SOLUTIONS MANUAL TO

Advanced Linear Algebra

by

Bruce N. Cooperstein

CRC Press
Taylor & Francis Group

SOLUTIONS MANUAL TO
Advanced Linear Algebra

by

Bruce N. Cooperstein

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper
Version Date: 20150506

International Standard Book Number-13: 978-1-4822-4888-3 (Ancillary)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Chapter 1

Vector Spaces

1.1. Fields

1. Let $z = a + bi, w = c + di$ with $a, b, c, d \in \mathbb{R}$. Then $|z|^2 = a^2 + b^2, |w|^2 = c^2 + d^2$ and $|z|^2|w|^2 = (a^2 + b^2)(c^2 + d^2) = a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2$. On the other hand, $zw = (ac - bd) + (ad + bc)I$ and $|zw|^2 = (ac - bd)^2 + (ad + bc)^2 = a^2c^2 + b^2d^2 - 2abcd + a^2d^2 + b^2c^2 + 2abcd = a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2$.

2. Part ii) of Theorem (1.1) follows from part i).

Part iii) of Theorem (1.1) $z\overline{z} = (a + bi)(a - bi) = a^2 - abi + (bi)a - (bi)(bi) = a^2 + b^2 = |z|^2$.

3. Since $\mathbb{C}$ is a field and $\mathbb{Q}[i]$ is a subset of $\mathbb{C}$ it suffices to prove the following:

(i) If $x, y \in \mathbb{Q}[i]$ then $x + y \in \mathbb{Q}[i]$;

(ii) If $x \in \mathbb{Q}[i]$ then $-x \in \mathbb{Q}[i]$;

(iii) If $x, y \in \mathbb{Q}[i]$ then $xy \in \mathbb{Q}[i]$; and

(iv) If $x \in \mathbb{Q}[i], x \neq 0$ then $\frac{1}{x} \in \mathbb{Q}[i]$.

(i) We can write $x = a + bi, y = c + di$ where $a, b, c, d \in \mathbb{Q}$. Then $a + c, b + d \in \mathbb{Q}$ and consequently, $x + y = (a + c) + (b + d)i \in \mathbb{Q}[i]$.

(ii) If $x = a + bi, a, b \in \mathbb{Q}$ then $-a, -b \in \mathbb{Q}$ and $-(a + bi) = -a - bi \in \mathbb{Q}[i]$.

(iii) If $x = a + bi, y = c + di$ with $a, b, c, d \in \mathbb{Q}$ then $ac, ad, bc, bd \in \mathbb{Q}$ and therefore $(a + bi)(c + di) = [ac - bd] + [ad + bc]i \in \mathbb{Q}[i]$.

(iv) Assume that $x = a + bi$ with $a, b \in \mathbb{Q}$ not both zero. Then in $\mathbb{C}$ we have $(a + bi)^{-1} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i$. However, if $a, b \in \mathbb{Q}$ then $a^2 + b^2 \in \mathbb{Q}$ whence $\frac{a}{a^2+b^2}$ and $\frac{b}{a^2+b^2} \in \mathbb{Q}$.

4. Write $z = a + bi, w = c + di$. Then $z + w = (a + c) + (b + d)i$. Then $\overline{z + w} = (a + c) - (b + d)i$. On the other hand, $\overline{z} = a - bi, \overline{w} = c - di$. Then $\overline{z} + \overline{w} = (a - bi) + (c - di) = (a + c) + [(-b) + (-d)]i = (a + c) + (-b - d)i = (a + c) - (b + d)i$.

5. This follows from part ii. of Theorem (1.1.1) since if c is a real number then $\overline{c} = c$. Therefore $\overline{cz} = \overline{c}\,\overline{z} = c\overline{z}$.

6. a) The addition table is symmetric which implies that addition is commutative. Likewise the multiplication table is symmetric from which we conclude that multiplication is commutative.

b) The entry in the row indexed by 0 and the column indexed the element i is i for $i \in \{0, 1, 2, 3, 4\}$.

c) Every row of the addition table contains a 0. If the row is headed by the element a and the column in which the 0 occurs is indexed by b then $a + b = 0$. This establishes the existence of a negative of a with respect to 0. Note that since there is only one zero in each row and column, the negative is unique.

d) The entry in the row indexed by 1 and the column indexed the element i is i for $i \in \{1, 2, 3, 4\}$.

e) Every row has a 1 in it. If the row is headed by the element a and the column in which the 1 occurs is indexed by c then $ac = 1$. This establishes the existence of a multiplicative inverse of a with respect to 1. Note that since each row and column has only a single 1, the multiplicative inverse of a non-zero element with respect to 1 is unique.

7. The additive inverse (negative) of 2 is 3. So add 3 to both sides

$$(3x+2)+3=4+3$$

$$3x+(2+3)=2$$

$$3x+0=2$$

$$3x=2$$

Now multiply by 2 since $2\cdot 3=1$.

$$2(3x)=2\cdot 2$$

$$(2\cdot 3)x=4$$

$$1\cdot x=4$$

$$x=4.$$

The unique solution is $x=4$.

8.

$$2x-(1+2i)=-ix+(2+3i)$$

Add $1+2i$ to both sides to obtain the equation

$$2x=-ix+(3+4i)$$

Add ix from both sides to obtain

$$(2+i)x=3+4i$$

Divide by $2+i$ to obtain the equation

$$x=\frac{3+4i}{2+i}$$

After multiplying the complex number on the right hand side by $\frac{2-i}{2-i}$ we obtain

$$x=\frac{(3+4i)(2-i)}{(2+i)(2-i)}$$

After performing the arithmetic on the right hand side we get

$$x=2+i$$

9. Existence of an additive inverse, associativity of addition, the neutral character of 0, the existence of multiplicative inverses for all non-zero elements, the associativity of multiplication, the neutral character of 1 with respect to multiplication.

1.2. The Space $\mathbb{F}^n$

1. $\begin{pmatrix} 2i \\ -2+2i \\ 4-2i \end{pmatrix}$

2. $\begin{pmatrix} 2 \\ 6 \\ -4 \end{pmatrix}$

3. $\begin{pmatrix} -6i \\ 2i \\ 8i \end{pmatrix}$

4. $\begin{pmatrix} 1+3i \\ 2 \\ -1+i \end{pmatrix}$

5. $\begin{pmatrix} -3+2i \\ -2-i \\ 1 \end{pmatrix}$

6. $\begin{pmatrix} 1+2i \\ 3+i \\ 5 \end{pmatrix}$

7. $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

8. $\begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}$

9. $\begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}$

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

11. $\boldsymbol{v} = \begin{pmatrix} 3-i \\ 3+i \end{pmatrix}$

12. $\boldsymbol{v} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

1.3. Introduction to Vector Spaces

1. Set $\boldsymbol{x} = \boldsymbol{0}$. Then $\boldsymbol{x} + \boldsymbol{x} = \boldsymbol{0} + \boldsymbol{x} = \boldsymbol{x}$. Now multiply by the scalar c to get

$$c(\boldsymbol{x} + \boldsymbol{x}) = c\boldsymbol{x}$$

After distributing we get

$$c\boldsymbol{x} + c\boldsymbol{x} = c\boldsymbol{x}.$$

Set $\boldsymbol{y} = -(c\boldsymbol{x})$ and add to both sides of the equation:

$$\boldsymbol{y} + (c\boldsymbol{x} + c\boldsymbol{x}) = \boldsymbol{y} + c\boldsymbol{x}$$

Use associativity on the left hand side to get

$$(\boldsymbol{y} + c\boldsymbol{x}) + c\boldsymbol{x} = \boldsymbol{y} + c\boldsymbol{x} = \boldsymbol{0}$$

$$\boldsymbol{0} + c\boldsymbol{x} = \boldsymbol{0}$$

$$c\boldsymbol{x} = \boldsymbol{0}.$$

2. Assume $c\boldsymbol{u} = \boldsymbol{0}$ but $c \neq 0$. We prove that $\boldsymbol{u} = \boldsymbol{0}$. Multiply by $\frac{1}{c}$ to get $\frac{1}{c}(c\boldsymbol{u}) = \frac{1}{c}\boldsymbol{0} = \boldsymbol{0}$ by part iii) of Theorem (1.4). On the other hand, $\frac{1}{c}(c\boldsymbol{u}) = (\frac{1}{c}c)\boldsymbol{u} = 1\boldsymbol{u} = \boldsymbol{u}$. Thus, $\boldsymbol{u} = \boldsymbol{0}$.

3. Since $\boldsymbol{v} + (-\boldsymbol{v}) = \boldsymbol{0}$ it follows that $\boldsymbol{v}$ is the negative of $-\boldsymbol{v}$, that is, $-(-\boldsymbol{v}) = \boldsymbol{v}$.

4. Add the negative, $-\boldsymbol{v}$, of the vector $\boldsymbol{v}$ to both sides of the equation

$$\boldsymbol{v} + \boldsymbol{x} = \boldsymbol{v} + \boldsymbol{y}$$

to obtain

$$(-\boldsymbol{v}) + [\boldsymbol{v} + \boldsymbol{x}] = (-\boldsymbol{v}) + [\boldsymbol{v} + \boldsymbol{y}].$$

By the associativity property we have

$$[(-\boldsymbol{v}) + \boldsymbol{v}] + \boldsymbol{x} = [(-\boldsymbol{v}) + \boldsymbol{v}] + \boldsymbol{y}.$$

Now use the axiom that states $(-\boldsymbol{v}) + \boldsymbol{v} = \boldsymbol{0}$ to get

$$\boldsymbol{0} + \boldsymbol{x} = \boldsymbol{0} + \boldsymbol{y}$$

Since $\boldsymbol{0} + \boldsymbol{x} = \boldsymbol{x}$ and $\boldsymbol{0} + \boldsymbol{y} = \boldsymbol{y}$ we conclude that

$$\boldsymbol{x} = \boldsymbol{y}.$$

5. Multiply on the left hand side by the scalar $\frac{1}{c}$:

$$\frac{1}{c}(c\boldsymbol{x}) = \frac{1}{c}(c\boldsymbol{y}).$$

Now make use of (M3) to get

$$[\frac{1}{c}c]\boldsymbol{x} = [\frac{1}{c}c]\boldsymbol{y}.$$

Since $\frac{1}{c}c = 1$ we get

$$1\boldsymbol{x} = 1\boldsymbol{y}$$

and then by (M4) we conclude

$$\boldsymbol{x} = \boldsymbol{y}.$$

6. Let $f, g, h \in \mathcal{M}(X, \mathbb{F})$ and $a, b \in \mathbb{F}$ be scalars. We show all the axioms hold.

(A1) For any $x \in X, (f+g)(x) = f(x)+g(x)$. However, addition in $\mathbb{F}$ is commutative and therefore $f(x)+g(x) = g(x) + f(x) = (g+f)(x)$. Thus, the functions $f+g$ and $g+f$ are identical.

(A2) For any $x \in X, [(f+g)+h](x) = (f+g)(x) + h(x) = [f(x)+g(x)] + h(x)$. This is just a sum of elements in $\mathbb{F}$. Addition in $\mathbb{F}$ is associative and therefore $[f(x)+g(x)]+h(x) = f(x)+[g(x)+h(x)] = f(x)+(g+h)(x) = [f+(g+h)](x)$. Thus, the functions $(f+g)+h$ and $f+(g+h)$ are identical.

(A3) $\mathbf{O}$ is an identity for addition: $(\mathbf{O}+f)(x) = \mathbf{O}(x) + f(x) = 0 + f(x) = f(x)$ and consequently, $\mathbf{O}+f = f$.

(A4) Let $-f$ denote the function from X to $\mathbb{F}$ such that $(-f)(x) = -f(x)$. Then $[(-f)+f](x) = (-f)(x) + f(x) = -f(x)+f(x) = 0$ and therefore $(-f)+f = \mathbf{O}$.

(M1) $[a(f+g)](x) = a[f+g)(x)] = a[f(x)+g(x)]$. Now, $a, f(x), g(x)$ are elements of $\mathbb{F}$ and the distributive axiom holds in $\mathbb{F}$ and therefore $a[f(x)+g(x)] = af(x) + ag(x) = (af)(x) + (ag)(x) = [(af)+(ag)](x)$. Thus, the functions $a(f+g)$ and $(af)+(ag)$ are identical.

(M2) $[(a+b)f](x) = (a+b)f(x)$. Now a, b and $f(x)$ are all elements of the field $\mathbb{F}$ where the distributive axiom holds. Therefore $(a+b)f(x) = af(x)+bf(x) = (af)(x)+(bf)(x) = [af+bf](x)$. This shows that the functions $(a+b)f$ and $af+bf$ are identical as required.

(M3) $[(ab)f](x) = (ab)f(x)$. Since $a, b, f(x)$ are in $\mathbb{F}$ and the multiplication in $\mathbb{F}$ is associative we have $(ab)f(x) = a[bf(x)] = a[(bf)(x)] = [a(bf)](x)$. Thus, the functions $(ab)f$ and $a(bf)$ are equal.

(M4) $(1f)(x) = 1f(x) = f(x)$ so $1f = f$.

7. Let $f, g, h \in \mathcal{M}(X, V)$ and $a, b \in \mathbb{F}$ be scalars. We show all the axioms hold.

(A1) For any $x \in X, (f+g)(x) = f(x)+g(x)$. However, addition in V is commutative and therefore $f(x)+g(x) = g(x)+f(x) = (g+f)(x)$. Thus, the functions $f+g$ and $g+f$ are identical.

(A2) For any $x \in X, [(f+g)+h](x) = (f+g)(x) + h(x) = [f(x)+g(x)] + h(x)$. This is just a sum of elements in V. Addition in V is associative and therefore $[f(x)+g(x)]+h(x) = f(x)+[g(x)+h(x)] = f(x)+(g+h)(x) = [f+(g+h)](x)$. Thus, the functions $(f+g)+h$ and $f+(g+h)$ are identical.

(A3) $\mathbf{O}$ is an identity for addition: $(\mathbf{O}+f)(x) = \mathbf{O}(x) + f(x) = \mathbf{0} + f(x) = f(x)$ and consequently, $\mathbf{O}+f = f$.

(A4) Let $-f$ denote the function from X to V such that $(-f)(x) = -f(x)$. Then $[(-f)+f](x) = (-f)(x) + f(x) = -f(x)+f(x) = \mathbf{0}$ and therefore $(-f)+f = \mathbf{O}$.

(M1) $[a(f+g)](x) = a[f+g)(x)] = a[f(x)+g(x)]$. Now, a is a scalar and $f(x), g(x) \in V$. By (M1) applied to V we have $a[f(x)+g(x)] = af(x)+ag(x) = (af)(x) + (ag)(x) = [af+ag](x)$. Therefore the functions $a(f+g)$ and $af+ag$ are equal.

(M2) $[(a+b)f](x) = (a+b)f(x)$. Now a, b are scalars and $f(x) \in V$. By axiom (M2) applied to V we have $(a+b)(f(x) = af(x)+bf(x) = (af)(x)+(bf)(x) = [af+bf](x)$. Thus, the functions $(a+b)f$ and $af+bf$ are equal.

(M3) $[(ab)f](x) = (ab)f(x)$. Since $a, b \in \mathbb{F}$ and $f(x) \in V$ we can apply (M3) for V and conclude that $[(ab)f](x) = (ab)f(x) = a[bf(x)] = a[(bf)(x)] = [a(bf)](x)$. Thus, $(ab)f = a(bf)$ as required.

(M4) Finally, $(1f)(x) = 1f(x) = f(x)$ by (M4) applied to V and therefore $1f = f$.

8. We demonstrate the axioms all hold. So let $\boldsymbol{u}, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3 \in U, \boldsymbol{v}, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \in V$ and $c, d \in \mathbb{F}$. We then have

$$(\boldsymbol{u}_1, \boldsymbol{v}_1) + (\boldsymbol{u}_2, \boldsymbol{v}_2) = (\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{v}_1 + \boldsymbol{v}_2).$$

$$(\boldsymbol{u}_2, \boldsymbol{v}_2) + (\boldsymbol{u}_1, \boldsymbol{v}_1) = (\boldsymbol{u}_2 + \boldsymbol{u}_1, \boldsymbol{v}_2 + \boldsymbol{v}_1).$$

Since addition in U and addition in V is commutative, $\boldsymbol{u}_1 + \boldsymbol{u}_2 = \boldsymbol{u}_2 + \boldsymbol{u}_1, \boldsymbol{v}_1 + \boldsymbol{v}_2 = \boldsymbol{v}_2 + \boldsymbol{v}_1$. Thus, the vectors are equal. This establishes (A1).

$$[(\boldsymbol{u}_1, \boldsymbol{v}_1) + (\boldsymbol{u}_2, \boldsymbol{v}_2)] + (\boldsymbol{u}_3, \boldsymbol{v}_3) =$$
$$([\boldsymbol{u}_1 + \boldsymbol{u}_2] + \boldsymbol{u}_3, [\boldsymbol{v}_1 + \boldsymbol{v}_2] + \boldsymbol{v}_3).$$
$$(\boldsymbol{u}_1, \boldsymbol{v}_1) + [(\boldsymbol{u}_2, \boldsymbol{v}_2) + (\boldsymbol{u}_3, \boldsymbol{v}_3)] =$$
$$([\boldsymbol{u}_1 + [\boldsymbol{u}_2 + \boldsymbol{u}_3\, \boldsymbol{v}_1 + [\boldsymbol{v}_2 + \boldsymbol{v}_3]).$$

Since addition in U and addition in V is associative, the results are identical. Thus, (A2) holds.

$$(\boldsymbol{u}, \boldsymbol{v}) + (\mathbf{0}_U, \mathbf{0}_W) = (\boldsymbol{u} + \mathbf{0}_U, \boldsymbol{v} + \mathbf{0}_V) = (\boldsymbol{u}, \boldsymbol{v}).$$

So, indeed $(\mathbf{0}_U, \mathbf{0}_V)$ is an identity for $U \times V$ with the addition as defined.

Moreover,

$$(\boldsymbol{u}, \boldsymbol{v}) + (-\boldsymbol{u}, -\boldsymbol{v}) = (\boldsymbol{u} + (-\boldsymbol{u}), \boldsymbol{v} + (-\boldsymbol{v})) =$$
$$(\boldsymbol{u} + [-\boldsymbol{u}], \boldsymbol{v} + [-\boldsymbol{v}]) = (\mathbf{0}_U, \mathbf{0}_V).$$

$$c[(\boldsymbol{u}_1, \boldsymbol{v}_1) + (\boldsymbol{u}_2, \boldsymbol{v}_2)] = c(\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{v}_1 + \boldsymbol{v}_2) =$$
$$(c[\boldsymbol{u}_1 + \boldsymbol{u}_2], c[\boldsymbol{v}_1 + \boldsymbol{v}_2]) = (c\boldsymbol{u}_1 + c\boldsymbol{u}_2, c\boldsymbol{v}_1 + c\boldsymbol{v}_2)$$

since the distributive property holds in U and V. On the other hand,

$$(c\boldsymbol{u}_1 + c\boldsymbol{u}_2, c\boldsymbol{v}_1 + c\boldsymbol{v}_2) =$$
$$(c\boldsymbol{u}_1, c\boldsymbol{v}_1) + (c\boldsymbol{u}_2, c\boldsymbol{v}_2) =$$
$$c(\boldsymbol{u}_1, \boldsymbol{v}_1) + c(\boldsymbol{u}_2, \boldsymbol{v}_2).$$

For the second distributive property we have

$$[c+d](\boldsymbol{u}, \boldsymbol{v}) = ([c+d]\boldsymbol{u}, [c+d]\boldsymbol{v}) = (c\boldsymbol{u}+d\boldsymbol{u}, c\boldsymbol{v}+d\boldsymbol{v}) =$$

$$(c\boldsymbol{u}, c\boldsymbol{v}) + (d\boldsymbol{u}, d\boldsymbol{v}) = c(\boldsymbol{u}, \boldsymbol{v}) + d(\boldsymbol{u}, \boldsymbol{v})$$

$$[cd](\boldsymbol{u}, \boldsymbol{v}) = ([cd]\boldsymbol{u}, [cd]\boldsymbol{v}) = (c[d\boldsymbol{u}], c[d\boldsymbol{v}]) =$$

$$c(d\boldsymbol{u}, d\boldsymbol{v}) = c[d\boldsymbol{u}].$$

Finally,

$$1(\boldsymbol{u}, \boldsymbol{v}) = (1\boldsymbol{u}, 1\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v}).$$

9. Let $f, g, h \in \prod_{i \in I} U_i, a, b \in \mathbb{F}$.

(A1) For $i \in I, (f+g)(i) = f(i)+g(i)$. Now $f(i), g(i) \in U_i$ and addition in U_i is commutative and so $f(i)+g(i) = g(i)+f(i) = (g+f)(i)$. Since $i \in I$ is arbitrary, $f+g = g+f$.

(A2) For $i \in I, [(f+g)+h](i) = (f+g)(i)+h(i) = [f(i)+g(i)]+h(i)$. Now $f(i), g(i), h(i) \in U_i$ and addition in U_i is associative. Therefore, $[f(i)+g(i)]+h(i) = f(i)+[g(i)+h(i)] = f(i)+[g+h](i) = [f_(g+h)](i)$. Since $i \in I$ is arbitrary, $(f+g)+h = f+(g+h)$.

(A3) $[\mathbf{O}+f])i) = \mathbf{O}(i)+f(i) = \mathbf{0}_i+f(i) = f(i)$. So, $\mathbf{O}+f = f$.

(A4) Let $-f$ denote the function from I to $\cup_{i \in I}$ such that $(-f)(i) = -f(i)$ for $i \in I$. Then $[(-f)+f](i) = (-f)(i)+f(i) = -f(i)+f(i) = \mathbf{0}_i$ and therefore $(-f)+f = \mathbf{O}$.

(M1) Since (M1) applies to U_i we have $[a(f+g)](i) = a[(f+g)(i)] = a[f(i)+g(i)] = af(i)+ag(i) = (af)(i)+(ag)(i) = [(af)+(ag)](i) = [af+ag](i)$. Thus, the functions $a(f+g)$ and $af+ag$ are equal.

(M2) $[(a+b)f](i) = (a+b)f(i)$. Now a, b are scalars and $f(i) \in U_i$. By axiom (M2) applied to U_i we have $(a+b)(f(i) = af(i)+bf(i) = (af)(i)+(bf)(i) = [af+bf](i)$. Thus, the functions $(a+b)f$ and $af+bf$ are equal.

(M3) $[(ab)f](i) = (ab)f(i)$. Since $a, b \in \mathbb{F}$ and $f(i) \in U_i$ we can apply (M3) for U_i and conclude that $[(ab)f](i) = (ab)f(i) = a[bf(i)] = a[(bf)(i)] = [a(bf)](i)$. Thus, $(ab)f = a(bf)$ as required.

(M4) Finally, $(1f)(i) = 1f(i) = f(i)$ by (M4) applied to U_i. Since $i \in I$ is arbitrary, $1f = f$.

10. (A1) Since for any two sets U, W we have $U \ominus W = W \ominus U$ we have $U + W = W + U$.

(A2) We remark that for three sets U, W, Z that $[U \ominus W] \ominus Z$ consists of those elements of $U \cup W \cup Z$ which are contained in 1 or 3 of these sets. This is also true of $U \ominus [W \ominus Z]$ and therefore we have $[U+W]+Z = U+[W+Z]$.

(A3) $\emptyset + U = (\emptyset \cup U) \setminus (\emptyset \cap U) = U \setminus \emptyset = U$. So, $\emptyset$ does act like a zero vector.

(A4) $U + U = (U \cup U) \setminus (U \cap U) = U \setminus U = \emptyset$. So, indeed, U is the negative of U.

(M1) $0 \cdot (U+W) = \emptyset$. Also, $0 \cdot U = 0 \cdot W = \emptyset$ and $\emptyset + \emptyset = \emptyset$.

$1 \cdot (U+W) = U+W$ and $1 \cdot U = W, 1 \cdot W = W$ and so $1 \cdot U + 1 \cdot W = U + W = 1 \cdot (U+W)$.

(M2) If $a = b$ then $a \cdot U = b \cdot U$ and $a \cdot U + b \cdot U = a \cdot U + a \cdot U = \emptyset$. On the other hand, $a + b = 0$ and $0 \cdot U = \emptyset$. Therefore we may assume $a = 0, b = 1$. Then $(a+b) \cdot U = 1 \cdot U = U$. On the other hand, $0 \cdot U + 1 \cdot U = \emptyset + U = U$.

(M3) If either $a = 0$ or $b = 0$ then both $(ab) \cdot U = \emptyset$ and $a \cdot (b \cdot U) = \emptyset$. Therefore we may assume $a = b = 1$. Clearly, $1 \cdot U = U = 1 \cdot (1 \cdot U)$.

(M4) This holds by the definition of $1 \cdot U$.

11. (A1) Since multiplication in $\mathbb{R}^+$ is commutative we have

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} =$$

$$\begin{pmatrix} a_1a_2 \\ b_1\boldsymbol{b}_2 \end{pmatrix} = \begin{pmatrix} a_2a_1 \\ b_2b_1 \end{pmatrix} =$$

$$\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} + \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}.$$

(A2) Since multiplication in $\mathbb{R}^+$ is associative we have

$$[\begin{pmatrix} a1 \\ b_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}] + \begin{pmatrix} a_3 \\ b_3 \end{pmatrix} =$$

$$\begin{pmatrix} a_1a_2 \\ b_1b_2 \end{pmatrix} + \begin{pmatrix} a_3 \\ b_3 \end{pmatrix} =$$

$$\begin{pmatrix} (a_1a_2)a_3 \\ (b_1b_2)b_3 \end{pmatrix} = \begin{pmatrix} a_1(a_2a_3) \\ b_1(b_2b_3) \end{pmatrix} =$$

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + [\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} + \begin{pmatrix} a_2 \\ b_3 \end{pmatrix}].$$

(A3) $\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1a \\ 1b \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$

(A4)

$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} \frac{1}{a} \\ \frac{1}{b} \end{pmatrix} =$$

$$\begin{pmatrix} a\frac{1}{a} \\ b\frac{1}{b} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(M1)

$$c[\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}] =$$

$$c\begin{pmatrix} a_1a_2 \\ b_1b_2 \end{pmatrix} = \begin{pmatrix} (a_1a_2)^c \\ (b_1b_2)^c \end{pmatrix} =$$

$$\begin{pmatrix} a_1^c a_2^c \\ b_1^c b_2^c \end{pmatrix} =$$

$$\begin{pmatrix} a_1^c \\ b_1^c \end{pmatrix} + \begin{pmatrix} a_2^c \\ b_2^c \end{pmatrix} =$$

$$c\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} + c\begin{pmatrix} a_2 \\ b_2 \end{pmatrix}.$$

(M2)

$$(c+d)\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a^{c+d} \\ b^{c+d} \end{pmatrix} =$$

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

$$c\begin{pmatrix} a \\ b \end{pmatrix} + d\begin{pmatrix} a \\ b \end{pmatrix}$$

(M3)

$$(cd)\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a^{cd} \\ b^{cd} \end{pmatrix} =$$

$$\begin{pmatrix} (a^d)^c \\ (b^d)^c \end{pmatrix} = c\begin{pmatrix} a^d \\ b^d \end{pmatrix} = c(d\begin{pmatrix} a \\ b \end{pmatrix}).$$

(M4) $1\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a^1 \\ b^1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$

1.4. Subspaces of Vector Spaces

1. In order for W to contain the zero vector there must exist a, b such that

$$\begin{array}{rcrcrcl} 2a & - & 3b & + & 1 & = & 0 \\ -2a & + & 5b & & & = & 0 \\ 2a & + & b & & & = & 0 \end{array}$$

which is equivalent to the linear system

$$\begin{array}{rcrcr} 2a & - & 3b & = & -1 \\ -2a & + & 5b & = & 0 \\ 2a & + & b & = & 0 \end{array}$$

which is inconsistent. Thus, W is not a subspace.

2. We show that W is not a subspace by proving it is not closed under scalar multiplication. Specifically, W contains the vector $f(1,1) = 1+2X^2$. However, we claim the vector $2+4X^2$ does not belong to W. If it did then there exist a, b such that $f(a,b) = ab+(a-b)X+(a+b)X^2 = 2+4X^2$. We must then have

$$\begin{array}{rcrcl} a & - & b & = & 0 \\ a & + & b & = & 4 \end{array}.$$

This has the unique solution $a = b = 2$. However, then $ab = 4 \neq 2$ as required.

3. Suppose $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in W$ so that $3x - 2y + 4z = 0$. Then

$$3(cx) - 2(cy) + 4(cz) = (3x - 2y + 4z)c = 0c = 0$$

which implies that $\begin{pmatrix} cx \\ cy \\ cz \end{pmatrix} \in W$ and W is closed under scalar multiplication.

Suppose $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \in W$ which implies that

$$3x_1 - 2y_1 + 4z_1 = 0 = 3x_2 - 2y_2 + 4z_2$$

It then follows that

$$3(x_1 + x_2) - 2(y_1 + y_2) + 4(z_1 + z_2) =$$

$$[3x_1 - 2y_1 + 4z_1] + [3x_2 - 2y_2 + 4z_2] = 0.$$

This implies that $\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix} \in W$ and W is closed under addition. Thus, W is a subspace.

4. Set $W = \cup_U \in \mathcal{F}U$. In order to show that W is a subspace we have to show that it is closed under addition and scalar multiplication.

Closed under scalar multiplication. Assume $\boldsymbol{x} \in W$ and c is a scalar. By the definition of W there is an $X \in \mathcal{F}$ such that $\boldsymbol{x} \in X$. Since X is a subspace of V, $c\boldsymbol{x} \in X$ and hence $c\boldsymbol{x} \in \cup_{U \in \mathcal{F}} U = W$.

Closed under addition. We need to show if $\boldsymbol{x}, \boldsymbol{y} \in W$ then $\boldsymbol{x} + \boldsymbol{y} \in W$. By the definition of W there are subspaces $X, Y \in \mathcal{F}$ such that $\boldsymbol{x} \in x, \boldsymbol{y} \in Y$. By our hypothesis there exists $Z \in \mathcal{F}$ such that $X \cup Y \subset Z$. It then follows that $\boldsymbol{x}, \boldsymbol{y} \in Z$. Since Z is a subspace of V, $\boldsymbol{x} + \boldsymbol{y} \in Z$. Then $\boldsymbol{x} + \boldsymbol{y} \in W$ and W is closed under addition as claimed.

5. Assume $\boldsymbol{u} \in U, \boldsymbol{w} \in W$ and c is a scalar. Then $c(\boldsymbol{u} + \boldsymbol{w}) = c\boldsymbol{u} + c\boldsymbol{w}$. Since U is a subspace and $\boldsymbol{u} \in U$ it follows that $c\boldsymbol{u} \in U$. Similarly, $c\boldsymbol{w} \in W$. Then

$$c(\boldsymbol{u} + \boldsymbol{w}) = c\boldsymbol{u} + c\boldsymbol{w} \in U + W.$$

6. *Closed under scalar multiplication.* Assume $\boldsymbol{x} \in U \cup W$ and c is a scalar. Then either $\boldsymbol{x} \in U$ in which case $c\boldsymbol{x} \in U$ or $\boldsymbol{x} \in W$ and $c\boldsymbol{x} \in W$. In either case, $c\boldsymbol{x} \in U \cup W$ and $U \cup W$ is closed under scalar multiplication.

Not closed under addition. Since U is not a subset of W there exists $\boldsymbol{u} \in U, \boldsymbol{u} \notin W$. Since W is not a subset of U there is a $\boldsymbol{w} \in W, \boldsymbol{w} \notin U$. We claim that $\boldsymbol{u} + \boldsymbol{w} \notin U \cup W$. For suppose to the contrary that $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w} \in U \cup W$. Suppose $\boldsymbol{v} \in U$. Then $\boldsymbol{w} = \boldsymbol{v} - \boldsymbol{u}$ is the difference of two vectors in U, whence $\boldsymbol{w} \in U$ contrary to assumption. Likewise, if $\boldsymbol{v} \in W$ then $\boldsymbol{u} = \boldsymbol{v} - \boldsymbol{w} \in W$, a contradiction. So, $\boldsymbol{u} + \boldsymbol{w}$ is not in $U \cup W$ and $U \cup W$ is not a subspace.

7. Let $W = \{\begin{pmatrix} a \\ 0 \end{pmatrix} | a \in \mathbb{R}\}, X = \{\begin{pmatrix} 0 \\ b \end{pmatrix} | b \in \mathbb{R}\}$ and $Y = \{\begin{pmatrix} c \\ c \end{pmatrix} | c \in \mathbb{R}\}$.

8. Take $X = \{\begin{pmatrix} a \\ 0 \end{pmatrix} | a \in \mathbb{R}\}, Y = \{\begin{pmatrix} 0 \\ b \end{pmatrix} | b \in \mathbb{R}\}, Z = \{\begin{pmatrix} c \\ c \end{pmatrix} | c \in \mathbb{R}\}$.

9. Clearly, $\mathcal{M}_{fin}(X, \mathbb{F})$ is a subset of $\mathcal{M}(X, \mathbb{F})$ so we have to show that it is closed under addition and scalar multiplication.

Closed under scalar multiplication. Suppose $f \in \mathcal{M}_{fin}(X, \mathbb{F})$ and c is a scalar. If $c = 0$ then cf is the zero map with support equal to the empty set, which belongs to $\mathcal{M}_{fin}(X, \mathbb{F})$. On the other hand, if $c \neq 0$ then $spt(cf) = spt(f)$ is finite and so $cf \in \mathcal{M}_{fin}(X, \mathbb{F})$.

Closed under addition. Assume $f, g \in \mathcal{M}_{fin}(X, \mathbb{F})$. If $x \notin spt(f) \cup spt(g)$ the $f(x) = g(x) = 0$ and $(f + g)(x) = 0$. This implies that $spt(f+g) \subset spt(f) \cup spt(g)$ and is therefore finite. Thus, $f + g \in \mathcal{M}_{fin}(X, \mathbb{F})$.

10. Set $W = \{f \in \mathcal{M}(X, \mathbb{F}) | f(y) = 0 \ \forall y \in Y\}$. We need to show W is closed under scalar multiplication and addition.

Closed under scalar multiplication. Suppose $f \in W$ and c is scalar, $y \in Y$. Then $(cf)(y) = cf(y) = c0 = 0$. Since y is arbitrary, $cf \in W$.

Closed under addition. Suppose $f, g \in W$ and $y \in Y$. Then $(f + g)(y) = f(y) + g(y) = 0 + 0 = 0$.

11. This is clearly not a subspace since the zero vector does not belong to it.

12. *Closed under scalar multiplication.* Suppose $f \in \sum_{i \in I} U_i$ so that $spt(f)$ is finite and c is a scalar. If $c = 0$ then cf is the zero vector of $\prod_{i \in I} U_i$ which has support the empty set and is in $\sum_{i \in I} U_i$. On the other hand, if $c \neq 0$ then $spt(cf) = spt(f)$ and so is finite and $cf \in \sum_{i \in I} U_i$.

Closed under addition. Suppose $f, g \in \sum_{i \in I} U_i$. As in the proof of 14, $spt(f + g) \subset spt(f) \cup spt(g)$ and so is a finite subset. Then $f + g \in \sum_{i \in I} U_i$.

13. Assume $\boldsymbol{x} \in X \cap Y + Z$. Then there are vectors $\boldsymbol{y} \in Y, \boldsymbol{z} \in Z$ such that $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$. Then $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}$. Since $Y \subset X, \boldsymbol{y} \in X$ and because X a subspace we can conclude that $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y} \in X$. In particular, $\boldsymbol{z} \in X \cap Z$. Thus, $\boldsymbol{x} \in Y + (X \cap Z)$. This proves that $X \cap (Y + Z) \subset Y + (X \cap Z)$.

Conversely, assume that $\boldsymbol{u} \in Y + (X \cap Z)$. Write $\boldsymbol{u} = \boldsymbol{y} + \boldsymbol{z}$ where $\boldsymbol{y} \in Y$ and $\boldsymbol{z} \in X \cap Z$. Then clearly, $\boldsymbol{u} \in Y + Z$. On the other hand, since $Y \subset X, \boldsymbol{z} \in X \cap Z$, and X is a subspace, we can conclude that $\boldsymbol{u} = \boldsymbol{y} + \boldsymbol{z} \in X$. Thus, $\boldsymbol{u} \in X \cap (Y + Z)$. This proves that $Y + (X \cap Z) \subset X \cap (Y + Z)$ and therefore we have equality.

14. We need to show that $\mathcal{M}_{odd}(\mathbb{R}, \mathbb{R})$ is closed under addition and scalar multiplication.

Closed under addition. Assume $f, g \in \mathcal{M}_{odd}(\mathbb{R}, \mathbb{R})$ so that $f(-x) = -f(x), g(-x) = -g(x)$. Then $(f + g)(-x) = f(-x) + g(-x) = -f(x) + -g(x) = -(f + g)(x)$.

Closed under scalar multiplication. Assume $f \in \mathcal{M}_{odd}(\mathbb{R}, \mathbb{R})$ and $c \in \mathbb{R}$. Then $(cf)(-x) = cf(-x) = c[-f(x)] = (-cf)(x)$.

1.5. Span and Independence

1. We need to show that $Span(X) \subset Span(Y)$ and, conversely, that $Span(Y) \subset Span(X)$. Since by hypothesis, $X \subset Span(Y)$ it then follows that $Span(X) \subset Span(Span(Y)) = Span(Y)$. In exactly the same way we conclude that $Span(Y) \subset Span(X)$ and we obtain the desired equality.

2. Clearly, $\boldsymbol{u}, c\boldsymbol{u} + \boldsymbol{v}$ are both linear combinations of $\boldsymbol{u}, \boldsymbol{v}$ and consequently, $\boldsymbol{u}, c\boldsymbol{u} + \boldsymbol{v} \in Span(\boldsymbol{u}, \boldsymbol{v})$. On the other hand, $\boldsymbol{v} = (-c)\boldsymbol{u} + (c\boldsymbol{u} + \boldsymbol{v})$ is a linear combination of $\boldsymbol{u}$ and $c\boldsymbol{u} + \boldsymbol{v}$. Whence, both $\boldsymbol{u}$ and $\boldsymbol{v}$ are linear combinations of $\boldsymbol{u}, c\boldsymbol{u} + \boldsymbol{v}$ and so $\boldsymbol{u}, \boldsymbol{v} \in Span(\boldsymbol{u}, c\boldsymbol{u} + \boldsymbol{v})$. Now by exercise 1 we have the equality $Span(\boldsymbol{u}, \boldsymbol{v}) = Span(\boldsymbol{u}, c\boldsymbol{u} + \boldsymbol{v})$.

3. Clearly $c\boldsymbol{u}, \boldsymbol{v} \in Span(\boldsymbol{u}, \boldsymbol{v})$. On the other hand, $\boldsymbol{u} = (\frac{1}{c})(c\boldsymbol{u}) + 0\boldsymbol{v}$ is in $Span(c\boldsymbol{u}, \boldsymbol{v})$. Thus, $\boldsymbol{u}, \boldsymbol{v} \in Span(c\boldsymbol{u}, \boldsymbol{v})$ and, using exercise 1, we get the equality $Span(\boldsymbol{u}, \boldsymbol{v}) = Span(c\boldsymbol{u}, \boldsymbol{v})$.

4. Since $\boldsymbol{v}_1, c_{12}\boldsymbol{v}_1 + \boldsymbol{v}_2, c_{13}\boldsymbol{v}_1 + c_{23}\boldsymbol{v}_2 + \boldsymbol{v}_3$ are all linear combinations of $\boldsymbol{v}_1, \boldsymbol{v}_2 1, \boldsymbol{v}_3$ it follows that $Span(\boldsymbol{v}_1, c_{12}\boldsymbol{v}_1 + \boldsymbol{v}_2, c_{13}\boldsymbol{v}_1 + c_{23}\boldsymbol{v}_2 + \boldsymbol{v}_3\}) \subset Span(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$.

On the other hand, $\boldsymbol{v}_2 = (-c_{12})\boldsymbol{v}_1 + [c_{12}\boldsymbol{v}_1 + \boldsymbol{v}_2$

$\boldsymbol{v}_3 = (-c_{13} + c_{12}c_{23})\boldsymbol{v}_2 + (-c_{23})(c_{12}\boldsymbol{v} + 1 + \boldsymbol{v}_2) + (c_{13}\boldsymbol{v}_1 + c_{23}\boldsymbol{v}_3 + \boldsymbol{v}_3)$.

Thus, $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ are all linear combinations of $\boldsymbol{v}_1, c_{12}\boldsymbol{v}_1 + \boldsymbol{v}_{24}$ and $c_{13}\boldsymbol{v}_1 + c_{23}\boldsymbol{v}_2 + \boldsymbol{v}_3$. It now follows from exercise 1 that $Span(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3) = Span(\boldsymbol{v}_1, c_{12}\boldsymbol{v}_1, c_{13}\boldsymbol{v}_1 + c_{23}\boldsymbol{v}_2 + \boldsymbol{v}_3)$.

5. If $\boldsymbol{v} = \boldsymbol{0}$ then $1\boldsymbol{v} = \boldsymbol{0}$ and therefore $\boldsymbol{0}$ is linear dependent. On the other hand, if $c\boldsymbol{v} = \boldsymbol{0}, c \neq 0$ then $\boldsymbol{v} = \boldsymbol{0}$.

6. Assume $\boldsymbol{v} = \gamma\boldsymbol{u}$. Since $\boldsymbol{v} \neq \boldsymbol{0}, \gamma \neq 0$. Then $(-\gamma)\boldsymbol{u} + 1\boldsymbol{v} = \boldsymbol{0}$ is a non-trivial dependence relation on $(\boldsymbol{u}, \boldsymbol{v})$ and consequently, $(\boldsymbol{u}, \boldsymbol{v})$ is linearly dependent.

Conversely, assume that $(\boldsymbol{u}, \boldsymbol{v})$ is linearly dependent and neither vector is $\boldsymbol{0}$. Suppose $a\boldsymbol{u} + b\boldsymbol{v} = \boldsymbol{0}$. If $a = 0$ then $b\boldsymbol{v} = \boldsymbol{0}$ which by exercise implies that $\boldsymbol{v} = \boldsymbol{0}$, a contradiction. So, $a \neq 0$. In exactly the same way, $b \neq 0$.

Now $\boldsymbol{v} = (-\frac{a}{b})\boldsymbol{u}, \boldsymbol{u} = (-\frac{b}{a})\boldsymbol{u}$ so $\boldsymbol{v}$ is a multiple of $\boldsymbol{u}$ and $\boldsymbol{u}$ is a multiple of $\boldsymbol{v}$.

7. Multiply all the non-zero vectors by 0 and the zero vector by 1. This gives a non-trivial dependence relation and so the sequence of vectors is linearly dependent.

8. If $i < j$ and $\boldsymbol{v}_i = \boldsymbol{v}_j$ then let $c_i = 1, c_j = -1$ and set all the other scalars equal to 0. In this way we obtain a non-trivial dependence relation.

9. Extend a non-trivial dependence relation on the vectors of S_0 to a dependence relation on S by setting the scalars equal to zero for every vector $\boldsymbol{v} \in S \setminus S_0$.

10. This is logically equivalent to exercise 9. Alternatively, assume that some subsequence of S is linearly dependent. Then by exercise 9 the sequence S is dependent, a contradiction.

11. Assume $Span(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \cap Span(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l) = \{\boldsymbol{0}\}$. Suppose $c_1, \ldots, c_k, d_1, \ldots, d_l$ are scalars and

$$c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k + d_1\boldsymbol{v}_1 + \cdots + d_l\boldsymbol{v}_l = \{\boldsymbol{0}\}.$$

Then

$$c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k = (-d_1)\boldsymbol{v}_1 + \cdots + (-d_l)\boldsymbol{v}_l.$$

The vector $c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k \in Span(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ while $d_1\boldsymbol{v}_1 + \cdots + d_l\boldsymbol{v}_l \in Span(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l)$. By hypothesis the only common vector is the zero vector. Therefore

$$c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k = \boldsymbol{0}$$

$$d_1\boldsymbol{v}_1 + \cdots + d_l\boldsymbol{v}_l = \boldsymbol{0}$$

However, since $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ is linearly independent we get $c_1 = \cdots = c_k = 0$. Similarly, $d_1 = \cdots = d_l = 0$. This implies that $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_l)$ is linearly independent.

On the other hand, suppose $\boldsymbol{0} \neq \boldsymbol{w} \in Span(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \cap Span(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l)$. Then there are scalars $c_1, \ldots, c_k, d_1, \ldots, d_l$ such that

$$\boldsymbol{w} = c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k$$

$$\boldsymbol{w} = d_1\boldsymbol{v}_1 + \cdots + d_l\boldsymbol{v}_l.$$

Note that since $\boldsymbol{w} \neq \boldsymbol{0}$ at least one c_i and d_j is non-zero. Now we have

$$c_1\boldsymbol{u}_1 + \cdots + c_k\boldsymbol{u}_k + (-d_1)\boldsymbol{v}_1 + \cdots + (-d_l)\boldsymbol{v}_l = \boldsymbol{w} - \boldsymbol{w} = \boldsymbol{0}$$

is a non-trivial dependence relation and therefore $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_l)$ is linearly dependent.

12. Since $\boldsymbol{w} \in Span(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v})$ there are scalars $c_1, \ldots, c_k, d$ such that

$$\boldsymbol{w} = c_1\boldsymbol{u}_1 + \ldots c_k\boldsymbol{u}_k + d\boldsymbol{v}.$$

If $d = 0$ then $\boldsymbol{w} \in Span(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ contrary to our hypothesis. Therefore, $d \neq 0$. But then

$$\boldsymbol{v} = (-\frac{c_1}{d})\boldsymbol{u}_1 + \cdots + (-\frac{c_k}{d})\boldsymbol{u}_k + \frac{1}{d}\boldsymbol{w}.$$

Thus, $\boldsymbol{v} \in Span(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{w})$ as required.

13. Let c_1, c_2, c_3 be scalars, not all zero, such that $c_1(\boldsymbol{v}_1 + \boldsymbol{w}) + c_2(\boldsymbol{v}_2 + \boldsymbol{w}) + c_3(\boldsymbol{v}_3 + \boldsymbol{w}) = \boldsymbol{0}$. Claim $c_1 + c_2 + c_3 \neq 0$. Suppose otherwise. Then $\boldsymbol{0} = c_1(\boldsymbol{v}_1 + \boldsymbol{w}) + c_2(\boldsymbol{v}_2 + \boldsymbol{w}) + c_3(\boldsymbol{v}_3 + \boldsymbol{w}) = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3 + (c_1 + c_2 + c_3)\boldsymbol{w} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3$, contrary to the hypothesis that $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ is linearly independent. It now follows that $\boldsymbol{w} = -\frac{1}{c_1+c_2+c_3}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3) \in Span(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$.

1.6. Bases and Finite Dimensional Vector Spaces

1. Let $\mathcal{B}$ be a basis of V. $\mathcal{B}$ has 4 elements since $dim(V) = 4$.

a) Suppose $\mathcal{S}$ spans V and $|\mathcal{S}| = 3$. Since $\mathcal{B}$ is a basis, in particular, $\mathcal{B}$ is linearly independent. Now by the Exchange theorem, $3 = |\mathcal{S}| \geq |\mathcal{B}| = 4$, a contradiction.

b) Suppose $\mathcal{I}$ is a linearly independent subset of V with five elements. Since $\mathcal{B}$ is a basis it spans V. By the exchange theorem, $5 = |\mathcal{I}| \leq |\mathcal{B}| = 4$, a contradiction.

2. If $U \subset W$ then, since $dim(U) = dim(W) = 3$ we must have $U = W$ by Theorem (1.22) which contradicts the assumption that $U \neq W$. So, U is not contained in W (and similarly, W is not contained in U). It now follows that the subspace $U + W$ of V contains U as well as vectors not in U. Then $4 = dim(V) \geq dim(U + W) > dim(U) = 3$ which implies that $dim(U + W) = 4$. By part ii) of Theorem (1.22), $U + W = V$.

Since W is not contained in U we know that $U \cap W$ is a proper subset of U and hence $dim(U \cap W) < dim(U) =$

3. It therefore suffices to prove that $dim(U \cap W) \geq 2$. Let $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ be a basis for U and $(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$ be a basis for W with $\boldsymbol{w}_3 \notin U$. If $Span(\boldsymbol{w}_1, \boldsymbol{w}_3) \cap U = \{\boldsymbol{0}\}$ then by Exercise (1.5.11) the sequence $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{w}_1, \boldsymbol{w}_3)$ is linearly independent which contradicts the assumption that $dim(V) = 4$. Thus, $Span(\boldsymbol{w}_1, \boldsymbol{w}_3) \cap U \neq \{\boldsymbol{0}\}$. Since $\boldsymbol{w}_1 \notin U$ any non-zero vector $a\boldsymbol{w}_1 + b\boldsymbol{w}_3$ which is in U must have $a \neq 0$. By multiplying by $\frac{1}{a}$ we can say there is an $c \in \mathbb{F}$ such that $\boldsymbol{w}_1 + c\boldsymbol{w} \in U$. In exactly the same way we can conclude that there must be a $d \in \mathbb{F}$ such that $\boldsymbol{w}_2 + d\boldsymbol{w}_3 \in U$. We claim that $(\boldsymbol{w}_1 + c\boldsymbol{w}_3, \boldsymbol{w}_2 + d\boldsymbol{w}_3)$ is linearly independent. Suppose $e_1, e_2 \in \mathbb{F}$ and

$$e_1(\boldsymbol{w}_1 + c\boldsymbol{w}_3) + e_2(\boldsymbol{w}_2 + d\boldsymbol{w}_3) = \boldsymbol{0}$$

Then $e_1\boldsymbol{w}_1 + e_2\boldsymbol{w}_2 + (ce_1 + de_2)\boldsymbol{w}_3 = \boldsymbol{0}$. However, since $(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$ is linearly independent we must have $e_1 = e_2 = 0$.

Since $(\boldsymbol{w}_1 + c\boldsymbol{w}_3, \boldsymbol{w}_2 + d\boldsymbol{w}_3)$ is a linearly independent sequence from $U \cap W$ we conclude that $dim(U \cap W) \geq 2$.

3. By Exercise (1.5.11) the sequence $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$ is linearly independent and so it suffices to show that $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$ is a spanning sequence. Toward that end consider an arbitrary vector $\boldsymbol{v} \in U + W$. By the definition of $U + W$ there are vectors $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in W$ such that $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$. Since $\boldsymbol{u} \in U$ and $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ is a basis for U there are scalars $a_1, a_2 \in \mathbb{F}$ such that $\boldsymbol{u} = a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2$. Similarly, there are scalars $b_1, b_2, b_3 \in \mathbb{F}$ such that $\boldsymbol{w} = b_1\boldsymbol{w}_1 + b_2\boldsymbol{w}_2 + b_3\boldsymbol{w}_3$. Thus, $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w} = a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + b_1\boldsymbol{w}_1 + b_2\boldsymbol{w}_2 + b_3\boldsymbol{w}_3$ and therefore $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$ is a spanning sequence as required.

4. We are assuming that $dim(V) = n, S = (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n)$ is a sequence of vectors from V and S spans V. We need to prove that S is linearly independent. For $1 \leq j \leq n$ let $S - \boldsymbol{v}_j$ denote the sequence obtained from S by deleting $\boldsymbol{v}_j$.

Suppose to the contrary that S is not linearly independent. Then for some $j, 1 \leq j \leq n, \boldsymbol{v}_j$ is a linear combination of $S - \boldsymbol{v}_j$ by Theorem (1.14). Then by Theorem (1.13), $V = Span(S) = Span(S - \boldsymbol{v}_j)$. But then by the exchange theorem no independent sequence of V can have more than $n - 1$ vectors which contradicts the assumption that bases have cardinality n.

5. We are assuming $dim(V) = n, S = (\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_m)$ with $m > n$ and $Span(S) = V$. We have to prove that some subsequence of S is a basis of V. Toward that end, let S_0 be a subsequence of S such that $Span(S_0) = V$ and the length of S_0 is as small as possible. We claim that S_0 is linearly independent and therefore a basis of V. Let $S_0 = (\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_k)$ and assume to the contrary that S_0 is linearly dependent. Then for some $j, 1 \leq j \leq k, \boldsymbol{w}_j$ is a linear combination of $S_0 - \boldsymbol{w}_j$ by Theorem (1.14). Then by Theorem (1.13), $V = Span(S_0) = Span(S_0 - \boldsymbol{w}_j)$. However, this contradicts the assumption that S_0 is a spanning subsequence of S of minimal length. Thus, S_0 is linearly independent and a basis as claimed.

6. Assume $dim(U \cap W) = l, dim(U) - l = m, dim(W) - l = n$ (so that $dim(U) = l + m$ and $dim(W) = l + n$). Choose a basis $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_l)$ for $U \cap W$. This can be extended to a basis of U. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ be a sequence of vectors from U such that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_l, \boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ is a basis for U. Likewise there is a sequence of vectors $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ from W such that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_l, \boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ is a basis for W. We claim that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_l, \boldsymbol{u}_1, \dots, \boldsymbol{u}_m, \boldsymbol{w}_1 \dots, \boldsymbol{w}_n)$ is a basis for $U + W$.

We first show that $\mathcal{S} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_l, \boldsymbol{u}_1, \dots, \boldsymbol{u}_m, \boldsymbol{w}_1 \dots, \boldsymbol{w}_n)$ is linearly independent. Suppose $a_i, 1 \leq i \leq l, b_j, 1 \leq j \leq m$ and $c_k, 1 \leq k \leq n$ are scalars such that

$$a_1\boldsymbol{x}_1 + \cdots + a_l\boldsymbol{x}_l + b_1\boldsymbol{u}_1 + \cdots + b_m\boldsymbol{u}_m +$$

$$c_1\boldsymbol{w}_1 + \cdots + c_n\boldsymbol{w}_n = \boldsymbol{0}.$$

We then have that

$$a_1\boldsymbol{x}_1 + \cdots + a_l\boldsymbol{x}_l + b_1\boldsymbol{u}_1 + \cdots + b_m\boldsymbol{u}_m =$$

$$-(c_1\boldsymbol{w}_1+\cdots+c_n\boldsymbol{w}_n) \quad (1.1)$$

Note the vector on the left hand side of Equation (1.1) belongs to U and the vector on the right hand side belongs to W. Therefore, since we have equality the vector belongs to $U\cap W = Span(\boldsymbol{x}_1,\dots,\boldsymbol{x}_l)$. However, since $(\boldsymbol{x}_1,\dots,\boldsymbol{x}_l,\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)$ is linearly independent, by Exercise (1.5.11) we must have $Span(\boldsymbol{x}_1,\dots,\boldsymbol{x}_l)\cap Span(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)=\{\boldsymbol{0}\}$. Consequently,

$$a_1\boldsymbol{x}_1+\cdots+a_l\boldsymbol{x}_l+b_1\boldsymbol{u}_1+\cdots+b_m\boldsymbol{u}_m=$$

$$-(c_1\boldsymbol{w}_1+\dots c_n\boldsymbol{w}_n)=\boldsymbol{0}.$$

However, since $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)$ is linearly independent this implies that $c_1=\cdots=c_n=0$. Also, since $(\boldsymbol{x}_1,\dots,\boldsymbol{x}_l,\boldsymbol{u}_1,\dots,\boldsymbol{u}_m)$ is linearly independent we get $a_1=\cdots=a_l=b_1=\cdots=b_m=0$. Thus, all a_i,b_j,c_k are zero and the sequence $\mathcal{S}$ is linearly independent.

We next show that $\mathcal{S}$ spans $U+W$. Suppose $\boldsymbol{v}$ is an arbitrary vector in $U+W$. Then there are vectors $\boldsymbol{u}\in U$ and $\boldsymbol{w}\in W$ such that $\boldsymbol{v}=\boldsymbol{u}+\boldsymbol{w}$. Since $(\boldsymbol{x}_1,\dots,\boldsymbol{x}_l,\boldsymbol{u}_1,\dots,\boldsymbol{u}_m)$ is a basis for U there are scalars $a_i, 1\le i\le l, b_j, 1\le j\le m$ such that

$$\boldsymbol{u}=a_1\boldsymbol{x}_1+\cdots+a_l\boldsymbol{x}_l+b_1\boldsymbol{u}_1+\cdots+b_m\boldsymbol{u}_m.$$

Since $(\boldsymbol{x}_1,\dots,\boldsymbol{x}_l,\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)$ is a basis for W there are scalars $c_i, 1\le i\le l, d_k, 1\le k\le n$ such that

$$\boldsymbol{w}=c_1\boldsymbol{x}_1+\cdots+c_l\boldsymbol{x}_l+d_1\boldsymbol{w}_1+\cdots+d_n\boldsymbol{w}_n.$$

But now we have

$$\boldsymbol{v}=\boldsymbol{u}+\boldsymbol{w}=$$

$$(a_1\boldsymbol{x}_1+\cdots+a_l\boldsymbol{x}_l+b_1\boldsymbol{u}_1+\cdots+b_m\boldsymbol{u}_m)+$$

$$(c_1\boldsymbol{x}_1+\cdots+c_l\boldsymbol{x}_l+d_1\boldsymbol{w}_1+\cdots+d_n\boldsymbol{w}_n)=$$

$$(a_1+c_1)\boldsymbol{x}_1+\cdots+(a_l+c_l)\boldsymbol{x}_l+$$

$$b_1\boldsymbol{u}_1+\cdots+b_m\boldsymbol{u}_m+d_1\boldsymbol{w}_1+\dots d_n\boldsymbol{w}_n.$$

Thus, $\mathcal{S}$ spans $U+W$ as required and $\mathcal{S}$ is a basis of $U+W$. It follows that $dim(U+W)=l+m+n$. Now

$$dim(U)+dim(W)=(l+m)+(l+n)=2l+m+n=$$

$$l+(l+m+n)=dim(U\cap W)+dim(U+W).$$

7. Let $(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3)$ be a basis for X and $(\boldsymbol{y}_1,\boldsymbol{y}_2,\boldsymbol{y}_3)$ be a basis for Y. If $X\cap Y=\{\boldsymbol{0}\}$ then $(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,\boldsymbol{y}_1,\boldsymbol{y}_2,\boldsymbol{y}_3)$ is linearly independent by Exercise 11 of Section (1.5) . But this contradicts $dim(V)=5$. Thus, $X\cap Y\ne\{\boldsymbol{0}\}$.

8. Since $U+W=V, dim(U+W)=dim(V)=n$. Making use of Exercise 8 we get

$$dim(U\cap W)=dim(V)-dim(U)-dim(W)=$$

$$n-k-(n-k)=0.$$

Since $dim(U\cap W)=0, U\cap W=\{\boldsymbol{0}\}$ as required.

9. Set

$$X=\{\begin{pmatrix}x_1\\x_2\\x_3\\0\\0\\0\end{pmatrix}|x_1,x_2,x_3\in\mathbb{F}\},$$

$$Y=\{\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\0\\0\end{pmatrix}|x_1,x_2,x_3,x_4\in\mathbb{F}\},$$

$$Z = \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ 0 \end{pmatrix} | x_1, x_2, x_3, x_4, x_5 \in \mathbb{F}\}$$

Note that $dim(X) = 3, dim(Y) = 4, dim(Z) = 5$.

a) Set

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\boldsymbol{w}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{w}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{w}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

b) Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ be as in a) and now set

$$\boldsymbol{w}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{w}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{w}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Note that $Span(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3) = X$ and that $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3 \notin X$. Therefore $Span(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3) \neq Span(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$. On the other hand, $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3 \in Y$. By the argument of Exercise 2 we have

$$dim[Span(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3) \cap Span(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)] = 2.$$

c) Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ be as in a) and now set

$$\boldsymbol{w}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{w}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{w}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

Set $V = Span(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3), W = Span(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$. Since $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3 \in Z, V+W \subset Z$ and therefore

$$dim(V + W) \leq dim(Z) = 5 \tag{1.2}$$

On the other hand, $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{w}_1, \boldsymbol{w}_2)$ is linearly independent and therefore

$$dim(V + W) \geq 5 \tag{1.3}$$

From Equation (1.2) and (1.3) we get

$$dim(V + W) = 5.$$

Now use Exercise 6:

$$dim(V) + dim(W) = dim(V \cap W) + dim(V + W)$$

Since $dim(V) = dim(W) = 3$ and $dim(V + W) = 5$ we conclude that $dim(V \cap W) = 1$ as required.

10. a) There are 8 non-zero vectors and any one can be the first vector in a basis. Suppose we choose the vector $\boldsymbol{v}$. The second vector must be linearly independent and therefore not a multiple of $\boldsymbol{v}$. There are 3 multiples: $\boldsymbol{0}, \boldsymbol{v}, -\boldsymbol{v}$. So there are 9 - 3 = 6 choices for the second vector. Thus, there are $8 \times 6 = 48$ bases.

b) There are 24 non-zero vectors and any can be the first vector in a basis. Suppose we choose the vector $\boldsymbol{v}$. The second vector must be linearly independent and therefore not a multiple of $\boldsymbol{v}$. There are 5 multiples:

$\mathbf{0}, \boldsymbol{v}, 2\boldsymbol{v}, 3\boldsymbol{v}, 4\boldsymbol{v}$. So there are 25 - 5 = 20 choices for the second vector. Thus, there are $24 \times 20 = 480$ bases.

c) There are $p^2 - 1$ non-zero vectors and any can be the first vector in a basis. Suppose we choose the vector $\boldsymbol{v}$. The second vector must be linearly independent and therefore not a multiple of $\boldsymbol{v}$. There are p multiples: $\mathbf{0}, \boldsymbol{v}, \dots, (p-1)\boldsymbol{v}$. So there are $p^2 - p$ choices for the second vector. Thus, there are $(p^2-1)(p^2-p)$ bases.

11. Let $dim(V) = n$ and $dim(U) = k$. If $U = V$ we can take $W = \{\mathbf{0}\}$ so we may assume $k < n$. Choose a basis $\mathcal{B}_U = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ for U. By part 1) of Theorem (1.24) we can expand $\mathcal{B}_U$ to a basis, $\mathcal{B} = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ for V. Set $W = Span(\boldsymbol{u}_{k+1}, \dots, \boldsymbol{u}_n)$. Then W is a complement to U.

12. If $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k) \subset W$ then $V = Span(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k) \subset Span(W) = W$, a contradiction.

13. First assume that $X \cap Y = \{\mathbf{0}\}$. Let $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k)$ be basis for X and $(\boldsymbol{y}_1, \dots, \boldsymbol{y}_k)$ a basis for Y. Set $\boldsymbol{u}_i = \boldsymbol{x}_i + \boldsymbol{y}_i$ and $U = Span(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$. Then $X \cap U = Y \cap U = \mathbf{0}$ and $X \oplus U = Y \oplus U = X \oplus Y$. Let W be a complement to $X + Y$ in V and set $Z = U \oplus W$.

Assume $X \cap Y \neq \{\mathbf{0}\}$ and let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_j)$ be a basis for $X \cap Y$. Set $s = k - j$ and let $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_s)$ be a sequence of vectors from X such that $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{x}_1, \dots, \boldsymbol{x}_s)$ is a basis of X and similarly let $(\boldsymbol{y}_1, \dots, \boldsymbol{y}_s)$ be a sequence of vectors from Y such that $(\boldsymbol{u}_1, \dots, \boldsymbol{v}_j, \boldsymbol{y}_1, \dots, \boldsymbol{v}_s)$ is a basis for Y. Set $\boldsymbol{u}_i = \boldsymbol{x}_i + \boldsymbol{y}_i, 1 \leq i \leq s$ and $U = Span(\boldsymbol{u}_1, \dots, \boldsymbol{u}_s)$. Then $X \cap U = Y \cap U = \{\mathbf{0}\}$ and $X \oplus U = Y \oplus U = X + Y$. Let W be a complement to $X + Y$ in V and set $Z = U \oplus W$.

1.7. Bases of Infinite Dimensional Vector Spaces

1. First we show that $\{\chi_x | x \in X\}$ is independent. If not, then there must exist a finite subset $\{\chi_{x_i} | 1 \leq i \leq n\}$ which is dependent. There there are scalars, c_i such that

$$f = \sum_{i=1}^{n} c_i \chi_{x_i}$$

is the zero function. In particular, for each i, $f(x_i) = 0$. However, $f(x_i) = c_i$. Thus, $c_1 = \cdots = c_n = 0$. Thus, $\{\chi_x | x \in X\}$ is linearly independent.

Now we need to show that $\{\chi_x | x \in X\}$ spans. Let $f \in \mathcal{M}_{fin}(X, \mathbb{F})$ be a non-zero function. Then $spt(f)$ is non-empty but finite. Suppose $spt(f) = \{x_1, \dots, x_n\}$. Set $c_i = f(x_i)$. Then the function f and $\sum_{i=1}^{n} c_i \chi_{x_i}$ are equal.

2. First of all suppose X is an $n-$dimensional vector space over $\mathbb{Q}$ say with basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$. Then there is a one-to-one set correspondence between V and $\mathbb{Q}^n$, namely taking $\sum_{i=1}^{n} a_i \boldsymbol{v}_i$ to $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. Therefore the cardinality of X is the same as the cardinality of $\mathbb{Q}^n$ which is the same as the cardinality of $\mathbb{Q}$.

Suppose $\mathcal{B}$ is a basis of $\mathbb{R}$ as a vector space over $\mathbb{Q}$. Let $\mathcal{P}_{fin}(\mathcal{B})$ be all the finite subsets of $\mathcal{B}$. Then

$$\mathbb{R} = \cup_{B \in \mathcal{P}_{fin}(\mathcal{B})} Span(B).$$

As shown above, for any $B \in \mathcal{P}_{fin}(\mathcal{B})$, the cardinality of $Span(B)$ is countable.

For an infinite set X the cardinality of the finite subsets of X is the same as the cardinality of X. Also, if Y is an infinite set and for each $y \in Y, S_y$ is a countable set then $\cup_{y \in Y} S_y$ has cardinality no greater then Y. It now follows that the cardinality of $\mathbb{R}$ is no greater than the cardinality of $\mathcal{B}$. Since $\mathcal{B}$ is a subset of $\mathbb{R}$ we conclude that $\mathcal{B}$ and $\mathbb{R}$ have the same cardinality.

3. Choose a basis $\mathcal{B}_U$. This can be extended to a basis $\mathcal{B}$ for V. Let $W = Span(\mathcal{B} \setminus \mathcal{B}_U)$. Then W is a complement to U in V.

4. Let $\mathcal{B}$ be a basis for V. Choose a subset X_n of $\mathcal{B}$ with cardinality n. Then $U_n = Span(\mathcal{B} \setminus X_n)$ satisfies $dim(V/U_n) = n$.

1.8. Coordinate Vectors

1a) Since $dim(\mathbb{F}_2[x])$ is 3 and there are three vectors in the sequence, by the half is good enough theorem it suffices to show the vectors are linearly independent.

Thus, suppose c_1, c_2, c_3 are scalars such that

$$c_1(1+x) + c_2(1+x^2) + c_3(1+2x-2x^2) = 0$$

After distributing and collecting terms we obtain

$$(c_1 + c_2 + c_3) + (c_1 + 2c_3)x + (c_2 - 2c_3)x^2 = 0$$

This gives rise to the homogeneous linear system

$$\begin{array}{ccccccc} c_1 & + & c_2 & + & c_3 & = & 0 \\ c_1 & & & + & 2c_3 & + & 0 \\ & & c_2 & - & 2c_3 & + & 0 \end{array}$$

This system has only the trivial solution $c_1 = c_2 = c_3 = 0$. Thus, the vectors are linearly independent as required.

1b) To compute $[1]_{\mathcal{F}}$ we need to determine c_1, c_2, c_3 such that

$$(c_1 + c_2 + c_3) + (c_1 + 2c_3)x + (c_2 - 2c_3)x^2 = 1$$

This gives rise to the linear system

$$\begin{array}{ccccccc} c_1 & + & c_2 & + & c_3 & = & 1 \\ c_1 & & & + & 2c_3 & = & 0 \\ & & c_2 & - & 2c_3 & = & 0 \end{array}$$

which has the unique solution $[1]_{\mathcal{F}} = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}$. In a similar manner $[x]_{\mathcal{F}} = \begin{pmatrix} 3 \\ -1 \\ -1 \end{pmatrix}$ and $[x^2]_{\mathcal{F}} = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$.

2. Set $\boldsymbol{c}_j = \begin{pmatrix} c_{1j} \\ c_{2j} \\ c_{3j} \end{pmatrix}$, $j = 1, 2, 3$. This means that

$$\boldsymbol{u}_j = c_{1j}\boldsymbol{v}_1 + c_{2j}\boldsymbol{v}_2 + c_{3j}\boldsymbol{v}_3.$$

If $[\boldsymbol{x}]_{\mathcal{B}_1} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ then $\boldsymbol{x} = a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 + a_3\boldsymbol{u}_3$.

It then follows that

$$\begin{gathered} \boldsymbol{x} = a_1(c_{11}\boldsymbol{v}_1 + c_{21}\boldsymbol{v}_2 + c_{31}\boldsymbol{v}_3) + \\ a_2(c_{12}\boldsymbol{v}_1 + c_{22}\boldsymbol{v}_2 + c_{32}\boldsymbol{v}_3) + \\ a_3(c_{13}\boldsymbol{v}_1 + c_{23}\boldsymbol{v}_2 + c_{33}\boldsymbol{v}_3) = \\ (a_1c_{11} + a_2c_{12} + a_3c_{13})\boldsymbol{v}_1 + \\ (a_1c_{21} + a_2c_{22} + a_3c_{23})\boldsymbol{v}_2 + \\ (a_1c_{31} + a_2c_{32} + a_3c_{33})\boldsymbol{v}_3. \end{gathered}$$

Consequently, $[\boldsymbol{x}]_{\mathcal{B}_2} =$

$$\begin{pmatrix} a_1c_{11} + a_2c_{12} + a_3c_{13} \\ a_1c_{21} + a_2c_{22} + a_3c_{23} \\ a_1c_{31} + a_2c_{32} + a_3c_{33} \end{pmatrix} =$$

$$a_1 \begin{pmatrix} c_{11} \\ c_{21} \\ c_{31} \end{pmatrix} + a_2 \begin{pmatrix} c_{12} \\ c_{22} \\ c_{32} \end{pmatrix} + a_3 \begin{pmatrix} c_{31} \\ c_{32} \\ c_{33} \end{pmatrix} =$$

$$(a_1\boldsymbol{c}_1 + a_2\boldsymbol{c}_2 + a_3\boldsymbol{c}_3)$$

3a) For $i = 0, 1, 2, 3$ we have $f_j(i) = 1$ if $i = j$ and 0 otherwise. We claim that implies that (f_1, f_2, f_3, f_4) is linearly independent. For suppose $c_1f_1 + c_2f_2 + c_3f_3 + c_4f_4$ is the zero function. Substituting i with $i = 0, 1, 2, 3$ we get $0 = c_{i+1}$. Thus, all the $c_i = 0$. Now since there are 4 vectors and $dim(\mathbb{R}_3[x]) = 4$ it follows that $\mathcal{F}$ is a basis for $\mathbb{R}_3[x]$.

b) Suppose $g = c_1f_1 + c_2f_2 + c_3f_3 + c_4f_4$. Then $g(i) = c_1f_i(i) + c_2f_2(i) + c_3f_3(i) + c_4f_4(i) = c_{i+1}$ as required.

4. $[1]_{\mathcal{F}} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, [x]_{\mathcal{F}} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}, [x^2]_{\mathcal{F}} = \begin{pmatrix} 0 \\ 1 \\ 4 \\ 9 \end{pmatrix}, [x^3]_{\mathcal{F}} = \begin{pmatrix} 0 \\ 1 \\ 8 \\ 27 \end{pmatrix}$.

5. Assume that $Span(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) = V$ and that $\boldsymbol{c} \in \mathbb{F}^n$. Let $\boldsymbol{x} \in V$ be the vector such that $[\boldsymbol{x}]_{\mathcal{B}} = \boldsymbol{c}$. By hypothesis $\boldsymbol{x} \in Span(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$. By Theorem (1.29), $\boldsymbol{c}$ is a linear combination of $([\boldsymbol{u}_1]_{\mathcal{B}}, [\boldsymbol{u}_2]_{\mathcal{B}}, \ldots, [\boldsymbol{u}_k]_{\mathcal{B}})$, equivalently, $\boldsymbol{c} \in Span([\boldsymbol{u}_1]_{\mathcal{B}}, [\boldsymbol{u}_2]_{\mathcal{B}}, \ldots, [\boldsymbol{u}_k]_{\mathcal{B}})$.

Conversely, assume that $Span([\boldsymbol{u}_1]_{\mathcal{B}}, \ldots, [\boldsymbol{u}_k]_{\mathcal{B}}) = \mathbb{F}^n$ and that $\boldsymbol{x} \in V$. Let $\boldsymbol{c} = [\boldsymbol{x}]_{\mathcal{B}}$. By hypothesis, $\boldsymbol{c}$ is a linear combination of $([\boldsymbol{u}_1]_{\mathcal{B}}, [\boldsymbol{u}_2]_{\mathcal{B}}, \ldots, [\boldsymbol{u}_k]_{\mathcal{B}})$. Then by Theorem (1.29) it follows that $\boldsymbol{x}$ is a linear combination of $(\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k)$.

6. This follows from Exercise 5 and the half is good enough theorem.

Chapter 2

Linear Transformations

2.1. Introduction to Linear Transformations

1. Set $\boldsymbol{e}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \boldsymbol{e}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \boldsymbol{e}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$, so that $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ is a basis of $\mathbb{F}^3$. Then $T(a\boldsymbol{e}_1 + b\boldsymbol{e}_2 + c\boldsymbol{e}_3) = a(1 + x + x^2) + b(1 - x^2) + c(-2 - x^2)$.

By Theorem (2.5) it follows that T is a linear transformation.

2. We show the additive property is violated: $T\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} = T\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$. It should then be the case that $T\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}0 & 0\\0 & 0\end{pmatrix}$. However, $T\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}$.

3. $T(a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix}) = a\begin{pmatrix}2\\0\\4\end{pmatrix} + b\begin{pmatrix}-3\\0\\5\end{pmatrix}$. Now by Theorem (2.5) it follows that T is a linear transformation.

4.

$$T(\begin{pmatrix}x_1\\y_1\end{pmatrix} + \begin{pmatrix}x_2\\y_2\end{pmatrix}) = T\begin{pmatrix}x_1 + x_2\\y_1 + y_2\end{pmatrix} =$$

$$\begin{pmatrix}e^{x_1+x_2}\\e^{y_1+y_2}\end{pmatrix} = \begin{pmatrix}e^{x_1}e^{x_2}\\e^{y_1}e^{y_2}\end{pmatrix} =$$

$$\begin{pmatrix}e^{x_1}\\e^{y_1}\end{pmatrix} + \begin{pmatrix}e^{x_2}\\e^{y_2}\end{pmatrix} = T\begin{pmatrix}x_1\\y_1\end{pmatrix} + T\begin{pmatrix}x_2\\y_2\end{pmatrix}.$$

$$T(c\begin{pmatrix}x\\y\end{pmatrix}) = T(\begin{pmatrix}cx\\cy\end{pmatrix}) = \begin{pmatrix}e^{cx}\\e^{cy}\end{pmatrix} =$$

$$\begin{pmatrix}(e^x)^c\\(e^y)^c\end{pmatrix} = c\begin{pmatrix}e^x\\e^y\end{pmatrix} = cT\begin{pmatrix}x\\y\end{pmatrix}.$$

5. Let $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U, c_1, c_2 \in \mathbb{F}$. Then $(T \circ S)(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2) = T(S(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2))$. Since S is a linear transformation $S(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2) = c_1S(\boldsymbol{u}_1) + c_2S(\boldsymbol{u}_2)$. It then follows that

$$T(S(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2)) = T(c_1S(\boldsymbol{u}_1) + c_2S(\boldsymbol{u}_2)).$$

Since T is a linear transformation

$$T(c_1S(\boldsymbol{u}_1) + c_2S(\boldsymbol{u}_2)) = c_1T(S(\boldsymbol{u}_1)) + c_2T(S(\boldsymbol{u}_2)) =$$

$$c_1(T \circ S)(\boldsymbol{u}_1) + c_2(T \circ S)(\boldsymbol{u}_2).$$

It now follows that $T \circ S$ is a linear transformation.

6. Let $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$. Then by the definition of $S + T$ we have

$$(S + T)(\boldsymbol{v}_1 + \boldsymbol{v}_2) = S(\boldsymbol{v}_1 + \boldsymbol{v}_2) + T(\boldsymbol{v}_1 + \boldsymbol{v}_2)$$

Since S, T are linear transformations

$$S(\boldsymbol{v}_1 + \boldsymbol{v}_2) = S(\boldsymbol{v}_1) + S(\boldsymbol{v}_2),$$
$$T(\boldsymbol{v}_1 + \boldsymbol{v}_2) = T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2).$$

We then have

$$S(\boldsymbol{v}_1 + \boldsymbol{v}_2) + T(\boldsymbol{v}_1 + \boldsymbol{v}_2) =$$
$$[S(\boldsymbol{v}_1) + S(\boldsymbol{v}_2)] + [T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2)] =$$
$$[S(\boldsymbol{v}_1) + T(\boldsymbol{v}_1)] + [S(\boldsymbol{v}_2) + T(\boldsymbol{v}_2)] =$$
$$(S+T)(\boldsymbol{v}_1) + (S+T)(\boldsymbol{v}_2).$$

Now let $\boldsymbol{v} \in V, c \in \mathbb{F}$. Then by the definition of $S+T$

$$(S+T)(c\boldsymbol{v}) = S(c\boldsymbol{v}) + T(c\boldsymbol{v})$$

Since S, T are linear we have $S(c\boldsymbol{v}) = cS(\boldsymbol{v}), T(c\boldsymbol{v}) = cT(\boldsymbol{v})$. Then

$$S(c\boldsymbol{v}) + T(c\boldsymbol{v}) = cS(\boldsymbol{v}) + cT(\boldsymbol{v}) =$$
$$c[S(\boldsymbol{v}) + T(\boldsymbol{v})] = c[(S+T)(\boldsymbol{v})].$$

7. Let $\boldsymbol{v} \in V$ and write $\boldsymbol{v} = \boldsymbol{x} + \boldsymbol{y}$ where $\boldsymbol{x} \in X$ and $\boldsymbol{y} \in Y$.

a) $P_1(\boldsymbol{v}) = \boldsymbol{x}$ and $P_1(\boldsymbol{x}) = \boldsymbol{x}$. Consequently, $P_1 \circ P_1(\boldsymbol{v}) = P_1(\boldsymbol{v})$. In exactly the same way, $P_2 \circ P_2 = P_2$.

b) Now $(P_1 + P_2)(\boldsymbol{v}) = P_1(\boldsymbol{v}) + P_2(\boldsymbol{v}) = \boldsymbol{x} + \boldsymbol{y} = \boldsymbol{v}$ and therefore $P_1 + P_2$ is the identity transformation of V.

c) Note that $P_1(\boldsymbol{y}) = P_2(\boldsymbol{x}) = \boldsymbol{0}$. In now follows that $(P_1 \circ P_2)(\boldsymbol{v}) = P_1(P_2(\boldsymbol{v})) = P_1(\boldsymbol{y}) = \boldsymbol{0}$ and therefore $P_1 \circ P_2 = 0_V$. In a similar fashion $P_2 \circ P_1 = 0_V$.

8. By Exercise 6, $P_1 + P_2 = I_V$. Then $T = I_V \circ T = (P_1 + P_2) \circ T = (P_1 \circ T) + (P_2 \circ T)$. Since $P_1 \circ T$ and $P_2 \circ T$ are linear transformations, by Lemma (2.2) $T = (P_1 \circ T) + (P_2 \circ T)$ is a linear transformation.

9. Let $\boldsymbol{v} \in V$ be an arbitrary vector and set $\boldsymbol{x} = P_1(\boldsymbol{v}), \boldsymbol{y} = P_2(\boldsymbol{v})$. Then $\boldsymbol{x} \in X, \boldsymbol{y} \in Y$ so $\boldsymbol{x} + \boldsymbol{y} \in X + Y$. By hypothesis, $P_1 + P_2 = I_V$ and therefore $\boldsymbol{v} = (P_1 + P_2)(\boldsymbol{v}) = P_1(\boldsymbol{v}) + P_2(\boldsymbol{v}) = \boldsymbol{x} + \boldsymbol{y}$. As $\boldsymbol{v}$ is arbitrary we conclude that $V = X + Y$. On the other hand, assume $\boldsymbol{v} \in X \cap Y$. Since $\boldsymbol{v} \in Y, P_2(\boldsymbol{v}) = \boldsymbol{v}$. Since $\boldsymbol{v} \in X, P_1(\boldsymbol{v}) = \boldsymbol{v}$. It then follows that $(P_1 \circ P_2)(\boldsymbol{v}) = P_1(P_2(\boldsymbol{v})) = P_1(\boldsymbol{v}) = \boldsymbol{v}$. However, we are assuming that $P_1 \circ P_2 = 0_V$ and therefore $\boldsymbol{v} = (P_1 \circ P_2)(\boldsymbol{v}) = \boldsymbol{0}$. Thus, $X \cap Y = \{\boldsymbol{0}\}$.

10. Let $\mathcal{B} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be a basis for V. Now $T(\mathcal{B}) = (T(\boldsymbol{v}_1), \dots, T(\boldsymbol{v}_n))$ is a sequence of n vectors in W and $dim(W) = m < n$. By the Exchange theorem there are scalars $c_1, \dots, c_n$, not all zero, such that

$$c_1 T(\boldsymbol{v}_1) + \cdots + c_n T(\boldsymbol{v}_n) = \boldsymbol{0}_W.$$

Since T is a linear transformation

$$c_1 T(\boldsymbol{v}_1) + \cdots + c_n T(\boldsymbol{v}_n) = T(c_1 \boldsymbol{v}_1 + \dots c_n \boldsymbol{v}_n).$$

Since not all $c_1, c_2 \dots, c_n$ are zero and $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is a basis, the vector $\boldsymbol{v} = c_1 \boldsymbol{v}_1 + \dots c_n \boldsymbol{v}_n \neq \boldsymbol{0}_V$ and satisfies $T(\boldsymbol{v}) = \boldsymbol{0}_W$.

11. $\Pi(\boldsymbol{v}_1 + \boldsymbol{v}_2) = (\boldsymbol{v}_1 + \boldsymbol{v}_2) + W = (\boldsymbol{v}_1 + W) + (\boldsymbol{v}_2 + W) = \Pi(\boldsymbol{v}_1) + \Pi(\boldsymbol{v}_2)$.

$\Pi(c\boldsymbol{v}) = (c\boldsymbol{v}) + W = c(\boldsymbol{v} + W) = c\Pi(\boldsymbol{v})$.

12. If $\boldsymbol{w}_j \in R(T)$ then there is a vector $\boldsymbol{v}_j \in V$ such that $T(\boldsymbol{v}_j) = \boldsymbol{w}_j$. Now let $\boldsymbol{w} \in W$ be arbitrary. Since $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$ is a spanning sequence for W there are scalars $c_1, \dots, c_m$ such that

$$\boldsymbol{w} = c_1 \boldsymbol{w}_1 + \dots c_m \boldsymbol{w}_m = c_1 T(\boldsymbol{v}_1) + \dots c_m T(\boldsymbol{v}_m).$$

Set $\boldsymbol{v} = c_1 \boldsymbol{v}_1 + \dots c_m \boldsymbol{v}_m$. Since T is a linear transformation

$$T(\boldsymbol{v}) = T(c_1 \boldsymbol{v}_1 + \cdots + c_m \boldsymbol{v}_m) =$$
$$c_1 T(\boldsymbol{v}_1) + \dots c_m T(\boldsymbol{v}_m) =$$
$$c_1 \boldsymbol{w}_1 + \dots c_m \boldsymbol{w}_m = \boldsymbol{w}.$$

Since $\boldsymbol{w}$ is arbitrary this proves that $R(T) = W$.

13. Since each $T(\boldsymbol{v}_i) \in R(T)$ and $R(T)$ is a subspace of W it follows that $Span(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \subset R(T)$ so we have to prove the reverse inclusion.

Assume $\boldsymbol{w} \in R(T)$. Then there exists $\boldsymbol{v} \in V$ such that $T(\boldsymbol{v}) = \boldsymbol{w}$. Since $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is a basis for V there are scalars $c_1, \dots, c_n$ such that $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \dots + c_n\boldsymbol{v}_n$. Then

$$\boldsymbol{w} = T(\boldsymbol{v}) = T(c_1\boldsymbol{v}_1 + \dots + c_n\boldsymbol{v}_n) = c_1T(\boldsymbol{v}_1) + \dots + c_nT(\boldsymbol{v}_n)$$

The last equality follows from Lemma (2.1). Clearly, $c_1T(\boldsymbol{v}_1) + \dots + c_nT(\boldsymbol{v}_n)$ is in $Span(T(\boldsymbol{v}_1), \dots, T(\boldsymbol{v}_n))$ which completes the proof.

14. Let $T : V \to W$ be a linear transformation. Let $T(\boldsymbol{v}_j) = a_{1j}\boldsymbol{w}_1 + \dots + a_{mj}\boldsymbol{w}_m$. Consider the linear transformation

$$S = \sum_{i,j} a_{ij}E_{ij}.$$

Claim that $S = T$. Towards this end it suffices to prove that $S(\boldsymbol{v}_j) = T(\boldsymbol{v}_j)$ for all j. Now since $E_{ik}(\boldsymbol{v}_j) = \boldsymbol{0}_W$ for $k \neq j$ we have

$$S(\boldsymbol{v}_j) = \sum_{i=1}^{m} a_{ij}E_{ij}(\boldsymbol{v}_j) =$$

$$\sum_{i=1}^{n} a_{ij}\boldsymbol{w}_i = a_{1j}\boldsymbol{w}_1 + \dots + a_{mj}\boldsymbol{w}_m = T(\boldsymbol{v}_j)$$

as required.

On the other hand, suppose $S = \sum_{i,j} a_{ij}E_{ij}$ is the zero function from V to W which takes every vector of V to the zero vector of W. In particular, $S(\boldsymbol{v}_j) = \boldsymbol{0}_W$. Thus, $S(\boldsymbol{v}_j) = \sum_{i=1}^{m} a_{ij}E_{ij}(\boldsymbol{v}_j) = \sum_{i=1}^{m} a_{ij}\boldsymbol{w}_i = \boldsymbol{0}_W$. However, since $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$ is a basis of W we must have $a_{ij} = \dots = a_{mj} = 0$. Since this holds for every $j, 1 \leq j \leq n$ we have all $a_{ij} = 0$ and the E_{ij} are linearly independent. With what we have proved above, $(E_{11}, E_{21}, \dots, E_{m1}, E_{12}, \dots, E_{m2}, \dots, E_{1n}, \dots, E_{mn})$ is a basis for $\mathcal{L}(V, W)$. It now follows that $dim(\mathcal{L}(V, W)) = mn$.

15. Let $\mathcal{X}$ consist of all pairs $(\mathcal{A}, \phi)$ where $\mathcal{A} \subset \mathcal{B}, \phi$ is a linear transformation from $Span(\mathcal{A}$ to W and ϕ restricted to $\mathcal{A}$ is equal to f restricted to $\mathcal{A}$. Order $\mathcal{X}$ as follows: $(\mathcal{A}, \phi) \leq (\mathcal{A}', \phi')$ if and only if $\mathcal{A} \subset \mathcal{A}'$ and ϕ' restricted to $Span(\mathcal{A})$ is equal to ϕ. We prove that every chain has an upper bound.

Thus, assume that $\mathcal{C} = \{(\mathcal{A}_i, \phi_i) | i \in I\}$ is a chain in $\mathcal{X}$. Set $\mathcal{A} = \cup_{i \in I}\mathcal{A}_i$. Define ϕ as follows: If $\boldsymbol{v} \in Span(\mathcal{A}_i)$ then $\phi(\boldsymbol{v}) = \phi_i(\boldsymbol{v})$. We need to show this is well defined. Suppose $\boldsymbol{v} \in Span(\mathcal{A}_i) \cap Span(\mathcal{A}_j)$ for $i, j \in I$. Since $\mathcal{C}$ is a chain either $\mathcal{A}_i \subset \mathcal{A}_j$ or $\mathcal{A}_j \subset \mathcal{A}_i$. Assume $\mathcal{A}_i \subset \mathcal{A}_j$. Also, since $\mathcal{C}$ is a chain, ϕ_j restricted to $\mathcal{A}_i$ is ϕ_i. It then follows that $\phi_j(\boldsymbol{v}) = \phi_i(\boldsymbol{v})$ since ϕ_i, ϕ_j are linear transformations.

Now by Zorn's lemma there exists a maximal element $(\mathcal{A}, \phi)$. We need to show that $\mathcal{A} = \mathcal{B}$. Suppose to the contrary that $\mathcal{A} \neq \mathcal{B}$ and let $\boldsymbol{v} \in \mathcal{B} \setminus \mathcal{A}$ and set $\mathcal{A}' = \mathcal{A} \cup \{\boldsymbol{v}\}$. Define $\phi' : Span(\mathcal{A}) \oplus Span(\boldsymbol{v})$ as follows: $\phi'(\boldsymbol{x} + c\boldsymbol{v}) = \phi(\boldsymbol{x}) = cf(\boldsymbol{v})$. Then ϕ' is linear and ϕ' restricted to $\mathcal{A}'$ is equal to f restricted to f. So, $(\mathcal{A}', \phi') \in \mathcal{X}$ which contradicts the maximality of $(\mathcal{A}, \phi)$. Therefore $\mathcal{A} = \mathcal{B}$ as claimed.

16. Let $c_1, \dots, c_k$ be scalars such that $c_1\boldsymbol{v}_1 + \dots + c_k\boldsymbol{v}_k = \boldsymbol{0}_V$. We need to show that $c_1 = \dots = c_k = 0$. Applying T we get

$$T(c_1\boldsymbol{v}_1 + \dots + c_k\boldsymbol{v}_k) = T(\boldsymbol{0}_V) = \boldsymbol{0}_W.$$

By Lemma (2.1)

$$T(c_1\boldsymbol{v}_1 + \dots + c_k\boldsymbol{v}_k) =$$

$$c_1T(\boldsymbol{v}_1) + \dots + c_kT(\boldsymbol{v}_k) = c_1\boldsymbol{w}_1 + \dots + c_k\boldsymbol{w}_k.$$

Since $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_k)$ is assumed to be linearly independent we have $c_1 = \dots = c_k = 0$ as required.

2.2. Range and Kernel of a Linear Transformation

1. There are three vectors in the basis for $R(T)$ and therefore $rank(T) = 3$. Since $dim(M_{23}(\mathbb{R}) = 6$ applying the rank-nullity theorem we get

$$rank(T) + nullity(T) = dim(M_{23}(\mathbb{R}))$$

$$3 + nullity(T) = 6$$

Thus, $nullity(T) = 3$.

2. $Ker(T) = Span((x-a)(x-b), x(x-a)(x-b))$. The nullity of T is 2. Since $dim(\mathbb{F}_3[x]) = 4$ by the rank-nullity theorem we have $dim(Ker(T)) = 2$. We can see that later directly by noting that $T(\frac{x-b}{a-b}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $T(\frac{x-a}{b-a}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

3. $T(a + bx + cx^2 + dx^3) =$

$$a\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + b\begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix} + c\begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} + d\begin{pmatrix} 2 \\ 1 \\ 1 \\ 2 \end{pmatrix}.$$

It follows that

$$R(T) = Span(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 2 \end{pmatrix}).$$

The first, second and fourth vectors are a basis for this subspace:

$$R(T) = Span(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 2 \end{pmatrix})$$

$Ker(T) = Span(2 - x + x^2), rank(T) = 3$, and $nullity(T) = 1$.

4. An arbitrary vector $a + bx + cx^2 \in \mathbb{F}_2[x]$ is the image of $\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$. This proves that T is surjective. T cannot be an isomorphism since $dim(M_{22}(\mathbb{F})) = 4 > 3 = dim(\mathbb{F}_2[x])$.

5. The only solution to the linear system

$$\begin{array}{ccccccc} a & + & b & & & = & 0 \\ a & - & 2b & - & 2c & = & 0 \\ & & b & + & c & = & 0 \\ a & + & 2b & + & c & = & 0 \end{array}$$

is the trivial one $a = b = c = 0$. Therefore $Ker(T) = \{\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\}$ which implies that T is one-to-one. On the other hand, since $dim(\mathbb{F}^3) = 3$ and $dim(M_{22}(\mathbb{F})) = 4$, the spaces cannot be isomorphic.

6. $R(T) = Span(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}) = \mathbb{F}^3$. Thus, T is surjective. By the half is good enough theorem for transformations, T is an isomorphism.

7. Let $w \in W$ be arbitrary. Since T is onto there exists an element $v \in V$ such that $T(v) = w$. Since S is onto there exists and element $u \in U$ such that $S(u) = v$. Then $(T \circ S)(u) = T(S(u)) = T(v) = w$. We have thus shown the existence of an element $u \in U$ such that $(T \circ S)(u) = w$. This proves that $T \circ S$ is surjective.

8. Suppose that $(T \circ S)(u) = (T \circ S)(u')$. Then $T(S(u)) = T(S(u'))$. Since T is one-to-one this implies that $S(u) = S(u')$. Since S is on-to-one we have $u = u'$. Thus, $T \circ S$ is one-to-one.

9. By Exercises 7 and 8 it follows that if S, t are isomorphisms then $T \circ S$ is a bijective function. By Exercise 5 of Section (2.1), $T \circ S$ is a linear transformation. Thus, $T \circ S$ is an isomorphism.

10. Let $\mathcal{B}_V = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis of V and set $\boldsymbol{w}_j = T(\boldsymbol{v}_j)$. Since T is injective and $\mathcal{B}_V$ is linearly independent $\mathcal{B}_W = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ is linearly independent by Theorem (2.11). Since V and W are isomorphic, $dim(W) = dim(V) = n$ so that by the half is good enough theorem, $\mathcal{B}_W$ is a basis for W. Now there exists a unique linear transformation $S : W \to V$ such $S(\boldsymbol{w}_j) = \boldsymbol{v}_j$. Note that $S \circ T$ is a linear operator on V such that $(S \circ T)(\boldsymbol{v}_j) = \boldsymbol{v}_j$ for all j. It then follows that $S \circ T = I_V$ and hence $S = T^{-1}$. This proves that T^{-1} is a linear transformation.

Alternative proof (applies even when V is infinite dimensional): Assume $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$. We need to show that $T^{-1}(\boldsymbol{w}_1+\boldsymbol{w}_2) = T^{-1}(\boldsymbol{w}_1)+T(\boldsymbol{w}_2)$. Set $\boldsymbol{v}_1 = T^{-1}(\boldsymbol{w}_1)$ and $\boldsymbol{v}_2 = T^{-1}(\boldsymbol{w}_2)$. Just as a reminder, $\boldsymbol{v}_i, i = 1, 2$ is the unique element of V such that $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$. Since T is linear, $T(\boldsymbol{v}_1 + \boldsymbol{v}_2) = T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2) = \boldsymbol{w}_1 + \boldsymbol{w}_2$. Therefore $T^{-1}(\boldsymbol{w}_1 + \boldsymbol{w}_2) = \boldsymbol{v}_1 + \boldsymbol{v}_2 = T^{-1}(\boldsymbol{w}_1) + T^{-1}(\boldsymbol{w}_2)$.

Now assume that $\boldsymbol{w} \in W$ and $c \in \mathbb{F}$. We need to show that $T^{-1}(c\boldsymbol{w}) = cT^{-1}(\boldsymbol{w})$. Set $\boldsymbol{v} = T^{-1}(\boldsymbol{w})$. Since T is linear $T(c\boldsymbol{v}) = cT(\boldsymbol{v}) = c\boldsymbol{w}$. We may therefore conclude that $T^{-1}(c\boldsymbol{w}) = c\boldsymbol{v} = cT^{-1}(\boldsymbol{w})$.

11. Let $\mathcal{B}_V = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis for V. We know that $R(T) = Span(T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n))$. Now let $\mathcal{B}_W = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ be a basis for W. By the exchange theorem, $dim(V) = n \geq m = \dim(W)$.

12. Let $\mathcal{B}_V = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis for V and $\mathcal{B}_W = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ be a basis for W. Since T is injective by Theorem (2.11), $T(\mathcal{B}_V) = (T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n))$ is linearly independent. By the exchange theorem, $dim(V) = n \leq m = dim(W)$.

13. Let $\mathcal{B}_W = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ be a basis for W. Since T is surjective, for each $i = 1, \ldots, m$ there exists a vector $\boldsymbol{v}_i$ such that $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$. By Theorem (2.6) there exists a unique linear transformation $S : W \to V$ such that $S(\boldsymbol{w}_i) = \boldsymbol{v}_i$. Now $T \circ S : W \to W$ is linear transformation and $(T \circ S)(\boldsymbol{w}_i) = T(S(\boldsymbol{w}_i)) = T(\boldsymbol{v}_i) = \boldsymbol{w}_i$. Since $T \circ S$ is a linear transformation and the identity when restricted to a basis of W it follows that $T \circ S = I_W$. (Note: If S is not one-to-one then there will be other choices of $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ and therefore T will not be unique).

14. If $dim(V) = dim(W)$ then T is an isomorphism by the half is good enough theorem for linear transformations and then T has a unique inverse (as a map) which is linear by Exercise 10. Therefore by Exercise 12 we may assume that $dim(V) < dim(W)$.

Let $\mathcal{B}_V = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis of V. Set $\boldsymbol{w}_i = T(\boldsymbol{v}_i), i = 1, \ldots, n$. Since T is one-to-one the sequence $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ is linearly independent. We can then extend $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ to a basis $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ for W. Now there exists a unique linear transformation $S : W \to V$ such that $S(\boldsymbol{w}_i) = \boldsymbol{v}_i$ for $i \leq n$ and $S(\boldsymbol{w}_i) = \boldsymbol{0}_V$ for $i > n$.

$S \circ T : V \to V$ is a linear transformation. Moreover, $(S \circ T)(\boldsymbol{v}_i) = S(T(\boldsymbol{v}_i)) = S(\boldsymbol{w}_i) = \boldsymbol{v}_i$. It then follows that $S \circ T = I_V$.

15. To show R is well-defined we need to prove if $T_1(\boldsymbol{v}) = T_1(\boldsymbol{v}')$ then $T_2(\boldsymbol{v}) = T_2(\boldsymbol{v}')$. If $T_1(\boldsymbol{v}) = T_1(\boldsymbol{v}')$ then $T(\boldsymbol{v} - \boldsymbol{v}') = \boldsymbol{0}$, that is, $\boldsymbol{v} - \boldsymbol{v}' \in Ker(T_1)$. Since by hypothesis, $Ker(T_1) = Ker(T_2), \boldsymbol{v} - \boldsymbol{v}' \in Ker(T_2)$. Consequently, $T_2(\boldsymbol{v} - \boldsymbol{v}') = \boldsymbol{0}$. It then follows that $T_2(\boldsymbol{v}) - T_2(\boldsymbol{v}') = \boldsymbol{0}$ and therefore $T_2(\boldsymbol{v}) = T_2(\boldsymbol{v}')$ as desired.

Now suppose $\boldsymbol{u}_1, \boldsymbol{u}_2 \in R(T_1)$ and c_1, c_2 are scalars. We need to show that $S(c_1\boldsymbol{u}_1+c_2\boldsymbol{u}_2) = c_1S(\boldsymbol{u}_1)+c_2S(\boldsymbol{u}_2)$. Toward that end, let $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ such that $T_1(\boldsymbol{v}_1) = \boldsymbol{u}_1, T_1(\boldsymbol{v}_2) = \boldsymbol{u}_2$. Then $T_1(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = c_1T_1(\boldsymbol{v}_1) + c_2T_1(\boldsymbol{v}_2) = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2$. Therefore $S(c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2) = T_2(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2)$. Since T_2 is linear, $T_2(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = c_1T_2(\boldsymbol{v}_1) + c_2T_2(\boldsymbol{v}_2) = c_1S(T_1(\boldsymbol{v}_1)) + c_2S(T_1(\boldsymbol{v}_2)) = c_1S(\boldsymbol{u}_1) + c_2S(\boldsymbol{u}_2)$.

16. Suppose first that $Ker(T) = \{\boldsymbol{0}\}$. Then T is injective and by the half is good enough theorem an isomorphism. Since the composition of injective functions is injective, T^k is also injective for all $k \in \mathbb{N}$. In particular, T^n and T^{n+1} and consequently, $Ker(T^n) = Ker(T^{n+1}) = \{\boldsymbol{0}\}$. Moreover, T^n and T^{n+1} are surjective and therefore

$Range(T^n) = Range(T^{n+1}) = V$. Therefore we may assume that T is not injective.

Let m be a natural number and suppose $\boldsymbol{v} \in Ket(T^m)$. Then $T^{m+1}(\boldsymbol{v}) = T(T^m(\boldsymbol{v})) = T(\mathbf{0}) = \mathbf{0}$. Thus, $Ker(T^m) \subset Ker(T^{m+1})$. Next, assume $\boldsymbol{w} \in Range(T^{m+1})$. Then there is a vector $\boldsymbol{x} \in V$ such that $\boldsymbol{w} = T^{m+1}(\boldsymbol{x}) = T^m(T(\boldsymbol{x})) \in Range(T^m)$. Thus, $Range(T^{m+1}) \subset Range(T^m)$.

It follows that $nullity(T^{m+1}) \geq nullity(T^m)$ and $rank(T^{m+1}) \leq rank(T^m)$. Consider the sequence of numbers $(nullity(T), nullity(T^2), \dots, nullity(T^n))$. Each is a natural number between 1 and n. If they are all distinct they are then decreasing and we must have $nullity(T^n) = n$, that is, T^n is the zero map. But then T^{n+1} is also the zero map and in this case $T^n = T^{n+1}$. In the case that they are not all distinct we must have some $m < n$ such that $nullity(T^m) = nullity(T^{m+1})$, so that $Ker(T^m) = Ker(T^{m+1})$. It then follows that $Ker(T^k) = Ker(T^m)$ for all $k \geq m$. In particular, $Ker(T^m) = Ker(T^n) = Ker(T^{n+1})$. By Theorem (2.9) it then follows that $rank(T^n) = rank(T^{n+1})$. Since $Range(T^{n+1}) \subset Range(T^n)$ we conclude that $Range(T^{n+1}) = Range(T^n)$.

17. Since $dim(Range(T^n)) + dim(Ker(T^n)) = dim(V)$ it suffices to prove that $Range(T^n) \cap Ker(T^n) = \{\mathbf{0}\}$. From the proof of Exercise 16, $Ker(T^k) = Ker(T^n)$ and $Range(T^k) = Range(T^n)$ for all $k \geq n$. Set $S = T^n$. It then follows that $Ker(S^2) = Ker(S)$ and $Range(S^2) = Range(S)$. So, assume that $\boldsymbol{v} \in Ker(S) \cap Range(S)$. Then $\boldsymbol{v} = S(\boldsymbol{w})$ for some $\boldsymbol{w} \in V$. Now $S^2(\boldsymbol{w}) = S(\boldsymbol{v}) = \mathbf{0}$ since $\boldsymbol{v} \in Ker(S)$. Thus, $\boldsymbol{w} \in Ker(S^2)$, whence $\boldsymbol{w} \in Ker(S)$. Thus, $\boldsymbol{v} = S(\boldsymbol{w}) = \mathbf{0}$ as required.

18. This follows immediately from Theorem (2.7) and the fact, established in the proof of Exercise 16 that $Range(T^2) \subset Range(T)$ and $Ker(T) \subset Ker(T^2)$.

19. a) Since $TS = \mathbf{0}_{V\to V}$ it follows that $Range(S) \subset Ker(T)$ from which we conclude that $rank(S) \leq nullity(T) = n - k$.

b) Let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n-k})$ be a basis for $Ker(T)$ and extend to a basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of V. There exists a unique operator S such that $S(\boldsymbol{v}_i) = \boldsymbol{v}_i$ if $1 \leq i \leq n-k$ and $S(\boldsymbol{v}_i) = \mathbf{0}$ if $i > n-k$. Then $Range(S) = Ker(T)$ so that $rank(S) = n-k$ and $TS = \mathbf{0}_{V\to V}$.

20. a) Since $ST = \mathbf{0}_{V\to V}$ it follows that $Range(T) \subset Ker(S)$. Consequently, $nullity(S) = dim(Ker(S)) \geq dim(Range(T)) = rank(T) = k$. It then follows that $rank(S) \leq n-k$.

b) Let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k)$ be a basis for $Range(T)$ and extend to a basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$. Let S be the unique linear operator such that $S(\boldsymbol{v}_i) = \mathbf{0}$ if $1 \leq i \leq k$ and $S(\boldsymbol{v}_i) = \boldsymbol{v}_i$ if $i > k$. Then $Range(S) = Span(\boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n)$ so that $rank(S) = n-k$. Since $Range(T) \subset Ker(S), ST = \mathbf{0}_{V\to V}$.

21. Let $dim(V) = n$. Since $T^2 = \mathbf{0}_{V\to V}$ it follows that $Range(T) \subset Ker(T)$ whence $rank(T) = k \leq nullity(T) = n-k$. It then follows that $2k \leq n$ so that $k \leq \frac{n}{2}$.

22. Let T be the unique linear operator such that $T(\boldsymbol{v}_i) = \boldsymbol{v}_{i+m}$ if $1 \leq i \leq m$ and $T(\boldsymbol{v}_i) = \mathbf{0}$ if $m+1 \leq i \leq n$.

2.3. Correspondence and Isomorphism Theorems

1. By the Theorem (2.19) we have

$$V/W = (X_1 + W)/W \cong X_1/(X_1 \cap W)$$

$$V/W = (X_2 + W)/W \cong X_2/(X_2 \cap W)$$

Thus, $X_1/(X_1 \cap W) \cong X_2/(X_2 \cap W)$.

2. If $V = X_1 \oplus W$ then by Theorem (2.19) $V/W = (X_1 + W)/W \cong X_1/(X_1 \cap W) = X_1$ since $X_1 \cap W = \{\mathbf{0}\}$. Similarly, $V/W \cong X_2$. Thus, $X_1 \cong X_2$.

3. Since f is not the zero map, by Theorem (2.17) it follows that $V/Ker(f) \cong Range(f)$. Since f is not the zero map, f must be surjective, that is, $Range(f) = \mathbb{F}$.

4. Let $\boldsymbol{u} \in U$. Then $T(\boldsymbol{u}) = \boldsymbol{u}$ which implies that $S(\boldsymbol{u}) = (T - I_V)(\boldsymbol{u}) = \boldsymbol{0}$.

Now assume that $\boldsymbol{v} \in V$ is arbitrary. Since $T(\boldsymbol{v} + U) = \boldsymbol{v} + U$ it follows that $S(\boldsymbol{v}) = (T - I_V)(\boldsymbol{v}) = T(\boldsymbol{v}) - \boldsymbol{v} \in U$. Then $S^2(\boldsymbol{v}) = S(S(\boldsymbol{v})) = \boldsymbol{0}$.

5 a) Assume $T(\boldsymbol{v}) = \boldsymbol{v}$. Then $(S + I_V)(\boldsymbol{v}) = S(\boldsymbol{v}) + I_V(\boldsymbol{v}) = S(\boldsymbol{v}) + \boldsymbol{v} = \boldsymbol{v}$ and therefore $S(\boldsymbol{v}) =$. Conversely, if $\boldsymbol{v} \in Ker(S)$ then $T(\boldsymbol{v}) = (S + I_V)(\boldsymbol{v}) = S(\boldsymbol{v}) + I_V(\boldsymbol{v}) = \boldsymbol{0} + \boldsymbol{v} = \boldsymbol{v}$.

b) Let $\boldsymbol{v} \in V$. Then

$$T(\boldsymbol{v}) - \boldsymbol{v} = T(\boldsymbol{v}) - I_V(\boldsymbol{v}) = (T - I_V)(\boldsymbol{v}) = S(\boldsymbol{v}).$$

Since S^2 is the zero map, $S(\boldsymbol{v}) \in Ker(S) = U$. Therefore $T(\boldsymbol{v}) - \boldsymbol{v} \in U$ which is equivalent to $T(\boldsymbol{v} + U) = \boldsymbol{v} + U$.

6. Define a map $T : U \oplus V \to (U/X) \oplus (V/Y)$ by

$$T(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u} + X, \boldsymbol{v} + Y).$$

This map is surjective and has kernel $X \oplus Y$. By Theorem (2.17 $(U \oplus V)/(X \oplus Y)$ is isomorphic to $U/X \oplus V$.

7 a.) Need to show that Γ is closed under addition and scalar multiplication. Suppose $\boldsymbol{v}, \boldsymbol{w} \in V$. Then $(\boldsymbol{v}, T(\boldsymbol{v}))$ and $(\boldsymbol{w}, T(\boldsymbol{w}))$ are two typical elements of Γ. Since T is linear

$$(\boldsymbol{v}, T(\boldsymbol{v})) + (\boldsymbol{w} + T(\boldsymbol{w})) = (\boldsymbol{v} + \boldsymbol{w}, T(\boldsymbol{v}) + T(\boldsymbol{w})) =$$

$$(\boldsymbol{v} + \boldsymbol{w}, T(\boldsymbol{v} + \boldsymbol{w})).$$

b) The subspace $V_1 = \{(\boldsymbol{v}, \boldsymbol{0}) | \boldsymbol{v} \in V\}$ is a complement to Γ in $V \oplus V$ and isomorphic to V. It then follows from Theorem (2.19) that

$$V/\Gamma = (V_1 + \Gamma)/\Gamma \cong V_1/(V_1 \cap \Gamma) = V_1$$

the last equality since $V_1 \cap \Gamma = \{\boldsymbol{0}\}$.

8. Since $U + W$ is a subspace of V, $(U + W)/W$ is a subspace of V/W. By hypothesis, $dim(V/W) = n$ and therefore $dim((U + W)/W) \leq n$.

By Theorem (2.19)

$$(U + W)/W \cong U/(U \cap W).$$

We therefore conclude that $dim(U/(U \cap W) \leq n$.

Now by Theorem (2.18)

$$(V/(U \cap W))/(U/(U \cap W) \cong V/U.$$

Thus, $dim(V/(U \cap W))/(U/(U \cap W) = dim(V/U) = m$. It now follows that $dim(V/(V \cap W)) \leq dim(U/(U \cap W)) + dim(V/U) \leq m + n$.

2.4. Matrix of a Linear Transformation

1. Assume T is onto. It follows from Exercise 13 of Section (2.1) that $Span(T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n)) = W$. By Exercise 5 of Section (1.8) it follows that $Span([T(\boldsymbol{v}_1)]_{\mathcal{B}_W}, \ldots, [T(\boldsymbol{v}_n)]_{\mathcal{B}_W}) = \mathbb{F}^m$. However, the coordinate vectors $[T(\boldsymbol{v}_j)]_{\mathcal{B}_W}, j = 1, \ldots, n$ are just the columns of A.

Conversely, assume the columns of A span $\mathbb{F}^m$. Then $([T(\boldsymbol{v}_1)]_{\mathcal{B}_W}, \ldots, [T(\boldsymbol{v}_n)]_{\mathcal{B}_W})$ spans $\mathbb{F}^m$. By Exericise 1 it follows that $(T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n))$ spans W.

2. Assume T is injective. Then by Theorem (2.11) $(T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_n))$ is linearly independent. Then by Theorem (1.30) $([T(\boldsymbol{v}_1)]_{\mathcal{B}_W}, \ldots, [T(\boldsymbol{v}_n)]_{\mathcal{B}_W})$ is linearly independent in $\mathbb{F}^n$. However, $([T(\boldsymbol{v}_1)]_{\mathcal{B}_W}, \ldots, [T(\boldsymbol{v}_n)]_{\mathcal{B}_W})$ is the sequence of columns of the matrix A.

Conversely, assume that the sequence of columns of the matrix A is linearly independent in $\mathbb{F}^n$.

By definition of the matrix A this means that $([T(\boldsymbol{v}_1)]_{\mathcal{B}_W}, \dots, [T(\boldsymbol{v}_n)]_{\mathcal{B}_W})$ is linearly independent in $\mathbb{F}^n$. By Theorem (1.30) we can conclude that $(T(\boldsymbol{v}_1), \dots, T(\boldsymbol{v}_n))$ is linearly independent in W. Finally, since $\mathcal{B}_V = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is a basis for V, by Theorem (2.11), it follows that T is injective.

3. The matrix $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is non-zero but $A^2 = 0_{2\times 2}$. Let T be the operator on $\mathbb{R}^2$ such that with respect to the standard basis $(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix})$ it has matrix A. Thus,

$$T(\begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} y \\ 0 \end{pmatrix}.$$

4. There are lots of examples. Here is one possible pair: $(A, B) = (\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix})$.

5. $\mathcal{M}_T(\mathcal{S}, \mathcal{S}) = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

6. Let $\mathcal{S}_n$ be the standard basis of $\mathbb{F}^n$ and $\mathcal{S}_m$ be the standard basis of $\mathbb{F}^m$. Set $A = \mathcal{M}_T(\mathcal{S}_n, \mathcal{S}_m)$. Then for any vector $\boldsymbol{v} \in \mathbb{F}^n, T(\boldsymbol{v}) = A\boldsymbol{v}$.

7. Let $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ such that $A = \mathcal{M}_T(\mathcal{S}_n, \mathcal{S}_m)$. Since the columns of A span $\mathbb{F}^m$ by Exercise 1 T is surjective. By Exercise 13 of Section (2.2) there is an linear transformation $S : \mathbb{F}^m \to \mathbb{F}^n$ such that $TS = I_{\mathbb{F}^m}$. Let $B = \mathcal{M}_S(\mathcal{S}^m, \mathcal{S}^n)$. It then follows that $BA = \mathcal{M}_{I_{\mathbb{F}^m}}(\mathcal{S}^m, \mathcal{S}^m) = I_m$.

8. Let $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ such that $A = \mathcal{M}_T(\mathcal{S}^n, \mathcal{S}^m)$. Since the sequence of columns of A is linearly independent, by Exercise 2 the operator T is injective. Then by Exercise 14 of Section (2.2) there is linear transformation $S : \mathbb{F}^m \to \mathbb{F}^n$ such that $TS = I_{\mathbb{F}^n}$. Set $B = \mathcal{M}_S(\mathcal{S}^m, \mathcal{S}^n)$. Then $AB = \mathcal{M}_{I_{\mathbb{F}^n}}(\mathcal{S}^n, \mathcal{S}^n) = I_n$.

9. The reduced echelon form of A is $\begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Since every row is non-zero the columns of A span $\mathbb{Q}^3$.

One such matrix is $\begin{pmatrix} 4 & -2 & -1 \\ -5 & 3 & 2 \\ 0 & 0 & 0 \\ 2 & -1 & -1 \end{pmatrix}$.

10. The reduced echelon form of A is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

Since the columns of this matrix are linearly independent the columns of A are linearly independent.

One such matrix is $\begin{pmatrix} 2 & -1 & -2 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix}$.

11. Since $A = \mathcal{M}_T(\mathcal{B}_V, \mathcal{B}_W)$ it follows that $[T(\boldsymbol{v})]_{\mathcal{B}_W} = A[\boldsymbol{v}]_{\mathcal{B}_V}$. Suppose $\boldsymbol{v} \in Ker(T)$ so that $T(\boldsymbol{v}) = \mathbf{0}_W$. Thus $A[\boldsymbol{v}]_{\mathcal{B}_V} = \mathbf{0}$ from which we conclude that $[\boldsymbol{v}]_{\mathcal{B}_V} \in null(A)$. On the other hand if $[\boldsymbol{v}]_{\mathcal{B}_V} \in null(A)$ then $[T(\boldsymbol{v})]_{\mathcal{B}_W} = \mathbf{0}$ from which we conclude that $T(\boldsymbol{v}) = \mathbf{0}_W$ and $\boldsymbol{v} \in Ker(T)$.

2.5. The Algebra of $\mathcal{L}(V, W)$ and $M_{mn}(\mathbb{F})$

1. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Then

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

Now let $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ be linearly independent in V and xtend to a basis $\mathcal{B} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for V. Let S be the linear operator on V such that $S(\boldsymbol{v}_1) = \boldsymbol{v}_1, S(\boldsymbol{v}_2) = \boldsymbol{v}_1 + \boldsymbol{v}_2$ and for $S(\boldsymbol{v}_i) = \boldsymbol{v}_i$ for $3 \le i \le n$.

Let T be the linear operator on V such that $T(\boldsymbol{v}_1) = \boldsymbol{v}_1 + \boldsymbol{v}_2, T(\boldsymbol{v}_2) = \boldsymbol{v}_2$ and $T(\boldsymbol{v}_i) = \boldsymbol{v}_i$ for $3 \le i \le n$.

Then $(ST)(\boldsymbol{v}_1) = S(T(\boldsymbol{v}_1)) = S(\boldsymbol{v}_1 + \boldsymbol{v}_2) = S(\boldsymbol{v}_1) + S(\boldsymbol{v}_2) = \boldsymbol{v}_1 + (\boldsymbol{v}_1 + \boldsymbol{v}_2) = 2\boldsymbol{v}_1 + \boldsymbol{v}_2$.

$(TS)(\boldsymbol{v}_1) = T(S(\boldsymbol{v}_1)) = T(\boldsymbol{v}_1) = \boldsymbol{v}_1 + \boldsymbol{v}_2 \neq (ST)(\boldsymbol{v}_1)$.

2. Let $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ be linearly independent and extend to a basis $\mathcal{B} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ for V. Let S be the operator on V such that $S(\boldsymbol{v}_1) = \boldsymbol{v}_1 + \boldsymbol{v}_2, S(\boldsymbol{v}_2) = \boldsymbol{v}_1 + \boldsymbol{v}_2$ and $S(\boldsymbol{v}_i) = \boldsymbol{0}_V$ for $3 \leq i \leq n$.

Let T be the operator on V such that $T(\boldsymbol{v}_1) = \boldsymbol{v}_1 - \boldsymbol{v}_2, T(\boldsymbol{v}_2) = \boldsymbol{v}_1 - \boldsymbol{v}_2$ and $T(\boldsymbol{v}_i) = \boldsymbol{0}_V$ for $3 \leq i \leq n$.

If $\boldsymbol{v} \in \mathcal{B} \setminus \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ then $(ST)(\boldsymbol{v}) = S(T(\boldsymbol{v})) = S(\boldsymbol{0}) = \boldsymbol{0}$. On the other hand

$$(ST)(\boldsymbol{v}_1) = S(T(\boldsymbol{v}_1)) = S(\boldsymbol{v}_1 - \boldsymbol{v}_2) =$$
$$S(\boldsymbol{v}_1) - S(\boldsymbol{v}_2) = (\boldsymbol{v}_1 + \boldsymbol{v}_2) - (\boldsymbol{v}_1 + \boldsymbol{v}_2) = \boldsymbol{0}.$$

$$(ST)(\boldsymbol{v}_2) = S(T(\boldsymbol{v}_2)) = S(\boldsymbol{v}_1 - \boldsymbol{v}_2) = \boldsymbol{0}.$$

3. Since $1\boldsymbol{a} = \boldsymbol{a}1$ the identity of A is in $C_A(\boldsymbol{a})$ and so $C_A(\boldsymbol{a})$ has an identity. We need to show that $C_A(\boldsymbol{a})$ is closed under addition, multiplication and scalar multiplication.

Suppose $\boldsymbol{b}, \boldsymbol{c} \in C_A(\boldsymbol{a})$. Then $(\boldsymbol{b} + \boldsymbol{c})\boldsymbol{a} = \boldsymbol{ba} + \boldsymbol{ca} = \boldsymbol{ab} + \boldsymbol{ac} = \boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c})$. So $\boldsymbol{b} + \boldsymbol{c} \in C_A(\boldsymbol{a})$.

We also have $(\boldsymbol{bc})\boldsymbol{a} = \boldsymbol{b}(\boldsymbol{ca}) = \boldsymbol{b}(\boldsymbol{ac}) = (\boldsymbol{ba})\boldsymbol{c} = (\boldsymbol{ab})\boldsymbol{c} = \boldsymbol{a}(\boldsymbol{bc})$ and so $\boldsymbol{bc} \in C_A(\boldsymbol{a})$. Finally, if d is a scalar we have $(d\boldsymbol{b})\boldsymbol{c} = d(\boldsymbol{ba}) = d(\boldsymbol{ab}) = \boldsymbol{a}(d\boldsymbol{b})$.

4. We prove the only non-zero ideal in $M_{nn}(\mathbb{F})$ is $M_{nn}(\mathbb{F})$. Suppose J is an ideal and A is a matrix with entries a_{ij} is in J and A is not the zero matrix. Then for some $a_{ij} \neq 0$. Let E_{ii} be the matrix with zeros in all entries except the $(i,i)-$entry which is a 1. The matrix $E_{ii}AE_{jj} = a_{ij}E_{ij}$ is in J since J is an ideal. Since $a_{ij} \neq 0$ we can multiply by the reciprocal and therefore $E_{ij} \in J$.

Now let P_{kl} be the matrix which is obtained from the identity matrix by exchanging the k and l columns (rows). If B is an $n \times n$ matrix then $P_{kl}B$ the matrix obtained from B by exchanging the k and l rows and BP_{kl} is the matrix obtained from B by exchanging the k and l columns. It then follows that $P_{ik}E_{ij}P_{jl} = E_{kl}$ is in J. However, the matrices E_{kl} span $M_{nn}(\mathbb{F})$ and it follows that $J = M_{nn}(\mathbb{F})$. Since $M_{nn}(\mathbb{F})$ and $\mathcal{L}(V,V)$ are isomorphic as algebras the only ideals in $\mathcal{L}(V,V)$ are the zero ideal and all of $\mathcal{L}(V,V)$.

5. Clearly the sum of two upper triangular matrices is upper triangular. Consider the product of two upper triangular matrices A and B. The $(i,j)-$entry of AB is the dot product of the i^{th} row of A with the j^{th} for of B:

$$\begin{pmatrix} 0 & \dots & 0 & a_{ii} & \dots & a_{in} \end{pmatrix} \begin{pmatrix} b_{1j} \\ \vdots \\ b_{jj} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

If $i > j$ then the product is zero and AB is upper triangular.

6. $\overline{U}_{nn}(\mathbb{F})$ is a subspace. We need only need to show that the diagonal elements of a product of a strict upper triangular matrix and a triangular matrix are zero. Suppose A is strictly upper triangular and B is upper triangular. Then the $(i,i)-$entry of AB is

$$\begin{pmatrix} 0 & \dots & 0 & a_{i+1,i+1} & \dots & a_{in} \end{pmatrix} \begin{pmatrix} b_{1j} \\ \vdots \\ b_{jj} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0.$$

Similarly the diagonal entries of BA are zero.

7. Assume $T \in \mathcal{L}(V,V)$ is not a unit. Then T is not injective by the half is good enough theorem and $Ker(T) \neq \{\boldsymbol{0}\}$. Let $\boldsymbol{v}$ be a non-zero vector in $Ker(T)$. Choose a basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ and let S be the operator such that $S(\boldsymbol{v}_i) = \boldsymbol{v}$ for all i. Then S is not the zero operator but $Range(S) = Span(\boldsymbol{v})$. It then follows that TS is the zero operator.

2.6. Invertible Transformations and Matrices

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

2. $S^{-1}(a+bx+cx^2) = (b-c) + (2a-b)x + (-3a+b+c)x^2$.

3. Set $\boldsymbol{w}_i = T(\boldsymbol{v}_i)$. Assume that $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ is a basis for W. Then there exists a unique linear transformation $S : W \to V$ such that $S(\boldsymbol{w}_i) = \boldsymbol{v}_i$. Then $ST(\boldsymbol{v}_i) = \boldsymbol{v}_i$ and since ST is linear, $ST = I_V$. Similarly, $TS = I_W$. So in this case T is an isomorphism. On the other hand, assume T is an isomorphism. Then T is injective and so $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ is linearly independent. Also, T is surjective. However, $Range(T) = Span(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$. Thus, $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ is a basis of W.

4. This follows from Exercise 3.

5. From Exercise 4 we need to count the number of bases $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ there are in $\mathbb{F}_2^3$. There are 7 non-zero vectors and $\boldsymbol{v}_1$ can be any of these vectors. $\boldsymbol{v}_2$ can be any vector except $\mathbf{0}$ and $\boldsymbol{v}_1$ and so there are 8 - 2 = 6 choices for $\boldsymbol{v}_2$. Having chosen $\boldsymbol{v}_1, \boldsymbol{v}_2$ we can choose $\boldsymbol{v}_3$ to be any vector not in $Span(\boldsymbol{v}_1, \boldsymbol{v}_2) = \{\mathbf{0}, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_1 + \boldsymbol{v}_2\}$. So, there are 8 - 4 = 4 choices for $\boldsymbol{v}_3$. So the number of bases is

$$7 \times 6 \times 4 = 168.$$

6. From Exercise 4 we need to count the number of bases $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ there are in $\mathbb{F}_3^3$. There are 26 non-zero vectors and $\boldsymbol{v}_1$ can be any of these vectors. $\boldsymbol{v}_2$ can be any vector except $\mathbf{0}$ and $\pm\boldsymbol{v}_1$ and so there are 27 - 3 = 24 choices for $\boldsymbol{v}_2$. Having chosen $\boldsymbol{v}_1, \boldsymbol{v}_2$ we can choose $\boldsymbol{v}_3$ to be any vector not in $Span(\boldsymbol{v}_1, \boldsymbol{v}_2)$. There are 9 vectors in $span(\boldsymbol{v}_1, \boldsymbol{v}_2)$ and so there are 27 - 9 = 18 choices for $\boldsymbol{v}_3$. So the number of bases is

$$26 \times 24 \times 18 = 11232 = 2^5 3^3 13.$$

7. This is the same as Exercise 7 of Section (2.5).

8. Assume S, T are invertible operators on the space V. Then

$$(ST)(T^{-1}S^{-1}) = S[T(T^{-1}S^{-1})] =$$
$$S[(TT^{-1})S^{-1}] = S[I_V S^{-1}] = SS^{-1} = I_V.$$

$$(T^{-1}S^{-1})(ST) = T^{-1}[S^{-1}(ST)] =$$
$$T^{-1}[(S^{-1}S)T] = T^{-1}[I_V T] = T^{-1}T = I_V.$$

9. By the distributive property, $\widehat{S}$ is an operator on $\mathcal{L}(V, V)$. Let $T \in \mathcal{L}(V, V)$. Then

$$\widehat{S^{-1}}\widehat{S}(T) = S^{-1}(ST) = (S^{-1}S)T = I_V T = T$$

$$\widehat{S}\widehat{S^{-1}}T = S(S^{-1}T) = (SS^{-1})T = I_V T = T.$$

So $\widehat{S}$ is an invertible operator with inverse $\widehat{S^{-1}}$.

10. $I_V - S)(I_V + S + \cdots + S^{k-1}) =$

$$I_V + S + \cdots + S^{k-1} - (S + \cdots + S^{k-1} + S^k) =$$

$$I_V - S^k = I_V.$$

A similar calculation gives

$$(I_V + S + \cdots + S^{k-1})(I_V - S) = I_V.$$

11. Every operator T is similar to itself via the identity: $I_V T I_V^{-1} = T$. Therefore the relation is reflexive.

Assume the operators S and T are similar via Q, that is, $T = QSQ^{-1}$. Then $S = Q^{-1}TQ = Q^{-1}T(Q^{-1})^{-1}$. So, setting $P = Q^{-1}$ we have $S = PTP^{-1}$. The relation of similarity is symmetric.

Assume $S = QRQ^{-1}$ and $T = PSP^{-1}$. Then

$$T = P(QRQ^{-1})P^{-1} = (PQ)R(Q^{-1}P^{-1}) =$$

$$(PQ)R(PQ)^{-1}$$

which implies that similarity is a transitive relation.

12. Suppose $T_2 = QT_1Q^{-1}$. Then

$$\mathcal{M}_{T_2}(\mathcal{B},\mathcal{B}) = \mathcal{M}_Q(\mathcal{B},\mathcal{B})\mathcal{M}_{T_1}(\mathcal{B},\mathcal{B})\mathcal{M}_{Q^{-1}}(\mathcal{B},\mathcal{B}).$$

Set $A = \mathcal{M}_Q(\mathcal{B},\mathcal{B})$ then $\mathcal{M}_{Q^{-1}}(\mathcal{B},\mathcal{B}) = A^{-1}$. Thus,

$$\mathcal{M}_{T_2}(\mathcal{B},\mathcal{B}) = A\mathcal{M}_{T_1}(\mathcal{B},\mathcal{B})A^{-1}$$

and so the matrices are similar.

13. We are assuming there is an invertible matrix A such that $\mathcal{M}_{T_2}(\mathcal{B},\mathcal{B}) = A\mathcal{M}_{T_1}(\mathcal{B},\mathcal{B})A^{-1}$. Let Q be the operator on V such that the matrix of Q with respect to $\mathcal{B}$ is A :

$$\mathcal{M}_Q(\mathcal{B},\mathcal{B}) = A.$$

Let $T' = QT_1Q^{-1}$. Then

$$\mathcal{M}_{T'}(\mathcal{B},\mathcal{B}) = \mathcal{M}_Q(\mathcal{B},\mathcal{B})\mathcal{M}_{T_1}(\mathcal{B},\mathcal{B})\mathcal{M}_{Q^{-1}}(\mathcal{B},\mathcal{B})$$

$$= \mathcal{M}_Q(\mathcal{B},\mathcal{B})\mathcal{M}_{T_1}(\mathcal{B},\mathcal{B})\mathcal{M}_Q(\mathcal{B},\mathcal{B})^{-1}$$

$$A\mathcal{M}_{T_1}(\mathcal{B},\mathcal{B})A^{-1} = \mathcal{M}_{T_2}(\mathcal{B},\mathcal{B})$$

It follows that $T' = T_2$ and since $T' = QT_1Q^{-1}$, T_2 and T_1 are similar.

14. $\mathcal{M}_{T_2}(\mathcal{B},\mathcal{B})$ is similar to $\mathcal{M}_{T_2}(\mathcal{B}',\mathcal{B}')$. Since $\mathcal{M}_{T_1}(\mathcal{B},\mathcal{B})$ is similar to $\mathcal{M}_{T_2}(\mathcal{B}',\mathcal{B}')$ by hypothesis, by transitivity, $\mathcal{M}_{T_1}(\mathcal{B},\mathcal{B})$ and $\mathcal{B}_{T_2}(\mathcal{B},\mathcal{B})$ are similar. Now by Exercise 13, T_1 and T_2 are similar operators.

Chapter 3

Polynomials

3.1. The Algebra of Polynomials

1. $x^2 + 1$.

2. Assume $F(x), G(x)$ are in J. Then there are $a_i(x), b_i(X), i = 1, 2$ such that

$$F(x) = a_1(x)f(x) + b_1(x)g(x),$$
$$G(x) = a_2(x)f(x) + b_2(x)g(x).$$

Then

$$F(x) + G(x) =$$
$$[a_1(x) + a_2(x)]f(x) + [b_1(x) + b_2(x)]g(x) \in J$$

3. Assume $F(x) \in J$ and $c(x) \in \mathbb{F}[x]$. We need to show $c(x)F(x) \in J$. Now there are $a(x), b(x) \in \mathbb{F}[x]$ such that $F(x) = a(x)f(x) + b(x)g(x)$. Then

$$c(x)F(x) = [c(x)a(x)]f(x) + [c(x)b(x)]g(x)$$

4. Suppose $d(x)$ is monic and has least degree in J. Want to show that every element of J is a multiple of $d(x)$. Let $f(x) \in J$. Apply the division algorithm to write $f(x) = q(x)d(x) + r(x)$ where either $r(x)$ is the zero polynomial or $deg(r(x)) < deg(d(x))$. Suppose to the contrary that $r(x) \neq 0$. Then $r(x) = f(x) - q(x)d(x)$. By the definition of an ideal $-q(x)d(x) \in J$. Since $f(x), -q(x)d(x) \in J$ we get $r(x) = f(x) - q(x)d(x) \in J$. However $deg(r(x)) < deg(d(x))$ which contradicts the minimality of the degree of $d(x)$. So, $r(x)$ is the zero polynomial and $d(x)$ divides $f(x)$.

Now suppose $d(x), d'(x)$ are both minimal degree and monic. Then $d(x)$ divides $d'(x)$ and $d'(x)$ divides $d(x)$. This implies there is a scalar $c \in \mathbb{F}$ such that $d'(x) = cd(x)$. Since both are monic, $c = 1$ and they are equal.

5. Let $a(x)$ and $b(x)$ be polynomials such that $a(x)f(x) + b(x)g(x) = d(x)$. We then have

$$a(x)f'(x)d(x) + b(x)g'(x)d(x) = d(x)$$

$$a(x)f'(x) + b(x)g'(x) = 1$$

We therefore conclude that $f'(x)$ and $g'(x)$ are relatively prime.

6. Write $f(x) = f'(x)d(x), g(x) = g'(x)d(x)$. Then f', g' are relatively prime. Let $l(x)$ be the least common multiple of $f(x)$ and $g(x)$.

Now $\frac{f(x)g(x)}{d(x)} = f'(x)g'(x)d(x)$ is divisible by $f(x)$ and $g(x)$ and therefore $l(x)$ divides $\frac{f(x)g(x)}{d(x)}$. On the other hand, let $l(x) = l'(x)d(x)$. Since $f(x) = f'(x)d(x)$ divides $l(x) = l'(x)d(x)$ we conclude that $f'(x)$ divides $l'(x)$. Similarly, $g'(x)$ divides $l'(x)$. Since $f'(x)g'(x)$

are relatively prime it then follows that $f'(x)g'(x)$ divides $l'(x)$. Consequently, $\frac{f(x)g(x)}{d(x)} = f'(x)g'(x)$ divides $l'(x)d(x) = l(x)$. It now follows that $\frac{f(x)g(x)}{d(x)}$ is a scalar multiple of $l(x)$.

7. If $l(x)$ and $l'(x)$ are both lcms of $f(x), g(x)$ then $l(x)|l'(x)$ and $l'(x)|l(x)$ so that $l'(x) = cl(x)$ for some $c \in \mathbb{F}$. Since both are monic, $c = 1$ and $l(x) = l'(x)$.

8. Without loss of generality we can assume that $f(x)$ is monic. Let $d(x)$ be the gcd of $f(x)$ and $g(x)$. Since $f(x)$ is irreducible and $d(x)$ divides $f(x)$ either $d(x) = 1$ or $d(x) = f(x)$. However, in the latter case, $f(x) = d(x)$ divides $g(x)$, contrary to the hypothesis. Thus, $d(x) = 1$ and $f(x), g(x)$ are relatively prime.

9. This follows from Exercise 6.

10. Let $d(x)$ be the greatest common divisor of $f(x)$ and $g(x)$ in $\mathbb{F}[x]$ and assume, to the contrary that $d(x) \neq g(x)$. Write $f(x) = f'(x), g(x) = g'(x)d(x)$. As in the proof of Exercise 6 there are $a(x), b(x) \in \mathbb{F}[x]$ such that $a(x)f'(x) + b(x)g'(x) = 1$. Since $\mathbb{F} \subset \mathbb{K}$ it follows that $f'(x)$ and $g'(x)$ are relatively prime in $\mathbb{K}[x]$. But then the gcd of $f(x)$ and $g(x)$ in $\mathbb{K}[x]$ is $d(x)$ contrary to the assumption that $g(x)$ divides $f(x)$ in $\mathbb{K}[x]$. We therefore have a contradiction and consequently, $d(x) = g(x)$ as required.

11. A polynomial $g(x)$ divides$f(x)$ if and only if $g(x)$ has a factorization $cp_1(x)^{f_1} \dots p_t(x)^{f_t}$ where $c \in \mathbb{F}$ and f_i are in $\mathbb{N} \cup \{0\}$ and $f_i \leq e_i$. There are then $e_i + 1$ choices for f_i and hence $(e_1 + 1) \times \cdots \times (e_t + 1)$ choices for $(f_1, \dots, f_t)$. For any such t there is a unique c such that the polynomial is monic.

3.2. Roots of Polynomials

1. Since non-real, complex roots of a real polynomial come in conjugate pairs the number of non-real, complex roots of a real polynomial is always even. Therefore, if the degree of a real polynomial is $2n+1$, an odd number, there must be a real root.

2. $x^4 + 5x^2 + 4$. Another, which is irreducible over the rational numbers is $x^4 + 2x^2 + 2$.

3. Note that complex conjugation satisfies

$$\overline{z+w} = \overline{z} + \overline{w}, \overline{zw} = \overline{z}\,\overline{w} \tag{3.1}$$

If λ is a root of $f(x)$ then

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0 \tag{3.2}$$

Taking the complex conjugate of both sides of (3.2) and using (3.1) we obtain

$$\overline{\lambda^n} + \overline{a_{n-1}\lambda^{n-1}} + \cdots + \overline{a_1\lambda} + \overline{a_0} = \overline{0}$$

$$\overline{\lambda^n} + \overline{a_{n-1}}\,\overline{\lambda^{n-1}} + \cdots + \overline{a_1}\,\overline{\lambda} + \overline{a_0} = 0$$

$$\overline{\lambda}^n + \overline{a_{n-1}}\,\overline{\lambda}^{n-1} + \cdots + \overline{a_1}\,\overline{\lambda} + \overline{a_0} = 0.$$

Thus, $\overline{\lambda}$ is a root of $\overline{f}(x)$.

4. $x^4 - 6x^3 + 15x^- 18x + 10$.

5. If $3 + 4i$ is a root of $f(x)$ then so is $\overline{3+4i} = 3 - 4i$. Then $(x - [3+4i])(x - [3-4i]) = x^2 - 3x + 25$ is a factor of $f(x)$. Likewise $x^2 - 6x + 25$ is a factor of $g(x)$.

6. Let $n = max\{deg(f(x), g(x)\}$. Then $f(x), g(x) \in \mathbb{F}_{(n)}[x]$ and we can write $f(x) = \sum_{i=0}^n a_i x^i, g(x) = \sum_{i=0}^n b_i x^i$. Then $f(x) + g(x) = \sum_{i=0}^n (a_i + b_i)x^i$. By the definition of D we have

$$D(f(x)) = \sum_{i=1}^n i a_i x^{i-1}, D(g(x)) = \sum_{i=1}^n i b_i x^{i-1}.$$

Adding $D(f(x)) + D(g(x))$ we get

$$D(f(x)) + D(g(x)) = \sum_{i=1}^n (ia_i + ib_i)x^{i-1} =$$

$$\sum_{i=1}^{n} i(a_i + b_i)x^{i-1} = D(f(x) + g(x)).$$

7. Assume $f(x) = \sum_{i=0}^{n} a_i x^i$. Then $cf(x) = \sum_{i=0}^{n}(ca_i)x^i$. We then have

$$D(cf(x)) = \sum_{i=1}^{n} i(ca_i)x^{i-1} =$$

$$\sum_{i=1}^{n} c(ia_i)x^{i-1} = c\sum_{i=1}^{n} ia_i x^{i-1} = cD(f(x)).$$

8. Let k, l be natural numbers. We first prove that $D(x^{k+l}) = D(x^k)x^l + x^k D(x^l)$. By the definition of D we have

$$D(x^k)x^l + x^k D(x^l) = kx^{k-1}x^l + x^k(lx^{l-1}) =$$

$$kx^{k+l-1} + lx^{k+l-1} = (k+l)x^{k+l-1} =$$

$$D(x^{k+l}) = D(x^k \cdot x^l).$$

We next prove if $g(x) \in \mathbb{F}[x]$ and k is a natural number then $D(x^k g(x)) = D(x^k)g(x) + x^k D(g(x))$. Write $g(x) = \sum_{i=0}^{n} b_i x^i$. Then $x^k g(x) = \sum_{i=0}^{m} b_i(x^k \cdot x^i)$. By Exercises 6 and 7 we have

$$D(\sum_{i=0}^{m} b_i(x^k \cdot x^i) = \sum_{i=0}^{n} b_i D(x^k \cdot x^i).$$

By Exercise 7 we have

$$\sum_{i=0}^{n} b_i D(x^k \cdot x^i) = \sum_{i=0}^{n} b_i[D(x^k)x^i + x^k D(x^i)] =$$

$$\sum_{i=0}^{n} b_i D(x^k) + \sum_{i=0}^{n} b_i x^k D(x^i) =$$

$$D(x^k)\sum_{i=0}^{n} b_i x^i + x^k \sum_{i=0}^{n} b_i D(x^i) =$$

$$D(x^k)g(x) + x^k D(g(x)).$$

Now write $f(x) = \sum_{j=0}^{m} a_j x^j$. Then $f(x)g(x) = \sum_{j=0}^{m} a_j x^j g(x)$. By Exercises 6 and 7 we have

$$D(\sum_{j=0}^{m} a_j x^j g(x)) = \sum_{j=0}^{m} a_j D(x^j g(x)).$$

Then by what we showed immediately above we have

$$\sum_{j=0}^{m} a_j D(x^j g(x)) = \sum_{j=0}^{m} a_j[D(x^j)g(x) + x^j D(g(x))] =$$

$$\sum_{j=0}^{m} a_j D(x^j)g(x) + \sum_{j=0}^{m} a_j x^j D(g(x)) =$$

$$[\sum_{j=0}^{m}(ja_j)x^{j-1}]g(x) + \sum_{j=0}^{m} a_j x^j D(g(x)) =$$

$$D(f(x))g(x) + f(x)D(g(x)).$$

9. Suppose α is a root of multiplicity at least two. Then $(x-\alpha)^2$ divides $f(x)$. Then we can write $f(x) = (x-\alpha)^2 g(x)$. Then $D(f(x)) = 2(x-\alpha)g(x) + (x-\alpha)^2 g'(x)$. Consequently $(x-\alpha)$ is a factor of $D(f(x))$.

On the other hand, suppose $f(x)$ has n distinct roots, $\alpha_1, \ldots, \alpha_n$. Then $f(x) = (x-\alpha_1)\ldots(x-\alpha_n)$. Then

$$D(f(x)) = \sum_{i=1}^{n} \frac{f(x)}{(x-\alpha_i)}$$

Evaluating $D(f(x))$ at α_i we obtain $(\alpha_i - \alpha_1)\ldots(\alpha_i - \alpha_{i-1})(\alpha_i - \alpha_{i+1})\ldots(\alpha_i - \alpha_n) \neq 0$. Therefore $(x-\alpha_i)$ is not a factor of $D(f(x))$ and $f(x), D(f(x))$ are relatively prime.

10. If $i \neq j$ then $(x-\alpha_i)$ is a factor of $F_j(x)$ and hence $f_j(x)$ and therefore $f_j(\alpha_i) = 0$. On the other hand, $f_j(\alpha_j) = \frac{F_j(\alpha_j)}{F_j(\alpha_j)} = 1$. Now suppose

$$\sum_{ij=1}^{n} c_j f_j(x) = 0$$

Evaluating at α_i we obtain

$$\sum_{j=1} c_j f_j(\alpha_i) = c_i f_i(\alpha_i) = c_i.$$

Thus, $c_i = 0$ for all i and $\mathcal{B} = (f_1, \ldots, f_n)$ is linearly independent in $\mathbb{F}_{n-1}[x]$. Since $dim(\mathbb{F}_{n-1}[x]) = n$, $\mathcal{B}$ is a basis.

11. Set $f(x) = \sum_{i=j}^{n} \beta_j f_j(x)$. $f(x)$ satisfies the conditions. Suppose $g(x)$ does also. Then α_j is a root of $f(x) - g(x)$ for all j. Since $f(x) - g(x) \in \mathbb{F}_{n-1}[x]$ if $f(x) - g(x) \neq 0$ then it has at most $n-1$ roots. Therefore $f(x) - g(x) = 0$ and $g(x) = f(x)$.

12. If $g(x) \in \mathbb{F}_{n-1}[x]$ and $g(\alpha_j) = \beta_j$ then $g(x) = \sum_{j=1}^{n} \beta_j f_j(x)$ by Exercise 8. Since the coefficient of $f_j(x)$ is $g(\alpha_j)$ we conclude that

$$[g(x)]_{\mathcal{B}} = \begin{pmatrix} g(\alpha_1) \\ \vdots \\ g(\alpha_n) \end{pmatrix}$$

13. The coordinate vector of the constant function 1 is $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. The coordinate vector of x^k is

$$[x^k]_{\mathcal{B}} = \begin{pmatrix} \alpha_1^k \\ \alpha_2^k \\ \vdots \\ \alpha_n^k \end{pmatrix}.$$

It follows that the change of basis matrix from $\mathcal{S}$ to $\mathcal{B}$ is

$$\mathcal{M}_{I_{\mathbb{F}_{n-1}[x]}}(\mathcal{S}, \mathcal{B}) = \begin{pmatrix} 1 & \alpha_1 & \alpha_2^2 & \ldots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \alpha_2^2 & \ldots & \alpha_2^{n-1} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & \alpha_n & \alpha_n^2 & \ldots & \alpha_n^{n-1} \end{pmatrix}$$

When $\alpha_1, \ldots, \alpha_n$ are distinct this matrix is invertible.

Chapter 4

Theory of a Single Linear Operator

4.1. Invariant Subspaces of an Operator

1. $T(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}) = \begin{pmatrix} x_1 \\ x_2 \\ x_1 + x_2 \end{pmatrix}$.

2. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ be a basis for U. Extend this to a basis $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ of V. Let T be the operator such that $T(\boldsymbol{u}_i) = \boldsymbol{u}_n$ for all i.

3. $x^2 - 3x + 2$.

4. $x^3 - 2x^2 - x + 2$.

5. $S(T(\boldsymbol{v})) = (ST)(\boldsymbol{v}) = (TS)(\boldsymbol{v}) = T(S(\boldsymbol{v})) = T(\lambda \boldsymbol{v}) = \lambda T(\boldsymbol{v})$.

6. Let $\widehat{T} : U \to U$ be defined by $\widehat{T}(\boldsymbol{u}) = T(\boldsymbol{u})$. Since T is invertible on V, in particular, T is injective. Then $\widehat{T}$ is injective. Since V is finite dimensional it follows that U is finite dimensional. Consequently by the half is good enough theorem $\widehat{T}$ is surjective. Now let $\boldsymbol{u} \in U$ be arbitrary. By what we have shown there is a $\boldsymbol{u}' \in U$ such that $T(\boldsymbol{u}') = \boldsymbol{u}$. Then $T^{-1}(\boldsymbol{u}) = T^{-1}(T(\boldsymbol{u}')) = (T^{-1}T)(\boldsymbol{u}') = I_V(\boldsymbol{u}') = \boldsymbol{u}' \in U$. Thus, U is T^{-1} invariant.

7. Suppose $\boldsymbol{v} \in E_1 \cap E_{-1}$. Then $T(\boldsymbol{v}) = \boldsymbol{v}$ and $T(\boldsymbol{v}) = -\boldsymbol{v}$. Then $\boldsymbol{v} = -\boldsymbol{v}$ or $2\boldsymbol{v} = \boldsymbol{0}$. Since $2 \neq 0, \boldsymbol{v} = \boldsymbol{0}$. So $E_1 \cap E_{-1} = \{\boldsymbol{0}\}$. So we need to show that $V = E_1 + E_{-1}$. Let $\boldsymbol{v} \in V$ be arbitrary. Set $\boldsymbol{x} = \frac{1}{2}(\boldsymbol{v} + T(\boldsymbol{v}))$ and $\boldsymbol{y} = \frac{1}{2}(\boldsymbol{v} - T(\boldsymbol{v}))$. We claim $\boldsymbol{x} \in E_1, \boldsymbol{y} \in E_{-1}$.

$$T(\boldsymbol{x}) = T(\frac{1}{2}(\boldsymbol{v} + T(\boldsymbol{v}))) = \frac{1}{2}T(\boldsymbol{v} + T(\boldsymbol{v}))$$

$$= \frac{1}{2}(T(\boldsymbol{v}) + T^2(\boldsymbol{v})) = \frac{1}{2}(T(\boldsymbol{v}) + \boldsymbol{v}) = \boldsymbol{x}.$$

$$T(\boldsymbol{y}) = T(\frac{1}{2}(\boldsymbol{v} - T(\boldsymbol{v}))) = \frac{1}{2}T(\boldsymbol{v} - T(\boldsymbol{v}))$$

$$= \frac{1}{2}(T(\boldsymbol{v}) - T^2(\boldsymbol{v})) = \frac{1}{2}(T(\boldsymbol{v}) - \boldsymbol{v}) = -\boldsymbol{y}.$$

Now $\boldsymbol{v} = \boldsymbol{x} + \boldsymbol{y} \in E_1 + E_{-1}$.

8. In addition to $\mathbb{R}^3$ and $\boldsymbol{0}$ the invariant subspaces are $Span(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix})$ and $\{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} | x_1 + x_2 + x_3 = 0\}$.

9. The $T-$invariant subspaces are $\{\boldsymbol{0}\}, Span(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}), Span(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}), \mathbb{R}^3$.

10. This follows since for any polynomials f, g we have $f(T) + g(T) = (f+g)(T)$ and $f(T)g(T) = h(T)$ where $h(x) = f(x)g(x)$.

11. Need to show that i) If $f(x), g(x) \in Ann(T, \boldsymbol{v})$ then $f(x) + g(x) \in Ann(T, \boldsymbol{v})$; and ii) If $f(x) \in Ann(T, \boldsymbol{v})$ and $h(x) \in \mathbb{F}[x]$ then $h(x)f(x) \in Ann(T, \boldsymbol{v})$.

i) $[(f+g)(T)](\boldsymbol{v}) = [f(T) + g(T)](\boldsymbol{v}) = f(T)(\boldsymbol{v}) + g(T)(\boldsymbol{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$.

ii) $[h(T)f(T)](\boldsymbol{v}) = h(T)(f(T)(\boldsymbol{v})) = h(T)(\mathbf{0}) = \mathbf{0}$.

12. Let $\boldsymbol{u} \in U, \boldsymbol{w} \in W$. Then $T(\boldsymbol{u}+\boldsymbol{w}) = T(\boldsymbol{u})+T(\boldsymbol{w})$. By hypothesis, $T(\boldsymbol{u}) \in U$ and $T(\boldsymbol{w}) \in W$ and therefore $T(\boldsymbol{u}) + T(\boldsymbol{w}) \in U + W$. Thus, $U + W$ is $T-$invariant.

Suppose $\boldsymbol{x} \in U \cap W$. Since $\boldsymbol{x} \in U$ and U is $T-$invariant $T(\boldsymbol{x}) \in U$. Similarly, $T(\boldsymbol{x}) \in W$ and hence $T(\boldsymbol{x}) \in U \cap W$.

13. Need to show that i) If $f(x), g(x) \in Ann(T)$ then $f(x) + g(x) \in Ann(T)$; and ii) If $f(x) \in Ann(T)$ and $h(x) \in \mathbb{F}[x]$ then $h(x)f(x) \in Ann(T)$.

i) Let $\boldsymbol{v}$ be an arbitrary vector. We then have

$$[(f+g)(T)](\boldsymbol{v}) = [f(T) + g(T)](\boldsymbol{v}) =$$
$$f(T)(\boldsymbol{v}) + g(T)(\boldsymbol{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0}.$$

Since $\boldsymbol{v}$ is an arbitrary vector $f + g \in Ann(T)$.

ii) Let $\boldsymbol{v}$ be an arbitrary vector. Then

$$[h(T)f(T)](\boldsymbol{v}) = h(T)(f(T)(\boldsymbol{v})) = h(T)(\mathbf{0}) = \mathbf{0}.$$

Thus, $h(x)f(x) \in Ann(T)$.

14. Assume T has an eigenvector $\boldsymbol{v}$ with eigenvalue λ. Then $\mu_{T,\boldsymbol{v}}(x) = x - \lambda$. Since $\mu_T(x)(\boldsymbol{v}) = \mathbf{0}$ it follows by Remark (4.4) that $(x - \lambda)$ divides $\mu_T(x)$.

15. Assume $\boldsymbol{v}$ is an eigenvector with eigenvalue λ. Then $T(\boldsymbol{v}) = \lambda\boldsymbol{v}$. We then have

$$[T(\boldsymbol{v})]_{\mathcal{B}} = [\lambda\boldsymbol{v}]_{\mathcal{B}} = \lambda[\boldsymbol{v}]_{\mathcal{B}}.$$

On the other hand,

$$[T(\boldsymbol{v})]_{\mathcal{B}} = \mathcal{M}_T(\mathcal{B}, \mathcal{B})[\boldsymbol{v}]_{\mathcal{B}}.$$

Thus, we conclude that $[\boldsymbol{v}]_{\mathcal{B}}$ is an eigenvector of the matrix $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ with eigenvalue λ.

Conversely, assume $[\boldsymbol{v}]_{\mathcal{B}}$ is an eigenvector of the matrix $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ with eigenvalue λ. Then

$$\mathcal{M}_T(\mathcal{B}, \mathcal{B})[\boldsymbol{v}]_{\mathcal{B}} = \lambda[\boldsymbol{v}]_{\mathcal{B}} = [\lambda\boldsymbol{v}]_{\mathcal{B}}.$$

On the other hand,

$$\mathcal{M}_T(\mathcal{B}, \mathcal{B}) = [T(\boldsymbol{v})]_{\mathcal{B}}.$$

Thus, we have

$$[T(\boldsymbol{v})]_{\mathcal{B}} = [\lambda\boldsymbol{v}]_{\mathcal{B}}.$$

It then follows that

$$T(\boldsymbol{v}) = \lambda\boldsymbol{v}.$$

16. This is a consequence of the fact that $f(A) = \mathcal{M}_{f(T)}(\mathcal{B}, \mathcal{B})$.

17. Let $f(x) \in \mathbb{F}[x]$ and let T be an operator on a finite dimensional vector space. Suppose $f(S) = \mathbf{0}_{V\to V}$. Then $f(\mathcal{M}_S(\mathcal{B}, \mathcal{B})) = \mathbf{0}_{nn}$ by Exercise 16. Then $f(\mathcal{M}_S(\mathcal{B}, \mathcal{B})^{tr}) = \mathbf{0}_{nn}$. However, $\mathcal{M}_{S'}(\mathcal{B}, \mathcal{B}) = \mathcal{M}_S(\mathcal{B}, \mathcal{B})^{tr}$ and consequently, $f(\mathcal{M}_{S'}(\mathcal{B}, \mathcal{B})) = \mathbf{0}_{nn}$. Then by Exercise 16, $f(S') = \mathbf{0}_{V\to V}$. This implies that $\mu_{S'}(x)$ divides $\mu_S(x)$. However, the argument can be reversed so that $\mu_S(x)$ divides $\mu_{S'}(x)$. Since both are monic they are equal.

18. Since $T(\boldsymbol{v}) = \lambda\boldsymbol{v}, \boldsymbol{v} = \frac{1}{\lambda}T(\boldsymbol{v})$. Then

$$T^{-1}(\boldsymbol{v}) = T^{-1}(\frac{1}{\lambda}T(\boldsymbol{v})) = \frac{1}{\lambda}T^{-1}(T(\boldsymbol{v}))$$
$$= \frac{1}{\lambda}(T^{-1}T)(\boldsymbol{v}) = \frac{1}{\lambda}I_V(\boldsymbol{v}) = \frac{1}{\lambda}\boldsymbol{v}.$$

19. This follows since $\boldsymbol{v}$ is an eigenvector of T^k with eigenvalue λ^k. Also, if $\boldsymbol{v}$ is an eigenvector for an operator S with eigenvalue λ then $\boldsymbol{v}$ is an eigenvector for cS with eigenvalue $c\lambda$. Finally, if $\boldsymbol{v}$ is an eigenvector for operators S_1, S_2 with respective eigenvalues λ_1, λ_2 then $\boldsymbol{v}$ is an

eigenvector of $S_1 + S_2$ with eigenvalue $\lambda_1 + \lambda_2$. Now if $f(x) = a_m x^m + \cdots + a_1 x + a_0$ the operators $a_k T^k$ has $\boldsymbol{v}$ as an eigenvector with eigenvalue $a_k \lambda^k$. Then add.

20. Apply $S^{-1}TS$ to $S^{-1}(\boldsymbol{v})$:

$$(S^{-1}TS)(S^{-1}(\boldsymbol{v}) = (S^{-1}T)(S(S^{-1}(\boldsymbol{v})) =$$
$$(S^{-1}T)(I_V(\boldsymbol{v}) = (S^{-1}T)(\boldsymbol{v}) =$$
$$S^{-1}(T(\boldsymbol{v})) = S^{-1}(\lambda\boldsymbol{v}) = \lambda S(\boldsymbol{v}).$$

So we applied $S^{-1}TS$ to the vector $S^{-1}(\boldsymbol{v})$ and the result is $\lambda S^{-1}(\boldsymbol{v})$. Thus, $S^{-1}(\boldsymbol{v})$ is an eigenvector of $S^{-1}TS$ with eigenvalue λ.

21. Note for any polynomial $f(x)$ that $f(ST)S = Sf(TS)$ and $f(TS)T = Tf(ST)$.

Let $f(x) = \mu_{ST}(x)$ and $g(x) = \mu_T(x)$. Then $f(ST) = 0_{V\to V}$. Then $f(ST)S = 0_{V\to V}$. By the above $Sf(TS) = 0_{V\to V}$ whence $(TS)f(TS) = 0_{V\to V}$. It then follows that $g(x)$ divides $xf(x)$. In exactly the same way $f(x)$ divides $xg(x)$. In particular, $f(x)$ and $g(x)$ have the same non-zero roots.

4.2. Cyclic Operators

1a) $T(\boldsymbol{z}) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, T^2(\boldsymbol{z}) = \begin{pmatrix} 3 \\ -4 \\ -2 \end{pmatrix}, T^3(\boldsymbol{z}) = \begin{pmatrix} 3 \\ -3 \\ -2 \end{pmatrix}$.

The matrix $\begin{pmatrix} 1 & 1 & 3 \\ 0 & -1 & -4 \\ 0 & 0 & -2 \end{pmatrix}$ is invertible and therefore $(\boldsymbol{z}, T(\boldsymbol{z}), T^2(\boldsymbol{x}))$ is a basis for $\mathbb{R}^3$.

$\mu_{T,\boldsymbol{z}}(x) = x^3 - x^2 + x - 1$.

b) $\mu_{T,\boldsymbol{u}}(x) = x - 1$.

2. $T(\boldsymbol{z}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, T^2(\boldsymbol{z}) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, T^3(\boldsymbol{z}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

So $\langle T, \boldsymbol{z}\rangle$ contains the standard basis of $\mathbb{R}^4$ and therefore $\langle T, \boldsymbol{z}\rangle = \mathbb{R}^4$. Now $T^4(\boldsymbol{z}) = \begin{pmatrix} -4 \\ 0 \\ -5 \\ 0 \end{pmatrix}$ and $\mu_{T,\boldsymbol{z}}(x) = x^4 + 5x^2 + 4 = (x^2+1)(x^2+4)$.

3. Let $\boldsymbol{z} \in V$ be any non-zero vector. Then $\langle T, \boldsymbol{z}\rangle$ is a $T-$invariant subspace which is not $\{\boldsymbol{0}\}$. Therefore $\langle T, \boldsymbol{z}\rangle = V$ and T is a cyclic operator.

4. There are lots of such operators. Here is a simple one

$$T(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ 2x_2 \\ 3x_3 \\ 4x_4 \end{pmatrix}.$$

If $\boldsymbol{z} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ then $\langle T, \boldsymbol{z}\rangle = \mathbb{R}^4$.

5. Since T is cyclic, $\mu_T(x)$ has degree 3. Suppose $\mu_T(x)$ has one real root so that $\mu_T(x) = (x-\lambda)g(x)$ where $g(x)$ is a real irreducible quadratic. Let $\boldsymbol{x}$ be a vector such that $\mathbb{R}^3 = \langle T, \boldsymbol{x}\rangle$ Then the $T-$invariant subspaces are $\mathbb{R}^3, \{\boldsymbol{0}\}, \langle T, (T-\lambda I)\boldsymbol{x}\rangle$ and $\langle T, g(T)(\boldsymbol{x})\rangle$.

We can suppose $\mu_T(x)$ has three real roots. If there is only one distinct root, say λ so that $\mu_T(x) = (x-\lambda)^3$. In this case there are four $T-$invariant subspaces.

Assume $\mu_T(x) = (x-\alpha)^2(x-\beta), \alpha \neq \beta$, then there are six $T-$invariant subspaces.

Assume $\mu_T(x) = (x-\alpha)(x-\beta)(x-\gamma)$, distinct. In this case there are eight $T-$invariant subspaces

6. Let T have the following matrix with respect to the standard basis

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

Here is another:

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

7. Let T have the following matrix with respect to the standard basis

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

8. Because T is cyclic, $\mu_T(x)$ has degree 4.

The number of $T-$invariant subspaces can be computed from a factorization of $\mu_T(x)$. The possibilities are:

$p(x)^2$ where $p(x)$ is a real quadratic irreducible. There are three $T-$invariant subspaces in this case.

$p(x)(x-\alpha)^2$ where $p(x)$ is a real irreducible quadratic and $\alpha \in \mathbb{R}$. There are six $T-$invariant subspaces in this case.

$p(x)(x-\alpha)(x-\beta)$ where $p(x)$ is a real irreducible quadratic and $\alpha \neq \beta$ are real numbers. There are eight $T-$invariant subspaces in this case.

$(x-\alpha)^4, \alpha \in \mathbb{R}$. There are five $T-$invariant subspaces in this case.

$(x-\alpha)^3(x-\beta)$ where $\alpha \neq \beta$ are real numbers. There are eight $T-$invariant subspaces in this case.

$(x-\alpha)^2(x-\beta)^2$ where $\alpha \neq \beta$ are real numbers. There are nine $T-$invariant subspaces in this case.

$(x-\alpha)^2(x-\beta)(x-\gamma)$ where α, β, γ are distinct real numbers. where $\alpha \neq \beta$ are real numbers. There are twelve $T-$invariant subspaces in this case.

$(x-\alpha)(x-\beta)(x-\gamma)(x-\delta)$ where $\alpha, \beta, \gamma, \delta$ are distinct real numbers. where $\alpha \neq \beta$ are real numbers. There are sixteen $T-$invariant subspaces in this case.

9. Let T be the operator on $\mathbb{R}^4$ which has the following matrix with respect to the standard basis:

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

10. Let T be the operator on $\mathbb{R}^4$ which has the following matrix with respect to the standard basis:

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

11. Let T be the operator on $\mathbb{R}^4$ which has the following matrix with respect to the standard basis:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

12. Set $\boldsymbol{v}_0 = \boldsymbol{v}$ and $\boldsymbol{v}_{i+1} = T^i(\boldsymbol{v})$ for $1 \leq i < n$. By our hypothesis, $(\boldsymbol{v}_0, \ldots, \boldsymbol{v}_{n-1})$ is a basis for V. Assume $S(\boldsymbol{v}) = a_0\boldsymbol{v}_0 + \ldots a_{n-1}\boldsymbol{v}_{n-1}$. Set $f(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} \in \mathbb{F}_{n-1}[x]$. We claim that $S = f(T)$. Since S and $f(T)$ are linear operators it suffices to show that $S(\boldsymbol{v}_i) = f(T)(\boldsymbol{v}_i)$ for $i = 0, \ldots, n-1$.

We have constructed $f(x)$ such that $f(T)(\boldsymbol{v}) = f(T)(\boldsymbol{v}_0) = S(\boldsymbol{v})$. Consider $f(T)(\boldsymbol{v}_i)$ for $1 \leq i \leq n-1$. $f(T)(\boldsymbol{v}_i) = f(T)(T^i(\boldsymbol{v})) = T^i f(T)(\boldsymbol{v}) = T^i S(\boldsymbol{v}) = ST^i(\boldsymbol{v}) = S(\boldsymbol{v}_i)$.

4.3. Maximal Vectors

1. a) $\mu_{T,\boldsymbol{e}_1}(x) = x^2 + 2x + 2$,
$\mu_{T,\boldsymbol{e}_2}(x) = x^3 - 2x - 4$,
$\mu_{T,\boldsymbol{e}_3}(x) = x^3 - 2x - 4$.
b) $\mu_T(x) = x^3 - 2x - 4$.
c) $\boldsymbol{e}_2, \boldsymbol{e}_3$ are maximal vectors.

2. $\mu_T(x) = (x-2)^2(x-1)$. $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are maximal vectors for T.

3. $\mu_T(x) = x^4 - x^3 - x^2 - x - 2 = (x-2)(x^3+x^2+x+1)$. $\boldsymbol{e}_1$ is a maximal vector.

4. Clearly $(\boldsymbol{v}_1)$ is linearly independent since an eigenvector is non-zero. Assume we have shown that $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_j)$ is linearly independent with $j < k$. We show that $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{j+1})$ is linearly independent. Suppose that

$$c_1\boldsymbol{v}_1 + \cdots + c_{j+1}\boldsymbol{v}_{j+1} = \mathbf{0} \quad (4.1)$$

Apply T to get

$$c_1T(\boldsymbol{v}_1) + \cdots + c_{j+1}T(\boldsymbol{v}_{j+1}) = \mathbf{0}$$

Using the fact that $T(\boldsymbol{v}_i) = \alpha_i\boldsymbol{v}_i$ we get

$$c_1\alpha_1\boldsymbol{v}_1 + \cdots + c_j\alpha_j\boldsymbol{v}_j + c_{j+1|}\alpha_{j+1}\boldsymbol{v}_{j+1} \quad (4.2)$$

Now multiply (4.1) by α_{j+1} and subtract it from (4.2) to get

$$c_1(\alpha_1 - \alpha_{j+1})\boldsymbol{v}_1 + \cdots + c_j(\alpha_j - \alpha_{j+1}) = \mathbf{0} \quad (4.3)$$

Thus, we obtain a dependence relation on $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_j)$. However, $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_j)$ is linearly independent and therefore

$$c_1(\alpha_1 - \alpha_{j+1}) = \cdots = c_j(\alpha_j - \alpha_{j+1}) = 0.$$

Since the α_i are distinct, for all $i < j+1, \alpha_i - \alpha_{j+1} \neq 0$. Therefore

$$c_1 = c_2 = \cdots = c_j = 0.$$

In then follows that the dependence relation in (4.1) is

$$c_{j+1}\boldsymbol{v}_{j+1} = \mathbf{0}$$

Since $\boldsymbol{v}_{j+1}$ is non-zero, $c_{j+1} = 0$.

5. Since $x^2+1, x+1, x-2$ are relatively prime it follows that $(x^2+1)(x+1)(x-2)$ divides $\mu_T(x)$. Since T is an operator on $\mathbb{R}^4$ the degree of $\mu_T(x)$ is at most four. We can then conclude that $\mu_T(x) = (x^2+1)(x+1)(x-2)$. Set $\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3$. Since $\mu_{T,\boldsymbol{v}_1}(x), \mu_{T,\boldsymbol{v}_2}(x)$ and $\mu_{T,\boldsymbol{v}_3}(x)$ are relatively prime, $\mu_{T,\boldsymbol{v}}(x)$ is the product $\mu_{T,\boldsymbol{v}_1}(x)\mu_{T,\boldsymbol{v}_2}(x)\mu_{T,\boldsymbol{v}_3}(x) = (x^2+1)(x+1)(x-2) = \mu_T(x)$. Thus, $\boldsymbol{v}$ is a maximal vector.

6. $\langle T, \boldsymbol{v}_1\rangle = Span(\boldsymbol{v}_1, \boldsymbol{v}_2)$ and every non-zero vector $\boldsymbol{v}$ in $\langle T, \boldsymbol{v}_1\rangle$ satisfies $\mu_{T,\boldsymbol{v}}(x) = x^2+1$ since x^2+1 is irreducible. $\boldsymbol{v}_3$ is an eigenvector with eigenvalue -1 and $\boldsymbol{v}_4$ is an eigenvector with eigenvalue 1. If $\boldsymbol{w} = \boldsymbol{w}_1 + \boldsymbol{w}_2 + \boldsymbol{w}_3$ with $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$ nonzero and $\boldsymbol{w}_1 \in Span(\boldsymbol{v}_1, \boldsymbol{v}_2), \boldsymbol{w}_2 \in Span(\boldsymbol{v}_3)$ and $\boldsymbol{w}_3 \in Span(\boldsymbol{v}_4)$ then $\mu_{T,\boldsymbol{w}}(x)$ is divisible by $(x^2+1)(x+1)(x-1)$ and is therefore a maximal vector. On the other hand, if $\boldsymbol{w}_1, \boldsymbol{w}_2$ are non-zero but $\boldsymbol{w}_3$ is zero then $\mu_{T,\boldsymbol{w}}(x) = (x^2+1)(x+1)$. If $\boldsymbol{w}_1, \boldsymbol{w}_3$ are non-zero but $\boldsymbol{w}_2 = \mathbf{0}$ then $\mu_{T,\boldsymbol{w}}(x) = (x^2+1)(x-1)$. If $\boldsymbol{w}_2, \boldsymbol{w}_3 \neq \mathbf{0}, \boldsymbol{w}_1 = \mathbf{0}$ then $\mu_{T,\boldsymbol{w}}(x) = (x+1)(x-1)$.

7. If $\boldsymbol{v} \neq \mathbf{0}$ then $\mu_{T,\boldsymbol{v}}(x)$ is a non-constant polynomial and divides $\mu_T(x)$ which is irreducible. Then $\mu_{T,\boldsymbol{v}}(x) = \mu_T(x)$ and $\boldsymbol{v}$ is a maximal vector.

8. Since $\mu_T(x)$ has degree 5 and $dim(\mathbb{F}_5^5) = 5$ the operator T is cyclic. Note that $x^5 - x = x(x-1)(x-2)(x-3)(x-4)$. Set $f_i(x) = \frac{x^5-x}{x-i}$ where $i = 0, 1, 2, 3, 4$ and let $\boldsymbol{v}$ be a maximal vector. Set $\boldsymbol{v}_i = f_i(T)\boldsymbol{v}$. Then $\boldsymbol{v}_i$ is an eigenvector with eigenvalue $i, i = 0, 1, 2, 3, 4$. Since these eigenvalues are distinct, $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_5)$ is independent by Exercise 4. By the half is good enough theorem $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_5)$ is a basis. A vector $c_1\boldsymbol{v}_1 + \cdots + c_5\boldsymbol{v}_5$

is a maximal vector if and only if $c_i \neq 0$ for all i. Thus, there are 4^5 maximal vectors.

4.4. Indecomposable Linear Operators

1. This operator is decomposable since it is not cyclic. It has minimal polynomial $x^2 + 2x + 1 = (x + 1)^2$.

2. This operator is indecomposable. $\mu_{T,\boldsymbol{e}_3}(x) = x^3 + 3x^2+3x+1 = (x+1)^2$. So, the operator is cyclic and the minimum polynomial is a power of the irreducible polynomial $x + 1$.

3. This operator is indecomposable. $\mu_{T,\boldsymbol{e}_1}(x) = x^3 - 6x^2 + 12x - 8 = (x - 2)^3$.

4. Set $f(x) = \mu_S(x)$. If T in in $\mathcal{P}(S)$ then there is a polynomial $g(x)$ with degree of $g(x)$ less then degree of $f(x)$ such that $T = g(S)$. Since $g(x)$ is not the zero polynomial, $f(x)$ is irreducible and $deg(g) < deg(f)$ it must be the case that $f(x), g(x)$ are relatively prime. Consequently, there are polynomials $a(x), b(x)$ such that $a(x)f(x) + b(x)g(x) = 1$. Substituting S for x we get $a(S)f(S) + b(S)g(S) = I_V$. Since $f(S) = 0_{V\to V}$ and $g(S) = T$ we have $b(S)T = I_V$. Thus, $b(S) \in \mathcal{P}(S)$ is an inverse to T.

5. Set $f_i(x) = \mu_{T,\boldsymbol{v}_i}(x)$. Then $f_i(x)$ divides $p(x)^m$ and so there is a natural number $m_i \leq m$ such that $f_i(x) = p(x)^{m_i}$. Let j be chosen such that $m_j \geq m_i$ for $i = 1, 2, \dots, n$. Then the gcd of $f_i(x)$ is $f_j(x) = p(x)^{m_j}$. However, the $gcd(f_1, \dots, f_n)$ is $\mu_T(x)$. Thus, $f_j(x) = \mu_T(x)$ and $\boldsymbol{v}_j$ is a maximal vector.

6. Since T is indecomposable, $\mu_T(x) = p(x)^m$ for some irreducible polynomial $p(x)$. By Exercise 5 there is a j such that $\boldsymbol{v}_j$ is maximal. Since T is indecomposable, T is cyclic. Therefore $V = \langle T, \boldsymbol{v}_j\rangle$.

7. Let $f(x) = \mu_T(x)$. Since T is indecomposable, $deg(f) = dim(V) = 2n + 1$. Since the degree of $f(x)$ is odd it has a real root a. Thus, $x-a$ divides $f(x)$. However, since $f(x)$ has single distinct irreducible factor it follows that $f(x) = (x - 1)^{2n+1}$.

8. Let $f(x) = \mu_T(x)$. Then either $f(x) = (x - a)^{2n}$ for some n or $f(x) = p(x)^n$ for some real irreducible quadratic polynomial. By the theory of cyclic operators in the first case the number of T-invariant subspaces is $2n + 1$ and in the latter case the number is $n + 1$.

9. Let $f(x) = \mu_T(x)$. Then either $f(x) = (x - a)^4$ or $g(x)^2$ where $g(x)$ is an irreducible polynomial of degree 2. In either case, let $\boldsymbol{v}$ be a maximal vector. In the first case, any vector $\boldsymbol{w}$ such that $\langle T, \boldsymbol{w}\rangle \neq V$ lies in $\langle T, (T - aI)(\boldsymbol{v})\rangle$ which has dimension three and contains p^3 vectors. Any other vector is maximal and hence in this case there are $p^4 - p^3$ maximal vectors.

Suppose $f(x) = g(x)^2$. Now if $V \neq \langle T, \boldsymbol{w}\rangle$ then $\boldsymbol{w}$ belongs to $\langle T, g(T)(\boldsymbol{v})\rangle$ which has dimension 2 and p^2 vectors. In this case there are $p^4 - p^2$ maximal vectors.

10. Suppose T is indecomposable. Let $\mu_T(x) = f(x) = p(x)^m$ where $p(x)$ is irreducible. Let $\boldsymbol{v}$ be a maximal vector. The only proper $T-$invariant subspaces of V are $\langle T, p(T)^k(\boldsymbol{v})\rangle$ for $1 \leq k \leq n$. Moreover, $\langle T, p(T)^{j+1}(\boldsymbol{v})\rangle$ is a subspace of $\langle T, p(T)^j(\boldsymbol{v})\rangle$ and hence $\langle T, p(T)(\boldsymbol{v})\rangle$ is the unique maximal proper $T-$invariant subspace.

On the other hand, suppose T is not indecomposable. Then there are $T-$invariant subspaces U and W such that $V = U \oplus W$. Let U' be a proper maximal $T-$invariant subspace of U (possibly $\{\mathbf{0}\}$) and similarly choose W' in W. Then $U \oplus W'$ and $U' \oplus W$ are two distinct proper maximal $T-$invariant subspaces.

4.5. Invariant Factors and Elementary Divisors of a Linear Operator

1. Set $d_i = dim(U_i)$. Then $(d_1, \dots, d_5) = (12, 22, 28, 34, 38)$.

2. The invariant factors, $d_i(x)$, ordered so $d_1|d_2|d_3|d_4$ are

$$d_1(x) = (x^2 - x + 1)^2(x^2 + 1)$$

$$d_2(x) = (x^2 - x + 1)^2(x^2 + 1)^2(x + 2)$$

$$d_3(x) = (x^2 - x + 1)^3(x^2 + 1)^2(x + 2)^2$$

$$d_4(x) = (x^2 - x + 1)^4(x^2 + 1)^3(x + 2)^2$$

$dim(V) = 44$.

3. The elementary divisors and the invariant factors are $x^2 + 1$ and $x^2 + 1$

4. There is a single elementary divisor and invariant factor which is $(x^2 + 1)^2$

5. The elementary divisors are $x^2 + 1, x + 1$ and $x - 1$. There is a single invariant factor, $x^4 - 1$.

6. The elementary divisors are $x, x, x - 1, x - 1$ and the invariant factors are $x^2 - x, x^2 - x$.

7. Assume the minimum polynomial of T is $p_1(x) \dots p_t(x)$ where $p_i(x)$ is irreducible and distinct. Let $V_i = \{\boldsymbol{v} \in V | p_i(T)\boldsymbol{v} = \boldsymbol{0}\}$. Then $V = V_1 \oplus \dots \oplus V_t$. If each V_i is completely reducible then so is V. Consequently, we may reduce to the case that the minimum polynomial is irreducible. In this case, there are vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_s$ such that $\mu_{T,\boldsymbol{v}_i}(x) = p(x)$ is irreducible and

$$V = \langle T, \boldsymbol{v}_1 \rangle \oplus \dots \oplus \langle T, \boldsymbol{v}_s \rangle.$$

Each space $\langle T, \boldsymbol{v}_j \rangle$ is $T-$irreducible. It follows from this that V is completely reducible.

8. Assume V is completely reducible and $\mu_T(x) = (x - \alpha_1) \dots (x - \alpha_t)$ where α_i are distinct. Set $V_i = \{\boldsymbol{v} \in V : (T - \alpha_i I_V)(\boldsymbol{v}) = \boldsymbol{0}\}$ Then

$$V = V_1 \oplus \dots \oplus V_t.$$

Let $\mathcal{B}_i$ be a basis for V_i and set $\mathcal{B} = \mathcal{B}_1 \sharp \dots \sharp \mathcal{B}_t$. Then $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ is a diagonal matrix.

Conversely, assume that T is diagonalizable. Then there exists a basis $\mathcal{B}$ consisting of eigenvectors. Let $\alpha_1, \dots, \alpha_t$ be the distinct eigenvectors and set $V_i = \{\boldsymbol{v} \in V | T(\boldsymbol{v}) = \alpha_i \boldsymbol{v}\}$. Then $V_1 + \dots + V_t$ is a direct sum. Since $\mathcal{B} \subset V_1 + \dots + V_t$

$$V = V_1 + \dots + V_t.$$

Consequently, $V = V_1 \oplus \dots \oplus V_t$. Thus, V is completely reducible. Now $\mu_T(x) = (x - \alpha_1) \dots (x - \alpha_t)$.

9. Let $dim(V) = n$ and set $V_i = \{\boldsymbol{v} \in V | p_i(x)^n(\boldsymbol{v}) = \boldsymbol{0}\}$ so that V_i is the p_i-Sylow subspace. If there are infinitely many $T-$invariant subspaces in some V_i then there are clearly infinitely many $T-$invariant subspaces.

Suppose, on the other hand that each V_i has finitely many $T-$invariant subspaces. Suppose U is a $T-$invariant subspace. Set $U = U \cap V_i$. Then

$$U = U_1 \oplus \dots \oplus U_t.$$

It follows from this that there are only finitely many $T-$invariant subspaces.

10. Let $dim(V) = n$ and $p_1(x), \dots, p_t(x)$ be the distinct irreducible factors of $\mu_T(x)$ Set $V_i = \{\boldsymbol{v} \in V | p_i(T)(\boldsymbol{v}) = \boldsymbol{0}\}$ so that V_i is the $p_i(x)-$Sylow subspace of V and

$$V = V_1 \oplus \dots \oplus V_t.$$

Then T is cyclic if and only if T restricted to each V_i is cyclic. Also, V has finitely many $T-$invariant subspaces if and only there are finitely many $T-$invariant subspaces in V_i for each i by Exercise 9. Thus, we may assume that $\mu_T(x) = p(x)^m$ for some irreducible polynomial $p(x)$. If T is cyclic in this case then the number of T-invariant subspaces is $m + 1$. On the other hand suppose T is not cyclic. Then there are at least two elementary divisors (invariant factors). Thus, there are vectors $\boldsymbol{v}_1, \boldsymbol{v}_2$ with $\mu_{T,\boldsymbol{v}_1}(x) = p(x)^{m_1}$ and $\mu_{T,\boldsymbol{v}_2}(x) = p(x)^{m_2}$ such that

$$\langle T, \boldsymbol{v}_1\rangle \cap \langle T, \boldsymbol{v}_2\rangle = \{\boldsymbol{0}\}.$$

Set $\boldsymbol{w}_1 = p(T)^{m_1-1}(\boldsymbol{v}_1), \boldsymbol{w}_2 = p(T)^{m_2-1}(\boldsymbol{v}_2)$. Then $\mu_{T,\boldsymbol{w}_1}(x) = \mu_{T,\boldsymbol{w}_2}(x) = p(x)$ and

$$\langle T, \boldsymbol{w}_1\rangle \cap \langle T, \boldsymbol{w}_2\rangle = \{\boldsymbol{0}\}.$$

Now each of the spaces $\langle T, \boldsymbol{w}_1 + c\boldsymbol{w}_2\rangle, c \in \mathbb{F}$ is distinct and $T-$invariant. Since $\mathbb{F}$ is infinite there are infinitely many $T-$invariant subspaces.

11. Let $V(p) = \{\boldsymbol{v} \in V | p(T)^m(\boldsymbol{v}) = \boldsymbol{0}\}$ and $V(q) = \{\boldsymbol{v} \in V | q(T)^n(\boldsymbol{v}) = \boldsymbol{0}\}$. Then $V = V(p) \oplus V(q)$. Since $a(x)p(x)^m + b(x)q(x)^n = 1$ it follows that $a(T)p(T)^m + b(T)q(T)^n = I_V$. However, $a(T)p(T)^n$ restricted to $V(p)$ is the zero map and consequently, $b(T)q(T)^n$ restricted to $V(p)$ is the identity. Then $b(T)q(T)^n p(T)$ restricted to $V(p)$ is identical to $p(T)$ restricted to $V(p)$. In a similar manner, $b(T)q(T)^n$ restricted to $V(q)$ is the zero map and $a(T)p(T)^m$ restricted to $V(q)$ is the identity map. Consequently $a(T)p(T)^m q(T)$ restricted to $V(q)$ is the same as $q(T)$ restricted to $V(q)$. Thus, $f(T)$ restricted to $V(p)$ is equal to the restriction of $p(T)$ to $V(p)$ and $f(T)$ restricted to $V(q)$ is equal to the restriction of $q(T)$ to $V(q)$.

Suppose now that $\boldsymbol{x} = \boldsymbol{v}_p + \boldsymbol{v}_q$ with $\boldsymbol{v}_p \in V(p)$ and $\boldsymbol{v}_q \in V(q)$. Then $f(T)(\boldsymbol{x}) = f(T)(\boldsymbol{v}_p + \boldsymbol{v}_q) = f(T)(\boldsymbol{v}_p) + f(T)(\boldsymbol{v}_q) = p(T)(\boldsymbol{v}_p) + q(T)(\boldsymbol{v}_q)$. If $l = max(m, n)$ then $f(T)^l(\boldsymbol{x}) = p(T)^l(\boldsymbol{v}_p) + q(T)^l(\boldsymbol{v}_q) = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$. Thus, $f(T)^l = 0_{V\to V}$ and $f(T)$ is nilpotent.

12. Let $p(x)^{e_1}, \ldots, p(x)^{e_t}$ be the elementary divisors of T where $e_1 \le e_2 \cdots \le e_t \le m$. Then there are vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_t$ such that $\mu_{T,\boldsymbol{v}_i}(x) = p(x)^{e_i}$ and

$$V = \langle T, \boldsymbol{v}_1\rangle \oplus \cdots \oplus \langle T, \boldsymbol{v}_t\rangle.$$

Now $U_1 =$

$$\langle T, p(T)^{e_1-1}(\boldsymbol{v}_1)\rangle \oplus \cdots \oplus \langle T, p(T)^{e_t-1}(\boldsymbol{v}_t)\rangle$$

and has dimension td which proves a).

Now assume $e_{k-1} < j-1 \le e_k$ and $e_{l-1} < j \le e_l$. Clearly $k \le l$.

Now

$$U_{j-1} = \langle T, \boldsymbol{v}_1\rangle \oplus \cdots \oplus \langle T, \boldsymbol{v}_{k-1}\rangle \oplus \langle T, p(T)^{e_k-j+1}(\boldsymbol{v}_j)\rangle \oplus \cdots \oplus \langle T, p(T)^{e_t-j+1}(\boldsymbol{v}_t)\rangle \quad (4.4)$$

$$U_j = \langle T, \boldsymbol{v}_1\rangle \oplus \cdots \oplus \langle T, \boldsymbol{v}_{l-1}\rangle \oplus \langle T, p(T)^{e_l-j}(\boldsymbol{v}_l)\rangle \oplus \cdots \oplus \langle T, p(T)^{e_t-j}(\boldsymbol{v}_t)\rangle \quad (4.5)$$

It follows from (4.4) that $m_{j-1} = dim(U_{j-1}) = e_1 d + \ldots e_{k-1}d + [t-k+1](j-1)d$. It follows from (4.5) that $m_j = dim(U_j) = e_1 d + \ldots e_{l-1}d + [t-l+1]jd$.

Suppose $k = l$ then $m_j - m_{j-1} = dim(U_j) - dim(U_{j-1}) = [t-l+1]d$ and $\frac{m_j - m_{j-1}}{d} = t - l + 1$ which is the number of e_i which are greater than or equal to j.

Suppose $k < l$. Then $e_k = \cdots = e_{l-1} = j-1$. Making use of this we get that

$$m_j = dim(U_j) = e_1 d + \ldots e_{k-1}d + (j-1)d[l-1-k] + [t-l+1]jd.$$

Then

$$m_j - m_{j-1} = (j-1)d[l-1-k] + [t-l+1]jd - [t-k+1](j-1)d = [t-l+1]d$$

and we again get $\frac{m_j - m_{j-1}}{d} = [t-l+1]$ which is equal to the number of e_i which are greater than or equal to j.

13. Since each invariant factor is the product of distinct elementary divisors, it follows that the characteristic polynomial of T is equal to the product of the elementary divisors of T. It therefore suffices to prove this in the case that $\mu_T(x) = p(x)^m$ for some irreducible polynomial $p(x)$ of degree d. In this case $dm = deg(p(x)^m) = deg(\chi_T(x)) = dim(V)$ so that $m = \frac{dim(V)}{d}$.

4.6. Canonical Forms

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

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

3. $\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & -2 \end{pmatrix}$

4. $\begin{pmatrix} 3 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 1 & -2 \end{pmatrix}$

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

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

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

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

8. $0_{44}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

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

9. The characteristic polynomial of a nilpotent operator on an $n-$dimensional vector space is x^n. It is completely reducible if and only if the minimum polynomial has distinct roots and consequently we must have $\mu_T(x) = x$. This implies that $T = 0_{V \to V}$.

10. Every vector of V satisfies $p(T)^e(\boldsymbol{v}) = \mathbf{0}$ and therefore $p(T)^e = 0_{V \to V}$.

11. Let A be a square matrix. We have seen that if $f(x) \in \mathbb{F}[x]$ then $f(A^{tr}) = f(A)^{tr}$. Since also $\mathcal{M}_{f(S)}(\mathcal{B}, \mathcal{B}) = f(\mathcal{M}_S(\mathcal{B}, \mathcal{B}))$ it follows that S and S' have the same minimum polynomial. Let $p_1(x), \dots, p_s(x)$ be the distinct irreducible polynomials dividing $\mu_S(x) = \mu_{S'}(x)$ and let $\mu_S(x) = p_1(x)^{m_1} \dots p_s(x)^{m_s}$. Let $V_i = \{\boldsymbol{v} \in V | p_i(S)^{m_i}(\boldsymbol{v}) = \mathbf{0}\}$ and $V_i' = \{\boldsymbol{v} \in V | p_i(S')^{m_i}(\boldsymbol{v}) = \mathbf{0}\}$. Note for any operator $T, nullity(T)$ and $nullity(\mathcal{M}_T(\mathcal{B}, \mathcal{B}))$ and equal and for any square matrix $nullity(A)$ and $nullity(A^{tr})$ are equal. It follows that $dim(V_i) = dim(V_i')$. By Exercise 12 of Section (4.5) we can determine the elementary divisors of S divisible by $p_i(x)$ by determining the dimensions of the $Ker(p_i(S)^k)$ for all

$k \leq m_i$ and likewise for the elementary divisors of S'. However, these numbers will all be the same and consequently, S and S' have the same elementary divisors and invariant factors.

12. $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}$

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

4.7. Linear Operators on Real and Complex Vector Spaces

1. i) implies ii). Assume T is completely reducible. Note that the assumption that T is completely reducible carries over to the restriction of T to any $T-$invariant subspace.

Set $f(x) = \mu_T(x)$ and $n = dim(V)$. Since $f(x) \in \mathbb{C}[x]$ $f(x)$ splits into linear factors. Let $\alpha_1, \ldots, \alpha_t$ be the distinct roots of $f(x)$. Set $V_i = \{\boldsymbol{v} \in V | (T - \alpha_i I_V)^n(\boldsymbol{v}) = \boldsymbol{0}\}$. Then $V = V_1 \oplus \cdots \oplus V_t$. We show that, in fact, $T(\boldsymbol{v}) = \alpha_i \boldsymbol{v}$ for $\boldsymbol{v} \in V_i$. Suppose to the contrary that there exists $\boldsymbol{v} \in V_i$ such that $\mu_{T,\boldsymbol{v}}(x) = (x - \alpha_i)^k$ with $k > 1$. Then $\langle T, \boldsymbol{v}\rangle$ is indecomposable. However, by the above remark $\langle T, \boldsymbol{v}\rangle$ is completely reducible, so we have a contradiction. Thus, every non-zero vector in V_i is an eigenvector with eigenvalue α_i. In now follows that $\mu_T(x) = (x - \alpha_1) \ldots (x - \alpha_t)$.

ii) implies iii). Assume $\mu_T(x) = (x - \alpha_1) \ldots (x - \alpha_t)$ with α_i distinct. Set $V_i = \{\boldsymbol{v} \in V | T(\boldsymbol{v}) = \alpha_i \boldsymbol{v}\}$. Then $V = V_1 \oplus \cdots \oplus V_t$. For each i, T restricted to V_i is $\alpha_i I_{V_i}$. Choose a basis $\mathcal{B}_i$ for V_i. Each vector in $\mathcal{B}_i$ is an eigenvector with eigenvalue α_i. Now set $\mathcal{B} = \mathcal{B}_1 \sharp \ldots \sharp \mathcal{B}_t$. This is a basis of eigenvectors.

iii) implies iv). Let $\mathcal{B}$ be a basis of eigenvectors. Then $\mu_T(\mathcal{B}, \mathcal{B})$ is a diagonal matrix which must be the Jordan canonical form of T.

iv) implies i). Let $\mathcal{B}$ be a basis for V such that $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ is the Jordan canonical form of T. Since the Jordan canonical form is diagonal this implies that the basis $\mathcal{B}$ consists of eigenvectors. Let $\alpha_1, \ldots, \alpha_t$ be the distinct eigenvalues and $V_i = \{\boldsymbol{v} \in V | T(\boldsymbol{v}) = \alpha_i \boldsymbol{v}\}$. Then $V = V_1 \oplus \cdots \oplus V_t$. Suppose U is a $T-$invariant subspace of V. Set $U_i = U \cap U_i$. Then

$$U = U_1 \oplus \cdots \oplus U_t.$$

Since T acts a scalar when restricted to V_i every subspace of V_i is $T-$invariant. Let W_i be any complement to U_i in V_i and set $W = W_1 \oplus \cdots \oplus W_t$. Then W is a $T-$invariant complement to U in V.

2. i) implies ii). Let α be an eigenvalue of T and let $(x - \alpha_i)^m$ be the exact power of $(x - \alpha)$ which divides $\mu_T(x)$. Let $\mu_T(x) = (x - \alpha)^m f(x)$. Then $f(x)$ and $x - \alpha$ are relatively prime. Set $V_\alpha = \{\boldsymbol{v} \in V | (T - \alpha I_V)(\boldsymbol{v}) = \boldsymbol{0}\}$ and $V_f = \{\boldsymbol{v} \in V | f(T)(\boldsymbol{v}) = \boldsymbol{0}\}$. Then $V = V_\alpha \oplus V_f$. It follows from our hypothesis that $V_f = \{\boldsymbol{0}\}$ and $f(x) = 1$.

If T has more than one elementary divisor then V can be decomposed into a direct sum, therefore by our hypothesis, there is only one elementary divisor and one Jordan block.

ii) implies i). If there is a single Jordan block of size n, say $J_n(\alpha)$, then T is cyclic and the minimum polynomial of T is $(x - \alpha)^n$ and T is indecomposable.

3. The minimal polynomial is $\mu_T(x) = (x - 1)(x^3 - 1)$. The characteristic polynomial is $(x - 1)(x^3 - 1)^2$.

The invariant factors are $(x - 1)(x^3 - 1)$ and $x^3 - 1$.

The elementary divisors are $x - 1, (x - 1)^2, x^2 + x + 1, x^2 + x + 1$.

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

5. Set $\omega = -\frac{1}{2}+i\frac{\sqrt{3}}{2}$ and $\omega^2 = \frac{1}{\omega} = -\frac{1}{2}-i\frac{\sqrt{3}}{2}$. Then the Jordan canonical form of T over the complex numbers is

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \omega & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \omega^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \omega^2 \end{pmatrix}$$

6. We define S, T on $\mathbb{C}^4$ by matrices with respect to the standard basis:

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Then $\chi_S(x) = \chi_T(x) = (x-1)^4$ And $\mu_S(x) = \mu_T(x) = (x-1)^2$. The elementary divisors (invariant factors) of S are $x-1, x-1, (x-1)^2$. The elementary divisors of T are $(x-1)^2, (x-1)^2$.

7. There are eight possibilities. They are:

$$J_2(0) \oplus J_3(-2i) \oplus J_1(0) \oplus J_1(0) \oplus J_1(0)$$

$$J_2(0) \oplus J_3(-2i) \oplus J_1(0) \oplus J_2(0))$$

$$J_2(0) \oplus J_3(-2i) \oplus J_1(0) \oplus J_1(0) \oplus J_1(-2i)$$

$$J_2(0) \oplus J_3(-2i) \oplus J_2(0) \oplus J_1(-2i)$$

$$J_2(0) \oplus J_3(-2i) \oplus J_1(0) \oplus J_2(-2i)$$

$$J_2(0) \oplus J_3(-2i) \oplus J_1(-2i) \oplus J_1(-2i) \oplus J_1(-2i)$$

$$J_2(0) \oplus J_3(-2i) \oplus J_1(-2i) \oplus J_2(-2i)$$

$$J_2(0) \oplus J_3(-2i) \oplus J_3(-2i)$$

8. The proof is by induction on $n = dim(V)$. Let $\alpha_1, \ldots, \alpha_s$ be the distinct eigenvalues of S and $\beta_1, \ldots, \beta_t$ the distinct eigenvalues of T. For α an eigenvalue of S set $V_{S,\alpha} = \{\boldsymbol{v} \in V | (S-\alpha I_V)^n(\boldsymbol{v}) = \boldsymbol{0}\}$. Since S and T commute, $V_{S,\alpha}$ is $T-$invariant. Likewise, define $V_{T,\beta}$ for $\beta = \beta_j, 1 \le j \le t$. Then $V_{T,\beta}$ is $S-$invariant.

As usual we have

$$V = V_{S,\alpha_1} \oplus \cdots \oplus V_{S,\alpha_s} = V_{T,\beta_1} \oplus \cdots \oplus V_{T,\beta_t}.$$

Suppose $s > 1$. Then we can apply induction and conclude that for each $i, 1 \le i \le s$ there is a basis $\mathcal{B}_i$ of V_{S,α_i} such that the matrix of S and T restricted to V_{S,α_i} with respect to $\mathcal{B}_i$ is in Jordan canonical form. Thus, we can assume that $s = 1$. In a similar way we can reduce to the case that $t = 1$. Let $\alpha = \alpha_1, \beta = \beta_1$.

Set $E_{S,\alpha} = \{\boldsymbol{v} \in V | S(\boldsymbol{v}) = \alpha\boldsymbol{v}\}$. Then $E_{S,\alpha}$ is $T-$invariant. It is then the case there must be a vector $\boldsymbol{v} \in E_{S,\alpha}$ which is also an eigenvector for T. Then $Span(\boldsymbol{v})$ is S and $T-$invariant. Let $\widehat{S}$ denote the transformation induced on $V/Span(\boldsymbol{v})$ by S and similarly define $\widehat{T}$. By the induction hypothesis there exists a basis $\widehat{\mathcal{B}} = (\widehat{\boldsymbol{v}}_1, \ldots, \widehat{\boldsymbol{v}}_{n-1})$ such that the matrix of $\widehat{S}$ and $\widehat{T}$ with respect to $\widehat{\mathcal{B}}$ is in Jordan canonical form. For $1 \le i \le n-1$ let $\boldsymbol{v}_{i+1}$ be a vector in V such that $Span(\boldsymbol{v}) + \boldsymbol{v}_{i+1} = \widehat{\boldsymbol{v}}_i$ and set $\boldsymbol{v}_1 = \boldsymbol{v}$. Then $\mathcal{B} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is a basis for

V and $\mathcal{M}_S(\mathcal{B},\mathcal{B})$ and $\mathcal{M}_T(\mathcal{B},\mathcal{B})$ are in Jordan canonical form.

9. $\begin{pmatrix} -2 & 0 & 0 & 0 \\ 1 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}$

10. Let $\mathcal{B}$ be a basis such that $M = \mathcal{M}_T(\mathcal{B},\mathcal{B})$ is in Jordan canonical form. Let A be the diagonal part of M and $B = M - A$. Then A is a diagonal matrix and B is a strictly lower triangular matrix and hence is nilpotent. Let D be the operator such that $\mathcal{M}_D(\mathcal{B},\mathcal{B}) = A$ and N the operator such that $\mathcal{M}_N(\mathcal{B},\mathcal{B}) = B$ (so that $N = T - D$). Then D is a diagonalizable operator and N is a nilpotent operator. Note for a Jordan block the diagonal part is a scalar matrix (scalar times the identity) which clearly commutes with the nilpotent part. From this it follows that $AB = BA$ and consequently $DN = ND$.

Suppose there exists polynomials $d(x), n(x)$ such that $d(T) = D, n(T) = N$. We use this to prove uniqueness.

Suppose also that D' is diagonalizable, N' is nilpotent, $D'N' = N'D'$ and $T = D' + N'$. Since $T = D' + N'$ we have $D'T = D'(D' + N') = (D')^2 + D'N' = (D')^2 + N'D' = (D' + N')D' = TD'$. Similarly, T and N' commute.

Since D and N are polynomials in T it follows that D and D', N and N' commute. From $D + N = T = D' + N'$ we conclude that

$$D - D' = N' - N \tag{4.6}$$

The operator on the left of (4.6) is the difference of two commuting diagonalizable operators and so is diagonalizable. The operator on the left hand side of (4.6) is the difference of two commuting nilpotent operators and therefore is nilpotent. However, the only nilpotent diagonalizable operator is the zero operator. Thus $D - D' = N - N' = 0_{V\to V}$ whence $D = D'$ and $N = N'$.

It therefore suffices to prove that there exist polynomials $d(x)$ and $n(x)$ such that $d(x) + n(x) = 1$ and $d(T)$ is diagonalizable, $n(T)$ is nilpotent.

Set $f(x) = \mu_T(x)$ and assume that $f(x) = (x - \alpha_1)^{m_1} \dots (x - \alpha_s)^{m_s}$ with α_i distinct. Set $V_i = \{\boldsymbol{v} \in V | (T - \alpha_i I_V)^{m_i} = \mathbf{0}\}$ so that

$$V = V_1 \oplus \cdots \oplus V_s.$$

Now let $f_i(x) = \frac{f(x)}{(x-\alpha_i)^{m_i}}$. Then $f_i(x)$ and $(x - \alpha_i)^{m_i}$ are relatively prime and consequently there exists polynomials $a_i(x), b_i(x)$ such that

$$a_i(x)f_i(x) + b_i(x)(x - \alpha_i)^{m_i} = 1.$$

Now $a_i(T)f_i(T)$ acts as the zero operator on $V_1 \oplus \cdots \oplus V_{i-1} \oplus V_{i+1} \oplus \cdots \oplus V_s$ and as the identity operator on V_i. Then $\alpha_i a_i(T)f_i(T)$ is the zero operator on $V_1 \oplus \cdots \oplus V_{i-1} \oplus V_{i+1} \oplus \cdots \oplus V_s$ and acts as scalar multiplication by α_i on V_i. Also, $a_i(T)f_i(T)(T - \alpha_i I_V)$ is the zero operator on $V_1 \oplus \cdots \oplus V_{i-1} \oplus V_{i+1} \oplus \cdots \oplus V_s$ and is nilpotent on V_i.

Now set $d(x) = \sum_{i=1}^s \alpha_i a_i(x)f_i(x)$ and $n(x) = \sum_{i=1}^s a_i(x)f_i(x)(x - \alpha_i)$. Then $d(T)$ acts a scalar multiplication by α_i on V_i for each i and $n(T)$ acts as $T - \alpha_i I_V$ on each V_i. Consequently, $d(T)$ is a diagonalizable operator, $n(T)$ is a nilpotent operator and $d(T) + n(T) = T$. Since $d(T), n(T)$ are polynomials in T we have $d(T)n(T) = n(T)d(T)$.

11. Assume first that T has no real eigenvectors. Let $p(x)$ be an irreducible factor of $\mu_T(x)$. Then $p(x)$ has no real roots and therefore $p(x)$ is a real quadratic. Suppose U is a $T-$invariant subspace. Let the elementary divisors of T restricted to U be $p_1(x)^{m_1}, \dots, p_s(x)^{m_s}$ (note that the $p_i(x)$ are not necessarily distinct. Each $p_i(x)$ is a real irreducible quadratic. Then the degree of $p_i(x)^{m_i}$ is $2m_i$ and the dimension of U is the sum $2m_1 + \cdots + 2m_s = 2(m_1 + \dots m_s)$.

12. Let T be the operator with matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ with respect to the standard basis. The minimum polynomial of T is x^2+1 which has no real roots. However, $T^2 = -I_{\mathbb{R}^2}$.

13. $TS = S^{-1}(ST)S$. Thus, TS and ST are similar. Therefore if ST is diagonalizable then so is TS.

Chapter 5

Inner Product Spaces

5.1. Inner Products

1. **Positive definite**. If $\boldsymbol{u} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ then

$$\boldsymbol{u} \cdot \boldsymbol{u} = a_1^2 + \cdots + a_n^n$$

which is non-negative since each square is non-negative and zero if and only if $a_1 = \cdots = a_n = 0$, that if, if and only if $\boldsymbol{u} = \boldsymbol{0}$.

Symmetry If $\boldsymbol{u} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \boldsymbol{v} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ then

$$\boldsymbol{u} \cdot \boldsymbol{v} = a_1 b_1 + \cdots + a_n b_n = b_1 a_1 + \cdots + b_n a_n = \boldsymbol{v} \cdot \boldsymbol{u}$$

since for all real numbers a, b we have commutativity of multiplication: $ab = ba$.

Additivity If $\boldsymbol{u} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \boldsymbol{v} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \boldsymbol{w} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ then

$$(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} =$$

$$(a_1 + b_1)c_1 + \cdots + (a_n + b_n)c_n =$$

$$(a_1 c_1 + b_1 c_1) + \ldots (a_n c_n + b_n c_n)$$

$$= (a_1 c_1 + \cdots + a_n c_n) + (b_1 c_1 + \cdots + b_n c_n) = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}.$$

Essentially additivity holds because multiplication distributes over addition in $\mathbb{R}$.

Homogeneity If $\boldsymbol{u} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \boldsymbol{v} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ and $\gamma \in \mathbb{R}$ then

$$(\gamma \boldsymbol{u}) \cdot \boldsymbol{v} = \begin{pmatrix} \gamma a_1 \\ \vdots \\ \gamma a_n \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

$$= (\gamma a_1) b_1 + \cdots + (\gamma a_n) b) n = \gamma(a_1 b_1) + \cdots + \gamma(a_n b_n)$$

$$= \gamma(a_1 b_1 + \cdots + a_n b_n) = \gamma(\boldsymbol{u} \cdot \boldsymbol{v}).$$

Essentially homogeneity holds since multiplication in $\mathbb{R}$ is associative.

2. $\langle \boldsymbol{u}, \gamma \boldsymbol{v} \rangle = \overline{\langle \gamma \boldsymbol{v}, \boldsymbol{u} \rangle} =$

$$\overline{\gamma \langle \boldsymbol{v}, \boldsymbol{u} \rangle} = \overline{\gamma} \overline{\langle \boldsymbol{v}, \boldsymbol{u} \rangle} = \overline{\gamma} \langle \boldsymbol{u}, \boldsymbol{v} \rangle.$$

3. $\langle \boldsymbol{u}, \boldsymbol{v} + \boldsymbol{w} \rangle = \overline{\langle \boldsymbol{v} + \boldsymbol{w}, \boldsymbol{u} \rangle} =$

$$\overline{\langle \boldsymbol{v}, \boldsymbol{u} \rangle + \langle \boldsymbol{w}, \boldsymbol{u} \rangle} = \overline{\langle \boldsymbol{v}, \boldsymbol{u} \rangle} + \overline{\langle \boldsymbol{w}, \boldsymbol{u} \rangle}$$

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{u}, \boldsymbol{w}\rangle.$$

4. **Positive definite** Let $\boldsymbol{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$. Then $\langle \boldsymbol{w}, \boldsymbol{w}\rangle =$

$$\langle \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \rangle = \alpha_1 w_1 \overline{w_1} + \cdots + \alpha_n w_n \overline{w_n} \quad (5.1)$$

Each $w_i\overline{w_i}$ is non-negative. Since $\alpha_i > 0$ also $\alpha_i w_i \overline{w_i} \geq 0$ and zero if and only if $w_i = 0$. Then each term of (5.1) is non-negative. Therefore, $\langle \boldsymbol{w}, \boldsymbol{w}\rangle$ is non-negative and zero if and only if each term is zero, if and only if $w_i = 0$ for all i, if and only if $\boldsymbol{w} = \boldsymbol{0}$.

Conjugate Symmetry Suppose $\boldsymbol{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ and $\boldsymbol{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$. For each term we have $\overline{\alpha_i u_i \bar{v}_i} = \alpha_i \overline{u_i \bar{v}_i} = \alpha_i v_i \bar{u}_i$ since α_i is real. This implies that $\langle \boldsymbol{v}, \boldsymbol{u}\rangle = \overline{\boldsymbol{u}, \boldsymbol{v}\rangle}$.

Additivity in the first argument This holds since multiplication distributes over addition in $\mathbb{F}$.

Homogeneity in the first argument

This holds since multiplication is associative and commutative: For each term we have

$$\alpha_i(\gamma u_i)v_i = \alpha_i[\gamma(u_i v_i)] =$$

$$(\alpha_i \gamma)(u_i v_i) = (\gamma \alpha_i)(u_i v_i) = \gamma[\alpha_i(u_i v_i)].$$

5. **Positive definite**

$\langle \boldsymbol{u}, \boldsymbol{u}\rangle = \langle S(\boldsymbol{u}), S(\boldsymbol{u})\rangle_{EIP} \geq 0$ since $\langle \ , \ \rangle$ is an inner product and is equal to zero if and only if $S(\boldsymbol{u}) = \boldsymbol{0}$. However, S is an invertible operator and therefore $S(\boldsymbol{u}) = \boldsymbol{0}$ if and only if $\boldsymbol{u} = \boldsymbol{0}$.

Conjuagate symmetry

$\overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle S(\boldsymbol{u}), S(\boldsymbol{v})\rangle_{EIP} = \overline{\langle S(\boldsymbol{v}), S(\boldsymbol{u})\rangle_{EIP}} = \langle \boldsymbol{v}, \boldsymbol{u}\rangle}$.

Additivity in the first argument

$\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{w}\rangle = \langle S(\boldsymbol{u}+\boldsymbol{v}), S(\boldsymbol{w})\rangle_{EIP}$ Since S is linear and $\langle \ , \ \rangle_{EIP}$ is additive in the first argument we have

$$\langle S(\boldsymbol{u}+\boldsymbol{v}), S(\boldsymbol{w})\rangle_{EIP} =$$

$$\langle S(\boldsymbol{u}) + S(\boldsymbol{v}), S(\boldsymbol{w})\rangle_{EIP} =$$

$$\langle S(\boldsymbol{u}), S(\boldsymbol{w})\rangle_{EIP} + \langle S(\boldsymbol{v}), S(\boldsymbol{w})\rangle_{EIP} =$$

$$\langle \boldsymbol{u}, \boldsymbol{w}\rangle + \langle \boldsymbol{v}, \boldsymbol{w}\rangle.$$

Homogeneity in the first argument

By the linearity of S and the homogeneity of $\langle \ , \ \rangle_{EIP}$ in the first argument we have

$$\langle \gamma \boldsymbol{u}, \boldsymbol{v}\rangle = \langle S(\gamma \boldsymbol{u}), S\boldsymbol{v})\rangle_{EIP} = \langle \gamma S(\boldsymbol{u}), S(\boldsymbol{v})\rangle_{EIP}$$

$$\gamma \langle S(\boldsymbol{u}), S(\boldsymbol{v})\rangle_{EIP} = \gamma \langle \boldsymbol{u}, \boldsymbol{v}\rangle.$$

6. **Positive definite**

Let A have entries a_{ij}. The $(i,i)-$entry of $\langle A, A\rangle$ is

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

It follows that $\langle A, A\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{a_{ij}}$. Since each $a_{ij}\overline{a_{ij}} \geq 0$ $\langle A, A\rangle \geq 0$. A term $a_{ij}\overline{a_{ij}} = 0$ if and only if $a_{ij} = 0$ and therefore the sum is zero if and only if $A = \boldsymbol{0}_{nn}$.

Conjugate symmetry

The $(j,j)-$entry of $A^{tr}\overline{B}$ is $\sum_{i=1}^{n} a_{ij}\overline{b_{ij}}$ and therefore $\langle A, B\rangle$ is

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

In a similar fashion

$$\langle B, A\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}\overline{a_{ij}}$$

Since $\overline{a_{ij}\overline{b_{ij}}} = b_{ij}\overline{a_{ij}}$ the conjugate symmetry follows.

Additivity in the first argument

This follows from the formula and the fact that multiplication distributes over addition in $\mathbb{F}$.

Homogeneity in the first argument

This follows from the formula and the fact that multiplication is associative in $\mathbb{F}$.

7. $\langle\ ,\ \rangle$ is an inner product.

Throughout $\boldsymbol{u} = (\boldsymbol{u}_1, \boldsymbol{u}_2), \boldsymbol{v} = (\boldsymbol{v}_1, \boldsymbol{v}_2), \boldsymbol{w} = (\boldsymbol{w}_1, \boldsymbol{w}_2)$ with $\boldsymbol{u}_1, \boldsymbol{v}_1, \boldsymbol{w}_1 \in V_1$ and $\boldsymbol{u}_2, \boldsymbol{v}_2, \boldsymbol{w}_2 \in V_2$.

Positive definite

Assume $\langle \boldsymbol{u}, \boldsymbol{u}\rangle = \langle \boldsymbol{u}_1, \boldsymbol{u}_1\rangle_1 + \langle \boldsymbol{u}_2, \boldsymbol{u}_2\rangle_2$. Each of these is non-negative. Moreover, we get zero if and only if

$$\langle \boldsymbol{u}_1, \boldsymbol{u}_1\rangle_1 = \langle \boldsymbol{u}_2, \boldsymbol{u}_2\rangle_2 = 0$$

if and only if

$$\boldsymbol{u}_1 = \boldsymbol{0}_{V_1}, \boldsymbol{u}_2 = \boldsymbol{0}_{V_2}.$$

Conjugate symmetry

$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{u}_1, \boldsymbol{v}_1\rangle_1 + \langle \boldsymbol{u}_2, \boldsymbol{v}_2\rangle_2 =$

$$\overline{\langle \boldsymbol{v}_1, \boldsymbol{u}_1\rangle_1} + \overline{\langle \boldsymbol{v}_2, \boldsymbol{u}_2\rangle_2} = \overline{\langle \boldsymbol{v}_1, \boldsymbol{u}_1\rangle_1 + \langle \boldsymbol{v}_2, \boldsymbol{u}_2\rangle_2}$$

$$= \overline{\langle \boldsymbol{v}, \boldsymbol{u}\rangle}$$

Additivity in the first argument

$$\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{u}_1+\boldsymbol{v}_1, \boldsymbol{w}_1\rangle_1 + \langle \boldsymbol{u}_2+\boldsymbol{v}_2, \boldsymbol{w}_2\rangle_2 =$$

$$\langle \boldsymbol{u}_1, \boldsymbol{w}_1\rangle_1 + \langle \boldsymbol{v}_1, \boldsymbol{w}_1\rangle_1 + \langle \boldsymbol{u}_2, \boldsymbol{w}_2\rangle_2 + \langle \boldsymbol{v}_2, \boldsymbol{w}_2\rangle_2 =$$

$$(\langle \boldsymbol{u}_1, \boldsymbol{w}_1\rangle_1 + \langle \boldsymbol{u}_2, \boldsymbol{w}_2\rangle_2) + (\langle \boldsymbol{v}_1, \boldsymbol{w}_1\rangle_1 + \langle \boldsymbol{v}_2, \boldsymbol{w}_2\rangle_2)$$

$$= \langle \boldsymbol{u}, \boldsymbol{w}\rangle + \langle \boldsymbol{v}, \boldsymbol{w}\rangle.$$

Homogeneity in the first argument

$$\langle \gamma\boldsymbol{u}, \boldsymbol{v}\rangle = \langle \gamma\boldsymbol{u}_1, \boldsymbol{v}_1\rangle_1 + \langle \gamma\boldsymbol{u}_2, \boldsymbol{v}_2\rangle_2 =$$

$$\gamma\langle \boldsymbol{u}_1, \boldsymbol{v}_1\rangle_1 + \gamma\langle \boldsymbol{u}_2, \boldsymbol{v}_2\rangle_2 =$$

$$\gamma(\langle \boldsymbol{u}_1, \boldsymbol{v}_1\rangle_1 + \langle \boldsymbol{u}_2, \boldsymbol{v}_2\rangle_2) = \gamma\langle \boldsymbol{u}, \boldsymbol{w}\rangle.$$

8. Assume that $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0}$. Then $\boldsymbol{v}\cdot\boldsymbol{v}_j = 0$ for all j. Using additivity and homogeneity in the first argument we get

$$c_1\langle \boldsymbol{v}_1, \boldsymbol{v}_j\rangle + \cdots + c_n\langle \boldsymbol{v}_n, \boldsymbol{v}_j\rangle = 0$$

This implies that $\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ is in the null space of the matrix A. It immediately follows that $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linearly independent if and only if $null(A) = \boldsymbol{0}$ if and only if A is invertible.

9. If $c_i > 0$ for all i then $\langle\ ,\ \rangle$ is an inner product by Exercise 4. Suppose some $c_i \leq 0$. Then $\langle \boldsymbol{e}_i, \boldsymbol{e}_i\rangle \leq 0$ which contradicts positive definiteness. Thus, if $\langle\ ,\ \rangle$ is an inner product, all $c_i > 0$.

10. First note that if $f, g \in V$ then $spt(f) \cup spt(g)$ is finite so there are only finitely many non-zero terms in $\sum_{i\in\mathbb{N}} f(i)g(i)$.

Positive definite

Let $I = spt(f)$. Then $\langle f, f\rangle = \sum_{i\in I} f(i)^2$. Each $f(i)^2 > 0$ for $i \in I$. Therefore $\langle f, f\rangle \geq$. Moreover we obtain a positive sum unless $I = \emptyset$, that is, unless f is the zero function.

Symmetry

This follows since for all $i \in spt(f) \cap spt(g)$, $f(i)g(i) = g(i)f(i)$ since multiplication of real numbers is commutative.

Additivity in the first argument

Set $J = [spt(f) \cap spt(h)] \cup [spt(g) \cap spt(h)]$. Let $f, g, h \in V$ and $i \in J$. Then

$$[f(i) + g(i)]h(i) = f(i)h(i) + g(i)h(i)$$

since multiplication distributes over addition in $\mathbb{R}$.

It then follows that $\langle f+g, h\rangle = \sum_{i\in J}[f(i)+g(i)]h(i) = \sum_{iinJ}[f(i)h(i) + g(i)h(i)] =$

$$\sum_{i\in J} f(i)h(i) + \sum_{i\in J} g(i)h(i) = \langle f, h\rangle + \langle g, h\rangle.$$

Homogeneity in the first argument

Set $I = spt(f) \cap spt(g)$. Then $\langle \gamma f, g\rangle =$

$$\sum_{i\in I}[\gamma f(i)]g(i) = \sum_{i\in I} \gamma(f(i)g(i)) =$$

$$\gamma \sum_{i\in I} f(i)g(i) = \gamma\langle f, g\rangle.$$

11. This is a real inner product.

Positive definite

Since $\langle \boldsymbol{v}, \boldsymbol{v}\rangle$ is a non-negative real number, $\langle \boldsymbol{v}, \boldsymbol{v}\rangle_{\mathbb{R}} = \langle \boldsymbol{v}, \boldsymbol{v}\rangle$ and so is non-negative and equal to zero if and only if $\boldsymbol{v} = \boldsymbol{0}$.

Symmetry

Suppose $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = a + bi$ with $a, b \in \mathbb{R}$. Then $\langle \boldsymbol{v}, \boldsymbol{w}\rangle_{\mathbb{R}} = a$. We then have $\langle \boldsymbol{w}, \boldsymbol{v}\rangle = \overline{\langle \boldsymbol{v}, \boldsymbol{w}\rangle} = \overline{a + bi} = a - bi$. Thus, $\langle \boldsymbol{v}, \boldsymbol{w}\rangle_{\mathbb{R}} = a$.

Additivity in the first argument

Suppose $\langle \boldsymbol{u}, \boldsymbol{w}\rangle = a+bi, \langle \boldsymbol{v}, \boldsymbol{w}\rangle = c+di$ with $a, b, c, d \in \mathbb{R}$. Then $\langle \boldsymbol{u}, \boldsymbol{w}\rangle_{\mathbb{R}} = a, \langle \boldsymbol{v}, \boldsymbol{w}\rangle_{\mathbb{R}} = c$.

$$\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{w}\rangle + \langle \boldsymbol{v}, \boldsymbol{w}\rangle =$$

$$(a + bi) + (c + di)_{=}(a + c) + (b + d)i.$$

Thus, $\langle \boldsymbol{u} + \boldsymbol{w}, \rangle_{\mathbb{R}} = a + b$ as required.

Homogeneity in the first argument

Assume $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = a + bi$ so that $\langle \boldsymbol{v}, \boldsymbol{w}\rangle_{\mathbb{R}} = a$. Let $\gamma \in \mathbb{R}$. Then $\langle \gamma\boldsymbol{v}, \boldsymbol{w}\rangle = \gamma\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \gamma(a + bi) = (\gamma a) + (\gamma b)i$. It follows that $\langle \gamma\boldsymbol{v}, \boldsymbol{w}\rangle_{\mathbb{R}} = \gamma a = \gamma\langle \boldsymbol{v}, \boldsymbol{w}\rangle_{\mathbb{R}}$.

5.2. The Geometry of Inner Product Spaces

1. Let $\boldsymbol{v}, \boldsymbol{w} \in \boldsymbol{u}^{\perp}$ so that $\langle \boldsymbol{v}, \boldsymbol{u}\rangle = \langle \boldsymbol{w}, \boldsymbol{u}\rangle = 0$.

By additivity in the first argument we have

$$\langle \boldsymbol{v} + \boldsymbol{w}, \boldsymbol{u}\rangle = \langle \boldsymbol{v}, \boldsymbol{u}\rangle + \langle \boldsymbol{w}, \boldsymbol{u}\rangle = 0 + 0 = 0.$$

Thus, $\boldsymbol{v} + \boldsymbol{w} \in \boldsymbol{u}^{\perp}$.

Now assume $\gamma \in \mathbb{F}$. By homogeneity in the first argument we have

$$\langle \gamma\boldsymbol{v}, \boldsymbol{u}\rangle = \gamma\langle \boldsymbol{v}, \boldsymbol{u}\rangle = \gamma \times 0 = 0.$$

So, if $\boldsymbol{v} \in \boldsymbol{u}^{\perp}$ and γ is a scalar then $\gamma\boldsymbol{v} \in \boldsymbol{u}^{\perp}$.

2. Define a map $f_{\boldsymbol{u}} : V \to \mathbb{F}$ by $f_{\boldsymbol{u}}(\boldsymbol{v}) = \langle \boldsymbol{v}, \boldsymbol{u}\rangle$. Since $\langle\ ,\ \rangle$ is additive and homogeneous in the first argument, the map $f_{\boldsymbol{u}}$ is a linear transformation. Assume $\boldsymbol{u} \neq \boldsymbol{0}$. Then for $a \in \mathbb{F}$, $f_{\boldsymbol{u}}(\frac{a}{\langle \boldsymbol{u}, \boldsymbol{u}\rangle}\boldsymbol{u}) = a$ and consequently, $f_{\boldsymbol{u}}$ is surjective. Now by the rank nullity theorem, $dim(Ker(f_{\boldsymbol{u}}) = n - 1$. However, $Ker(f_{\boldsymbol{u}})) = \boldsymbol{u}^{\perp}$.

3. Assume $\boldsymbol{w} \in W \cap W^{\perp}$. Then $\boldsymbol{w} \perp \boldsymbol{w}$ that is, $\langle \boldsymbol{w}, \boldsymbol{w}\rangle = 0$. By positive definiteness, $\boldsymbol{w} = \boldsymbol{0}$.

4. $\langle ax^2 + bx + c, x^2 + x + 1\rangle = \int_0^1 (ax^2 + bx + c)(x^2 + x + 1)dx =$

$$\int_0^1 [ax^4 + (a+b)x^3 + (a+b+c)x^2 + (b+c)x + c]dx =$$

$$(a\frac{x^5}{5} + (a+b)\frac{x^4}{4} + (a+b+c)\frac{x^3}{3} + (b+c)\frac{x^2}{2} + cx|_0^1 =$$

$$\frac{a}{5} + \frac{a+b}{4} + \frac{a+b+c}{3} + \frac{b+c}{2} + c$$

This reduces to finding the solutions to the homogeneous equation

$$47a + 65b + 110c = 0.$$

A basis for $(x^2+x+1)^\perp$ is $(\frac{110}{47}x^2 - 1, \frac{65}{47}x^2 - 1)$.

5. $d(A,B) = \sqrt{\langle A-B, A-B\rangle}$. $A - B = \begin{pmatrix} -4 & -3 \\ 3 & -4 \end{pmatrix}$.

We have to compute the trace of the product $(A-B)^{tr}(A-B) = \begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix}$. Thus, $d(A,B) = 5\sqrt{2}$.

6. $\langle A, I_2\rangle$ is equal to the trace of A^{tr} which is the same as the trace of A. So, the orthogonal complement to I_2 consists of all matrices with trace zero. A basis for this is

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

7. The subspace of diagonal matrices in $M_{22}(\mathbb{R})$ has basis $(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix})$.

The orthogonal complement to $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ consists of all matrices $\begin{pmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. The orthogonal complement to the matrix $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ consists of all matrices $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{pmatrix}$. Thus, the orthogonal complement to the space of diagonal matrices in $M_{22}(\mathbb{R})$ consists of all matrices with zeros on the diagonal.

8. $d(x^2, x) = \sqrt{\langle x^2 - x, x^2 - x\rangle}$.

$\langle x^2 - x, x^2 - x\rangle = \int_0^1 (x^2 - x)^2 dx =$

$\int_0^1 (x^4 - 2x^3 + x^2)dx = [\frac{x^5}{5} - 2\frac{x^4}{4} + \frac{x^3}{3}]|_0^1 =$

$$\frac{1}{5} - \frac{2}{4} + \frac{1}{3} = \frac{1}{15}.$$

$$d(x^2, x) = \sqrt{\frac{1}{15}} = \frac{\sqrt{15}}{15}.$$

9. $\langle \boldsymbol{u} - \frac{\langle \boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{v}\|^2}\boldsymbol{v}, \boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{v}\rangle - \frac{\langle \boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{v}\|^2}\langle \boldsymbol{v}, \boldsymbol{v}\rangle =$

$\langle \boldsymbol{u}, \boldsymbol{v}\rangle - \frac{\langle \boldsymbol{u},\boldsymbol{v}\rangle}{\|\boldsymbol{v}\|^2} \parallel \boldsymbol{v} \parallel^2 = \langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0.$

10. $\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{u}\rangle + \langle \boldsymbol{u}, \boldsymbol{v}\rangle + \overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} + \langle \boldsymbol{v}, \boldsymbol{v}\rangle.$

$$\langle \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{u}-\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{u}\rangle - \langle \boldsymbol{u}, \boldsymbol{v}\rangle - \overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} + \langle \boldsymbol{v}, \boldsymbol{v}\rangle.$$

$$\langle \boldsymbol{u}+i\boldsymbol{v}, \boldsymbol{u}+i\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{u}\rangle - i\langle \boldsymbol{u}, \boldsymbol{v}\rangle + i\overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} + \langle \boldsymbol{v}, \boldsymbol{v}\rangle.$$

$$\langle \boldsymbol{u}-i\boldsymbol{v}, \boldsymbol{u}-i\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{u}\rangle + i\langle \boldsymbol{u}, \boldsymbol{v}\rangle - i\overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} + \langle \boldsymbol{v}, \boldsymbol{v}\rangle.$$

We then have

$$\parallel \boldsymbol{u}+\boldsymbol{v} \parallel^2 - \parallel \boldsymbol{u}-\boldsymbol{v} \parallel^2 + i \parallel \boldsymbol{u}+i\boldsymbol{v} \parallel^2 - i \parallel \boldsymbol{u}-i\boldsymbol{v} \parallel^2 =$$

$$(\langle \boldsymbol{u}, \boldsymbol{u}\rangle - \langle \boldsymbol{u}, \boldsymbol{u}\rangle + i\langle \boldsymbol{u}, \boldsymbol{u}\rangle - i\langle \boldsymbol{u}, \boldsymbol{u}\rangle) +$$

$$(\langle \boldsymbol{u}, \boldsymbol{v}\rangle - (-\langle \boldsymbol{u}, \boldsymbol{v}\rangle) - i^2\langle \boldsymbol{u}, \boldsymbol{v}\rangle - i^2\langle \boldsymbol{u}, \boldsymbol{v}\rangle) +$$

$$\overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} + \overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} + i^2\overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} + i^2\overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} +$$

$$(\langle \boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{v}, \boldsymbol{v}\rangle + i\langle \boldsymbol{v}, \boldsymbol{v}\rangle - i\langle \boldsymbol{v}, \boldsymbol{v}\rangle) =$$

$$4\langle \boldsymbol{u}, \boldsymbol{v}\rangle.$$

11. Set $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \frac{x_2}{\sqrt{2}} \\ \vdots \\ \frac{x_n}{\sqrt{n}} \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} y_1 \\ \sqrt{2}y_2 \\ \vdots \\ \sqrt{n}y_n \end{pmatrix}$. Apply Cauchy-Schwartz:

$$(\boldsymbol{x}\cdot\boldsymbol{y})^2 \le (\boldsymbol{x}\cdot\boldsymbol{x})(\boldsymbol{y}\cdot\boldsymbol{y}).$$

The left hand side is

$$(\sum_{i=1}^n x_i y_i)^2$$

while the right hand side is

$$(\sum_{i=1}^n \frac{x_i^2}{i})(\sum_{i=1}^n i y_i^2).$$

12. a) Since $d(\boldsymbol{u},\boldsymbol{v}) = \sqrt{\langle \boldsymbol{u}-\boldsymbol{v},\boldsymbol{u}-\boldsymbol{v}\rangle}$ we have $d(\boldsymbol{u},\boldsymbol{v}) \ge 0$ and equals zero if and only if $\boldsymbol{u}-\boldsymbol{v}=\boldsymbol{0}$, if and only if $\boldsymbol{u}=\boldsymbol{v}$.

b) This follows since $\langle -\boldsymbol{x},-\boldsymbol{x}\rangle = \langle \boldsymbol{x},\boldsymbol{x}\rangle$ for any vector $\boldsymbol{x}$.

c) $d(\boldsymbol{u},\boldsymbol{w}) = \parallel \boldsymbol{u}-\boldsymbol{w} \parallel = \parallel (\boldsymbol{u}-\boldsymbol{v})+(\boldsymbol{v}-\boldsymbol{w}) \parallel$

$\le \parallel \boldsymbol{u}-\boldsymbol{v} \parallel + \parallel \boldsymbol{v}-\boldsymbol{w} \parallel = d(\boldsymbol{u},\boldsymbol{v})+d(\boldsymbol{v},\boldsymbol{w})$ by Theorem (5.5).

13. Denote the 2×2 identity by I_2 and the 2×2 all one matrix by J_2. Easy calculations give $\parallel I_2 \parallel = \sqrt{2}$,

$\parallel J_2 \parallel = 2\sqrt{2}$ and $\langle I_2, J_2\rangle = 2$. Thus,

$$\frac{\langle I_2,J_2\rangle}{\parallel I_2 \parallel\parallel J_2 \parallel} = \frac{2}{\sqrt{2}2\sqrt{2}} = \frac{2}{4} = \frac{1}{2}.$$

The angle is $\frac{\pi}{3}$.

14. If $\parallel \boldsymbol{u}+\boldsymbol{v} \parallel = \parallel \boldsymbol{u} \parallel + \parallel \boldsymbol{v} \parallel$ then $\boldsymbol{u}\perp\boldsymbol{v}$. Then $(c\boldsymbol{u})\perp(d\boldsymbol{v})$. Then by the general Pythagorean theorem

$$\parallel c\boldsymbol{u}+d\boldsymbol{v} \parallel^2 = \parallel c\boldsymbol{u} \parallel^2 + \parallel d\boldsymbol{v} \parallel^2 = c^2 \parallel \boldsymbol{u} \parallel^2 + d^2 \parallel \boldsymbol{v} \parallel^2 .$$

15. It is a consequence of the assumption that $\parallel \boldsymbol{u} \parallel_1 =$

$\parallel \boldsymbol{u} \parallel_2$ for all vectors $\boldsymbol{u}$. Then $\parallel \boldsymbol{u}+\boldsymbol{v} \parallel_1^2 =$

$\parallel \boldsymbol{u}+\boldsymbol{v} \parallel_2^2$. It then follows that

$$2\langle \boldsymbol{u},\boldsymbol{v}\rangle_1 = 2\langle \boldsymbol{u},\boldsymbol{v}\rangle_2$$

and so we get that $\langle\ ,\ \rangle_1$ and $\langle\ ,\ \rangle_2$ are identical.

16. Let a be the scalar such that $\boldsymbol{z} = \boldsymbol{y} - a\boldsymbol{x}$ is orthogonal to $\boldsymbol{x}$. Then $\boldsymbol{y} = a\boldsymbol{x}+\boldsymbol{z}$ and $\langle \boldsymbol{y},\boldsymbol{y}\rangle = \langle a\boldsymbol{x}+\boldsymbol{z}, a\boldsymbol{x}+\boldsymbol{z}\rangle = |a|^2 + \langle \boldsymbol{z},\boldsymbol{z}\rangle \ge a^2$. On the other hand, $\langle \boldsymbol{y},\boldsymbol{x}\rangle\langle \boldsymbol{x},\boldsymbol{y}\rangle = \langle a\boldsymbol{x}+\boldsymbol{z},\boldsymbol{x}\rangle\langle \boldsymbol{x}, a\boldsymbol{x}+\boldsymbol{z}\rangle = a\overline{a} = |a|^2$.

5.3. Orthonormal Sets and the Gram-Schmidt Process

1. $\boldsymbol{x}_2 = \boldsymbol{w}_2 - \frac{\langle \boldsymbol{w}_2,\boldsymbol{x}_1\rangle}{\langle \boldsymbol{x}_1,\boldsymbol{x}_1}\boldsymbol{x}_1$. Then

$$\langle \boldsymbol{x}_2,\boldsymbol{x}_1\rangle = \langle \boldsymbol{w}_2 - \frac{\langle \boldsymbol{w}_2,\boldsymbol{x}_1\rangle}{\langle \boldsymbol{x}_1,\boldsymbol{x}_1}\boldsymbol{x}_1,\boldsymbol{x}_1\rangle =$$

$$\langle \boldsymbol{w}_2,\boldsymbol{x}_1\rangle - \frac{\langle \boldsymbol{w}_2,\boldsymbol{x}_1\rangle}{\langle \boldsymbol{x}_1,\boldsymbol{x}_1\rangle}\langle \boldsymbol{x}_1,\boldsymbol{x}_1\rangle =$$

$$\langle \boldsymbol{w}_2,\boldsymbol{x}_1\rangle - \langle \boldsymbol{w}_2,\boldsymbol{x}_1\rangle = 0.$$

2. $\boldsymbol{x}_3 = \boldsymbol{w}_3 - \frac{\langle \boldsymbol{w}_3,\boldsymbol{x}_1\rangle}{\langle \boldsymbol{x}_1,\boldsymbol{x}_1\rangle}\boldsymbol{x}_1 - \frac{\langle \boldsymbol{w}_3,\boldsymbol{x}_2\rangle}{\langle \boldsymbol{x}_2,\boldsymbol{x}_2\rangle}\boldsymbol{x}_2$.

Making use of additivity and homogeneity in the first argument and the fact that $\boldsymbol{x}_1\perp\boldsymbol{x}_2$ we get

$$\langle \boldsymbol{w}_3,\boldsymbol{x}_1\rangle = \langle \boldsymbol{w}_3 - \frac{\langle \boldsymbol{w}_3,\boldsymbol{x}_1\rangle}{\langle \boldsymbol{x}_1,\boldsymbol{x}_1\rangle}\boldsymbol{x}_1 - \frac{\langle \boldsymbol{w}_3,\boldsymbol{x}_2\rangle}{\langle \boldsymbol{x}_2,\boldsymbol{x}_2\rangle}\boldsymbol{x}_2,\boldsymbol{x}_1\rangle =$$

$$\langle \boldsymbol{w}_3,\boldsymbol{x}_1\rangle - \frac{\langle \boldsymbol{w}_3,\boldsymbol{x}_1\rangle}{\langle \boldsymbol{x}_1,\boldsymbol{x}_1\rangle}\langle \boldsymbol{x}_1,\boldsymbol{x}_1\rangle = \langle \boldsymbol{w}_3,\boldsymbol{x}_1\rangle - \langle \boldsymbol{w}_3,x_1\rangle = 0.$$

$\langle \boldsymbol{w}_3,\boldsymbol{x}_2\rangle$ is computed in exactly the same way.

3. Let $\boldsymbol{u}\in U$ and $\boldsymbol{x}\in W^\perp$. Since $U\subset W, \boldsymbol{u}\perp\boldsymbol{x}$. Since $\boldsymbol{u}$ is arbitrary, $\boldsymbol{x}\in U^\perp$. Since $\boldsymbol{x}$ is arbitrary we have $W^\perp\subset U^\perp$.

4. Let $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_k)$ be basis for W and extend to a basis $(\boldsymbol{w}_1,\dots,\boldsymbol{w}_n)$ for V. Apply the Gram-Schmidt process to obtain an orthonormal basis $(\boldsymbol{x}_1,\dots,\boldsymbol{x}_n)$ such that for all $j\le n, Span(\boldsymbol{x}_1,\dots,\boldsymbol{x}_j) = Span(\boldsymbol{w}_1,\dots,\boldsymbol{w}_j)$. Then, in particular, $(\boldsymbol{x}_1,\dots,\boldsymbol{x}_k)$ is an orthonormal basis of W.

Clearly $Span(\boldsymbol{w}_{k+1}, \dots, \boldsymbol{w}_n)$ is contained in $W^\perp$ and has dimension $n-k$. Thus, $dim(W^\perp) \geq n-k$. Since $W \cap W^\perp = \{\boldsymbol{0}\}, dim(W^\perp) \leq n-k$. Consequently we have equality.

5. This follows from the fact that $W \cap W^\perp = \{\boldsymbol{0}\}$ and $dim(W) + dim(W^\perp) = n$.

6. Assume $dim(V) = n, dim(W) = k$. Then $dim(W^\perp) = n-k$ by Theorem (5.11). By another application of Theorem (5.11) we get that $dim((W^\perp)^\perp) = k$. On the other hand, $W \subset (W^\perp)^\perp$. Since they have the same dimension we get equality.

7. Let $\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{w} \in U + W$ and $\boldsymbol{y} \in U^\perp \cap W^\perp$. Then $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{u} + \boldsymbol{w}, \boldsymbol{y} \rangle = \langle \boldsymbol{u}, \boldsymbol{y} \rangle + \langle \boldsymbol{w}, \boldsymbol{y} \rangle = 0 + 0 = 0$. Thus, $U^\perp \cap W^\perp \subset (U+W)^\perp$.

On the other hand, since $U \subset U + W, (U+W)^\perp \subset U^\perp$ by Exercise 3. Similarly, $(U+W)^\perp \subset W^\perp$. Thus $(U+W)^\perp \subset U^\perp \cap W^\perp$ and we have equality.

Now apply the first part to $U^\perp + W^\perp$:

$$(U^\perp + W^\perp)^\perp = (U^\perp)^\perp \cap (W^\perp)^\perp = U \cap W.$$

It now follows from Exercise 6 that $(U \cap W)^\perp = U^\perp + W^\perp$.

8. For each $j, Span(\boldsymbol{w}_1, \dots, \boldsymbol{w}_j) = Span(\boldsymbol{v}_1, \dots, \boldsymbol{v}_j)$. This implies that $[\boldsymbol{w}_j]_{\mathcal{B}'}$ has the form $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{jj} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Therefore, the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}', \mathcal{M}_{I_V}(\mathcal{B}, \mathcal{B}')$ is upper triangular. Reversing the roles of the bases, it also follows that $\mathcal{M}_{I_V}(\mathcal{B}', \mathcal{B})$ is upper triangular.

9. $(1, x - \frac{1}{2}, x^2 - x + \frac{1}{6})$.

10. Extend $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k)$ to an orthonormal basis of V. Set $\langle \boldsymbol{u}, \boldsymbol{v}_i \rangle = c_i$ so that $\boldsymbol{u} = c_1\boldsymbol{v}_1 + \dots + c_n\boldsymbol{v}_n$. Then

$$\sum_{i=1}^{k} |\langle \boldsymbol{u}, \boldsymbol{v}_i \rangle|^2 = \sum_{i=1}^{k} |c_i|^2.$$

On the other hand

$$\| \boldsymbol{u} \|^2 = \sum_{i=1}^{n} |c_i|^2 \geq \sum_{i=1}^{k} |c_i|^2.$$

We get equality if and only if $c_{k+1} = \dots = c_n = 0$ which occurs if and only if $\boldsymbol{u}$ is in the span of $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k)$.

11. $J_2^\perp$ consists of all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $a + b + c + d = 0$ and has basis $(\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix})$. Applying Gram-Schmidt to this basis we obtain the following orthogonal basis:

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

The first matrix has norm $\sqrt{3}$ the second has norm $\sqrt{\frac{3}{2}}$ and the last has norm $\sqrt{\frac{4}{3}}$. Dividing the respective vectors by these numbers gives an orthonormal basis.

12. Set $a_i = \langle \boldsymbol{x}, \boldsymbol{v}_i \rangle, b_i = \langle \boldsymbol{y}, \boldsymbol{v}_i \rangle$. Since $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$is an orthonormal basis we have

$$\boldsymbol{x} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i, \boldsymbol{y} = \sum_{i=1}^{n} b_i \boldsymbol{v}_i.$$

Again, since $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is an orthonormal basis,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{n} a_i \overline{b_i} = \sum_{i=1}^{n} \langle \boldsymbol{x}, \boldsymbol{v}_i \rangle \overline{\langle \boldsymbol{y}, \boldsymbol{v}_i \rangle}.$$

5.4. Orthogonal Complements and Projections

1. $Proj_W(\boldsymbol{u}) = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 3 \end{pmatrix}, Proj_{W^\perp}(\boldsymbol{u}) = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$

2. $Proj_W(J_2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Proj_{W^\perp}(J_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

3. $Proj_W(x^3) = \frac{3}{2}x^2 - \frac{3}{5}x + \frac{1}{20}$.

4. $\frac{5}{3}$.

5. $\frac{\sqrt{30}}{3}$

6. $\frac{2\sqrt{15}}{5}$

7. $\frac{1}{35}(-20x^2 + 48x + 6)$.

8. Extend $\mathcal{B}$ to an orthonormal basis $\mathcal{B}' = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$. We need to show that $Q[\boldsymbol{w}_i]_{\mathcal{S}} = [\boldsymbol{w}_i]_{\mathcal{S}}$ for $i \leq k$ and $\boldsymbol{0}_n$ if $i > k$. Since $\mathcal{S}$ is an orthonormal basis, for any two vectors $\boldsymbol{u}, \boldsymbol{v}$ we have $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = [\boldsymbol{u}]_{\mathcal{S}} \cdot [\boldsymbol{v}]_{\mathcal{S}}$. Now for $1 \leq i < j \leq k$ we have $\langle \boldsymbol{w}_i, \boldsymbol{w}_j \rangle = 0$ which implies that $[\boldsymbol{w}_i]_{\mathcal{S}} \cdot [\boldsymbol{w}_j]_{\mathcal{S}} = 0$ and $[\boldsymbol{w}_i]_{\mathcal{S}} \cdot [\boldsymbol{w}_i]_{\mathcal{S}} = 1$. This implies that $A^{tr}[\boldsymbol{w}_i]_{\mathcal{S}} = \boldsymbol{e}_i$ the i^{th} standard basis vector of $\mathbb{R}^n$. Whence $Q[\boldsymbol{w}_i]_{\mathcal{S}} = AA^{tr}[\boldsymbol{w}_i]_{\mathcal{S}} = A\boldsymbol{e}_i = [\boldsymbol{w}_i]_{\mathcal{S}}$ as required. On the other hand, if $1 \leq i \leq k, k+1 \leq j \leq n$ then $A^{tr}[\boldsymbol{w}_j]_{\mathcal{S}} = \boldsymbol{0}_n$ and, consequently, $Q[\boldsymbol{w}_j]_{\mathcal{S}} = \boldsymbol{0}_n$.

9. $Q^{tr} = (AA^{tr})^{tr} = (A^{tr})^{tr}A^{tr} = AA^{tr} = Q$, so Q is symmetric. Let $\boldsymbol{v} \in V$ and write $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{u}$ where $\boldsymbol{w} \in W, \boldsymbol{u} \in W^\perp$. Then $Proj_W^2(\boldsymbol{v}) = Proj_W(\boldsymbol{w}+\boldsymbol{u}) = Proj_W(\boldsymbol{w}) = \boldsymbol{w} = Proj_W(\boldsymbol{v})$. It follows from this that $Q^2 = Q$ since Q is the matrix of $Proj_W$ with respect to $\mathcal{S}$.

10. Let $\boldsymbol{u}, \boldsymbol{v}$ be in V. As stated in Exercise 8, $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = [\boldsymbol{u}]_{\mathcal{S}} \cdot [\boldsymbol{v}]_{\mathcal{S}}$. Now suppose $\boldsymbol{w} \in W = Range(T)$ and $\boldsymbol{u} \in Ker(T)$. We claim that $\langle \boldsymbol{w}, \boldsymbol{u} \rangle = 0$. By the above remark it suffices to show that $[\boldsymbol{w}]_{\mathcal{S}} \cdot [\boldsymbol{u}]_{\mathcal{S}} = 0$ which is equivalent to $[\boldsymbol{w}]_{\mathcal{S}}^{tr}[\boldsymbol{u}]_{\mathcal{S}} = 0$. Since $\boldsymbol{w} \in Range(T)$ there exists a vector $\boldsymbol{v} \in V$ such that $\boldsymbol{w} = T(\boldsymbol{v})$. Then $[\boldsymbol{w}]_{\mathcal{S}} = [T(\boldsymbol{v})]_{\mathcal{S}} = Q[\boldsymbol{v}]_{\mathcal{S}}$. Since $\boldsymbol{u} \in Ker(T), Q[\boldsymbol{u}]_{\mathcal{S}} = \boldsymbol{0}_n$. We now use this to compute $[\boldsymbol{w}]_{\mathcal{S}}^{tr}[\boldsymbol{u}]_{\mathcal{S}} = 0$:

$$[\boldsymbol{w}]_{\mathcal{S}}^{tr}[\boldsymbol{u}]_{\mathcal{S}} = (Q[\boldsymbol{v}]_{\mathcal{S}})^{tr}[\boldsymbol{u}]_{\mathcal{S}} = [\boldsymbol{v}]^{tr}Q^{tr}[\boldsymbol{u}]_{\mathcal{S}}$$

$$= [\boldsymbol{v}]_{\mathcal{S}}^{tr}(Q[\boldsymbol{u}]_{\mathcal{S}}) = [\boldsymbol{v}]_{\mathcal{S}}^{tr}\boldsymbol{0}_n = \boldsymbol{0}_n.$$

It now follows that $W' = Ker(T) \subset W^\perp$. However, by the rank nullity theorem $dim(W') = n - dim(W)$. By Theorem (5.11), $dim(W^\perp) = n - dim(W)$ and therefore $W' = W^\perp$.

Now let $\boldsymbol{w} \in W$. It remains to show that $T(\boldsymbol{w}) = \boldsymbol{w}$. Since $Q^2 = Q$ it follows that $T^2 = T$. Let $\boldsymbol{u} \in V$ such that $T(\boldsymbol{u}) = \boldsymbol{w}$. Then $T(\boldsymbol{w}) = T^2(\boldsymbol{u}) = T(\boldsymbol{u}) = \boldsymbol{w}$.

11. Assume $W \perp U$ and let $\boldsymbol{v} \in V$. Set $\boldsymbol{w} = Proj_W(\boldsymbol{v})$. Then $\boldsymbol{w} \in W$. Since $\boldsymbol{w} \perp U$ it follows that $Proj_U(\boldsymbol{w}) = \boldsymbol{0}$. Conversely, assume $Proj_U \circ Proj_W = 0_{V \to V}$. Let $\boldsymbol{w} \in W$. Then $Proj_W(\boldsymbol{w}) = \boldsymbol{w}$. Then $Proj_U(\boldsymbol{w}) = \boldsymbol{0}$ which implies that $\boldsymbol{w} \in Ker(Proj_U) = U^\perp$. Since $\boldsymbol{w}$ is arbitrary we have $W \subset U^\perp$ and therefore $W \perp U$.

12. Let $\boldsymbol{w} \in W, \boldsymbol{v} \in W^\perp$ such that $\boldsymbol{u} = \boldsymbol{w} + \boldsymbol{v}$. Then $Proj_W(\boldsymbol{v}) = \boldsymbol{w}$ and $\langle Proj_W(\boldsymbol{u}), Proj_W(\boldsymbol{u}) \rangle = \langle \boldsymbol{w}, \boldsymbol{w} \rangle$. On the other hand, since $\boldsymbol{w} \perp \boldsymbol{v}$ by the Pythagorean theorem $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = \langle \boldsymbol{w} + \boldsymbol{v}, \boldsymbol{w} + \boldsymbol{v} \rangle = \langle \boldsymbol{w}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq \langle \boldsymbol{w}, \boldsymbol{w} \rangle$. Moreover, we get equality if and only if $\boldsymbol{v} = \boldsymbol{0}$ if and only if $\boldsymbol{u} = \boldsymbol{w} \in W$.

13. Let $\boldsymbol{w} \in W, \boldsymbol{v} \in W^\perp$ such that $\boldsymbol{u} = \boldsymbol{w} + \boldsymbol{v}$. Then $dist(\boldsymbol{u}, W) = \langle \boldsymbol{v}, \boldsymbol{v} \rangle$. By the Pythagorean theorem, $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = \langle \boldsymbol{w} + \boldsymbol{v}, \boldsymbol{w} + \boldsymbol{v} \rangle = \langle \boldsymbol{w}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq \langle \boldsymbol{v}, \boldsymbol{v} \rangle$. Moreover, we have equality if and only if $\boldsymbol{w} = \boldsymbol{0}$ if and only if $\boldsymbol{u} = \boldsymbol{v} \in W^\perp$.

5.5. Dual Spaces

1. $\begin{pmatrix} 1 & 2 & 1 & 2 \\ 2 & 3 & 1 & 3 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 \end{pmatrix}$ is invertible with inverse

$\begin{pmatrix} -5 & 3 & 0 & 1 \\ -1 & 1 & -1 & 0 \\ -2 & 1 & 0 & 1 \\ 5 & -3 & 1 & -1 \end{pmatrix}$. This implies that the sequence of column vectors are a basis for $\mathbb{R}^4$.

Set

$$g_1 = -5f_1 + 3f_2 + f_4$$

$$g_2 = -f_1 + f_2 - f_3$$

$$g_3 = -2f_1 + f_2 + f_4$$

$$g_4 = 5f_1 - 3f_2 + f_3 - f_4$$

(g_1, g_2, g_3, g_4) is the basis of $(\mathbb{R}^4)'$ which is dual to $\mathcal{B}$.

2. Since $dim(\mathcal{L}(V, W)) = dim(\mathcal{L}(W', V'))$ it suffices to show that the map $T \to T'$ is injective. Let $T \in \mathcal{L}(V, W)$ be a non-zero map. We need to show that T' is non-zero. Since T is non-zero there exists $\boldsymbol{v} \in V$ such that $\boldsymbol{w} = T(\boldsymbol{v}) \neq \mathbf{0}_W$. Extend $\boldsymbol{w}$ to a basis $(\boldsymbol{w} = \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ for W. Define $g : W \to \mathbb{F}$ by $g(\sum_{i=1}^m c_i\boldsymbol{w}_i) = a_i$. Now $[T'(g)](\boldsymbol{v}) = [g \circ T](\boldsymbol{v}) = g(T(\boldsymbol{v})) = g(\boldsymbol{w}) = 1$. Thus, $T'(g) \neq \mathbf{0}_{W'}$.

3. Assume T is one-to-one and let $f \in V'$ be non-zero. Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1})$ be a basis for $Ker(f)$ and extend to a basis $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ for V such that $f(\boldsymbol{v}_n) = 1$. Set $\boldsymbol{w}_i = T(\boldsymbol{v}_i), 1 \leq i \leq n$. Since T is one-to-one, $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ is linearly independent. Extend to a basis $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ fo W. Define $g : W \to \mathbb{F}$ by $g(\sum_{i=1}^m a_i\boldsymbol{w}_i) = a_1$. Then $T'(g)(\boldsymbol{v}_1) = (g \circ T)(\boldsymbol{v}_1) = g(T(\boldsymbol{v}_1)) = g(\boldsymbol{w}_1) = 1$. Also, for $i > 1, T'(g)(\boldsymbol{v}_i)(g \circ T)(\boldsymbol{v}_i) = g(T(\boldsymbol{v}_i)) = g(\boldsymbol{w}_i) = 0$. Thus, $T'(g) = f$ and T' is onto.

Conversely, assume that T' is onto and let $\boldsymbol{v} \in V$ be a non-zero vector. Let $(\boldsymbol{v} = \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis of V. And let $f : V \to \mathbb{F}$ be defined by $f(\sum_{i=1}^n a_i\boldsymbol{v}_i) = a_1$. Since T' is onto there exists $g \in W'$ such that $T'(g) = f$. Then $(T'(g))(\boldsymbol{v}) = f(\boldsymbol{v}) = 1$. Whence $g(T(\boldsymbol{v})) = 1$ which clearly implies that $T(\boldsymbol{v}) \neq \mathbf{0}_W$. Since $\boldsymbol{v}$ is arbitrary we can conclude that $Ker(T) = \{\mathbf{0}_V\}$ and T is injective.

Now assume that T is onto and let $g \in W'$ be non-zero. Then there exists $\boldsymbol{w} \in W$ such that $g(\boldsymbol{w}) = 1$. Let $(\boldsymbol{w} = \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ be a basis for W such that $(\boldsymbol{w}_2, \ldots, \boldsymbol{w}_m)$ is a basis for $Ker(g)$. Since T is onto, there exists $\boldsymbol{v} \in V$ such that $T(\boldsymbol{v}) = \boldsymbol{w}$. Now $[T'(g)](\boldsymbol{v}) = g(T(\boldsymbol{v})) = g(\boldsymbol{w}) = 1$. In particular, $T'(g)$ is not the zero vector in V', equivalently, g is not in $Ker(T')$. Since g is arbitrary in W' it follows that T' is injective.

Conversely, assume that $Range(T) \neq W$. Let $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k)$ be a basis for $Range(T)$ and extend this to a basis $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ for W. Now define $g : W \to \mathbb{F}$ by $g(\sum_{i=1}^m a_i\boldsymbol{w}_i = a_m$. Note that $Range(T) \subset Ker(T)$. This implies that $T'(g) = \mathbf{0}_{V'}$ and therefore T' is not one-to-one.

4. This follows immediately from Exercise 3.

5. Set $k = rank(T)$ and let $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k)$ be a basis for $Range(T)$. Extend to a basis $\mathcal{B}_W = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ for W and let $\mathcal{B}_{W'} = (g_1, \ldots, g_m)$ be the basis of W' dual to $\mathcal{B}_W$. We claim that $(T'(g_1), \ldots, T'(g_k))$ is a basis for $Range(T')$.

Suppose $j > k$. Then $Range(T) \subset Ker(g_j)$ and therefore $T'(g_j) = 0_{V'}$. Thus, $Range(T') \subset Span(T'(g_1), \ldots, T'(g_k))$. Since each $T'(g_i)$ is in $Range(T')$ we have equality. It remains to show that $(T'(g_1), \ldots, T'(g_k))$ is linearly independent.

For $1 \leq i \leq k$ let $\boldsymbol{v}_i \in V$ such that $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$. Then $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$ is linearly independent since $(T(\boldsymbol{v}_1), \ldots, T(\boldsymbol{v}_k)) = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k)$ is linearly independent. Extend to a basis $\mathcal{B}_V = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ where $(\boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n)$ is a basis for $Ker(T)$. Let $\mathcal{B}_{V'} = (f_1, \ldots, f_n)$ be the basis of V' which is dual to $\mathcal{B}_V$. We claim that $T'(g_i) = f_i$ from which the result will follow.

Suppose $j > k$. Then $\boldsymbol{v}_j \in Ker(T)$ and $[T'(g_i)](\boldsymbol{v}_j) = 0$. Suppose $j \leq k, j \neq i$. Then $[T'(g_i)](\boldsymbol{v}_j) = g_i(T(\boldsymbol{v}_j)) = g_i(\boldsymbol{w}_j) = 0$. Finally, $[T'(g_i)](\boldsymbol{v}_i) = g_i(T(\boldsymbol{v}_i)) = g_i(\boldsymbol{w}_i) = 1$.

6. Assume $dim(V) = n, dim(W) = m$ and $rank(T) = k$. By the rank-nullity theorem, $nullity(T) = n - k$. By Exercise 5, $rank(T') = k$. Again by the rank-nullity theorem, $nullity(T') = m - k$. Thus, $nullity(T) = nullity(T')$ if and only if $n - k = m - k$ if and only if $n = m$.

7. Let $\boldsymbol{v} \neq \boldsymbol{0}_V$ and $(\boldsymbol{v} = \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis for V and $(g_1, \ldots, g_n)$ the basis of V' dual to $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. Then $g_1(\boldsymbol{v}) = 1$. Since $(f_1, \ldots, f_n)$ is a basis for V' there exists scalars $c_1, \ldots, c_n$ such that

$$g = c_1 f_1 + \cdots + c_n f_n.$$

Then $1 = g(\boldsymbol{v}) = c_1 f_1(\boldsymbol{v}) + \ldots c_n f_n(\boldsymbol{v})$. This implies for some $i, f_i(\boldsymbol{v}) \neq 0$. It then follows that $T(\boldsymbol{v}) \neq \boldsymbol{0}_n$. Thus, T is injective whence an isomorphism by the half is good enough theorem.

8. Since T is an isomorphism by Exercise 4, T' is an isomorphism. Since $(\pi_1, \ldots, \pi_n)$ is a basis for $(\mathbb{F}^n)'$ we can conclude that $(T'(\pi_1), \ldots, T'(\pi_n)) = (f_1, \ldots, f_n)$ is a basis for V'.

9. We give an indirect proof. We first establish the existence of a natural isomorphism between V and $(V')'$.

Thus, let $\boldsymbol{v} \in V$. Define a map $F_{\boldsymbol{v}} : V' \to \mathbb{F}$ by $F_{\boldsymbol{v}}(f) = f(\boldsymbol{v})$. Claim that the map $\boldsymbol{v} \to F_{\boldsymbol{v}}$ is a linear transformation.

Let $\boldsymbol{v}, \boldsymbol{w} \in V$. We need to prove that $F_{\boldsymbol{v}+\boldsymbol{w}} = F_{\boldsymbol{v}} + F_{\boldsymbol{w}}$. Let $f \in V'$. Then $F_{\boldsymbol{v}+\boldsymbol{w}}(f) = f(\boldsymbol{v} + \boldsymbol{w}) = f(\boldsymbol{v}) + f(\boldsymbol{w}) = F_{\boldsymbol{v}}(f) + F_{\boldsymbol{w}}(f) = (F_{\boldsymbol{v}} + F_{\boldsymbol{w}})(f)$.

Now let $\boldsymbol{v} \in V, c \in \mathbb{F}$. We need to prove that $F_{c\boldsymbol{v}} = cF_{\boldsymbol{v}}$. Now for $f \in V'$ we have $F_{c\boldsymbol{v}}(f) = f(c\boldsymbol{v}) = cf(\boldsymbol{v}) = cF_{\boldsymbol{v}}(f) = (cF_{\boldsymbol{v}})(f)$.

Now let $\boldsymbol{v} \in V$ be non-zero. We have oftentimes seen that there exists $f \in V'$ such that $f(\boldsymbol{v}) \neq 0$. Then $F_{\boldsymbol{v}}(f) = f(\boldsymbol{v}) \neq 0$. Therefore the linear map $\boldsymbol{v} \to F_{\boldsymbol{v}}$ has trivial kernel and is injective. Since $dim(V) = dim((V')$, in fact, this map is an isomorphism.

Now let $(F_1, \ldots, F_n)$ be the basis of $(V')'$ which is dual to $(f_1, \ldots, f_n)$ and let $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ be the basis of V such that $F_{\boldsymbol{x}_i} = F_i$. This sequence satisfies the requirements of the exercise.

10. Suppose $f, g \in U', c \in \mathbb{F}$ and $\boldsymbol{u} \in U$. Then $(f + g)(\boldsymbol{u}) = f(\boldsymbol{u}) + g(\boldsymbol{u}) = 0 + 0 = 0$. Since $\boldsymbol{u} \in U$ is arbitrary, $U \subset Ker(f + g)$.

$(cf)(\boldsymbol{u}) = c[f(\boldsymbol{u})] = c \times 0 = 0$. Thus, $U \subset Ker(cf)$ and $U^\perp$ is a subspace of V'.

Let $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ be a basis for U and extend to a basis $\mathcal{B} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ for V. Let $(g_1, \ldots, g_n)$ be the basis of V' which is dual to $\mathcal{B}$. Then $U^\perp = Span(g_{k+1}, \ldots, g_n)$ and has dimension $n - k$.

11. Since $U, W \subset U + W$ it follows that if $f \in (U + W)'$ then $f \in U' \cap W'$. On the other hand, suppose $f \in U' \cap W'$ and $\boldsymbol{v} \in U + W$. Then there are $\boldsymbol{u} \in U, \boldsymbol{w} \in W$ such that $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$. Then $f(\boldsymbol{u} + \boldsymbol{w}) = f(\boldsymbol{u}) + f(\boldsymbol{w}) = 0 + 0 = 0$ and $f \in (U + W)'$. Thus, we have equality.

Now suppose $f \in U', g \in W'$ and $\boldsymbol{v} \in U \cap W$. Then $(f + g)(\boldsymbol{v}) = f(\boldsymbol{v}) + g(\boldsymbol{v}) = 0 + 0 = 0$. Thus, $f + g \in (U \cap W)'$. This shows that $U' + W' \subset (U \cap W)'$. We complete this with a dimension argument. Let $dim(U) = k, dim(W) = l$ and $dim(U \cap W) = j$. Then $dim(U + W) = k + l - j$. From Exercise 10 it follows that $dim((U + W)^\perp) = n - k - l + j$. By the first part, $U' \cap W' = (U + W)'$. Thus, $dim(U' \cap W') = n - k - l + j$.

Again by Exercise 10, $dim(U') = n - k, dim(W') = n - l$. Then

$$dim(U' + W') = dim(U') + dim(W') - dim(U' \cap W') =$$

$$(n-k) + (n-l) - (n-k-l+j) = n - j = dim((U \cap W)').$$

12. Clearly π is linear and injective. Since $dim((U \oplus W)') = dim(U \oplus W) = dim(U) + dim(W) = dim(U') + dim(W') = dim(U' \oplus W')$, π is an isomorphism.

13. Let $f \in \mathcal{L}(X, \mathbb{F})$. Then $(S \circ T)'(f) = (S \circ T) \circ f = S \circ (T \circ f) = S'(T'(f)) = (S' \circ T')(f)$.

14. Let $f \in U'$ and $\boldsymbol{u} \in U$. Then $[T'(f)](\boldsymbol{u}) = f(T(\boldsymbol{u}))$. Since U is $T-$invariant and $\boldsymbol{u} \in U, T(\boldsymbol{u}) \in U$. Since $f \in U'$ we conclude that $f(T(\boldsymbol{u})) = 0$. Since $\boldsymbol{u}$ is arbitrary, $T'(f) \in U'$.

15. Let $\mathcal{B} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis for V and $\mathcal{B}' = (f_1, \ldots, f_n)$ be the basis of V' dual to $\mathcal{B}$. Then $\mathcal{M}_{T'}(\mathcal{B}', \mathcal{B}') = \mathcal{M}_T(\mathcal{B}, \mathcal{B})^{tr}$. It follows from this that the operators $T \in \mathcal{L}(V, V)$ and $T' \in \mathcal{L}(V', V')$ have the same minimum polynomial (and elementary divisors and invariant factors).

16i) Suppose $g \in Ker(T')$ and $\boldsymbol{w} = T(\boldsymbol{v})$. Then $g(\boldsymbol{w}) = g(T(\boldsymbol{v})) = [T'(g)[\boldsymbol{v}] = 0$. Since $\boldsymbol{w}$ is arbitrary, $g \in Range(T)'$ and $Ker(T') \subset Range(T)'$. On the other hand, suppose $g \in Range(T)'$. Let $\boldsymbol{v} \in V$. Then $[T'(g)](\boldsymbol{v}) = g(T(\boldsymbol{v}))$. Since $T(\boldsymbol{v}) \in Range(T)$ and $g \in Range(T)'$ we have $g(T(\boldsymbol{v})) = 0$. Since $\boldsymbol{v} \in V$ is arbitrary, $T'(g)$ is the zero vector in V' and $g \in Ker(T')$. Thus, we have equality.

ii) Assume $f = T'(g) \in Range(T')$ and $\boldsymbol{v} \in Ker(T)$. Then $f(\boldsymbol{v}) = [T'(g)](\boldsymbol{v}) = g(T(\boldsymbol{v})) = g(\boldsymbol{0}_W) = 0$. Thus, $f \in Ker(T)'$. Now we can complete this with dimension arguments: By Exercise 5, $dim(Range(T')) = dim(Range(T))$. By Exercise 10, $dim(Ker(T)') = dim(V) - dim(Ker(T)) = dim(Range(T))$.

iii) Let $\boldsymbol{v} \in Ker(T)$ and $f = T'(g)$. Then $f(\boldsymbol{v}) = [T'(g)](\boldsymbol{v}) = g(T(\boldsymbol{v})) = g(\boldsymbol{0}_W) = 0$. Since f and $\boldsymbol{v}$ are arbitrary we have $Ker(T) \subset Range(T')'$. Once again we get equality by a dimension argument.

iv) Let $\boldsymbol{w} = T(\boldsymbol{v})$ and $g \in Ker(T')$. Then $g(\boldsymbol{w}) = g(T(\boldsymbol{v})) = [T'(g)](\boldsymbol{v}) = 0$ since $T'(g)$ is the zero vector in V'. This implies that $Range(T) \subset Ker(T')'$. Equality follows from dimension arguments.

5.6. Adjoints

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

2. $-420x^2 + 396x - 60$.

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

4. Let $\boldsymbol{v} \in V$ and $\boldsymbol{x} \in X$. Then

$$\langle (TS)(\boldsymbol{v}), \boldsymbol{x} \rangle_X = \langle \boldsymbol{v}, (TS)^*(\boldsymbol{x}) \rangle_V$$

On the other hand,

$$\langle (TS)(\boldsymbol{v}), \boldsymbol{x} \rangle_X = \langle T(S(\boldsymbol{v})), \boldsymbol{x} \rangle_X =$$

$$\langle S(\boldsymbol{v}), T^*(\boldsymbol{x}) \rangle_W = \langle \boldsymbol{v}, S^*(T^*(\boldsymbol{x})) \rangle_V$$

Thus, for all $\boldsymbol{v} \in V$ and $\boldsymbol{x} \in X$ we have

$$\langle \boldsymbol{v}, (TS)^*(\boldsymbol{x}) \rangle_V = \langle \boldsymbol{v}, (S^*T^*)(\boldsymbol{x}) \rangle_V$$

This implies that $(TS)^*(\boldsymbol{x}) = (S^*T^*)(\boldsymbol{x})$ for all $\boldsymbol{x} \in X$ whence $(TS)^* = S^*T^*$.

5. Let $\boldsymbol{v} \in V, \boldsymbol{w} \in W$. Then

$$\langle T(\boldsymbol{v}), \boldsymbol{w} \rangle_V = \langle \boldsymbol{v}, T^*(\boldsymbol{w}) \rangle_V =$$

$$\langle (T^*)^*(\boldsymbol{v}), \boldsymbol{w} \rangle_W.$$

Since this holds for all $\boldsymbol{w} \in W$ we must have $(T^*)^*(\boldsymbol{v}) = T(\boldsymbol{v})$ for all $\boldsymbol{v} \in V$ whence $(T^*)^* = T$.

6. Assume $T(\boldsymbol{u}) = \lambda \boldsymbol{u}$. Then for all $\boldsymbol{v} \in V$ we have

$$\langle (T - \lambda)(\boldsymbol{u}), \boldsymbol{v} \rangle = 0.$$

This implies that

$$\langle \boldsymbol{u}, (T - \lambda I_V)^*(\boldsymbol{v}) \rangle = \langle \boldsymbol{u}, (T^* - \overline{\lambda} I_V)(\boldsymbol{v}) \rangle = 0.$$

This implies that $Range(T - \overline{\lambda} I_V) \subset \boldsymbol{u}^\perp$ is a proper subspace of V. It must then be the case that $Ker(T - \overline{\lambda} I_V) \neq \{\boldsymbol{0}\}$ which implies that there exists an eigenvector with eigenvalue $\overline{\lambda}$.

7. Let $\boldsymbol{w} \in W$ and assume that $T^*(\boldsymbol{w}) = \boldsymbol{0}_V$. Then for all $\boldsymbol{v} \in V$

$$\langle \boldsymbol{v}, T^*(\boldsymbol{w})\rangle_V = 0.$$

Using the definition of T^* we have for all $\boldsymbol{v} \in V$ that

$$\langle T(\boldsymbol{v}), \boldsymbol{w}\rangle_W = 0$$

Since T is invertible, in particular, $Range(T) = W$. This implies that $\boldsymbol{w} = \boldsymbol{0}_W$ and consequently, T^* is injective, whence invertible. Exercise 3 applied to $TT^{-1} = I_V = T^{-1}T$ yields $(T^*)^{-1} = (T^{-1})^*$.

8. Suppose $(T^*T)(\boldsymbol{v}) = \boldsymbol{0}_V$. It then follows that by the definition of T^* that

$$\langle (T(\boldsymbol{v}), T(\boldsymbol{v})\rangle_W = \langle \boldsymbol{v}, (T^*T)(\boldsymbol{v})\rangle_V = 0.$$

Since $\langle\ ,\ \rangle_W$ is positive definite we can conclude that $T(\boldsymbol{v}) = \boldsymbol{0}_W$. However, T injective implies that $\boldsymbol{v} = \boldsymbol{0}_V$. Thus, T^*T is injective. Since V is finite dimensional it follows that T^*T is bijective.

9. It follows from part i) of Theorem (5.22), if $T : V \to W$ is surjective then T^* is injective. By Exercise 5, $(T^*)^* = T$. Now it follows from Exercise 8 that $TT^* : W \to W$ is bijective.

10. Let $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in U^\perp$. Since U is $T-$invariant we have

$$\langle T(\boldsymbol{u}), \boldsymbol{w}\rangle = 0.$$

Making use of the definition of T^* we then get for all $\boldsymbol{u} \in U$

$$\langle \boldsymbol{u}, T^*(\boldsymbol{w})\rangle = 0.$$

This implies that $T^*(\boldsymbol{w}) \in U^\perp$.

11. Suppose $(T^*T)(\boldsymbol{v}) = \boldsymbol{0}$. Then

$$\langle T(\boldsymbol{v}), T(\boldsymbol{v})\rangle = \langle \boldsymbol{v}, T^*(T(\boldsymbol{v}))\rangle = 0.$$

By positive definiteness we have $T(\boldsymbol{v}) = \boldsymbol{0}$.

12. $S^*(\boldsymbol{x}, \boldsymbol{y}) = (-\boldsymbol{y}, \boldsymbol{x})$.

13. Let $\mathcal{B}_V$ be an orthonormal basis of V and $\mathcal{B}_W$ be an orthonormal basis of W. Set $A = \mathcal{M}_T(\mathcal{B}_V, \mathcal{B}_W)$ and $A^* = \mathcal{M}_{T^*}(\mathcal{B}_W, \mathcal{B}_V)$. By Theorem (5.23), $A^* = \overline{A}^{tr}$. Since $rank(T) = rank(A), rank(T^*) = rank(A^*)$ and $rank(A) = rank(\overline{A}^{tr})$ it follows that $rank(T) = rank(T^*)$.

14. Assume S exists. Then $1 = \langle \boldsymbol{v}_1, \boldsymbol{v}_1\rangle = \langle \boldsymbol{v}_1, S^*(\boldsymbol{y})\rangle = \langle S(\boldsymbol{v}_1), \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$. This proves that $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 1$. Conversely, assume $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 1$. Let $(\boldsymbol{x}_2, \dots, \boldsymbol{x}_n)$ be a basis for $\boldsymbol{y}^\perp$. Since $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 1 \neq 0$, $\boldsymbol{x} \notin \boldsymbol{y}^\perp$ so that $(\boldsymbol{x}, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n)$ is linearly independent. Set $\boldsymbol{x}_1 = \boldsymbol{x}$ and define $S : V \to V$ so that $S(\boldsymbol{v}_i) = \boldsymbol{x}_i$. Then S is invertible and $S(\boldsymbol{v}_1) = \boldsymbol{x}_1 = \boldsymbol{x}$. It remains to show that $S^*(\boldsymbol{y}) = \boldsymbol{v}_1$. Let $2 \leq j \leq n$. Then $0 = \langle S(\boldsymbol{v}_j), \boldsymbol{y}\rangle = \langle \boldsymbol{v}_j, S^*(\boldsymbol{y})\rangle$. Consequently, $S^*(\boldsymbol{y}) = Span(\boldsymbol{v}_2, \dots, \boldsymbol{v}_n)^\perp = Span(\boldsymbol{v}_1)^\perp$. There is then a scalar α such that $S^*(\boldsymbol{y}) = \alpha\boldsymbol{v}_1$. However, $1 = \langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle S(\boldsymbol{v}_1), \boldsymbol{y}\rangle = \langle \boldsymbol{v}_1, S^*(\boldsymbol{y})\rangle = \langle \boldsymbol{v}_1, \alpha\boldsymbol{v}_1\rangle = \overline{\alpha}$. Thus, $\alpha = 1$ as required.

5.7. Normed Vector Spaces

1. a) $\left\| \begin{pmatrix} -4 \\ 2 \\ -1 \\ -2 \end{pmatrix} \right\|_1 = 9$, $\left\| \begin{pmatrix} -4 \\ 2 \\ -1 \\ -2 \end{pmatrix} \right\|_2 = 5$, $\left\| \begin{pmatrix} -4 \\ 2 \\ -1 \\ -2 \end{pmatrix} \right\|_\infty = 4$.

b) $\left\| \begin{pmatrix} 3 \\ -6 \\ 0 \\ 2 \end{pmatrix} \right\|_1 = 11$, $\left\| \begin{pmatrix} 3 \\ -6 \\ 0 \\ 2 \end{pmatrix} \right\|_2 = 7$, $\left\| \begin{pmatrix} 3 \\ -6 \\ 0 \\ 2 \end{pmatrix} \right\|_\infty = 6$.

2. Let $\boldsymbol{x} = \begin{pmatrix} -4 \\ 2 \\ -1 \\ 2 \end{pmatrix}$, $\boldsymbol{y} = \begin{pmatrix} 3 \\ -6 \\ 0 \\ 2 \end{pmatrix}$. Then $\boldsymbol{x} - \boldsymbol{y} = \begin{pmatrix} -7 \\ 8 \\ -1 \\ 0 \end{pmatrix}$.
Then $d_1(\boldsymbol{x}, \boldsymbol{y}) = \| \boldsymbol{x} - \boldsymbol{y} \|_1 = 16$, $d_2(\boldsymbol{x}, \boldsymbol{y}) = \| \boldsymbol{x} - \boldsymbol{y} \|_2 = \sqrt{114}$, $d_\infty(\boldsymbol{x}, \boldsymbol{y}) = \| \boldsymbol{x} - \boldsymbol{y} \|_\infty = 8$.

3. 1

4. 1) If $\boldsymbol{x} = \boldsymbol{y}$ then $d(\boldsymbol{x},\boldsymbol{y}) = \parallel \boldsymbol{0} \parallel = 0$. On the other hand, if $d(\boldsymbol{x},\boldsymbol{y}) = 0$ then $\parallel \boldsymbol{x}-\boldsymbol{y} \parallel = 0$, whence $\boldsymbol{x}-\boldsymbol{y} = \boldsymbol{0}$ so $\boldsymbol{x} = \boldsymbol{y}$.

2) This holds since $d(\boldsymbol{y},\boldsymbol{x}) = \parallel \boldsymbol{y}-\boldsymbol{x} \parallel = \parallel (-1)(\boldsymbol{x}-\boldsymbol{y}) \parallel = |-1| \parallel \boldsymbol{x}-\boldsymbol{y} \parallel = \parallel \boldsymbol{x}-\boldsymbol{y} \parallel = d(\boldsymbol{x},\boldsymbol{y})$.

3) $d(\boldsymbol{x},\boldsymbol{z}) = \parallel \boldsymbol{x}-\boldsymbol{z} \parallel = \parallel (\boldsymbol{x}-\boldsymbol{y})+(\boldsymbol{y}-\boldsymbol{z}) \parallel \leq \parallel \boldsymbol{x}-\boldsymbol{y} \parallel + \parallel \boldsymbol{y}-\boldsymbol{z} \parallel = d(\boldsymbol{x},\boldsymbol{y}) + d(\boldsymbol{y},\boldsymbol{z})$.

5. Let $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. If $\boldsymbol{x} = \boldsymbol{0}$ then clearly $\parallel \boldsymbol{x} \parallel_\infty = max\{|x_1|,\dots,|x_n|\} = 0$. On the other hand if $\parallel \boldsymbol{x} \parallel_\infty = max\{|x_1|,\dots,|x_n|\} = 0$ then $x_i = 0$ for all i and $\boldsymbol{x} = \boldsymbol{0}$.

Let c be a scalar. Then $\parallel c\boldsymbol{x} \parallel_\infty = max\{|cx_1|,\dots,|cx_n|\} = max\{|c||x_1|,\dots,|c||x_n|\} = |c|max\{|x_1|,\dots,|x_n|\} = |c| \parallel \boldsymbol{x} \parallel_\infty$.

Now $\parallel \boldsymbol{x}+\boldsymbol{y} \parallel_\infty = max\{|x_1+y_1|,\dots,|x_n+y_n|\}$. Now $|x_i+y_i| \leq |x_i|+|y_i| \leq \parallel \boldsymbol{x} \parallel_\infty + \parallel \boldsymbol{y} \parallel -\infty$. Consequently, $max\{|x_1+y_1|,\dots,|x_n+y_n|\} \leq \parallel \boldsymbol{x} \parallel_\infty + \parallel \boldsymbol{y} \parallel_\infty$.

6. Let $B_r^2(\boldsymbol{x})$ denote the open ball centered at $\boldsymbol{x}$ with radius r with respect to the l_2-norm and $B_r^\infty(\boldsymbol{x})$ the open ball centered at $\boldsymbol{x}$ with radius r with respect to the l_∞-norm. We must show the following: i) For an arbitrary point $\boldsymbol{y} \in B_r^2(\boldsymbol{0})$ there is a positive s such that $B_s^\infty(\boldsymbol{y}) \subset B_r^(\boldsymbol{0})$; and ii) For an arbitrary point $\boldsymbol{z} \in B_r^\infty(\boldsymbol{0})$ there is a positive t such that $B_t^2(\boldsymbol{z}) \subset B_r^\infty(\boldsymbol{0})$.

i) Set $s = \frac{r-\|\boldsymbol{y}\|_2}{2\sqrt{2}}$. Suppose $\boldsymbol{u} \in B_s^\infty(\boldsymbol{y})$. Then $\parallel \boldsymbol{u} \parallel_2 = \parallel \boldsymbol{y}+(\boldsymbol{u}-\boldsymbol{y}) \parallel_2 \leq \parallel \boldsymbol{y} \parallel_2 + \parallel \boldsymbol{u}-\boldsymbol{y} \parallel_2$. If $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ then $\parallel \boldsymbol{u}-\boldsymbol{y} \parallel_2^2 = (u_1-y_1)^2 + (u_2-y_2)^2 < 2s^2$

$$= \frac{(r- \parallel \boldsymbol{y} \parallel)^2}{4}$$

Consequently, $\parallel \boldsymbol{u}-\boldsymbol{y} \parallel \leq \frac{r-\|\boldsymbol{y}\|_2}{2}$. Then, $\parallel \boldsymbol{u} \parallel_2 \leq \parallel \boldsymbol{y} \parallel_2 + \frac{r-\|\boldsymbol{y}\|_2}{2} = \frac{r+\|\boldsymbol{y}\|_2}{2} < r$.

ii) Assume $\boldsymbol{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in B_r^\infty(\boldsymbol{0})$. Set $t = \frac{r-\|\boldsymbol{z}\|_\infty}{2}$ and assume $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in B_t^2(\boldsymbol{z})$. We need to prove that $\boldsymbol{v} \in B_r^\infty(\boldsymbol{0})$. Now $\parallel \boldsymbol{v} \parallel_\infty = \parallel \boldsymbol{z}+(\boldsymbol{v}-\boldsymbol{z}) \parallel_\infty \leq \parallel \boldsymbol{z} \parallel_\infty + \parallel \boldsymbol{v}-\boldsymbol{z} \parallel_\infty$. Note that $\parallel \boldsymbol{v}-\boldsymbol{z} \parallel_\infty \leq \parallel \boldsymbol{v}-\boldsymbol{z} \parallel_2$. Therefore, $\parallel \boldsymbol{v} \parallel_\infty \leq \parallel \boldsymbol{z} \parallel_\infty + \frac{r-\|\boldsymbol{z}\|_\infty}{2} = \frac{r+\|\boldsymbol{z}\|_\infty}{2} < r$.

7. Let $\{\boldsymbol{x}_k\}_{k=1}^\infty$ be a Cauchy sequence with respect to the l_1-norm. Assume $\boldsymbol{x}_k = \begin{pmatrix} x_{1k} \\ \vdots \\ x_{nk} \end{pmatrix}$. We claim for each $i, 1 \leq i \leq n$ that $\{x_{ik}\}_{k=1}^\infty$ is a Cauchy sequence in $\mathbb{R}$ of $\mathbb{C}$. Thus, let ϵ be a positive real number. Since $\{\boldsymbol{x}_k\}_{k=1}^\infty$ is a Cauchy sequence with respect to the l_1-norm there exists a natural number $N = N(\epsilon)$ such that if $p,q \geq N$ then $\parallel \boldsymbol{x}_q - \boldsymbol{x}_p \parallel_1 < \epsilon$. Since $\parallel \boldsymbol{x}_q - \boldsymbol{x}_p \parallel_1 = \sum_{i=1}^n |x_{iq} - x_{ip}|$ it follows for each $i, 1 \leq i \leq n$ and $p,q \geq N$ that $|x_{iq} - x_{ip}| < \epsilon$ as required. Since $\mathbb{R}$ and $\mathbb{C}$ are complete we can conclude for each i the sequence $\{x_{ik}\}_{k=1}^\infty$ converges. Set $x_i = \lim_{k\to\infty} x_{ik}$ and set $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. We claim that $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$. Toward that objective, assume ϵ is a positive real number. Since $\lim_{k\to\infty} x_{ik} = x_i$ there is a natural number M_i such that if $k \geq M_i$ then $|x_i - x_{ik}| < \frac{\epsilon}{n}$. Now set $M = max\{M_1,\dots,M_n\}$ and suppose $k \geq M$. Then $k \geq M_i$ so that $|x_i - x_{ik}| < \frac{\epsilon}{n}$ and therefore $\parallel \boldsymbol{x} - \boldsymbol{x}_k \parallel_1 = \sum_{i=1}^n |x_i - x_{ik}| < n \times \frac{\epsilon}{n} = \epsilon$ so that $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$ as claimed.

8. Let $\{\boldsymbol{x}_k\}_{k=1}^\infty$ be a Cauchy sequence with respect to the l_∞-norm. Assume $\boldsymbol{x}_k = \begin{pmatrix} x_{1k} \\ \vdots \\ x_{nk} \end{pmatrix}$. We claim that for each $i, 1 \leq i \leq n$ that $\{x_{ik}\}_{k=1}^\infty$ is a Cauchy sequence in $\mathbb{R}$ or $\mathbb{C}$. Thus, let ϵ be a positive real number. Since $\{\boldsymbol{x}_k\}_{k=1}^\infty$ is a Cauchy sequence with respect to the l_∞-norm there exists a natural number $N = N(\epsilon)$ such that if $p,q \geq N$ then $\parallel \boldsymbol{x}_q - \boldsymbol{x}_p \parallel_\infty < \epsilon$. Since $\parallel \boldsymbol{x}_q - \boldsymbol{x}_p \parallel_\infty = max\{|x_{1q} - x_{1p},\dots,|x_{nq} - x_{np}|\}$ we can conclude for each $i, 1 \leq i \leq n$ and $p,q \geq N$ that $|x_{iq} - x_{ip}| < \epsilon$

establishing our claim. Since $\mathbb{R}$ and $\mathbb{C}$ are complete we can conclude for each i the sequence $\{x_{ik}\}_{k=1}^{\infty}$ converges. Set $x_i = \lim_{k\to\infty} x_{ik}$ and set $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. We claim that $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$. Toward that objective, assume ϵ is a positive real number. Since $\lim_{k\to\infty} x_{ik} = x_i$ there is a natural number M_i such that if $k \geq M_i$ then $|x_i - x_{ik}| < \epsilon$. Now set $M = max\{M_1, \ldots, M_n\}$ and assume that $k \geq M$. Then $k \geq M_i$ so that $|x_i - x_{ik}| < \epsilon$. Since this is true for each $i, 1 \leq i \leq n$ it follows that $\| \boldsymbol{x} - \boldsymbol{x}_k \|_{\infty} = max\{|x_1 - x_{1k}|, \ldots |x_n - x_{nk}|\} < \epsilon$ so that, indeed, $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}$.

9. i) The relation fo equivalence is reflexive: Take $c = d = 1$.

ii) The relation of equivalence is symmetric: Assume $\| \cdot \|$ is equivalent to $\| \cdot \|'$ and let c, d be positive real numbers such that for every $\boldsymbol{x}, c \| \boldsymbol{x} \|' \leq \| \boldsymbol{x} \| \leq d \| \boldsymbol{x} \|'$. Then $\frac{1}{d} \| \boldsymbol{x} \| \leq \| \boldsymbol{x} \|' \leq \frac{1}{c} \| \boldsymbol{x} \|$ for every $\boldsymbol{x}$.

iii) The relation of equivalence is transitive: Assume $\| \cdot \|$ is equivalent to $\| \cdot \|'$ and $\| \cdot \|'$ is equivalent to $\| \cdot \|^*$. Let a, b be positive real numbers such that for every $\boldsymbol{x}$, $a \| \boldsymbol{x} \|' \leq \| \boldsymbol{x} \| \leq b \| \boldsymbol{x} \|'$ and let c, d be positive real numbers such that for every $\boldsymbol{x}$ we have $c \| \boldsymbol{x} \|^* \leq \| \boldsymbol{x} \|' \leq d \| \boldsymbol{x} \|^*$.

Set $e = ac$ and $f = bd$. Then for every $\boldsymbol{x}, e \| \boldsymbol{x} \|^* \leq \| \boldsymbol{x} \| \leq f \| \boldsymbol{x} \|^*$.

10. $\| \boldsymbol{e}_1 + \boldsymbol{e}_2 \|_p^2 = \| \boldsymbol{e}_1 - \boldsymbol{e}_2 \|_p^2 = 2^{\frac{2}{p}}$. Therefore

$$\| \boldsymbol{e}_1 + \boldsymbol{e}_2 \|_p^2 + \| \boldsymbol{e}_1 - \boldsymbol{e}_2 \|_p^2 = 2 \times 2^{\frac{2}{p}}.$$

On the other hand, $2(\| \boldsymbol{e}_1 \|_p^2 + \| \boldsymbol{e}_2 \|_p^2) = 4$.

If $p = 2$ then $2 \times 2^{\frac{2}{p}} = 2 \times 2 = 4$ and we have equality. Conversely, assume we have equality;

$$2 \times 2^{\frac{2}{p}} = 4$$

so that $2^{\frac{2}{p}} = 2$ so that $\frac{2}{p} = 1$ and $p = 2$, as asserted.

11. Let $m = max\{\| \boldsymbol{e}_1 \|, \ldots, \| \boldsymbol{e}_n \|\}$ and assume $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is in S_1^{∞}. Then $\| \boldsymbol{x} \| \leq \sum_{i=1}^n \| x_i \boldsymbol{e}_i \|$ by the triangle inequality. $\sum_{i=1}^n \| x_i \boldsymbol{e}_i \| = \sum_{i=1}^n |x_i| \| \boldsymbol{e}_i \| \leq \sum_{i=1}^n m|x_i| \leq nm$ since $|x_i| \leq 1$. Thus, S_1^{∞} is bounded. It remains to show that it is S_1^{∞} is closed.

Chapter 6

Linear Operators on Inner Product Spaces

6.1. Self-Adjoint Operators

1. This follows immediately since $(S+T)^* = S^*+T^* = S+T$.

2. $(\gamma T)^* = \overline{\gamma} T^* = \gamma T$ since $\gamma \in \mathbb{R}$ and T is self-adjoint.

3i) $R^* = [\frac{1}{2}(T+T^*)]^* = \frac{1}{2}[T^*+(T^*)^*] = \frac{1}{2}[T^*+T] = R$.

$S^* = (\frac{1}{2}i[-T+T^*])^* = \frac{1}{2}\overline{i}[-T^*+(T^*)^*] = -\frac{1}{2}i[-T^*+T] = \frac{1}{2}i[-T+T^*] = S$.

ii) $R+iS = \frac{1}{2}[T+T^*] + \frac{1}{2}i^2[-T+T^*] = \frac{1}{2}[T+T^*] - \frac{1}{2}[-T+T^*] = \frac{1}{2}T + \frac{1}{2}T^* + \frac{1}{2}T - \frac{1}{2}T^* = T$.

iii) If $T = R_1 + iS_1$ is self-adjoint then $T^* = R_1^*, +(iS_1)^* = R_1^* + \overline{i}S_1^* = R_1 - iS_1$.

Then $T+T^* = 2R_1, R_1 = \frac{1}{2}[T+T^*] = R$. Then $S_1 = T - R_1 = T - R = S$.

4. $T^* = R - iS$. Suppose $RS = SR$. Then

$$TT^* = (R+iS)(R-iS) =$$

$$R^2 + (iS)R - R(iS) + S^2 = R^2 + S^2 =$$

$$(R-iS)(R+iS) = T^*T$$

and so T is normal.

Conversely, assume that T is normal so that $TT^* = T^*T$. Then

$$TT^* = R^2 + i[SR - RS] + S^2 =$$

$$T^*T = R^2 + i[RS - SR] + S^2.$$

It follows that $SR - RS = RS - SR$ from which we conclude that $2SR = 2RS, SR = RS$.

5. The dimension is n^2 as a real vector space.

6. Assume $(ST)^* = ST$. $(ST)^* = T^*S^* = TS$. Thus, $TS = ST$. On the other hand, assume $ST = TS$. Then $(ST)^* = T^*S^* = TS = ST$.

7. Let $\mathcal{B} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be an orthonormal basis for V. Let $S(\boldsymbol{v}_j) = T(\boldsymbol{v}_j) = \boldsymbol{v}_j$ for $3 \leq j \leq n$. Set $S(\boldsymbol{v}_1) = \boldsymbol{v}_2, S(\boldsymbol{v}_2) = \boldsymbol{v}_1; T(\boldsymbol{v}_1) = \boldsymbol{v}_1 + 2\boldsymbol{v}_2, T(\boldsymbol{v}_2) = 2\boldsymbol{v}_1 + 3\boldsymbol{v}_2$.

Since $\mathcal{M}_S(\mathcal{B},\mathcal{B}), \mathcal{M}_T(\mathcal{B},\mathcal{B})$ are real symmetric the operators S and T are self-adjoint by Theorem (6.1). However, $\mathcal{M}_{ST}(\mathcal{B},\mathcal{B}) \neq \mathcal{M}_{TS}(\mathcal{B},\mathcal{B})$.

8. $\parallel T(\boldsymbol{v}) \parallel^2 = \langle T(\boldsymbol{v}), T(\boldsymbol{v})\rangle = \langle \boldsymbol{v}, T^*T(\boldsymbol{v})\rangle$. Since T is normal, $\langle \boldsymbol{v}, T^*T(\boldsymbol{v})\rangle = \langle \boldsymbol{v}, TT^*(\boldsymbol{v})\rangle = \langle T^*(\boldsymbol{v}), T^*(\boldsymbol{v})\rangle = \parallel T^*(\boldsymbol{v}) \parallel^2$.

9. This follows from Exercise 8.

10. By Exercise 9 we have $Ker(T) = Ker(T^*)$. Then $Range(T^*) = Ker(T)^\perp = Ker(T^*)^\perp = Range(T)$ by Theorem (5.22).

11. If $TT^* = T^2$ then for $\boldsymbol{v} \in V, \langle \boldsymbol{v}, TT^*(\boldsymbol{v})\rangle = \langle \boldsymbol{v}, T^2(\boldsymbol{v})\rangle$. This implies that $\langle T^*(\boldsymbol{v}), T^*(\boldsymbol{v})\rangle = \langle T^*(\boldsymbol{v}), T(\boldsymbol{v})\rangle$ from which we conclude that

$\langle T^*(\boldsymbol{v}), T^*(\boldsymbol{v}) - T(\boldsymbol{v})\rangle = 0$ for all $\boldsymbol{v} \in V$. Suppose $T(\boldsymbol{v}) = \mathbf{0}$. Then $\langle T^*(\boldsymbol{v}), T^*(\boldsymbol{v})\rangle = 0$ so that by positive definiteness, $T^*(\boldsymbol{v}) = \mathbf{0}$. Thus, $Ker(T) \subset Ker(T^*)$. However, by Exercise 13 of Section (5.6), $rank(T) = rank(T^*)$. Then $nullity(T) = nullity(T^*)$. This implies that $Ker(T) = Ker(T^*)$.

Set $U = Ker(T) = Ker(T^*)$. By Theorem (5.22), $Range(T) = Ker(T^*)^\perp = Ker(T)^\perp = Range(T^*)$. Set $W = Range(T)$. Note that $U \cap W = U \cap U^\perp = \{\mathbf{0}\}$. Since $dim(U) + dim(U^\perp) = dim(V)$ we have $V = U \oplus W$. Let S be T restricted to $W = Range(T) = Range(T^*)$. Note for $\boldsymbol{u}, \boldsymbol{w} \in W, \langle S(\boldsymbol{u}), \boldsymbol{w}\rangle = \langle T(\boldsymbol{u}), \boldsymbol{w}\rangle = \langle \boldsymbol{u}, T^*(\boldsymbol{w})\rangle = \langle \boldsymbol{u}, S^*(\boldsymbol{w})\rangle$. Since $T^*(\boldsymbol{w}) \in W$ it must be the case that S^* is the restriction of T^* to W. It now follows that $S^2 = SS^*$. Since S is invertible, $S = S^*$. We have therefore shown that T restricted to $W = Range(T)$ is equal to T^* restricted to W. However, T restricted to $U = Ker(T)$ is equal to T^* restricted to U (both are the zero map). We can now conclude $T = T^*$ and T is self-adjoint.

12. Let $U = Ker(T)$ and assume that U is proper in V. By Exercise 9, $Ker(T^*). = U$ Since U is T^*-invariant by Theorem (5.22) it follows that $U^\perp$ is $T-$invariant. Since T is nilpotent and T leaves $U^\perp$ invariant, $Ker(T_{|_{U^\perp}})$ is not just the zrero vector. But this contradicts $U \cap U^\perp = \{\mathbf{0}\}$. Thus, $U = V$.

13. $(T - \lambda I_V)^* = T^* - \overline{\lambda} I_V$. Since T and T^* commute and λI_V and $\overline{\lambda} I_V$ commute with all operators, $T - \lambda I_V$ and $T^* - \overline{\lambda} I_V$ commute and $T - \lambda I_V$ is normal.

14. i) implies ii). Assume T is normal. Then by Exercise 9, $Ker(T^*) = Ker(T) = W$. By Theorem (5,22), $U = Range(T) = Ker(T^*)^\perp = Ker(T)^\perp = W^\perp$.

ii) implies iii). Let $\mathcal{B}_U$ be an orthonormal basis for U and $\mathcal{B}_W$ be an orthonormal basis for W. Also, set $k = dim(U), l = dim(W)$. Then $\mathcal{B} = \mathcal{B}_U \sharp \mathcal{B}_W$ is an orthonormal basis for V. Let A denote $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ and $A^* = \mathcal{M}_{T^*}(\mathcal{B}, \mathcal{B})$.

Since $T = Proj_{(U,W)}, A = \begin{pmatrix} I_k & 0_{k\times l} \\ 0_{l\times k} & 0_{l\times l} \end{pmatrix}$. Since A is a real symmetric matrix, $A^* = \overline{A}^{tr} = A$. Consequently, $T^* = T$.

iii) implies i). A self-adjoint operator is always normal so there is nothing to prove.

6.2. Spectral Theorems

1. Let $\alpha_1, \dots, \alpha_s$ be the distinct eigenvalues of T and denote the minimum polynomial, $\mu_T(x) = (x-\alpha_1)\dots(x-\alpha_s)$ of T by $F(x)$ and set $F_i(x) = \frac{F(x)}{x-\alpha_i}$. Also, let $V_i = \{\boldsymbol{v} \in V | T(\boldsymbol{v}) = \alpha_i \boldsymbol{v}\}$ and $W_i = V_1 \oplus \dots \oplus V_{i-1} \oplus V_{i+1} \oplus \dots \oplus V_s$ so that $V = V_i \perp W_i$.

Now $(x - \alpha_i)$ and $F_i(x)$ are relatively prime so there are polynomials $a_i(x), b_i(x)$ such that $a_i(x)(x - \alpha_i) + b_i(x)F_i(x) = 1$. Now let $\boldsymbol{v}_i \in V_i, \boldsymbol{w}_i \in W_i$. Since $(T - \alpha_i I_V)(\boldsymbol{v}_i) = \mathbf{0}$ we have

$$b_i(T)f_i(T)(\boldsymbol{v}_i) =$$

$$[a_i(T)(T - \alpha_i I_V) + b_i(T)f_i(T)](\boldsymbol{v}_i) =$$

$$I_V(\boldsymbol{v}_i) = \boldsymbol{v}_i.$$

On the other hand

$$b_i(T)F_i(T)(\boldsymbol{w}_i) = \mathbf{0}$$

Now set $g_i(x) = \overline{\alpha_i} b_i(x) F_i(x)$. Then

$$g_i(T)(\boldsymbol{v}_i) = \overline{\alpha_i}\boldsymbol{v}_i$$

$$g_i(T)(\boldsymbol{w}_i) = \mathbf{0}.$$

Now set $g(x) = g_1(x) + \dots + g_s(x)$. Then $g(T) = T^*$.

2. i) implies ii). Let $\alpha_1, \dots, \alpha_s$ be the distinct eigenvalues of T and set $V_i = \{\boldsymbol{v} \in V | T(\boldsymbol{v}) = \alpha_i \boldsymbol{v}\}$. Then $V = V_1 \perp \dots \perp V_s$. Note that the eigenvalues of T^* are

$\overline{\alpha_1}, \ldots, \overline{\alpha_s}$ and $\{\boldsymbol{v} \in V | T^*(\boldsymbol{v}) = \overline{\alpha_i}\boldsymbol{v}\} = V_i$. Note that every subspace of V_i is $T-$ and T^*-invariant.

Now let U be a $T-$invariant subspace and set $U_i = U \cap V_i$. Then $U = U_1 \perp \cdots \perp U_s$. It follows by the above remark that each U_i is T^*-invariant and therefore U is $T-$invariant.

ii) implies iii). Assume U is $T-$invariant. By hypothesis, U is T^*-invariant. Then by Exercise 10 of Section (5.6) $U^\perp$ is $T-$invariant.

iii) implies i). Since $U + U^\perp = V, U \cap U^\perp = \{\mathbf{0}\}$. This implies that T is completely reducible. Since the field is the complex numbers this also implies that T is diagonalizable. Now let $\alpha_1, \ldots, \alpha_s$ be the distinct eigenvalues of T and set $V_i = \{\boldsymbol{v} \in V | T(\boldsymbol{v}) = \alpha_i \boldsymbol{v}\}$. We then have that $V = V_1 \oplus \cdots \oplus V_s$. Further, let V_i' denote $V_1 \oplus \cdots \oplus V_{i-1} \oplus V_{i+1} \oplus \cdots \oplus V_s$ so that $V = V_i \oplus V_i'$.

Note that if X is a $T-$invariant subspace then $X = X_1 \oplus \cdots \oplus X_s$ where $X_i = X \cap V_i$.

Now set $W = V_j^\perp$. Then $W = W_1 \oplus \cdots \oplus W_s$ where $W_i = W \cap V_i$. However, $W_j = W \cap V_j = \{\mathbf{0}\}$ and therefore $W \subset V_j'$. However, since $V = V_j \oplus V_j^\perp = V_j \oplus V_j'$ we must have $dim(W) = dim(V_j')$. Consequently, $V_j^\perp = W = V_j'$. In particular, for $i \neq j, V_i \subset V_j'$, that is, $V_i \perp V_j$. Now let $\mathcal{B}_i$ be an orthonormal basis for $V_i, 1 \leq i \leq s$ and set $\mathcal{B} = \mathcal{B}_1 \sharp \ldots \sharp \mathcal{B}_s$. Then $\mathcal{B}$ is an orthonormal basis for V and $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ is diagonal. It follows from Theorem (6.3) that T is normal.

3. $\overline{\begin{pmatrix} 4 & -i \\ i & 4 \end{pmatrix}}^{tr} = \begin{pmatrix} 4 & i \\ -i & 4 \end{pmatrix}^{tr} = \begin{pmatrix} 4 & -i \\ i & 4 \end{pmatrix}$. Thus, $T^* = T$.

With respect to the orthonormal basis $\left(\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-i}{\sqrt{2}} \end{pmatrix}\right)$ the matrix of T is $\begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}$.

4. $\left(\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}\right)$.

5. Assume that $b = c$. Then $\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ is an eigenvector with eigenvalue $b = c$. Then with respect to the orthonormal basis $\left(\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}\right)$ the matrix of T is diagonal with diagonal entries a, b, b so that by the spectral theorem T is self-adjoint.

Conversely, assume that T is self-adjoint. Since $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ are eigenvectors it must be the case that $Span\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}\right)^\perp = Span\left(\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}\right)$ is $T-$invariant, equivalently, $\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ is an eigenvector, say with eigenvalue d. However, $\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. Applying T we have

$$T\left(\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}\right) = T\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right) = T\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}\right) + T\left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right)$$

$$d\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = b\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + c\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$$d\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right) = b\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + c\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$$d\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + d\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = b\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + c\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$$(d-b)\begin{pmatrix}1\\-1\\0\end{pmatrix}+(d-c)\begin{pmatrix}0\\1\\-1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

Since $(\begin{pmatrix}1\\-1\\0\end{pmatrix},\begin{pmatrix}0\\1\\-1\end{pmatrix})$ is linearly independent, $b=c=d$.

6. Assume T is self-adjoint. Since $(\begin{pmatrix}1\\1\\1\\1\end{pmatrix},\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix},\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix})$ are eigenvectors it follows that $Span(\begin{pmatrix}1\\1\\1\\1\end{pmatrix},\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix},\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix})^{\perp}=Span(\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix})$ is $T-$invariant, equivalently, $\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}$ is an eigenvector for T.

Conversely, assume $\begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}$ is an eigenvector of T. by normalizing (dividing each vector by 2) we obtain an orthonormal basis of eigenvectors. We need to know that the corresponding eigenvalues are all real. If they are not all distinct then since three are distinct and real the fourth must be real. Therefore we may assume the eigenvalues are all distinct. Then the minimum polynomial of T has degree four and the eigenvalues are the roots of this polynomial. Since T is a real operator the minimum polynomial is a real polynomial. If it had a complex root then it would have to have a second. However, since three of the roots are real it must then be the case that the fourth root is real.

7. It T is self-adjoint we have seen that the eigenvalues are all real. Assume that T is normal and all its eigenvalues are real. Let $\mathcal{B}$ be an orthonormal basis such that $\mathcal{M}_T(\mathcal{B},\mathcal{B})$ is diagonal. Then $\mathcal{M}_T(\mathcal{B},\mathcal{B})$ is a real diagonal matrix. In particular, $\mathcal{M}_{T^*}(\mathcal{B},\mathcal{B})=\overline{\mathcal{M}_T(\mathcal{B},\mathcal{B})}^{tr}=\mathcal{M}_T(\mathcal{B},\mathcal{B})$ from which it follows that $T^*=T$.

8. Let $\alpha_1,\dots,\alpha_s$ be the distinct eigenvalue of S and $\beta_1,\dots,\beta_t$ the distinct eigenvalues of T. Set $V_i=\{\boldsymbol{v}\in V|S(\boldsymbol{v})=\alpha_i\boldsymbol{v}\}$. Then $V=V_1\oplus\cdots\oplus V_s$. Claim each V_i is $T-$invariant: let $\boldsymbol{v}\in V_i$. Then $S(T(\boldsymbol{v}))=(ST)(\boldsymbol{v})=(TS)(\boldsymbol{v})=T(S(\boldsymbol{v}))=T(\alpha_i\boldsymbol{v})=\alpha_iT(\boldsymbol{v})$.

Now set $W_j=\{\boldsymbol{u}\in V|T(\boldsymbol{u})=\beta_j\boldsymbol{u}\}$ so that $V=W_1\oplus\cdots\oplus W_t$. Now each W_j is $S-$invariant. Since V_i is $T-$invariant it follows that $V_i=V_{i1}\oplus\cdots\oplus V_{it}$ where $V_{ij}=V_i\cap W_j$. Note that if either $i\neq i'$ or $j\neq j'$ then $V_{ij}\perp V_{i'j'}$. Thus, V is the orthogonal direct sum of $V_{ij}, 1\le i\le s, 1\le j\le t$. Let $\mathcal{B}_{ij}$ be an orthonormal basis of V_{ij}. Order the collection of bases $\mathcal{B}_{ij}$ lexicographically. Then $\sharp_{i,j}\mathcal{B}_{ij}$is an orthonormal basis of V. Thus, there exists an orthonormal basis of eigenvectors for each of S and T.

9. If T is invertible there is nothing to prove so assume that $Ker(T)\neq\{\mathbf{0}\}$. Let $\alpha_1=0,\alpha_2,\dots,\alpha_s$ be the distinct eigenvalues of T. Set $V_i=\{\boldsymbol{v}\in V|T(\boldsymbol{v})=\alpha_i\boldsymbol{v}\}$. Since T is normal it is diagonalizable and $V=V_1\oplus\cdots\oplus V_t$. Let $i>1$ and $\boldsymbol{v}\in V_i$. Then $T^k(\boldsymbol{v})=\alpha_i^k\boldsymbol{v}\neq\mathbf{0}$ and therefore $\boldsymbol{v}\in Range(T^k)$. Thus, $Range(T^k)=V_2\oplus\cdots\oplus V_t$. Since $Range(T^k)=Range(T), rank(T^k)=rank(T)$. Then by the rank-nullity theorem $nullity(T^k)=nullity(T)$. However, $Ker(T)\subset Ker(T^k)$ and therefore we have equality: $Ker(T)=Ker(T^k)$.

10. If T is completely reducible on a complex space then there exists a basis $\mathcal{B}=(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ of eigenvectors. Define $\langle\sum_{i=1}a_i\boldsymbol{v}_i,\sum_{j=1}^n b_i\boldsymbol{v}_i\rangle=\sum_{i=1}^n a_i\overline{b_i}$. With respect to this inner product $\mathcal{B}$ is an orthonormal basis. By the Spectral theorem T is normal.

11. Let $\mathcal{B}_U$ be an orthonormal basis of U consisting of eigenvectors of $T_{|_U}$ and $\mathcal{B}_{U^\perp}$ be an orthonormal basis of $U^\perp$ consisting of eigenvectors of $T_{|_{U^\perp}}$. Since $T_{|_U}$ is self-adjoint, for each vector $\boldsymbol{u}\in\mathcal{B}_U, T(\boldsymbol{u})=\alpha\boldsymbol{u}$ for some

real number α. Since $T_{|_{U^\perp}}$ is self-adjoint, for each vector $\boldsymbol{v} \in \mathcal{B}_{U^\perp}, T(\boldsymbol{v}) = \beta\boldsymbol{v}$ for some real number $\boldsymbol{v}$.

Now $\mathcal{B} = \mathcal{B}_U \sharp \mathcal{B}_{U^\perp}$ is an orthonormal basis of V consisting of eigenvectors of T. Thus T is normal. Since all the eigenvalues are real it follows that T is self-adjoint.

12. This is false. Consider the operator which satisfies $T(\boldsymbol{e}_1) = \boldsymbol{e}_1, T(\boldsymbol{e}_2) = \boldsymbol{e}_2, T(\boldsymbol{e}_3) = 2\boldsymbol{e}_3, T(\boldsymbol{e}_4) = 2\boldsymbol{e}_4$ where $\boldsymbol{e}_i$ is the i^{th} standard basis vector of $\mathbb{R}^4$ equipped with the dot product. T is self-adjoint. Now let $U = Span(\boldsymbol{e}_1, \boldsymbol{e}_3)$ and $W = Span(\boldsymbol{e}_1 + \boldsymbol{e}_2, \boldsymbol{e}_3 + \boldsymbol{e}_4)$. Then W is $T-$invariant (the vectors $\boldsymbol{e}_1 + \boldsymbol{e}_2, \boldsymbol{e}_3 + \boldsymbol{e}_4$ are eigenvectors). Moreover, $\mathbb{R}^4 = U \oplus W$. However, $U^\perp = Span(\boldsymbol{e}_2, \boldsymbol{e}_4)$.

13. If T is self-adjoint then $U = Range(T) = Ker(T)^\perp = W^\perp$, whence $W = U^\perp$. Conversely, assume that $W = Ker(T)^\perp$. Choose an orthonormal basis $\mathcal{B}_U$ of U and an orthonormal basis $\mathcal{B}_W$ of W. Since $U \perp W$ it follows that $\mathcal{B} = \mathcal{B}_U \sharp \mathcal{B}_W$ is an orthonormal basis of V. Now $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ is diagonal with diagonal entries 0 and 1 from which we conclude that T is self-adjoint. Alternatively, the operator T is equal to $Proj_{(U,W)}$. By Exercise 14 of Section (6.1), T is self-adjoint if and only if $W = U^\perp$.

14. Since T is skew-Hermitian, $T^* = -T$. Assume $\boldsymbol{v}$ is an eigenvector with eigenvalue $\alpha \in \mathbb{C} \setminus \mathbb{R}$. Let $\boldsymbol{v}$ be a unit vector with eigenvalue $\alpha \neq 0$. Then

$$\alpha = \alpha\langle \boldsymbol{v}, \boldsymbol{v}\rangle = \langle \alpha\boldsymbol{v}, \boldsymbol{v}\rangle = \langle T(\boldsymbol{v}), \boldsymbol{v}\rangle$$

$$= \langle \boldsymbol{v}, T^*(\boldsymbol{v})\rangle = \langle \boldsymbol{v}, -T(\boldsymbol{v})\rangle = \langle \boldsymbol{v}, -\alpha\boldsymbol{v}\rangle$$

$$= -\overline{\alpha}\langle \boldsymbol{v}, \boldsymbol{v}\rangle = -\overline{\alpha}.$$

Thus, $\alpha + \overline{\alpha} = 0$ which implies that the real part of α is zero and α a pure imaginary number.

15. Let $\mathcal{B} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a orthonormal basis of V such that $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ is $diag\{\alpha_1, \ldots, \alpha_n\}$. Assume $[\boldsymbol{u}]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. Then $[T(\boldsymbol{u})]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 c_1 \\ \vdots \\ \alpha_n c_n \end{pmatrix}$ and $\langle T(\boldsymbol{u}), \boldsymbol{u}\rangle = \alpha_1|c_1|^2 + \ldots \alpha_n|c_n|^2 \in \mathbb{R}$.

6.3. Normal Operators on Real Inner Product Spaces

In all the following $\mathcal{S}_n$ is the standard orthonormal basis of $\mathbb{R}^n$.

1. Let T be the operator on $\mathbb{R}^4$ such that $\mathcal{M}_T(\mathcal{S}_4, \mathcal{S}_4) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & -2 & 0 \end{pmatrix}$.

2. Let T be the operator on $\mathbb{R}^4$ such that $\mathcal{M}_T(\mathcal{S}_4, \mathcal{S}_4) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$.

3. By Lemma (6.5) there is an orthonormal basis $\mathcal{S} = (\boldsymbol{v}_1, \boldsymbol{v}_2)$ and real numbers α, β with $\beta \neq 0$ such that the matrix of T with respect to $\mathcal{S}$ is $A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$.

Let $A' = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$. Then A' is the matrix of T^* with respect to $\mathcal{S}$. It therefore suffices to prove that there is a linear polynomial $f(x)$ such that $A' = f(A)$. Set $f(x) = -x + 2\alpha$.

4. Since T is normal it is completely reducible. Since the minimum polynomial is a real irreducible quadratic the minimal $T-$invariant subspaces have dimension 2. It follows that there are $T-$invariant subspaces $U_1, \ldots, U_s$ each of dimension 2 such that $V = U_1 \perp \cdots \perp U_s$. Assume that the roots of $\mu_T(x)$ are $\alpha \pm i\beta$ with $\beta \neq 0$.

Set $A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$. There are then orthonormal bases $\mathcal{B}_j$ of U_j such that $\mathcal{M}_{T|U_j}(\mathcal{B}_j, \mathcal{B}_j) = A$. If we set $\mathcal{B} = \mathcal{B}_1 \sharp \dots \sharp \mathcal{B}_s$ then $\mathcal{B}$ is an orthonormal basis of V and $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ is the block diagonal matrix with s blocks equal to A.

Now set $A' = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$. Then the matrix of T^* with respect to $\mathcal{B}$ is the block diagonal matrix with s blocks equal to A'. As we saw in Exercise 3, there is a real linear polynomial $f(x)$ such that $A' = f(A)$. It then follows that $\mathcal{M}_{T^*}(\mathcal{B}, \mathcal{B}) = f(\mathcal{M}_T(\mathcal{B}, \mathcal{B}))$ from which we conclude that $T^* = f(T)$.

5. Since T is completely reducible there are distinct real numbers $\alpha_1, \dots, \alpha_s$ and distinct real irreducible quadratics $p_1(x), \dots, p_t(x)$ such that $\mu_T(x) = (x - \alpha_1) \dots (x - \alpha_s) p_1(x) \dots p_t(x)$. Let the roots of $p_j(x)$ be $a_j \pm i b_j$ where $a_j, b_j \in \mathbb{R}$. Also, for $1 \le j \le s$ set $g_j(x) = \frac{\mu_T(x)}{(x - \alpha_j)}$ and for $1 \le k \le t$ set $h_k = \frac{\mu_T(x)}{p_k(x)}$. Further, for $1 \le j \le s$ set $U_j = \{\boldsymbol{v} \in V | T(\boldsymbol{v}) = \alpha_j \boldsymbol{v}\}$ and for $1 \le k \le t$ set $W_k = \{\boldsymbol{v} \in V | p_k(T)(\boldsymbol{v}) = \mathbf{0}\}$. Then

$$V = U_1 \perp \dots \perp U_s \perp W_1 \perp \dots \perp W_t.$$

Let $\mathcal{S}_j$ be an orthonormal basis of $U_j, 1 \le j \le s$ and $\mathcal{S}_k$ an orthonormal basis of W_k such that the matrix of T restricted to W_k with respect to $\mathcal{S}_k$ is block diagonal with blocks equal to $A_k = \begin{pmatrix} a_k & -b_k \\ b_k & a_k \end{pmatrix}$.

Set $A'_k = \begin{pmatrix} a_k & b_k \\ -b_k & a_k \end{pmatrix}$. As we have seen in Exercise 3 there is a linear polynomial $f_k(x)$ such that $A'_k = f_k(A_k)$.

Now, for each $j, (x - \alpha_j)$ and $g_j(x)$ are relatively prime and consequently there are polynomials $c_j(x), d_j(x)$ such that $c_j(x)(x - \alpha_j) + d_j(x) g_j(x) = 1$. Set $F_j(x) = d_j(x) g_j(x)$.

Also, for each $k, p_k(x)$ and $h_k(x)$ are relatively prime and so there are polynomials $C_k(x)$ and $D_k(x)$ such that $C_k(x) p_k(x) + D_k(x) h_k(x) = 1$. Set $G_k(x) = f_k(x) D_k(x) h_k(x)$.

Now set $f(x) = F_1(x) + \dots + F_s(x) + G_1(x) + \dots + G_t(x)$. Then $f(T) = T^*$.

6. By Exercise 5 there exists a polynomial $f(x)$ such that $T^* = f(T)$. Since $(T^*)^* = T$ there is also a polynomial $g(x)$ such that $T = g(T^*)$. Now assume that $ST = TS$. Then S commutes with $f(T) = T^*$. Likewise, if S commutes with T^* then S commutes with $g(T^*) = T$.

7. The hypothesis implies that T is cyclic. Since $TS = ST$ it follows that there is a polynomial $f(x)$ such that $S = f(T)$. Since $\mu_T(x)$ is quadratic, $f(x)$ is linear which implies that $S \in Span(T, I_V)$.

8. Under the given hypothesis, there are real distinct irreducible quadratic polynomials $p_1(x), \dots, p_s(x)$ such that $\mu_T(x) = p_1(x) \dots p_s(x)$ and if $U_j = \{\boldsymbol{v} \in V | p_j(T)(\boldsymbol{v}) = \mathbf{0}\}$ then $dim(U_j) = 2$. Moreover, $V = U_1 \perp \dots \perp U_s$. If U is a $T-$invariant subspace then $U = (U \cap U_1) \oplus \dots \oplus (U \cap U_s)$. It therefore suffices to show that each U_j is $S-$invariant. Since $ST = TS$ it follows that S commutes with $p_j(T)$. Assume now that $\boldsymbol{v} \in U_j$. Then $p_j(T)(S(\boldsymbol{v})) = (p_j(T)S)(\boldsymbol{v}) = (S p_j(T))(\boldsymbol{v}) = S(p_j(T)(\boldsymbol{v})) = S(\mathbf{0}) = \mathbf{0}$. Thus, $S(\boldsymbol{v}) \in U_j$.

9. Continue with the notation of Exercise 8. It suffices to prove that S restricted to each U_j is normal. However, this follows from Exercise 7.

10. Under the given hypotheses, T is a cyclic operator. Therefore $dim(C(T)) = dim(V) = deg(\mu_T(x))$ which is even.

11. We claim that $dim(C(T)) = 8$. Let $\mathcal{S} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$ be an orthonormal basis such that $\mathcal{M}_T(\mathcal{S}, \mathcal{S}) = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$. The vector $S(\boldsymbol{v}_1)$ can be chosen arbitrarily. Then $S(\boldsymbol{v}_2) = T(S(\boldsymbol{v}_1)) - S(\boldsymbol{v}_1)$. Likewise, $S(\boldsymbol{v}_3)$ can be chosen arbitrarily and $S(\boldsymbol{v}_4) = T(S(\boldsymbol{v}_3)) - S(\boldsymbol{v}_3)$.

12. Since T is skew-symmetric it is normal since T commutes with $-T$. Note that for any vector $\boldsymbol{v}$ we have $\langle T(\boldsymbol{v}), \boldsymbol{v}\rangle = \langle \boldsymbol{v}, T^*(\boldsymbol{v})\rangle = \langle \boldsymbol{v}, -T(\boldsymbol{v})\rangle$ which implies that $2\langle T(\boldsymbol{v}), \boldsymbol{v}\rangle = 0$. Thus, $\boldsymbol{v} \perp T(\boldsymbol{v})$. This implies that if $\boldsymbol{v}$ is an eigenvector then $T(\boldsymbol{v}) = \boldsymbol{0}$. Since T is invertible there are no eigenvectors. Thus, a minimal $T-$invariant subspace has dimension 2. Let $p(x)$ be a real irreducible quadratic polynomial dividing $\mu_T(x)$ with roots $\alpha \pm i\beta$ with $\alpha, \beta \in \mathbb{R}, \beta \neq 0$.

Let U be a 2-dimensional $T-$invariant subspace of V such that $p(T)$ restricted to U is the zero operator. Let $\mathcal{S} = (\boldsymbol{u}_1, \boldsymbol{u}_2)$ be an orthonormal basis of U such that the matrix of T with respect to $\mathcal{S}$ is $A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$. Since T is a skew-symmertic operator the matrix A is skew-symmetric. But this implies that $\alpha = 0$ and the roots of $p(x)$ are purely imaginary.

6.4. Unitary and Orthogonal Operators

1. Suppose $T(\boldsymbol{v}) = \boldsymbol{0}$. Then $0 = \parallel T(\boldsymbol{v}) \parallel = \parallel \boldsymbol{v} \parallel$. By positive definiteness we conclude that $\boldsymbol{v} = \boldsymbol{0}$ and so $Ker(T) = \{\boldsymbol{0}\}$ and T is injective. Since V is finite dimensional from the half is good enough theorem T is bijective. Let $\boldsymbol{v} \in V$ and set $\boldsymbol{u} = T^{-1}(\boldsymbol{v})$. Then $\boldsymbol{v} = T(\boldsymbol{u})$. Therefore, $\parallel T^{-1}(\boldsymbol{v}) \parallel = \parallel \boldsymbol{u} \parallel = \parallel T(\boldsymbol{u}) \parallel$ since T is an isometry. However, $T(\boldsymbol{u}) = \boldsymbol{v}$ so we have shown that $\parallel T^{-1}(\boldsymbol{v}) \parallel = \parallel \boldsymbol{v} \parallel$ and T^{-1} is an isometry.

2. Let S, T be isometries of V and let $\boldsymbol{v}$ be an arbitrary vector in V. Since S is an isometry, $\parallel S(T(\boldsymbol{v})) \parallel = \parallel T(\boldsymbol{v}) \parallel$. Since T is an isometry, $\parallel T(\boldsymbol{v}) \parallel = \parallel \boldsymbol{v} \parallel$. Thus, $\parallel (ST)(\boldsymbol{v}) \parallel = \parallel S(T(\boldsymbol{v})) \parallel = \parallel \boldsymbol{v} \parallel$ and ST is an isometry.

3. Let $\boldsymbol{u}$ be an arbitrary vector and assume $[\boldsymbol{u}]_{\mathcal{S}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ so that $\boldsymbol{u} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$. Then

$$\parallel \boldsymbol{u} \parallel^2 = \sum_{j=1}^{n} \parallel c_j \parallel^2 .$$

By the definition of $T, T(\boldsymbol{u}) = \sum_{j=1}^n \lambda_j c_j \boldsymbol{v}_j$. Then

$$\parallel T(\boldsymbol{u}) \parallel^2 = \sum_{j=1}^{n} \parallel \lambda_j c_j \parallel^2 =$$

$$\sum_{j=1}^{n} |\lambda_j|^2 \parallel c_j \parallel^2 = \sum_{j=1}^{n} \parallel c_j \parallel^2 = \parallel \boldsymbol{u} \parallel^2 .$$

4. We recall that for a real finite dimensional vector space and $\mathcal{S}$ an orthonormal basis, $\mathcal{M}_{T^*}(\mathcal{S}, \mathcal{S}) = \mathcal{M}_T(\mathcal{S}, \mathcal{S})^{tr}$. So, assume T is an isometry. Then $T^* = T^{-1}$ whence $A^{-1} = \mathcal{M}_{T^{-1}}(\mathcal{S}, \mathcal{S}) = \mathcal{M}_{T^*}(\mathcal{S}, \mathcal{S}) = A^{tr}$.

Conversely, assume $A^{-1} = A^{tr}$. Since $A^{-1} = \mathcal{M}_{T^{-1}}(\mathcal{S}, \mathcal{S})$ and $A^{tr} = \mathcal{M}_{T^*}(\mathcal{S}, \mathcal{S})$ we can conclude that $T^{-1} = T^*$ and therefore T is an isometry.

5. Let S be the operator such that $S(\boldsymbol{u}_j) = \boldsymbol{v}_j$. Then S is a unitary operator and, consequently, $\mathcal{M}_S(\mathcal{S}_1, \mathcal{S}_1)$ is a unitary matrix. However, $\mathcal{M}_{I_V}(\mathcal{S}_2, \mathcal{S}_1) = \mathcal{M}_S(\mathcal{S}_1, \mathcal{S}_1)$ and so the change of basis matrix, $\mathcal{M}_{I_V}(\mathcal{S}_2, \mathcal{S}_1)$, is a unitary matrix.

6. Let S be the operator such that $S(\boldsymbol{u}_j) = \boldsymbol{v}_j$. Then S is an orthogonal operator and $\mathcal{M}_S(\mathcal{S}_1, \mathcal{S}_1)$ is an orthogonal matrix. However, $\mathcal{M}_{I_V}(\mathcal{S}_2, \mathcal{S}_1) = \mathcal{M}_S(\mathcal{S}_1, \mathcal{S}_1)$ and consequently, $\mathcal{M}_{I_V}(\mathcal{S}_2, \mathcal{S}_1)$ is an orthogonal matrix.

7. Let $V = \mathbb{C}^n$ equipped with the usual inner product, S be the operator on V given by multiplication by A and let $\mathcal{S} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ be the standard (orthonormal basis). Assume $A\overline{A}^{tr} = \overline{A}^{tr}A$. Since $\mathcal{M}_{S^*}(\mathcal{S}, \mathcal{S}) = \overline{A}^{tr}$ it follows that S is normal. By the complex spectral theorem there exists an orthonormal basis $\mathcal{S}'$ consisting of eigenvectors for S, equivalently, so that $\mathcal{M}_S(\mathcal{S}', \mathcal{S}')$ is a diagonal matrix. Set $Q = \mathcal{M}_{I_V}(\mathcal{S}', \mathcal{S})$. By Exercise 5, Q is a unitary matrix. Then $Q^{-1}AQ = \mathcal{M}_S(\mathcal{S}', \mathcal{S}')$ is a diagonal matrix.

Conversely, assume there is a unitary matrix Q such that $Q^{-1}AQ$ is diagonal. Let T be the operator such

$\mathcal{M}_T(\mathcal{S},\mathcal{S}) = Q$. T is a unitary operator and therefore if $T(\boldsymbol{e}_j) = \boldsymbol{u}_j, \mathcal{S}' = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ is an orthonormal basis. Moreover, the vectors $\boldsymbol{u}_j$ are eigenvectors for S and, consequently, by the complex spectral theorem S is normal, that is, $SS^* = S^*S$. Since $A = \mathcal{M}_S(\mathcal{S},\mathcal{S})$ and $\overline{A}^{tr} = \mathcal{M}_{S^*}(\mathcal{S},\mathcal{S})$ it follows that $A\overline{A}^{tr} = \overline{A}^{tr}A$.

8. Let $V = \mathbb{R}^n$ equipped with the dot product, let S be the operator on V giben by multiplication by A and let $\mathcal{S} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ be the standard (orthonormal basis). Assume $A = A^{tr}$. Since $\mathcal{M}_{S^*}(\mathcal{S},\mathcal{S}) = A^{tr}$ it follows that S is self-adjoint. By the real spectral theorem there exists an orthonormal basis $\mathcal{S}'$ consisting of eigenvectors for S, equivalently, so that $\mathcal{M}_S(\mathcal{S}',\mathcal{S}')$ is a diagonal matrix. Set $Q = \mathcal{M}_{I_V}(\mathcal{S}',\mathcal{S})$. By Exercise 6, Q is an orthogonal matrix. Then $Q^{tr}AQ = Q^{-1}AQ = \mathcal{M}_S(\mathcal{S}',\mathcal{S}')$ is a diagonal matrix.

Conversely, assume there is an orthogonal matrix Q such that $Q^{tr}AQ = Q^{-1}AQ$ is diagonal. Let T be the operator such $\mathcal{M}_T(\mathcal{S},\mathcal{S}) = Q$. T is an orthogonal operator and therefore if $T(\boldsymbol{e}_j) = \boldsymbol{u}_j, \mathcal{S}' = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ is an orthonormal basis. Moreover, the vectors $\boldsymbol{u}_j$ are eigenvectors for S and, consequently, by the real spectral theorem, S is self-adjoint, that is, $S^* = S$ Since $A = \mathcal{M}_S(\mathcal{S},\mathcal{S})$ and $A^{tr} = \mathcal{M}_{S^*}(\mathcal{S},\mathcal{S})$ it follows that $A^{tr} = A$.

9. Assume T is a real isometry. Since $TT^* = T^*T = I_V$ it follows that T is a normal operator. Consequently, T is completely reducible and we can express V as an orthogonal sum $U_1 \perp \cdots \perp U_s \perp W_1 \perp \cdots \perp W_t$ of $T-$invariant subspspaces where $dim(U_j) = 1, dim(W_k) = 2$ and T restricted to W_k does not contain an eigenvector. Moreover, there is an orthonormal basis of $\mathcal{S}_k$ of W_k such that the matrix of T restricted to W_k with respect to $\mathcal{S}_k$ has the form $\begin{pmatrix} \alpha_k & -\beta_k \\ \beta_k & \alpha_k \end{pmatrix}$ with $\beta_k > 0$.

Let $\boldsymbol{u}_j \in U_j$ be a vector of norm 1. Since U_j is $T-$invariant, in particular, $\boldsymbol{u}_j$ is an eigenvector of T. Suppose $T(\boldsymbol{u}_j) = a_j\boldsymbol{u}_j$. Then $1 = \parallel \boldsymbol{u}_j \parallel = \parallel T(\boldsymbol{u}_j) \parallel = \parallel a_j\boldsymbol{u}_j \parallel = |a_j| \parallel \boldsymbol{u}_j \parallel = |a_j|$. Thus, $a_j = \pm 1$.

Finally, assume that $\mathcal{S}_k = (\boldsymbol{v}_{1k}, \boldsymbol{v}_{2k})$. Then $T(\boldsymbol{v}_{1k}) = \alpha_k\boldsymbol{v}_{1k} + \beta_k\boldsymbol{v}_{2k}$. Since T is an isometry

$$1 = \parallel \boldsymbol{v}_{1k} \parallel^2 = \parallel T(\boldsymbol{v}_{1k} \parallel^2 =$$

$$\parallel \alpha_k\boldsymbol{v}_{1k} + \beta_k\boldsymbol{v}_{2k} \parallel^2 = \alpha_k^2 + \beta_k^2.$$

Since $\beta_k > 0$ there exists a unique $\theta_k, 0 < \theta_k < \pi$ such that $\alpha_k = cos\ \theta_k, \beta_k = sin\ \theta_k$. Now by setting $\mathcal{S} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_s)\sharp\mathcal{S}_1\sharp\ldots\sharp\mathcal{S}_t$ we obtain an orthonormal basis of V and the matrix of T is block diagonal with entries $a_j, 1 \leq j \leq s$ and 2×2 blocks $A_k = \begin{pmatrix} cos\ \theta_k & -sin\ \theta_k \\ sin\ \theta_k & cos\ \theta_k \end{pmatrix}$ for $1 \leq k \leq t$.

Conversely, assume that there exists an orthonormal basis $\mathcal{S}$ such that $\mathcal{M}_T(\mathcal{S},\mathcal{S})$ is block diagonal and each block is either 1×1 with entry ± 1 or 2×2 of the form $\begin{pmatrix} cos\ \theta & -sin\ \theta \\ sin\ \theta & cos\ \theta \end{pmatrix}$ for some $\theta, 0 \leq \theta < \pi$.

Assume the basis has been ordered so that the first s blocks are 1×1 and the remaining blocks, $A_1, \ldots, A_t$ are 2×2 and suppose $A_k = \begin{pmatrix} cos\ \theta_k & -sin\ \theta_k \\ sin\ \theta_k & cos\ \theta_k \end{pmatrix}$. Suppose $\mathcal{S} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_s)\sharp\mathcal{S}_1\sharp\ldots\sharp\mathcal{S}_t$ where $\mathcal{S}_k = (\boldsymbol{v}_{1k}, \boldsymbol{v}_{2k})$ consists of two orthogonal unit vectors. By our assumption on the $\mathcal{M}_T(\mathcal{S},\mathcal{S})$ $T(\boldsymbol{u}_j) = a_j\boldsymbol{u}_j$ where $a \in \{-1, 1\}$. Also, $T(\boldsymbol{v}_{1k}) = cos\ \theta_k\boldsymbol{v}_{1k} + sin\ \theta_k\boldsymbol{v}_{2k}, T(\boldsymbol{v}_{2k}) = -sin\ \theta_k\boldsymbol{v}_{1k} + cos\ \theta_k\boldsymbol{v}_{2k}$. Then $T(\mathcal{S}_k) = (T(\boldsymbol{v}_{1k}), T(\boldsymbol{v}_{2k}))$ is an orthonormal basis for $Span(\boldsymbol{v}_{1k}, \boldsymbol{v}_{2k})$. Thus, $(T(\boldsymbol{u}_1), \ldots, T(\boldsymbol{u}_s))\sharp T(\mathcal{S}_1)\sharp\ldots\sharp T(\mathcal{S}_t)$ is an orthonormal basis of V and T is an isometry.

10. If T is an isometry then $T^*T = TT^* = I_V$. If T is self-adjoint then $T^* = T$ and therefore $T^2 = I_V$. It follows that the minimum polynomial of T divides $x^2 - 1 = (x-1)(x+1)$. Consequently, the eigenvalues of T are all ± 1. Since T is self-adjoint there exists an orthonormal basis $\mathcal{S}$ such that $\mathcal{M}_T(\mathcal{S},\mathcal{S})$ is diagonal. Since the eigenvalues of T are all ± 1 the diagonal entries of $\mathcal{M}_T(\mathcal{S},\mathcal{S})$ are all ± 1.

11. Assume that T is a self-adjoint operator. Then there exists an orthonormal basis $\mathcal{S} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ consisting

of eigenvectors of T. Moreover, if $T(\boldsymbol{v}_j) = a_j\boldsymbol{v}_j$ then $a_j \in \mathbb{R}$. Suppose now that $T^2 = I_V$. Then the minimum polynomial of T divides $x^2 - 1$ and all the $a_j \in \{1, -1\}$. Then $(T(\boldsymbol{v}_1), \dots, T(\boldsymbol{v}_n))$ is an orthonormal basis and T is an isometry.

Conversely, assume that T is an isometry. It follows from Exercise 10 that $T^2 = I_V$.

12. Let T be the operator on $\mathbb{C}^2$ which has matrix $\begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}$ with respect to the standard basiis.

13. Assume U is $T-$invariant. Since T is an isometry, by Exercise 1, T is bijective. In particular, T restricted to U is injective and since U is $T-$invariant, T restricted to U is bijective. Assume $\boldsymbol{w} \in U^\perp$ and $\boldsymbol{u} \in U$ is arbitrary. We need to prove that $T(\boldsymbol{w}) \perp \boldsymbol{u}$. Since T restricted to U is bijective there exists $\boldsymbol{v} \in U$ such that $T(\boldsymbol{v}) = \boldsymbol{u}$. Now

$$\langle T(\boldsymbol{w}), \boldsymbol{u}\rangle = \langle T(\boldsymbol{w}), T(\boldsymbol{v})\rangle = \langle T^*T(\boldsymbol{w}), \boldsymbol{u}\rangle.$$

Since T is an isometry, $T^*T = I_V$ and therefore

$$\langle T^*T(\boldsymbol{w}), \boldsymbol{u}\rangle = \langle \boldsymbol{w}, \boldsymbol{u}\rangle = 0.$$

14. Assume A is upper triangular and a unitary matrix. Then the diagonal entries of A are non-zero since A is invertible. The inverse of an upper triangular matrix is upper triangular. On the other hand, since A is unitary, $A^{-1} = \overline{A}^{tr}$ is lower triangular. So A^{-1} is both upper and lower triangular and hence diagonal and therefore A is diagonal.

15. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ be an orthonormal basis of U_1 and set $\boldsymbol{u}'_j = R(\boldsymbol{u}_j)$. Then $(\boldsymbol{u}'_1, \dots, \boldsymbol{u}'_k)$ is an orthonormal basis of U_2. Extend $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ to an orthonormal basis $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$ of V and extend $(\boldsymbol{u}'_1, \dots, \boldsymbol{u}'_k)$ to an orthonormal basis $(\boldsymbol{u}'_1, \dots, \boldsymbol{u}'_n)$ of V. Let S be the operator on V such that $S(\boldsymbol{u}_j) = \boldsymbol{u}'_j$ for $1 \leq j \leq n$. Then S restricted to U_1 is equal to R and since S takes an orthonormal basis of V to an orthonormal basis of V, S is an isometry.

16. Since the dimension of V is odd, S must have an eigenvector $\boldsymbol{v}$. Since $\| \boldsymbol{v} \| = \| S(\boldsymbol{v}) \|$ the eigenvalue of $\boldsymbol{v}$ is ± 1. Then $S^2(\boldsymbol{v}) = \boldsymbol{v}$.

17. Let $\mathcal{S}_U$ be an orthonormal basis of U and $\mathcal{S}_{U^\perp}$ be an orthonormal basis of $U^\perp$. Then $\mathcal{S} = \mathcal{S}_U \cup \mathcal{S}_{U^\perp}$ is an orthonormal basis of V. $\mathcal{M}_T(\mathcal{S}, \mathcal{S})$ is a diagonal matrix with ± 1 on the diagonal and therefore T is a isometry and self-adjoint.

18. Set $\mathcal{S} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4) = (\begin{pmatrix} 1\\1\\1\\1 \end{pmatrix}, \begin{pmatrix} 1\\1\\-1\\-1 \end{pmatrix}, \begin{pmatrix} 1\\-1\\1\\-1 \end{pmatrix}, \begin{pmatrix} 1\\-1\\-1\\1 \end{pmatrix})$ and $\mathcal{S}' = (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4) = (\begin{pmatrix} 1\\1\\1\\1 \end{pmatrix}, \begin{pmatrix} 1\\1\\-1\\-1 \end{pmatrix}, \begin{pmatrix} 1\\-1\\1\\-1 \end{pmatrix}, \begin{pmatrix} 1\\0\\0\\0 \end{pmatrix})$.

A necessary and sufficient condition for an operator Q to satisfy $Q^{-1}SQ = T$ is that $Qi(\boldsymbol{v}_j) = a_j\boldsymbol{u}_j$ with $a_j \neq 0$. If Q is an isometry then we must have $Q(\boldsymbol{v}_j) = \pm\boldsymbol{u}_j$ for $j = 1, 2, 3$ it order to preserve norms and $Q(\boldsymbol{v}_4) = \pm\frac{1}{2}\boldsymbol{u}_4$. However, since $\langle \boldsymbol{v}_3, \boldsymbol{v}_4\rangle = 1$ we must have $\langle Q(\boldsymbol{v}_3), Q(\boldsymbol{v}_4)\rangle = 1$ whereas $Q(\boldsymbol{v}_3), Q(\boldsymbol{v}_4)$ are orthogonal.

6.5. Positive Operators, Polar Decomposition and Singular Value Decomposition

1. Assume S is a positive operator and $S^2 = T$. Then $ST = TS$. Since both are self-adjoint there exists an orthonormal basis $\mathcal{S} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ consisting of eigenvectors for both S and T. Let $S(\boldsymbol{v}_j) = a_j\boldsymbol{v}_j$ and $T(\boldsymbol{v}_j) = b_j\boldsymbol{v}_j$ and assume the notation has been chosen so that $b_j \neq 0$ for $j \leq k, b_j = 0$ for $j > k$ so that $Ker(T) = Span(\boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n)$. Now $b_j\boldsymbol{v}_j = T(\boldsymbol{v}_j) = S^2(\boldsymbol{v}_j) = a_j^2\boldsymbol{v}_j$. If follows that if $j > k$ then $a_j = 0$. If

$j \leq k$ then $a_j, b_j > 0$ and $a_j^2 = b_j$ so that a_j is uniquely determined. Thus, S is unique.

2. Since T is normal there exists an orthonormal basis $\mathcal{S} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ consisting of eigenvectors of T. Assume $T(\boldsymbol{v}_j) = a_j\boldsymbol{v}_j$. Let $b_j \in \mathbb{C}$ such that $b_j^2 = a_j$ and let S be the operator such that $S(\boldsymbol{v}_j) = b_j\boldsymbol{v}_j$. Then $S^2 = T$.

3. Under the hypothesis there is an orthonormal basis $\mathcal{S} = (\boldsymbol{v}_1, \boldsymbol{v}_2)$ such that $\mathcal{M}_T(\mathcal{S}, \mathcal{S}) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ with $b > 0$. Set $r = \sqrt{a^2 + b^2}, A = \frac{a}{r}, B = \frac{b}{R}$. Then $A^2 + B^2 = 1$ and $B > 0$ so that there exists a unique $\theta, 0 < \theta < \pi$ such that $A = cos\ \theta, B = sin\ \theta$. Let S be the operator such that $\mathcal{M}_S(\mathcal{S}, \mathcal{S}) = \begin{pmatrix} rcos\ \frac{\theta}{2} & -rsin\ \frac{\theta}{2} \\ rsin\ \frac{\theta}{2} & rcos\ \frac{\theta}{2} \end{pmatrix}$. Then $S^2 = T$.

4. Under the hypothesis there are $T-$invariant subspaces $U_1, \ldots, U_n$ each of dimension two such that for $j \neq k, U_j \perp U_k$ and $V = U_1 \oplus \cdots \oplus U_n$. By Exercise 3, there exists an operator S_j such that $Ker(S_j) = U_1 \oplus \cdots \oplus U_{j-1} \oplus U_{j+1} \oplus \cdots \oplus U_n$ and such that S_j^2 is the restriction of T to U_j. Set $S = S_1 + \cdots + S_n$. Then $S^2 = T$.

5. Let S, T be positive operators. Then $(S+T)^* = S^* + T^* = S + T$ and so $S + T$ is self-adjoint. Now let $\boldsymbol{v} \in V$. Then $\langle S(\boldsymbol{v}), \boldsymbol{v}\rangle \geq 0$ and $\langle T(\boldsymbol{v}), \boldsymbol{v}\rangle$. It now follows that

$$\langle (S+T)(\boldsymbol{v}), \boldsymbol{v}\rangle = \langle S(\boldsymbol{v}) + T(\boldsymbol{v}), \boldsymbol{v}\rangle =$$

$$\langle S(\boldsymbol{v}), \boldsymbol{v}\rangle + \langle T(\boldsymbol{v}), \boldsymbol{v}\rangle \geq 0 + 0 = 0.$$

6. Since T is a positive operator it is self-adjoint. Since c is a real number cT is self-adjoint. Now let $\boldsymbol{v} \in V$. We then have

$$\langle (cT)(\boldsymbol{v}), \boldsymbol{v}\rangle = \langle cT(\boldsymbol{v}), \boldsymbol{v}\rangle = c\langle T(\boldsymbol{v}), \boldsymbol{v}\rangle \geq 0$$

the latter inequality since $\langle T(\boldsymbol{v}), \boldsymbol{v}\rangle \geq 0$ and $c > 0$.

7. Since T is a positive operator there exists an orthonormal basis $\mathcal{S} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ of eigenvectors such that if $T(\boldsymbol{v}_j) = a_j\boldsymbol{v}_j$ then $a_j \geq 0$. Assume that T is invertible. Then all $a_j > 0$. Let $\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n \neq \boldsymbol{0}$. Then

$$\langle T(\boldsymbol{v}), \boldsymbol{v}\rangle = \langle \sum_{j=1}^n a_jc_j\boldsymbol{v}_j, \sum_{j=1}^n c_j\boldsymbol{v}_j\rangle =$$

$$\sum_{j=1}^n a_jc_j^2.$$

Since $a_j > 0$ and $c_j \in \mathbb{R}, a_jc_j^2 \geq 0$. On the other hand, some $c_j > 0$ and therefore $\sum_{j=1}^n a_jc_j^2 > 0$.

On the other hand, suppose T is not invertible and $\boldsymbol{v} \in Ker(T)$. Then $\langle T(\boldsymbol{v}), \boldsymbol{v}\rangle = 0$.

8. Since T is an invertible positive operator there exists an orthonormal basis $\mathcal{S} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ such that $T(\boldsymbol{v}_j) = a_j\boldsymbol{v}_j$ with $a_j \in \mathbb{R}^+$. Let S be the operator such that $S(\boldsymbol{v}_j) = \frac{1}{a_j}\boldsymbol{v}_j$. Then $S = T^{-1}$ since $ST(\boldsymbol{v}_j) = \boldsymbol{v}_j$ for all j. Since $\frac{1}{a_j} > 0$ it follows that $T^{-1} = S$ is a positive operator.

9. 1) [,] is positive definite: Assume $\boldsymbol{v} \neq \boldsymbol{0}$. Then $[\boldsymbol{v}, \boldsymbol{v}] = \langle T(\boldsymbol{v}), \boldsymbol{v}\rangle > 0$ by Exercise 7.

2) [,] is additive in the first argument: $[\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}] = \langle T(\boldsymbol{v}_1 + \boldsymbol{v}_2), \boldsymbol{w}\rangle = \langle T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2), \boldsymbol{w}\rangle = \langle T(\boldsymbol{v}_1), \boldsymbol{w}\rangle + \langle T(\boldsymbol{v}_2), \boldsymbol{w}\rangle = [\boldsymbol{v}_1, \boldsymbol{w}] + [\boldsymbol{v}_2, \boldsymbol{w}]$.

3) [,] is homogeneous in the first argument: $[c\boldsymbol{v}, \boldsymbol{w}] = \langle T(c\boldsymbol{v}), \boldsymbol{w}\rangle = \langle cT(\boldsymbol{v}), \boldsymbol{w}\rangle = c\langle T(\boldsymbol{v}), \boldsymbol{w}\rangle = c[\boldsymbol{v}, \boldsymbol{w}]$.

4) [.] satisfies conjugate symmetry: $[\boldsymbol{w}, \boldsymbol{v}] = \langle T(\boldsymbol{w}), \boldsymbol{v}\rangle = \langle \boldsymbol{w}, T^*(\boldsymbol{v})\rangle$. However, since T is a positive operator it is self-adjoint and $T^* = T$. Thus

$$\langle \boldsymbol{w}, T^*(\boldsymbol{v})\rangle = \langle \boldsymbol{w}, T(\boldsymbol{v})\rangle =$$

$$\overline{\langle T(\boldsymbol{v}), \boldsymbol{w}\rangle} = \overline{[\boldsymbol{v}, \boldsymbol{w}]}.$$

10. $[S(\boldsymbol{v}), \boldsymbol{w}] = \langle T(S(\boldsymbol{v})), \boldsymbol{w}\rangle = \langle S(\boldsymbol{v}), T(\boldsymbol{w})\rangle$ since T is self-adjoint. We therefore need to show that $[\boldsymbol{v}, (T^{-1}S^*T)(\boldsymbol{w})] = \langle S(\boldsymbol{v}), T(\boldsymbol{w})\rangle$. By the definition of [,] we have

$$[\boldsymbol{v},(T^{-1}S^*T)(\boldsymbol{w})] = \langle T(\boldsymbol{v}),(T^{-1}S^*T)(\boldsymbol{w})\rangle =$$
$$\langle \boldsymbol{v}, T(T^{-1}S^*T)(\boldsymbol{w})\rangle =$$
$$\langle \boldsymbol{v}, (TT^{-1})(S^*T)(\boldsymbol{v})\rangle =$$
$$\langle \boldsymbol{v}, S^*(T(\boldsymbol{w})\rangle =$$
$$\langle S(\boldsymbol{v}), T(\boldsymbol{w})\rangle$$

as was required.

11. Define $[\ ,\ ]$ by $[\boldsymbol{v},\boldsymbol{w}] = \langle T(\boldsymbol{v}),\boldsymbol{w}\rangle$. By Exercise 9 this is an inner product on V. Set $S = RT$. By Exercise 10 the adjoint, $S^\star$, of S with respect to $[\ ,\ ]$ is $S^\star = T^{-1}S^*T$. Now $S^* = (RT)^* = T^*R^* = TR$ since both R and T are self-adjoint. Thus, $S^\star = (RT)^\star = T^{-1}(TR)T = RT$. Consequently, RT is self-adjoint with respect to $[\ ,\ ]$ and therefore is diagonalizable with real eigenvalues. The operator TR is similar to RT since $TR = T(RT)T^{-1}$ and since RT is diagonalizable with real eigenvalues so is TR.

12. Assume T is a positive operator and let $\mathcal{S} = (\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ be an orthonormal basis for V such that $T(\boldsymbol{v}_j) = a_j\boldsymbol{v}_j$ where $a_j \in \mathbb{R}^+ \cup \{0\}$. Now assume that T is an isometry. Then we must have $|a_j| = 1$ for all j, whence, $a_j = 1$ for all j and $T = I_V$.

13. Since S, T are self-adjoint and $ST = TS$ it follows that ST is self-adjoint. Then there exists an orthonormal basis $\mathcal{S} = (\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ consisting of eigenvectors for S and for T. Set $S(\boldsymbol{v}_j) = a_j\boldsymbol{v}_j, T(\boldsymbol{v}_j) = b_j\boldsymbol{v}_j$. Since S,T are positive, $a_j, b_j \geq 0$. Then $a_jb_j \geq 0$. Now $\mathcal{S}$ is an orthonormal basis for V consisting of eigenvectors for ST and $ST(\boldsymbol{v}_j) = a_jb_j \geq 0$. It follows that ST is a positive operator.

14. Let the operators S and T on $\mathbb{R}^2$ be defined as multiplication by the following matrices, respectively:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}.$$

15. Assume T is invertible. Then T^*T is invertible and hence so is $\sqrt{T^*T}$. Then $S = T\sqrt{T^*T}^{-1}$ is unique. On the other hand if T is not invertible then T^*T is not invertible and neither is $\sqrt{T^*T}$. In this case there are infinitely many isometries S which extend R and so S is not unique.

16. Since T is not invertible, S is not unique. One solution is

$$\begin{pmatrix} \frac{1}{3} & \frac{-\sqrt{3}-1}{3} & \frac{-\sqrt{3}+1}{3} \\ \frac{\sqrt{3}-1}{3} & \frac{1}{3} & \frac{-\sqrt{3}-1}{3} \\ \frac{\sqrt{3}+1}{3} & \frac{\sqrt{3}+1}{3} & \frac{1}{3} \end{pmatrix}$$

17. Let $V = \mathbb{C}^n, W = \mathbb{C}^m$ and $T : V \to W$ be the operator such that $T(\boldsymbol{v}) = A\boldsymbol{v}$. Let $\mathcal{S}_V$ denote the standard basis of V and $\mathcal{S}_W$ the standard basis of W.

By Theorem (6.12) there are orthonormal bases $\mathcal{B}_V = (\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ and $\mathcal{B}_W = (\boldsymbol{u}_1,\dots,\boldsymbol{u}_m)$ such that $T(\boldsymbol{v}_j) = s_j\boldsymbol{w}_j$ for $1 \leq j \leq r$ and $T(\boldsymbol{v}_j) = \boldsymbol{0}_W$ if $j > r$. Let $P = \mathcal{M}_{I_V}(\mathcal{B}_V, \mathcal{S}_V)$ and $Q = \mathcal{M}_{I_W}(\mathcal{S}_W, \mathcal{B}_W)$. Assume $1 \leq j \leq r$. Then

$$QAP\boldsymbol{e}_j^V = QA\boldsymbol{v}_j = Q(T(\boldsymbol{v}_j)) = Q(s_j\boldsymbol{w}_j) =$$
$$s_jQ\boldsymbol{w}_j = s_j\boldsymbol{e}_j^W.$$

If $j > r$ then

$$QAP\boldsymbol{e}_j^V = QA\boldsymbol{v}_j = Q(T(\boldsymbol{v}_j)) = Q\boldsymbol{0}_W = \boldsymbol{0}_W.$$

Thus, QAP has the required form.

18. We have shown that $Ker(T) = Ker(T^*T)$ and similarly, $Ker(T^*) = Ker(TT^*)$. Since T^*T is a self-adjoint operator it is diagonalizable and $V = Ker(T^*T) \oplus Range(T^*T) = Ker(T) \oplus Range(T^*T)$. Similarly, $V = Ker(T^*) \oplus Range(TT^*)$. Note that since $nullity(T) = nullity(T^*)$ it follows that $dim(Range(T^*T)) = dim(Range(TT^*))$. Let S denote the restriction of the operator T to $Range(T^*T)$. Suppose $\boldsymbol{v} \in Range(T^*T)$. Then there is a vector $\boldsymbol{u} \in V$ such that $\boldsymbol{v} = (T^*T)(\boldsymbol{u})$. Then $S(\boldsymbol{v}) =$

$T(\boldsymbol{v}) = T([T^*T])(\boldsymbol{u}) = (TT^*)(T(\boldsymbol{u})) \in Range(TT^*)$. Thus, $Range(S) \subset Range(TT^*)$. However, S is injective since $Ker(T) \cap Range(T^*T) = \{\mathbf{0}\}$. Since $dim(Range(T^*T)) = dim(Range(TT^*))$, in fact, S is an isomorphism.

Now assume that $\boldsymbol{v} \in Range(T^*T)$ is an eigenvector with eigenvalue a. We claim that $S(\boldsymbol{v})$ is an eigenvector of TT^* with eigenvalue a. Thus, we must apply TT^* to $S(\boldsymbol{v})$ and obtain $aS(\boldsymbol{v})$.

$$(TT^*)(S(\boldsymbol{v}) = (TT^*)(T(\boldsymbol{v})) = [T(T^*T)](\boldsymbol{v}) =$$

$$T[(T^*T)(\boldsymbol{v})] = T(a\boldsymbol{v}) = aT(\boldsymbol{v}) = aS(\boldsymbol{v})$$

as required.

19. Assume $rank(T) = k$. Since T is semi-positive there exists an orthonormal basis $\mathcal{B} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ of eigenvectors for T with $T(\boldsymbol{v}_j) = a_j\boldsymbol{v}_j$ with $a_j > 0$ for $1 \leq j \leq k$ and $T(\boldsymbol{v}_j) = \mathbf{0}$ for $j > k$. Now $T^*T = T^2$ since T is self-adjoint. Now $T^2(\boldsymbol{v}_j) = a_j^2\boldsymbol{v}_j$ for $1 \leq j \leq k$ and $T^2(\boldsymbol{v}_j) = \mathbf{0}$ for $j > k$. Then the singular values of T are $\sqrt{a_j^2}, 1 \leq j \leq k$. However, since $a_j > 0$, $\sqrt{a_j^2} = a_j$.

20. Assume $SP = PS$. Multiplying on the left and on the right by S^{-1} we get $PS^{-1} = S^{-1}P$.

Now, $(SP)^* = P^*S^* = PS^{-1} = S^{-1}P$. Then $(SP)^*(SP) = (PS^{-1})(SP) = P^2$. On the other hand, $(SP)(SP)^* = (SP)(PS^{-1}) = (PS)(S^{-1}P) = P^2$ and SP is normal.

Conversely, assume SP is normal. Then $P^2 = (SP)^*(SP) = (SP)(SP)^* = SP^2S^{-1}$. Thus, S commutes with P^2 and therefore leaves invariant each eigenspace of P^2. However, since P is positive the eigenspaces of P and the eigenspaces of P^2 are the same. Therefore S leaves the eigenspaces of P invariant. Let $\boldsymbol{v}$ be an eigenvector of P with eigenvalue a. Then $(SP)(\boldsymbol{v}) = S(P(\boldsymbol{v})) = S(a\boldsymbol{v}) = aS(\boldsymbol{v})$. On the other hand, $(PS)(\boldsymbol{v}) = P(S(\boldsymbol{v})) = aS(\boldsymbol{v})$ since $S(\boldsymbol{v})$ is an eigenvector with eigenvalue a.

Chapter 7

Trace and Determinant of a Linear Operator

7.1. Trace of a Linear Operator

1. Let A have entries $a_{ij}, 1 \leq i,j \leq n$ and B have entries $b_{ij}, 1 \leq i,j \leq n$. Then $A+B$ has entries $a_{ij}+b_{ij}, 1 \leq i,j \leq n$. Then traces are given by

$$Trace(A) = a_{11} + \cdots + a_{nn}$$

$$Trace(B) = b_{11} + \cdots + b_{nn}$$

$$Tr(A+B) = (a_{11}+b_{11}) + \cdots + (a_{nn}+b_{nn}) =$$

$$[a_{11} + \cdots + a_{nn}] + (b_{11} + \cdots + b_{nn}] =$$

$$Trace(A) + Trace(B).$$

2. Let A have entries $a_{ij}, 1 \leq i,j \leq n$. Then the entries of cA are $ca_{ij}, 1 \leq i,j \leq n$. The traces are given by

$$Trace(A) = a_{11} + \cdots + a_{nn}$$

$$Trace(cA) = (ca_{11}) + \cdots + (ca_{nn}) =$$

$$c[a_{11} + \cdots + a_{nn}] = cTrace(A).$$

3. Assume P is an invertible $n \times n$ matrix and C is an $n \times n$ matrix and set $D = P^{-1}CP$. We need to show that $Trace(D) = Trace(C)$. Set $A = CP$ and $B = P^{-1}$. By Theorem (7.1), $Trace(AB) = Trace(BA)$. However, $AB = (CP)P^{-1} = C(PP^{-1}) = CI_n = C$ whereas $BA = P^{-1}CP = D$.

4. The matrices $\mathcal{M}_T(\mathcal{B},\mathcal{B})$ and $\mathcal{M}_T(\mathcal{B}',\mathcal{B}')$ are similar so by Corollary (7.1) they have the same trace.

5. Let $\mathcal{B}$ be a basis for V. Then $Tr(ST) = Trace(\mathcal{M}_{ST}(\mathcal{B},\mathcal{B}))$. However, $\mathcal{M}_{ST}(\mathcal{B},\mathcal{B}) = \mathcal{M}_S(\mathcal{B},\mathcal{B})\mathcal{M}_T(\mathcal{B},\mathcal{B})$ so that

$$Tr(ST) = Trace(\mathcal{M}_S(\mathcal{B},\mathcal{B})\mathcal{M}_T(\mathcal{B},\mathcal{B}))$$

In exactly the same way

$$Tr(TS) = Trace(\mathcal{M}_T(\mathcal{B},\mathcal{B})\mathcal{M}_S(\mathcal{B},\mathcal{B}))$$

By Theorem (7.1), $Trace(\mathcal{M}_S(\mathcal{B},\mathcal{B})\mathcal{M}_T(\mathcal{B},\mathcal{B}) = Trace(\mathcal{M}_T(\mathcal{B},\mathcal{B})\mathcal{M}_S(\mathcal{B},\mathcal{B}))$.

6. Let $\mathcal{B}$ be a basis for V. Then $Tr(cT) = Trace(\mathcal{M}_{cT}(\mathcal{B},\mathcal{B})) = Trace(c\mathcal{M}_T(\mathcal{B},\mathcal{B}))$. However, by Exercise 2, $Trace(c\mathcal{M}_T(\mathcal{B},\mathcal{B})) = cTrace(\mathcal{M}_T(\mathcal{B},\mathcal{B})) = cTr(T)$.

7. Since $(x_1+x_2+x_3)^2 - (x_1^2+x_2^2+x_3^2) = 2(x_1x_2 + x_1x_3 + x_2x_3)$ we conclude that $x_1x_2 + x_1x_3 + x_2x_3 = 0$.

Next note that

$$3x_1x_2x_3 = (x_1+x_2+x_3)(x_1x_2+x_1x_3+x_2x_3) -$$

$$(x_1+x_2+x_3)(x_1^2+x_2^2+x_3^2) + (x_1^3+x_2^3+x_3^3) = 0$$

We may therefore assume that at least one of x_1, x_2, x_3 is zero and by symmetry that $x_3 = 0$. Since $x_1x_2 + x_1x_2 + x_2x_3 = 0$ we further conclude that $x_1x_2 = 0$. But then

$$(x_1 \pm x_2)^2 = x_1^2 + x_2^2 \pm 2x_1x_2 = 0$$

from which we have

$$x_1 + x_2 = x_1 - x_2 = 0$$

and $x_1 = x_2 = 0$.

8. Since A is similar to an upper triangular matrix we can assume that A is upper triangular. Let $\alpha_1, \alpha_2, \alpha_3$ be the diagonal entries of A. Then the diagonal entries of A^2 are $\alpha_1^2, \alpha_2^2, \alpha_3^2$ and the diagonal entries of A^3 are $\alpha_1^3, \alpha_2^3, \alpha_3^3$. It follows that

$$0 = Tr(A) = \alpha_1 + \alpha_2 + \alpha_3$$

$$0 = Tr(A^2) = \alpha_1^2 + \alpha_2^2 + \alpha_3^2$$

$$0 = Tr(A^3) = \alpha_1^3 + \alpha_2^3 + \alpha_3^3.$$

From Exercise 7, $\alpha_1 = \alpha_2 = \alpha_3 = 0$ and A is nilpotent.

9. This depends on proving the only solution in complex numbers to the systems of equation

$$\begin{array}{ccccccc} x_1 & + & \dots & + & x_n & = & 0 \\ x_1^2 & + & \dots & + & x_n^2 & = & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \\ x_1^n & + & \dots & + & x_n^n & = & 0 \end{array}$$

is $x_1 = x_2 = \dots = x_n$.

The crux is to show that $x_1x_2\dots x_n = 0$ from which it follows that one of the variables is zero which, by symmetry can be taken to be x_n and then apply induction.

For the former, consult an abstract algebra book in which it is proven that the homogeneous polynomials $x_1 + \dots + x_n, x_1^2 + \dots + x_n^2, \dots, x_1^n + \dots + x_n^n$ generate the ring of homogeneous polynomials in $x_1, x_2, \dots, x_n$. In particular, $x_1x_2\dots x_n$.

10. Let $\mathcal{B}$ be a basis for V. Set $A = \mathcal{M}_T(\mathcal{B}, \mathcal{B})$. We need to prove that A is the $n \times n$ zero matrix. Now $Trace(\mathcal{M}_S(\mathcal{B}, \mathcal{B})A) = 0$ for every operator S on V. However, $\mathcal{M}_S(\mathcal{B}, \mathcal{B})$ ranges over all the matrices in $M_{nn}(\mathbb{F})$. Let E_{ij} be the matrix which has one non-zero entry, a one in the $(i, j)-$position. Let a_{ij} be the entry of A in the $(i, j)-$position. If $k \neq i$ then the k^{th} row of $E_{ij}A$ is zero, whereas the i^{th} row consists of the j^{th} row of A. and therefore $Trace(E_{ij}A) = a_{ji}$. Thus, for all $i, j, a_{ij} = 0$ and A is the zero matrix as required.

11. Since all the eigenvalues of A are real there exists a non-singular matrix Q such that $Q^{-1}AQ$ is upper triangular. Assume the diagonal entries of $Q^{-1}AQ$ are $a_1, \dots, a_n$. Then the diagonal entries of $(Q^{-1}AQ)^2 = Q^{-1}A^2Q$ are $a_1^2, \dots, a_n^2$. Then $Trace(A^2) = Trace(Q^{-1}A^2Q) = a_1^2 + \dots + a_n^2 \geq 0$.

12. Since $T^2 = T$ the minimum polynomial of T divides $x^2 - x$ and all the eigenvalues of A are 0 or 1. Let $\mathcal{B}$ be a basis for V and set $A = \mathcal{M}_T(\mathcal{B}, \mathcal{B})$. There exists a non-singular matrix Q such that $Q^{-1}AQ$ is upper triangular with diagonal entries 0 or 1. Then $Tr(T) = Trace(A) = Trace(Q^{-1}AQ)$ is a sum of the diagonal entries, whence a non-negative integer.

13. Let $\mathcal{B}$ be an orthonormal basis of V and set $A = \mathcal{M}_T(\mathcal{B}, \mathcal{B})$. Then $\mathcal{M}_{T^*}(\mathcal{B}, \mathcal{B}) = A^{tr}$. Since A and A^{tr} have the same diagonal entries, $Trace(A) = Trace(A^{tr})$. Thus, $Tr(T) = Trace(A) = Trace(A^{tr}) = Tr(T^*)$.

14. Let $\mathcal{B}$ be an orthonormal basis of V and set $A = \mathcal{M}_T(\mathcal{B}, \mathcal{B})$. Then $\mathcal{M}_{T^*}(\mathcal{B}, \mathcal{B}) = \overline{A}^{tr}$. It now follows that $Tr(T^*) = Trace(\overline{A}^{tr}) = Trace(\overline{A}) = \overline{Trace(A)} = \overline{Tr(T)}$.

15. $Tr : \mathcal{L}(V, V) \to \mathbb{F}$ is a non-zero linear transformation and consequently it is onto the one-dimensional space $\mathbb{F}$. $sl(V) = Ker(Tr)$ and so is a subspace and by the rank-nullity theorem, $dim(sl(V)) = dim(\mathcal{L}(V, V)) - 1 = n^2 - 1$.

16. If $S = T^*T$ then S is a semi-positive operator which implies that it has real eigenvalues which are all non-negative and, since it is self-adjoint, there is a orthonormal basis $\mathcal{B}$ such that $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ is diagonal. Let

$A = \mathcal{M}_T(\mathcal{B}, \mathcal{B}) = diag\{a_1, \ldots, a_n\}$. Since a_j are eigenvalues all $a_i \geq 0$. Then $Tr(T) = Trace(A) = a_1 + \cdots + a_n \geq 0$. On the other hand, the trace is zero if and only if $a_1 = \cdots = a_n = 0$. However, this implies A is the zero matrix, whence T is the zero operator.

17. We do a proof by induction on $n = dim(V)$. If $n = 1$ there is nothing to do. So assume the result is true for operators on spaces of dimension $n-1$ and assume $dim(V) = n$. If $T = \mathbf{0}_{V \to V}$ then any basis works so we may assume $T \neq \mathbf{0}_{V \to V}$. We first claim that there is a vector $\boldsymbol{v}$ such that $T(\boldsymbol{v}) \notin Span(\boldsymbol{v})$. Otherwise, for each $\boldsymbol{v}$ there is $\lambda_{\boldsymbol{v}} \in \mathbb{F}$ such that $T(\boldsymbol{v}) = \lambda_{\boldsymbol{v}} \boldsymbol{v}$. We claim that $\lambda_{\boldsymbol{v}}$ is independent of the vector $\boldsymbol{v}$. Of course, if $\boldsymbol{w} = c\boldsymbol{v}$ then $T(\boldsymbol{w}) = cT(\boldsymbol{v}) = c\lambda_{\boldsymbol{v}}\boldsymbol{v} = \lambda_{\boldsymbol{v}}(c\boldsymbol{v}) = \lambda_{\boldsymbol{v}}\boldsymbol{w}$. Suppose on the other hand that $(\boldsymbol{v}, \boldsymbol{w})$ is linearly independent. Then $\lambda_{\boldsymbol{v}+\boldsymbol{w}}\boldsymbol{v} + \lambda_{\boldsymbol{v}+\boldsymbol{w}}\boldsymbol{w} = \lambda_{\boldsymbol{v}+\boldsymbol{w}}(\boldsymbol{v}+\boldsymbol{w}) = T(\boldsymbol{v}+\boldsymbol{w}) = T(\boldsymbol{v}) + T(\boldsymbol{w}) = \lambda_{\boldsymbol{v}}\boldsymbol{v} + \lambda_{\boldsymbol{w}}\boldsymbol{w}$. Then $\lambda_{\boldsymbol{v}} = \lambda_{\boldsymbol{v}+\boldsymbol{w}} = \lambda_{\boldsymbol{w}}$. Thus, $T = \lambda I_V$ for some $\lambda \in \mathbb{F}$. Then $Tr(T) = n\lambda = 0$. Since the characteristic of $\mathbb{F}$ is zero, $\lambda = 0$, contrary to our assumption that $T \neq \mathbf{0}_{V \to V}$.

Now choose $\boldsymbol{v}$ such that $T(\boldsymbol{v}) \notin Span(\boldsymbol{v})$ and set $\boldsymbol{v}_1 = \boldsymbol{v}$ and $\boldsymbol{v}_2 = T(\boldsymbol{v})$ and extend to a basis $\mathcal{B}_1 = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)$. Let $T(\boldsymbol{v}_j) = \sum_{i=1}^n a_{ij}\boldsymbol{v}_i$ and set $\boldsymbol{a}_j = [T(\boldsymbol{v}_j)]_{\mathcal{B}}$. Note that $\boldsymbol{a}_1 = \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Since $Tr(T) = 0$ we must have $\sum_{i=2}^n a_{ii} = 0$. Now define an operator S on V such that $S(\boldsymbol{v}_1) = \boldsymbol{v}_1$ and $S(\boldsymbol{v}_j) = \sum_{i=2}^n a_{ij}\boldsymbol{v}_i$. Note that $W = Span(\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)$ is S-invariant and that $Tr(S_{|W}) = \sum_{i=2}^n a_{ii} = 0$. Set $S' = S_{|W}$. By the inductive hypothesis, there is a basis $\mathcal{B}_W = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{n-1})$ for W such that the diagonal entries of $\mathcal{M}_{S'}(\mathcal{B}_W, \mathcal{B}_W)$ are all zero. Set $\boldsymbol{u}_1 = \boldsymbol{v}_1, \boldsymbol{u}_j = \boldsymbol{w}_{j-1}$ for $2 \leq j \leq n$ and $\mathcal{B} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$. Since $T(\boldsymbol{v}_j) - S(\boldsymbol{v}_j) \in Span(\boldsymbol{v}_1)$ for $2 \leq j \leq n$ it follows that $T(\boldsymbol{u}_j) - S(\boldsymbol{u}_j) \in Span(\boldsymbol{v}_1)$ for $2 \leq i \leq n$. Consequently, for $2 \leq j \leq n$, $[T(\boldsymbol{u}_j)]_{\mathcal{B}} - [S(\boldsymbol{u}_j)]_{\mathcal{B}} \in Span(\boldsymbol{e}_1)$ for $2 \leq j \leq n$ where $\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

However, the (j,j)-entry of $[S(\boldsymbol{u}_j)]_{\mathcal{B}}$ is zero, whence the (j,j)-entry of $[T(\boldsymbol{u}_j)]_{\mathcal{B}}$ is zero. Thus, all the diagonal entries of $\mathcal{M}_T(\mathcal{B}, \mathcal{B})$ are zero as required.

18. Let Z be the set of $n \times n$ matrices all of whose diagonal entries are zero. Then Z is a subspace of $M_{nn}(\mathbb{F})$ of dimension $n^2 - n$. For a matrix B define a map $ad(B) : M_{nn}(\mathbb{F}) \to M_{nn}(\mathbb{F})$ by $ad(B)(C) = BC - CB$. This is a linear map and $Ker(ad(B)) = C(B)$ the subalgebra of $M_{nn}(\mathbb{F})$ which commutes with B. Note that if the minimum polynomial of B has degree n then $dim(C(B)) = n$ and $Range(ad(B))$ has dimension $n^2 - n$.

Now let $a_1, \ldots, a_n$ be distinct elements of $\mathbb{F}$ and let B be the diagonal matrix with entries $a_1, \ldots, a_n$. Then the minimum polynomial of B is $(x - a_1) \ldots (x - a_n)$ has degree n and therefore $dim(Range(ad(B)) = n^2 - n$. On the other hand, for any matrix C the diagonal entries of $BC - CB$ are all zero. Thus, $Range(ad(B)) = Z$. Thus, for $A \in Z$ then there is a matrix C such that $A = BC - CB$.

19. By Exercise 17 there exists a basis $\mathcal{B}$ such that $A = \mathcal{M}_T(\mathcal{B}, \mathcal{B})$ has diagonal entries all zero. By Exercise 18 there are matrices B, C such that $A = BC - CB$. Let $R, S \in \mathcal{L}(V, V)$ such that $\mathcal{M}_R(\mathcal{B}, \mathcal{B}) = B, \mathcal{M}_S(\mathcal{B}, \mathcal{B}) = S$. Then $T = RS - SR$.

7.2. Determinants

1. Let B be the matrix obtained from A by exchanging the first and i^{th} rows. Let the entries of A be a_{kl} and the entries of B be b_{kl}. Let A_{kl} be the matrix obtained from A by deleting the k^{th} row and l^{th} column, with B_{kl} defined similarly. Set $M_{kl} = det(A_{kl}), M'_{kl} = det(B_{kl}), C_{kl} = (-1)^{k+l}M_{kl}$ and $C'_{kl} = (-1)^{k+l}M'_{kl}$.

Since B is obtained from A by exchanging the first and i^{th} rows, $det(B) = -det(A)$. By Exercise 1

$$det(B) = b_{11}C'_{11} + \cdots + b_{1n}C'_{1n}.$$

Note that $b_{1j} = a_{ij}$. Also, the matrix B_{1j} is obtained from the matrix A_{ij} by moving the first row to the $(i-1)^{st}$ row. Note that this can be obtained by exchanging the first and second row of A_{1j} then exchanging the second and third rows of that matrix, and continuing until we exchange the $(i-2)^{nd}$ and $(i-1)^{st}$ rows. Thus, there are $i-2$ exchanges which implies that $det(B_{1j}) = (-1)^{i-2}det(A_{ij}) = (-1)^i det(A_{ij})$. Thus,

$$M'_{1j} = (-1)^i M_{ij}.$$

It then follows that

$$C'_{1j} = (-1)^{1+j}M'_{1j} =$$

$$(-1)^{1+j}(-1)^i M_{ij} = (-1)^{1+i+j}M_{ij} = -C_{ij}.$$

Putting this together we get

$$-det(A) = det(B) = a_{i1}(-C_{i1}) + \cdots + a_{in}(-C_{in})$$

Multiplying by -1 we get

$$det(A) = a_{i1}C_{i1} + \cdots + a_{in}C_{in}$$

2. Let $B = A^{tr}$. Denote the $(i,j)-$entry of B by b_{ij} and the $(i,j)-$cofactor by C'_{ij}. By Exercise 3, $det(B) = det(A)$. By exercise 2 for any j

$$det(B) = b_{j1}C'_{j1} + b_{j2}C'_{j2} + \cdots + b_{jn}C'_{jn}.$$

Since B is the transpose of $A, b_{ji} = a_{ij}$ and $C'_{ji} = C_{ij}$. Thus,

$$det(A) = det(B) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

3. Let $\mathcal{B}$ be an orthonormal basis of V. Then $det(T) = det(\mathcal{M}_T(\mathcal{B},\mathcal{B}))$ and $det(A^*) = det(\mathcal{M}_{T^*}(\mathcal{B},\mathcal{B}))$. However, if $A = \mathcal{M}_T(\mathcal{B},\mathcal{B})$ then $\mathcal{M}_{T^*}(\mathcal{B},\mathcal{B}) = \overline{A}^{tr}$. Then $det(T^*) = det(\overline{A}^{tr}) = det(\overline{A}) = \overline{det(T)}$.

4. If $\boldsymbol{v} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$ then $J_n\boldsymbol{v}$ is the vector all of whose entries are equal to $c_1+\cdots+c_n$. Thus, $J_n j_n = \begin{pmatrix} n \\ n \\ \vdots \\ n \end{pmatrix} = nj_n$ and j_n is an eigenvector with eigenvalue n. This proves i.

If follows from the previous paragraph that $J_n\boldsymbol{v}_i = \boldsymbol{0}$. The sequence $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{n-1})$ is linearly independent and spans a subspace of dimension $n-1$ contained in $null(J_n)$. Since $null(J_n)$ is a proper subspace of $\mathbb{R}^n$, $Span(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{n-1}) = null(J_n)$. This proves ii).

iii) Since $j_n \notin Span(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{n-1})$ the sequence $\mathcal{B} = (\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{n-1},j_n)$ is linearly independent. Since there are n vectors and $dim(\mathbb{R}^n) = n$, $\mathcal{B}$ is a basis for $\mathbb{R}^n$.

5. $Aj_n = (aI_n + bJ_n)j_n =$

$$(aI_n)j_n + (bJ_n)j_n = aj_n + (bn)j_n = [a+bn]j_n.$$

Thus, j_n is an eigenvector of A with eigenvalue $a+bn$.

On the other hand, $A\boldsymbol{v}_i = (aI_n + bJ_n)\boldsymbol{v}_n =$

$$(aI_n)\boldsymbol{v}_i + (bJ_n)\boldsymbol{v}_i = a\boldsymbol{v}_i.$$

This shows that each $\boldsymbol{v}_i$ is an eigenvector with eigenvalue a. Thus, $\mathcal{B} = (\boldsymbol{v}_1,\ldots,\boldsymbol{v}_{n-1},j_n)$ is a basis of eigenvectors for A with eigenvalues a with multiplicity $n-1$ and $a+bn$ with multiplicity 1. This implies that A is similar to the diagonal matrix with $n-1$ entries equal to a and one entry equal to $a+bn$. The determinant of this diagonal matrix is the product of the diagonal entries and is equal to $a^{n-1}(a+bn)$. Since A is similar to this diagonal matrix, $det(A)=a^n(a+bn)$.

6. We proceed by induction starting with $n = 2$ as the base case. Direct computation shows that $det(\begin{pmatrix} 1 & 1 \\ \alpha_1 & \alpha_2 \end{pmatrix}) = \alpha_2 - \alpha_1$. Thus, assume the result holds for $n-1$. We will use properties of determinants of matrices to compute the determinant of

$$\begin{pmatrix} 1 & 1 & \dots & 1 & \\ \alpha_1 & \alpha_2 & \dots & \alpha_n & \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 & \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \dots & \alpha_n^{n-1} & \end{pmatrix}$$

We begin working with the transpose

$$\begin{pmatrix} 1 & \alpha_1 & \dots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \dots & \alpha_2^{n-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & \alpha_n & \dots & \alpha_n^{n-1} \end{pmatrix}.$$

Subtract the first row from all the other rows to get the following matrix which has the same determinant:

$$\begin{pmatrix} 1 & \alpha_1 & \dots & \alpha_1^{n-1} \\ 0 & \alpha_2-\alpha_1 & \dots & \alpha_2^{n-1}-\alpha_1^{n-1} \\ \vdots & \vdots & \dots & \vdots \\ 0 & \alpha_n-\alpha_1 & \dots & \alpha_n^{n-1}-\alpha_1^{n-1} \end{pmatrix}$$

By Exercise 2 we can use a cofactor in the first column to compute the determinant. Since the only non-zero entry is in the (1,1)-position the determinant is equal to the determinant of the following $(n-1)\times(n-1)-$matrix

$$\begin{pmatrix} \alpha_2-\alpha_1 & \alpha_2^2-\alpha_1^2 & \dots & \alpha_2^{n-1}-\alpha_1^{n-1} \\ \alpha_3-\alpha_1 & \alpha_3^2-\alpha_1^2 & \dots & \alpha_3^{n-1}-\alpha_1^{n-1} \\ \vdots & \vdots & \dots & \vdots \\ \alpha_n-\alpha_1 & \alpha_n^2-\alpha_1^2 & \dots & \alpha_n^{n-1}-\alpha_1^{n-1} \end{pmatrix}$$

We can factor $\alpha_k-\alpha_1$ from $(k-1)^{st}$ row to get the determinant is
$\prod_{k=2}^n(\alpha_k-\alpha_1)$ times the determinant of

$$\begin{pmatrix} 1 & \alpha_2+\alpha_1 & \dots & \alpha_2^{n-2}+\alpha_2^{n-3}\alpha_1+\cdots+\alpha_1^{n-2} \\ 1 & \alpha_3+\alpha_1 & \dots & \alpha_3^{n-2}+\alpha_3^{n-3}\alpha_1+\cdots+\alpha_1^{n-2} \\ \vdots & \vdots & \dots & \vdots \\ 1 & \alpha_n+\alpha_1 & \dots & \alpha_n^{n-2}+\alpha_n^{n-3}\alpha_1+\cdots+\alpha_1^{n-2} \end{pmatrix}$$

We need to compute the determinant of the matrix

$$\begin{pmatrix} 1 & \alpha_2+\alpha_1 & \dots & \alpha_2^{n-2}+\alpha_2^{n-3}\alpha_1+\cdots+\alpha_1^{n-2} \\ 1 & \alpha_3+\alpha_1 & \dots & \alpha_3^{n-2}+\alpha_3^{n-3}\alpha_1+\cdots+\alpha_1^{n-2} \\ \vdots & \vdots & \dots & \vdots \\ 1 & \alpha_n+\alpha_1 & \dots & \alpha_n^{n-2}+\alpha_n^{n-3}\alpha_1+\cdots+\alpha_1^{n-2} \end{pmatrix}$$

Take its transpose. Then subtract α_1 times the $(n-2)^{nd}$ row from the $(n-1)^{st}$ row, the α_1 times the $(n-3)^{rd}$ row from the $(n-2)^{st}$ row and continue, subtracting α_1 times the second row from the third row and α_1 times the first row from the second row. The matrix obtained is

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha_2 & \alpha_3 & \dots & \alpha_n \\ \alpha_2^2 & \alpha_3^2 & \dots & \alpha_n^2 \\ \vdots & \vdots & \dots & \vdots \\ \alpha_2^{n-2} & \alpha_3^{n-2} & \dots & \alpha_n^{n-2} \end{pmatrix}$$

By the induction hypothesis, the determinant of this matrix is $\prod_{2\le j<k\le n}(\alpha_k-\alpha_j)$. Multiplying this by $\prod_{1<j\le n}(\alpha_j-\alpha_1)$ we obtain the desired formula.

7. Construct the matrix A' from the matrix A by replacing the i^{th} row with the j^{th} row. Note that the $(i,k)-$cofactor of A' is the same as the $(i,k)-$cofactor of A which we denote by C_{ik}.

Since two rows of A' are identical, $det(A')=0$. On the other hand, computing the determinant of A' using the cofactor expansion in the i^{th} row of A' we get

$$0 = det(A') = a_{j1}C_{i1} + a_{j2}C_{i2} + \cdots + a_{jn}C_{in}$$

8. The $(i,i)-$entry of $AAdj(A)$ is equal to

$$a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

By Exercise 2 this is $det(A)$. On the other hand if $i \neq j$ then the $(i,j)-$entry is

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \cdots + a_{in}C_{jn}$$

which is zero by Exercise 7.

Putting these together we have $AAdj(A) = det(A)I_n$.

9. Since the entries of A and A^{-1} are integers we have $det(A), det(A^{-1})$ are integers. Since $det(A)det(A^{-1}) = 1$ either $det(A) = det(A^{-1}) = 1$ of $det(A) = det(A^{-1}) = -1$.

10. Since A is an integer matrix, $Adj(A)$ is an integer matrix. By Exercise 9, $AAdj(A) = det(A)I_n = \epsilon I_n$ with $\epsilon \in \{-1,1\}$. Set $B = \epsilon Adj(A)$, an integer matrix. Then $AB = I_n$ from which it follows that $BA = I_n$ and $B = A^{-1}$.

11. If T is a Hermitian operator then there exists a basis $\mathcal{B}$ of V such that $\mathcal{M}_T(\mathcal{B},\mathcal{B})$ is diagonal with real entries. Then $det(T) = det(\mathcal{M}_T(\mathcal{B},\mathcal{B}))$ is a real number.

12. T^*T is a positive operator, whence diagonalizable with non-negative real eigenvalues. Therefore $det(T^*T)$ is non-negative. If T is not invertible then $det(T^*T) = 0$. On the other hand, if T is invertible then T^*T is invertible and hence $det(T^*T) \neq 0$ whence $det(T^*T) > 0$.

13. Let $\mathcal{B}$ be an orthonormal basis for V and set $A = \mathcal{M}_T(\mathcal{B},\mathcal{B})$. Then A is an orthogonal matrix, that is, $A^{-1} = A^{tr}$. Then

$$1 = det(I_n) = det(AA^{-1}) = det(AA^{tr}) =$$

$$det(A)det(A^{tr}) = det(A)^2.$$

Thus, $det(T) = det(A) \in \{-1,1\}$.

14. Let $\mathcal{B}$ be an orthonormal basis of V and set $A = \mathcal{M}_T(\mathcal{B},\mathcal{B})$. Then A is a unitary matrix, that is, $A^{-1} = \overline{A}^{tr}$. Then $det(A^{-1}) = \overline{det(A)}$. Then $1 = det(I_n) = det(AA^{-1}) = det(A)det(A^{-1}) = det(A)\overline{det(A)} = \| det(A) \|^2$. Thus, $\| det(T) \| = \| det(A) \| = 1$.

15. Assume $dim(V) = n = 2k+1$. Recall this means that $T^* = -T$. Note that for a scalar $c, det(cT) = c^n det(T)$. Therefore $det(T) = det(T^*) = det(-T) = (-1)^{2k+1}det(T) = -det(T)$. Thus, $det(T) = 0$.

16. The assumption implies that the columns of the matrix are linearly dependent:

$$\boldsymbol{v}_1 - \boldsymbol{v}_2 + \boldsymbol{v}_3 - \cdots - \boldsymbol{v}_{2k} + \boldsymbol{v}_{2k+1} = \mathbf{0}.$$

Thus, the matrix is not invertible and $det(A) = 0$.

17. Let A be such a matrix. Since we are not concerned about signs, by multiplying the first row by -1, if necessary, we can assume that the (1,1)-entry is 1. Use Gaussian elimination to make all the other entries in the first column zero and denote this matrix by B. Then $det(B) = \pm det(A)$. Let b_{ij} denote the entries of B. Then for $i,j \geq 2, b_{ij} \in \{-2,0,2\}$. Write $b_{ij} = 2c_{ij}$ where $c_{ij} \in \{-1,0,1\}$ so that

$$B = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 2c_{22} & 2c_{23} & \dots & 2c_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 2c_{n2} & 2c_{n3} & \dots & 2c_{nn} \end{pmatrix}$$

Then $det(B) = 2^{n-1}det(C)$ where

$$C = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & c_{22} & c_{23} & \dots & c_{2n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & c_{n2} & c_{n3} & \dots & c_{nn} \end{pmatrix}$$

Since C is an integer matrix $det(C)$ is an integer.

18. Since at most one row has no zeros we can assume that all rows below the first have at least one zero and therefore there at least $n-1$ zeros. On the other hand we claim that the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 1 & \dots & 0 \end{pmatrix}$$

is invertible.

After subtracting the first row from each subsequent row we get the matrix

$$B = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & -1 \end{pmatrix}$$

Then $det(A) = det(B) = \pm 1$.

19. Player 2 should apply the following strategy: if Player 1 places an entry in position (i, j) with $i > 2$ then Player 2 places an arbitrary entry in a position (k, l) with $k > 2$. If Player 1 puts a number b in position (i, j) with $i \in \{1, 2\}$ then Player 2 puts the same number b in position (k, j) where $\{i, k\} = \{1, 2\}$. When the matrix is filled the first and second rows will be identical and the matrix will have determinant zero.

20. Taking determinants we get $det(AB) = det(A)det(B) = det(BA)$. On the other hand $det(-BA) = (-1)^{2k1+1}det(BA) = -det(BA)$. Therefore $det(AB) = -det(AB)$ and so is zero. Then either $det(A) = 0$ or $det(B) = 0$ so either A is not invertible or B is not invertible.

7.3. Uniqueness of the Determinant

1. The proof is by induction on $j - i$. If $j - i = 1$ this follows from the definition of an alternating form. Assume the result for t and assume that $\boldsymbol{u}_i = \boldsymbol{u}_j$ with $j - i = t+1$. By Lemma (7.10)

$$f(\boldsymbol{u}_1, \dots, \boldsymbol{u}_i, \dots, \boldsymbol{u}_{j-1}, \boldsymbol{u}_j, \dots, \boldsymbol{u}_m) =$$
$$-f(\boldsymbol{u}_1, \dots, \boldsymbol{u}_i, \dots, \boldsymbol{u}_j, \boldsymbol{u}_{j-1} \dots, \boldsymbol{u}_m)$$

However, by the inductive hypothesis $f(\boldsymbol{u}_1, \dots, \boldsymbol{u}_i, \dots, \boldsymbol{u}_j, \boldsymbol{u}_{j-1} \dots, \boldsymbol{u}_m) = 0$.

2. We previously proved that every invertible matrix is a product of elementary matrices. This implies that every invertible operator is a product of elementary operators.

3. We demonstrate this for $m = 2$ the proof of general m is similar. Let $f, g \in \mathcal{L}(V^2, W)$. We need to show that $f + g \in \mathcal{L}(V^2, W)$. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{u} \in V$. Then

$$(f + g)(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{u}) = f(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{u}) + g(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{u}) =$$
$$[f(\boldsymbol{v}_1, \boldsymbol{u}) + f(\boldsymbol{v}_2, \boldsymbol{u})] + [g(\boldsymbol{v}_1, \boldsymbol{u}) + g(\boldsymbol{v}_2, \boldsymbol{u})] =$$
$$[f(\boldsymbol{v}_1, \boldsymbol{u}) + g(\boldsymbol{v}_1, \boldsymbol{u})] + [f(\boldsymbol{v}_2, \boldsymbol{u}) + g(\boldsymbol{v}_2, \boldsymbol{u})] =$$
$$(f + g)(\boldsymbol{v}_1, \boldsymbol{u}) + (f + g)(\boldsymbol{v}_2, \boldsymbol{u})$$

That $(f+g)(\boldsymbol{u}, \boldsymbol{v}_1+\boldsymbol{v}_2) = (f+g)(\boldsymbol{u}, \boldsymbol{v}_1)+(f+g)(\boldsymbol{u}, \boldsymbol{v}_2)$ is proved in exactly the same way.

Now let $\boldsymbol{u}, \boldsymbol{v} \in V$ and c is a scalar. We need to prove $(f + g)(c\boldsymbol{u}, \boldsymbol{v}) = (f + g)(\boldsymbol{u}, \boldsymbol{v}) = c(f + g)(\boldsymbol{u}, \boldsymbol{v})$.

$$(f + g)(c\boldsymbol{u}, \boldsymbol{v}) = f(c\boldsymbol{u}, \boldsymbol{v}) + g(c\boldsymbol{u}, \boldsymbol{v}) =$$
$$cf(\boldsymbol{u}, \boldsymbol{v}) + cg(\boldsymbol{u}, \boldsymbol{v}) = c[f(\boldsymbol{u}, \boldsymbol{v}) + g(\boldsymbol{u}, \boldsymbol{v})] =$$
$$c[(f + g)(\boldsymbol{u}, \boldsymbol{v})] = c(f + g)(\boldsymbol{u}, \boldsymbol{v})$$

Likewise

$$(f + g)(\boldsymbol{u}, c\boldsymbol{v}) = f(\boldsymbol{u}, c\boldsymbol{v}) + g(c\boldsymbol{u}, c\boldsymbol{v}) =$$
$$cf(\boldsymbol{u}, \boldsymbol{v}) + cg(\boldsymbol{u}, \boldsymbol{v}) = c[f(\boldsymbol{u}, \boldsymbol{v}) + g(\boldsymbol{u}, \boldsymbol{v})] =$$
$$c[(f + g)(\boldsymbol{u}, \boldsymbol{v})] = c(f + g)(\boldsymbol{u}, \boldsymbol{v})$$

Now assume $f \in \mathcal{L}(V^2, W)$ and c is a scalar. We need to prove that $cf \in \mathcal{L}(V^2, W)$. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{u} \in V$. Then

$$(cf)(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{u}) = c[f(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{u})] =$$
$$c[f(\boldsymbol{v}_1, \boldsymbol{u}) + f(\boldsymbol{v}_2, \boldsymbol{u})] = cf(\boldsymbol{v}_1, \boldsymbol{u}) + cf(\boldsymbol{v}_2, \boldsymbol{u}) =$$
$$(cf)(\boldsymbol{v}_1, \boldsymbol{u}) + (cf)(\boldsymbol{v}_2, \boldsymbol{u}).$$

In exactly the same way we prove that $(cf)(\boldsymbol{u}, \boldsymbol{v}_1+\boldsymbol{v}_2) = (cf)(\boldsymbol{u}, \boldsymbol{v}_1) + (cf)(\boldsymbol{u}, \boldsymbol{v}_2)$.

Finally, we need to show that if d is a scalar and $\boldsymbol{u}, \boldsymbol{v} \in V$ then $(cf)(d\boldsymbol{u}, \boldsymbol{v}) = (cf)(\boldsymbol{u}, d\boldsymbol{v}) = d[(cf)(\boldsymbol{u}, \boldsymbol{v})]$.

$$(cf)(d\boldsymbol{u}, \boldsymbol{v}) = c[f(d\boldsymbol{u}, \boldsymbol{v})] = c[df(\boldsymbol{u}, \boldsymbol{v})] =$$
$$(cd)f(\boldsymbol{u}, \boldsymbol{v}) = (dc)f(\boldsymbol{u}, \boldsymbol{v}) = d[cf(\boldsymbol{u}, \boldsymbol{v})] =$$
$$d[(cf)(\boldsymbol{u}, \boldsymbol{v})].$$

Similarly,

$$(cf)(\boldsymbol{u}, d\boldsymbol{v}) = c[f(\boldsymbol{u}, d\boldsymbol{v})] = c[df(\boldsymbol{u}, \boldsymbol{v})] =$$
$$(cd)f(\boldsymbol{u}, \boldsymbol{v}) = (dc)f(\boldsymbol{u}, \boldsymbol{v}) = d[cf(\boldsymbol{u}, \boldsymbol{v})] =$$
$$d[(cf)(\boldsymbol{u}, \boldsymbol{v})].$$

4. Again we prove this for $m = 2$. The general case is similar. We need to prove the following: i) If $f, g \in Alt(V^2, W)$ and $\boldsymbol{u} \in V$ then $(f+g)(\boldsymbol{u}, \boldsymbol{u}) = \boldsymbol{0}_W$; and ii) If $f \in Alt(V^2, W)$, c is a scalar and $\boldsymbol{u} \in V$ then $(cf)(\boldsymbol{u}, \boldsymbol{u}) = \boldsymbol{0}_W$.

i) $(f+g)(\boldsymbol{u}, \boldsymbol{u}) = f(\boldsymbol{u}, \boldsymbol{u}) + g(\boldsymbol{u}, \boldsymbol{u})$. Since $f, g \in Alt(V^2, W)$ we have

$$f(\boldsymbol{u}, \boldsymbol{u}) + g(\boldsymbol{u}, \boldsymbol{u}) = \boldsymbol{0}_W + \boldsymbol{0}_W = \boldsymbol{0}_W.$$

ii) $(cf)(\boldsymbol{u}, \boldsymbol{u}) = c[f(\boldsymbol{u}, \boldsymbol{u})] = c\boldsymbol{0}_W = \boldsymbol{0}_W$.

5. Assume $m > n = dim(V)$ and $f \in Alt(V^m, W)$. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ be a sequence from V. Since m is greater than the dimension of V the sequence is linearly dependent. By Lemma (7.11) $f(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m) = \boldsymbol{0}_W$.

6. We prove this for $(i, j) = (1, 2)$. We first show that f_{12} is alternating:

$$f_{12}(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}, \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}) = a_1a_2 - a_2a_1 = 0.$$

We prove additive in the first varialbe:

$$f_{12}(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix}) =$$
$$(a_1 + b_1)c_2 - (a_2 + b_2)c_1 =$$
$$(a_1c_2 + b_1c_2) - (a_2c_1 + b_2c_1) =$$
$$(a_1c_2 - a_2c_1) + (b_1c_2 - b_2c_1) =$$

$$f_{12}(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix}) + f_{12}(\begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix})$$

A similar argument shows that f_{12} is additive in the second argument.

We now prove that f_{12} has the scalar property in the first variable

$$f_{12}(c\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}) = f_{12}(\begin{pmatrix} ca_1 \\ ca_2 \\ ca_3 \\ ca_4 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix}) =$$

$$(ca_1)b_2 - (ca_2)b_1 = c(a_1b_2) - c(a_2b_1) =$$

$$c[a_1b_2 - a_2b_1] = cf_{12}(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix})$$

The scalar property in the second variable is proved similarly.

7. Let $\boldsymbol{e}_i$ denote the i^{th} standard basis vector of $\mathbb{F}^4$ and set $\boldsymbol{e}_{ij} = (\boldsymbol{e}_i, \boldsymbol{e}_j)$ for $1 \leq i, j \leq 4$. Then any alternating form $f \in Alt(V^2, \mathbb{F})$ is uniquely determined by its values on $\boldsymbol{e}_{ij}$. Next note that $f_{ij}(\boldsymbol{e}_{kl}) = \delta_{ik}\delta_{jl}$. This implies that $(f_{12}, \dots, f_{34})$ is linearly independent: Suppose $\sum_{1 \leq i < j \leq 4} a_{ij} f_{ij} = 0$. Evaluating at $\boldsymbol{e}_{kl}$ we find that $a_{kl} = 0$.

We next show that $(f_{12}, \dots, f_{34})$ spans $\mathcal{L}(V^2, \mathbb{F})$. Let $f \in Alt(V^2, \mathbb{F})$ and let a_{ij} be equal to $f(\boldsymbol{e}_{ij})$. Set $g = a_{12}f_{12} + \cdots + a_{34}f_{34}$. Then $g \in Alt(V^2, \mathbb{F})$. Moreover, $g(\boldsymbol{e}_{ij}) = a_{ij} = f(\boldsymbol{e}_{ij})$ which implies that $f = g$.

8. This follows since the determinant is an alternating map of the columns of the matrix.

9. Again let $\boldsymbol{e}_i$ be the i^{th} standard basis vector. Let $\boldsymbol{e}_{ijk}$ denote the ordered triple $(\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k)$ where $1 \leq i < j < k \leq 4$. Then any alternating form $f \in Alt(V^3, \mathbb{F})$ is uniquely determined by its values on $(\boldsymbol{e}_{123}, \boldsymbol{e}_{124}, \boldsymbol{e}_{134}, \boldsymbol{e}_{234})$.

Next note that $g_l(\boldsymbol{e}_{ijk}) = 1$ if $\{i, j, k, l\} = \{1, 2, 3, 4\}$ and is zero otherwise. As in Exercise 8 this implies that (g_1, g_2, g_3, g_4) is linearly independent. We show that it spans $Alt(V^3, \mathbb{F})$.

Let $f \in Alt(V^3, \mathbb{F})$ and set $a_{ijk} = f(\boldsymbol{e}_{ijk})$ and $g = a_{234}g_1 + a_{134}g_2 + a_{124}g_3 + a_{123}g_4$. Then $g(\boldsymbol{e}_{ijk}) = a_{ijk}$ and therefore $g = f$. Thus, (g_1, g_2, g_3, g_4) spans $Alt(V^3, \mathbb{F})$ and so is a basis.

Chapter 8

Bilinear Maps and Forms

8.1. Basic Properties of Bilinear Maps

1. Let $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V, \boldsymbol{w} \in W$. Since each f_i is bilinear we have $f_i(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) = f_i(\boldsymbol{v}_1, \boldsymbol{w}) + f_i(\boldsymbol{v}_2, \boldsymbol{w})$. We then have

$$F(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) = \sum_{i=1}^{s} f_i(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) =$$

$$\sum_{i=1}^{s} [f_i(\boldsymbol{v}_1, \boldsymbol{w}) + f_i(\boldsymbol{v}_2, \boldsymbol{w})] =$$

$$\sum_{i=1}^{s} f_i(\boldsymbol{v}_1, \boldsymbol{w}) + \sum_{i=1}^{s} f_i(\boldsymbol{v}_2, \boldsymbol{w}) =$$

$$F(\boldsymbol{v}_1, \boldsymbol{w}) + F(\boldsymbol{v}_2, \boldsymbol{w})$$

For $\boldsymbol{v} \in V, \boldsymbol{w}_1, \boldsymbol{w}_2 \in W$ that $F(\boldsymbol{v}, \boldsymbol{w}_1 + \boldsymbol{w}_2) = F(\boldsymbol{v}, \boldsymbol{w}_1) + F(\boldsymbol{v}, \boldsymbol{w}_2)$ is proved in exactly the same way.

Now let $\boldsymbol{v} \in V, \boldsymbol{w} \in W$ and $c \in \mathbb{F}$. Since each f_i is bilinear, $f_i(c\boldsymbol{v}, \boldsymbol{w}) = cf_i(\boldsymbol{v}, \boldsymbol{w}) = f_i(\boldsymbol{v}, c\boldsymbol{w})$. We then have

$$F(c\boldsymbol{v}, \boldsymbol{w}) = \sum_{i=1}^{s} f_i(c\boldsymbol{v}, \boldsymbol{w}) = \sum_{i=1}^{s} cf_i(\boldsymbol{v}, \boldsymbol{w}) =$$

$$c \sum_{i=1}^{s} f_i(\boldsymbol{v}, \boldsymbol{w}) = cF(\boldsymbol{v}, \boldsymbol{w}).$$

$$F(\boldsymbol{v}, c\boldsymbol{w}) = \sum_{i=1}^{s} f_i(\boldsymbol{v}, c\boldsymbol{w}) = \sum_{i=1}^{s} cf_i(\boldsymbol{v}, \boldsymbol{w}) =$$

$$c \sum_{i=1}^{s} f_i(\boldsymbol{v}, \boldsymbol{w}) = cF(\boldsymbol{v}, \boldsymbol{w}).$$

2. Let $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{F}^m, \boldsymbol{w} \in \mathbb{F}^n$. Then

$$f(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) = (\boldsymbol{v}_1 + \boldsymbol{v}_2)^{tr} A\boldsymbol{w} = (\boldsymbol{v}_1^{tr} + \boldsymbol{v}_2^{tr}) A\boldsymbol{w} =$$

$$\boldsymbol{v}_1^{tr} A\boldsymbol{w} + \boldsymbol{v}_2^{tr} A\boldsymbol{w} = f(\boldsymbol{v}_1, \boldsymbol{w}) + f(\boldsymbol{v}_2, \boldsymbol{w})$$

If $\boldsymbol{v} \in \mathbb{F}^m, \boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathbb{F}^n$ then

$$f(\boldsymbol{v}, \boldsymbol{w}_1 + \boldsymbol{w}) = \boldsymbol{v}^{tr} A(\boldsymbol{w}_1 + \boldsymbol{w}_2) =$$

$$\boldsymbol{v}^{tr} A\boldsymbol{w}_1 + \boldsymbol{v}^{tr} A\boldsymbol{w}_2 = f(\boldsymbol{v}, \boldsymbol{w}_1) + f(\boldsymbol{v}, \boldsymbol{w}_2).$$

Assume $\boldsymbol{v} \in \mathbb{F}^m, \boldsymbol{w} \in \mathbb{F}^n$ and $c \in \mathbb{F}$ then

$$f(c\boldsymbol{v}, \boldsymbol{w}) = (c\boldsymbol{v})^{tr} A\boldsymbol{w} = c\boldsymbol{v}^{tr} A\boldsymbol{w} =$$

$$c(\boldsymbol{v}^{tr} A\boldsymbol{w}) = cf(\boldsymbol{v}, \boldsymbol{w})$$

$$f(\boldsymbol{v}, c\boldsymbol{w}) = \boldsymbol{v}^{tr} A(c\boldsymbol{w}) = c(\boldsymbol{v}^{tr} A\boldsymbol{w}) = cf(\boldsymbol{v}, \boldsymbol{w}).$$

3. By additivity in the first variable we have

$$f(\boldsymbol{v}, \boldsymbol{w}) = f(\sum_{i=1}^{m} c_i \boldsymbol{v}_i, \sum_{j=1}^{n} d_j \boldsymbol{w}_j) =$$

$$\sum_{i=1}^{m} f(c_i \boldsymbol{v}_i, \sum_{j=1}^{n} d_j \boldsymbol{w}_j)$$

By additivity in the second variable we have

$$= \sum_{i=1}^{m} (\sum_{j=1}^{n} f(c_i \boldsymbol{v}_i, d_j \boldsymbol{w}_j)$$

By homogeneity we have $f(c_i\boldsymbol{v}_i, d_j\boldsymbol{w}_j) = (c_id_j)f(\boldsymbol{v}_i, \boldsymbol{w}_j) = (c_id_j)a_{ij}$ and consequently

$$f(\boldsymbol{v}, \boldsymbol{w}) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_i d_j a_{ij} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}^{tr} A \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}$$

4. Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ be a basis for $V, (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ a basis for W and $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s)$ a basis for X. Let $f_{klt} : V \times W \to X$ be defined by

$$f_{klt}(\sum_{i=1}^{m} a_i \boldsymbol{v}_i, \sum_{j=1}^{n} b_j \boldsymbol{w}_j) = a_k b_l \boldsymbol{x}_t.$$

Then $\{f_{klt} | 1 \leq k \leq m, 1 \leq l \leq n, 1 \leq t \leq s\}$ is a basis for $B(V, W; X)$.

5. For $\boldsymbol{w} \in W$ denote by F the map from V to $\mathbb{F}$ given by $F(\boldsymbol{w})(\boldsymbol{v}) = f(\boldsymbol{v}, \boldsymbol{w})$. This is a linear transformation from W to $V' = \mathcal{L}(V, \mathbb{F})$. Since $dim(V') = dim(V) = m$ by the rank-nullity theorem $dim(Ker(F)) \geq n - m$. The result follows since $Rad_R(f) = Ker(F)$.

6. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Set $V = \mathbb{F}^2$ and define $f : V \times V \to \mathbb{F}$ by $f(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}^{tr} A \boldsymbol{w}$. Then $Rad_L(f) = \{\begin{pmatrix} 0 \\ a \end{pmatrix} | a \in \mathbb{F}\}$ and $Rad_R(f) = \{\begin{pmatrix} b \\ 0 \end{pmatrix} | b \in \mathbb{F}\}$.

7. Let $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Set $V = \mathbb{F}^3$ and define $f : V \times V \to \mathbb{F}$ by $f(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}^{tr} A \boldsymbol{w}$. Then $Rad_R(f) = Rad_L(f) = Span(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix})$. However,

$$f(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}) = 0$$

$$f(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = 1.$$

8. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Set $V = \mathbb{F}^2$ and define $f : V \times V \to \mathbb{F}$ by $f(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}^{tr} A \boldsymbol{w}$. Then

$$f(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}) = 0$$

$$f(\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}) = 1$$

9. Define $f^+ : V \times V \to \mathbb{F}$ by $f^+(\boldsymbol{v}, \boldsymbol{w}) = \frac{1}{2}[f(\boldsymbol{v}, \boldsymbol{w}) + f(\boldsymbol{w}, \boldsymbol{v})]$. Then f^+ is symmetric. Similarly, define $f^- : V \times V \to \mathbb{F}$ by $f^-(\boldsymbol{v}, \boldsymbol{w}) = \frac{1}{2}[f(\boldsymbol{v}, \boldsymbol{w}) - f(\boldsymbol{w}, \boldsymbol{v})]$. Then f^- is alternating and $f^+ + f^- = f$.

10. Since $A = I_m A I_n$ every $m \times n$ matrix is equivalent to itself and the relation of equivalence is reflexive.

Suppose B is equivalent to A. Then there exists an invertible $m \times m$ matrix R and an invertible $n \times n$ matrix Q such that $B = RAQ$. Then $A = R^{-1}BQ^{-1}$ and so A is equivalent to B.

Finally, assume B is equivalent to A and C is equivalent to B. Then there are invertible $m \times m$ matrices R_1, R_2 and invertible $n \times n$ matrices Q_1, Q_2 such that

$$B = R_1 A Q_1, C = R_2 B Q_2.$$

Then $C = (R_2R_1)A(Q_1Q_2)$. Since the product of invertible two invertible matrices is invertible, R_2R_1 and Q_1Q_2 are invertible and therefore C is equivalent to A.

11. Assume the $m \times n$ matrix A has rank r. We prove that A is equivalent to the matrix $M^r_{m\times n} = \begin{pmatrix} I_r & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix}$. The result will follow from this.

Use Gaussian elimination to obtain the reduced echelon form of A. Since A has rank r there is an $r \times (n-r)$ matrix C such that the reduced echelon form of A is

$$B = \begin{pmatrix} I_r & C \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix}.$$

This implies that there are elementary matrices $E_1, \dots, E_s$ such that $B = E_s \dots E_1 A$. Set $R = E_s \dots E_1$. Then R is an invertible matrix. Next consider $B^{tr} =$

$$\begin{pmatrix} I_r & 0_{r\times(m-r)} \\ C^{tr} & 0_{(n-r)\times(m-r)} \end{pmatrix}$$

The reduced echelon form of this matrix is $\begin{pmatrix} I_r & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix}$. Thus, as above, there is an invertible $n \times n$ matrix P such that $PB^{tr} = M^r_{n\times m}$. Then $RAP^{tr} = BP^{tr} = (PB^{tr})^{tr} = M^r_{n\times m} = M^r_{m\times n}$ as claimed.

12. Since $I_n^{tr} A I_n = I_n A I_n = A$ every $n \times n$ matrix is congruent to itself and the relation of congruence is reflexive.

Assume B is congruent to A so that there is an invertible $n \times n$ matrix P such that $B = P^{tr}AP$. Set $Q = P^{-1}$. Then $Q^{tr} = (P^{-1})^{tr} = (P^{tr})^{-1}$. Then $A = (P^{tr})^{-1}BP^{-1} = Q^{tr}BQ$ and so A is congruent to B. This implies the relation is symmetric.

Finally, assume that B is congruent to A and C is congruent to B. Then there are invertible matrices P, Q such that $B = P^{tr}AP, C = Q^{tr}BQ$. Substituting in the latter expression we get

$$C = Q^{tr}(P^{tr}AP)Q = [Q^{tr}P^{tr}]A[PQ] = [PQ]^{tr}A[PQ].$$

Since P, Q are invertible $n \times n$ matrices, the product PQ is invertible. Thus, C is congruent to A and the relation is transitive.

13. Let $\mathcal{B}_V = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$ be a basis for V and $\mathcal{B}_W = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ be a basis for W. Let $a_{ij} = f(\boldsymbol{v}_i, \boldsymbol{w}_j)$ and let A be the $m \times n$ matrix with entries a_{ij}. Set $r = rank(A)$.

Let $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$. Then $\boldsymbol{v} \in Rad_L(f)$ if and only if $[\boldsymbol{v}]_{\mathcal{B}_V}$ is the null space of A^{tr} and $\boldsymbol{w} \in Rad_R(f)$ if and only if $[\boldsymbol{w}]_{\mathcal{B}_W}$ is in the null space of A. Since the $rank(A^{tr}) = rank(A)$ we have $dim(Rad_L(f)) = nullity(A^{tr}) = m - r$ and $dim(Rad_R(f)) = nullity(A) = n - r$. Then $dim\ (V/Rad_L(f)) = m-(m-r) = r = n-(n-r) = dim\ (W/Rad_R(f))$.

14. By Exercise 13, $dim(V) - dim(Rad_L(f)) = dim(W) - dim(Rad_R(f))$. Since $dim(Rad_L(f)) = dim(Rad_R(f)) = 0$ it follows that $dim(V) = dim(W)$.

15. Since f is alternating, $f(\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}) = f(\boldsymbol{u}, \boldsymbol{u}) = f(\boldsymbol{v}, \boldsymbol{v}) = 0$. However, by bilinearity $0 = f(\boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}) = f(\boldsymbol{u}, \boldsymbol{u}) + f(\boldsymbol{u}, \boldsymbol{v}) + f(\boldsymbol{v}, \boldsymbol{u}) + f(\boldsymbol{v}, \boldsymbol{v}) = f(\boldsymbol{u}, \boldsymbol{v}) + f(\boldsymbol{v}, \boldsymbol{u})$ from which the result follows.

16. Note that since f is non-degenerate, $dim(V) = dim(W)$. Let $F : W \to V'$ be the map given by $F(\boldsymbol{w})(\boldsymbol{v}) = f(\boldsymbol{v}, \boldsymbol{w})$. Then F is a linear transformation. Since f is non-degenerate, $Ker(F) = Rad_R(f) = \{\boldsymbol{0}_W\}$ so F is injective. Since $dim(V') = dim(V) = dim(W)$ it is then the case that F is an isomorphism. Let $g_i : V \to \mathbb{F}$ be the linear form given by $g_i(\sum_{j=1}^n a_j \boldsymbol{v}_j) = a_i$ and let $\boldsymbol{w}_i \in W$ such that $F(\boldsymbol{w}_i) = g_i$. Then $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ is the required basis.

8.2. Symplectic Space

1i) S^{-1} is an isomorphism of W onto V. We need to show if $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$ then $\langle S^{-1}(\boldsymbol{w}_1), S^{-1}(\boldsymbol{w}_2)\rangle_V = \langle \boldsymbol{w}_1, \boldsymbol{w}_2\rangle_W$. Set $\boldsymbol{v}_i = S^{-1}(\boldsymbol{w}_i)$. Then $S(\boldsymbol{v}_i) = \boldsymbol{w}_i$. Since

S is an isometry we have

$$\langle S^{-1}(\boldsymbol{w}_1), S^{-1}(\boldsymbol{w}_2)\rangle_V = \langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle_V$$

$$= \langle S(\boldsymbol{v}_1), S(\boldsymbol{v}_2)\rangle_W = \langle \boldsymbol{w}_1, \boldsymbol{w}_2\rangle.$$

ii) Let $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$. Then

$$\langle \boldsymbol{w}_1, \boldsymbol{w}_2\rangle_W = \langle (S(\boldsymbol{w}_1), S(\boldsymbol{w}_2)\rangle_V =$$

$$\langle T(S(\boldsymbol{w}_1)), T(S(\boldsymbol{w}_2))\rangle_X = \langle (TS)(\boldsymbol{w}_1), (TS)(\boldsymbol{w}_2)\rangle_X$$

2. Let $\boldsymbol{u} \in U$ be arbitrary and $\boldsymbol{w}_1, \boldsymbol{w}_2 \in U^\perp$. Then $\langle \boldsymbol{w}_1, \boldsymbol{u}\rangle = \langle \boldsymbol{w}_2, \boldsymbol{u}\rangle = 0$. Then

$$\langle \boldsymbol{w}_1 + \boldsymbol{w}_2, \boldsymbol{u}\rangle = \langle \boldsymbol{w}_1, \boldsymbol{u}\rangle =$$

$$\langle \boldsymbol{w}_1, \boldsymbol{u}\rangle + \langle \boldsymbol{w}_2, \boldsymbol{u}\rangle = 0 + 0 = 0.$$

Since $\boldsymbol{u}$ is arbitrary, $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in U^\perp$.

Next assume $\boldsymbol{u} \in U, \boldsymbol{v} \in V^\perp$ and $c \in \mathbb{F}$. Then $f(c\boldsymbol{v}, \boldsymbol{u}) = c(\boldsymbol{u}, \boldsymbol{v}) = c \times 0 = 0$ and $c\boldsymbol{v} \in U^\perp$.

3. Clearly $U \subset (U^\perp)^\perp$. By part i) of Lemma (8.12) we have $dim(V) = dim(U) + dim(U^\perp)$. Applying this to $U^\perp$ we also get $dim(V) = dim(U^\perp) + dim((U^\perp)^\perp)$. Consequently, $dim((U^\perp)^\perp) = dim(V) - dim(U^\perp) = dim(U)$. Since $U \subset (U^\perp)^\perp$ it follows that $U = (U^\perp)^\perp$.

4) Suppose $\boldsymbol{w} \in Rad(U^\perp)$. Then $\boldsymbol{w} \in U^\perp$ and $\boldsymbol{w} \in (U^\perp)^\perp = U$. However, since U is non-degenerate, $U \cap U^\perp = \{\boldsymbol{0}\}$ and so $\boldsymbol{w} = \boldsymbol{0}$.

5. Set $k = dim(U)$. By part i) of Lemma (8.12) we have $dim(V) = dim(U) + dim(U^\perp)$. However, since U is totally isotropic, $U \subset U^\perp$ so that $k = dim(U) \leq dim(U^\perp)$. Then $2n = dim(V) = dim(U) + dim(U^\perp) \geq 2dim(U) = 2k$. Therefore, $k \leq n$.

6. Since U is totally isotropic, $U \subset U^\perp$. Since $\langle \boldsymbol{u}, \boldsymbol{w}\rangle = 0$ for all $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in U^\perp$, in fact, $U \subset Rad(U^\perp)$. On the other hand, $Rad(U^\perp) = U^\perp \cap (U^\perp)^\perp = U^\perp \cap U \subset U$.

7. Set $T = T_{(\boldsymbol{v},c)}$. We compute:

$$\langle T(\boldsymbol{u}), T(\boldsymbol{w})\rangle = \langle \boldsymbol{u} + c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\boldsymbol{v}, \boldsymbol{w} + c\langle \boldsymbol{w}, \boldsymbol{v}\rangle\boldsymbol{v}\rangle =$$

$$\langle \boldsymbol{u}, \boldsymbol{w}\rangle + \langle \boldsymbol{u}, c\langle \boldsymbol{w}, \boldsymbol{v}\rangle\boldsymbol{v}\rangle +$$

$$\langle c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\boldsymbol{v}, \boldsymbol{w}\rangle + \langle c\langle \boldsymbol{u}, \boldsymbol{w}\rangle\boldsymbol{v}, c\langle \boldsymbol{w}, \boldsymbol{v}\rangle\boldsymbol{v}\rangle =$$

$$\langle \boldsymbol{u}, \boldsymbol{w}\rangle = c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{w}, \boldsymbol{v}\rangle +$$

$$c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{v}, \boldsymbol{w}\rangle + c^2\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{w}, \rangle\langle \boldsymbol{v}, \boldsymbol{v}\rangle.$$

Since $\langle\ ,\ \rangle$ is alternating $\langle \boldsymbol{v}, \boldsymbol{v}\rangle = 0$ and $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = -\langle \boldsymbol{w}, \boldsymbol{v}\rangle$. Thus,

$$c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{w}, \boldsymbol{v}\rangle + c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{v}, \boldsymbol{w}\rangle + c^2\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{w}, \rangle\langle \boldsymbol{v}, \boldsymbol{v}\rangle =$$

$$\langle \boldsymbol{u}, \boldsymbol{w}\rangle + c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{w}, \boldsymbol{v}\rangle - c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{w}, \boldsymbol{v}\rangle =$$

$$\langle \boldsymbol{u}, \boldsymbol{w}\rangle$$

8. Set $T = T_{(\boldsymbol{v},c)}$ and $S = T_{(\boldsymbol{w},d)}$. For $\boldsymbol{u} \in V$ we have $(TS)(\boldsymbol{u}) =$

$$\boldsymbol{u} + d\langle \boldsymbol{u}, \boldsymbol{w}\rangle\boldsymbol{w} + c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\boldsymbol{v} + (cd)\langle \boldsymbol{u}, \boldsymbol{w}\rangle\langle \boldsymbol{w}, \boldsymbol{v}\rangle\boldsymbol{v}.$$

On the other hand, $(ST)(\boldsymbol{u}) =$

$$\boldsymbol{u} + c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\boldsymbol{v} + d\langle \boldsymbol{u}, \boldsymbol{w}\rangle\boldsymbol{w} + (dc)\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{v}, \boldsymbol{w}\rangle\boldsymbol{w}$$

If $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0$ then $(ST)(\boldsymbol{u}) = (TS)(\boldsymbol{u}) = \boldsymbol{u} + c\langle \boldsymbol{u}, \boldsymbol{v}\rangle\boldsymbol{v} + d\langle \boldsymbol{u}, \boldsymbol{w}\rangle\boldsymbol{w}$ and S and T commute.

Conversely, assume $ST = TS$. Then $\langle \boldsymbol{u}, \boldsymbol{w}\rangle\langle \boldsymbol{w}, \boldsymbol{v}\rangle\boldsymbol{v} = \langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{v}, \boldsymbol{w}\rangle\boldsymbol{w}$ for every vector $\boldsymbol{u}$. If $(\boldsymbol{v}, \boldsymbol{w})$ is linearly dependent then clearly $\boldsymbol{v} \perp \boldsymbol{w}$. Assume $(\boldsymbol{v}, \boldsymbol{w})$ is linearly independent. Then we must have $\langle \boldsymbol{u}, \boldsymbol{w}\rangle\langle \boldsymbol{w}, \boldsymbol{v}\rangle = 0 = \langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{v}, \boldsymbol{w}\rangle$ for every vector $\boldsymbol{u}$. Since a vector space of dimension n cannot be the union of two subspaces of dimension $n - 1$, particular, $V \neq \boldsymbol{v}^\perp \cup \boldsymbol{w}^\perp$. Choose $\boldsymbol{u} \in V$ such that $\boldsymbol{v} \not\perp \boldsymbol{v} \not\perp \boldsymbol{w}$. It then must be the case that $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0$.

9. The number of non-zero vectors is $q^{2n}-1$. There are this many choices for $\boldsymbol{u}$. There are q^{2n-1} vectors in $\boldsymbol{u}^{\perp}$ and hence $q^{2n}-q^{2n-1}$ vectors $\boldsymbol{x}$ such that $\langle \boldsymbol{u},\boldsymbol{x}\rangle \neq 0$. For any such vector $\boldsymbol{x}$ there is a unique vector $\boldsymbol{v}$ in $Span(\boldsymbol{x})$ such that $\langle \boldsymbol{u},\boldsymbol{v}\rangle = 1$. Since there are $q-1$ non-zero vectors in $Span(\boldsymbol{x})$ there are $\frac{q^{2n}-q^{2n-1}}{q-1} = q^{2n-1}$ vectors $\boldsymbol{v}$ such that $\langle \boldsymbol{u},\boldsymbol{v}\rangle = 1$. Thus, the number of such pairs is $q^{2n-1}(q^{2n}-1)$.

10. We proceed by induction to prove that the number of hyperbolic bases in a non-degenerate $2n-$dimensional symplectic space over $\mathbb{F}_q$ is $q^{n^2}\prod_{i=1}^n (q^{2i}-1)$.

The base case, $n=1$ follows from Exercise 9. Suppose there are $q^{n^2}\prod_{i=1}^n (q^{2i}-1)$ hyperbolic bases in a non-degenerate $2n-$dimensional symplectic space over $\mathbb{F}_q$. and assume that $(V,\langle\ ,\ \rangle)$ is a non-degenerate symplectic space of dimension $2n+2$. By Exercise 9 there are $q^{2n+1}(q^{2n+2}-1)$ pairs $(\boldsymbol{u},\boldsymbol{v})$ with $\langle \boldsymbol{u},\boldsymbol{v}\rangle = 1$. Set $W = Span(\boldsymbol{u},\boldsymbol{v})^{\perp}$. Then W is a non-degenerate space of dimension $2n$. By the inductive hypothesis there are $q^{n^2}\prod_{i=1}^n (q^{2i}-1)$ hyperbolic bases $\mathcal{B}$ in W. We obtain a hyperbolic basis for V by adjoining $(\boldsymbol{u},\boldsymbol{v})$ to $\mathcal{B}$. Since there are $q^{2n+1}(q^{2n+2}-1)$ choices of $(\boldsymbol{u},\boldsymbol{v})$ and $q^{n^2}\prod_{i=1}^n (q^{2i}-1)$ choices of $\mathcal{B}$ there are

$$q^{2n+1}(q^{2n+2}-1)q^{n^2}\prod_{i=1}^{n}(q^{2i}-1) =$$

$$q^{n^2+2n+1}\prod_{i=1}^{n+1}(q^{2i}-1) =$$

$$q^{(n+1)^2}\prod_{i=1}^{n+1}(q^{2i}-1)$$

Now suppose $dim(V) = 2n$. If we fix a hyperbolic basis $\mathcal{B}$ then any isometry T takes $\mathcal{B}$ to a hyperbolic basis $T(\mathcal{B}')$. On the other hand, T is uniquely determined by its image on a basis. Therefore there is a one-to-one correspondence between isometries and hyperbolic bases and this implies that $|Sp(V)| = q^{n^2}\prod_{i=1}^n (q^{2i}-1)$.

11. Clearly, $U \subset (U^{\perp})^{\perp}$. By Lemma (8.12) i) $dim(V) = dim(U) + dim(U^{\perp}) = dim(U^{\perp}) + dim((U^{\perp})^{\perp})$.Therefore $dim((U^{\perp})^{\perp}) = dim(U)$ so we have equality.

8.3. Quadratic Forms and Orthogonal Space

1i) Let $\boldsymbol{v} \in V_2$ and set $\boldsymbol{u} = S^{-1}(\boldsymbol{v})$. Then $S(\boldsymbol{u}) = \boldsymbol{v}$. Thus,

$$\phi_1(S^{-1}(\boldsymbol{v})) = \phi_1(\boldsymbol{u}) = \phi_2(S(\boldsymbol{u})) = \phi_2(\boldsymbol{v}).$$

ii) Let $\boldsymbol{u} \in V_1$. Then $\phi_1(\boldsymbol{u}) = \phi_2(S(\boldsymbol{u})) = \phi_3(T(S(\boldsymbol{u}))) = \phi_3((TS)(\boldsymbol{u}))$.

2. Set $n = dim(V)$. A vector $\boldsymbol{v} \in Rad(V)$ if and only if $[\boldsymbol{v}]_{\mathcal{B}}$ is in the null space of the matrix A. Therefore $dim(Rad(V)) = nullity(A)$. Then the rank of (V,ϕ) is equal to $dim(V) - dim(Rad(V)) = n - dim(Rad(V)) = n - nullity(A) = rank(A)$.

3. $\rho(\boldsymbol{y}) = \boldsymbol{y} - 2\frac{\langle \boldsymbol{y},\boldsymbol{x}\rangle}{\langle \boldsymbol{x},\boldsymbol{x}\rangle}\boldsymbol{x} = \boldsymbol{y}$ since $\langle \boldsymbol{y},\boldsymbol{x}\rangle = 0$. On the other hand,

$$\rho(\boldsymbol{x}) = \boldsymbol{x} - 2\frac{\langle \boldsymbol{x},\boldsymbol{x}\rangle}{\langle \boldsymbol{x},\boldsymbol{x}\rangle}\boldsymbol{x} = \boldsymbol{x} - 2\boldsymbol{x} = -\boldsymbol{x}.$$

4. Choose a singular vector $\boldsymbol{u}$ and let $\boldsymbol{u}_{\infty} = \boldsymbol{u}$. Let $\boldsymbol{v}$ be a vector in $\boldsymbol{u}^{\perp}$ but not a multiple of $\boldsymbol{u}$ so that $\boldsymbol{u}^{\perp} = Span(\boldsymbol{u},\boldsymbol{v})$. The vector $\boldsymbol{v}$ is non-singular and, in fact, for every $c \in \mathbb{F}, \boldsymbol{v}_c = c\boldsymbol{u} + \boldsymbol{v} \in \boldsymbol{u}^{\perp}$ and is non-singular. Then $\boldsymbol{v}_c^{\perp}$ is a non-singular two dimensional subspace and contains the singular vector $\boldsymbol{u}$. There is then a second singular vector which we denote by $\boldsymbol{u}_c$.

On the other hand, suppose $\boldsymbol{w}$ is a singular vector, $Span(\boldsymbol{w}) \neq Span(\boldsymbol{u})$. Then $\boldsymbol{w}^{\perp}$ has dimension two and $\boldsymbol{w}^{\perp} \neq \boldsymbol{u}^{\perp} = Span(\boldsymbol{u},\boldsymbol{v})$. It follows that $\boldsymbol{w}^{\perp}$ intersects $Span(\boldsymbol{u},\boldsymbol{v})$ in a one-dimensional subspace different from $\boldsymbol{u}$ and therefore contains a vector $a\boldsymbol{u} + b\boldsymbol{v}$ with $b \neq 0$. Then $\boldsymbol{w}$ is orthogonal to $\frac{a}{b}\boldsymbol{u} + \boldsymbol{v}$ and hence $\boldsymbol{w} = \boldsymbol{u}_c$

where $c = \frac{a}{b}$. Thus, we have a one-to-one correspondence between $\mathbb{P}(V)$ and $\mathbb{F} \cup \{\infty\}$.

5. Replacing $\boldsymbol{u}_2$ and $\boldsymbol{v}_2$ by scalar multiples, if necessary, we can assume that $\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = 1$. By Lemma (8.25) there exists an isometry T such that $T(\boldsymbol{u}_1) = \boldsymbol{v}_1$. Set $\boldsymbol{v}_2' = T(\boldsymbol{u}_2)$. Then since T is an isometry $\langle \boldsymbol{v}_1, \boldsymbol{v}_2' \rangle = \langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle = \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = 1$. Now by Lemma (8.26) there exists an isometry S such that $S(\boldsymbol{v}_1) = \boldsymbol{v}_1, S(\boldsymbol{v}_2') = \boldsymbol{v}_2$. Then (ST) is an isometry and $(ST)(\boldsymbol{u}_1) = T(\boldsymbol{v}_1) = \boldsymbol{v}_1, (ST)(\boldsymbol{u}_2) = S(\boldsymbol{v}_2') = \boldsymbol{v}_2$.

6. Let U, W be totally singular subspaces which are not properly contained in a totally isotropic subspace. Assume $dim(U) = k, dim(W) = l$ with $k \leq l$. Choose any subspace X of W of dimension k. Since $dim(U) = dim(X) = k$ U and X are isomorphic. Since U and X are totally singular any isomorphism $S : X \to U$ is an isometry. By Witt's theorem there is an isometry of $\widehat{S}$ such that $\widehat{S}$ restricted to X is S. Now set $W' = \widehat{S}(W)$. Then W' is a totally singular subspace which contains U. By our assumption that U is a maximal totally singular subspace it follows that $W' = U$ and so $dim(U) = dim(W') = dim(W)$.

7. The proof is by induction on $n = dim(V)$. Suppose $n = 1$. Let $V = Span(\boldsymbol{v})$ and T an isometry, $T \neq I_V$. Then $T(\boldsymbol{v}) = -\boldsymbol{v}$ and $T = \rho_{\boldsymbol{v}}$. Assume now that $dim(V) = n + 1$ and the result has been proved for spaces of dimension n. Let $\boldsymbol{v} \in V$ be a non-singular vector and T an isometry. Set $\boldsymbol{w} = T(\boldsymbol{v})$. Then $\phi(\boldsymbol{w}) = \phi(T(\boldsymbol{v})) = \phi(\boldsymbol{v}) \neq 0$. By the proof of Theorem (8.11) there is an isometry S which is a product of at most two reflections such that $S(\boldsymbol{v}) = \boldsymbol{w}$. Set $T' = S^{-1}T$. Then $T'(\boldsymbol{v}) = \boldsymbol{v}$ and T' leaves $\boldsymbol{w}^\perp$ invariant. The subspace $\boldsymbol{v}^\perp$ is non-degenerate. By the inductive hypothesis there are non-singular vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_t$ from $\boldsymbol{v}^\perp$ such that T' restricted to $\boldsymbol{v}^\perp$ is the product of the reflections $\rho_i = \rho_{\boldsymbol{u}_i}$ restricted to $\boldsymbol{v}^\perp$. It then follows that $T = S\rho_1 \ldots \rho_t$, a product of reflections.

8. By Exercise 7 an isometry is a product of reflections. Since $det(\rho_{\boldsymbol{x}}) = -1$ for $\boldsymbol{x}$ non-singular it follows that the determinant of an isometry is ± 1.

9a) $\phi(T_{(\boldsymbol{u},\boldsymbol{v})}(\boldsymbol{z})) = \phi(\boldsymbol{z} + \langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - \langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v}) =$

$$\phi(\boldsymbol{z}) + \phi(\langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - \langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v}) + \langle \boldsymbol{z}, \langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - \langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v} \rangle$$

Since $\langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - \langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v}$ is a singular vector we get

$$= \phi(\boldsymbol{z}) + \langle \boldsymbol{z}, \langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - \langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v} \rangle =$$

$$\phi(\boldsymbol{z}) = \langle \boldsymbol{z}, \boldsymbol{v} \rangle \langle \boldsymbol{z}, \boldsymbol{u} \rangle - \langle \boldsymbol{z}, \boldsymbol{u} \rangle \langle \boldsymbol{z}, \boldsymbol{v} \rangle =$$

$$\phi(\boldsymbol{z}).$$

b) If $\boldsymbol{z} \in Span(\boldsymbol{u}, \boldsymbol{v})^\perp$ then $\langle \boldsymbol{z}, \boldsymbol{u} \rangle = \langle \boldsymbol{z}, \boldsymbol{v} \rangle = 0$ and $T_{(\boldsymbol{u},\boldsymbol{v})}(\boldsymbol{z}) = \boldsymbol{z}$.

c) Since $(T_{(\boldsymbol{u},\boldsymbol{v})} - I_V)(\boldsymbol{z}) = \langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - \langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v} \in Span(\boldsymbol{u}, \boldsymbol{v})$ we clearly have $Range(T_{(\boldsymbol{u},\boldsymbol{v})} - I_V) \subset Span(\boldsymbol{u}, \boldsymbol{v})$. On the other hand, there exist vectors $\boldsymbol{z}_1, \boldsymbol{z}_2$ such that $\langle \boldsymbol{z}_1, \boldsymbol{u} \rangle = 0, \langle \boldsymbol{z}_1, \boldsymbol{v} \rangle = 1$ and $\langle \boldsymbol{z}_2, \boldsymbol{u} \rangle = 1, \langle \boldsymbol{z}_2, \boldsymbol{v} \rangle = 0$.

We then have $(T_{(\boldsymbol{u},\boldsymbol{v})} - I_V)(\boldsymbol{z}_1) = \langle \boldsymbol{z}_1, \boldsymbol{v} \rangle \boldsymbol{u} = \boldsymbol{u}$. Similarly, $(T_{(\boldsymbol{u},\boldsymbol{v})} - I_V)(\boldsymbol{z}_2) = -\boldsymbol{v}$.

10a) Set $T = T_{(\boldsymbol{u},-c\boldsymbol{v})}$ and $S = T_{(\boldsymbol{u},c\boldsymbol{v})}$. Then

$$S(\boldsymbol{z}) = \boldsymbol{z} + c\langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - c\langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v}.$$

Set $\boldsymbol{w} = c\langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - c\langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v}$. Then $\boldsymbol{w}$ is orthogonal to $\boldsymbol{u}, \boldsymbol{v}$ and consequently, $T(\boldsymbol{w}) = \boldsymbol{w}$. Then

$$(TS)(\boldsymbol{z}) = T(\boldsymbol{z} + \boldsymbol{w}) = T(\boldsymbol{z}) + \boldsymbol{w} =$$

$$\boldsymbol{z} - c\langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} + c\langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v} + \boldsymbol{w} = \boldsymbol{z} - \boldsymbol{w} + \boldsymbol{w} = \boldsymbol{z}.$$

b) Set $S = T_{(\boldsymbol{u},c\boldsymbol{v})}, T = T_{(\boldsymbol{u},d\boldsymbol{v})}$. Then

$$T(\boldsymbol{z}) = \boldsymbol{z} + d\langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - d\langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v}.$$

Set $\boldsymbol{w} = d\langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - d\langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v}$. Then $S(\boldsymbol{w}) = \boldsymbol{w}$. Therefore

$$(ST)(\boldsymbol{z}) = S(\boldsymbol{z} + \boldsymbol{w}) = \boldsymbol{z} + c\langle \boldsymbol{z}, \boldsymbol{v} \rangle \boldsymbol{u} - c\langle \boldsymbol{z}, \boldsymbol{u} \rangle \boldsymbol{v} + \boldsymbol{w} =$$

$$\boldsymbol{z} + c\langle \boldsymbol{z}, \boldsymbol{v}\rangle \boldsymbol{u} - c\langle \boldsymbol{z}, \boldsymbol{u}\rangle \boldsymbol{v} + d\langle \boldsymbol{z}, \boldsymbol{v}\rangle \boldsymbol{u} - d\langle \boldsymbol{z}, \boldsymbol{u}\rangle \boldsymbol{v} =$$
$$\boldsymbol{z} + (c+d)\langle \boldsymbol{z}, \boldsymbol{v}\rangle \boldsymbol{u} - (c+d)\langle \boldsymbol{z}, \boldsymbol{u}\rangle \boldsymbol{v} =$$
$$T_{(\boldsymbol{u},(c+d)\boldsymbol{v})}(\boldsymbol{z}).$$

11. Set $T = T_{(\boldsymbol{x},\boldsymbol{y})}$ then

$$T(\boldsymbol{z}) = \boldsymbol{z} + \langle \boldsymbol{z}, \boldsymbol{y}\rangle \boldsymbol{x} - \langle \boldsymbol{z}, \boldsymbol{x}\rangle \boldsymbol{y} =$$
$$\boldsymbol{z} + \langle \boldsymbol{z}, c\boldsymbol{u} + d\boldsymbol{v}\rangle (a\boldsymbol{u} + b\boldsymbol{v}) - \langle \boldsymbol{z}, a\boldsymbol{u} + b\boldsymbol{v}\rangle (c\boldsymbol{u} + d\boldsymbol{v}) =$$
$$\boldsymbol{z} + (ac)\langle \boldsymbol{z}, \boldsymbol{u}\rangle \boldsymbol{u} + (bd)\langle \boldsymbol{z}, \boldsymbol{v}\rangle \boldsymbol{v} + (ad)\langle \boldsymbol{z}, \boldsymbol{v}\rangle \boldsymbol{u} +$$
$$(bc)\langle \boldsymbol{z}, \boldsymbol{u}\rangle \boldsymbol{v} - (ac)\langle \boldsymbol{z}, \boldsymbol{u}\rangle \boldsymbol{u} +$$
$$(bd)\langle \boldsymbol{z}, \boldsymbol{v}\rangle \boldsymbol{v} - (ad)\langle \boldsymbol{z}, \boldsymbol{u}\rangle \boldsymbol{v} - (bd)\langle \boldsymbol{z}, \boldsymbol{v}\rangle \boldsymbol{u} =$$
$$\boldsymbol{z} + (ad - bc)\langle \boldsymbol{z}, \boldsymbol{v}\rangle \boldsymbol{u} - (ad - bc)\langle \boldsymbol{z}, \boldsymbol{u}\rangle \boldsymbol{u} =$$
$$T_{(\boldsymbol{u},(ad-bc)\boldsymbol{v})}(\boldsymbol{z}).$$

12. We compute $T_{(\boldsymbol{u},d\boldsymbol{v})}(\boldsymbol{w})$:

$$T_{(\boldsymbol{u},d\boldsymbol{v})}(\boldsymbol{w}) = \boldsymbol{w} + \langle \boldsymbol{w}, d\boldsymbol{v}\rangle \boldsymbol{u} - \langle \boldsymbol{w}, \boldsymbol{u}\rangle (d\boldsymbol{v}) =$$

$$\boldsymbol{w} + d\langle \boldsymbol{w}, \boldsymbol{v}\rangle \boldsymbol{u} = \boldsymbol{w} + d\boldsymbol{u}.$$

Thus, we must take $d = c$.

13. Assume $\boldsymbol{x} \in \boldsymbol{u}^\perp$. Then $\phi(\delta_{\boldsymbol{u},\boldsymbol{v}}(\boldsymbol{x})) = \phi(\boldsymbol{x} + \langle \boldsymbol{x}, \boldsymbol{v}\rangle_\phi \boldsymbol{u}) = \phi(\boldsymbol{x}) + \langle \boldsymbol{x}, \boldsymbol{v}\rangle_\phi^2 \phi(\boldsymbol{u}) + \langle \boldsymbol{x}, \langle \boldsymbol{x}, \boldsymbol{v}\rangle_\phi \boldsymbol{u}\rangle_\phi = \phi(\boldsymbol{x})$ since $\boldsymbol{u}$ is singular and $\boldsymbol{x} \perp \boldsymbol{u}$.

14. Since $V = Span(\boldsymbol{u}, \boldsymbol{w}) \oplus Span(\boldsymbol{u}, \boldsymbol{w})^\perp$ there is a vector $\boldsymbol{z} \in Span(\boldsymbol{u}, \boldsymbol{w})^\perp$ and scalars a and b such that $D(\boldsymbol{w}) = a\boldsymbol{w} + b\boldsymbol{u} + \boldsymbol{z}$. Note that $\phi(a\boldsymbol{w} + b\boldsymbol{u} + \boldsymbol{z}) = ab + \phi(\boldsymbol{z})$. Since $0 = \phi(\boldsymbol{w}) = \phi(D(\boldsymbol{w}))$ we have $ab + \phi(\boldsymbol{z}) = 0$.

Next note that $a = \langle a\boldsymbol{w} + b\boldsymbol{u} + \boldsymbol{z}, \boldsymbol{u}\rangle_\phi = \langle D(\boldsymbol{w}), \boldsymbol{u}\rangle_\phi = \langle \boldsymbol{w}, \boldsymbol{u}\rangle_\phi = 1$.

Now suppose $\boldsymbol{x} \in Span(\boldsymbol{u}, \boldsymbol{v})^\perp$. Then $D(\boldsymbol{x}) = \delta_{\boldsymbol{u},\boldsymbol{v}}(\boldsymbol{x}) = \boldsymbol{x}$. Consequently,

$$\langle \boldsymbol{w}, \boldsymbol{x}\rangle_\phi = \langle D(\boldsymbol{w}), D(\boldsymbol{x})\rangle_\phi = \langle \boldsymbol{w} + b\boldsymbol{u} + \boldsymbol{z}, \boldsymbol{x}\rangle_\phi =$$

$$\langle \boldsymbol{w}, \boldsymbol{x}\rangle_\phi + \langle \boldsymbol{z}, \boldsymbol{x}\rangle_\phi.$$

Therefore $\langle \boldsymbol{z}, \boldsymbol{x}\rangle_\phi = 0$ for every $\boldsymbol{x} \in Span(\boldsymbol{u}, \boldsymbol{v})^\perp$ so that $\boldsymbol{z} \in Span(\boldsymbol{u}, \boldsymbol{v})$ in which case we can assume that $\boldsymbol{x}$ is a multiple of $\boldsymbol{v}$. Thus, there is a scalar c such that $D(\boldsymbol{w}) = \boldsymbol{w} + c\boldsymbol{v} - c^2\phi(\boldsymbol{v})\boldsymbol{u}$. Now let $\boldsymbol{x} \in \boldsymbol{u}^\perp$ be chosen such that $\langle \boldsymbol{v}, \boldsymbol{x}\rangle_\phi = 1$. Then $D(\boldsymbol{x}) = \delta_{\boldsymbol{u},\boldsymbol{v}}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{u}$. We then have

$$\langle \boldsymbol{w}, \boldsymbol{x}\rangle_\phi = \langle \boldsymbol{w} + c\boldsymbol{v} - c^2\phi(\boldsymbol{v})\boldsymbol{u}, \boldsymbol{x} + \boldsymbol{u}\rangle_\phi =$$
$$\langle \boldsymbol{w}, \boldsymbol{x}\rangle_\phi + c\langle \boldsymbol{v}, \boldsymbol{x}\rangle_\phi + \langle \boldsymbol{w}, \boldsymbol{u}\rangle_\phi =$$
$$\langle \boldsymbol{w}, \boldsymbol{x}\rangle_\phi + c + 1.$$

We can finally conclude that $c = -1$ and $D(\boldsymbol{w}) = \boldsymbol{w} - \boldsymbol{v} + \phi(\boldsymbol{v})\boldsymbol{u}$ as claimed.

15. By Exercise 14 it suffices to prove for $\boldsymbol{x} \in \boldsymbol{u}^\perp$ that $\delta_{\boldsymbol{u},\boldsymbol{v}}\delta_{\boldsymbol{u},\boldsymbol{w}}(\boldsymbol{x}) = \delta_{\boldsymbol{u},\boldsymbol{v}+\boldsymbol{w}}(\boldsymbol{x})$.

$$\delta_{\boldsymbol{u},\boldsymbol{v}}\delta_{\boldsymbol{u},\boldsymbol{w}}(\boldsymbol{x}) = \delta_{\boldsymbol{u},\boldsymbol{v}}(\boldsymbol{x} + \langle \boldsymbol{x}, \boldsymbol{w}\rangle_\phi \boldsymbol{u}) =$$
$$\delta_{\boldsymbol{u},\boldsymbol{v}}(\boldsymbol{x}) + \delta_{\boldsymbol{u},\boldsymbol{v}}(\langle \boldsymbol{x}, \boldsymbol{w}\rangle_\phi \boldsymbol{u}) =$$
$$\boldsymbol{x} + \langle \boldsymbol{x}, \boldsymbol{v}\rangle_\phi \boldsymbol{u} + \langle \boldsymbol{x}, \boldsymbol{w}\rangle_\phi \boldsymbol{u} =$$
$$\boldsymbol{x} + (\langle \boldsymbol{x}, \boldsymbol{v}\rangle_\phi + \langle \boldsymbol{x}, \boldsymbol{w}\rangle_\phi)\boldsymbol{u} =$$
$$\boldsymbol{x} + \langle \boldsymbol{x}, \boldsymbol{v} + \boldsymbol{w}\rangle_\phi \boldsymbol{u} = \delta_{\boldsymbol{u},\boldsymbol{v}+\boldsymbol{w}}(\boldsymbol{x}).$$

16. By Lemma (8.21), V has an orthogonal basis $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. Order the basis vectors such that $\phi(\boldsymbol{v}_1), \ldots, \phi(\boldsymbol{v}_r) \neq 0$ and $\phi(\boldsymbol{v}_k) = 0$ for $r < k \leq n$. Set $a_i = \phi(\boldsymbol{v}_i)$ for $1 \leq i \leq r$ and let b_i be chosen such that $b_i^2 = a_i$. Replacing $\boldsymbol{v}_i$, if necessary, with $\frac{1}{b_i}\boldsymbol{v}_i$ we can assume that $\phi(\boldsymbol{v}_i) = 1$ for $1 \leq i \leq r$ and $\phi(\boldsymbol{v}_i) = 0$ for $r < i \leq n$. We have shown that the matrix of a quadratic form over the field $\mathbb{F}$ is congruent to one and only one of the matrices $\begin{pmatrix} I_r & \mathbf{0}_{r\times n-r} \\ \mathbf{0}_{n-r\times r} & \mathbf{0}n - r \times n - r \end{pmatrix}$. Thus, two orthogonal spaces of dimension n over $\mathbb{F}$ are isometric if and only if they have the same rank.

17. Let M be a subspace such that $\phi(\boldsymbol{u}) > 0$ for all non-zero vectors $\boldsymbol{u} \in M$. Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ be an orthogonal basis of M which exists by Lemma (8.21). As in Exercise 16 we can assume that $\phi(\boldsymbol{v}_i) = 1$ for $1 \le i \le m$. Now suppose $M_1, M_2 \in \mathcal{P}$ are maximal with dimensions $m_i, i = 1, 2$. Let $(\boldsymbol{v}_{1i}, \ldots, \boldsymbol{v}_{m_i,i})$ be an orthonormal basis of M_i. We claim that $m_1 = m_2$. Suppose to the contrary that $m_1 \neq m_2$. We may assume and without loss of generality that $m_1 < m_2$. Let σ be the linear map from $U = Span(\boldsymbol{v}_{12}, \ldots, \boldsymbol{v}_{m_1,2})$ to M_1 such that $\sigma(\boldsymbol{v}_{j2}) = \boldsymbol{v}_{j1}$ for $1 \le j \le m_1$. Then σ is an isometry. By Witt's theorem there exists an isometry S of V such that S restricted to U is σ. Set $M_2' = S(M_2)$. Then $S(M_2) \in \mathcal{P}$ and $S(M_2)$ properly contains M_1, a contradiction. Thus, $m_1 = m_2$ and the map S is an isometry of V which takes M_2 to M_1.

18. By Lemma (8.21) there is an orthogonal basis $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ for V. First suppose that at least two of $(\boldsymbol{v}_1), \phi(\boldsymbol{v}_2), \phi(\boldsymbol{v}_3)$ are squares, say $\phi(\boldsymbol{v}_1) = a^2, \phi(\boldsymbol{v}_2) = b^2$. Then replacing $\boldsymbol{v}_1$ by $\frac{1}{a}\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ by $\frac{1}{b}\boldsymbol{v}_2$ we can assume that $\phi(\boldsymbol{v}_1) = \phi(\boldsymbol{v}_2) = 1$.

The following is well-known and proved in a first course in abstract algebra (it depends on the fact that the multiplicative group of the field $\mathbb{F}_q$ is cyclic): for any $c \in \mathbb{F}_q$ there are $d, e \in \mathbb{F}_q$ such that $c = d^2 + e^2$. In particular, this applies to $-\phi(\boldsymbol{v}_3)$. Now if $d^2 + e^2 = -\phi(\boldsymbol{v}_3)$ then $\phi(d\boldsymbol{v}_1 + e\boldsymbol{v}_2 + \boldsymbol{v}) = 0$.

On the other hand if $\phi(\boldsymbol{v}_1)$ and $\phi(\boldsymbol{v}_2)$ are non-squares in $\mathbb{F}_q$ then replace ϕ with $d\phi$ where d is a non-square. Then we can apply the above. Since the singular vectors of ϕ and $d\phi$ are the same we are done.

19. The proof is by induction on n. If $n \le 3$ then this follows from Exercise 18. So assume that $dim(V) = n > 3$ and the result is true for all spaces of dimension less than n. Since $dim(V) > 3$ by Exercise 18 there is a singular vector $\boldsymbol{v}$. By Lemma (8.24) there exists a singular vector $\boldsymbol{w}$ such that $\langle \boldsymbol{v}, \boldsymbol{w} \rangle_\phi = 1$. Set $U = Span(\boldsymbol{v}, \boldsymbol{w})^\perp$ which is non-degenerate of dimension $n-2$. Set $h = \lfloor \frac{n-3}{2} \rfloor$. By the inductive hypothesis there exists a totally singular subspace M of $U, dim(M) = h$. Then $M' = M \oplus Span(\boldsymbol{v})$ is a totally singular subspace of $V, dim(M') = \lfloor \frac{n-1}{2} \rfloor$.

20. When $m = 1$ there are two singular subspaces of dimension one, each with $q - 1$ non-zero vectors so there are altogether $2(q-1)$ singular vectors when $m = 1$ which is equal to $(q^m - 1)(q^{m-1} + 1)$. Assume we have shown that in a space of dimension $2m$ (maximal Witt index) there are $(q^m - 1)(q^{m-1} + 1)$ singular vectors and that $dim(V) = 2(m+1)$. Let $\boldsymbol{v}, \boldsymbol{w}$ be singular vectors with $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = 1$ and set $U = Span(\boldsymbol{v}, \boldsymbol{w})^\perp$. Let $\Delta(\boldsymbol{v})$ be the singular vectors $\boldsymbol{u}$ such that $\boldsymbol{u} \perp \boldsymbol{v}$ and $Span(\boldsymbol{u}) \neq Span(\boldsymbol{v})$ and $\Gamma(\boldsymbol{v})$ the singular vectors $\boldsymbol{z}$ such that $\boldsymbol{v} \not\perp \boldsymbol{z}$. Count $|\Gamma(\boldsymbol{v})|$ first. Note that $\boldsymbol{v}^\perp$ has dimension $2m - 1$ and therefore there are $q^{2m} - q^{2m-1}$ vectors in $V \setminus \boldsymbol{v}^\perp$. Let $\boldsymbol{x}$ be any vector such $\boldsymbol{x} \not\perp \boldsymbol{v}$. Then in $Span(\boldsymbol{v}, \boldsymbol{x})$ there are $q^2 - q$ which are non-orthogonal to $\boldsymbol{v}$. Consequently, the number of non-degenerate two dimensional subspace which contain $\boldsymbol{v}$ is $\frac{q^{2m} - q^{2m-1}}{q^2 - q}$. Any such two dimensional subspace contains $q - 1$ singular vectors which are not orthogonal to $\boldsymbol{v}$ and we can conclude that

$$|\Gamma(\boldsymbol{v})| = (q-1)\frac{q^{2m} - q^{2m-1}}{q^2 - q} = q^{2m-1} - q^{2m-2}.$$

Now suppose $\boldsymbol{u} \in \boldsymbol{v}^\perp, Span(\boldsymbol{u}) \neq Span(\boldsymbol{v})$. Then $Span(\boldsymbol{u}, \boldsymbol{v})$ is a totally singular two dimensional subspace and has $q^2 - q$ such vectors. Now $\boldsymbol{w}^\perp \cap Span(\boldsymbol{u}, \boldsymbol{v})$ is a one-dimensional singular subspace contained in U. By the inductive hypothesis there are $(q^m - 1)(q^{m-1} + 1)$ singular vectors in U and therefore $\frac{(q^m-1)(q^{m-1}+1)}{q-1}$ such subspaces. It now follows that

$$|\Delta(\boldsymbol{v})| = (q^2 - q)\frac{(q^m - 1)(q^{m-1} + 1)}{q - 1} =$$

$$q(q^m - 1)(q^{m-1} + 1)$$

Finally, there are $q - 1$ singular vectors in $Span(\boldsymbol{v})$. Thus, the number of singular vectors is

$$(q-1) + q(q^m - 1)(q^{m-1} + 1) + q^{2m-1} - q^{2m-2}$$

After simplifying we get $(q^{m+1}-1)(q^m+1)$.

21. This follows from the proof of Exercise 20.

22. It follows from Exercises 20 and 21 that the number of hyperbolic pairs is $(q^m-1)(q^{m-1}+1)q^{2m-2}$. The result follows from a straight forward induction on m.

23. Need to do induction on m. Assume $m=1$. Let $\boldsymbol{x}$ be an element of V such that $\phi(\boldsymbol{x})=1$ and let $\boldsymbol{y}\in \boldsymbol{x}^\perp$ and set $d=\phi(\boldsymbol{y})$. There are then $2(q+1)$ pairs $(\boldsymbol{u},\boldsymbol{v})$ such that $\phi(\boldsymbol{u})=1, \phi(vv)=d$. For any such pair the linear map $\tau(a\boldsymbol{x}+b\boldsymbol{y})=a\boldsymbol{u}+b\boldsymbol{v}$ is an isometry and every isometry arises this way. Thus, $|O(V,\phi)|=2(q+1)$. Now assume that $m>1$. We can show by induction that the number of singular vectors is $(q^m+1)(q^{m-1}-1)$ and the number of hyperbolic pairs is $q^{2m-2}(q^m+1)(q^{m-1}-1)$. Then we can show by induction that the number of sequences $(\boldsymbol{x}_1,\boldsymbol{y}_1,\dots,xx_{m-1},\boldsymbol{y}_{m-1})$ such that $\langle \boldsymbol{x}_i,\boldsymbol{y}\rangle_\phi=1$ for $1\le i\le m-1$ and $\boldsymbol{x}_i\perp \boldsymbol{x}_j, \boldsymbol{x}_i\perp \boldsymbol{y}_j, \boldsymbol{y}_i\perp \boldsymbol{y}_j$ for $i\neq j$ is $q^{2\binom{m}{2}}(q^m+1)(q-1)\Pi_{i=2}^{m-1}(q^{2i}-1)$. The order of $O(V,\phi)$ is obtained by multiplying this number by $2(q+1)$ to get $2q^{2\binom{m}{2}}(q^m+1)\Pi_{i=1}^{m-1}(q^{2i}-1)$.

8.4. Orthogonal Space, Characteristic Two

1. Let $\boldsymbol{v}$ be a singular vector in (V,ϕ). By Lemma (8.29) there is a singular vector $\boldsymbol{w}$ such that $\langle \boldsymbol{v},\boldsymbol{w}\rangle_\phi=1$ and then $(\boldsymbol{v},\boldsymbol{w})$ is a basis for V. Suppose $\boldsymbol{x}=a\boldsymbol{v}+b\boldsymbol{w}$ is in $Rad(V,\phi)$, Then, in particular, $\langle \boldsymbol{x},\boldsymbol{v}\rangle_\phi=0$. However, $\langle \boldsymbol{x},\boldsymbol{v}\rangle_\phi=\langle a\boldsymbol{v}+b\boldsymbol{w},\boldsymbol{v}\rangle_\phi=b$, whence $b=0$. Similarly, $a=0$ and $\boldsymbol{x}=\boldsymbol{0}$.

2. Suppose to the contrary that $dim(M_1)\neq dim(M_2)$. Then without loss of generality we can assume $dim(M_1)=m_1<m_2=dim(M_2)$. Let M_2' be a subspace of $M_2, dim(M_2')=m_1$ and choose bases $(\boldsymbol{v}_{1,,i},\dots,\boldsymbol{v}_{m_1,i})$ for $M_i, i=1,2$. Let σ be the linear transformation from M_1 to M_2' such that $\sigma(\boldsymbol{v}_{2,i})=\boldsymbol{v}_{1,i}$. Then σ is an isometry. By Witt's theorem there exist an isometry S of V such that S restricted to $M_2'=\sigma$. But then $S(M_2)$ is totally singular subspace of V and $S(M_2)$ properly contains M_1, a contradiction.

3. By Witt's theorem there exists an isometry S on V such that $S(X)=Y$. Then $S(X^\perp)=Y^\perp$ so that $X^\perp$ and $Y^\perp$ are isometric.

4. This follows immediately from Theorem (8.15)

5. $Span\begin{pmatrix}0\\0\\1\end{pmatrix}$ and $Span\begin{pmatrix}1\\1\\0\end{pmatrix}$.

6. Set $\mathcal{B}=\mathcal{B}_X\cup\mathcal{B}_Y$. Let $J=\begin{pmatrix}\boldsymbol{0}_m & I_m\\ I_m & \boldsymbol{0}_m\end{pmatrix}$. Suppose T is an operator on V and $\mathcal{M}_T(\mathcal{B},\mathcal{B})=M$. Then T is an isometry if and only if $M^{tr}JM=J$. Now let S be an operator on V and assume that $S(X)=X, S(Y)=Y$. Then $\mathcal{M}_S(\mathcal{B},\mathcal{B})=\begin{pmatrix}M_X & \boldsymbol{0}_m\\ \boldsymbol{0}_m & M_Y\end{pmatrix}$. From the above it follows that S is an isometry if and only if $M_X^{tr}M_Y=I_m$.

7. Let c be a scalar and $\boldsymbol{v}=\boldsymbol{v}_1+\boldsymbol{v}_2$ a vector in $V_1\oplus V_2$. Then

$$\begin{aligned}\phi(c\boldsymbol{v}) &= \phi(c(\boldsymbol{v}_1+vv_2))\\ &= \phi(c\boldsymbol{v}_1+c\boldsymbol{v}_2)\\ &= \phi_1(c\boldsymbol{v}_1)+\phi_2(c\boldsymbol{v}_2)\\ &= c^2\phi_1(\boldsymbol{v}_1)+c^2\phi_2(\boldsymbol{v}_2)\\ &= c^2[\phi_1(\boldsymbol{v}_1)+\phi_2(\boldsymbol{v}_2)\\ &= c^2\phi(\boldsymbol{v}_1+\boldsymbol{v}_2)=c^2\phi(\boldsymbol{v})\end{aligned}$$

Assume $\boldsymbol{v}=\boldsymbol{v}_1+\boldsymbol{v}_2$ and $\boldsymbol{w}=\boldsymbol{w}_1+\boldsymbol{w}_2$ with $\boldsymbol{v}_1,\boldsymbol{w}_1\in V_1$ and $\boldsymbol{v}_2,\boldsymbol{w}_2\in V_2$. Then

$$\langle \boldsymbol{v}_1+\boldsymbol{v}_2,\boldsymbol{w}_1+\boldsymbol{w}_2\rangle_\phi=$$
$$\phi([\boldsymbol{v}_1+\boldsymbol{v}_2]+[\boldsymbol{w}_1+\boldsymbol{w}_2])-\phi(\boldsymbol{v}_1+\boldsymbol{v}_2)-\phi(\boldsymbol{w}_1+\boldsymbol{w}_2)=$$
$$\phi([\boldsymbol{v}_1+\boldsymbol{w}_1]+[\boldsymbol{v}_2+\boldsymbol{w}_2]-\phi(\boldsymbol{v}_1+\boldsymbol{v}_2)-\phi(\boldsymbol{w}_1+\boldsymbol{w}_2)=$$
$$\phi_1(\boldsymbol{v}_1+\boldsymbol{w}_1)+\phi_2(\boldsymbol{v}_2+\boldsymbol{w}_2)-$$

$$\phi_1(\boldsymbol{v}_1) + \phi_2(\boldsymbol{v}_2) + \phi_1(\boldsymbol{w}_1) + \phi_2(\boldsymbol{w}_2)) =$$
$$[\phi_1(\boldsymbol{v}_1 + \boldsymbol{w}_1) - \phi_1(\boldsymbol{v}_1) - \phi_1(\boldsymbol{w}_1)] +$$
$$[\phi_2(\boldsymbol{v}_2 + \boldsymbol{w}_2) - \phi_2(\boldsymbol{v}_2) - \phi_2(\boldsymbol{w}_2)] =$$
$$\langle \boldsymbol{v}_1, \boldsymbol{w}_1\rangle_{\phi_1} + \langle \boldsymbol{v}_2, \boldsymbol{w}_2\rangle_{\phi_2}.$$

8. $H_1 \perp H_2$ is non-degenerate space of dimension four and Witt index 2. It therefore suffices to prove that $E_1 \perp E_2$ has Witt index two. $E_1 \perp E_2$ is isometric to $(\mathbb{F}^4, \phi)$ where $\phi(\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}) = a^2 + ab + b^2\delta + c^2 + d^2 + d^2\delta$. Check that $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \perp \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ and $\phi\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \phi(\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = 0.$

8.5. Real Quadratic Forms

1. $(\pi, \sigma) = (1, 0)$.

2. $(\pi, \sigma) = (2, 1)$.

3. $(\begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix})$.

4. $(\begin{pmatrix} \sqrt{2} \\ -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ 1 \\ 1 \end{pmatrix})$.

5. The number of congruence classes is equal to the number of triples $(\pi, \nu, \zeta) \in \mathbb{N}^3$ such that $\pi + \nu + \zeta = n$. This is $\binom{n+1}{2}$.

6a) Let m be the Witt index. We know the rank, $\rho = n$ and $\pi + \nu = n$. Also, $m = min\{\pi, \nu\}$ and therefore $(\pi, \nu) = (m, n-m)$ or $(n-m, m)$. Since n is odd one of π, ν is even the other is odd and therefore one of $m, n-m$ is even and the other is odd. If $det(A) < 0$ then ν is odd and if $det(A) > 0$ then ν is even. Thus, m and the sign of $det(A)$ determine π and ν. Note, for each m there are two classes of isometry forms.

b) If m is the Witt index and $m < \frac{n}{2}$ then we can have $(\pi, \nu) = (m, n-m)$ or $(n-m, m)$. Thus, there are two isometry classes.

c) If $m = \frac{n}{2}$ then $(\pi, \nu) = (m, m)$ and there is a unique isometry class of forms.

7. That $[\ ,\ ]$ is bilinear follows from the bilinearity of $\langle\ ,\ \rangle$ and the linearity of T. We therefore have to show that $([\boldsymbol{x}, \boldsymbol{y}] = [\boldsymbol{y}, \boldsymbol{x}]$. However $[\boldsymbol{x}, \boldsymbol{y}] = \langle \boldsymbol{x}, T(\boldsymbol{y})\rangle = \langle T(\boldsymbol{x}), \boldsymbol{y}\rangle$ since T is self-adjoint. Since $\langle\ ,\ \rangle$ is symmetric, we have $\langle T(\boldsymbol{x}), \boldsymbol{y}\rangle = \langle \boldsymbol{y}, T(\boldsymbol{x})\rangle = [\boldsymbol{y}, \boldsymbol{x}]$ as required.

8. Fix $\boldsymbol{y} \in V$ and define $F_{\boldsymbol{y}} : V \to \mathbb{R}$ by $F_{\boldsymbol{y}}(\boldsymbol{x}) = [\boldsymbol{x}, \boldsymbol{y}]$. Then $F_{\boldsymbol{y}} \in V' = \mathcal{L}(V, \mathbb{R})$. Then there exists a unique vector $T(\boldsymbol{y}) \in V$ such that $F_{\boldsymbol{y}}(\boldsymbol{x}) = \langle \boldsymbol{x}, T(\boldsymbol{y})\rangle$. We have defined a function $T : V \to V$ such that $[\boldsymbol{x}, \boldsymbol{y}] = \langle \boldsymbol{x}, T(\boldsymbol{y})\rangle$. We have to show that T is linear and then a symmetric operator.

Let $\boldsymbol{y}_1, \boldsymbol{y}_2 \in V$. Then $\langle \boldsymbol{x}, T(\boldsymbol{y}_1, +\boldsymbol{y}_2)\rangle = [\boldsymbol{x}, \boldsymbol{y}_1 + \boldsymbol{y}_2] =$

$$[\boldsymbol{x}, \boldsymbol{y}_1] + [\boldsymbol{x}, \boldsymbol{y}_2] = \langle \boldsymbol{x}, T(\boldsymbol{y}_1)\rangle + \langle \boldsymbol{x}, T(\boldsymbol{y}_2)\rangle =$$
$$\langle \boldsymbol{x}, T(\boldsymbol{y}_1) + T(\boldsymbol{y}_2)\rangle.$$

It follows that $T(\boldsymbol{y}_1 + \boldsymbol{y}_2) = T(\boldsymbol{y}_1) + T(\boldsymbol{y}_2)$.

Now suppose c is a scalar and $\boldsymbol{y} \in V$. Then $[\boldsymbol{x}, c\boldsymbol{y}] = \langle \boldsymbol{x}, T(c\boldsymbol{y})\rangle$. On the other hand, $[\boldsymbol{x}, c\boldsymbol{y}] = c[\boldsymbol{x}, \boldsymbol{y}] = \langle \boldsymbol{x}, T(\boldsymbol{y})\rangle = \langle \boldsymbol{x}, cT(\boldsymbol{y})\rangle$ and consequently, $T(c\boldsymbol{y}) = cT(\boldsymbol{y})$. Thus, T is an operator on V.

Since $[\ ,\ ]$ is a symmetric bilinear form we have $\langle \boldsymbol{x}, T(\boldsymbol{y})\rangle = [\boldsymbol{x}, \boldsymbol{y}] = [\boldsymbol{y}, \boldsymbol{x}] = \langle \boldsymbol{y}, T(\boldsymbol{x})\rangle = \langle T(\boldsymbol{x}), \boldsymbol{y}\rangle$, the latter equality since $\langle\ ,\ \rangle$ is symmetric. This implies that $T = T^*$ and T is a symmetric operator.

9. i) implies ii). Since A is symmetric there exists an orthonormal basis $\mathcal{B} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ of $\mathbb{R}^n$ consisting of eigenvectors. Assume $A\boldsymbol{v}_i = a_i$. Since $\boldsymbol{v}^{tr}A\boldsymbol{v} > 0$ for all $\boldsymbol{v} \in \mathbb{R}^n$, in particular, $\boldsymbol{v}_i^{tr}A\boldsymbol{v}_i = a_i \parallel \boldsymbol{v}_i \parallel^2 > 0$ which

implies that $a_i > 0$. Set $\boldsymbol{u}_i = \frac{1}{\sqrt{a_i}}\boldsymbol{v}_i$ and let Q be the matrix with columns equal to the $\boldsymbol{u}_i$. Then $Q^{tr}AQ = I_n$.

ii) implies iii). Assume A is congruent to the identity matrix. Then there is an invertible matrix P such that $P^{tr}AP = I_n$. Then $(P^{-1})^{tr}I_nP^{-1} = A$. Set $Q = P^{-1}$. Then $A = Q^{tr}Q$.

iii) implies i). Let $\boldsymbol{v} \in \mathbb{R}^n$. Then $\boldsymbol{v}^{tr}A\boldsymbol{v} = (\boldsymbol{v}^{tr}Q^{tr})(Q\boldsymbol{v}) = (Q\boldsymbol{v})^{tr}(Q\boldsymbol{v}) = \| Q\boldsymbol{v} \|^2 > 0$ unless $Q\boldsymbol{v} = \mathbf{0}$. Since Q is invertible, $Q\boldsymbol{v} = \mathbf{0}$ if and only if $\boldsymbol{v} = \mathbf{0}$.

Chapter 9

Sesquilinear Forms and Unitary Spaces

9.1. Basic Properties of Sesquilinear Forms

1. Let $\boldsymbol{v} \in V$ and $c \in \mathbb{F}$. Then $(T \circ S)(c\boldsymbol{v}) = T(S(c\boldsymbol{v})) = T(\sigma(c)S(\boldsymbol{v})) = \tau(\sigma(c))T(S(\boldsymbol{v})) = (\tau \circ \sigma)(c)[T \circ S](\boldsymbol{v})$. Thus, $T \circ S$ is a $\tau \circ \sigma$-semilinear map.

2. Let $f, g \in S_\sigma(V)$. We need to prove that $f + g \in S_\sigma(V)$. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w} \in V$. Then

$$(f+g)(\boldsymbol{v}_1+\boldsymbol{v}_2, \boldsymbol{w}) = f(\boldsymbol{v}_1+\boldsymbol{v}_2, \boldsymbol{w}) + g(\boldsymbol{v}_1+\boldsymbol{v}_2, \boldsymbol{w}) =$$

$$[f(\boldsymbol{v}_1, \boldsymbol{w}) + f(\boldsymbol{v}_2, \boldsymbol{w})] + [g(\boldsymbol{v}_1, \boldsymbol{w}) + g(\boldsymbol{v}_2, \boldsymbol{w})] =$$

$$[f(\boldsymbol{v}_1, \boldsymbol{w}) + g(\boldsymbol{v}_1, \boldsymbol{w})] + [f(\boldsymbol{v}_2, \boldsymbol{w}) + g(\boldsymbol{v}_2, \boldsymbol{w})] =$$

$$(f+g)(\boldsymbol{v}_1, \boldsymbol{w}) + (f+g)(\boldsymbol{v}_2, \boldsymbol{w}).$$

That $(f+g)(\boldsymbol{w}, \boldsymbol{v}_1+\boldsymbol{v}_2) = (f+g)(\boldsymbol{w}, \boldsymbol{v}_1) + (f+g)(\boldsymbol{w}, \boldsymbol{v}_2)$ is proved in exactly the same way.

Now suppose $\boldsymbol{v}, \boldsymbol{w} \in V$ and $a \in \mathbb{F}$. Then

$$(f+g)(a\boldsymbol{v}, \boldsymbol{w}) = f(a\boldsymbol{v}, \boldsymbol{w}) + g(a\boldsymbol{v}, \boldsymbol{w}) =$$

$$af(\boldsymbol{v}, \boldsymbol{w}) + ag(\boldsymbol{v}, \boldsymbol{w}) = a[f(\boldsymbol{v}, \boldsymbol{w}) + g(\boldsymbol{v}, \boldsymbol{w})] =$$

$$a[(f+g)(\boldsymbol{v}, \boldsymbol{w})].$$

Also,

$$(f+g)(\boldsymbol{v}, a\boldsymbol{w}) = f(\boldsymbol{v}, a\boldsymbol{w}) + g(\boldsymbol{v}, a\boldsymbol{w}) =$$

$$\sigma(a)f(\boldsymbol{v}, \boldsymbol{w}) + \sigma(a)g(\boldsymbol{v}, \boldsymbol{w}) =$$

$$\sigma(a)[f(\boldsymbol{v}, \boldsymbol{w}) + g(\boldsymbol{v}, \boldsymbol{w})] =$$

$$\sigma(a)[(f+g)(\boldsymbol{v}, \boldsymbol{w})].$$

Thus, $f + g \in S_\sigma(V)$.

Now assume $f \in S_\sigma(V), c \in \mathbb{F}$. We must show that $cf \in S_\sigma(V)$. Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w} \in V$. Then

$$(cf)(\boldsymbol{v}_1+\boldsymbol{v}_2, \boldsymbol{w}) = c[f(\boldsymbol{v}_1+\boldsymbol{v}_2, \boldsymbol{w})] =$$

$$c[f(\boldsymbol{v}_1, \boldsymbol{w}) + f(\boldsymbol{v}_2, \boldsymbol{w})] = cf(\boldsymbol{v}_1, \boldsymbol{w}) + cf(\boldsymbol{v}_2, \boldsymbol{w}) =$$

$$(cf)(\boldsymbol{v}_1, \boldsymbol{w}) + (cf)(\boldsymbol{v}_2, \boldsymbol{w}).$$

That $(af)(\boldsymbol{w}, \boldsymbol{v}_1+\boldsymbol{v}_2) = (af)(\boldsymbol{w}, \boldsymbol{v}_1) + (af)(\boldsymbol{w}, \boldsymbol{v}_2)$ is proved in exactly the same way.

Finally, let $a \in \mathbb{F}, \boldsymbol{v}, \boldsymbol{w} \in V$. Then

$$(cf)(a\boldsymbol{v}, \boldsymbol{w}) = cf(a\boldsymbol{v}, \boldsymbol{w}) = caf(\boldsymbol{v}, \boldsymbol{w}) =$$

$$acf(\boldsymbol{v}, \boldsymbol{w}) = a[cf(\boldsymbol{v}, \boldsymbol{w})] = a[(cf)(\boldsymbol{v}, \boldsymbol{w})].$$

$$(cf)(\boldsymbol{v}, a\boldsymbol{w}) = cf(\boldsymbol{v}, a\boldsymbol{w}) = c\sigma(a)f(\boldsymbol{v}, \boldsymbol{w}) =$$

$$(\sigma(a)c)f(\boldsymbol{v}, \boldsymbol{w}) = \sigma(a)[cf(\boldsymbol{v}, \boldsymbol{w})] =$$

$$\sigma(a)[(cf)(\boldsymbol{v},\boldsymbol{w})].$$

3. Assume $A = \mathcal{M}_f(\mathcal{B},\mathcal{B}) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$, $[\boldsymbol{u}]_{\mathcal{B}} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, and $[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. Then

$$f(\boldsymbol{u},\boldsymbol{v}) = f(\sum_{i=1}^n b_i\boldsymbol{v}_i, \sum_{j=1}^n c_j\boldsymbol{v}_j) =$$

$$\sum_{i=1}^n\sum_{j=1}^n b_i a_{ij}\sigma(c_j) =$$

$$[\boldsymbol{u}]_{\mathcal{B}}^{tr}A\sigma([\boldsymbol{v}]_{\mathcal{B}}).$$

4. Let $\mathcal{B} = (\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ be a basis. Denote by g_i the σ-semilinear map from V to $\mathbb{F}$ given by $g_i(\boldsymbol{w}) = f(\boldsymbol{v}_i,\boldsymbol{w})$. We claim that $(g_1,\dots,g_n)$ is linearly independent. For suppose $\sum_{i=1}^n a_ig_i = \boldsymbol{0}_{V\to\mathbb{F}}$. Then for all $\boldsymbol{w}\in V$,

$$\sum_{i=1}^n a_ig_i(\boldsymbol{w}) = \sum_{i=1} a_if(\boldsymbol{v}_i,\boldsymbol{w}) = 0$$

Consequently, $f(\sum_{i=1}^n a_i\boldsymbol{v}_i,\boldsymbol{w}) = 0$. Thus, $\sum_{i=1}^m a_i\boldsymbol{v}_i \in Rad_L(f) = \{\boldsymbol{0}_V\}$. Since $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n)$ is linearly independent it follows that $a_1 = \dots = a_n = 0$. Thus, $(g_1,\dots,g_n)$ is linearly independent as claimed. Since the dimension of the space of σ-semilinear maps from V to $\mathbb{F}$ is equal to $dim(V) = n$ it follows that $(g_1,\dots,g_n)$ is a basis for this space. Since F is a σ-semilinear map from V to $\mathbb{F}$ there are unique scalars $b_1,\dots,b_n$ such that $F = \sum_{i=1}^n b_ig_i$. Set $\boldsymbol{v} = b_1\boldsymbol{v}_1+\dots+b_n\boldsymbol{v}_n$. Then $F(\boldsymbol{w}) = f(\boldsymbol{v},\boldsymbol{w})$ for all $\boldsymbol{w}\in V$.

5. i. Assume f is Hermitian and $1\le i,j\le n$. Then $a_{ji} = f(\boldsymbol{v}_j,\boldsymbol{v}_i) = \sigma(f(\boldsymbol{v}_i,\boldsymbol{v}_j)) = \sigma(a_{ij}) = \overline{a_{ij}}$. Thus, $A^{tr} = \overline{A}$. Conversely, if $A^{tr} = \overline{A}$ then $f(\boldsymbol{v}_j,\boldsymbol{v}_i) = \sigma(f(\boldsymbol{v}_i,\boldsymbol{v}_j))$. Suppose $\boldsymbol{v} = b_1\boldsymbol{v}_1+\dots+b_n\boldsymbol{v}_n$ and $\boldsymbol{w} = c_1\boldsymbol{v}_1+\dots+c_n\boldsymbol{v}_n$. Then

$$f(\boldsymbol{w},\boldsymbol{v}) = [\boldsymbol{w}]_{\mathcal{B}}^{tr}A\sigma([\boldsymbol{v}]_{\mathcal{B}}).$$

Since $f(\boldsymbol{w},\boldsymbol{v})$ is a scalar we have

$$[\boldsymbol{w}]_{\mathcal{B}}^{tr}A\sigma([\boldsymbol{v}]_{\mathcal{B}}) = \sigma([\boldsymbol{v}]_{\mathcal{B}}^{tr}A^{tr}[\boldsymbol{w}]_{\mathcal{B}} =$$
$$\sigma([\boldsymbol{v}]_{\mathcal{B}}^{tr}\sigma(A^{tr})\sigma([\boldsymbol{w}]_{\mathcal{B}}) = \sigma(f(\boldsymbol{v},\boldsymbol{w})).$$

ii. This is proved exactly as i.

6. Let g_i be the σ-semilinear map such that $g_i(\boldsymbol{v}_i) = 1$ and $g_i(\boldsymbol{v}_j) = 0$ for $j\ne i$. By Lemma (9.5) there is vector $\boldsymbol{v}_i'$ such that $f(\boldsymbol{v}_i',\boldsymbol{w}) = g_i(\boldsymbol{w})$. Thus, $f(\boldsymbol{v}_i',\boldsymbol{v}_i) = g_i(\boldsymbol{v}_i) = 1$ and $f(\boldsymbol{v}_i',\boldsymbol{v}_j) = g_i(\boldsymbol{v}_j) = 0$.

7. Assume $\sigma^2 = I_{\mathbb{F}}$. Then f is Hermitian, hence reflexive. To see this, let $\boldsymbol{v} = \sum_{i=1}^n a_i\boldsymbol{v}_i$ and $\boldsymbol{w} = \sum_{i=1}^n b_i\boldsymbol{v}_i'$. Then

$$f(\boldsymbol{v},\boldsymbol{w}) = \sum_{i=1}^n a_i\sigma(b_i).$$

On the other hand,

$$f(\boldsymbol{w},\boldsymbol{v}) = \sum_{i=1} b_i\sigma(a_i).$$

If $\sigma^2 = I_{\mathbb{F}}$ then $f(\boldsymbol{w},\boldsymbol{v}) = \sigma(f(\boldsymbol{v},\boldsymbol{w}))$ so f is Hermitian and reflexive.

Assume $dim(V)\ge 2$ and $\sigma^2\ne I_{\mathbb{F}}$. Let $a\in\mathbb{F},\sigma^2(a)\ne a$. Then

$$f(\sigma(a)\boldsymbol{v}_1+\boldsymbol{v}_2,\boldsymbol{v}_1-a\boldsymbol{v}_2) = \sigma(a)-\sigma(a) = 0.$$

On the other hand,

$$f\boldsymbol{v}_1-a\boldsymbol{v}_2,\sigma(a)\boldsymbol{v}_1+\boldsymbol{v}_2) = \sigma^2(a)-a\ne 0.$$

8. We can view $\mathbb{F}$ as a two dimensional vector space over $\mathbb{E}$. The map $tr_{\mathbb{F}/\mathbb{W}}$ is a linear map. Therefore either $tr_{\mathbb{F}/\mathbb{E}}$ is the zero map or it is surjective. Suppose the characteristic of $\mathbb{F}$ is not two. Then the map $a\to -a$ is not an automorphism of $\mathbb{F}$ and there exists an a such that $a+\sigma(a)\ne 0$. If follows that in this case the map is surjective. Suppose the characteristic is two. Since $\sigma\ne I_{\mathbb{F}}$ there

is an a such that $a \neq \sigma(a)$ which implies that $a+\sigma(a) \neq 0$ and once again the map is surjective.

9.2. Unitary Space

1. Suppose $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linearly dependent. Then for some k, $\boldsymbol{v}_k$ is a linear combination of $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1})$. Say $\boldsymbol{v}_k = a_1\boldsymbol{v}_1 + \cdots + a_{k-1}\boldsymbol{v}_{k-1}$. Then $f(\boldsymbol{v}_k, \boldsymbol{v}_k) = f(\boldsymbol{v}_k, a_1\boldsymbol{v}_1 + \cdots + a_{k-1}\boldsymbol{v}_{k-1}) = a_1 f(\boldsymbol{v}_k, \boldsymbol{v}_1) + \cdots + a_{k-1} f(\boldsymbol{v}_k, \boldsymbol{v}_{k-1}) = 0$ which contradicts the assumption that $f(\boldsymbol{v}_i, \boldsymbol{v}_i) \neq 0$ for $1 \leq i \leq n$.

2. If T is an isometry then $f(T(\boldsymbol{v}_i), T(\boldsymbol{v}_j)) = f(\boldsymbol{w}_i, \boldsymbol{w}_j)$ for all i and j. Conversely, assume that $f(\boldsymbol{w}_i, \boldsymbol{w}_j) = f(\boldsymbol{v}_i, \boldsymbol{v}_j)$ for all i and j. Assume $\boldsymbol{v} = \sum_{i=1}^n a_i \boldsymbol{v}_m \boldsymbol{w} = \sum_{i=1} b_i \boldsymbol{v}_i$. Then $f(\boldsymbol{v}, \boldsymbol{w}) =$

$$f(\sum_{i=1^n} a_i \boldsymbol{v}_i, \sum_{i=1}^n b_i \boldsymbol{v}_i) = \sum_{i=1}^n a_i \overline{b_i} f(\boldsymbol{v}_i, \boldsymbol{v}_i).$$

On the other hand, $T(\boldsymbol{v}) = \sum_{i=1}^n a_i T(\boldsymbol{v}_i) = \sum_{i=1}^n a_i \boldsymbol{w}_i$ and, similarly, $T(\boldsymbol{w}) = \sum_{i=1}^n b_i \boldsymbol{w}_i$. Then

$$f(T(\boldsymbol{v}), T(\boldsymbol{w})) = f(\sum_{i=1}^n a_i \boldsymbol{w}_i, \sum_{i=1}^n b_i \boldsymbol{w}_i) =$$

$$\sum_{i=1}^n a_i \overline{b_i} f(\boldsymbol{w}_i, \boldsymbol{w}_j) = f(\boldsymbol{v}, \boldsymbol{w}).$$

3. Since f is non-degenerate there exists a vector $\boldsymbol{u}$ such that $f(\boldsymbol{v}, \boldsymbol{u}) \neq 0$. Now the result follows from Lemma (9.14).

4. Let $\boldsymbol{x} = a\boldsymbol{v}_1 + b\boldsymbol{w}_1, \boldsymbol{y} = c\boldsymbol{v}_1 + d\boldsymbol{w}_1$. Then $f(\boldsymbol{x}, \boldsymbol{y}) = a\sigma(d) + b\sigma(c)$. By the definition of T we have $T(\boldsymbol{x}) = a\boldsymbol{v}_2 + b\boldsymbol{w}_2, T(\boldsymbol{y}) = c\boldsymbol{v}_2 + d\boldsymbol{w}_2$. We then have $f(T(\boldsymbol{x}), T(\boldsymbol{y})) = f(a\boldsymbol{v}_2 + b\boldsymbol{w}_2, c\boldsymbol{v}_2 + d\boldsymbol{w}_2) = a\sigma(d) + b\sigma(c)$.

5. Suppose to the contrary that $dim(U_1) \neq dim(U_2)$. Then without loss of generality we can assume $dim(U_1) = l < m = dim(U_2)$. Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l)$ be a basis for U_2 and $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ a basis for U_2. Let τ be the linear transformation from $Span(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_l) \to U_2$ such that $\tau(\boldsymbol{w}_i) = \boldsymbol{v}_i$. Then τ is an isometry. By Witt's theorem there is an isometry T of V such that T restricted to $Span(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_l)$ is τ. Set $U_2' = T(U_2)$. Then U_2' is a totally singular subspace of V and U_2' properly contains U_1, contradicting the assumption that U_1 is a maximal totally singular subspace.

6. By Witt's theorem there is an isometry T of V such that $T(U_1) = U_2$. Let $\boldsymbol{x} \in U_2$ and $\boldsymbol{y} \in T(U_1)^\perp$. We claim that $f(\boldsymbol{x}, \boldsymbol{y}) = 0$ from which it will follow that $T(U_1^\perp) \subset U_2^\perp$. Since $U_2 = T(U_1)$ there is a vector $\boldsymbol{u} \in U_1$ such that $\boldsymbol{x} = T(\boldsymbol{u})$. Since $\boldsymbol{y} \in T(U_1^\perp)$ there is a $\boldsymbol{v} \in U_1^\perp$ such that $\boldsymbol{y} = T(\boldsymbol{v})$. Now $f(\boldsymbol{x}, \boldsymbol{y}) = f(T(\boldsymbol{u}), T(\boldsymbol{v})) = T(\boldsymbol{u}, \boldsymbol{v}) = 0$. Now $dim(U_1^\perp) = dim(V) - dim(U_1) = dim(V) - dim(U_2) = U_2^\perp$. Since $dim(T(U_1^\perp)) = dim(U_1^\perp)$ it now follows that $T(U_1^\perp) = U_2$.

7. Let $\boldsymbol{x}$ be an anisotropic vector and let $\boldsymbol{y} \in \boldsymbol{x}^\perp$. Since N is surjective, we can assume that $f(\boldsymbol{x}, \boldsymbol{x}) = 1$ and $f(\boldsymbol{y}, \boldsymbol{y}) = -1$. Then $\boldsymbol{v} = \boldsymbol{x} - \boldsymbol{y}$ is isotropic.

8. We prove this by induction on $n \geq 2$ (there is nothing to prove for $n = 1$). The base case is covered by Exercise 7. Assume that $n > 2$ and that the result has been proved for all non-degenerate unitary spaces (W, g) where $dim(W) < n$. and that $dim(V) = n$. By Exercise 7 the space is isotropic. By Exercise 3 there exists a hyperbolic pair, $(\boldsymbol{x}, \boldsymbol{y})$. $U = Span(\boldsymbol{x}, \boldsymbol{y})$ is non-degenerate and therefore $U^\perp$ is a non-degenerate subspace of dimension $n-2$. By the inductive hypothesis there exists an totally isotropic subspace $M, dim(M) = \lfloor \frac{n-2}{2} \rfloor$. Then $M' = M + Span(\boldsymbol{x})$ is a totally isotropic subspace and $dim(M') = \lfloor \frac{n-2}{2} \rfloor + 1 = \lfloor \frac{n}{2} \rfloor$.

9. Let I be the set of all isotropic vectors and set $U = Span(I)$. If we can show that $Span(I) = V$ then I contains a basis of V and are done. Fix an isotropic vector $\boldsymbol{x}$ and let $\boldsymbol{y} \in V$ be arbitrary. If $\boldsymbol{y}$ is isotropic then $\boldsymbol{y} \in Span(I)$ and there is nothing to prove. So assume $\boldsymbol{y}$ is anisotropic. If $f(\boldsymbol{x}, \boldsymbol{y}) \neq 0$ then $Span(\boldsymbol{x}, \boldsymbol{y})$ is non-degenerate and by Corollary (9.3) there is an isotropic

vector $\boldsymbol{z} \in Span(\boldsymbol{x}, \boldsymbol{y})$ such that $f(\boldsymbol{x}, \boldsymbol{z}) = 1$. Then $\boldsymbol{y} \in Span(\boldsymbol{x}, \boldsymbol{y}) = Span(\boldsymbol{x}, \boldsymbol{z}) \subset Span(I)$. We may therefore assume that $\boldsymbol{x} \perp \boldsymbol{y}$. Now $\boldsymbol{y}^{\perp}$ is non-degenerate and contains $\boldsymbol{x}$ so there is an isotropic vector $\boldsymbol{u} \in \boldsymbol{y}^{\perp}$ such that $f(\boldsymbol{x}, \boldsymbol{u}) = 1$ and then $U = Span(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u})$ is a non-degenerate three dimensional subspace. Let W be a two dimensional subspace of U containing $\boldsymbol{x}$ such that $W \neq Span(\boldsymbol{x}, \boldsymbol{y}), Span(\boldsymbol{x}, \boldsymbol{u})$. Then W is non-degenerate and by Corollary (9.3) there is an isotropic vector $\boldsymbol{z} \in W$ such that $f(\boldsymbol{x}, \boldsymbol{z}) = 1$ and $W = Span(\boldsymbol{x}, \boldsymbol{z})$. Then $\boldsymbol{y} \in U = Span(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{w}) \subset Span(I)$.

10. The proof is by induction on $n = dim(V)$. If $dim(V) = 1$ then any non-zero vector is an orthogonal basis. So assume $n > 1$ and the result is true for all non-degenerate spaces of dimension $n-1$. Since (V, f) is non-degenerate there exists a vector $\boldsymbol{v}$ such that $f(\boldsymbol{v}, \boldsymbol{v}) \neq 0$. Set $U = \boldsymbol{v}^{\perp}$, a non-degenerate subspace of dimension $n - 1$. By the inductive hypothesis there exists an orthogonal basis $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1})$ for U. Set $\boldsymbol{v}_n = \boldsymbol{v}$. Then $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is an orthogonal basis for V.

Chapter 10

Tensor Products

10.1. Introduction to Tensor Products

1. Assume V_1, V_2 are finite dimensional and suppose $\mathcal{B}_1 = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m), \mathcal{B}_2 = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$. If $\widehat{f}$ exists then we must have

$$\widehat{f}(\sum_{i=1}^{m} a_i \boldsymbol{v}_i, \sum_{j=1}^{n} b_j \boldsymbol{w}_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j f(\boldsymbol{v}_i, \boldsymbol{w}_j).$$

In fact, defining $\widehat{f}$ in this way gives a bilinear function with $\widehat{f}$ restricted to $\mathcal{B}_1 \times \mathcal{B}_2$ equal to f. On the other hand, if V_1 or V_2 is not finite dimensional, then for any finite subset $\mathcal{B}_1'$ of $\mathcal{B}_1$ and $\mathcal{B}_2'$ of $\mathcal{B}_2$ we can use this to define a bilinear map on $Span(\mathcal{B}_1') \times Span(\mathcal{B}_2')$. One can then use Zorn's lemma to show that these can be extended to all of $V_1 \times V_2$. We omit the details.

,

2. Define a map $\theta : V_1 \times V_2$ to $V_2 \otimes V_1$ by $\theta(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{y} \otimes \boldsymbol{x}$. This map is bilinear. Since $V_1 \otimes V_2$ is universal for $V_1 \times V_2$ there exists a linear transformation $T : V_1 \otimes V_2 \to V_2 \otimes V_1$ such that $T(\boldsymbol{x} \otimes \boldsymbol{y}) = \boldsymbol{y} \otimes \boldsymbol{x}$. In a similar way we get a linear transformation $S : V_2 \otimes V_1 \to V_1 \otimes V_2$ such that $S(\boldsymbol{y} \otimes \boldsymbol{x}) = \boldsymbol{x} \otimes \boldsymbol{y}$. Then $ST = I_{V_1 \otimes V_2}$ and $TS = I_{V_2 \otimes V_1}$.

3. Let $\boldsymbol{x}, \boldsymbol{x}_1, \boldsymbol{x}_2 \in V_1, \boldsymbol{y}, \boldsymbol{y}_1, \boldsymbol{y}_2 \in V_2$ and $c \in \mathbb{F}$. Since $f_1 \in \mathcal{L}(V_1, \mathbb{F})$ we have

$$f(\boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y}) = f_1(\boldsymbol{x}_1 + \boldsymbol{x}_2) f_2(\boldsymbol{y}) =$$

$$[f_1(\boldsymbol{x}_1) + f_1(\boldsymbol{x}_2)] f_2(\boldsymbol{y})$$

By the distributive property in $\mathbb{F}$ we can conclude that

$$[f_1(\boldsymbol{x}_1) + f_1(\boldsymbol{x}_2)] f_2(\boldsymbol{y}) = f_1(\boldsymbol{x}_1) f_2(\boldsymbol{y}) + f_1(\boldsymbol{x}_2) f_2(\boldsymbol{y}) =$$

$$f(\boldsymbol{x}_1, \boldsymbol{y}) + f(\boldsymbol{x}_2, \boldsymbol{y})$$

In exactly the same way we can prove that $f(\boldsymbol{x}, \boldsymbol{y}_1 + \boldsymbol{y}_2) = f(\boldsymbol{x}, \boldsymbol{y}_1) + f(\boldsymbol{x}, \boldsymbol{y}_2)$.

Now to the scalar property: $f(c\boldsymbol{x}, \boldsymbol{y}) = f_1(c\boldsymbol{x}) f_2(\boldsymbol{y}) = [c f_1(\boldsymbol{x}) f_2(\boldsymbol{y})] = c f(\boldsymbol{x}, \boldsymbol{y})$. Similarly, $f(\boldsymbol{x}, c\boldsymbol{y}) = f_1(\boldsymbol{x}) f_2(c\boldsymbol{y}) = f_1(\boldsymbol{x})[c f_2(\boldsymbol{y})] = [f_1(\boldsymbol{x}) c] f_2(\boldsymbol{y}) = [c f_1(\boldsymbol{x})] f_2(\boldsymbol{y}) = c[f_1(\boldsymbol{x}) f_2(\boldsymbol{y}) = c f(\boldsymbol{x}, \boldsymbol{y})$.

4. Let $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ be linearly independent in V and $(\boldsymbol{w}_1, \boldsymbol{w}_2)$ be linearly independent in W and set $\boldsymbol{x} = \boldsymbol{v}_1 \otimes \boldsymbol{w}_1 + \boldsymbol{v}_2 \otimes \boldsymbol{w}_2$. Then $\boldsymbol{x}$ is indecomposable. Suppose to the contrary that there are vectors $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$ such that $\boldsymbol{x} = \boldsymbol{v} \otimes \boldsymbol{w}$. We can assume that $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ and $(\boldsymbol{w}_1, \boldsymbol{w}_2)$ can each be extended to an independent sequence $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$, respectively, such that $\boldsymbol{v} \in Span(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ and $\boldsymbol{w} \in Span(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$. Write $\boldsymbol{v} = a_1\boldsymbol{v}_1 + \cdots + a_m\boldsymbol{v}_m, \boldsymbol{w} = b_1\boldsymbol{w}_1 + \cdots + b_n\boldsymbol{w}_n$. Then

$$\boldsymbol{v} \otimes \boldsymbol{w} = \sum_{i,j} a_i b_j \boldsymbol{v}_i \otimes \boldsymbol{w}_j.$$

Since $\{\boldsymbol{v}_i \otimes \boldsymbol{w}_j | 1 \le i \le m, 1 \le j \le n\}$ is linearly independent we must have $a_1 b_2 = 0 = a_2 b_1$ and $a_1 b_1 = 1 = a_2 b_2$ which gives a contradiction.

5. Set $W' = Span(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$, a finite dimensional subspace of W. Let $(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_s)$ be a basis for W'. We can express each $\boldsymbol{w}_j$ as a linear combination of $(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_s) : \boldsymbol{w}_j = \sum_{i=1}^{s} a_{ij}\boldsymbol{z}_i$. Now $\sum_{j=1}^{n} \boldsymbol{x}_j \otimes \boldsymbol{w}_j = \sum_{j=1}^{n}\sum_{i=1}^{s} a_{ji}\boldsymbol{x}_j \otimes \boldsymbol{z}_i$. However, $\{\boldsymbol{x}_j \otimes \boldsymbol{z}_i | 1 \le j \le n, 1 \le i \le s\}$ is linearly indendent and therefore each $a_{ji} = 0$ whence $\boldsymbol{w}_i = \boldsymbol{0}_W$ for all i.

6. Before proceeding we claim that for any basis $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s)$ of V and any $\boldsymbol{z} \in Z$ there are vectors $\boldsymbol{y}_i \in W$ such that $\boldsymbol{z} = f(\boldsymbol{x}_1, \boldsymbol{y}_1) + \cdots + f(\boldsymbol{x}_s, \boldsymbol{y}_s)$. Thus, assume that $\boldsymbol{z} = f(\boldsymbol{v}_1, \boldsymbol{w}_1) + \cdots + f(\boldsymbol{v}_m, \boldsymbol{w}_m)$. We can express $\boldsymbol{v}_j$ in terms of the basis $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_s)$: $\boldsymbol{v}_j = \sum_{i=1}^{s} a_{ij}\boldsymbol{x}_i$. Now set $\boldsymbol{y}_i = \sum_{j=1}^{m} a_{ij}\boldsymbol{w}_j$. Then by the bilinearity of f it follows that

$$\boldsymbol{z} = f(\boldsymbol{v}_1, \boldsymbol{w}_1) + \cdots + f(\boldsymbol{v}_m, \boldsymbol{w}_m) =$$

$$f(\boldsymbol{x}_1, \boldsymbol{y}_1) + \cdots + f(\boldsymbol{x}_s, \boldsymbol{y}_s).$$

Now by hypothesis b) it follows for $\boldsymbol{z} \in Z$ there are unique $\boldsymbol{y}_i \in W$ such that $\boldsymbol{z} = f(\boldsymbol{x}_1, \boldsymbol{y}_1) + \cdots + f(\boldsymbol{x}_s, \boldsymbol{y}_s)$.

Now assume that $g : V \times W \to X$ is a bilinear form. We need to show that there exists a unique linear map $\widehat{g} : Z \to X$ such that $\widehat{g} \circ f = g$. Clearly the only possible way to define g is as follows: suppose $\boldsymbol{z} = \sum_{i=1}^{s} f(\boldsymbol{x}_i, \boldsymbol{y}_i)$. Then $\widehat{g}(\boldsymbol{z}) = \sum_{i=1}^{s} g(\boldsymbol{x}_i, \boldsymbol{y}_i)$. By the uniqueness of expression for $\boldsymbol{z}$, $\widehat{g}$ is a well-defined function. However, we need to demonstrate that it is linear. Thus, let $\boldsymbol{z}, \boldsymbol{z}' \in Z$. We need to prove that $\widehat{g}(\boldsymbol{z} + \boldsymbol{z}') = \widehat{g}(\boldsymbol{z}) + \widehat{g}(\boldsymbol{z}')$.

Assume that $\boldsymbol{z} = \sum_{i=1}^{s} f(\boldsymbol{x}_i, \boldsymbol{y}_i)$ and $\boldsymbol{z}' = \sum_{i=1}^{s} f(\boldsymbol{x}_i, \boldsymbol{y}_i')$. Since f is bilinear we have $\boldsymbol{z} + \boldsymbol{z}' = \sum_{i=1}^{s} f(\boldsymbol{x}_i, \boldsymbol{y}_i + \boldsymbol{y}_i')$. Then

$$\widehat{g}(\boldsymbol{z} + \boldsymbol{z}') = \sum_{i=1}^{s} g(\boldsymbol{x}_i, \boldsymbol{y}_i + \boldsymbol{y}_i')$$

By the bilinearity of g we have

$$\sum_{i=1}^{s} g(\boldsymbol{x}_i, \boldsymbol{y}_i + \boldsymbol{y}_i') = \sum_{i=1}^{s}(g(\boldsymbol{x}_i, \boldsymbol{y}_i) + g(\boldsymbol{x}_i, \boldsymbol{y}_i') =$$

$$\sum_{i=1}^{s} g(\boldsymbol{x}_i, \boldsymbol{y}_i) + \sum_{i=1}^{s} g(\boldsymbol{x}_i, \boldsymbol{y}_i') =$$

$$\widehat{g}(\boldsymbol{z}) + \widehat{g}(\boldsymbol{z}').$$

Next assume that $\boldsymbol{z} \in Z$ and $c \in \mathbb{F}$. If $\boldsymbol{z} = \sum_{i=1}^{s} f(\boldsymbol{x}_i, \boldsymbol{y}_i)$ then by the bilinearity of f we have $c\boldsymbol{z} = \sum_{i=1}^{s} f(\boldsymbol{x}_i, c\boldsymbol{y}_i)$ so that

$$\widehat{g}(c\boldsymbol{z}) = \sum_{i=1}^{s} g(\boldsymbol{x}_i, c\boldsymbol{y}_i) = \sum_{i=1}^{s} cg(\boldsymbol{x}_i, \boldsymbol{y}_i) =$$

$$c\sum_{i=1}^{s} g(\boldsymbol{x}_i, \boldsymbol{y}_i) = c\widehat{g}(\boldsymbol{z}).$$

7. By the universal property of the tensor product we know for all $f \in B(V, W; Z)$ there exists a linear map $\widehat{f} : V \otimes W \to Z$. Let θ denote the map $f \to \widehat{f}$ so that θ is a map from $B(V, W; Z)$ to $\mathcal{L}(V \otimes W, Z)$. We need to prove that θ is linear and bijective.

We first show that θ is additive: Assume $f, g \in B(V, W; Z)$. Then $\widehat{f}, \widehat{g}$ are the uniique linear maps from $V \otimes W$ to Z such that $\widehat{f}(\boldsymbol{v} \otimes \boldsymbol{w}) = f(\boldsymbol{v}, \boldsymbol{w})$ and $\widehat{g}(\boldsymbol{v} \otimes \boldsymbol{w}) = g(\boldsymbol{v}, \boldsymbol{w})$. Now $\widehat{f} + \widehat{g}$ is a linear map from $V \otimes W$ to Z. Computing $(\widehat{f} + \widehat{g})(\boldsymbol{v} \otimes \boldsymbol{w})$ we get

$$\widehat{f}(\boldsymbol{v}\otimes\boldsymbol{w}) + \widehat{g}(\boldsymbol{v}\otimes\boldsymbol{w}) = f(\boldsymbol{v}, \boldsymbol{w}) + g(\boldsymbol{v}, \boldsymbol{w}) = (f+g)(\boldsymbol{v}, \boldsymbol{w}).$$

By uniqueness, $\widehat{f + g} = \widehat{f} + \widehat{g}$.

We next show homogeneity holds: if c is a scalar then $\widehat{cf} = c\widehat{f}$.

$$\widehat{cf}(\boldsymbol{v} \otimes \boldsymbol{w}) = (cf)(\boldsymbol{v}, \boldsymbol{w}) = cf(\boldsymbol{v}, \boldsymbol{w}) = c\widehat{f}(\boldsymbol{v} \otimes \boldsymbol{w})$$

as required.

It remains to show that θ is an isomorphism. Injectivity is easy: If $f, g \in B(V, W; Z)$ are distinct then there exists

$\boldsymbol{v} \in V, \boldsymbol{w} \in W$ such that $f(\boldsymbol{v}, \boldsymbol{w}) \neq g(\boldsymbol{v}, \boldsymbol{w})$. Then $\widehat{f}(\boldsymbol{v} \otimes \boldsymbol{w}) = f(\boldsymbol{v}, \boldsymbol{w}) \neq g(\boldsymbol{v}, \boldsymbol{w}) = \widehat{g}(\boldsymbol{v} \otimes \boldsymbol{w})$.

It remains to show surjectivity. Assume $T \in \mathcal{L}(V \otimes W, Z)$. Define $t : V \times W \to Z$ by $t(\boldsymbol{v}, \boldsymbol{w}) = T(\boldsymbol{v} \otimes \boldsymbol{w})$. Since T is linear and the tensor product is bilinear it follows that t is bilinear. Thus, $t \in B(V, W; Z)$. Clearly $\widehat{t} = T$ and so θ is surjective.

8. The map $\mu : \mathbb{F} \times V \to V$ given by $\mu(c, \boldsymbol{v}) = c\boldsymbol{v}$ is bilinear. We therefore have a linear transformation from $\mathbb{F} \otimes V$ to V such that $c \otimes \boldsymbol{v} = c\boldsymbol{v}$. Denote this map by T. Define $S : V \to \mathbb{F} \otimes V$ by $S(\boldsymbol{v}) = 1 \otimes \boldsymbol{v}$. This map is linear. Consider the composition $ST(c \otimes \boldsymbol{v}) = S(c\boldsymbol{v}) = 1 \otimes (c\boldsymbol{v}) = c \otimes \boldsymbol{v}$ so that $ST = I_{\mathbb{F} \otimes V}$. Similarly, $TS(\boldsymbol{v}) = T(1 \otimes \boldsymbol{v}) = \boldsymbol{v}$ and $TS = I_V$.

9. To avoid confusion we denote the tensor product of X and Y by $X \otimes' Y$ (as well as products of elements in this space). Define $\theta : X \times Y \to V \otimes W$ by $\theta(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x} \otimes \boldsymbol{y}$ (since this is in $V \otimes W$ there is no prime). By the universal property for the tensor product there exists a linear map $\widehat{\theta} : X \otimes' Y \to V \otimes W$ such that $\widehat{\theta}(\boldsymbol{x} \otimes' \boldsymbol{y}) = \theta(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x} \otimes \boldsymbol{y}$ (there is no $'$ in the latter expression since this is the tensor product of the two elements in $V \otimes W$). Since Z is the subspace of $V \otimes W$ generated by all elements $\boldsymbol{x} \otimes \boldsymbol{y}$, $Range(\widehat{\theta}) = Z$. We need to prove that $\widehat{\theta}$ is injective. Suppose $\widehat{\theta}(\boldsymbol{u}) = \boldsymbol{0}_{V \otimes W}$ where $\boldsymbol{u} \in X \otimes' Y$. Then there exists a linearly independent sequence $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m)$ in X and a linearly independent sequence $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ in Y and scalars a_{ij} such that $\boldsymbol{u} = \sum_{ij} a_{ij} \boldsymbol{x}_i \otimes' \boldsymbol{y}_j'$.

We then have $\theta(\sum_{i,j} a_{ij} \boldsymbol{x}_i \otimes' \boldsymbol{y}_j) = \boldsymbol{0}_{V \otimes W}$. However, $\theta(\sum_{i,j} a_{ij} \boldsymbol{x}_i \otimes' \boldsymbol{y}_j) = \sum_{i,j} a_{ij} \boldsymbol{x}_i \otimes \boldsymbol{y}_j$ (now this tensor product is in $V \otimes W$). The vectors $\{\boldsymbol{x}_i \otimes \boldsymbol{y}_j \in V \otimes W | 1 \leq i \leq m, 1 \leq j \leq n\}$ are linearly independent and therefore $a_{ij} = 0$ and $\boldsymbol{u} = \boldsymbol{0}_{X \otimes' Y}$.

10. Let $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)$ be a basis of $Y_1 \cap Y_2$. Let $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_l)$ be a basis for Y_1 and $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_m)$ be a basis for Y_2. Suppose $\boldsymbol{u} \in (V \otimes Y_1) \cap (V \otimes Y_2)$. Since $\boldsymbol{u} \in V \otimes Y_1$ there are unique vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_l \in V$ such that $\boldsymbol{u} = \boldsymbol{u}_1 \otimes \boldsymbol{y}_1 + \cdots + \boldsymbol{u}_k \otimes \boldsymbol{y}_k + \boldsymbol{v}_1 \otimes \boldsymbol{w}_1 + \cdots + \boldsymbol{v}_l \otimes \boldsymbol{w}_l$.

On the other hand since $\boldsymbol{u} \in V \otimes Y_2$ there are unique vectors $\boldsymbol{u}_1', \ldots, \boldsymbol{u}_k', \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ such that $\boldsymbol{u} = \boldsymbol{u}_1' \otimes \boldsymbol{y}_1 + \cdots + \boldsymbol{u}_k' \otimes \boldsymbol{y}_k + \boldsymbol{x}_1 \otimes \boldsymbol{z}_1 + \cdots + \boldsymbol{x}_m \otimes \boldsymbol{z}_m$. However, this implies that

$$(\boldsymbol{u}_1 - \boldsymbol{u}_1') \otimes \boldsymbol{y}_1 + \cdots + (\boldsymbol{u}_k - \boldsymbol{u}_k') \otimes \boldsymbol{y}_k +$$

$$\boldsymbol{v}_1 \otimes \boldsymbol{w}_1 + \cdots + \boldsymbol{v}_l \otimes \boldsymbol{w}_l - \boldsymbol{x}_1 \otimes \boldsymbol{z}_1 - \cdots - \boldsymbol{x}_m \otimes \boldsymbol{z}_m = \boldsymbol{0}_{V \otimes W}.$$

Since $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_l, \boldsymbol{z}_1, \ldots, \boldsymbol{z}_m)$ is linearly independent by Exercise 5 we have $\boldsymbol{u}_1 - \boldsymbol{u}_1' = \cdots = \boldsymbol{u}_k - \boldsymbol{u}_k' = \boldsymbol{v}_1 = \cdots = \boldsymbol{v}_l = \boldsymbol{x}_1 = \cdots = \boldsymbol{x}_m = \boldsymbol{0}_V$. Thus, $\boldsymbol{u} \in V \otimes (Y_1 \cap Y_2)$.

10.2. Properties of Tensor Products

1. If π fixes one of 1, 2 or 3 this follows from the fact that $V \otimes W$ is isomorphic to $W \otimes V$ by Exercise 1 of Section (10.1). We illustrate the proof for $\pi = (123)$. Let $f : V_1 \times V_2 \times V_3 \to V_2 \otimes V_3 \otimes V_1$ given by $f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \boldsymbol{y} \otimes \boldsymbol{z} \otimes \boldsymbol{x}$. This is a 3-linear map. By the universal property of the tensor product we have a linear map $\widehat{f} : V_1 \otimes V_2 \otimes V_3 \to V_2 \otimes V_3 \otimes V_1$ such that $\widehat{f}(\boldsymbol{x} \otimes \boldsymbol{y} \otimes \boldsymbol{z}) = \boldsymbol{y} \otimes \boldsymbol{z} \otimes \boldsymbol{z}$.

Likewise can define a trilinear map $g : V_2 \times V_3 \times V_1 \to V_1 \otimes V_2 \otimes V_3$ given by $g(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{x}) = \boldsymbol{x} \otimes \boldsymbol{y} \otimes \boldsymbol{z}$. There is then a linear map $\widehat{g} : V_2 \otimes V_3 \otimes V_1$ such that $\widehat{g}(\boldsymbol{y} \otimes \boldsymbol{z} \otimes \boldsymbol{x}) = \boldsymbol{x} \otimes \boldsymbol{y} \otimes \boldsymbol{z}$.

$\widehat{g}\widehat{f} = I_{V_1 \otimes V_2 \otimes V_3}$ and $\widehat{f}\widehat{g} = I_{V_2 \otimes V_3 \otimes V_1}$.

2. If $\boldsymbol{v}_i \in V_i$ then $S_1 \otimes \cdots \otimes S_m)(\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_m) = S_1(\boldsymbol{v}_1) \otimes \cdots \otimes S_m(\boldsymbol{v}_m) \in R_1 \otimes \cdots \otimes R_m$. $V_1 \otimes \cdots \otimes V_m$ is generated by all vectors of the form $\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_m$ it follows that $Range(S_1 \otimes \cdots \otimes S_m) \subset R_1 \otimes \cdots \otimes R_m$. We need to prove the reverse inclusion.

Suppose $\boldsymbol{r}_i \in R_i = Range(S_i)$. Let $\boldsymbol{v}_i \in V_i$ such that $S(\boldsymbol{v}_i) = \boldsymbol{r}_i$. Then $(S_1 \otimes \cdots \otimes S_m)(\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_m) =$

$S(\boldsymbol{v}_1) \otimes \cdots \otimes S(\boldsymbol{v}_m) = \boldsymbol{r}_1 \otimes \cdots \otimes \boldsymbol{r}_m$ and consequently, $R_1 \otimes \cdots \otimes R_m \subset R$.

3. We do a proof by induction on $m \geq 2$. For the base case let $dim(V_i) = n_i$ and $dim(K_i) = k_i$ for $i = 1, 2$. Let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{k_1})$ be a basis for K_1 and extend to a basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n_1})$ for V_1. Similarly, let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{k_2})$ be a basis for K_2 and extend to a basis $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n_2})$ for V_2. Note that the sequence $(S_1(\boldsymbol{v}_{k_1+1}), \dots, S_1(\boldsymbol{v}_{n_1}))$ in W_1 is linearly independent as is the sequence $(S_2(\boldsymbol{u}_{k_2+1}), \dots, S_2(\boldsymbol{u}_{n_2}))$ in W_2. As a consequence the set of vectors $\{S_1(\boldsymbol{v}_i) \otimes S_2(\boldsymbol{u}_j) | k_1 + 1 \leq i \leq n_1, k_2 + 1 \leq j \leq n_2\}$ is linearly independent.

Assume now that $\boldsymbol{x} = \sum_{i,j} a_{ij} \boldsymbol{v}_i \otimes \boldsymbol{u}_j \in K$. Applying $S_1 \otimes S_2$ we get

$$(S_1 \otimes S_2)(\sum_{i,j} a_{ij} \boldsymbol{v}_i \otimes \boldsymbol{v}_j =$$

$$\sum_{i,j} a_{ij} S_1(\boldsymbol{v}_i) \otimes S(\boldsymbol{v}_j) =$$

$$\sum_{i=k_1+1}^{n_1} \sum_{j=k_2+1}^{n_2} a_{ij} S(\boldsymbol{u}_i) \otimes S(\boldsymbol{v}_j) = \boldsymbol{0}_{W_1 \otimes W_2}.$$

Since $\{S_1(\boldsymbol{v}_i) \otimes S_2(\boldsymbol{u}_j) | k_1 + 1 \leq i \leq n_1, k_2 + 1 \leq j \leq n_2\}$ is linearly independent for $k_1 + 1 \leq i \leq n_1, k_2 + 1 \leq j \leq n_2, a_{ij} = 0$. This implies that $\boldsymbol{u} \in K_1 \otimes V_2 + V_1 \otimes K_2 = X_1 + X_2$.

Assume the result is true for tensor product of m spaces and that we have $S_i : V_i \to W_i$ for $1 \leq i \leq m$. Set $V_2' = V_2 \otimes \cdots \otimes V_{m+1}$ and $S_2' = S_2 \otimes \cdots \otimes S_{m+1}$. Then $V_1 \otimes \cdots \otimes V_{m+1}$ is equal to $V_1 \otimes V_2'$ and $S_1 \otimes \cdots \otimes S_{m+1} = S_1 \otimes S_2'$. Set $K_2' = Ker(S_2')$. Also set $Y_2 = K_2 \otimes V_3 \otimes \cdots \otimes V_{m+1}, Y_i = V_2 \otimes \cdots \otimes V_{i-1} \otimes K_i \otimes V_{i+1} \otimes \cdots \otimes V_{m+1}$. By the inductive hypothesis $K_2' = Y_2 + \cdots + Y_{m+1}$. Also, by the base case, $K = K_1 \otimes V_2' + V_1 \otimes K_2'$. However, $V_1 \otimes K_2' = V_1 \otimes (Y_2 + \cdots + Y_{m+1}) = X_2 + \cdots + X_{m+1}$ while $K_1 \otimes V_2' = X_1$. This completes the proof.

4. Let $S : \mathbb{F}^l \to \mathbb{F}^k$ be the transformation such that $S(\boldsymbol{v}) = A\boldsymbol{v}$ and $T : \mathbb{F}^n \to \mathbb{F}^m$ be the transformation such that $T(\boldsymbol{w}) = B\boldsymbol{w}$. Then $A \otimes B$ is the matrix of the transformation $S \otimes T : \mathbb{F}^l \otimes \mathbb{F}^n \to \mathbb{F}^k \otimes \mathbb{F}^m$ with respect to the bases obtained by taking tensor products of the standard basis vectors in lexicographical order. Then $rank(A \otimes B) = dim(Range(S \otimes T))$. On the other hand, $rank(A) = dim(Range(S))$ and $rank(B) = dim(Range(T))$. By Exercise 2 the range of $S \otimes T$ is $Range(S) \otimes Range(T)$ which has dimension $dim(Range(S)) \times dim(Range(T))$.

5. Let $dim(V) = m, dim(W) = n$. Assume neither S nor T is nilpotent. Let $f(x) = \mu_S(x)$ and $g(x) = \mu_T(x)$. Then $f(x)$ does not divide x^m and $g(x)$ does not divide x^n. Let $p(x) \neq x$ be an irreducible factor of $f(x)$ and $q(x) \neq x$ be an irreducible factor of $g(x)$. Let $\boldsymbol{x} \in V$ such that $\mu_{S,\boldsymbol{x}}(x) = f(x)$ and $\boldsymbol{y} \in W$ such that $\mu_{T,\boldsymbol{y}}(x) = q(x)$. Set $X = \langle S, \boldsymbol{x} \rangle$ and $Y = \langle T, \boldsymbol{y} \rangle$. Then $X \otimes Y$ is invariant under $S \otimes I_W$ and $I_V \otimes T$ and therefore under $S \otimes T = (S \otimes I_W)(I_V \otimes T)$. Set $Z = X \otimes Y$ and denote by $\overline{S \otimes T}$ the restriction of $S \otimes T$ to Z. It follows from Exercise 2 that $\overline{S \otimes T}$ is surjective, therefore injective, on Z. In particular, $Ker(\overline{S \otimes T})$ is trivial. However, if $S \otimes T$ is nilpotent then for any invariant subspace U of $V \otimes W$ the kernel of the restriction of $S \otimes T$ to U must be non-trivial. Thus, $S \otimes T$ is not nilpotent.

6. Let $\boldsymbol{v}_i \in V$ be an eigenvector of S with eigenvalue α_i and $\boldsymbol{w}_j$ an eigenvector of T with eigenvalue β_j. Then $\boldsymbol{v}_i \otimes \boldsymbol{w}_j$ is an eigenvector of $S \otimes T$ with eigenvalue $\alpha_i \beta_j$. Consequently, $S \otimes T$ is diagonalizable. On the other hand, since the eigenvalues $\alpha_i \beta_j$ are all distinct, the minimum polynomial of $S \otimes T$ is

$$\prod_{i=1}^{m} \prod_{j=1}^{n} (x - \alpha_i \beta_j)$$

has degree mn and therefore $S \otimes T$ is cyclic.

7. For any cyclic diagonalizable operator $S : V \to V$ the operator $S \otimes S : V \otimes V \to V \otimes V$ will not be cyclic. For example, let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be given by multiplication by $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. Then

$$A \otimes A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

So, the eigenvalue 2 occurs with algebraic multiplicity 2 and the operator is not cyclic.

8. Let $\boldsymbol{v} \in V$ such that $\mu_{S,\boldsymbol{v}}(x) = (x-\alpha)^k$ and let $\boldsymbol{w} \in W$ such that $\mu_{T,\boldsymbol{w}}(x) = (x-\beta)^l$. It suffices to prove that $S \otimes T$ restricted to $\langle S, \boldsymbol{v}\rangle \otimes \langle T, \boldsymbol{w}\rangle$ has minimum polynomial dividing $(x-\alpha\beta)^{kl}$.

Set $\boldsymbol{v}_1 = \boldsymbol{v}$ and for $i < k$ set $\boldsymbol{v}_{i+1} = (S - \alpha I_V)\boldsymbol{v}_i$. Similarly, set $\boldsymbol{w}_1 = \boldsymbol{w}$ and for $j < l, \boldsymbol{w}_{j+1} = (T - \beta I_W)\boldsymbol{w}_j$. $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k)$ is a basis for $\langle S, \boldsymbol{v}\rangle$ and $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_l)$ is a basis for $\langle T, \boldsymbol{w}\rangle$. Let $\overline{S}$ be the restriction of S to $\langle S, \boldsymbol{v}\rangle$ and $\overline{T}$ the restriction of T to $\langle T, \boldsymbol{w}\rangle$. The matrix of $\overline{S}$ with respect to $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_k)$ is the $k \times k$ matrix

$$A = \begin{pmatrix} \alpha & 0 & 0 & \dots & 0 \\ 1 & \alpha & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \alpha \end{pmatrix}$$

The matrix of $\overline{T}$ with respect to $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_l)$ is

$$B = \begin{pmatrix} \beta & 0 & 0 & \dots & 0 \\ 1 & \beta & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & \beta \end{pmatrix}$$

The matrix $A \otimes B$ is a lower triangular $kl \times kl$ with $\alpha\beta$ on the diagonal. This implies the result.

9. Let $c, d \in \mathbb{K}$ and $\widehat{\boldsymbol{v}} = \sum_{i=1}^n a_i \otimes_{\mathbb{F}} \boldsymbol{v}_i$. Then

$$(c+d)\widehat{\boldsymbol{v}} = (c+d)\sum_{i=1}^n a_i \otimes_{\mathbb{F}} \boldsymbol{v}_i = \sum_{i=1}^n (c+d)a_i \otimes_{\mathbb{F}} \boldsymbol{v}_i =$$

$$\sum_{i=1}^n (ca_i + da_i) \otimes_{\mathbb{F}} \boldsymbol{v}_i = \sum_{i=1}^n ca_i \otimes_{\mathbb{F}} \boldsymbol{v}_i + \sum_{i=1}^n da_i \otimes_{\mathbb{F}} \boldsymbol{v}_i =$$

$$c\sum_{i=1}^n a_i \otimes_{\mathbb{F}} \boldsymbol{v}_i + d\sum_{i=1}^n a_i \otimes_{\mathbb{F}} \boldsymbol{v}_i =$$

$$c\widehat{\boldsymbol{v}} + d\widehat{\boldsymbol{v}}.$$

We need also to compute $(cd)\widehat{\boldsymbol{v}}$:

$$(cd)\widehat{\boldsymbol{v}} = (cd)\sum_{i=1}^n a_i \otimes_{\mathbb{F}} \boldsymbol{v}_i = \sum_{i=1}^n (cd)a_i \otimes_{\mathbb{F}} \boldsymbol{v}_i =$$

$$\sum_{i=1}^n c(da_i) \otimes_{\mathbb{F}} \boldsymbol{v}_i = c\sum_{i=1}^n da_i \otimes_{\mathbb{F}} \boldsymbol{v}_i =$$

$$c[d\sum_{i=1}^n a_i \otimes_{\mathbb{F}} \boldsymbol{v}_i]$$

10. That $\widehat{\mathcal{B}}$ is linearly independent follows from Exercise 5 of Section (10.1). We show that $\widehat{\mathcal{B}}$ spans $V_{\mathbb{K}}$. Since every element of $V_{\mathbb{K}}$ is a sum of elements of the form $c \otimes_{\mathbb{F}} \boldsymbol{w}$ for $\boldsymbol{w} \in V$ it suffices to prove that this element belong to $Span(\widehat{\mathcal{B}})$. Since $\mathcal{B}$ is a basis for V there exist scalars $a_i \in \mathbb{F}$ such that $\boldsymbol{w} = a_1\boldsymbol{v}_1 + \dots + a_n\boldsymbol{v}_n$. Then

$$c \otimes_{\mathbb{F}} \boldsymbol{w} = c \otimes_{\mathbb{F}} (a_1\boldsymbol{v}_1 + \dots + a_n\boldsymbol{v}_n) =$$

$$(ca_1) \otimes_{\mathbb{F}} \boldsymbol{v}_1 + \dots (ca_n) \otimes_{\mathbb{F}} \boldsymbol{v}_n \in Span(\widehat{\mathcal{B}}).$$

11. Let $dim(V) = m, dim(W) = n$. Then by Exercise 11, $dim_{\mathbb{K}}(V_{\mathbb{K}}) = m, dim_{\mathbb{K}}(W_{\mathbb{K}}) = n$. Then $dim(\mathcal{L}(V,W)) = mn = dim_{\mathbb{K}}(\mathcal{L}(V_{\mathbb{K}}, W_{\mathbb{K}}) = dim_{\mathbb{K}}(\mathcal{L}(V,W)_{\mathbb{K}}$. Thus, $\mathcal{L}(V_{\mathbb{K}}, W_{\mathbb{K}})$ and $\mathcal{L}(V,W)_{\mathbb{K}}$ are isomorphic as spaces over $\mathbb{K}$.

12. Let $\mathcal{B}_1 = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ be a basis for V_1 and $\mathcal{B}_2 = (\boldsymbol{w}_1, \dots, \boldsymbol{w}_n)$ be a basis for V_2. Let $\mathcal{B}'$ be the basis of $V_1 \otimes V_2$ consisting of the set $\{\boldsymbol{u}_i \otimes \boldsymbol{w}_j | 1 \leq i \leq m, 1 \leq j \leq n\}$ ordered lexicographically. Let $A = \mathcal{M}_{S_1}(\mathcal{B}_1, \mathcal{B}_1)$ and $B = \mathcal{M}_{S_2}(\mathcal{B}_2, \mathcal{B}_2)$ so that $\mathcal{M}_{S_1 \otimes S_2}(\mathcal{B}', \mathcal{B}') = A \otimes B$. Assume the $(k,l)-$entry of A is a_{ij} and the $(k,l)-$entry of B is b_{kl}. Then the diagonal

entries of $A \otimes B$ are $a_{ii}b_{jj}$ where $1 \leq i \leq m, 1 \leq j \leq n$. The sum is then $(a_{11} + \cdots + a_{mm})(b_{11} + \cdots + b_{nn}) = Trace(A)Trace(B)$.

13. Suppose E is an elimination matrix and is upper triangular. Then $E \otimes I_n$ is also upper triangular with 1's on the diagaonal and therefore $det(E \otimes I_n) = 1 = 1^n$. Similarly, if E is lower triangular then $det(E \otimes I_n) = 1$.

Suppose E is an exchange matrix, specifically obtained by exchanging the i and j rows of the identity matrix I_m. Then $E \otimes I_n$ can be obtained from the identity matrix by exchanging rows $n(i-1)+k$ with row $n(j-1)+k$ for $1 \leq k \leq n$. Therefore $det(E \otimes I_n) = (-1)^n = det(E)^n$.

Finally, assume that E is a scaling matrix obtained by multiplying the i row of I_m by c. Then $E \otimes I_n$ is obtained from the $mn-$identity matrix by multiplying rows $n(i-1)+k$ by c for $1 \leq k \leq n$ and is therefore a diagonal matrix with n diagonal entries equal to c and the remaining equal to 1. Then $det(E \otimes I_n) = c^n = det(E)^n$.

14. If either S_1 or S_2 is singular then so is $S_1 \otimes S_2$ and then the result clearly holds. Therefore we may assume that S_i is invertible for $i = 1, 2$.

Let $\mathcal{B}_1 = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ be a basis for V_1 and $\mathcal{B}_2 = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ be a basis for V_2. Let $\mathcal{B}'$ be the basis of $V_1 \otimes V_2$ consisting of the set $\{\boldsymbol{u}_i \otimes \boldsymbol{w}_j | 1 \leq i \leq m, 1 \leq j \leq n\}$ ordered lexicographically. Let $A = \mathcal{M}_{S_1}(\mathcal{B}_1, \mathcal{B}_1)$ and $B = \mathcal{M}_{S_2}(\mathcal{B}_2, \mathcal{B}_2)$ so that $\mathcal{M}_{S_1 \otimes S_2}(\mathcal{B}', \mathcal{B}') = A \otimes B$. Then $det(S_1 \otimes S_2) = det(A \otimes B)$.

Let $E_1, \ldots, E_k$ be $m \times m$ elementary matrices such that $S_1 = E_1 E_2 \ldots E_k$ and let $F_1, \ldots, F_l$ be $n \times n$ elementary matrices such that $S_2 = F_1 F_2 \ldots F_l$. Then $A \otimes B = (E_1 \otimes I_n) \ldots (E_k \otimes I_n)(I_m \otimes F_1) \ldots (I_m \otimes F_l)$. Since the determinant is multiplicative we have

$$det(A \otimes B) =$$

$$det(E_1 \otimes I_n) \ldots det(E_k \otimes I_n) det(I_m \otimes F_1) \ldots det(I_m \otimes F_l).$$

By Exercise 14, $det(E_i \otimes I_n) = det(E_i)^n$ and $det(I_m \otimes F_j) = det(F_j)^m$. Then

$$det(A \otimes B) =$$

$$det(E_1)^n \ldots det(E_k)^n det(F_1)^m \ldots det(F_l)^m =$$

$$[det(E_1) \ldots det(E_k)]^n [det(F_1) \ldots det(F_l)]^m =$$

$$det(A)^n det(B)^m.$$

10.3. The Tensor Algebra

1. Suppose $f_1, f_2 \in \oplus_{i \in I} V_i$. Set $I_1 = spt(f_1), I_2 = spt(f_2)$. Let $J = I_1 \cap I_2$, $J' = \{j \in I | f_2(j) = -f_1(j)\}$ and $J^* = J \setminus J'$. Further set $I_1' = I_1 \setminus J^*, I_2' = I_2 \setminus J^*$. Then $spt(f_1 + f_2) = I_1 \cup I_2 \setminus J' = I_1' \cup I_2' \cup J^*$. We now compute $G(f_1 + f_2)$:

$$G(f_1 + f_2) = \sum_{i \in spt(f_1+f_2)} g_i([f_1 + f_2](i)) =$$

$$\sum_{i \in I_1' \cup I_2' \cup J^*} g_i([f_1 + f_2](i)) =$$

$$\sum_{i \in I_1'} g_i([f_1 + f_2](i)) + \sum_{i \in I_2'} g_i([f_1 + f_2](i)) +$$

$$\sum_{i \in J^*} g_i([f_1 + f_2](i)) =$$

$$\sum_{i \in I_1'} g_i(f_1(i) + f_2(i)) + \sum_{i = I_2'} g_i(f_1(i) + f_2(i)) +$$

$$\sum_{i \in J^*} g_i(f_1(i) + f_2(i)) =$$

$$\sum_{i \in I_1'} g_i(f_1(i)) + \sum_{i \in I_2'} g_i(f_2(i)) + \sum_{i \in J^*} g_i(f_1(i) + f_2(i)) =$$

$$\sum_{i \in I_1'} g_i(f_1(i)) + \sum_{i \in I_2'} g_i(f_2(i)) +$$

$$\sum_{i \in J^*} [g_i(f_1(i)) + g_i(f_2(i)) \qquad (10.1)$$

For $i \in J'$ we have $f_1(i) + f_2(i) = 0$ and therefore

$$\sum_{i \in J'} g_i(f_1(i) + f_2(i)) = \mathbf{0}_W \tag{10.2}$$

Adding (10.2) to (10.1) we obtain

$$\sum_{i \in I_1'} g_i(f_1(i)) + \sum_{i \in I_2'} g_i(f_2(i)) +$$

$$\sum_{i \in J^*} [g_i(f_1(i)) + g_i(f_2(i)) + \sum_{i \in J'} g_i(f_1(i) + f_2(i)) =$$

$$\sum_{i \in I_1'} g_i(f_1(i)) + \sum_{i \in I_2'} g_i(f_2(i)) +$$

$$\sum_{i \in J^*} [g_i(f_1(i)) + g_i(f_2(i)) +$$

$$\sum_{i \in J'} g_i(f_1(i)) + \sum_{i \in J'} g_i(f_2(i)) \tag{10.3}$$

Rearranging and combining the terms in (10.3) we get

$$\sum_{i \in I_1' \cup J^* \cup J'} g_i(f_1(i)) + \sum_{i \in I_2' \cup J^* \cup J'} g_i(f_2(i)) \tag{10.4}$$

Since $spt(f_1) = I_1' \cup J^* \cup J'$ and $spt(f_2) = I_2' \cup J^* \cup J'$ we get (10.4) is equal to

$$\sum_{i \in spt(f_1)} g_i(f_1(i)) + \sum_{i \in spt(f_2)} g_i(f_2(i)) = G(f_1) + G(f_2)$$

Now suppose $f \in \oplus_{i \in I} V_i$ and $0 \neq c \in \mathbb{F}$ is a scalar. Then $spt(cf) = spt(f)$. Now

$$G(cf) = \sum_{i \in spt(cf)} g_i([cf](i)) = \sum_{i \in spt(f)} g_i([cf](i)) =$$

$$\sum_{i \in spt(f)} g_i(cf(i)) = \sum_{i \in spt(f)} c g_i(f(i)) =$$

$$c \sum_{i \in spt(f)} g_i(f(i)) = cG(f).$$

2. Since $S : V \to W$ is surjective, each $S \otimes \cdots \otimes S : T_k(V) \to T_k(W)$ is surjective by part i) of Lemma (10.2). It then follows by Lemma (10.3) that $\mathcal{T}(S)$ is surjective.

3. Since $S : V \to W$ is injective, each $S \otimes \cdots \otimes S : T_k(V) \to T_k(W)$ is injective by part ii) of Lemma (10.2). It then follows by Lemma (10.3) that $\mathcal{T}(S)$ is injective.

4. This follows immediately from part iv. of Lemma (10.2).

5. The eigenvalues are: 8, 27, 125 (with multiplicity 1) 12, 20, 18, 50, 45, 75 (with multiplicity 3) and 30 (with multiplicity 6).

6. Let $S(\boldsymbol{v}_1) = 2\boldsymbol{v}_1, S(\boldsymbol{v}_2) = 3\boldsymbol{v}_2$. Then $\boldsymbol{v}_1 \otimes \boldsymbol{v}_2, \boldsymbol{v}_2 \otimes \boldsymbol{v}_1$ are both eigenvectors of $\mathcal{T}_2(S)$ with eigenvalue 6. Thus, $\mathcal{T}_2(S)$ is not cyclic.

7. This is false unless $RS = SR = 0_{V \to V}$. Even taking $R = S = I_V$ gives a counterexample. In that case $R + S = 2I_V$ and $\mathcal{T}(R) + \mathcal{T}(S) = 2I_{\mathcal{T}(V)}$ and every vector in $\mathcal{T}(V)$ is an eigenvector with eigenvalue 2 for $2I_{\mathcal{T}(V)}$. On the other hand, vectors in $\mathcal{T}_2(V)$ are eigenvectors of $\mathcal{T}_2(2I_V)$ with eigenvalue 4.

8. This is false. For example, let V have dimension 3 with basis $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$. Set $X = Span(\boldsymbol{v}_1, \boldsymbol{v}_2)$ and $Y = Span(\boldsymbol{x}_3)$ and $S = Proj_{(X,Y)}$. Then $Ker(S) = Y$. Now $K_2 = Ker(\mathcal{T}_2(S)) = V \otimes Y + Y \otimes V$ has dimension 6 and therefore $\mathcal{T}_2(V)/K_2$ has dimension 3. On the other hand, $dim(V/Y) = 2$ and $\mathcal{T}_2(V/Y) = 4$.

9. This is an immediate consequence of the definition of the tensor product.

10. Assume $S^l = 0_{V \to V}$. We claim that $\mathcal{T}_k(S)^{kl-k+1} = 0_{\mathcal{T}_k(V) \to \mathcal{T}_k(V)}$. To see this note that $\mathcal{T}_k(S) = (S \otimes I_V \otimes \cdots \otimes I_V)(I_V \otimes S \otimes \cdots \otimes I_V) \ldots (I_V \otimes I_V \otimes \cdots \otimes S)$ where in each case there is tensor product of k maps, one is S and all the others are I_V. These maps commute and therefore $\mathcal{T}_k(S)^{kl-k+1} =$

$$\sum S^{j_1} \otimes \cdots \otimes S^{j_k} \tag{10.5}$$

where the sum is over all non-negative sequences $(j_1, \ldots, j_k)$ such that $j_1 + \cdots + j_k = kl - k + 1$. By the generalized pigeon-hole principle some $j_i \geq l$ which implies that $S^{j_i} = 0_{V\to V}$ and hence each term in (10.5) is zero.

11. The formula is $Tr(\mathcal{T}_k(S)) = Tr(S)^k$. The proof is by induction on k. By Exercise 13 of Section (10.2) $Tr(S \otimes S) = Tr(S)^2$. Assume the result is true for k: $Tr(\mathcal{T}_k(S)) = Tr(S)^k$. Again by Exercise 13 of Section (10.2) $Tr(\mathcal{T}_{k+1}(S)) = Tr(\mathcal{T}_k(S) \otimes S) = Tr(\mathcal{T}_k(S))Tr(S) = Tr(S)^k Tr(S) = Tr(S)^{k+1}$.

12. We claim that $det(\mathcal{T}_k(S)) = det(S)^{kn^{k-1}}$. We do induction on $k \geq 1$. When $k = 1$ the result clearly holds. Also, for $k = 2$ we can apply Exercise 15 of Section (10.2) to obtain $det(S \otimes S) = det(S)^n det(S)^n = set(S)^{2n}$ as required.

Assume now that $det(\mathcal{T}_k(S)) = det(S)^{kn^{k-1}}$. Now $\mathcal{T}_{k+1}(S) = \mathcal{T}_k(S) \otimes S$. The dimension of $\mathcal{T}_k(V) = n^k$. Then by Exercise 15 we have

$$det(\mathcal{T}_{k+1}(S)) = det(\mathcal{T}_k(S) \otimes S) =$$
$$det(\mathcal{T}_k(S))^n det(S)^{n^k} =$$
$$(det(S)^{kn^{k-1}})^n det(S)^{n^k} =$$
$$det(S)^{kn^k} det(S)^{n^k} = det(S)^{(k+1)n^k}.$$

10.4. The Symmetric Algebra

1. $\mathcal{S} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a sequence of vectors. We will denote by $\otimes(\mathcal{S})$ the product $\boldsymbol{v}_1 \otimes \cdots \otimes \boldsymbol{v}_n$. Also, for a permuation π of $\{1, \ldots, n\}$ we let $\pi(\mathcal{S}) = (\boldsymbol{v}_{\pi(1)}, \ldots, \boldsymbol{v}_{\pi(n)})$. We therefore have to prove that $\otimes(\mathcal{S}) - \otimes(\pi(\mathcal{S})) \in \mathcal{I}$ for all sequences $\mathcal{S}$ and permuations π. We first prove that the result for permuations (i, j) which exchange a pair $i < j$ and leave all other k in $\{1, \ldots, n\}$ fixed. We prove this by induction on $j - i$.

For the base case, $j = i + 1$ we have $\boldsymbol{v}_i \otimes \boldsymbol{v}_{i+1} - \boldsymbol{v}_{i+1} \otimes \boldsymbol{v}_i \in \mathcal{I}$ and then for every $\boldsymbol{x} \in \mathcal{T}_{i-1}(V), \boldsymbol{y} \in \mathcal{T}_{n-i-1}(V), \boldsymbol{x} \otimes [\boldsymbol{v}_i \otimes \boldsymbol{v}_{i+1} - \boldsymbol{v}_{i+1} \otimes \boldsymbol{v}_i] \otimes \boldsymbol{y} \in \mathcal{I}$. Therefore $\otimes(\mathcal{S}) - \otimes((i, i+1)\mathcal{S}) \in \mathcal{I}$.

Now assume that $i < k$ and for all j with $i < j < k$ $\otimes(\mathcal{S}) - \otimes((i, j)(\mathcal{S})) \in \mathcal{I}$. It suffices to prove that

$$\boldsymbol{v}_i \otimes \boldsymbol{v}_{i+1} \otimes \ldots \boldsymbol{v}_{k-1} \otimes \boldsymbol{v}_k -$$
$$\boldsymbol{v}_k \otimes \boldsymbol{v}_{i+1} \otimes \cdots \otimes \boldsymbol{v}_{k-1} \otimes \boldsymbol{v}_i \in \mathcal{I}.$$

By the induction hypothesis each of the following are in $\mathcal{I}$:

$$\boldsymbol{v}_i \otimes \boldsymbol{v}_{i+1} \otimes \cdots \otimes \boldsymbol{v}_{k-1} \otimes \boldsymbol{v}_k -$$
$$\boldsymbol{v}_{i+1} \otimes \boldsymbol{v}_i \otimes \boldsymbol{v}_{i+2} \otimes \cdots \otimes \boldsymbol{v}_k$$

$$\boldsymbol{v}_{i+1} \otimes \boldsymbol{v}_i \otimes \boldsymbol{v}_{i+2} \otimes \cdots \otimes \boldsymbol{v}_k -$$
$$\boldsymbol{v}_{i+1} \otimes \boldsymbol{v}_k \otimes \boldsymbol{v}_{i+2} \otimes \ldots \boldsymbol{v}_{k-1} \otimes \boldsymbol{v}_i$$

$$\boldsymbol{v}_{i+1} \otimes \boldsymbol{v}_k \otimes \boldsymbol{v}_{i+2} \otimes \cdots \otimes \boldsymbol{v}_{k-1} \otimes \boldsymbol{v}_i -$$
$$\boldsymbol{v}_k \otimes \boldsymbol{v}_{i+1} \otimes \boldsymbol{v}_{i+2} \otimes \cdots \otimes \boldsymbol{v}_{k-1} \otimes \boldsymbol{v}_i.$$

The sum of these is

$$\boldsymbol{v}_i \otimes \boldsymbol{v}_{i+1} \otimes \ldots \boldsymbol{v}_{k-1} \otimes \boldsymbol{v}_k -$$
$$\boldsymbol{v}_k \otimes \boldsymbol{v}_{i+1} \otimes \cdots \otimes \boldsymbol{v}_{k-1} \otimes \boldsymbol{v}_i$$

which is therefore in $\mathcal{I}$.

We now use the fact that every permutation can be written as a product of transpositions - permutations of type (i, j). We do induction on the number of transpositions needed to express the permutation π. We have already proved the base case. Suppose we have shown if π that can be expressed as a product of k transpositions then $\otimes(\mathcal{S}) - \otimes(\pi(\mathcal{S})) \in \mathcal{I}$ and assume that $\tau = \sigma_{k+1} \ldots \sigma_1$ a product of transpostions. Set $\gamma = \sigma_k \ldots \sigma_1$. Then $\otimes(\mathcal{S}) - \otimes(\tau(\mathcal{S}))$

$$[\otimes(\mathcal{S}) - \otimes(\gamma(\mathcal{S}))] + [\otimes(\gamma(\mathcal{S})) - \otimes(\sigma_{k+1}(\gamma(\mathcal{S})))].$$

$[\otimes(\mathcal{S}) - \otimes(\gamma(\mathcal{S}))]$ is in $\mathcal{I}$ by the inductive hypothesis and $[\otimes(\gamma(\mathcal{S})) - \otimes(\sigma_{k+1}(\gamma(\mathcal{S})))] \in \mathcal{I}$ by the base case. Thus, $\otimes(\mathcal{S}) - \otimes(\tau(\mathcal{S})) \in \mathcal{I}$ as required.

2. We use the identification of $Sym(V)$ with $\mathbb{F}[x_1, \ldots, x_n]$. Under this identification $Sym_k(V)$ corresponds to the subspace of homogeneous polynomial of degree k. A basis for this consists of all $x_1^{e_1} \ldots x_n^{e_n}$ where $e_1, \ldots, e_n$ are non-negative and $e_1 + \cdots + e_n = k$. This is equal to the number ways that n bins can be filled with k balls (where some bins can be empty). This is a standard combinatorics problem. The answer is $\binom{k+n-1}{k} = \binom{k+n-1}{n-1}$.

3. Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis of V consisting of eigenvectors of T with $T(\boldsymbol{v}_i) = \alpha_i$. Under the correspondence of $Sym_k(V)$ with the homogeneous polynomials of $\mathbb{F}[\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n]$ of degree k the monomial $\boldsymbol{v}_1^{e_1} \ldots \boldsymbol{v}_n^{e_n}$ where $(e_1, \ldots, e_n)$ is a sequence of non-negative integers with $e_1 + \cdots + e_n = k$ is an eigenvector with eigenvalue $\alpha_1^{e_1} \ldots \alpha_n^{e_n}$.

4. The eigenvalues are 1, 2, 8, 16 with multiplicity 1 and 4 with multiplicity 2. This operator is not cyclic.

5. $a_3^2 - a_2$.

10.5. Exterior Algebra

1. Minic the proof of Exercise 1 of Section (10.4).

2. We let $[1, n] = \{1, \ldots, n\}$ and $S_n = \{\pi : [1, n] \to [1, n] | \pi \text{ is bijective}\}$ and $S_n^* = S_n \setminus \{I_{[1,n]}$. Set $W_0 = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k)$. Denote by $S_n(W_0)$ the set $\{(\boldsymbol{v}_{\pi(1)}, \ldots, \boldsymbol{v}_{\pi(k)}) | \pi \in S_n\}$ and set $[\mathcal{B}^k]' = \mathcal{B}^k \setminus S_n(W_0)$. Also, for $W = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k) \in \mathcal{B}^k$ set $\otimes(W) = \boldsymbol{w}_1 \otimes \cdots \otimes \boldsymbol{w}_k$. Finally, for $W \in S_n(W_0)$ and $\pi \in S_n$ let $(W, \pi) = \otimes(W) - sgn(\pi) \otimes (\pi(W))$.

A typical spanning vector in$\mathcal{J}_k$ has the form $\boldsymbol{x}_1 \otimes \cdots \otimes \boldsymbol{x}_i \otimes \boldsymbol{y} \otimes \boldsymbol{y} \otimes \boldsymbol{z}_1 \otimes \cdots \otimes \boldsymbol{z}_j$ where $i+j+2 = k$. Express each of $\boldsymbol{x}_r, \boldsymbol{y}, \boldsymbol{z}_s$ as a linear combination of the basis $\mathcal{B}$. Then $\boldsymbol{x}_1 \otimes \cdots \otimes \boldsymbol{x}_i \otimes \boldsymbol{y} \otimes \boldsymbol{y} \otimes \boldsymbol{z}_1 \otimes \cdots \otimes \boldsymbol{z}_j$ is a linear combination of $\otimes(W), W \in [\mathcal{B}^k]'$ and $(W, \tau), W \in S_n(W_0), \tau \in S_n$.

3. We continue with the notation introduced in the solution of Exercise 2 with $k = n$ so that $W_0 = \mathcal{B} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. Our first claim is that

$$Span((W, \pi) | W \in S_n(\mathcal{B}), \pi \in S_n^*) =$$
$$Span((\mathcal{B}, \pi) | \pi \in S_n^*).$$

Since $(\mathcal{B}, \pi) \in Span((W, \pi) | W \in S_n(\mathcal{B}), \pi \in S_n^*)$, we need only prove the inclusion

$$Span((W, \pi) | W \in S_n(\mathcal{B}), \pi \in S_n^*) \subset Span((\mathcal{B}, \pi) | \pi \in S_n^*)$$

, equivalently,

$$(W, \pi) \in Span((\mathcal{B}, \pi) | \pi \in S_n^*).$$

Suppose $W = \sigma(\mathcal{B})$. Then $(\mathcal{B}, \sigma)$ and $(\mathcal{B}, \pi\sigma) \in Span(\mathcal{B}, \pi) | \pi \in S_n^*)$. Taking the difference we obtain

$$(\mathcal{B}, \sigma) - (\mathcal{B}, \pi\sigma) =$$
$$[\otimes(\mathcal{B}) - sgn(\sigma)(\otimes(\sigma(\mathcal{B}))] -$$
$$[\otimes(\mathcal{B}) - sgn(\pi\sigma)(\otimes(\pi\sigma(\mathcal{B}))] =$$

$$sgn(\pi\sigma)(\otimes(\pi(W)) - sgn(\sigma)(\otimes(W)) =$$

$$sgn(\sigma)sgn(\pi)(\otimes(\pi(W)) - sgn(\sigma)(\otimes(W)) =$$

$$-sgn(\sigma)(W, \pi)$$

which establishes the claim.

Next note that $Span(\otimes(W) | W \in [\mathcal{B}^n]') \cap Span(\otimes(W) | W \in S_n(\mathcal{B})) = \{\mathbf{0}\}$. Therefore, if $\otimes(\mathcal{B}) \in \mathcal{J}_n$ then, in fact, $\otimes(\mathcal{B})$ is a linear combination of $(\mathcal{B}, \pi), \pi \in S_n^*$.

Suppose $\otimes(\mathcal{B}) \in Span((\mathcal{B}, \pi) | \pi \in S_n^*)$. Let $\sigma \in S_n^*$. Then $-(\mathcal{B}, \sigma) + \otimes(\mathcal{B}) \in Span((\mathcal{B}, \pi) | \pi \in S_n^*)$. However, $-(\mathcal{B}, \sigma) + \otimes(\mathcal{B})$ is $\pm \otimes (\sigma(\mathcal{B}))$ and consequently, $\otimes(\sigma(\mathcal{B})) \in Span((\mathcal{B}, \pi) | \pi \in S_n^*)$. Since the set $\{\otimes(\sigma(\mathcal{B})) | \sigma \in S_n\}$ is linear independent this will imply that $Span((\mathcal{B}, \pi) | \pi \in S_n^*)$ has dimension $n!$ On the other hand, there are only $n! - 1$ generating vectors so $dim(Span(\mathcal{B},\pi)|\pi \in S_n^*) \leq n!$ Thus, $\otimes(\mathcal{B}) \notin Span((\mathcal{B},\pi)|\pi \in S_n^*)$.

This implies that $\wedge^n(V)$ has dimension at least one. On the other hand it has been established that $dim(\wedge^n(V))$ has dimension at most one and therefore exactly one. Since $\wedge^n(V)$ has dimension one with basis $\boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_n$ there is a unique map $F : \wedge^n(V) \to \mathbb{F}$ such that $F(\boldsymbol{v}_1\wedge\cdots\wedge\boldsymbol{v}_n) = 1$. By the universal property of $\wedge^n(V)$ this implies that there is a unique alternating $n-$linear form f on V such that $f(\mathcal{B}) = 1$.

4. Let $(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_k)$ be a sequence of vectors from V. Then

$$\wedge^k(SR)(\boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_k) =$$
$$(SR)(\boldsymbol{w}_1) \wedge \cdots \wedge (SR)(\boldsymbol{w}_k) =$$
$$S(R(\boldsymbol{w}_1)) \wedge \cdots \wedge S(R(\boldsymbol{w}_k)) =$$
$$\wedge^k(S)(R(\boldsymbol{w}_1) \wedge \cdots \wedge R(\boldsymbol{w}_k)) =$$
$$\wedge^k(S)(\wedge^k(R)(\boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_k)) =$$
$$(\wedge^k(S) \wedge^k (R))(\boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_k).$$

5. Let $\mathcal{B}_V = (\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)$ be a basis of V. For a sequence $(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_k)$ we let $\wedge(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_k) = \boldsymbol{u}_1 \wedge \cdots \wedge \boldsymbol{u}_k$. We continue with the notation of Exercises 2 and 3. To show that S is injective it suffices to show that the image of a basis of $\wedge^k(V)$ is linearly independent in $\wedge^k(W)$. Now $\{\wedge^k(\Phi)|\Phi \in S_n(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_k)\}$ is a basis for $\wedge^k(V)$. Since S is injective, $(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_n) = (S(\boldsymbol{v}_1),\ldots,S(\boldsymbol{v}_n))$ is linearly independent in W. Extend to a basis $(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_m)$ for W. Then $\{\wedge^k(\Psi)|\Psi \in S_m(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_k)\}$ is a basis for $\wedge^k(W)$. In particular, $\{\wedge^k(\Psi)|\Psi \in S_n(S(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_k))\} = \{\wedge^k(\Psi)|\Psi \in S_n((S(\boldsymbol{v}_1),\ldots,S(\boldsymbol{v}_k))\}$ is linearly independent. However,

$$\wedge^k((S(\boldsymbol{v}_{\pi(1)}),\ldots,\boldsymbol{v}_{\pi(k)}) = S(\boldsymbol{v}_{\pi(1)}) \wedge\cdots\wedge S(\boldsymbol{v}_{\pi(k)}) =$$

$$\wedge^k(S)(\boldsymbol{v}_{\pi(1)} \wedge \cdots \wedge \boldsymbol{v}_{\pi(k)})$$

Thus, $\{\wedge^k(S)(\wedge^k(\Phi))|\Phi \in S_n(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_k)\}$ is linearly independent and $\wedge^k(S)$ is injective.

6. If V is $n-$dimensional and S is nilpotent then $S^n = 0_{V\to V}$. Since $\wedge(S)^n = \wedge(S^n) = 0_{\wedge(V)\to\wedge(V)}$. Thus, $\wedge(S)$ is nilpotent.

7. Let $(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n)$ be a basis of V such that $S(\boldsymbol{v}_i) = \alpha_i\boldsymbol{v}_i$. Then $\boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_k}$ is an eigenvector of $\wedge^k(S)$ with eigenvalue $\alpha_{i_1}\ldots\alpha_{i_k}$. Since $\{\boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_k}|1 \leq i_1 < \cdots < i_k \leq n\}$ is a basis for $\wedge^k(V)$ it follows that $\wedge^k(S)$ is diagonalizable.

8. $det(\wedge^k(S)) = det(S)^{\binom{n-1}{k-1}}$. This can be proved by showing it holds for elementary operators (matrices).

9. Let $S : \mathbb{R}^4 \to \mathbb{R}^4$ be the operator with matrix $\begin{pmatrix} 0 & 1 & 0 & 0\\ -1 & 0 & 0 & 0\\ 0 & 0 & 3 & 4\\ 0 & 0 & -3 & 4\end{pmatrix}$.

Then the eigenvalues of S are $\pm i, 3 \pm 4i$. On the other hand, the eigenvalues of $\wedge^2(S)$ are 1, 25, $-4+3i, 4+3i, -4-3i, 4-3i$.

10. Let $(\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3,\boldsymbol{v}_4)$ be linearly independent and set $\boldsymbol{x} = \boldsymbol{v}_1\wedge\boldsymbol{v}_2+\boldsymbol{v}_3\wedge\boldsymbol{v}_4$. Then $\boldsymbol{x}\wedge\boldsymbol{x} = 2(\boldsymbol{v}_1\wedge\boldsymbol{v}_2\wedge\boldsymbol{v}_3\wedge\boldsymbol{v}_4)$.

11. Since multiplication is distributive ϕ is additive in each variable. Since $(c\boldsymbol{w}) \wedge \boldsymbol{v} = \boldsymbol{w} \wedge (c\boldsymbol{v}_= c(\boldsymbol{w} \wedge \boldsymbol{v})$, in fact, ϕ is bilinear. In light of bilinearity, to show that ϕ is symmetric it suffices to show that it is for decomposable vectors, that is, vectors of the form $\boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_{2n}$. However,

$$(\boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_{2n}) \wedge (\boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_{2n}) =$$
$$(-1)^{4n^2}(\boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_{2n}) \wedge (\boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_{2n})$$

from which it follows that

$$\phi(vv_1 \wedge \cdots \wedge \boldsymbol{v}_{2n}, \boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_{2n}) =$$
$$\phi(\boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_{2n}, \boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_{2n}).$$

We now need to show that ϕ is non-degenerate. For a subset $\alpha = \{i_1 < \cdots < i_{2n}\}$ of $[1,4n]$ let $\boldsymbol{v}_\alpha = \boldsymbol{v}_{i_1} \wedge$

$\cdots \wedge \boldsymbol{v}_{i_{2n}}$. Also denote by α' the complementary subset $\alpha' = [1, 4n] \setminus \alpha$. Note that $\boldsymbol{v}_\alpha \wedge \boldsymbol{v}_\beta = \boldsymbol{0}_{\wedge^{2n}(V)}$ if $\beta \neq \alpha'$ and $\boldsymbol{v}_\alpha \wedge \boldsymbol{v}_{\alpha'} = \pm \boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_{4n}$.

Now suppose that $\boldsymbol{x} = \sum_\alpha c_\alpha \boldsymbol{v}_\alpha$ where the sum is taken over all subsets α of $[1, 4n]$ of size $2n$ and that $\boldsymbol{x} \neq \boldsymbol{0}_{\wedge^{2n}(V)}$. Then some $c_\alpha \neq 0$. Then $\boldsymbol{x} \wedge \boldsymbol{v}_{\alpha'} = \pm c_\alpha$ and $\phi(\boldsymbol{x}, \boldsymbol{v}_{\alpha'}) \neq 0$. In particular, $\boldsymbol{x}$ is not in the radical. Since $\boldsymbol{x}$ is arbitrary the radical of ϕ is just the zero vector.

12. We need to show that there are totally singular subspaces of dimension 3. Note that the singular vectors are precisely the decomposable vectors. Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_4)$ be a basis for V. Then $Span(\boldsymbol{v}_1 \wedge \boldsymbol{v}_2, \boldsymbol{v}_1 \wedge \boldsymbol{v}_3, \boldsymbol{v}_1 \wedge \boldsymbol{v}_4)$ is a totally singular subspace of dimension 3. A second class of singular subspaces of dimension 3 is represented by $]'pan(\boldsymbol{v}_1 \wedge \boldsymbol{v}_2, \boldsymbol{v}_1 \wedge \boldsymbol{v}_3, \boldsymbol{v}_2 \wedge \boldsymbol{v}_3)$.

13. As in the proof of Exercise 11, the form is bilinear. Now we have

$$(\boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{v}_n) \wedge (\boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_n) =$$
$$(-1)^{n^2} (\boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_n) \wedge (\boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{w}_n) =$$
$$-(\boldsymbol{w}_1 \wedge \cdots \wedge \boldsymbol{w}_n) \wedge (\boldsymbol{v}_1 \wedge \cdots \wedge \boldsymbol{w}_n$$

Suppose now that $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{2n})$ is a basis of V. For a subset $\alpha = \{i_1 < \cdots < i_n\}$ of $[1, 2n]$ set $\boldsymbol{v}_\alpha = \boldsymbol{v}_{i_1} \wedge \cdots \wedge \boldsymbol{v}_{i_n}$. As in Exercise 18, if $\beta \neq \alpha' = [1, 2n] \setminus \alpha$ then $\boldsymbol{v}_\alpha \wedge \boldsymbol{v}_\beta = \boldsymbol{0}_{\wedge^{2n}(V)}$.

Let $\boldsymbol{x} = c_\alpha \boldsymbol{v}_\alpha + c_{\alpha'} \boldsymbol{v}_{\alpha'}$. Then $\boldsymbol{x}^2 = \boldsymbol{x} \wedge \boldsymbol{x} =$

$$c_\alpha c_{\alpha'} \boldsymbol{v}_\alpha \wedge \boldsymbol{v}_{\alpha'} + c_{\alpha'} c_\alpha \boldsymbol{v}_{\alpha'} \wedge \boldsymbol{v}_\alpha =$$

$$c_\alpha c_{\alpha'} \boldsymbol{v}_\alpha \wedge \boldsymbol{v}_{\alpha'} - c_{\alpha'} c_\alpha \boldsymbol{v}_\alpha \wedge \boldsymbol{v}_{\alpha'} = \boldsymbol{0}_{\wedge^{2n}(V)}.$$

It follows from the previous two paragraphs that $\boldsymbol{x} \wedge \boldsymbol{x} = \boldsymbol{0}_{\wedge^{2n}(V)}$ for any vector $\boldsymbol{x} \in \wedge^n(V)$ and therefore ϕ is alternating. The proof that ϕ is non-degenerate is exactly as in Exercise 11.

14. $x^6 + 14x^4 + 96x^3 - 128x - 32$.

15. $x^3 + 6x^2 - 9$.

16. $x^6 - 3x^4 - 27x^3 - 9x^2 + 27$.

10.6. Clifford Algebra

1. Let $\boldsymbol{v}$ be a non-zero vector in V. Set $a = \phi(\boldsymbol{v}) < 0$. Let $c = \sqrt{-a}$. By replacing $\boldsymbol{v}$, if necessary, by $\frac{1}{c}\boldsymbol{v}$ we can assume that $\phi(\boldsymbol{v}) = -1$. Since $dim(V) = 1, V = Span(\boldsymbol{v})$. Then $\mathcal{T}(V)$ is isomorphic to $\mathbb{R}[x]$. The ideal $\mathcal{I}_\phi$ is generated by $\boldsymbol{v} \otimes \boldsymbol{v} + 1$ which corresponds to the polynomial $x^2 + 1$. Thus, $\mathcal{T}(V)/\mathcal{I}_\phi$ is isomorphic to $\mathbb{C}$.

2. Let $\alpha = \{i_1 < \cdots < i_k\}$ and assume $j \notin \alpha$. Then $\boldsymbol{v}_j \boldsymbol{v}_{i_s} = -\boldsymbol{v}_{i_s} \boldsymbol{v}_j$ for $1 \leq s \leq k$. It follows by induction on k that $\boldsymbol{v}_j \boldsymbol{v}_\alpha = (-1)^k \boldsymbol{v}_\alpha \boldsymbol{v}_j$.

3. Continue with the notation of 2 and assume $i_s = j$. Then $\boldsymbol{v}_j \boldsymbol{v}_\alpha = \boldsymbol{v}_j \boldsymbol{v}_{i_1} \ldots \boldsymbol{v}_{i_s} \ldots \boldsymbol{v}_{i_k} =$

$$(-1)^{s-1} \boldsymbol{v}_{i_1} \ldots \boldsymbol{v}_{i_{s-1}} \boldsymbol{v}_j \boldsymbol{v}_{i_s} \ldots \boldsymbol{v}_{i_k} =$$
$$(-1)^{s-1} \boldsymbol{v}_{i_1} \ldots \boldsymbol{v}_{i_{s-1}} \boldsymbol{v}_j \boldsymbol{v}_j \ldots \boldsymbol{v}_{i_k} =$$
$$(-1)^{s-1} \boldsymbol{v}_{i_1} \ldots \boldsymbol{v}_{i_{s-1}} \phi(\boldsymbol{v}_j) \boldsymbol{v}_{s+1} \ldots \boldsymbol{v}_{i_k} =$$
$$(-1)^{s-1} \phi(\boldsymbol{v}_j) \boldsymbol{v}_{\alpha \setminus \{j\}}.$$

4. Assume first that $\mathcal{I}$ is a homogeneous ideal of $A = A^0 \oplus A^1$. Then $\mathcal{I} = \mathcal{I} \cap A^0 \oplus \mathcal{I} \cap A^1$ from which it immediately follows that $\mathcal{I}$ is generated as an ideal by $(\mathcal{I} \cap A^0) \cup (\mathcal{I} \cap A^1)$.

Conversely, assume X is a set of homogeneous elements of A and $\mathcal{I}$ consists of all elements of the form $z = b_1 x_1 c_1 + \cdots + b_k x_k c_k + d_1 y_1 e_1 + \ldots d_l y_l e_l$ where $x_i \in X \cap A^0, y_i \in X \cap A^1, b_i, c_i, d_i, e_i \in A$. We need to show the homogeneous parts of zR are in $\mathcal{I}$. Write each b_i as $b_{i0} + b_{i1}$ where $b_{it} \in A^t$ and similarly for c_i, d_i and e_i. Set

$$z_0 = \sum_{i=1}^{k} (b_{i0} x_i c_{i0} + b_{i1} x_i c_{i1}) + \sum_{i=1}^{l} (d_{i0} y_i e_{i1} + d_{i1} y_i e_{i0})$$

$$z_1 = \sum_{i=1}^{k}(b_{i0}x_ic_{i1} + b_{i1}x_ic_{i0}) + \sum_{i=1}^{l}(d_{i0}y_ie_{i0} + d_{i1}y_ie_{i1})$$

Then $z_0 \in A^0, z_1 \in A^1$ and $z = z_0 + z_1$. Now each of $b_{is}x_ic_{it}$ and $d_{is}y_ie_{it}$ belongs to $\mathcal{I}$ and therefore $z_0, z_1 \in \mathcal{I}$ and $\mathcal{I}$ is a homogeneous ideal.

5. Let $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ be an orthogonal basis for V. As in the solution to Exercise 1 we can assume that $\phi(\boldsymbol{v}_1) = \phi(\boldsymbol{v}_2) = -1$. Set $i = \boldsymbol{v}_1, j = \boldsymbol{v}_2$ and $k = \boldsymbol{v}_1\boldsymbol{v}_2$ then $(1, i, j, k)$ is a basis for $C(V)$. Note that $i^2 = \phi(\boldsymbol{v}_1) = -1 = \phi(\boldsymbol{v}_2) = j^2$. Since $\boldsymbol{v}_1 \perp \boldsymbol{v}_2, ij = \boldsymbol{v}_1\boldsymbol{v}_2 = -\boldsymbol{v}_2\boldsymbol{v}_1 = -ji$. By Exercise 3, $ki = (ij)i = i(-ij) = -i^2j = j$. Simiarly, $jk = i = -kj$. It follows that $C(V)$ is isomorphic to the division ring of quaterions.

6. Let $(\boldsymbol{x}, \boldsymbol{y})$ be a hyperbolic basis for V. Then $\phi(\boldsymbol{x} + \boldsymbol{y}) = \phi(\boldsymbol{x}) + \phi(\boldsymbol{y}) + \langle \boldsymbol{x}, \boldsymbol{y}\rangle_\phi = 0 + 0 + 1$. Therefore $\boldsymbol{x}^2 + \boldsymbol{y}^2 + \boldsymbol{xy} + \boldsymbol{yx} = 1$. Since $(1, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{xy})$ is a basis for $C(V)$ and $Span(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{xy}, \boldsymbol{yx})$ contains $\{1, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{xy}\}$ also $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{xy}, \boldsymbol{yx})$ is a basis. Denote by $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix}$ the vector $a_{11}\boldsymbol{xy} + a_{12}\boldsymbol{x} + a_{21}\boldsymbol{y} + a_{22}\boldsymbol{yx}$. Let's determine the product of $a_{11}\boldsymbol{xy} + a_{12}\boldsymbol{y} + a_{21}\boldsymbol{x} + a_{22}\boldsymbol{yx}$ with $b_{11}\boldsymbol{xy} + b_{12}\boldsymbol{x} + b_{21}\boldsymbol{y} + b_{22}\boldsymbol{yx}$. We note the following:

$$\boldsymbol{x}^2 = \boldsymbol{y}^2 = \boldsymbol{x}(\boldsymbol{xy}) = \boldsymbol{y}(\boldsymbol{yx}) = (\boldsymbol{yx})\boldsymbol{x} = (\boldsymbol{xy})\boldsymbol{y} = 0$$
$$\boldsymbol{x}(\boldsymbol{yx}) = \boldsymbol{x} = (\boldsymbol{xy})\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{xy}) = \boldsymbol{y} = (\boldsymbol{yx})\boldsymbol{y}$$
$$(\boldsymbol{xy})^2 = \boldsymbol{xy}, (\boldsymbol{yx})^2 = \boldsymbol{yx}.$$

Now the product of $a_{11}\boldsymbol{xy} + a_{12}\boldsymbol{x} + a_{21}\boldsymbol{y} + a_{22}\boldsymbol{yx}$ and $b_{11}\boldsymbol{xy} + b_{12}\boldsymbol{x} + b_{21}\boldsymbol{y} + b_{22}\boldsymbol{yx}$ is

$$(a_{11}b_{11} + a_{12}b_{21})\boldsymbol{xy} +$$
$$(a_{11}b_{12} + a_{12}b_{22})\boldsymbol{x} +$$
$$(a_{21}b_{11} + a_{22}b_{21}\boldsymbol{y} +$$
$$(a_{21}b_{12} + a_{22}b_{22})\boldsymbol{yx}.$$

It follows that the map from $a_{11}\boldsymbol{xy} + a_{12}\boldsymbol{x} + a_{21}\boldsymbol{y} + a_{22}\boldsymbol{yx}$ to $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix}$ is an isomorphism of algebras and $C(V)$ is isomorphic to $M_{22}(\mathbb{F})$.

Chapter 11

Linear Groups and Groups of Isometries

11.1. Linear Groups

1. The order of $GL(V)$ is equal to the number of bases of V. When $dim(V) = n$ and the field is $\mathbb{F}_q$ this is

$$q^{\binom{n}{2}}\prod_{i=1}^{n}(q^i-1).$$

Since $SL(V)$ is the kernel of $det : GL(V) \to \mathbb{F}_q^*$ and the determinant is surjective, $|SL(V)| = |GL(V)|/|\mathbb{F}_q^*| =$

$$q^{\binom{n}{2}}\prod_{i=1}^{n}(q^i-1) \times \frac{1}{q-1} = q^{\binom{n}{2}}\prod_{i=2}^{n}(q^i-1).$$

2. Let $j = dim(U_1 \cap U_2)$. Let $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_j)$ be a bass for $U_1 \cap U_2$ and $(\boldsymbol{y}_1, \dots, \boldsymbol{y}_j)$ a basis for $W_1 \cap W_2$. Set $t = k - j$. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_t)$ be vectors in U_1 such that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_j, \boldsymbol{u}_1, \dots, \boldsymbol{u}_t$ is a basis for U_1 and $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_t)$ vectors from U_2 such that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_j, \boldsymbol{v}_1, \dots, \boldsymbol{v}_t)$ is a basis for U_2. Let $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_t)$ be vectors from W_1 such that $(\boldsymbol{y}_1, \dots, \boldsymbol{y}_j, \boldsymbol{w}_1, \dots, \boldsymbol{w}_t)$ is a basis of W_1 and $(\boldsymbol{z}_1, \dots, \boldsymbol{z}_t)$ be vectors from W_2 such that $(\boldsymbol{y}_1, \dots, \boldsymbol{y}_j, \boldsymbol{z}_1, \dots, \boldsymbol{z}_t)$ is a basis for W_2. Now $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_j, \boldsymbol{u}_1, \dots, \boldsymbol{u}_t, \boldsymbol{v}_1, \dots, \boldsymbol{v}_t)$ is a basis for $U_1 + U_2$ and $(\boldsymbol{y}_1, \dots, y_j, \boldsymbol{w}_1, \dots, \boldsymbol{w}_t, \boldsymbol{z}_1, \dots, \boldsymbol{z}_t)$ is a basis for $W_1 + W_2$. Set $s = n - [k + t]$ and let $(\boldsymbol{p}_1, \dots, \boldsymbol{p}_s)$ be sequence of vectors such that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_j, \boldsymbol{u}_1, \dots, \boldsymbol{u}_t, \boldsymbol{v}_1, \dots, \boldsymbol{v}_t, \boldsymbol{p}_1, \dots, \boldsymbol{p}_s)$ is a basis for V and $(\boldsymbol{q}_1, \dots, \boldsymbol{q}_s)$ a sequence of vector such that $(\boldsymbol{y}_1, \dots, y_j, \boldsymbol{w}_1, \dots, \boldsymbol{w}_t, \boldsymbol{z}_1, \dots, \boldsymbol{z}_t, \boldsymbol{q}_1, \dots, \boldsymbol{q}_s)$ is also a basis for V. Now let S be the linear transformation such that

$$S(\boldsymbol{x}_i) = \boldsymbol{y}_i \text{ for } 1 \le i \le j$$
$$S(\boldsymbol{u}_i) = \boldsymbol{w}_i \text{ for } 1 \le i \le j$$
$$S(\boldsymbol{v}_i) = \boldsymbol{z}_i \text{ for } 1 \le i \le j$$
$$S(\boldsymbol{p}_i) = \boldsymbol{q}_i \text{ for } 1 \le i \le s$$

Then S is an invertible operator, $S(U_1) = W_1, S(U_2) = W_2$.

3. Since every one dimensional subspace is an intersection of the k-dimensional subspaces which contain it we can conclude that every vector of V is an eigenvector. Then by the proof of Lemma (11.1), $T \in Z(GL(V))$.

4. We may assume that $H_1 \neq H_2$. Then $dim(H_1 \cap H_2) = n-2$ and $P \subset H_1 \cap H_2$. Let $P = Span(\boldsymbol{x}_1)$ and extend to a basis $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2})$ for $H_1 \cap H_2$. Let x_{n-1} be a vector such that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1})$ is a basis for H_1 and let $\boldsymbol{x}_n$ be a vector in H_2 such that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2}, \boldsymbol{x}_n)$ is a basis for H_2. Note that $\boldsymbol{x}_{n-1} \notin H_2$ and $\boldsymbol{x}_n \notin H_1$. Let $S \in \chi(P, H_1)$ and $T \in \chi(P, H_2)$. For $1 \le i \le n-2, S(\boldsymbol{x}_i = T(\boldsymbol{x}_i) = \boldsymbol{x}_i$ so, in particular, $ST(\boldsymbol{x}_i) = TS(\boldsymbol{x}_i) = \boldsymbol{x}_i$. It there suffices to prove that $ST(\boldsymbol{x}_j) = TS(\boldsymbol{x}_j)$ for $j = n-1, n$. Now $S(\boldsymbol{x}_{n-1}) = \boldsymbol{x}_{n-1}$ and there is a scalar, a, such that $S(\boldsymbol{x}_n) = \boldsymbol{x}_n + a\boldsymbol{x}_1$. Similarly, $T(\boldsymbol{x}_n) = \boldsymbol{x}_n$ and there is scalar, b, such that $T(\boldsymbol{x}_{n-1}) = \boldsymbol{x}_{n-1} + b\boldsymbol{x}_1$. Now

$$TS(\boldsymbol{x}_{n-1}) = T(\boldsymbol{x}_{n-1}) = \boldsymbol{x}_{n-1} + b\boldsymbol{x}_1,$$

$$TS(\boldsymbol{x}_n) = T(\boldsymbol{x}_n + a\boldsymbol{x}_1) = \boldsymbol{x}_n + a\boldsymbol{x}_1.$$

$$ST(\boldsymbol{x}_{n-1}) = S(\boldsymbol{x}_{n-1}) = \boldsymbol{x}_{n-1} + b\boldsymbol{x}_1,$$

$$ST(\boldsymbol{x}_n) = S(\boldsymbol{x}_n) = \boldsymbol{x}_n + a\boldsymbol{x}_1.$$

Since ST and TS agree on a basis, $ST = TS$.

5. Set $\boldsymbol{x}'_{n-1} = b\boldsymbol{x}_{n-1} - a\boldsymbol{x}_n$ and $H = Span(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2}, \boldsymbol{x}'_{n-1})$. Then $ST(\boldsymbol{x}'_{n-1}) = \boldsymbol{x}'_{n-1}$. Since $ST(\boldsymbol{x}_n) = \boldsymbol{x}_n + a\boldsymbol{x}_1$ it follows that $ST = TS \in \chi(P, H)$.

6. Let $P_i = Span(\boldsymbol{x}_i), i = 1, 2$. Extend $\boldsymbol{x}_1, \boldsymbol{x}_2)$ to a basis $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1})$ of H and then to a basis $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ of V. Let $S \in \chi(P_1, H)$ and $T \in \chi(P_2, H)$. Now $S(\boldsymbol{x}_j) = \boldsymbol{x}_j = T(\boldsymbol{x}_j)$ for $1 \leq j \leq n - 1$. Moreover, there are scalars, a, b such that $S(\boldsymbol{x}_n) = \boldsymbol{x}_n + a\boldsymbol{x}_1$ and $T(\boldsymbol{x}_n) = \boldsymbol{x}_n + b\boldsymbol{x}_2$. Then

$$ST(\boldsymbol{x}_n) = S(\boldsymbol{x}_n + b\boldsymbol{x}_2) = \boldsymbol{x}_n + a\boldsymbol{x}_1 + b\boldsymbol{x}_2, TS(\boldsymbol{x}_n) =$$

$$T(\boldsymbol{x}_n + a\boldsymbol{x}_1) = \boldsymbol{x}_n + b\boldsymbol{x}_2 + a\boldsymbol{x}_1.$$

Since ST and TS agree on a basis, $ST = TS$.

7. Set $\boldsymbol{x}'_1 = a\boldsymbol{x}_1 + b\boldsymbol{x}_2$ and $P' = Span(\boldsymbol{x}'_1)$. Now $ST = TS$ is the identity when restricted to H and $ST(\boldsymbol{x}_n) = \boldsymbol{x}_n + \boldsymbol{x}'_1$ and therefore $ST \in \chi(P', H)$.

8. Let $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2})$ be a basis of $H_1 \cap H_2$ and let $\boldsymbol{y}_i \in P_i, i = 1, 2$ be nonzero vectors. Since $\boldsymbol{y}_1 \notin H_2$, in particular, $\boldsymbol{y}_2 \notin H_1 \cap H_2$ so that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2}, \boldsymbol{y}_1)$ is a linearly independent and therefore a basis of H_1. Similarly, $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2}, \boldsymbol{y}_2)$ is a basis of H_2. Set $W = Span(\boldsymbol{y}_1, \boldsymbol{y}_2)$. Note that if S is in the subgroup of $GL(V)$ generated by $\chi(P_1, H_1)$ and $\chi(P_2, H_2)$ then $S(\boldsymbol{v}) = \boldsymbol{v}$ for every vector $\boldsymbol{v} \in H_1 \cap H_2$. Consequently, the map $S \to S_{|W}$ is an injective homomorphism. Set $\mathcal{B}_W = (\boldsymbol{y}_1, \boldsymbol{y}_2)$. Denote by $\pi(S)$ the restriction of S to W.

Now assume $S \in \chi(P_1, H_1)$. Then $S(\boldsymbol{y}_1) = \boldsymbol{y}_1$ and there is a scalar a such that $S(\boldsymbol{y}_2) = a\boldsymbol{y}_1 + \boldsymbol{y}_2$. Therefore $\mathcal{M}_{\pi(S)}(\mathcal{B}_W, \mathcal{B}_w) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$. If $T \in \chi(P_2, H_2)$ then $T(\boldsymbol{y}_2) = \boldsymbol{y}_2$ and there is a scalar b such that $T(\boldsymbol{y}_1) = \boldsymbol{y}_1 + b\boldsymbol{y}_2$. Therefore $\mathcal{M}_{\pi(T)}(\mathcal{B}_W, \mathcal{B}_W) = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$. It follows that $\langle \chi(P_1, H_2), \chi(P_2, H_2) \rangle$ is isomorphic to $SL_2(\mathbb{F}) = SL(W)$.

9. First assume that $P_1 + P_2 \subset H_1 \cap H_2$. Assume $P_i = Span(\boldsymbol{x}_i), i = 1, 2$. Then $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is linearly independent. Extend to a basis $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2})$ of $H_1 \cap H_2$. Let $\boldsymbol{y}_i \in H_i, i = 1, 2$ such that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2}, \boldsymbol{y}_i)$ is a basis of H_i. Let $S_i \in \chi(P_i, H_i), i = 1, 2$. Then $S_i(\boldsymbol{x}_j) = \boldsymbol{x}_j$ for $i = 1, 2$ and $1 \leq j \leq n - 2$ Moreover, $S_i(\boldsymbol{y}_i) = \boldsymbol{y}_i$ for $i = 1, 2$. On the other hand, there are scalars, a_1, a_2 such that $S_1(\boldsymbol{y}_2) = a_1\boldsymbol{x}_1 + \boldsymbol{y}_1$ and $S_2(\boldsymbol{y}_1) = \boldsymbol{y}_1 + a_2\boldsymbol{x}_2$. To prove that $S_1S_2 = S_2S_1$ we need only show that S_1S_2 agree on $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$.

$$S_1S_2(\boldsymbol{y}_1) = S_1(\boldsymbol{y}_1 + a\boldsymbol{x}_2) = \boldsymbol{y}_1 + a_2\boldsymbol{x}_2,$$

$$S_2S_1(\boldsymbol{y}_1) = S_2(\boldsymbol{y}_1) = \boldsymbol{y}_1 + a_2\boldsymbol{x}_2.$$

$$S_1S_2(\boldsymbol{y}_2) = S_1(\boldsymbol{y}_2) = a_1\boldsymbol{x}_1 + \boldsymbol{y}_2,$$

$$S_2S_1(\boldsymbol{y}_2) = S_2(a_1\boldsymbol{x}_1 + \boldsymbol{y}_2) = a_1\boldsymbol{x}_1 + \boldsymbol{y}_2.$$

So we must now show if $P_1 + P_2$ is not contained in $H_1 \cap H_2$ then $\chi(P_1, H_1)$ and $\chi(P_2, H_2)$ do not commute. By Exercise 8 we can assume that either $P_1 \subset H_2$ or $P_2 \subset H_1$. Without loss of generality assume $P_1 \subset H_2$. Let $P_1 = Span(\boldsymbol{x}), P_2 = Span(\boldsymbol{y})$. Since $\boldsymbol{x} \in H_1 \cap H_2$ there is a basis $(\boldsymbol{x} = \boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2})$ for $H_1 \cap H_2$. Since $\boldsymbol{y} \notin H_1 \cap H_2$ the sequence $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-2}, \boldsymbol{y})$ is basis of H_2. Let $\boldsymbol{x}_{n-1}$ be a vector in H_1 such that $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n-1})$ is a basis for H_1. Now let $S_i \in \chi(P_i, H_i)$ for $i = 1, 2$. Then $S_1(\boldsymbol{x}_j) = \boldsymbol{x}_j$ for $1 \leq j \leq n - 1$ and there is a scalar, a_1 such that $S_1(\boldsymbol{y}) = \boldsymbol{y} + a_1\boldsymbol{x}_1$. On the other hand, $S_2(\boldsymbol{x}_j) = \boldsymbol{x}_j$ for $1 \leq j \leq n - 2, S(\boldsymbol{y}) = \boldsymbol{y}$ and there is a scalar, a_2 such that $S_2(\boldsymbol{x}_{n-1}) = \boldsymbol{x}_{n-1} + a_2\boldsymbol{x}_2$. We now compute $S_1S_2(\boldsymbol{x}_{n-1})$ and $S_2S_1(\boldsymbol{x}_{n-1})$.

$$S_1S_2(\boldsymbol{x}_{n-1}) = S_1(\boldsymbol{x}_{n-1} + a_2\boldsymbol{y}) = \boldsymbol{x}_{n-1} + a_2\boldsymbol{y}_2 + a_1a_2\boldsymbol{x},$$

$$S_2S_1(\boldsymbol{x}_{n-1}) = S_2(\boldsymbol{x}_{n-1}) = \boldsymbol{x}_{n-1} + a_2\boldsymbol{y}.$$

Thus, $S_1S_2 \neq S_2S_1$.

10. Let $T \in \chi(P, H)$ and $\boldsymbol{y} = S(\boldsymbol{x}), \boldsymbol{x} \in H$. Then $(STS^{-1})(\boldsymbol{y}) = (STS^{-1})(S(\boldsymbol{x})) = ST(\boldsymbol{x}) = S(\boldsymbol{x}) = \boldsymbol{y}$.

On the other hand, assume $\boldsymbol{w} \in V, \boldsymbol{w} \notin S(H)$. Set $\boldsymbol{u} = S^{-1}(\boldsymbol{w})$ so that $\boldsymbol{u} \notin H$. Now $(STS^{-1})(\boldsymbol{w}) - \boldsymbol{w} = (STS^{-1})(S(\boldsymbol{u})) - S(\boldsymbol{u}) = S(T(\boldsymbol{u}) - \boldsymbol{u})$. Since $T \in \chi(P, H), T(\boldsymbol{u}) - \boldsymbol{u} \in P$ and therefore $S(T(\boldsymbol{u}) - \boldsymbol{u}) \in S(P)$ so that $STS^{-1} \in \chi(S(P), S(H))$. This proves that $S\chi(P, H)S^{-1} \subset \chi(S(P), S(H))$. Since this also applies to S^{-1} we get the reverse inclusion and equality.

11.2. Symplectic Groups

1. If $\boldsymbol{y} \in \boldsymbol{x}^\perp$ then $T_{\boldsymbol{x},c}(\boldsymbol{y}) = \boldsymbol{y}$. On the other hand, if $\boldsymbol{y} \notin \boldsymbol{x}^\perp$ then $(T_{\boldsymbol{x},c} - I_V)(\boldsymbol{y}) = cf(\boldsymbol{y}, \boldsymbol{x})\boldsymbol{x} \in \boldsymbol{x}^\perp$. Thus, $T_{\boldsymbol{x},c} \in \chi(Span(\boldsymbol{x}), \boldsymbol{x}^\perp)$ and, therefore, is a transvection.

2. If $\boldsymbol{y} \in \boldsymbol{x}^\perp$ then $T_{\boldsymbol{x},c}T_{\boldsymbol{x},d}(\boldsymbol{y}) = \boldsymbol{y} = T_{\boldsymbol{x},c+d}(\boldsymbol{y})$. On the other hand, assume $f(\boldsymbol{y}, \boldsymbol{x}) = 1$. Then

$$T_{\boldsymbol{x},c}T_{\boldsymbol{x},d}(\boldsymbol{y}) = T_{\boldsymbol{x},c}(\boldsymbol{y} + d\boldsymbol{x}) = \boldsymbol{y} + d\boldsymbol{x} + c\boldsymbol{y} =$$

$$\boldsymbol{y} + (c + d)\boldsymbol{x} = T_{\boldsymbol{x},c+d}(\boldsymbol{y}).$$

3. If $\boldsymbol{y} \in \boldsymbol{x}^\perp$ then $T_{b\boldsymbol{x},c}(\boldsymbol{y}) = \boldsymbol{y} = T_{\boldsymbol{x},b^2c}(\boldsymbol{y})$. On the other hand, assume $f(\boldsymbol{y}, \boldsymbol{x}) = 1$ then

$$T_{b\boldsymbol{x},c}(\boldsymbol{y}) = \boldsymbol{y} + cf(\boldsymbol{u}, b\boldsymbol{x})(b\boldsymbol{x}) = \boldsymbol{y} + b^2c\boldsymbol{x} = T_{\boldsymbol{x},b^2c}(\boldsymbol{y}).$$

4. If $\boldsymbol{x} \perp \boldsymbol{y}$ then $Span(\boldsymbol{x}) + Span(\boldsymbol{y}) \subset \boldsymbol{x}^\perp \cap \boldsymbol{y}^\perp$. By Exercise 9 of Section (11.1) $\chi(\boldsymbol{x}) = \chi(Span(\boldsymbol{x}), \boldsymbol{x}^\perp)$ and $\chi(Span(\boldsymbol{y}), \boldsymbol{y}^\perp) = \chi(\boldsymbol{y})$ commute.

5. This is an instance of Exercise 10 of Section (11.1).

6. Let $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\boldsymbol{y}_1, \dots, \boldsymbol{y}_n)$ be a hyperbolic basis and $T \in \Psi(\boldsymbol{x}_1)$ act as follows:

$$T(\boldsymbol{x}_1) = \boldsymbol{x}_1$$

$$T(\boldsymbol{y}_1) = \boldsymbol{y}_1 + \sum_{k=2}^{n}(a_k\boldsymbol{x}_k + b_k\boldsymbol{y}_k) + \gamma\boldsymbol{x}_1$$

$$T(\boldsymbol{x}_j) = \boldsymbol{x}_j - b_j\boldsymbol{x}_1 \text{ for } 2 \leq j \leq n$$
$$T(\boldsymbol{y}_j) = \boldsymbol{y}_j + a_j\boldsymbol{x}_1 \text{ for } 2 \leq j \leq n$$

If $S \in Sp(V)$ set $\boldsymbol{x}'_j = S(\boldsymbol{x}_j)$ and $\boldsymbol{y}'_j = S(\boldsymbol{y}_j)$. If $T' = STS^{-1}$ then

$$T'(\boldsymbol{x}'_1) = \boldsymbol{x}'_1$$

$$T'(\boldsymbol{y}'_1) = \boldsymbol{y}'_1 + \sum_{k=2}^{n}(a_k\boldsymbol{x}'_k + b_k\boldsymbol{y}'_k) + \gamma\boldsymbol{x}'_1$$

$$T'(\boldsymbol{x}'_j) = \boldsymbol{x}'_j - b_j\boldsymbol{x}'_1 \text{ for } 2 \leq j \leq n$$
$$T'(\boldsymbol{y}'_j) = \boldsymbol{y}'_j + a_j\boldsymbol{x}'_1 \text{ for } 2 \leq j \leq n$$

Thus, $STS^{-1} = T' \in \Psi(\boldsymbol{x}'_1)$.

7. This follows from Exercise 6.

8. We continue with the notation of Exercise 6. For $\boldsymbol{a} = \begin{pmatrix} a_2 \\ \vdots \\ a_n \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} b_2 \\ \vdots \\ b_n \end{pmatrix}$ and $\gamma \in \mathbb{F}$ denote by $T(\boldsymbol{a}, \boldsymbol{b}, \gamma)$ the operator with the action as in Exercise 6. Let also $\boldsymbol{c} = \begin{pmatrix} c_2 \\ \vdots \\ c_n \end{pmatrix}, \boldsymbol{d} = \begin{pmatrix} d_2 \\ \vdots \\ d_n \end{pmatrix} \in \mathbb{F}^{n-1}$ and $\delta \in \mathbb{F}$. It is a straightforward calculation to see that

$$T(\boldsymbol{a}, \boldsymbol{b}, \gamma)T(\boldsymbol{c}, \boldsymbol{d}, \delta) =$$

$$T(\boldsymbol{a} + \boldsymbol{c}, \boldsymbol{b} + \boldsymbol{d}, \delta + \gamma - \boldsymbol{b}^{tr}\boldsymbol{d} + \boldsymbol{a}^{tr}\boldsymbol{c}) =$$

$$T(\boldsymbol{c}, \boldsymbol{d}, \delta)T(\boldsymbol{a}, \boldsymbol{b}, \gamma).$$

9. Let $X = Span(\boldsymbol{x}), Y_1 = Span(\boldsymbol{y})$ and $Y_2 = Span(\boldsymbol{z})$ where $f(\boldsymbol{y}, \boldsymbol{x}) = 1 = f(\boldsymbol{z}, \boldsymbol{x})$. Set $\boldsymbol{x}_1 = \boldsymbol{x}$ and $\boldsymbol{y}_1 = \boldsymbol{y}$ and extend $(\boldsymbol{x}_1, \boldsymbol{y}_1)$ to a hyperbolic basis

$(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$. Write $\boldsymbol{z}$ as a linear combination of this basis and note that since $f(\boldsymbol{z}, \boldsymbol{x}) = f(\boldsymbol{y}, \boldsymbol{x}) = 1$ the coefficient of $\boldsymbol{y}_1$ is 1. Thus,

$$\boldsymbol{z} = \boldsymbol{y}_1 + \sum_{k=2}^{n} (a_k \boldsymbol{x}_k + b_k \boldsymbol{y}_k) + \gamma \boldsymbol{x}_1.$$

By Lemma (11.15) if $T \in \Psi(\boldsymbol{x})$ and $T(\boldsymbol{y}) = T(\boldsymbol{y}_1) = \boldsymbol{z}$ then T is unique.

10. Let $(\boldsymbol{x} = \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ be a hyperbolic basis of V. Set $\mathcal{S}_0 = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ and for $2 \le j \le n$ set $\mathcal{S}_j = (\mathcal{S}_0 \cup \{\boldsymbol{y}_j\}) \setminus \{\boldsymbol{x}_j\}$ and set $M_j = Span(\mathcal{S}_j)$. Then

$$\cap_{j=2}^{n} M_j = Span(\boldsymbol{x}_1).$$

11. This is the same as Exercise 10 of Section (8.2).

12. By definition 0 is an identity element. Also every element $\boldsymbol{v} \in V$ is its own inverse with respect to 0. The definition of addition for $\alpha, \beta \in V \setminus \{0\}$ is symmetric in α and β and therefore is commutative. It remains to show the addition is associative. If $\alpha, \beta \in V \setminus \{0\}$ then clearly $(\alpha + \beta) + 0 = \alpha + (\beta + 0)$. Thus, we can assume our three elements are not zero. So let $\alpha, \beta, \gamma \in [1, 6]^{\{2\}}$. There are several cases to consider: a) $|\alpha \cup \beta \cup \gamma| = 3$. Then $\alpha + \beta = \gamma, \beta + \gamma = \alpha, \gamma + \beta = \alpha$ and $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) = 0$. b) $\alpha \cup \beta \cup \gamma = [1, 6]$. In this case we also have $\alpha + \beta = \gamma, \beta + \gamma = \alpha, \gamma + \alpha = \beta$ and $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) = 0$. c) $|\alpha \cup \beta \cup \gamma| = 1$. Then $|\alpha \cup \beta \cup \gamma| = 4$ and $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) = [1, 6] \setminus (\alpha \cup \beta \cup \gamma)$. d) $\alpha \cap \beta \cap \gamma = \emptyset$ and $|\alpha \cup \beta \cup \gamma| =$. Without loss of generality we can assume that $\alpha \cap \beta = \emptyset, \gamma \subset \alpha \cup \beta$. Then $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) = (\alpha \cup \beta) \setminus \gamma$. e) $|\alpha \cup \beta \cup \gamma| = 5$. Then we can assume that $\alpha \cap \beta = \emptyset$ and $|\alpha \cap \gamma| = 1, \beta \cap \gamma = \emptyset$. Let $[1, 6] \setminus (\alpha \cup \beta \cup \gamma) = \{i\}$ and $\alpha \cap \gamma = \{j\}$. Then $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) = \{i, j\}$.

13. If f is bilinear then by its definition it is alternating since $f(\boldsymbol{v}, \boldsymbol{v}) = 0$ for every $\boldsymbol{v} \in V$. Now $f(\boldsymbol{v}, \boldsymbol{w}) = f(\boldsymbol{v} + 0, \boldsymbol{w}) = f(\boldsymbol{v}, \boldsymbol{w}) + 0 = f(\boldsymbol{v}, \boldsymbol{w}) + f(0, \boldsymbol{w})$. Also, for $\boldsymbol{v}, \boldsymbol{w} \in V, f(\boldsymbol{v} + \boldsymbol{w}, 0) = f(\boldsymbol{v}, 0) + f(\boldsymbol{w}, 0)$ since both sides are zero. We may therefore assume that the three vectors are $\alpha, \beta, \gamma \in [1, 6]^{\{2\}}$. Again there are several cases to consider: a) $\alpha = \beta$. Then $f(\alpha + \beta, \gamma) = f(0, \gamma) = 0$. On the other hand, $f(\alpha, \gamma) + f(\beta, \gamma) = f(\alpha, \gamma) + f(\alpha, \gamma) = 0$. b) $\alpha = \gamma$ and $\alpha \cap \beta = \emptyset$. Then $(\alpha + \beta) \cap \gamma = (\alpha + \beta) \cap \alpha = \emptyset$. Then $f(\alpha + \beta, \gamma) = f(\alpha + \beta, \alpha) = 0$. On the other hand, $f(\alpha, \gamma) + f(\beta, \gamma) = f(\alpha, \alpha) + f(\beta, \alpha) = 0 + 0$, the latter since $\beta \cap \alpha = \emptyset$. In the remaining cases we can now assume that α, β, γ are distinct. c) $\alpha \cap \beta = \emptyset, \gamma \subset \alpha \cup \beta$. Then $(\alpha + \beta) \cap \gamma = \emptyset$ and $f(\alpha + \beta, \gamma) = 0$. On the other hand under these assumptions, $|\alpha \cap \gamma| = |\beta \cap \gamma| = 1$, and then $f(\alpha, \gamma) = f(\beta, \gamma) = 1$ so that $f(\alpha, \gamma) + f(\beta, \gamma) = 0$. d) $\alpha \cup \beta \cup \gamma = [1, 6]$. Then $\alpha + \beta = \gamma$ so that $f(\alpha + \beta, \gamma) = 0 = f(\alpha, \gamma) = f(\beta, \gamma)$. Consequently, $f(\alpha + \beta, \gamma) = f(\alpha, \gamma) + f(\beta, \gamma)$. e) $\alpha \cap \beta = \emptyset = \alpha \cap \gamma, |\beta \cap \gamma| = 1$. Then $|(\alpha + \beta) \cap \gamma| = 1$ and $f(\alpha + \beta, \gamma) = 1$. On the other hand $f(\alpha, \gamma) = 0, f(\beta, \gamma) = 1$ and we again get equality. We can now assume that α, β, γ are distinct and $|\alpha \cap \beta| = 1$. f) We now treat the case that $\gamma \subset \alpha \cup \beta$, that is, $\gamma = \alpha + \beta$. Then $f(\alpha, \gamma) = f(\beta, \gamma) = 1$ and $f(\alpha + \beta, \gamma) = f(\gamma, \gamma) = 0$ and we have the required equality. g) $|\alpha \cap \beta \cap \gamma| = 1$. In this case $(\alpha + \beta) \cap \gamma = \emptyset$ so $f(\alpha + \beta, \gamma) = 0$ whereas $f(\alpha, \gamma) = f(\beta, \gamma) = 1$. h) $\alpha \cap \beta \cap \gamma = \emptyset, |\alpha \cup \beta \cup \gamma| = 4$. Then we may assume that $\alpha \cap \gamma = \emptyset, \gamma$ intersects β and $\alpha + \beta$. Then $f(\alpha + \beta, \gamma) = 1, f(\alpha, \gamma) = 0, f(\beta, \gamma) = 1$ and we have equality. i) Finally, we have the case where γ is disjoint from $\alpha \cup \beta$. In this case $f(\alpha + \beta, \gamma) = 0 = f(\alpha, \gamma) + f(\beta, \gamma)$ and we are done.

14. Since $\pi \in S_6$ fixes 0, we have for any $\alpha \in [1, 6]^{\{2\}}$ that $f(\pi(\alpha), 0) = 0 = f(\alpha, 0)$. Suppose $\alpha, \beta \in [1, 6]^{\{2\}}$. Then $\alpha \cap \beta = \emptyset$ if and only if $\pi(\alpha) \cap \pi(\beta) = \emptyset$ and therefore $f(\pi(\alpha), \pi(\beta)) = f(\alpha, \beta)$. Thus, S_6 acts as isometries of the symplectic space (V, f). By Exercise 11, $|Sp_4(2)| = 2^4(2^4 - 1)(2^2 - 1) = 16 \times 15 \times 3 = 2^4 \times 3^2 \times 5 = 6! = |S_6|$. Therefore, S_6 is isomorphic to $Sp_4(2)$.

11.3. Orthogonal Groups, Characteristic Not Two

1. If $\boldsymbol{x} \in \boldsymbol{u}^{\perp} \cap \boldsymbol{y}^{\perp}$ then $\boldsymbol{x} \in \boldsymbol{z}^{\perp}$ and $\rho_{\boldsymbol{z}}\rho_{\boldsymbol{y}}(\boldsymbol{x}) = \boldsymbol{x}$. On the other hand, $\tau_{\boldsymbol{u},\boldsymbol{y}}(\boldsymbol{x}) = \boldsymbol{x} + \langle \boldsymbol{x}, \boldsymbol{y}\rangle_\phi \boldsymbol{u} = \boldsymbol{x}$. It now suffices to show that $\rho_{\boldsymbol{z}}\rho_{\boldsymbol{y}}(\boldsymbol{y}) = \tau_{\boldsymbol{u},\boldsymbol{y}}(\boldsymbol{y}) = \boldsymbol{y} + \langle \boldsymbol{y}, \boldsymbol{y}\rangle_\phi \boldsymbol{u}$. Set $a = \langle \boldsymbol{y}, \boldsymbol{y}\rangle_\phi$ so that $\boldsymbol{z} = \frac{a}{2}\boldsymbol{u} + \boldsymbol{y}$. Now $\langle \boldsymbol{z}, \boldsymbol{z}\rangle_\phi = \langle \boldsymbol{y}, \boldsymbol{y}\rangle_\phi = \langle, \boldsymbol{z}, \boldsymbol{y}\rangle_\phi$. Now $\rho_{\boldsymbol{z}}(\boldsymbol{y}) =$

$$\boldsymbol{y} - 2[\frac{\langle \boldsymbol{y}, \boldsymbol{z}\rangle_\phi}{\langle \boldsymbol{z}, \boldsymbol{z}\rangle_\phi}\boldsymbol{z}] =$$

$$\boldsymbol{y} - 2\boldsymbol{z} = \boldsymbol{y} - 2(\frac{a}{2}\boldsymbol{u} + \boldsymbol{y}) =$$

$$\boldsymbol{y} - a\boldsymbol{u} - 2\boldsymbol{y} = -\boldsymbol{y} - \langle \boldsymbol{y}, \boldsymbol{y}\rangle_\phi \boldsymbol{u}.$$

It then follows that $\rho_{\boldsymbol{z}}\rho_{\boldsymbol{y}}(\boldsymbol{y}) = \rho_{\boldsymbol{z}}(-\boldsymbol{y}) = \boldsymbol{y} + \langle \boldsymbol{y}, \boldsymbol{y}\rangle_\phi \boldsymbol{u} = \tau_{\boldsymbol{u},\boldsymbol{y}}(\boldsymbol{y})$.

2. Assume $\boldsymbol{w} = \boldsymbol{v} + c\boldsymbol{u}$ for some scalar c. Let $\boldsymbol{x} \in \boldsymbol{u}^{\perp}$. Then

$$\tau_{\boldsymbol{u},\boldsymbol{w}}(\boldsymbol{x}) = \boldsymbol{x} + \langle \boldsymbol{x}, \boldsymbol{w}\rangle_\phi \boldsymbol{u} =$$

$$\boldsymbol{x} + \langle \boldsymbol{x}, \boldsymbol{v} + c\boldsymbol{u}\rangle_\phi \boldsymbol{u} =$$

$$\boldsymbol{x} + \langle \boldsymbol{x}, \boldsymbol{v}\rangle_\phi \boldsymbol{u} = \tau_{\boldsymbol{u},\boldsymbol{v}}(\boldsymbol{x}).$$

Since $\tau_{\boldsymbol{u},\boldsymbol{v}}$ and $\tau_{\boldsymbol{u},\boldsymbol{w}}$ agree on $\boldsymbol{u}^{\perp}$ they are identical. Converely, assume $\tau_{\boldsymbol{u},\boldsymbol{v}} = \tau_{\boldsymbol{u},\boldsymbol{w}}$. Let $\boldsymbol{x} \in \boldsymbol{u}^{\perp}$. Then $\langle \boldsymbol{x}, \boldsymbol{v}\rangle_\phi = \langle \boldsymbol{x}, \boldsymbol{w}\rangle_\phi$ so that $\langle \boldsymbol{x}, \boldsymbol{v} - \boldsymbol{w}\rangle_\phi = 0$. Consequently, $\boldsymbol{v} - \boldsymbol{w} \in Rad(\boldsymbol{u}^{\perp}) = Span(\boldsymbol{u})$.

3. We need to prove if $\boldsymbol{v} \in \boldsymbol{u}^{\perp}$ is a singular vector then $\tau_{\boldsymbol{u},\boldsymbol{v}}$ is in the subgroup of $T_{\boldsymbol{u}}$ generated by all $\tau_{\boldsymbol{u},\boldsymbol{x}}$ where $\boldsymbol{x} \in \boldsymbol{u}^{\perp}$ and $\boldsymbol{x}$ is non-singular. Let $\boldsymbol{x} \in \boldsymbol{u}^{\perp}$ such that $\langle \boldsymbol{v}, \boldsymbol{x}\rangle_\phi = \langle \boldsymbol{x}, \boldsymbol{x}\rangle_\phi$. Since $\boldsymbol{x} \not\perp \boldsymbol{v}$ the subspace $Span(\boldsymbol{v}, \boldsymbol{x})$ is non-degenerate. Let $\boldsymbol{y}$ be a vector in $Span(\boldsymbol{v}, \boldsymbol{x})$ such that $\boldsymbol{y} \perp \boldsymbol{x}$ and $\langle \boldsymbol{v}, \boldsymbol{y}\rangle_\phi = \langle \boldsymbol{y}, \boldsymbol{y}\rangle_\phi$. It is then the case that $\boldsymbol{v} = \boldsymbol{x} + \boldsymbol{y}$. By Lemma (11.25) $\tau_{\boldsymbol{u},\boldsymbol{v}} = \tau_{\boldsymbol{u},\boldsymbol{x}}\tau_{\boldsymbol{u},\boldsymbol{y}}$.

4. Without loss of generality we can assume $\langle \boldsymbol{u}, \boldsymbol{v}\rangle_\phi = 1$. Suppose $\boldsymbol{x}$ is a singular vector and $\langle \boldsymbol{u}, \boldsymbol{x}\rangle_\phi = 1$. By Lemma (11.28) there is a unique element $\gamma \in T_{\boldsymbol{u}}$ such that $\gamma(\boldsymbol{v}) = \boldsymbol{x}$. Then $\gamma T_{\boldsymbol{v}}\gamma^{-1} = T_{\gamma(\boldsymbol{v})} = T_{\boldsymbol{x}}$. It follows that the subgroup generated by $T_{\boldsymbol{u}}, T_{\boldsymbol{v}}$ contains $T_{\boldsymbol{x}}$ for every singular vector $\boldsymbol{x}$. Consequently, the subgroup generated by $T_{\boldsymbol{u}}, T_{\boldsymbol{v}}$ contains $\Omega(V)$. Since $T_{\boldsymbol{u}}, T_{\boldsymbol{v}}$ are subgroups of $\Omega(V)$ we have equality.

5. Let H denote the subgroup of $\Omega(V)$ generated by $\chi(l_1) \cup \chi(l_2) \cup \chi(l_3) \cup \chi(l_4)$. We first point out that $\boldsymbol{x}_1^{\perp} = Span(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}_2)$. We claim that $T_{\boldsymbol{x}_1}$ is generated by $\chi(l_1) \cup \chi(l_4)$. Let $H_{\boldsymbol{x}_1}$ be the subgroup of $T_{\boldsymbol{x}_1}$ generated by $\chi(l_1) \cup \chi(l_4)$. We need to prove if $\boldsymbol{w} \in \boldsymbol{x}_1^{\perp}$ then $\tau_{\boldsymbol{x}_1,\boldsymbol{w}} \subset H_{\boldsymbol{x}_1}$. Since $\boldsymbol{w} \in \boldsymbol{x}_1^{\perp} = Span(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}_2)$ there are scalars a, b, c such that $\boldsymbol{w} = a\boldsymbol{x}_1 + b\boldsymbol{x}_2 + c\boldsymbol{y}_2$. By Exercise 2, $\tau_{\boldsymbol{x}_1,\boldsymbol{w}} = \tau_{\boldsymbol{x}_1,b\boldsymbol{x}_2+c\boldsymbol{y}_2}$ so we can assume $\boldsymbol{w} = b\boldsymbol{x}_2 + c\boldsymbol{y}_2$. By Lemma (11.25), $\tau_{\boldsymbol{x}_1,b\boldsymbol{x}_2+c\boldsymbol{y}_2} = \tau_{\boldsymbol{x}_1,b\boldsymbol{x}_2}\tau_{\boldsymbol{x}_1,c\boldsymbol{y}_2} \in H_{\boldsymbol{x}_1}$. In exactly the same way, $T_{\boldsymbol{x}_2}$ is the subgroup generated by $\chi(l_1) \cup \chi(l_2), T_{\boldsymbol{y}_1}$ is equal to the subgroup generated by $\chi(l_2) \cup \chi(l_3)$, and $T_{\boldsymbol{y}_2}$ is generated by $\chi(l_3) \cup \chi(l_4)$.

Suppose now that $\boldsymbol{u} = c\boldsymbol{x}_1 + \boldsymbol{x}_2$ is a vector in l_1. Set $\sigma = \tau_{\boldsymbol{x}_1,c\boldsymbol{x}_2}$. Then $\sigma(\boldsymbol{x}_2) = \boldsymbol{x}_2 = \langle \boldsymbol{x}_2, \boldsymbol{x}_1\rangle_\phi \boldsymbol{y}_2 + \langle \boldsymbol{x}_2, \boldsymbol{y}_2\rangle_\phi (c\boldsymbol{x}_1) = \boldsymbol{x}_2 + c\boldsymbol{x}_1$. It therefore follows that $\sigma T_{\boldsymbol{x}_2}\sigma^{-1} = T_{\boldsymbol{u}}$ is contained in H. In exactly the same way if $\boldsymbol{u} \in l_2 \cup l_3 \cup l_4$ then $T_{\boldsymbol{u}} \subset H$.

Now suppose $\boldsymbol{u}$ is an arbitrary singular vector. We must show that $T_{\boldsymbol{u}} \subset H$. By the above we can assume that $\boldsymbol{u} \notin l_1 \cup l_2 \cup l_3 \cup l_4$. Let $a\boldsymbol{x}_1 + \boldsymbol{x}_2$ be a vector in l_1 such that $a\boldsymbol{x}_1 + \boldsymbol{x}_2 \perp \boldsymbol{u}$ and let $b\boldsymbol{y}_2 + \boldsymbol{y}_1$ be a vector in l_3 such that $b\boldsymbol{y}_2 + \boldsymbol{y}_1 \perp \boldsymbol{u}$. It must be the case that $a\boldsymbol{x}_1 + \boldsymbol{x}_2 \perp b\boldsymbol{y}_2 + \boldsymbol{y}_1$ so that $b = -a$. Set $\boldsymbol{z}_1 = a\boldsymbol{x}_1 + \boldsymbol{x}_2$ and $\boldsymbol{z}_2 = \boldsymbol{y}_1 - a\boldsymbol{y}_2$. Then $\boldsymbol{u} \in Span(\boldsymbol{z}_1, \boldsymbol{z}_2)$. Since $T_{\boldsymbol{u}} = T_{c\boldsymbol{u}}$ for any nonzero scalar, c, we can assume that $\boldsymbol{u} = \boldsymbol{z}_1 + d\boldsymbol{z}_2$ for some scalar d. Now set $\gamma = \tau_{\boldsymbol{y}_2,d\boldsymbol{z}_2}$. Then $\gamma(\boldsymbol{z}_1) = \boldsymbol{z}_1 + d\boldsymbol{z}_2 = \boldsymbol{u}$. Then $T_{\boldsymbol{u}} = \gamma T_{\boldsymbol{z}_1}\gamma^{-1} \subset H$. It now follows that $\Omega(V) \subset H$ and since $H \subset \Omega(V)$ we have equality.

6. Let $\sigma \in L_1$ and $\boldsymbol{u} \in l_1$. We claim that $\sigma(\boldsymbol{u}) \in l_1$. It suffices to prove this for $\sigma \in \chi(l_2) \cup \chi(l_4)$. Suppose $\boldsymbol{u} = \boldsymbol{x}_1$. If $\sigma \in \chi(l_2)$ then $\sigma(\boldsymbol{u}) = \sigma(\boldsymbol{x}_1) = \boldsymbol{x}_1$. On the other hand, if $\sigma \in \chi(l_4)$, say $\sigma = \tau_{\boldsymbol{x}_2,a\boldsymbol{y}_1}$ then

$\sigma(\boldsymbol{u}) = \sigma(\boldsymbol{x}_1) = \tau_{\boldsymbol{x}_2,a\boldsymbol{y}_1}(\boldsymbol{x}_1) = \boldsymbol{x}_1 + \langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle_\phi a\boldsymbol{y}_1 + \langle \boldsymbol{x}_1, a\boldsymbol{y}_1\rangle_\phi \boldsymbol{x}_2 = \boldsymbol{x}_1 + a\boldsymbol{x}_2 \in l_1$. Similarly, if $\boldsymbol{u} = \boldsymbol{x}_2$ then $\sigma(\boldsymbol{x}_2) \in l_1$. Suppose $\boldsymbol{u} = \boldsymbol{x}_1 + a\boldsymbol{x}_2$. Then $\sigma(\boldsymbol{u}) = \sigma(\boldsymbol{x}_1) + a\sigma(\boldsymbol{x}_2) \in l_1$ since $\sigma(\boldsymbol{x}_1$ and $\sigma(\boldsymbol{x}_2) \in l_1$.

Suppose $\sigma = \tau_{\boldsymbol{x}_1,a\boldsymbol{y}_2}$. We determine the matrix of σ restricted to l_1 with respect to the basis $(\boldsymbol{x}_1, \boldsymbol{x}_2)$. By what we have shown this is $\begin{pmatrix} 1 & a \\ 0 & 1\end{pmatrix}$. Similarly, the matrix of $\tau_{\boldsymbol{x}_2,b\boldsymbol{y}_1}$ with respect to $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is $\begin{pmatrix} 1 & 0 \\ b & 1\end{pmatrix}$. It follows that the restriction of L_1 to l_1 is isomorphic to $SL_2(\mathbb{F})$. However, this map is injective and therefore L_1 is isomorphic to $SL_2(\mathbb{F})$. Similarly, L_2 is isomorphic to $SL_2(\mathbb{F})$.

7. Since $l_1 \cap l_4 = Span(\boldsymbol{x}_1), \chi(l_1)$ and $\chi(l_4)$ commute. Since $l_1 \cap l_2 = Span(\boldsymbol{x}_2), \chi(l_1)$ and $\chi(l_2)$ commute. Since $l_2 \cap l_3 = Span(\boldsymbol{y}_1)$ it follows that $\chi(l_2)$ and $\chi(l_3)$ commute. Finally, since $l_3 \cap l_4 = Span(\boldsymbol{y}_2)$ we can conclude that $\chi(l_3)$ and $\chi(l_4)$ commute. It now follows that L_1 and L_2 commute.

8. We have seen above that L_1 leaves l_1 invariant and therefore for $\sigma \in L_1, \sigma(B) = B$. If $\sigma \in \chi(l_1)$ acts trivially on l_1 then $\sigma(B) = B$. On the other hand, if $\sigma \in \chi(l_3)$ then $\sigma(B) \cap B = \emptyset$. It follows that B is a block of imprimitivity.

9. The Witt index of W is zero, that is, there are no singular vectors in W. In particular, for $\boldsymbol{v} = a\boldsymbol{x} + \boldsymbol{y}, \phi(a\boldsymbol{x} + \boldsymbol{y}) = a^2 + d \neq 0$. Since $a \in \mathbb{F}$ is arbitrary, there are no roots in $\mathbb{F}$ of the quadratic polynomial $X^2 + d$ which implies that $X^2 + d$ is irreducible in $\mathbb{F}[X]$.

10. Let $\boldsymbol{u}' = \begin{pmatrix} 1 & 0 \\ 0 & 0\end{pmatrix}, \boldsymbol{v}' = \begin{pmatrix} 0 & 0 \\ 0 & -1\end{pmatrix}$. Then $\boldsymbol{u}', \boldsymbol{v}'$ are singular vectors in (M, q) and $\langle \boldsymbol{u}', \boldsymbol{v}'\rangle_q = 1$. The orthogonal complement of U' in M is $\{\begin{pmatrix} 0 & \alpha \\ \overline{\alpha} & 0\end{pmatrix} | \alpha \in \mathbb{K}\}$. Set $\boldsymbol{x}' = \begin{pmatrix} 0 & 1 \\ 1 & 0\end{pmatrix}$ so that $q(\boldsymbol{x}') = 1$. Set $\boldsymbol{y}' = \begin{pmatrix} 0 & \omega \\ -\omega & 0\end{pmatrix}$. Then $q(\boldsymbol{y}') = -\omega^2 = d$ and

$$\langle \boldsymbol{x}', \boldsymbol{y}'\rangle_q =$$

$$q(\boldsymbol{x}' + \boldsymbol{y}') - q(\boldsymbol{x}') - q * \boldsymbol{y}') =$$

$$det\begin{pmatrix} 1+\omega & 0 \\ 0 & 1-\omega\end{pmatrix} -$$

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

$$det\begin{pmatrix} \omega & 0 \\ 0 & -\omega\end{pmatrix} =$$

$$(1+\omega)(1-\omega) - 1 + \omega^2 = 0.$$

It now follows that $q(a\boldsymbol{u}' + b\boldsymbol{v}' + c\boldsymbol{x}' + d\boldsymbol{y}') = ab + c^2 + e^2 d$ and the linear map that takes $\boldsymbol{u} \to \boldsymbol{u}', \boldsymbol{v} \to \boldsymbol{v}', \boldsymbol{x} \to \boldsymbol{x}', \boldsymbol{y} \to \boldsymbol{y}'$ is an isometry.

11. It suffices to prove that $A \cdot m \in M$ for $A = \begin{pmatrix} 1 & \beta \\ 0 & 1\end{pmatrix}$ and $A = \begin{pmatrix} 1 & 0 \\ \gamma & 1\end{pmatrix}$ since these matrices generate $SL_2(\mathbb{K})$ Thus, assume $m = \begin{pmatrix} a & \alpha \\ \overline{\alpha} & b\end{pmatrix}$ where $a, b \in \mathbb{F}$ and $\alpha \in \mathbb{K}$. If $A = \begin{pmatrix} 1 & \beta \\ 0 & 1\end{pmatrix}$ then

$$A \cdot m = \begin{pmatrix} a & a\beta + \alpha \\ a\overline{\beta} + \overline{\alpha} & a\beta\overline{\beta} + \alpha\overline{\beta} + \overline{\alpha}\beta + b\end{pmatrix}.$$

Since $\overline{a\beta + \alpha} = a\overline{\beta} + \overline{\alpha}$ and $a\beta\overline{\beta} + \alpha\overline{\beta} + \overline{\alpha}\beta + b \in \mathbb{F}$ it follows that $A \cdot m \in M$. In a similar fashion if $A = \begin{pmatrix} 1 & 0 \\ \beta & 1\end{pmatrix}$ then $A \cdot m \in M$.

12. If $A \in SL_2(\mathbb{K})$ and $m_1, m_2 \in M$ then $T_A(m_1 + m_2) = \overline{A}^{tr}(m_1 + m_2)A = \overline{A}^{tr} m_1 A + \overline{A}^{tr} m_2 A = T_A(m_1) + T_A(m_2)$.

If $A \in SL_2(\mathbb{K}), m \in M$ and $c \in \mathbb{F}$ then $T_A(cm) = \overline{A}^{tr}(cm)A = c(\overline{A}^{tr} mA = cT_A(m)$. Thus, T_A is a linear operator on M. Since $det(A) = 1$ also $det(\overline{A}^{tr}) = 1$. Then $det(\overline{A}^{tr} mA) = det(\overline{A}^{tr}) det(m) det(A) = det(A)$. Consequently, $q(T_A(m)) = -det(T_A(m)) = -det(m) = q(m)$. Thus, T_A is an isometry of (M, q).

13. Since the characteristic of $\mathbb{F}$ is not two, the center of $SL_2(\mathbb{K}) = \{\pm I_2\}$. This acts trivially on M and therefore

is in the kernel of the action. On the other hand, since $|\mathbb{K}| > 3, PSL_2(\mathbb{K})$ is simple. This implies the either $Range(T) = PSL_2(\mathbb{K})$ or the action is trivial, which it is not. Thus, $Range(T)$ is isomorphic to $PSL_2(\mathbb{K})$.

Let $\boldsymbol{z} = \begin{pmatrix} 0 & \alpha \\ \overline{\alpha} & 0 \end{pmatrix}$. It is straightforward to check that the action of $\tau_{\boldsymbol{u}',\boldsymbol{z}}$ is the same as T_A where $A = \begin{pmatrix} 1 & \overline{\alpha} \\ 0 & 1 \end{pmatrix}$. Thus, $Range(T)$ contains $T_{\boldsymbol{u}'}$. By Remark (11.5) $\Omega(M, q)$ is generated by $T_{\boldsymbol{u}'} \cup T_{\boldsymbol{v}'}$ which are both contained in $Range(T)$ and therefore, $Range(T)$ is isomorphic $PSL_2(\mathbb{K})$.

11.4. Unitary Groups

1. Assume τ restricted to W is a transvection of W, say $\tau \in \chi(X, X^{\perp} \cap W)$ where X is an isotropic subspace of W. Let $\widehat{\tau}$ be defined as follows: $\widehat{\tau}(\boldsymbol{w} + \boldsymbol{u}) = \tau(\boldsymbol{w}) + \boldsymbol{u}$ where $\boldsymbol{w} \in W$ and $\boldsymbol{u} \in W^{\perp}$. Then $\widehat{\tau} \in \chi(X, X^{\perp}) \in \Omega(V)$. If $T \in \Omega(W)$ express T as a product, $\tau_1 \dots \tau_k$ where $\tau_i \in \chi(X_i, X_i^{\perp} \cap W)$. Prove that $\widehat{T} = \widehat{\tau_1} \dots \widehat{\tau_k} \in \Omega(V)$.

2. Let $\boldsymbol{u}, \boldsymbol{v}$ be isotropic vectors such that $f(\boldsymbol{u}, \boldsymbol{v}) = 1$ and set $\mathcal{B} = (\boldsymbol{u}, \boldsymbol{v})$. Let $\mathbb{F}_4 = \{0, 1, \omega, \omega + 1 = \omega^2\}$. The six anistropic vectors are $\boldsymbol{u} + \omega\boldsymbol{v}, \omega\boldsymbol{u} + \omega^2\boldsymbol{v}, \omega^2\boldsymbol{u} + \boldsymbol{v}, \boldsymbol{u} + \omega^2\boldsymbol{v}, \omega\boldsymbol{u} + \boldsymbol{v}, \omega^2\boldsymbol{u} + \omega\boldsymbol{v}$. If $T \in SU(V)$ then the matrix of T with respect to $\mathcal{B}$ is one of the following:

$$I_2, \begin{pmatrix} 1 & \omega \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \omega^2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & \omega \\ \omega^2 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 0 & \omega \\ \omega^2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & \omega \\ \omega^2 & 0 \end{pmatrix}.$$

Apart from I_2 none of these leave the vector $\boldsymbol{u} + \omega\boldsymbol{v}$ fixed and so $SU(V)$ is transitive on the six vectors.

3. Let $\boldsymbol{w}$ be an anistropic vector in $Span(\boldsymbol{u}, \boldsymbol{v})^{\perp}$ and set $W = Span(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$. Then W is non-degenerate. By Lemma (11.37) there is an isometry T in $\Omega(W)$ such that $T(\boldsymbol{u}) = \omega\boldsymbol{u}, T(\boldsymbol{v}) = \omega\boldsymbol{v}$. Let τ be the isometry of V such that τ restricted to W is T and τ restricted to $W^{\perp}$ is the identity. Then $\tau \in \Omega(V)$.

4. Suppose first that $Span(\boldsymbol{u}) = Span(\boldsymbol{v})$. Let $\boldsymbol{w}$ be a singular vector such that $f(\boldsymbol{u}, \boldsymbol{w}) = 1$. By Exercise 3 there is a $\tau \in \Omega(V)$ such that $\tau(\boldsymbol{u}) = \omega\boldsymbol{u}, \tau(\boldsymbol{w}) = \omega\boldsymbol{w}$. Then either $\tau(\boldsymbol{u}) = \boldsymbol{v}$ or $\tau^2(\boldsymbol{u}) = \boldsymbol{v}$. We may therefore assume that $Span(\boldsymbol{u}) \neq Span(\boldsymbol{v})$. Next assume that $f(\boldsymbol{u}, \boldsymbol{v}) \neq 0$. Let $\boldsymbol{v}' \in Span(\boldsymbol{v})$ such that $f(\boldsymbol{u}, \boldsymbol{v}') = 1$ and set $W = Span(\boldsymbol{u}, \boldsymbol{v})$. Then there is a $\sigma \in \Omega(W)$ such that σ restricted to $W^{\perp}$ is the identity and $\sigma(\boldsymbol{u}) = \omega^2\boldsymbol{v}'$. It then follows by Exercise 3 that we can find a $\gamma \in \Omega(V)$ such that $\gamma(\boldsymbol{v}') = \omega(\boldsymbol{v}')$. Then one of the isometries $\sigma, \gamma\sigma, \gamma^2\sigma$ takes $\boldsymbol{u}$ to $\boldsymbol{v}$. Finally assume $Span(\boldsymbol{u}) \neq Span(\boldsymbol{v})$ and $\boldsymbol{u} \perp \boldsymbol{v}$. There exists an isotropic vector $\boldsymbol{x}$ such that $\boldsymbol{u} \not\perp \boldsymbol{x} \not\perp \boldsymbol{v}$. By what we have shown there exists $\tau_1, \tau_2 \in \Omega(V)$ such that $\tau_1(\boldsymbol{u}) = \boldsymbol{x}, \tau_2(\boldsymbol{x}) = \boldsymbol{v}$. Set $\tau = \tau_2\tau_1$. Then $\tau \in \Omega(V)$ and $\tau(\boldsymbol{u}) = \boldsymbol{v}$.

5. Let $\boldsymbol{u}$ be an isotropic vector and set $U = Span(\boldsymbol{u})$. If X is a non-degenerate two dimensional subspace containing $\boldsymbol{u}$ then $|I_1(X)| = 3$, one of which is U. The total number of two dimensional subspaces containing U is 21 and the number of two dimensional subspaces containing U and contained in $\boldsymbol{u}^{\perp}$ is five. Therefore there are 16 non-degenerate two dimensional subspaces containing U and $16 \times 2 = 32$ one spaces Y in $I_1(V)$ such that $Y \not\perp U$.

Now let $Y \in I_1(V)$ with $Y \not\perp U$ and assume Z is a totally isotropic two dimensional space containing U. Then $|I_1(Z)| = 5$, one of which is U. On the other hand, $Y^{\perp} \cap Z \in I_1(U^{\perp} \cap Y^{\perp})$. Now $U^{\perp} \cap Y^{\perp}$ is a non-degenerate two dimensional subspace and contains three one dimensional subspaces in $I_1(V)$. Thus, there exists exactly three totally isotropic subspaces of dimension two containing U. We can now conclude that there are 3×4 isotropic one spaces W in $U^{\perp}, W \neq U$. We therefore have $1 + 12 + 32 = 45$ one spaces in $I_1(V)$.

6. This was proved in the course of Exercise 5.

7. This was also proved in the course of Exercise 5.

8. Proved in Exercise 5.

9. $f(a\boldsymbol{x}_1+b\boldsymbol{x}_2+c\boldsymbol{y}_2+d\boldsymbol{y}_1, a\boldsymbol{x}_1+b\boldsymbol{x}_2+c\boldsymbol{y}_2+d\boldsymbol{y}_1) =$

$$a\overline{d}+d\overline{a}+b\overline{c}+c\overline{b} =$$

$$Tr(a\overline{d})+Tr(b\overline{c}).$$

Thus, $a\boldsymbol{x}_1+b\boldsymbol{x}_2+c\boldsymbol{y}_2+d\boldsymbol{y}_1$ is isotropic if and only if $Tr(a\overline{d})+Tr(b\overline{c})=0$.

10. If X is an anistropic one dimensional space then $X^\perp$ is a non-degenerate three dimensional subspace which contains 21 one dimensional subspaces. Moreover, if $W = X^\perp$ then $I_1(W) = 9$. Therefore $\mathcal{P}\cap L_1(W) = 21-9 = 12$,

11. If X, Y are orthogonal and anistropic then $X+Y$ is non-degenerate as is $X^\perp\cap Y^\perp$. A non degenerate two dimensional subspace contains 3 isotropic one dimensional subspaces and two anistropic one dimensional subspaces. Moreover, the two anistropic one dimensional subspaces of a non degenerate two dimensional subspace are orthogonal.

12. $|\mathcal{P}| = 85-45 = 40$. Suppose $X\in\mathcal{P}$. As we have seen in Exercise 10 there are 12 elements $Y\in\mathcal{P}$ with $Y\perp X$. Then there are two $Z\in\mathcal{P}$ with $X\perp Z\perp Y$. Therefore there are $40\times 12\times 2\times 1$ four-tuples (X,Y,Z,W) from $\mathcal{P}$ such that they are mutually orthogonal. Since there are 4! permutations of $\{X,Y,Z,W\}$ the number of subsets of cardinality four of mutually orthogonal elements of $\mathcal{P}$ is $\frac{40\times 12\times 2\times 1}{4\times 3\times 2\times 1} = 40$.

13. Assume $i\neq j$ and let $\{i,j,k,m\} = \{1,2,3,4\}$. Then $\mathcal{P}\cap L_1(X_i^\perp\cap X_j^\perp) = \{X_k, X_m\}$. Consequently, if $Y\in\mathcal{P}\setminus l$ then Y is orthogonal to as most one of X_i. Note that $|\mathcal{P}\setminus l| = 36$. For each i there are 12 elements $Z\in\mathcal{P}$ such that $Z\perp X_i$. Three of these are X_j, X_k, X_m and so there are nine elements Z of $\mathcal{P}\setminus l$ such that $Z\perp X_i$. This accounts for 9×4 elements of $\mathcal{P}\setminus l$, which is all of them.

Chapter 12

Additional Topics

12.1. Operator and Matrix Norms

1. Let $\| \quad \|$ be a norm on $\mathbb{F}^n$ and let $\boldsymbol{x}$ be a non-zero vector in $\mathbb{F}^n$. Then $I_n\boldsymbol{x} = \boldsymbol{x}$ and $\| I_n\boldsymbol{x} \| = \| \boldsymbol{x} \|$. Consequently, for every non-zero vector, $\boldsymbol{x}$, $\frac{\|I_n\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = 1$. Therefore, $\| I_n \|' = 1$.

2. $\| I_n \|_F = Trace(I_n^{tr} I_n)^{\frac{1}{2}} = n^{\frac{1}{2}} = \sqrt{n}$. By Exercise 1, $\| \quad \|_F$ is not induced by any norm on $\mathbb{F}^n$.

3. $\| A \|_F = (12^2 + 7^2 + 2^2)^{\frac{1}{2}} = \sqrt{193}, \| A \|_{1,1} = max\{14, 7\} = 14, \| A \|_{\infty,\infty} = max\{19, 2\} = 19.$

$$A^{tr}A = \begin{pmatrix} 12 & 2 \\ 7 & 0 \end{pmatrix}^{tr} \begin{pmatrix} 12 & 2 \\ 7 & 0 \end{pmatrix} =$$

$$\begin{pmatrix} 12 & 7 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 12 & 2 \\ 7 & 0 \end{pmatrix} =$$

$$\begin{pmatrix} 193 & 24 \\ 24 & 4 \end{pmatrix}.$$

The characteristic polynomial of $A^{tr}A$ is $X^2 - 197X + 196 = (X-1)(X-196)$. The eigenvalues are 1 and 196. Thus, $\rho(A^{tr}A) = 196$ and $\| A \|_{2,2} = \sqrt{196} = 14$.

4 $A^{tr}A = A^2 = \begin{pmatrix} 11 & 7 & 7 \\ 7 & 11 & 7 \\ 7 & 7 & 11 \end{pmatrix}$. $Trace(A^{tr}A) = 33$.
Therefore $\| A \|_F = \sqrt{33}$.

$\| A \|_{1,1} = \| A \|_{\infty,\infty} = 5.$

The eigenvalues of $A^{tr}A$ are 4 with multiplicity two and 25 with multiplicity 1. Then $\rho(A^{tr}A) = 25$ and $\| A \|_{2,2} = \sqrt{25} = 5$.

5. By Lemma (12.1) there is a positive real number M such that $\| T(\boldsymbol{x}) \|_W \leq M \| \boldsymbol{x} \|_V$. Suppose now that $\boldsymbol{x}_0 \in V, T(\boldsymbol{x}_0) = \boldsymbol{y}_0$ and ϵ is a positive real number. We have to show that there exists $\delta > 0$ such that if $\| \boldsymbol{x} - \boldsymbol{x}_0 \|_V < \delta$ then $\| T(\boldsymbol{x}) - T(\boldsymbol{x}_0) \|_W < \epsilon$. Let $\gamma = max\{M, 1\}$ and set $\delta = \frac{\epsilon}{\gamma}$. Now suppose $\| \boldsymbol{x} - \boldsymbol{x}_0 \|_V < \delta$ then $\| T(\boldsymbol{x}) - T(\boldsymbol{x}_0) \|_W = \| T(\boldsymbol{x} - \boldsymbol{x}_0) \|_W < M \| \boldsymbol{x} - \boldsymbol{x}_0 \|_V < \epsilon M \leq M$.

6. Since $I_n^2 = I_n$ we have $\| I_n \| = \| I_n \cdot I_n \| \leq$

$$\| I_n \| \cdot \| I_n \|$$

from which we conclude that $\| I_n \| \geq 1$.

12.2. Moore-Penrose Inverse

1. Since $\mu_P(x) = x^2 - x$ it follows that $P^2 = P$, whence $P^3 = P$. Thus, if $X = P$ then (PI1) and (PI2) hold. Since P is Hermitian, $P^* = P$ and therefore $(P^2)^* = P^2$. It follows if $X = P$ then (PI3) and (PI4) hold so $P^\dagger = P$.

2. If $X = diag\{\frac{1}{d_1}, \dots, \frac{1}{d_r}, 0, \dots, 0\}$ then $DXD = D$ and $XDX = X$ so (PI1) and (PI2) hold. Since $AX = XA = diag\{1, \dots, 1, 0, \dots, 0\}$ (rank r), $(AX)^* =$

$AX, (XA)^* = XA$ so also (PI3) and (PI4) hold. Thus, $X = diag\{\frac{1}{d_1}, \ldots, \frac{1}{d_r}, 0 \ldots, 0\}$ is the Moore-Penrose inverse of D.

3. Set $X = \frac{1}{\|\boldsymbol{v}\|}(\overline{a_1}, \ldots, \overline{a_n})$. Then $X\boldsymbol{v} = 1$ and $\boldsymbol{v}X\boldsymbol{v} = \boldsymbol{v}, X\boldsymbol{v}X = X$ so (PI1) and (PI2) hold. Since $X\boldsymbol{v}$ is a real scalar, $(X\boldsymbol{v})^* = X\boldsymbol{v}$ and (PI4) holds. On the other hand, $\boldsymbol{v}X$ is the $n \times n$ matrix whose (i,j)-entry is $\frac{1}{\|\boldsymbol{v}\|}a_i\overline{a_j}$ which is a Hermitian matrix: $(\boldsymbol{v}X)^* = \boldsymbol{v}X$ so (PI3) holds.

4. $AA^{-1}A = A$ so (PI1) holds. $A^{-1}AA^{-1} = A^{-1}$ so (PI2) holds. Since $AA^{-1} = A^{-1} = I_n = I_n^*$, (PI3) and (PI4) hold. Thus, $A^\dagger = A^{-1}$.

5. Set $X = C^*(CC^*)^{-1}$. Then $CX = I_r$ so that $XCX = X$ and $CXC = C$ so that (PI1) and (PI2) hold. Since $CX = I_r$, clearly (PI4) holds. On the other hand, $XC = C^*(CC^*)^{-1}C$ and $(XC)^* = C^*([CC^*])^{-1})^*(C^*)^* = C^*(CC^*)^{-1}C = XC$.

6. $P^2 = (AA^\dagger)^2 = AA^\dagger AA^\dagger$. By (PI1), $AA^\dagger A = A$ and therefore $(AA^\dagger)^2 = AA^\dagger = P$. By (PI3), $P^* = P$ so that $(P^2)^* = P^2$.

7. $(I_m - P)^2 = (I_m - P)I_m - (I_m - P)P = (I_m - P) + (P^2 - P) = I_m - P$. Also, $(I_M - P)^* = I_m^] * - P^* = I_m - P$.

8. Set $X = (A^\dagger)^*$. We want to show that $X = (A^*)^\dagger$.
$A^*XA^* = A^*(A^\dagger)^*A^* = (AA^\dagger A)^* = A^*$ so (PI1) holds.

$XA^*X = (A^\dagger)^*A^*(A^\dagger)^* = (A^\dagger AA^\dagger)^* = (A^\dagger)^* = X$. Thus, (PI2) holds.

$(A^*X)^* = [A^*(A^\dagger)^*]^* = A^\dagger A$. However, we then have $A^*X = [(AA^*)^*]^* = (A^\dagger A) = A^\dagger A$ by (PI4) for A. Thus, (PI3) holds for X relative to A^*.

$(XA^*)^* = AX^* = A[(A^\dagger)^*]^* = AA^\dagger$. Since $AA^\dagger$ is Hermitian it follows that XA^* is Hermitian and (PI4) holds.

9. Set $B = A^*A$ and $X = A^\dagger(A^*)^\dagger = A^\dagger(A^\dagger)^*$ by Exercise 7. We show that X is the Moore-Penrose inverse of B.

$$BXB = (A^*A)[A^\dagger(A^\dagger)^*](A^*A) =$$
$$(A^*A)[A^\dagger\{(A^\dagger)^*A^*\}A =$$
$$(A^*A)[A^\dagger(AA^\dagger)^*\}A = (A^*A)[A^\dagger(AA^\dagger]A =$$
$$(A^*AA^\dagger)(AA^\dagger A) = (A^*)(AA^\dagger A) = A^*A = B.$$

Thus, (PI1) holds.

$$XBX = [A^\dagger(A^\dagger)^*](A^*A)[A^\dagger(A^\dagger)^*] =$$
$$A^\dagger(AA^\dagger)^*AA^\dagger(A^*)^\dagger =$$
$$A^\dagger AA^\dagger(A^*)^\dagger = A^\dagger(A^*)^\dagger = X.$$

Consequently, (PI2) holds.

$$(BX)^* = [(A^*A)(A^\dagger\{A^\dagger\}^*]^* =$$
$$[A^*(AA^\dagger)\{A^\dagger\}^*]^* =$$
$$(\{A^\dagger\}^*)^*(AA^\dagger)^*(A^*)^* =$$
$$A^\dagger AA^\dagger A = A^\dagger A.$$

Since $A^\dagger A$ is Hermitian it follows that $(BX)^*$ is Hermitian, whence, BX is Hermitian and (PI3) holds.

$$(XB)^* = \{[A^\dagger(A^\dagger)^*][A^*A]\}^* =$$
$$\{A^\dagger[(A^\dagger)^*A^*]A\}^* =$$
$$\{A^\dagger([AA^\dagger)^*]A\}^* =$$
$$[A^\dagger(AA^\dagger)A]^* = A^\dagger A$$

which is Hermitian and (PI4) holds.

10. Since $A = AA^\dagger A = A(A^\dagger A)$ by (PI4) we have $A^* = [A(A^\dagger A)]^* = (A^\dagger A)^*A^* = (A^\dagger A)A^*$. On the other hand, writing $A = (AA^\dagger)A$ and using (PI3) we get $A^* = [(AA^\dagger)A]^* = A^*(AA^\dagger)^* = A^*(A^\dagger A)$.

11. $(A^*A)^\dagger A^* = [A^\dagger(A^*)^\dagger]A^*$ by Exercise 8. By Exercise 7 we have

$$A^\dagger(A^*)^\dagger]A^* = [A^\dagger(A^\dagger)^*]A^* =$$

$A^\dagger(AA^\dagger A)^* = A^\dagger(AA^\dagger)$ by (PI3). By (PI2) this is equal to $A^\dagger$ as desired.

12. $A^\dagger A = (A^*A)^\dagger(A^*A)$ by Exercise 11. Now A^*A is Hermitian and therefore so is $(A^*A)^\dagger$ by Exercise 8. Moreover, $A^\dagger A$ is Hermitian by (PI4). It follows that

$$A^\dagger A = (A^\dagger A)^* = \{[(A^*A)^\dagger](A^*A)\}^* =$$

$$(A^*A)^*[(A^*A)^\dagger]^* = (A^*A)(A^*A)^\dagger.$$

Since $A^*A = AA^*$ we have

$$(A^*A)(A^*A)^\dagger. = (AA^*)(AA^*)^\dagger =$$

$$A[A^*(AA^*)^\dagger] = AA^\dagger$$

by Exercise 11.

13. The proof is by induction on n. The base case is trivial. Assume that $(A^n)^\dagger = (A^\dagger)^n$. We must show that $(A^{n+1})^\dagger = (A^\dagger)^{n+1}$. We show that for $X = (A^\dagger)^{n+1}$ that (PI1) = (PI4) hold. We make use of Exercise 12: $A^\dagger A = AA^\dagger$. As a consequence we have

$$A^{n+1}(A^\dagger)^{n+1}A^{n+1} = (AA^\dagger A)(A^n(A^\dagger)^nA^n).$$

By the inductive hypothesis, $A^n(A^\dagger)^nA^n = A^n$. Furthermore, $AA^\dagger A = A$. Therefore

$$(AA^\dagger A)(A^n(A^\dagger)^nA^n) = AA^n = A^{n+1}.$$

Thus, (PI1) holds.

$$(A^\dagger)^{n+1}A^{n+1}(A^\dagger)^{n+1} = (A^\dagger AA^\dagger)[(A^\dagger)^nA^n(A^\dagger)^n] =$$

$$A^\dagger(A^\dagger)^n = (A^\dagger)^{n+1}.$$

Thus, (PI2) holds.

$$(A^\dagger)^{n+1}A^{n+1} = (A^\dagger A)[(A^\dagger)^nA^n] =$$

$$(A^\dagger A)(A^\dagger A)^n = (A^\dagger A)^{n+1}.$$

Since $A^\dagger A$ is Hermitian it follows that $(A^\dagger A)^{n+1}$ is Hermitian. This proves (PI3). In a similar fashion we have

$$A^{n+1}(A^\dagger)^{n+1} = [AA^\dagger]^{n+1}.$$

Since $AA^\dagger$ is Hermitian so is $(AA^\dagger)^{n+1}$ and (PI4) holds.

14. With $B = \lambda A$ and $X = \frac{1}{\lambda}A^\dagger$ we show (PI1) - (PI4) hold.

$$BXB = (\lambda A)(\frac{1}{\lambda}A^\dagger)(\lambda A) =$$

$$\lambda(AA^\dagger A) = \lambda A = B.$$

$$XBX = (\frac{1}{\lambda}A^\dagger)(\lambda A)(\frac{1}{\lambda}A^\dagger) =$$

$$\frac{1}{\lambda}(A^\dagger AA^\dagger) = \frac{1}{\lambda}A^\dagger = X.$$

$$BX = (\lambda A)(\frac{1}{\lambda}A^\dagger) = AA^\dagger = (AA^\dagger)^*.$$

$$XB = (\frac{1}{\lambda}A^\dagger)\lambda A) = A^\dagger A = (A^\dagger A)^*.$$

15. Assume $A^* = A^\dagger$. Then $(A^*A)^2 = (A^\dagger A)^2 = A^\dagger A = A^*A$ by Exercise 6. Converely, assume that $(A^*A)^2 = A^*A$. Then it is straightforward to see that $(A^*A)^\dagger = A^*A$. From Exercise 11 it follows that $A^\dagger = A^*AA^*$. It then follows that $A = A(A^*AA^*)A = A(A^*A)^2 = AA^*A$. Thus, (PI1) holds with $X = A^*$. Since $A = AA^*A$ taking adjoints we get $A^* = AA^*A$ so (PI2) holds as well. Since both AA^* and A^*A are Hermitian (PI3) and (PI4) are satisfied. Therefore, in fact, $A^\dagger = A^*$.

12.3. Nonnegative Matrices

1. Assume $a^k_{ij} \neq 0$ and $a^m_{jl} \neq 0$. The (i,l)-entry of $A^{k+m} = A^k A^m$ is $\sum_{p=1}^n a^k_{ip} a^m_{pl} \geq a^k_{ij} a^m_{jl} > 0$.

2. Assume $j \in \Delta(i)$ and $l \in \Delta(j)$. Then $(i,j),(j,l) \in \Delta$. By Exercise 1, $(i,l) \in \Delta$ so that $l \in \Delta(i)$.

3. Let $l \in I$. Suppose $a_{ml} \neq 0$ so that $m \in \Delta(l)$. Then by Exercise 2, $m \in \Delta(i) = I$. Consequently, $A\boldsymbol{e}_l \in Span(\boldsymbol{e}_j | j \in I)$ as was to be shown.

4. Let $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$ and $B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \dots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$ and assume that $\{(i,j) | a_{ij} \neq 0\} = \{(i,j) | b_{ij} \neq 0\}$. Then $a^k_{ij} \neq 0$ if and only if $b^k_{ij} \neq 0$. The result follows from this.

5. $(I_n + A)^{n-1} = \sum_{j=0}^{n-1} \binom{n-1}{j} I^j_n A^{n-1-j}$. Assume $i \neq j$. Since $(I_n + A)^{n-1} > 0$ for some $j, 1 \leq j \leq n-1$ the (i,j)-entry of A^j is positive. Fix $i, 1 \leq i \leq n$ and let $1 \leq l \leq n, l \neq i$. Then for some $j, a^j_{il} > 0$ and for some $k, a^k_{ji} > 0$. It then follows that $a^{j+k}_{ii} > 0$. Thus, A is irreducible.

6. Let the entries of A be a_{ij}. It suffices to prove if A is a positive $m \times n$ matrix, $\boldsymbol{x}$ is a non-zero nonnegative matrix then $A\boldsymbol{x} \neq \boldsymbol{0}$. Let $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and assume that $x_j > 0$. The i^{th} entry of $A\boldsymbol{x}$ is $\sum_{k=1}^n a_{ik} x_k > a_{ij} x_j > 0$.

7. If $\rho(A) = 0$ then all the eigenvalues of A are zero and A is a nilpotent matrix, whence $A^n = \boldsymbol{0}_{nn}$. However, if $A^k > 0$ then $(A^k)^n > 0$ which contradicts $(A^k)^n = (A^n)^k = \boldsymbol{0}_{nn}$.

8. Assume $\boldsymbol{x} > 0$ is an eigenvector for A, say $A\boldsymbol{x} = \lambda\boldsymbol{x}$. Since A is non-negative and $\boldsymbol{x} > 0, A\boldsymbol{x} > 0$. In particular, $\lambda > 0$. Thus, $\rho(A) > 0$.

9. Assume $A\boldsymbol{d} = \lambda\boldsymbol{d}$. Let a_{ij} be the (i,j)-entry of A. Since $\boldsymbol{d}$ is an eigenvector with eigenvalue $\rho(A)$ we have for each i,

$$\sum_{j=1}^n a_{ij} d_j = \rho(A) d_i.$$

Now the (i,j)-entry of $D^{-1}AD$ is $\frac{d_j}{d_i} a_{ij}$. Then

$$\sum_{j=1}^n b_{ij} = \sum_{j=1}^n \frac{d_j}{d_i} a_{ij} =$$

$$\frac{1}{d_i} \sum_{j=1}^n a_{ij} d_j =$$

$$\frac{1}{d_i}(\lambda d_i) = \lambda).$$

It follows that the matrix $\frac{1}{\lambda} D^{-1}AD$ is a row stochastic matrix. By Theorem (12.22) it follows that $\rho(\frac{1}{\lambda} D^{-1}AD) = 1$. However, $\rho(\frac{1}{\lambda} D^{-1}AD) = \frac{\rho(D^{-1}AD)}{\lambda)} = \frac{\rho(A)}{\lambda}$. Thus, $\lambda = \rho(A)$.

10. By Theorem (12.19) there are positive vectors $\boldsymbol{x}, \boldsymbol{y}$ such that $\| \boldsymbol{x} \|_1 = 1 = \boldsymbol{y}^{tr}\boldsymbol{x}, A\boldsymbol{x} = \rho\boldsymbol{x}, A^{tr}\boldsymbol{y} = \rho\boldsymbol{y}$. By Theorem (12.21), $\lim_{k\to\infty} [\frac{1}{\rho} A]^k = \boldsymbol{x}\boldsymbol{y}^{tr}$, a rank one positive matrix. It then follows for some natural number k that $A^k > 0$.

11. Assume $\omega \in \mathbb{C}, |\omega| = 1$ and $\omega z_i = |z_i|$ for all i. Then $|z_1 + \dots + z_n| = |\omega||z_1 + \dot{+} z_n| = |\omega(z_1) + \dots + z_n)| = |\omega z_1 + \dots + \omega z_n| = ||z_1| + \dots + |z_n|| = |z_1| + \dots |z_n|$.

We prove the converse by induction on n. Assume $n = 2$. Let $\omega \in \mathbb{C}, |\omega| = 1$, such that $\omega z_1 = a \in \mathbb{R}^+$. We need to prove that $\omega z_2 = |z_2|$, equivalently, $\omega z_2 \in \mathbb{R}^+$. Assume $\omega z_2 = b + ci$ where $b, c \in \mathbb{R}$. Then $|z_1 + z_2|^2 = |(a+b) + ci|^2 = (a+b)^2 + c^2 = a^2 + 2ab + b^2 + c^2$. On the other hand, $(|z_1| + |z_2|)^2 = [a + \sqrt{b^2 + c^2}]^2 = a^2 + b^2 + c^2 + 2a\sqrt{b^2 + c^2}$. If $c \neq 0$ or $b < 0$ then $2ab \neq 2a\sqrt{b^2 + c^2}$. Therefore $c = 0$ and $b > 0$.

Now assume that $n > 2$ and the result is true for $n-1$ complex numbers $z_1, \dots, z_{n-1}$. Assume $\omega \in \mathbb{C}, |\omega| = 1$ so that $\omega z_1 \in \mathbb{R}^+$. Replacing z_i with ωz_i, if necessary, we may assume that $z_1 \in \mathbb{R}^+$ and we need to prove that $z_i \in \mathbb{R}^+$ for every i. Suppose $|z_2+\cdots+z_n| < |z_2|+\cdots+|z_n|$. Then $|z_1+\cdots+z_n| \le |z_1|+|z_2+\cdots+z_n| < |z_1|+\cdots+|z_n|$, a contradiction. Therefore $|z_1+(z_2+\cdots+z_n)| = |z_1|+|z_2+\cdots+z_n| = |z_1|+\ldots|z_n|$. It follows by the case for $n=2$ that $z_2+\cdots+z_n \in \mathbb{R}^+$. By the induction hypothesis, there is $\gamma \in \mathbb{C}, |\gamma|=1$ such that $\gamma z_i = |z_i|$ for $2 \le i \le n$. Then $\gamma(z_2+\cdots+z_n) = \gamma z_2+\cdots+\gamma z_n \in \mathbb{R}^+$ which implies that $\gamma = 1$.

12. Assume first that $A > 0$ and that $A\boldsymbol{x} = \lambda\boldsymbol{x}$ and $|\lambda| = \rho(A)$. Then there is a $\omega \in \mathbb{C}, |\omega| = 1$ such that $\omega\boldsymbol{x} = |\boldsymbol{x}| > 0$. Since $A\boldsymbol{x} = \lambda\boldsymbol{x}$ we have $|A\boldsymbol{x}| = |\lambda\boldsymbol{x}| = |\lambda||\boldsymbol{x}| = \rho(A)|\boldsymbol{x}|$. It then follows that $A|\boldsymbol{x}| = \rho(A)\boldsymbol{x}$ to that $|\boldsymbol{x}|$ is a positive vector. Now assume A is nonnegative and primitive, $\lambda \in Spec(A), \lambda \ne \rho(A)$ and that $A\boldsymbol{x} = \lambda\boldsymbol{x}$. By Exercise 10 there is a k such that $A^k > 0$. Then $A^k\boldsymbol{x} = \lambda^k\boldsymbol{x}$ and so $\lambda^k \in Spec(A^k)$. Moreover, $\rho(A^k) = \rho(A)^k$. By the case just proved, either $\lambda^k = \rho(A)^k$ of $|\lambda|^k < |\rho(A)|^k$ so that $|\lambda| < |\rho(A)|$. However, in the first case the geometric multiplicity of $\rho(A)^k$ is greater than one which contradicts Theorem (12.19).

13. Let $\boldsymbol{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$. Then $\langle \boldsymbol{p}, \boldsymbol{j}_n \rangle = \sum_{j=1}^n p_n$. Thus, if $\boldsymbol{p} \ge 0$ then $\boldsymbol{p}$ is a probability vector if and only if $\langle \boldsymbol{p}, \boldsymbol{j}_n \rangle = 1$.

14. Assume $\boldsymbol{p}_j = \begin{pmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{pmatrix}$. Since $(s_1, \ldots, s_t)$ is a sequence of nonnegative numbers, $s_j p_{ij} \ge 0$ for every i and j. Then $\sum_{j=1}^t p_{ij}s_j \ge 0$ so that $s_1\boldsymbol{p}_1 + \cdots + s_t\boldsymbol{p}_t$ is a nonnegative vector. Now

$$\sum_{i=1}^n \sum_{j=1}^t s_j p_{ij} = \sum_{j=1}^t \sum_{i=1}^n s_j p_{ij} =$$
$$\sum_{j=1} s_j \sum_{i=1}^n p_{ij} = \sum_{j=1} s_j = 1.$$

15. By Lemma (12.7) (proved in Exercise 16) every column of $s_1A_1 + \cdots + s_tA_t$ is a probability vector and therefore $s_1A_1 + \cdots + s_tA_t$ is column stochastic. If the matrices are bi-stochastic then the argument applies to $A_1^{tr}, \ldots, A_t^{tr}$.

16. If the columns of A are $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$ and $\boldsymbol{p} = \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$ then $A\boldsymbol{p} = s_1\boldsymbol{c}_1 + \cdots + s_n\boldsymbol{c}_n$. Now the result follows by Lemma (12.7), Exercise 16.

17. If follows from Exercise 18 and the definition of matrix multiplication that the product of two column stochastic matrices is stochastic. If A is stochastic then by induction it follows that A^k is stochastic.

18. We first prove if $\boldsymbol{p}, \boldsymbol{q}$ are probability vectors and $\langle \boldsymbol{p}, \boldsymbol{q} \rangle = 1$ then $\boldsymbol{p} = \boldsymbol{q} = \boldsymbol{e}_i$ for some i. Assume to the contrary that $\boldsymbol{p} \ne \boldsymbol{e}_i$ for all i. Then there exists j such that $0 < p_j < 1$. Then $p_jq_j < q_j$ from which we conclude that $\sum_{j=1}^n p_jq_j < \sum_{j=1} q_j = 1$, a contradiction. Let the columns of A^{tr} be $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ and the columns of A^{-1} be $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. Since $AA^{-1} = I_n$ it follows that $\langle \boldsymbol{a}_i, \boldsymbol{b}_i \rangle = 1$ so that $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ and A is a permutation matrix.

19. Assume A has more than n entries and therefore some row has at least two entries. Without loss of generality we can assume the first row has two entries, which we point out are both less than one. Assume they are in columns i and j. Since $a_{1i} < 1$ there must be an s such that $a_{si} \ne 0$. Since $a_{1j} < 1$ there must be a t such that $a_{tj} \ne 0$. Now there are two columns with at least 2 non-zero entries. Since each column sums to 1 there must be at least one nonzero entry in each column and we have shown this matrix must have at least $n+2$ nonzero entries.

20. Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Since the matrix is bistochastic, $a+b = a+c$ so that $b=c$ and the matrix is symmetric. Also $b+d = 1 = a+b$ so $a=d$.

21. Since A is reducible it is permutation similar to a block matrix of the form $\begin{pmatrix} A_1 & B_{st} \\ \boldsymbol{0}_{ts} & A_2 \end{pmatrix}$ where A_1 is an $s \times s$ matrix and A_2 is a $t \times t$ matrix with $s+t=n$.

The matrix A_1 is column stochastic and the matrix A_2 is row stochastic. Now the sum of each column of A_1 is n and there are s columns so the sum of all the entries in A_1 is ns. Since A is row stochastic, each row of A sums to n. Let the rows of A_1 be $\boldsymbol{a}_{12}, \ldots, \boldsymbol{a}_{1s}$. The sum of the entries in each $\boldsymbol{a}_{1i}$ is at most n. Suppose some row does not sum to n. Then the sum of all the entries in all the rows of A_1 sums to less than ns, a contradiction. Thus, each row of A_1 sums to n and A_1 is a bistochastic matrix. Consequently, $B_{st} = \mathbf{0}_{st}$ and then A_2 is bistochastic.

12.4. Location of Eigenvalues

1. Since A is a stochastic matrix, in particular, A is real matrix. Since A is stochastic, A^{tr} is row stochastic. If $\delta = min\{a_{ii} | 1 \leq i \leq n\}$ then $C'_i(A) = R'_i(A^{tr}) \leq 1-\delta$ for all i. Suppose $z \in \mathbb{C}$ and $|z-\delta| \leq C'_i(A) \leq 1-\delta$. On the other hand, $|z - a_{ii}| \leq |z-\delta|$.

2. This implies that A is strictly diagonally dominant, consequently, A is invertible.

3. Since $\Gamma_i(A) \cap \Gamma_j(A) = \emptyset$ for all $i \neq j$ it follows from Theorem (12.26) the eigenvalues of A are distinct and it follows that A is diagonalizable.

4. By Theorem (12.26) each $\Gamma_i(A)$ contains an eigenvalue, whence, no $\Gamma_i(A)$ contains two. Now assume $a \in \mathbb{R}$ and $w, \overline{w}$ are complex conjugates. Then $|a-w| = |a-\overline{w}|$. If A has a complex eigenvalue, w, then also $\overline{w}$ is an eigenvalue since the characteristic polynomial is real. Suppose $w \in \Gamma_i(A)$ so that $|w - a_{ii}| \leq R'_i(A)$. But then $|\overline{w} - a_{ii}| \leq R'_i(A)$ and $\overline{w} \in \Gamma_i(A)$, a contradiction. Thus, the eigenvalues of A are all real.

5. Since each of the matrices $Q^{-1}AQ$ is similar to A we have $Spec(Q^{-1}AQ) = Spec(A)$. Therefore $Spec(A) = Spec(Q^{-1}AQ) \subset \Gamma(Q^{-1}AQ)$ and, consequently, contained in the intersection of all $\Gamma(Q^{-1}AQ)$.

6. This is proved just like Exercise 4.

7. Denote the columns of A by $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n$. Let $I = \{i_1 < \cdots < i_k\}$. We prove that the sequence $(\boldsymbol{c}_{i_1}, \ldots, \boldsymbol{c}_{i_k})$ is linearly independent. Let $A_{I,I}$ be the $k \times k$ matrix whose (s,t)-entry is a_{i_s, i_t}. Then $A_{I,I}$ is strictly diagonally dominant and therefore by Theorem (12.32) invertible. This implies that the sequence of columns of $A_{I,I}$ is linear independent which, in turn implies that the sequence $(\boldsymbol{c}_{i_1}, \ldots, \boldsymbol{c}_{i_k})$ is linearly independent. Thus, $rank(A) \geq k$.

8. Assume to the contrary that for all i, $|a_{ii}| \leq C'_i(A)$, we will obtain a contradiction. Since A is strictly diagonally dominant it follows that $\sum_{i=1}^n |a_{ii}| > \sum_{i=1}^n R'_i(A) = \sum_{i \neq j} |a_{ij}| = \sum_{i=1}^n C'_i(A) \geq \sum_{i=1}^n |a_{ii}|$, a contradiction.

9. Since multiplying the i^{th} row by -1 changes both the sign of $det(A)$ and a_{ii} we can assume that all $a_{ii} > 0$ and prove that $det(A) > 0$. The proof is by induction on $n \geq 2$. If $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ with $a_{11} > |a_{12}|$ and $a_{22} > |a_{21}|$ then $det(A) = a_{11}a_{22} - a_{21}a_{12} \geq a_{11}a_{22} - |a_{12}||a_{21}| > 0$.

Now assume the result is true for all strictly diagonally dominant $(n-1) \times (n-1)$ real matrices with positive diagonal entries and assume that A is an $n \times n$ strictly diagonally dominant real matrix with positive diagonal entries. Let a_{ij} be the (i,j)-entry of A. We add multiples of the first column to the other columns in order to obtain zeros in the first row. This does not change the determinant. After performing these operations the entry in the (i,j)-entry with $1 < i,j$ is $b_{ij} = a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}}$. We claim that the $(n-1) \times (n-1)$ B matrix with (k,l)-entry equal to $b_{k+1,l+1}$ is strictly diagonally dominant with positive diagonal entries. Now the diagonal entries are

$$a_{ii} - \frac{a_{1i}}{a_{11}} a_{i1} \geq a_{ii} - \left|\frac{a_{1i}}{a_{11}} a_{i1}\right| > a_{ii} - |a_{i1}| > 0$$

which establishes our second claim. We now must show that

$$a_{ii} - \frac{a_{1i}a_{i1}}{a_{11}} > \sum_{j \geq 2, j \neq i} \left|a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}}\right|.$$

We illustrate with $i = 2$, the other cases follow in exactly the same way. We thus have to show that

$$a_{22} - \frac{a_{12}a_{21}}{a_{11}} > \sum_{j=3}^{n} |a_{2j} - \frac{a_{21}a_{1j}}{a_{11}}| \leq$$

$$\sum_{j=3}^{n} |a_{2j}| + \sum_{j=3}^{n} |\frac{a_{21}a_{1j}}{a_{11}}|.$$

It therefore suffices to show that

$$a_{22} > \sum_{j=3}^{n} |a_{2j}| + \sum_{j=3}^{n} |\frac{a_{21}a_{1j}}{a_{11}}| + |\frac{a_{12}a_{21}}{a_{11}}.$$

Since A is strictly diagonally dominant it suffices to prove that $\sum_{j=2}^{n} |\frac{a_{21}a_{1j}}{a_{11}}| < |a_{21}|$. Now

$$\sum_{j=2}^{n} |\frac{a_{21}a_{1j}}{a_{11}}| =$$

$$|a_{21}| \times \frac{1}{a_{11}} \sum_{j=2}^{n} |a_{1j}| < |a_{21}|$$

since $a_{11} > \sum_{j=2}^{n} |a_{1j}|$.

Now $det(A) = a_{11}det(B)$. Since B is strictly diagonally dominant with positive diagonal entries, $det(B) > 0$ by the inductive hypothesis. Hence $det(A) > 0$.

12.5. Functions of Matrices

1i. For any $n \times n$ matrices A, B, $(Q^{-1}AQ)(Q^{-1}BQ) = Q^{-1}[AB]Q$. Then result follows by a straightforward induction on k.

ii. $[Q^{-1}(M_1 + M_2)]Q = [Q^{-1}M_1 + Q^{-1}M_2]Q = [Q^{-1}M_1]Q + [Q^{1}M_2]Q = Q^{-1}M_1Q + Q^{-1}M_2Q$.

2. Assume $f(x) = a_0 + a_1x + \cdots + a_mx^m$. Now $f(B) = a_0I_n + a_1B + \cdots + a_mB^m$. Then $Q^{-1}f(B)Q = Q^{-1}[a_0I_n + a_1B + \cdots + a_mB^m]Q$. By repeated application of Lemma (12.9) ii. it follows that $Q^{-1}f(B)Q = Q^{-1}(a_0I_n)Q + Q^{-1}(a_1B)Q + \cdots + Q^{-1}(a_mB^m)Q = a_0I_n + a_1Q^{-1}BQ + \cdots + a_mQ^{-1}B^mQ$. By (12.9) i, this is equal to $a_0I_m + a_1Q^{-1}BQ + \cdots + a_m(Q^{-1}BQ)^m = f(Q^{-1}BQ)$.

3. There exists an invertible matrix Q such that $Q^{-1}AQ$ is upper triangular. Since $exp(Q^{-1}AQ) = Q^{-1}exp(A)Q$ we can, without loss of generality, assume that A is upper triangular. Now if $A = \begin{pmatrix} \lambda_1 & a_{12} \dots & a_{1n} & \\ 0 & \lambda_2 & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$ then $exp(A)$ is upper triangular with $e^{\lambda_1}, \dots, e^{\lambda_n}$ on the diagonal. Consequently,

$$\chi_{exp(A)} = (x - e^{\lambda_1}) \dots (x - exp(\lambda_n)).$$

4. It follows from Theorem (12.36) and the definition of determinant that $det(exp(A)) = e^{\lambda_1} \dots e^{\lambda_n} = e^{\lambda_1 + \cdots + \lambda_n} = e^{Trace(A)}$.

5. For an $n \times n$ complex matrix we have $(A^n)^* = (A^*)^n$. We also have $(A+B)^* = A^* + B^*$. Also for a real scalar $c, (cA)^* = cA^*$. It then follows that if $f(x) \in \mathbb{R}[x]$ and A is an $n \times n$ complex matrix then $f(A)^* = f(A^*)$.

Set $f_n(x) = \sum_{j=0}^{n} \frac{1}{j!}x^j$. Then $f_n(A)^* = f_n(A^*)$. It then follows that

$$exp(A^*) = \lim_{n\to\infty} f_n(A^*) = \lim_{n\to\infty} f_n(A)^* = exp(A)^*.$$

6. By Exercise 5, $exp(A)^* = exp(A^*) = exp(A)$.

7. In general, if $AB = BA$ then $exp(A)exp(B) = exp(A + B) = exp(B + A) = exp(B)exp(A)$. In particular, $exp(A)exp(A)^* = exp(A)exp(A^*) = exp(A + A^*) = exp(A^* + A) = exp(A^*)exp(A) = exp(A)^*exp(A)$.

Chapter 13

Applications of Linear Algebra

13.1. Least Squares

1. By Lemma (13.1), if X is a $\{1,3\}$ inverse of A then $AX = AA^\dagger$. Conversely, suppose $AX = AA^\dagger$. Then $AXA = AA^\dagger A = A$ by (PI1) for the Moore-Penrose inverse. Also, $(AX)^* = (AA^\dagger)^* = AA^\dagger = AX$.

2. Since $\boldsymbol{z} = X\boldsymbol{b}$ where X is a $\{1,3\}$ inverse of A we have $A\boldsymbol{z} = AX\boldsymbol{b} = AA^\dagger\boldsymbol{b}$ which is a least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$. Therefore, if $A\boldsymbol{y} = A\boldsymbol{z}$ then $\boldsymbol{y}$ is a least square solution to $A\boldsymbol{x} = \boldsymbol{b}$.

3. Since $A = BC, A^* = C^*B^*$. Substituting this for A^* in the normal equation $A^*A\boldsymbol{x} = A^*\boldsymbol{b}$ we get

$$C^*B^*A\boldsymbol{x} = C^*B^*\boldsymbol{b}.$$

It then follows that $C^*(B^*A\boldsymbol{x} - B^*\boldsymbol{b}) = \boldsymbol{0}_n$. However, since C^* is an $n \times r$ matrix with rank r it must be the case that $B^*A\boldsymbol{x} - B^*\boldsymbol{b} = \boldsymbol{0}_r$ which implies that $B^*A\boldsymbol{x} = B^*\boldsymbol{b}$.

4. Clearly the columns of A are not multiple of each other and consequently, the sequence is linearly independent. The reduced echelon form of the matrix $\begin{pmatrix} 1 & 1 & 9 \\ 1 & -3 & 3 \\ -2 & 2 & -6 \end{pmatrix}$ is I_3 and therefore $\boldsymbol{b} \notin col(A)$.

$A^* = \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 2 \end{pmatrix}$. The matrix version of the normal equations is

$$\begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -3 \\ -2 & 2 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 1 & 1 & -2 \\ 1 & -3 & 2 \end{pmatrix} \begin{pmatrix} 9 \\ 3 \\ -6 \end{pmatrix}.$$

$$\begin{pmatrix} 6 & -6 \\ -6 & 14 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 24 \\ 6 \end{pmatrix},$$

$$\boldsymbol{x} = \begin{pmatrix} \frac{31}{4} \\ \frac{15}{4} \end{pmatrix}.$$

5. Clearly the columns of A are not multiples of each other and consequently, the sequence is linearly independent. The reduced echelon form of the matrix $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & -2 \\ 1 & 3 & 7 \end{pmatrix}$ is I_3 so $\boldsymbol{b} \notin col(A)$.

$A^* = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \end{pmatrix}$ so the matrix version of the normal equations is

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 3 & 1 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix}.$$

$$\begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 6 \\ 21 \end{pmatrix},$$

$$\boldsymbol{x} = \begin{pmatrix} -7 \\ \frac{9}{2} \end{pmatrix}.$$

6. Clearly the columns of A are not multiple of each other and consequently, the sequence is linearly independent. The reduced echelon form of the matrix $\begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 1 & 3 \\ 1 & -1 & 18 \end{pmatrix}$ is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ so $\boldsymbol{b} \notin col(A)$.

$A^* = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$. The matrix version of the normal equations is

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \boldsymbol{x} =$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 3 \\ 18 \end{pmatrix}.$$

$$\begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 24 \\ -12 \end{pmatrix},$$

$$\boldsymbol{x} = \begin{pmatrix} 20 \\ -16 \end{pmatrix}.$$

7. The reduced echelon form of A is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ so the sequence of columns of A is linearly independent. The reduced echelon form of the matrix $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 4 \end{pmatrix}$ is I_4. Therefore $\boldsymbol{b} \notin col(A)$. The matrix version of the normal equations is

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \boldsymbol{x} =$$

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

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 4 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 6 \\ -2 \\ 8 \end{pmatrix},$$

$$\boldsymbol{x} = \begin{pmatrix} \frac{3}{2} \\ -\frac{5}{4} \\ \frac{5}{4} \end{pmatrix}.$$

8. The reduced echelon form of the matrix $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 1 & 2 & 0 & -1 \\ 2 & 1 & 3 & 0 \end{pmatrix}$ is $\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. It follows that the sequence of columns of A is linearly dependent and that $\boldsymbol{b} \notin col(A)$.

The matrix version the normal equations is

$$\begin{pmatrix} 6 & 4 & 8 \\ 4 & 6 & 2 \\ 8 & 2 & 14 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} -1 \\ -3 \\ 3 \end{pmatrix}.$$

Set $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \\ 2 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \end{pmatrix}$. Then $A = BC$ is full rank decomposition of A. Then $A^\dagger = C^\dagger B^\dagger$.

$$B^\dagger = \begin{pmatrix} \frac{3}{10} & -\frac{1}{5} & -\frac{1}{10} & \frac{2}{5} \\ -\frac{1}{5} & \frac{3}{10} & \frac{2}{5} & -\frac{1}{10} \end{pmatrix},$$

$C^\dagger = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{5}{3} \\ \frac{1}{3} & -\frac{1}{6} \end{pmatrix}$.

$$A^\dagger = \begin{pmatrix} \frac{1}{3} & \frac{1}{30} & \frac{1}{5} & \frac{1}{5} \\ -\frac{1}{15} & \frac{11}{60} & \frac{3}{20} & \frac{1}{20} \\ \frac{2}{15} & -\frac{7}{60} & -\frac{1}{10} & \frac{3}{20} \end{pmatrix}.$$

Set $\boldsymbol{z} = A^\dagger \boldsymbol{b} = \begin{pmatrix} -\frac{1}{10} \\ -\frac{11}{20} \\ \frac{7}{20} \end{pmatrix}$. The general least square solution is $\boldsymbol{z} + \boldsymbol{y}$ where $\boldsymbol{y} \in col(I_3 - A^\dagger A) = Span \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$.

9. The reduced echelon form of $(A\ \boldsymbol{b})$ is

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

Thus, $\boldsymbol{b} \notin col(A)$ and the sequence of columns of A is linearly dependent.

The matrix version of the normal equations is

$$\begin{pmatrix} 4 & -1 & 3 & 5 \\ -1 & 4 & 3 & -5 \\ 3 & 3 & 6 & 0 \\ 5 & -5 & 0 & 10 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 6 \\ 3 \\ 9 \\ 3 \end{pmatrix}.$$

Let $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{pmatrix}$. Then $A = BC$ is a full rank decomposition of A. Then $A^\dagger = C^\dagger B^\dagger$.

$C^\dagger = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} \end{pmatrix}$. $B^\dagger = \begin{pmatrix} \frac{4}{15} & \frac{1}{15} & \frac{1}{5} & \frac{1}{3} & -\frac{1}{5} \\ \frac{1}{15} & \frac{4}{15} & -\frac{1}{5} & \frac{1}{3} & -\frac{1}{5} \end{pmatrix}$.

$$A^\dagger = \begin{pmatrix} \frac{4}{45} & \frac{1}{45} & \frac{1}{15} & \frac{1}{9} & -\frac{1}{15} \\ \frac{1}{45} & \frac{4}{45} & -\frac{1}{15} & \frac{1}{9} & \frac{1}{15} \\ \frac{1}{9} & \frac{1}{9} & 0 & \frac{2}{9} & 0 \\ \frac{1}{15} & -\frac{1}{15} & \frac{2}{15} & 0 & -\frac{2}{15} \end{pmatrix}.$$

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

Set $\boldsymbol{z} = A^\dagger \boldsymbol{b} = \begin{pmatrix} \frac{3}{5} \\ \frac{2}{5} \\ 1 \\ \frac{1}{5} \end{pmatrix}$. Then the general least square solution is $\boldsymbol{z} + \boldsymbol{y}$ where $\boldsymbol{y}$ is the column space of $I_4 - A^\dagger A$ which is equal to $Span \left(\begin{pmatrix} 1 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \end{pmatrix} \right)$.

10. $\begin{pmatrix} 1 & 1 \\ 2 & 8 \\ -2 & 5 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 3 & 9 \\ 0 & 3 \end{pmatrix}$. This leads to the equation $\begin{pmatrix} 3 & 9 \\ 0 & 3 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 2 \\ 7 \\ -5 \end{pmatrix}$.

After performing the multiplication on the right we get

$$\begin{pmatrix} 3 & 9 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix},$$

$$\boldsymbol{x} = \begin{pmatrix} -\frac{13}{3} \\ \frac{5}{3} \end{pmatrix}.$$

11. Let $Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix}$.

Then $A = QR$ where $R = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 5 & 4 \\ 0 & 0 & 1 \end{pmatrix}$. Then $R^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{10} & -\frac{9}{10} \\ 0 & \frac{1}{5} & -\frac{9}{5} \\ 0 & 0 & 1 \end{pmatrix}$

The solution is $\boldsymbol{x}' = R^{-1}(Q^{tr}\boldsymbol{b} = \begin{pmatrix} 4 \\ 1 \\ -2 \end{pmatrix}$

12. $y = 2.06 + 3.01x$.

13. $y = 3.51 - 356x$.

14. $y = 2.92 - 1.88x + 1.20x^2$.

15. $y = .26 + .87x + .35x^2$.

16. $y = .35e^{1.55t}$.

17. $y = 2.53e^{-.14t}$.

13.2. Error Correcting Codes

1. For an arbitrary word $\boldsymbol{z} = (c_1 \ \dots \ c_n)$ let $spt(\boldsymbol{z}) = \{i|c_i \neq 0)$ so that $wt(\boldsymbol{z}) = |spt(\boldsymbol{z})|$. For words $\boldsymbol{x}$ and $\boldsymbol{y}$ we have $spt(\boldsymbol{x}+\boldsymbol{y}) \subset spt(\boldsymbol{x}) \cup spt(\boldsymbol{y})$ and the result follows from this.

2. $d(\boldsymbol{x}, \boldsymbol{z}) = wt(\boldsymbol{x} - \boldsymbol{z}) = wt([\boldsymbol{x} - \boldsymbol{y}] + [\boldsymbol{y} - \boldsymbol{z}]) \leq wt(\boldsymbol{x} - \boldsymbol{y}) + wt(\boldsymbol{y} - \boldsymbol{z})$ by Theorem (13. 8) (Exercise 1). By the definition of distance

$$wt(\boldsymbol{x} - \boldsymbol{y}) + wt(\boldsymbol{y} - \boldsymbol{z}) = d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}\boldsymbol{z}).$$

3. Part i. Assume $d(\boldsymbol{w}, \boldsymbol{x}) = t$ where $\boldsymbol{w} = (a_1 \ \dots \ a_n)$. Then $\boldsymbol{x}$ and $\boldsymbol{w}$ differ in exactly t places. There are $\binom{n}{t}$ ways to pick those places. If i is such a place then the i^{th} component of $\boldsymbol{x}$ is not a_i and we have $q-1$ choices for the i^{th} component of $\boldsymbol{x}$. Thus, the number of such $\boldsymbol{x}$ is $\binom{n}{t}(q-1)^t$.

ii. The closed ball, $B_r(\boldsymbol{w})$ is the disjoint union of $\{\boldsymbol{w}, \{\boldsymbol{x}|d(\boldsymbol{w}, \boldsymbol{x}) = 1\}, \dots, \{\boldsymbol{x}|d(\boldsymbol{w}, \boldsymbol{x}) = r\}$. The result now follows by part i.

4. For a vector $\boldsymbol{v}$ set $\phi(\boldsymbol{v}) = \boldsymbol{v} \cdot \boldsymbol{v}$. This is a quadratic form. A self-dual code is a totally singular subspace. If q is odd then $(\mathbb{F}_q^n, \phi)$ is non degenerate and the Witt index is at most $\lfloor \frac{n}{2} \rfloor$. If q is even and n is even then $(\mathbb{F}_q^n, \phi)$ is a non degenerate hyperbolic orthogonal space and so has Witt index $\frac{n}{2}$. On the other hand, if q is even and n is odd then the orthogonal space $(\mathbb{F}_q^n, \phi)$ is nonsingular: it has a radical of dimension one (the space spanned by the all one vector) which is a non-singular vector. In this case the Witt index is $\frac{n-1}{2}$.

5. a) Let $\boldsymbol{z}$ denote the all one vector. Then $\boldsymbol{z} \in \mathcal{H}$. The map $\boldsymbol{x} \to \boldsymbol{z}+\boldsymbol{x}$ is a bijection from the collection of words of length t to words of length $7-t$.

b) Since the minimum weight is 3 there are no words of weight 1 or 2 and by part a) no words of weight 5 or 6. So the weight of a word in $\mathcal{H}$ is in $\{0, 3, 4, 7\}$. There is a single word of weight 0 and weight 7. There are equally many words of weight 3 and 4 and their total is 14 so there are 7 of each.

6. The parity check extends $\boldsymbol{0}_7$ to $\boldsymbol{0}_8$ and extends the all one vector of length 7 to the all one vector of length 8. Each of the 7 words of weight 4 are extended by adding a component equal to zero so each of these gives rise to a word of weight 4. On the other hand each of the 7 words of weight 3 are extended by adding a component equal to one so each of these also gives rise to a word of weight 4. Therefore, in all, there are 14 words of weight four in $\overline{\mathcal{H}}$.

7. $\boldsymbol{x} \cdot \boldsymbol{x} = wt(\boldsymbol{x}) \times 1_{\mathbb{F}_q}$ is equal to one of $\boldsymbol{w}(\boldsymbol{x})$ is odd and zero if $wt(\boldsymbol{x})$ is even.

8. $\boldsymbol{x} \cdot \boldsymbol{y} = |spt(\boldsymbol{x}) \cap spt(\boldsymbol{y})| \times 1_{\mathbb{F}_q}$. Therefore $\boldsymbol{x} \cdot \boldsymbol{y} = 0$ if $spt(\boldsymbol{x}) \cap spt(\boldsymbol{y})|$ has an even number of elements and is one if $|spt(\boldsymbol{x}) \cap spt(\boldsymbol{y})$ is odd.

9. Since $wt(\boldsymbol{x})$ is even for every word in $\overline{\mathcal{H}}$ it follows by Exercise 6 that $\boldsymbol{x} \cdot \boldsymbol{x} = 0$ for every $\boldsymbol{x} \in \overline{\mathcal{H}}$. Clearly if $\boldsymbol{x} = \boldsymbol{0}$ and $\boldsymbol{y}$ is arbitrary then $\boldsymbol{x} \cdot \boldsymbol{y} = 0$. On the other hand, since $\boldsymbol{x} \cap \boldsymbol{z} = \boldsymbol{x}$ it follows from Exercise 7 that $\boldsymbol{z} \cdot \boldsymbol{x} = 0$. We may therefore assume that $wt(\boldsymbol{x}) = wt(\boldsymbol{y}) = 4$ and prove that $\boldsymbol{x} \cdot \boldsymbol{y} = 0$. By Exercise 7 we need to prove that $|spt(\boldsymbol{x}) \cap spt(\boldsymbol{y})|$ is even. Suppose $|spt(\boldsymbol{x}) \cap spt(\boldsymbol{y})| = 1$. Then $wt(\boldsymbol{x}+\boldsymbol{y}) = 6$ which contradicts Exercise 5. Likewise, if $|spt(\boldsymbol{x}) \cap sit(\boldsymbol{y})| = 3$ then $wt(\boldsymbol{x}+\boldsymbol{y}) = 2$, again a contradiction.

10. Let $\boldsymbol{x}, \boldsymbol{w}$ be code words. Suppose $\boldsymbol{y} \in B_3(\boldsymbol{w}) \cap B_3(\boldsymbol{x})$. Then $d(\boldsymbol{w}, \boldsymbol{x}) \leq d(\boldsymbol{w}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{x}) \leq 3+3 = 6$, a contradiction. Therefore $B_3(\boldsymbol{w}) \cap B_3(\boldsymbol{x}) = \emptyset$. By part ii. of Theorem (13.11), $|B_3(\boldsymbol{w})| = 1+23+\binom{23}{2}+\binom{23}{3} = 1+23+253+1771 = 2048 = 2^{11}$. Since $\mathcal{C}$ has dimension 12, the number of words in $\mathcal{C}$ is 2^{12}. Since the balls of radius 3 centered at the code words are disjoint it follows that

$$|\cup_{\boldsymbol{w}\in\mathcal{C}} B_3(\boldsymbol{w})| = |\mathcal{C}| \times |B_3(\boldsymbol{w})| = 2^{12} \times 2^{11} = 2^{23} = |\mathbb{F}_2^{23}|.$$

Thus, $\{B_3(\boldsymbol{w}) | \boldsymbol{w} \in \mathcal{C}\}$ is a partition of $\mathbb{F}_2^{23}$.

11. The number of nonzero vectors in $\mathbb{F}^n$ is $q^n - 1$. Since $Span(\boldsymbol{x})$ contains $q-1$ nonzero vectors the number of one dimensional subspaces of $\mathbb{F}_q^n$ is $\frac{q^n-1}{q-1} = t$. Thus, there are t columns in the parity check matrix $H(n,q)$ so the length of the Hamming (n,q)-code is t.

We next claim that $rank(H(n,q)) = n$. Since there are n rows it follows that $rank(H(n,q)) \leq n$. On the other hand if $X_1, \dots, X_n$ are one dimensional subspaces, $X_1 + \cdots + X_n = \mathbb{F}_q^n$ and $X_i = Span(\boldsymbol{x}_i)$ for $1 \leq i \leq n$ then $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ is a basis of $\mathbb{F}_q^n$. Therefore, $rank(H(n,q)) \geq n$. Since the Hamming code is the null space of $H(n,q)$ we can conclude it has dimension $t-n$.

We now prove the assertion about the minimum distance. Since any two columns are linearly independent the minimum distance is at least 3. However if X, Y, Z are three distinct one dimensional subspaces such that $X + Y = X + W = Y + W$ and $X = Span(\boldsymbol{x}), Y = Span(\boldsymbol{y})$, and $W = Span(\boldsymbol{w})$ then $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w})$ is linearly dependent. Consequently, there are words of weight three. Thus, the minimum weight is exactly three.

It now follows that the balls $B_1(\boldsymbol{w})$ where $\boldsymbol{w}$ is in the Hamming code are disjoint since the minimum distance is 3. By Theorem (13.11) the number of vector in such a ball is $1 + t(q-1) = 1 + \frac{q^n-1}{q-1} \times (q-1) = q^n$. Since there are q^{t-n} code words we get

$$|\cup_{\boldsymbol{w}\in Ham(n,q)} B_1(\boldsymbol{w})| = |Ham(n,q)| \times |B_1(\boldsymbol{w})| = q^{t-n} \times q^n = q^t.$$

Thus, $\{B_1(\boldsymbol{w}) | \boldsymbol{w} \in Ham(n,q)\}$ is a partition of $\mathbb{F}_q^t$ and $Ham(n,q)$ is a perfect 1-error correcting code.

13.3. Ranking Web Pages

1. The matrix is $\begin{pmatrix} A & \mathbf{0}_{5\times 4} \\ \mathbf{0}_{4\times 5} & B \end{pmatrix}$ where $A = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 1 \\ \frac{1}{3} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 1 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & \frac{1}{3} & 0 & 0 \\ 1 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{2} & 0 \end{pmatrix}$.

2. The last column is a zero column and therefore its entries do not add up to 1.

3. $\widehat{L} = \begin{pmatrix} A & \mathbf{0}_{5\times 3} & \frac{1}{9}\boldsymbol{j}_5 \\ \mathbf{0}_{4\times 5} & B' & \frac{1}{9}\boldsymbol{j}_4 \end{pmatrix}$. where $\boldsymbol{j}_5$ is the all one 5-vector, $\boldsymbol{j}_4$ is the all one 4-vector and $B' = \begin{pmatrix} 0 & \frac{1}{3} & 0 \\ 1 & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{2} \end{pmatrix}$.

4. Since $Span(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4, \boldsymbol{e}_5)$ is invariant the matrix is reducible.

5. $Span(\begin{pmatrix} \frac{12}{31} \\ \frac{4}{31} \\ \frac{2}{31} \\ \frac{6}{31} \\ \frac{7}{31} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix})$.

6.

$$\begin{pmatrix}
\frac{1}{36} & \frac{29}{72} & \frac{1}{36} & \frac{10}{36} & \frac{28}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{4}{36} \\
\frac{10}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{9} \\
\frac{1}{36} & \frac{29}{72} & \frac{1}{36} & \frac{29}{72} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{9} \\
\frac{10}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{9} \\
\frac{10}{36} & \frac{1}{36} & \frac{1}{36} & \frac{29}{72} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{9} \\
\frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{10}{36} & \frac{1}{36} & \frac{1}{9} \\
\frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{7}{9} & \frac{1}{36} & \frac{29}{72} & \frac{1}{9} \\
\frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{10}{36} & \frac{1}{36} & \frac{1}{9} \\
\frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{1}{36} & \frac{10}{36} & \frac{29}{72} & \frac{1}{9}
\end{pmatrix}$$