n - 1$, and suppose that V has dimension n, with basis $\{u_1, u_2, \ldots, u_n\}$. By the induction hypothesis,

$$\{\alpha_{11}u_1, \alpha_{21}u_1 + \alpha_{22}u_2, \ldots, \alpha_{n-1,1}u_1 + \cdots + \alpha_{n-1,n-1}u_{n-1}\} \tag{2.3}$$

is a basis for $S = \mathrm{sp}\{u_1, u_2, \ldots, u_{n-1}\}$. We now show that (2.2) is a basis for V by showing that each $v \in V$ can be uniquely represented as a linear combination of the vectors in (2.2). So let v be any vector in V, say

$$v = \beta_1 u_1 + \beta_2 u_2 + \cdots + \beta_n u_n.$$

Notice that

$$\begin{aligned}\beta_n u_n = \beta_n \alpha_{nn}^{-1} (\alpha_{n1}u_1 + \cdots + \alpha_{nn}u_n) - \\ \left(\beta_n \alpha_{nn}^{-1}\alpha_{n1}u_1 + \cdots + \beta_n \alpha_{nn}^{-1}\alpha_{n,n-1}u_{n-1}\right).\end{aligned}$$

The vector

$$\beta_n \alpha_{nn}^{-1}\alpha_{n1}u_1 + \cdots + \beta_n \alpha_{nn}^{-1}\alpha_{n,n-1}u_{n-1}$$

belongs to S and, by the induction hypothesis, can be written uniquely as

$$\gamma_1(a_{11}u_1) + \gamma_2(\alpha_{21}u_1 + \alpha_{22}u_2) + \cdots + \gamma_{n-1}(\alpha_{n-1,1}u_1 + \cdots + \alpha_{n-1,n-1}u_{n-1}).$$

Also,

$$\begin{aligned}&\beta_1 u_1 + \cdots + \beta_{n-1}u_{n-1} \\ &= \delta_1(a_{11}u_1) + \delta_2(\alpha_{21}u_1 + \alpha_{22}u_2) + \cdots + \\ &\quad \delta_{n-1}(\alpha_{n-1,1}u_1 + \cdots + \alpha_{n-1,n-1}u_{n-1}).\end{aligned}$$

Putting this all together, we obtain

$$\begin{aligned}v &= \beta_1 u_1 + \cdots + \beta_{n-1}u_{n-1} + \beta_n u_n \\ &= \delta_1(a_{11}u_1) + \delta_2(\alpha_{21}u_1 + \alpha_{22}u_2) + \cdots + \\ &\quad \delta_{n-1}(\alpha_{n-1,1}u_1 + \cdots + \alpha_{n-1,n-1}u_{n-1}) + \\ &\quad \beta_n \alpha_{nn}^{-1} (\alpha_{n1}u_1 + \cdots + \alpha_{nn}u_n) - (\gamma_1(a_{11}u_1) + \gamma_2(\alpha_{21}u_1 + \alpha_{22}u_2) \\ &\quad + \cdots + \gamma_{n-1}(\alpha_{n-1,1}u_1 + \cdots + \alpha_{n-1,n-1}u_{n-1}) \\ &= (\delta_1 - \gamma_1)(a_{11}u_1) + (\delta_2 - \gamma_2)(\alpha_{21}u_1 + \alpha_{22}u_2) + \cdots + \\ &\quad (\delta_{n-1} - \gamma_{n-1})(\alpha_{n-1,1}u_1 + \cdots + \alpha_{n-1,n-1}u_{n-1}) + \\ &\quad \beta_n \alpha_{nn}^{-1} (\alpha_{n1}u_1 + \cdots + \alpha_{nn}u_n).\end{aligned}$$

This shows that each v can be written as a linear combination of the vectors in (2.2). Uniqueness follows from the induction hypothesis and the fact that there is only one way to write $\beta_n u_n$ as a multiple of $\alpha_{n1}u_1 + \cdots + \alpha_{nn}u_n$ plus a vector from S.

12. We wish to prove that every nontrivial subspace of a finite-dimensional vector space has a basis (and hence is finite-dimensional). Let V be a finite-dimensional vector space, let $\dim(V) = n$, and suppose S is a nontrivial subspace of V. Since S is nontrivial, it contains a nonzero vector s_1. Then either $\{v_1\}$ spans S, or there exists $v_2 \in S \setminus \text{sp}\{v_1\}$. In the first case, $\{v_1\}$ is a basis for S (since any set containing a single nonzero vector is linearly independent by Exercise 2.5.2). In the second case, $\{v_1, v_2\}$ is linearly independent by Exercise 2.5.4. Either $\{v_1, v_2\}$ spans S, in which case it is a basis for S, or we can find $v_3 \in S \setminus \text{sp}\{v_1, v_2\}$. We continue to add vectors in this fashion until we obtain a basis for S. We know that the process will end with a linearly independent spanning set for S, containing at most n vectors, because $S \subset V$, and any linearly independent set with n vectors spans all of V by Theorem 45.

13. Let F be a finite field with q distinct elements, and let n be positive integer, $n \geq q$. We wish to prove that $\dim(\mathcal{P}_n(F)) = q$ by showing that $\{1, x, \ldots, x^{q-1}\}$ is a basis for $\mathcal{P}_n(F)$. We have already seen (in Exercise 2.6.11) that $\{1, x, \ldots, x^{q-1}\}$ is linearly independent. Consider the following vectors in F^q:

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \; v_2 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_q \end{bmatrix}, \; v_3 = \begin{bmatrix} \alpha_1^2 \\ \alpha_2^2 \\ \vdots \\ \alpha_q^2 \end{bmatrix}, \; \ldots, \; v_q = \begin{bmatrix} \alpha_1^{q-1} \\ \alpha_2^{q-1} \\ \vdots \\ \alpha_q^{q-1} \end{bmatrix}.$$

The equation $c_1 v_1 + c_2 v_2 + \cdots + c_q v_q = 0$ is equivalent to the q equations

$$c_1 \cdot 1 + c_2 \alpha_i + \cdots + c_q \alpha_i^{q-1} = 0, \; i = 1, 2, \ldots, q,$$

which collectively are equivalent to the statement

$$c_1 \cdot 1 + c_2 x + \cdots + c_q x^{q-1} = 0 \text{ for all } x \in F.$$

Since $\{1, x, \ldots, x^{q-1}\}$ is linearly independent, this equation implies that $c_1 = c_2 = \cdots = c_q = 0$, and hence we have shown that $\{v_1, v_2, \ldots, v_q\}$ is linearly independent in F^q. Since we know that $\dim(F^q) = q$, Theorem 45 implies that $\{v_1, v_2, \ldots, v_q\}$ also spans F^q. Now let p be any polynomial in $\mathcal{P}_n(F)$, and define

$$u = \begin{bmatrix} p(\alpha_1) \\ p(\alpha_2) \\ \vdots \\ p(\alpha_q) \end{bmatrix} \in F^q.$$

Since $\{v_1, v_2, \ldots, v_q\}$ spans F^q, there exist scalars $c_1, c_2, \ldots, c_q \in F$ such that $c_1 v_1 + c_2 v_2 + \cdots + c_q v_q = u$. This last equation is equivalent to

$$c_1 \cdot 1 + c_2 \alpha_i + \cdots + c_q \alpha_i^{q-1} = p(\alpha_i), \; i = 1, 2, \ldots, q,$$

and hence to

$$c_1 \cdot 1 + c_2 x + \cdots + c_q x^{q-1} = p(x) \text{ for all } x \in F.$$

This shows that $p \in \text{sp}\{v_1, v_2, \ldots, v_q\}$, and hence that $\{v_1, v_2, \ldots, v_q\}$ is a basis for $\mathcal{P}_n(F)$. Thus $\dim(\mathcal{P}_n(F)) = q$, as desired.

14. Let V be an n-dimensional vector space over a field F, and suppose S and T are subspaces of V satisfying $S \cap T = \{0\}$. Suppose that $\{s_1, s_2, \ldots, s_k\}$ is a basis for S, $\{t_1, t_2, \ldots, t_\ell\}$ is a basis for T, and $k + \ell = n$. We wish to prove that $\{s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_\ell\}$ is a basis for V. This follows immediately from Theorem 45, since we have already shown in Exercise 2.5.15 that $\{s_1, s_2, \ldots, s_k, t_1, t_2, \ldots, t_\ell\}$ is linearly independent.

15. Let V be a vector space over a field F, and let $\{u_1, \ldots, u_n\}$ be a basis for V. Let $v_1, \ldots, v_k$ be vectors in V, and suppose

$$v_j = \alpha_{1,j} u_1 + \ldots + \alpha_{n,j} u_n, \; j = 1, 2, \ldots, k.$$

Define the vectors $x_1, \ldots, x_k$ in F^n by

$$x_j = (\alpha_{1,j}, \ldots, \alpha_{n,j}),\ j = 1, 2, \ldots, k.$$

(a) We first prove that $\{v_1, \ldots, v_k\}$ is linearly independent if and only if $\{x_1, \ldots, x_k\}$ is linearly independent. We will do this by showing that $c_1v_1 + \cdots + c_kv_k = 0$ in V is equivalent to $c_1x_1 + \cdots + c_kx_k = 0$ in F^n. Then the first equation has only the trivial solution if and only if the second equation does, and the result follows. The proof is a direct manipulation, for which summation notation is convenient:

$$\begin{aligned}
\sum_{j=1}^{k} c_j v_j = 0 &\Leftrightarrow \sum_{j=1}^{k} c_j \left(\sum_{i=1}^{n} \alpha_{ij} u_i \right) \\
&\Leftrightarrow \sum_{j=1}^{k} \sum_{i=1}^{n} c_j \alpha_{ij} u_i = 0 \\
&\Leftrightarrow \sum_{i=1}^{n} \sum_{j=1}^{k} c_j \alpha_{ij} u_i = 0 \\
&\Leftrightarrow \sum_{i=1}^{n} \left(\sum_{j=1}^{k} c_j \alpha_{ij} \right) u_i = 0.
\end{aligned}$$

Since $\{u_1, \ldots, u_n\}$ is linearly independent, the last equation is equivalent to

$$\sum_{j=1}^{k} c_j \alpha_{ij} = 0,\ i = 1, 2, \ldots, n,$$

which, by definition of x_j and of addition in F^n, is equivalent to

$$\sum_{j=1}^{k} c_j x_j = 0.$$

This completes the proof.

(b) Now we show that $\{v_1, \ldots, v_k\}$ spans V if and only if $\{x_1, \ldots, x_k\}$ spans F^n. Since each vector in V can be represented uniquely as a linear combination of $u_1, \ldots, u_n$, there is a one-to-one correspondence between V and F^n:

$$w = c_1u_1 + \cdots + c_nu_n \in V \ \longleftrightarrow\ x = (c_1, \ldots, c_n) \in F^n.$$

Mimicking the manipulations in the first part of the exercise, we see that

$$\sum_{j=1}^{k} c_j v_j = w \ \Leftrightarrow\ \sum_{j=1}^{k} c_j x_j = x.$$

Thus the first equation has a solution for every $v \in V$ if and only if the second equation has a solution for every $x \in F^n$. The result follows.

16. Consider the polynomials $p_1(x) = -1 + 3x + 2x^2$, $p_2(x) = 3 - 8x - 4x^2$, and $p_3(x) = -1 + 4x + 5x^2$ in $\mathcal{P}_2$. We wish to use the result of the previous exercise to determine if $\{p_1, p_2, p_3\}$ is linearly independent. The standard basis for $\mathcal{P}_2$ is $\{u_1, u_2, u_3\} = \{1, x, x^2\}$. In terms of this basis, p_1, p_2, p_3 correspond to

$$x_1 = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix},\ x_2 = \begin{bmatrix} 3 \\ -8 \\ -4 \end{bmatrix},\ x_3 = \begin{bmatrix} -1 \\ 4 \\ 5 \end{bmatrix} \in \mathbf{R}^3,$$

respectively. A direct calculation shows that $c_1x_1 + c_2x_2 + c_3x_3 = 0$ has only the trivial solution. Therefore $\{x_1, x_2, x_3\}$ and $\{p_1, p_2, p_3\}$ are both linearly independent.

2.8 Polynomial interpolation and the Lagrange basis

1. (a) The Lagrange polynomials for the interpolation nodes $x_0 = 1$, $x_1 = 2$, $x_3 = 3$ are

$$\begin{aligned} L_0(x) &= \frac{(x-2)(x-3)}{(1-2)(1-3)} = -\frac{1}{2}(x-2)(x-3), \\ L_1(x) &= \frac{(x-1)(x-3)}{(2-1)(2-3)} = -(x-1)(x-3), \\ L_2(x) &= \frac{(x-1)(x-2)}{(3-1)(3-2)} = \frac{1}{2}(x-1)(x-2). \end{aligned}$$

   (b) The quadratic polynomial interpolating $(1,0)$, $(2,2)$, $(3,1)$ is

$$\begin{aligned} p(x) &= 0L_0(x) + 2L_1(x) + L_2(x) \\ &= -2(x-1)(x-3) + \frac{1}{2}(x-1)(x-2) \\ &= -\frac{3}{2}x^2 + \frac{13}{2}x - 5. \end{aligned}$$

2. (a) The Lagrange polynomials for the interpolation nodes $x_0 = -2$, $x_1 = -1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$ are

$$\begin{aligned} L_0(x) &= \frac{(x+1)x(x-1)(x-2)}{(-2+1)(-2-0)(-2-1)(-2-2)} = \frac{1}{24}x(x+1)(x-1)(x-2), \\ L_1(x) &= \frac{(x+2)x(x-1)(x-2)}{(-1+2)(-1-0)(-1-1)(-1-2)} = -\frac{1}{6}x(x+2)(x-1)(x-2), \\ L_2(x) &= \frac{(x+2)(x+1)(x-1)(x-2)}{(0+2)(0+1)(0-1)(0-2)} = \frac{1}{4}(x+2)(x+1)(x-1)(x-2), \\ L_3(x) &= \frac{(x+2)(x+1)x(x-2)}{(1+2)(1+1)(1-0)(1-2)} = -\frac{1}{6}x(x+2)(x+1)(x-2), \\ L_4(x) &= \frac{(x+2)(x+1)x(x-1)}{(2+2)(2+1)(2-0)(2-1)} = \frac{1}{24}x(x+2)(x+1)(x-1). \end{aligned}$$

   (b) Using the Lagrange basis, we find the interpolating polynomial passing through the points $(-2,10)$, $(-1,-3)$, $(0,2)$, $(1,7)$, $(2,18)$ to be

$$\begin{aligned} p(x) &= 10L_0(x) - 3L_1(x) + 2L_2(x) + 7L_3(x) + 18L_4(x) \\ &= \frac{5}{12}x(x+1)(x-1)(x-2) + \frac{1}{2}x(x+2)(x-1)(x-2) + \\ &\quad \frac{1}{2}(x+2)(x+1)(x-1)(x-2) - \frac{7}{6}x(x+2)(x+1)(x-2) + \\ &\quad \frac{3}{4}x(x+2)(x+1)(x-1). \end{aligned}$$

   A tedious calculation shows that $p(x) = x^4 - x^3 - x^2 + 6x + 2$, which is the same result obtained in Example 48.

3. Consider the data $(1,5)$, $(2,-4)$, $(3,-4)$, $(4,2)$. We wish to find the cubic polynomial interpolating these points.

   (a) Using the standard basis, we write $p(x) = c_0 + c_1x + c_2x^2 + c_3x^4$. The equations

$$p(1) = 5,\ p(2) = -4,\ p(3) = -4,\ p(4) = 2$$

are equivalent to the system

$$\begin{aligned} c_0 + c_1 + c_2 + c_3 &= 5, \\ c_0 + 2c_1 + 4c_2 + 8c_3 &= -4, \\ c_0 + 3c_1 + 9c_2 + 27c_3 &= -4, \\ c_0 + 4c_1 + 16c_2 + 64c_3 &= 2. \end{aligned}$$

Gaussian elimination yields

$$c_0 = 26,\ c_1 = -28,\ c_2 = \frac{15}{2},\ c_3 = -\frac{1}{2},$$

and thus

$$p(x) = 26 - 28x + \frac{15}{2}x^2 - \frac{1}{2}x^3.$$

(b) The Lagrange polynomials for these interpolation nodes are

$$\begin{aligned} L_0(x) &= \frac{(x-2)(x-3)(x-4)}{(1-2)(1-3)(1-4)} = -\frac{1}{6}(x-2)(x-3)(x-4), \\ L_1(x) &= \frac{(x-1)(x-3)(x-4)}{(2-1)(2-3)(2-4)} = \frac{1}{2}(x-1)(x-3)(x-4), \\ L_2(x) &= \frac{(x-1)(x-2)(x-4)}{(3-1)(3-2)(3-4)} = -\frac{1}{2}(x-1)(x-2)(x-4), \\ L_2(x) &= \frac{(x-1)(x-2)(x-3)}{(4-1)(4-2)(4-3)} = \frac{1}{6}(x-1)(x-2)(x-3), \end{aligned}$$

and the interpolating polynomial is

$$\begin{aligned} p(x) &= 5L_0(x) - 4L_1(x) - 4L_2(x) + 2L_3(x) \\ &= -\frac{5}{6}(x-2)(x-3)(x-4) - 2(x-1)(x-3)(x-4) + \\ &\quad 2(x-1)(x-2)(x-4) + \frac{1}{3}(x-1)(x-2)(x-3). \end{aligned}$$

A tedious calculation shows that this is the same polynomial computed in the first part.

4. Let $\{L_0, L_1, \ldots, L_n\}$ be the Lagrange basis constructed on the interpolation nodes $x_0, x_1, \ldots, x_n \in F$. We wish to prove that, for all $p \in \mathcal{P}_n(F)$,

$$p(x) = p(x_0)L_0(x) + p(x_1)L_1(x) + \ldots + p(x_n)L_n(x)$$

Let $q(x)$ be the polynomial on the right. By the definition of the Lagrange polynomials, we see that $q(x_i) = p(x_i)$ for $i = 0, 1, \ldots, n$. If we define the polynomial $r(x) = p(x) - q(x)$, then we see that r has $n+1$ roots:

$$r(x_i) = p(x_i) - q(x_i) = p(x_i) - p(x_i) = 0,\ i = 0, 1, \ldots, n.$$

However, r is a polynomial of degree at most n, and therefore this is impossible unless r is the zero polynomial (compare the discussion on page 45 in the text). This shows that $q = p$, as desired.

5. We wish to write $p_2(x) = 2 + x - x^2$ as a linear combination of the Lagrange polynomials constructed on the nodes $x_0 = -1$, $x_1 = 1$, $x_2 = 3$. The graph of p passes through the points $(-1, p(-1))$, $(1, p(1))$, $(3, p(3))$, that is, $(-1, 0)$, $(1, 2)$, $(3, -4)$. The Lagrange polynomials are

$$\begin{aligned} L_0(x) &= \frac{(x-1)(x-3)}{(-1-1)(-1-3)} = \frac{1}{8}(x-1)(x-3), \\ L_1(x) &= \frac{(x+1)(x-3)}{(1+1)(1-3)} = -\frac{1}{4}(x+1)(x-3), \\ L_2(x) &= \frac{(x+1)(x-1)}{(3+1)(3-1)} = \frac{1}{8}(x+1)(x-1), \end{aligned}$$

and therefore,

$$\begin{aligned} p(x) &= 0L_0(x) + 2L_1(x) - 4L_2(x) \\ &= -\frac{1}{2}(x+1)(x-3) - \frac{1}{2}(x+1)(x-1). \end{aligned}$$

6. Let F be a field and suppose $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ are points in F^2. We wish to show that the polynomial interpolation problem has at most one solution, assuming the interpolation nodes $x_0, x_1, \ldots, x_n$ are distinct. This follows from reasoning we have seen before. If $p, q \in \mathcal{P}_n(F)$ both interpolate the given data, then $r = p - q$ is a nonzero polynomial of degree at most n having $n+1$ roots $(x_0, x_1, \ldots, x_n)$. The only polynomial of degree n (or less) having $n+1$ roots is the zero polynomial; thus $p = q$ and there is at most one interpolating polynomial.

7. Let F be a field and suppose $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ are points in F^2. We wish to show that the polynomial interpolation problem has at most one solution, assuming the interpolation nodes $x_0, x_1, \ldots, x_n$ are distinct. Suppose $p, q \in \mathcal{P}_n(F)$ both interpolate the data, and let $\{L_0, L_1, \ldots, L_n\}$ be the basis for $\mathcal{P}_n(F)$ of Lagrange polynomials for the given interpolation nodes. Both p and q can be written in terms of this basis:

$$p = \sum_{i=0}^{n} \alpha_i L_i, \;\; q = \sum_{i=0}^{n} \beta_i L_i.$$

Now, we know that the Lagrange polynomials satisfy

$$L_i(x_j) = \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases}$$

It follows that

$$p(x_j) = \sum_{i=0}^{n} \alpha_i L_i(x_j) = \alpha_j, \;\; q(x_j) = \sum_{i=0}^{n} \beta_i L_i(x_j) = \beta_j.$$

But, since p and q interpolate the given data, $p(x_j) = q(x_j) = y_j$. This shows that $\alpha_j = \beta_j$, $j = 0, 1, \ldots, n$, and hence that $p = q$.

8. Suppose $x_0, x_1, \ldots, x_n$ are distinct real numbers. We wish to prove that, for any real numbers $y_0, y_1, \ldots, y_n$, the system

$$\begin{aligned} c_0 + c_1 x_0 + c_2 x_0^2 + \ldots + c_n x_0^n &= y_0, \\ c_0 + c_1 x_1 + c_2 x_1^2 + \ldots + c_n x_1^n &= y_1, \\ \vdots \qquad\qquad\qquad\qquad & \qquad \vdots \\ c_0 + c_1 x_n + c_2 x_n^2 + \ldots + c_n x_n^n &= y_n \end{aligned}$$

has a unique solution $c_0, c_1, \ldots, c_n$. This follows immediately from our work on interpolating polynomials: $c_0, c_1, \ldots, c_n$ solves the given system if and only if $p(x) = c_0 + c_1 x + \cdots + c_n x^n$ interpolates the data $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$. Since there is a unique interpolating polynomial $p \in \mathcal{P}_n$, and since this polynomial can be uniquely represented in terms of the standard basis $\{1, x, \ldots, x^n\}$, it follows that the given system of equations has a unique solution.

9. We wish to represent every function $f : \mathbf{Z}_2 \to \mathbf{Z}_2$ by a polynomial in $\mathcal{P}_1(\mathbf{Z}_2)$. There are exactly four different functions $f : \mathbf{Z}_2 \to \mathbf{Z}_2$, as defined in the following table:

x	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$
0	0	1	0	1
1	0	0	1	1

We have $f_1(x) = 0$ (the zero polynomial), $f_2(x) = 1 + x$, $f_3(x) = x$, and $f_4(x) = 1$.

10. The following table defines three functions mapping $\mathbf{Z}_3 \to \mathbf{Z}_3$. We wish to find a polynomial in $\mathcal{P}_2(\mathbf{Z}_3)$ representing each one.

x	$f_1(x)$	$f_2(x)$	$f_3(x)$
0	1	0	2
1	2	0	2
2	0	1	1

We can compute f_1, f_2, and f_3 as interpolating polynomials. The Lagrange polynomials for the given interpolation nodes are

$$\begin{aligned} L_0(x) &= \frac{(x-1)(x-2)}{(0-1)(0-2)} = 2x^2 + 2, \\ L_1(x) &= \frac{x(x-2)}{(1-0)(1-2)} = 2x^2 + 2x, \\ L_2(x) &= \frac{x(x-1)}{(2-0)(2-1)} = 2x^2 + x. \end{aligned}$$

(Here we have used the arithmetic of $\mathbf{Z}_3$ to simplify the polynomials: $-1 = 2$, $-2 = 1$, $2^{-1} = 2$, etc.). We then have

$$\begin{aligned} f_1(x) &= L_0(x) + 2L_1(x) = 2x^2 + 1 + x^2 + x = 1 + x, \\ f_2(x) &= L_2(x) = 2x^2 + x, \\ f_3(x) &= 2L_0(x) + 2L_1(x) + L_2(x) = x^2 + 2 + x^2 + x + 2x^2 + x \\ &= x^2 + 2x + 2. \end{aligned}$$

11. Consider a secret sharing scheme in which five individuals will receive information about the secret, and any two of them, working together, will have access to the secret. Assume that the secret is a two-digit integer, and that p is chosen to be 101. The degree of the polynomial will be one, since then the polynomial will be uniquely determined by two data points. Let us suppose that the secret is $N = 42$ and we choose the polynomial to be $p(x) = N + c_1 x$, where $c_1 = 71$ (recall that c_1 is chosen at random). We also choose the five interpolation nodes at random to obtain $x_1 = 9$, $x_2 = 14$, $x_3 = 39$, $x_4 = 66$, and $x_5 = 81$. We then compute

$$\begin{aligned} y_1 &= p(x_1) = 42 + 71 \cdot 9 = 75, \\ y_2 &= p(x_2) = 42 + 71 \cdot 14 = 26, \\ y_3 &= p(x_3) = 42 + 71 \cdot 39 = 84, \\ y_4 &= p(x_4) = 42 + 71 \cdot 66 = 82, \\ y_5 &= p(x_5) = 42 + 71 \cdot 81 = 36 \end{aligned}$$

(notice that all arithmetic is done modulo 101). The data points, to be distributed to the five individuals, are $(9, 75)$, $(14, 26)$, $(39, 84)$, $(66, 82)$, $(81, 36)$.

12. An integer N satisfying $1 \leq N \leq 256$ represents a secret to be shared among five individuals. Any three of the individuals are allowed access to the information. The secret is encoded in a polynomial p, according to the secret sharing scheme described in Section 2.8.1, lying in $\mathcal{P}_2(\mathbf{Z}_{257})$. Suppose three of the individuals get together, and their data points are $(15, 13)$, $(114, 94)$, and $(199, 146)$. We wish to determine the secret. We begin by finding the Lagrange polynomials for the interpolation nodes 15, 114,

and 199:

$$L_0(x) = \frac{(x-114)(x-199)}{(15-114)(15-199)} = 58x^2 + 93x + 205,$$
$$L_1(x) = \frac{(x-15)(x-199)}{(114-15)(114-199)} = 74x^2 + 98x + 127,$$
$$L_2(x) = \frac{(x-15)(x-114)}{(199-15)(199-114)} = 125x^2 + 66x + 183.$$

The interpolating polynomial p is

$$p(x) = 13L_0(x) + 94L_1(x) + 146L_2(x).$$

We need only compute $p(0)$:

$$p(0) = 13L_0(0) + 94L_1(0) + 146L_2(0) = 13 \cdot 205 + 94 \cdot 127 + 146 \cdot 183 = 201.$$

Thus the secret is $N = 201$. (The polynomial $p(x)$ simplifies to $p(x) = 3x^2 + 11x + 201$).

13. We wish to solve the following interpolation problem: Given $v_1, v_2, d_1, d_2 \in \mathbf{R}$, find $p \in \mathcal{P}_3$ such that

$$p(0) = v_1,\ p(1) = v_2,\ p'(0) = d_1,\ p'(1) = d_2.$$

(a) If we represent p as $p(x) = c_0 + c_1x + c_2x^2 + c_3x^3$, then the given conditions

$$\begin{aligned} p(0) &= v_1, \\ p'(0) &= d_1, \\ p(1) &= v_2, \\ p'(1) &= d_2 \end{aligned}$$

are equivalent to the system

$$\begin{aligned} c_0 &= v_1, \\ c_1 &= d_1, \\ c_0 + c_1 + c_2 + c_3 &= v_2, \\ c_1 + 2c_2 + 3c_3 &= d_2. \end{aligned}$$

It is straightforward to solve this system:

$$c_0 = v_1,\ c_1 = d_1,\ c_2 = 3v_2 - 3v_1 - 2d_1 - d_2,\ c_3 = 2v_1 - 2v_2 + d_1 + d_2.$$

(b) We now find the special basis $\{q_1, q_2, q_3, q_4\}$ of $\mathcal{P}_3$ satisfying the following conditions:

$$\begin{aligned} q_1(0) = 1,\ q_1'(0) = 0,\ q_1(1) = 0,\ q_1'(1) = 0, \\ q_2(0) = 0,\ q_2'(0) = 1,\ q_2(1) = 0,\ q_2'(1) = 0, \\ q_3(0) = 0,\ q_3'(0) = 0,\ q_3(1) = 1,\ q_3'(1) = 0, \\ q_4(0) = 0,\ q_4'(0) = 0,\ q_4(1) = 0,\ q_4'(1) = 1. \end{aligned}$$

We can use the result of the first part of this exercise to write down the solutions immediately:

$$\begin{aligned} q_1(x) &= 1 - 3x^2 + 2x^3, \\ q_2(x) &= x - 2x^2 + x^3, \\ q_1(x) &= 3x^2 - 2x^3, \\ q_1(x) &= -x^2 + x^3. \end{aligned}$$

In terms of the basis $\{q_1, q_2, q_3, q_4\}$, the solution to the interpolation problem is

$$p(x) = v_1q_1(x) + d_1q_2(x) + v_2q_3(x) + d_2q_4(x).$$

14. We are given $n+1$ interpolation nodes, $x_0, x_1, \ldots, x_n$. Given $v_0, v_1, \ldots, v_n \in \mathbf{R}$, $d_0, d_1, \ldots, d_n \in \mathbf{R}$, we wish to find $p \in \mathcal{P}_{2n+1}$ such that

$$p(x_i) = v_i,\ p'(x_i) = d_i,\ i = 0, 1, \ldots, n.$$

We define (in terms of the Lagrange polynomials $L_0, L_1, \ldots, L_n$ for the nodes $x_0, x_1, \ldots, x_n$)

$$A_i(x) = (1 - 2(x - x_i)L_i'(x_i))\, L_i^2(x),$$
$$B_i(x) = (x - x_i)L_i^2(x)$$

for $i = 0, 1, \ldots, n$.

(a) Since each Lagrange polynomial L_i has degree exactly n, we see that B_i has degree $2n+1$ for each $i = 0, 1, \ldots, n$, while A_i has degree $2n+1$ if $L_i'(x_i) \neq 0$ and degree $2n$ if $L_i'(x_i) = 0$. Thus, in every case, $A_i, B_i \in \mathcal{P}_{2n+1}$ for all $i = 0, 1, \ldots, n$.

(b) If $j \neq i$, then $L_i(x_j) = 0$ and

$$A_i(x_j) = (1 - 2(x_j - x_i)L_i'(x_i))L_i^2(x_j) = (1 - 2(x_j - x_i)L_i'(x_i)) \cdot 0 = 0.$$

If $j = i$, then $L_i(x_j) = 1$ and

$$A_i(x_j) = (1 - 2(x_i - x_i)L_i'(x_i))L_i^2(x_i) = 1 \cdot 1 = 1.$$

Therefore,

$$A_i(x_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Also,

$$A_i'(x) = 2(1 - 2(x - x_j)L_i'(x_i))L_i(x)L_i'(x) - 2L_i'(x_i)L_i^2(x).$$

If $j \neq i$, then $L_i(x_j) = 0$ and therefore $A_i'(x_j) = 0$ (since both terms contain a factor of $L_i(x_j)$). If $j = i$, then $L_i(x_j) = 1$ and $A_i'(x_j)$ simplifies to

$$A_i'(x_j) = 2L_i'(x_j) - 2L_i'(x_j) = 0.$$

Thus $A_i'(x_j) = 0$, $j = 0, 1, \ldots, n$.

(c) Since either $x_j - x_i = 0$ (if $j = i$) or $L_i(x_j) = 0$ (if $j \neq i$), it follows that $B_i(x_j) = 0$ for all $j = 0, 1, \ldots, n$. Now,

$$B_i'(x) = L_i^2(x) + 2(x - x_i)L_i(x)L_i'(x).$$

If $j \neq i$, then

$$B_i'(x_j) = L_i^2(x_j) + 2(x_j - x_i)L_i(x_j)L_i'(x_j) = 0 - 0 = 0$$

since $L_i(x_j) = 0$. Also,

$$B_i'(x_i) = L_i^2(x_i) + 2(x_i - x_i)L_i(x_i)L_i'(x_i) = 1 - 0 = 1.$$

Therefore,

$$B_i'(x_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

(d) We now wish to prove that $\{A_0, \ldots, A_n, B_0, \ldots, B_n\}$ is a basis for $\mathcal{P}_{2n+1}$. Since $\dim(\mathcal{P}_{2n+1}) = 2n+2$ and the proposed basis contains $2n+2$ elements, it suffices to prove that the set spans $\mathcal{P}_{2n+1}$. Given any $p \in \mathcal{P}_{2n+1}$, define $v_j = p(x_j)$, $d_j = p'(x_j)$, $j = 0, 1, \ldots, n$. Then

$$q(x) = \sum_{i=0}^{n} v_i A_i(x) + \sum_{i=0}^{n} d_i B_i(x)$$

agrees with p at each x_j, and q' agrees with p' at each x_j (see below). Define $r = p - q$. Then r is a polynomial of degree at most $2n+1$, and $r(x_j) = r'(x_j) = 0$ for $j = 0, 1, \ldots, n$. Using elementary properties of polynomials, this implies that $r(x)$ can be factored as

$$r(x) = f(x)(x - x_0)^2(x - x_1)^2 \cdots (x - x_n)^2,$$

where $f(x)$ is a polynomial. But then $\deg(r(x)) \geq 2n + 2$ unless $f = 0$. Since we know that $\deg(r(x)) \leq 2n + 1$, it follows that $f = 0$, in which case $r = 0$ and hence $q = p$. Thus $p \in \text{sp}\{A_0, \ldots, A_n, B_0, \ldots, B_n\}$. This proves that the given set is a basis for $\mathcal{P}_{2n+1}$.

Using the properties of $A_0, A_1, \ldots, A_n$, $B_0, B_1, \ldots, B_n$ derived above, the solution to the Hermite interpolation problem is

$$p(x) = \sum_{i=0}^{n} v_i A_i(x) + \sum_{i=0}^{n} d_i B_i(x).$$

To verify this, notice that

$$p(x_j) = \sum_{i=0}^{n} v_i A_i(x_j) + \sum_{i=0}^{n} d_i B_i(x_j).$$

Every term in the second sum vanishes, as do all terms in the first sum except $v_j A_j(x_j) = v_j \cdot 1 = v_j$. Thus $p(x_j) = v_j$. Also,

$$p'(x_j) = \sum_{i=0}^{n} v_i A_i'(x_j) + \sum_{i=0}^{n} d_i B_i'(x_j).$$

Now every term in the first sum vanishes, as do all terms in the second sum except $d_j B_j'(x_j) = d_j \cdot 1 = d_j$. Therefore $p'(x_j) = d_j$, and p is the desired interpolating polynomial.

2.9 Continuous piecewise polynomial functions

1. The following table shows the maximum errors obtained in approximating $f(x) = e^x$ on the interval $[0, 1]$ by polynomial interpolation and by piecewise linear interpolation, each on a uniform grid with n nodes.

n	Poly. interp. err.	PW linear interp. err.
1	$2.1187 \cdot 10^{-1}$	$2.1187 \cdot 10^{-1}$
2	$1.4420 \cdot 10^{-2}$	$6.6617 \cdot 10^{-2}$
3	$9.2390 \cdot 10^{-4}$	$3.2055 \cdot 10^{-2}$
4	$5.2657 \cdot 10^{-5}$	$1.8774 \cdot 10^{-2}$
5	$2.6548 \cdot 10^{-6}$	$1.2312 \cdot 10^{-2}$
6	$1.1921 \cdot 10^{-7}$	$8.6902 \cdot 10^{-3}$
7	$4.8075 \cdot 10^{-9}$	$6.4596 \cdot 10^{-3}$
8	$1.7565 \cdot 10^{-10}$	$4.9892 \cdot 10^{-3}$
9	$5.8575 \cdot 10^{-12}$	$3.9692 \cdot 10^{-3}$
10	$1.8119 \cdot 10^{-13}$	$3.2328 \cdot 10^{-3}$

For this example, polynomial interpolation is very effective.

2. The following table shows the maximum errors obtained in approximating $f(x) = 1/(1 + x^2)$ on the interval $[-5, 5]$ by polynomial interpolation and by piecewise linear interpolation, each on a uniform grid with n nodes.

n	Poly. interp. err.	PW linear interp. err.
1	$9.6154 \cdot 10^{-1}$	$9.6154 \cdot 10^{-1}$
2	$6.4615 \cdot 10^{-1}$	$4.1811 \cdot 10^{-1}$
3	$7.0701 \cdot 10^{-1}$	$7.3529 \cdot 10^{-1}$
4	$4.3836 \cdot 10^{-1}$	$1.8021 \cdot 10^{-1}$
5	$4.3269 \cdot 10^{-1}$	$5.0000 \cdot 10^{-1}$
6	$6.1695 \cdot 10^{-1}$	$6.2304 \cdot 10^{-2}$
7	$2.4736 \cdot 10^{-1}$	$3.3784 \cdot 10^{-1}$
8	1.0452	$6.3898 \cdot 10^{-2}$
9	$3.0028 \cdot 10^{-1}$	$2.3585 \cdot 10^{-1}$
10	1.9156	$6.7431 \cdot 10^{-2}$

Here we see that polynomial interpolation is not effective (cf. Figure 2.6 in the text). Also, we see that piecewise linear interpolation is much more effective with n even, since then there is an interpolation node at $x = 0$, which is the peak of the graph of f.

3. We now repeat the previous two exercises, using piecewise quadratic interpolation instead of piecewise linear interpolation. We use a uniform grid of $2n + 1$ nodes.

   (a) Here the function is $f(x) = e^x$ on $[0, 1]$.

n	Poly. interp. err.	PW quad. interp. err.
1	$1.4416 \cdot 10^{-2}$	$1.4416 \cdot 10^{-2}$
2	$5.2637 \cdot 10^{-5}$	$2.2079 \cdot 10^{-3}$
3	$1.1921 \cdot 10^{-7}$	$7.0115 \cdot 10^{-4}$
4	$1.7565 \cdot 10^{-10}$	$3.0632 \cdot 10^{-4}$
5	$1.8119 \cdot 10^{-13}$	$1.6017 \cdot 10^{-4}$

   Once again, polynomial interpolation is quite effective for this function. We only go to $n = 5$, since for larger n we reach the limits of the finite precision arithmetic (in standard double precision).

   (b) Here $f(x) = 1/(1 + x^2)$ on $[-5, 5]$.

n	Poly. interp. err.	PW quad. interp. err.
1	$6.4615 \cdot 10^{-1}$	$6.4615 \cdot 10^{-1}$
2	$4.3818 \cdot 10^{-1}$	$8.5472 \cdot 10^{-2}$
3	$6.1667 \cdot 10^{-1}$	$2.3570 \cdot 10^{-1}$
4	1.0452	$9.7631 \cdot 10^{-2}$
5	1.9156	$8.5779 \cdot 10^{-2}$
6	3.6629	$7.4693 \cdot 10^{-2}$
7	7.1920	$3.4671 \cdot 10^{-2}$
8	$1.4392 \cdot 10^{1}$	$4.7781 \cdot 10^{-2}$

   For $n > 8$, it is not possible to solve the linear system defining the polynomial interpolant accurately in standard double precision arithmetic. Once again, for this example, we see the advantage of using piecewise polynomials to approximate f.

4. Let $x_0, x_1, \ldots, x_n$ define a uniform mesh on $[a, b]$ (that is, $x_i = a+ih$, $i = 0, 1, \ldots, n$, where $h = (b-a)/n$). We wish to prove that

$$\max_{x \in [a,b]} \frac{|(x - x_0)(x - x_1) \cdots (x - x_n)|}{(n + 1)!} \leq \frac{h^{n+1}}{2(n + 1)}.$$

   Let $x \in [a, b]$ be given. If x equals any of the nodes $x_0, x_1, \ldots, x_n$, then the left-hand side is zero, and thus the inequality holds. So let us assume that $x \in (x_{i-1}, x_i)$ for some $i = 1, 2, \ldots, n$. The distance from x to the nearer of x_{i-1}, x_i is at most $h/2$ and the distance to the further of the two is at most h. If we then list the remaining $n - 1$ nodes in order of distance from x (nearest to furthest), we see that the

distances are at most $2h, 3h, \ldots, nh$. Therefore,

$$\begin{aligned}|(x-x_0)(x-x_1)\cdots(x-x_n)| = \ & |x-x_0||x-x_1|\cdots|x-x_n| \\ & \leq \left(\frac{h}{2}\right)(h)(2h)\cdots(nh) = \frac{1}{2}n!h^{n+1}.\end{aligned}$$

Therefore,

$$\frac{|(x-x_0)(x-x_1)\cdots(x-x_n)|}{(n+1)!} \leq \frac{1}{2}\frac{n!h^{n+1}}{(n+1)!}\ \frac{h^{n+1}}{2(n+1)}.$$

Since this holds for each $x \in [a,b]$, the proof is complete.

5. We wish to derive a bound on the error in piecewise quadratic interpolation in the case of a uniform mesh. We use the notation of the text, and consider an arbitrary element $[x_{2i-2}, x_{2i}]$ (with $x_{2i} - x_{2i-2} = h$). Let f belonging to $C^3[a,b]$ be approximated on this element by the quadratic q that interpolates f at $x_{2i-2}, x_{2i-1}, x_{2i}$. Then the remainder term is given by

$$f(x) - q(x) = \frac{f^{(3)}(c_x)}{3!}(x-x_{2i-2})(x-x_{2i-1})(x-x_{2i}),\ x \in [a,b],$$

where $c_x \in [a,b]$. We assume $|f^{(3)}(x)| \leq M$ for all $x \in [a,b]$. We must maximize

$$|(x-x_{2i-2})(x-x_{2i-1})(x-x_{2i})|$$

on $[a,b]$. This is equivalent, by a simple change of variables, to maximizing $|p(x)|$ on $[0,h]$, where $p(x) = x(x-h/2)(x-h)$. This is a simple problem in single-variable calculus, and we can easily verify that

$$|p(x)| \leq \frac{h^3}{12\sqrt{3}} \text{ for all } x \in [0,h].$$

Therefore,

$$|f(x) - q(x)| \leq \frac{|f^{(3)}(c_x)|}{3!} \cdot \frac{h^3}{12\sqrt{3}} = \frac{|f^{(3)}(c_x)|}{72\sqrt{3}}h^3 \text{ for all } x \in [x_{2i-2}, x_{2i}].$$

Since this is valid over every element in the mesh, we obtain

$$|f(x) - q(x)| \leq \frac{M}{72\sqrt{3}}h^3 \text{ for all } x \in [a,b].$$