A Solutions Manual for Linear Algebra, 4th Edition by *Friedberg, Insel, & Spence*

Kyle Mickelson

August 2022

Contents

1 Vector Spaces

Here i will put the introduction

1.1 Introduction

why

1.2 Vector Spaces

Problem 1.2.1
a.) True. Axioms of vector space.

b.) False. By Corollary 1 of Theorem 1.1, the zero vector is unique.

c.) False. Consider $x = \vec{0}$.

d.) False.

e.) True. This is a column vector (nx1 matrix).

f.) False. n columns and m rows.

g.) False. Place $0 \in \mathbb{F}$ as coefficients.

h.) False. Not necessarily.

i.) True.

j.) True.

k.) True. $f, g \in \mathcal{F}(S, F)$, if $f(s) = g(s)$ for all $s \in S \implies f = g$.

Problem 1.2.2
The zero vector in $M_{3\times 4}(\mathbb{F})$ is the matrix with $0 \in \mathbb{F}$ in all entries.

$$\vec{0} = O = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Problem 1.2.3
$M_{13} = 3$, $M_{21} = 4$, and $M_{22} = 5$.

Problem 1.2.4
a.)

$$\begin{pmatrix} 6 & 3 & 2 \\ -4 & 3 & 9 \end{pmatrix}$$

b.)

$$\begin{pmatrix} 1 & -1 \\ 3 & -5 \\ 3 & 8 \end{pmatrix}$$

c.)

$$\begin{pmatrix} 8 & 20 & -12 \\ 4 & 0 & 28 \end{pmatrix}$$

d.)

$$\begin{pmatrix} 30 & -20 \\ -15 & 10 \\ -5 & -40 \end{pmatrix}$$

e.) $2x^4 + x^3 + 2x^2 - 2x + 10$

f.) $-x^3 + 7x^2 + 4$

g.) $10x^7 + -30x^4 + 40x^2 - 15x$

h.) $3x^5 - 6x^3 + 12x + 6$

Problem 1.2.5

$$\begin{pmatrix} 8 & 3 & 1 \\ 3 & 0 & 0 \\ 3 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 9 & 1 & 4 \\ 3 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 17 & 4 & 5 \\ 6 & 0 & 0 \\ 4 & 1 & 0 \end{pmatrix}$$

Problem 1.2.6

$$2\begin{pmatrix} 4 & 2 & 1 & 3 \\ 5 & 1 & 1 & 4 \\ 3 & 1 & 2 & 6 \end{pmatrix} = \begin{pmatrix} 8 & 4 & 2 & 6 \\ 10 & 2 & 2 & 8 \\ 6 & 2 & 4 & 12 \end{pmatrix}$$

$$\begin{pmatrix} 8 & 4 & 2 & 6 \\ 10 & 2 & 2 & 8 \\ 6 & 2 & 4 & 12 \end{pmatrix} - \begin{pmatrix} 5 & 3 & 1 & 2 \\ 6 & 2 & 1 & 5 \\ 1 & 0 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 1 & 4 \\ 4 & 0 & 1 & 3 \\ 5 & 2 & 1 & 9 \end{pmatrix}$$

So the sum of the entries is the number of suites sold in the June sale, 34.

Problem 1.2.7

Consider $\mathcal{F}(\{0,1\},\mathbb{R})$. Let $f(t) = 2t + 1$, $g(t) = 1 + 4t - 2t^2$, and $h(t) = t^t + 1$.

First we see $f(0) = 1$ and $f(1) = 3$. Also $g(0) = 1$ and $g(1) = 3$. Hence $f(t) = g(t)$ for all $t \in \{0, 1\}$ and so $f = g$. Now see that $h(0) = 2$ and $h(1) = 6$. Then $(f + g)(0) = 2 = h(0)$ and $(f + g)(1) = 6 = h(0)$, and so $f + g = h$.

Problem 1.2.8
Let V be a vector space and $x, y \in V$ and $a, b \in \mathbb{F}$. By VS7 and VS8, we have $(a + b)(x + y) = (a + b)x + (a + b)y = ax + bx + ay + by$. Since the choice of vectors and scalars was arbitrary, the equality holds for all $x, y \in V$ and $a, b \in \mathbb{F}$.

Problem 1.2.9
Corollary 1 of Theorem 1.1: Suppose $\vec{0}$ and $\vec{0}'$ are both vectors satisfying VS3. Now let $x \in V$. Then we have $\vec{0} + x = \vec{0}' + x \implies \vec{0} = \vec{0}'$ by Theorem 1.1, and hence the zero vector is unique. ■

Corollary 2 of Theorem 1.1: Let V be a vector space. Suppose y and y' are vectors satisfying VS4 for $x \in V$. Then $y + x = 0 = y' + x \implies y = y'$ by Theorem 1.1, and hence the additive inverse vector is unique. ■

Theorem 1.2 (c): We have $a(\vec{0} + \vec{0}) = a\vec{0} + a\vec{0} = a\vec{0} = \vec{0}$ by VS7. ■

Problem 1.2.10
Let V be the set of all differentiable real-valued functions on the real line. Define addition and multiplication by the familiar operations for functions. Note the set is closed since addition of differentiable functions produces differentiables functions and differentiable functions scaled by a nonzero constant are again differentiable. Since function addition is commutative and associative, and obeys the distributive laws for scalars and vectors, in addition to having additiive inverses and the zero function, V is a vector space by definition.

Problem 1.2.11
Let $V = \{\vec{0}\}$ and define $\vec{0} + \vec{0} = \vec{0}$ and $c\vec{0} = \vec{0}$ for all $c \in \mathbb{F}$. Then VS1 and VS2 are satisfied trivially. Also, VS3 is satisfied since $\vec{0}$ acts as the zero vector. VS4 is satisfied as $\vec{0}$ satisfies this condition. Since $1 \cdot \vec{0} = \vec{0}$ by the operation, then VS5 is also satisfied. VS6,VS7, and VS8 follow trivially. Hence V is a vector space (the zero vector space). ■

Problem 1.2.12
Let V be the set of even functions with real values, with usual addition and scalar multiplication of functions. Since addition of even functions is again even, and scaling an even function retains an even function, then the set V is closed under the operations. By commutativity and associativity of function addition, VS1 and VS2 are satisfied. $f(0) = 0 \in V$ is the zero vector and if $f(t) \in V$, then $-f(t) \in V$ so that VS3 and VS4 are satisfied. Also $1 \cdot f(t) = f(t)$, so VS5 holds. Let $a, b \in \mathbb{F}$, then $(ab)f(t) = a(bf(t)$. Also $a(f + g)(t) = (af + ag)(t) = af(t) + ag(t)$. Also $(a + b)f(t) = af(t) + bf(t)$ and so VS6, VS7, and VS8 all hold. Hence V is a vector space. ■

Problem 1.2.13
Let $(a_1, b_1), (a_2, b_2) \in V$. Then $(a_1, b_1) + (a_2, b_2) = (a_1 + b_1, a_2 b_2) \neq (a_2, b_2) + (a_1, b_1) = (a_2 + b_2, a_1 b_2)$, and hence VS1 fails to hold since the addition operation is not commutative. Hence V is not a vector space. $\blacksquare$

Problem 1.2.14
V is indeed a vector space over $\mathbb{R}$ since for any $r \in \mathbb{R}$, and $\vec{v} \in V$, we have $r\vec{v} \in V$. ($i \in \mathbb{C}$ does not cause problems here because vectors are closed under scalar multiplication by $\mathbb{R}$). The verification that V is a vector space follows exactly from Example 1 in the text. $\blacksquare$

Problem 1.2.15
On the other hand, in this case V is not a vector space over the field $\mathbb{C}$, since for the scalar $i \in \mathbb{C}$, and for any $\vec{v} \in V$, $i\vec{v} \notin V$ since $i \notin \mathbb{R}$. Hence the set is not closed under scalar multiplication and cannot be a vector space. $\blacksquare$

Problem 1.2.16
Since $\mathbb{Q}$ is also a field, and the fact that none of the vector space axioms are affected by the choice of field, $M_{m \times n}(\mathbb{Q})$ is indeed a vector space by the same reasoning that $M_{m \times n}(\mathbb{R})$ is a vector space. $\blacksquare$

Problem 1.2.17
Let $(a_1, a_2) \in V$, and $1 \in \mathbb{F}$. From the problem, we know that $1 \cdot (a_1, a_2) = (a_1, 0) \neq (a_1, a_2)$, and thus VS5 fails to hold. Therefore V cannot be a vector space. $\blacksquare$

Problem 1.2.18
Long and boring and follows from previous exercises.

Problem 1.2.19
Follow from previous exercises.

Problem 1.2.20
Let V be the set of sequences $\{a_n\}$ of real numbers with the operations defined in the problem. Since addition of sequences is commutative and associative, VS1 and VS2 hold. 0 acts as the zero vector, and is the zero sequence, and also if $\{a_n\} \in V$, then $\{-a_n\} \in V$ is the additive inverse. Hence VS3 and VS4 also hold. We see that $1 \cdot \{a_n\} = \{1 \cdot a_n\} = \{a_n\}$ and so VS5 holds. VS6, VS7, and VS8 all hold similarly given the defined operations. Hence V is a vector space as desired. $\blacksquare$

Problem 1.2.21
Let V and W be vector spaces over a field $\mathbb{F}$. Define $Z = \{(v, w) : v \in V, w \in W\}$

with operations given in the problem. We see immediately that:

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2) = (v_2 + v_1, w_2 + w_1) = (v_2, w_2) + (v_1, w_1)$$

Which follows from commutativity of addition in V and W. Hence VS1 is satisfied for Z. Now

$$((v_1, w_1) + (v_2, w_2)) + (v_3, w_3) = (v_1 + v_2, w_1 + w_2) + (v_3, w_3) =$$

$$(v_1 + v_2 + v_3, w_1 + w_2 + w_3) = (v_1, w_1) + (v_2 + v_3, w_2 + w_3) =$$

$$(v_1, w_1) + ((v_2, w_2) + (v_3, w_3))$$

And hence VS2 holds in Z, following from associativity of addition in V and W. Now note that the zero vector in Z is simply $(\vec{0}_V, \vec{0}_W) \in Z$, and the additive inverse is $(-v_1, -w_1) \in Z$. Hence VS3 and VS4 also hold. Now:

$$1 \cdot (v_1, w_1) = (1 \cdot v_1, 1 \cdot w_1) = (v_1, w_1)$$

And so VS5 also holds. The checks of VS6, VS7, and VS8 follow in the same way as previously. Hence Z is a vector space over $\mathbb{F}$ as desired. ■

Problem 1.2.22
We know that the field $\mathbb{Z}/2\mathbb{Z} = \{0, 1\}$. Hence in the vector space $M_{m\times n}(\mathbb{Z}/2\mathbb{Z})$, we know that there are mn entries in each matrix. Since each entry can have one of two options for scalar from the field, there must be 2^{mn} different unique matrices in this vector space.

1.3 Subspaces

Problem 1.3.1
a.) False. Not necessarily. W must be a vector space over the same field as V.

b.) False. $\vec{0} \notin \emptyset$.

c.) True. Consider $W = \{\vec{0}\}$.

d.) False. The intersection of two subspaces of V is a subspace of V.

e.) True. An $n \times n$ diagonal matrix has exactly n entries on the diagonal.

f.) False. The trace is the sum of the diagonal entries.

g.) False. $\mathbb{R}^2 \subseteq W$, but $W \nsubseteq \mathbb{R}^2$.

Problem 1.3.2

a.)

$$A^T = \begin{pmatrix} -4 & 2 \\ 5 & -1 \end{pmatrix}^T = \begin{pmatrix} -4 & 5 \\ 2 & -1 \end{pmatrix},\ (A) = -5$$

b.)

$$\begin{pmatrix} 0 & 8 & -6 \\ 3 & 4 & 7 \end{pmatrix}^T = \begin{pmatrix} 0 & 3 \\ 8 & 4 \\ -6 & 7 \end{pmatrix}$$

c.)

$$\begin{pmatrix} -3 & 9 \\ 0 & -2 \\ 6 & 1 \end{pmatrix}^T = \begin{pmatrix} -3 & 0 & 6 \\ 9 & -2 & 1 \end{pmatrix}$$

d.)

$$A^T = \begin{pmatrix} 10 & 0 & -8 \\ 2 & -4 & 3 \\ -5 & 7 & 6 \end{pmatrix}^T = \begin{pmatrix} 10 & 2 & -5 \\ 0 & -4 & 7 \\ -8 & 3 & 6 \end{pmatrix},\ (A) = 12$$

e.)

$$\begin{pmatrix} 1 & -1 & 3 & 5 \end{pmatrix}^T = \begin{pmatrix} 1 \\ -1 \\ 3 \\ 5 \end{pmatrix}$$

f.)

$$\begin{pmatrix} -2 & 5 & 1 & 4 \\ 7 & 0 & 1 & -6 \end{pmatrix}^T = \begin{pmatrix} -2 & 7 \\ 5 & 0 \\ 1 & 1 \\ 4 & -6 \end{pmatrix}$$

g.)

$$\begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix}^T = \begin{pmatrix} 5 & 6 & 7 \end{pmatrix}$$

h.)

$$A^T = \begin{pmatrix} -4 & 0 & 6 \\ 0 & 1 & -3 \\ 6 & -3 & 5 \end{pmatrix}^T = \begin{pmatrix} -4 & 0 & 6 \\ 0 & 1 & -3 \\ 6 & -3 & 5 \end{pmatrix},\ (A) = 2$$

Problem 1.3.3

Let $A, B \in M_{m\times n}(\mathbb{F})$ and $a, b \in \mathbb{F}$. Then $(aA + bB)^T_{ij} = (aA_{ij} + bB_{ij})^T = aA_{ji} + bB_{ji} = aA^T + bB^T$ as desired. ■

Problem 1.3.4

Let $A \in M_{m\times n}(\mathbb{F})$. Then $(A^T)^T = (A^T_{ij})^T = (A_{ji})^T = A_{ij} = A$. ■

Problem 1.3.5
Let A be a square matrix. Then $(A + A^T)_{ij} = A_{ij} + A_{ji}$. Now we see that $(A + A^T)^T_{ij} = A^T_{ij} + (A^T_{ij})^T = A_{ji} + A_{ij}$ by Problem 1.3.4. Therefore $(A + A^T) = (A + A^T)^T$, by commutativity of entry-wise addition of matrices, and by definition $A + A^T$ is a symmetric matrix. ■

Problem 1.3.6
Let $A, B \in M_{n\times n}(\mathbb{F})$. Then $(aA + bB) = \sum_{i,j=1}^{n}(aA_{ii} + bB_{jj}) = a\sum_{i=1}^{n} A_{ii} + b\sum_{j=1}^{n} B_{jj} = a(A) + b(B)$. ■

Problem 1.3.7
Let A be a diagonal matrix. Then by definition $A^T_{ij} = A_{ji}$. But since all entries in A are on the diagonal, we have $A^T_{ii} = A_{ii}$, and hence $A^T = A$. Therefore all diagonal matrices are symmetric. ■

Problem 1.3.8
a.) $W_1 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : a_1 = 3a_2, a_3 = -a_2\}$. By inspection, this set is a line in $\mathbb{R}^3$ that contains the origin. Hence it is a subspace of $\mathbb{R}^3$.

b.) $W_2 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : a_1 = a_3 + 2\}$. This set does not contain the zero vector and hence is not a subspace of $\mathbb{R}^3$.

c.) $W_3 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : 2a_1 - 7a_2 + a_3 = 0\}$. By inspection, this set is a plane in $\mathbb{R}^3$ that contains the origin. Hence it is a subspace of $\mathbb{R}^3$.

d.) $W_4 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : a_1 - 4a_2 - a_3 = 0\}$. By inspection, this set is a plane in $\mathbb{R}^3$ that contains the origin. Hence it is a subspace of $\mathbb{R}^3$.

e.) $W_5 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : a_1 + 2a_2 - 3a_3 = 1\}$. This set is clearly not a subspace of $\mathbb{R}^3$ since it does not contain the zero vector.

f.) $W_6 = \{(a_1, a_2, a_3) \in \mathbb{R}^3 : 5a_1^2 - 3a_2^2 + 6a_3^2 = 0\}$. By inspection, this set is not a plane or a line in $\mathbb{R}^3$ that contains the origin. Hence it is not a subspace of $\mathbb{R}^3$.

Problem 1.3.9
The set $W_1 \cap W_3$ is the interseciton of a line and a plane in $\mathbb{R}^3$, and hence is the origin (a subspace). The same goes for $W_1 \cap W_4$. For $W_3 \cap W_4$, we see two planes intersecting, again a line through the origin and hence a subspace.

Problem 1.3.10
The set $W_2 = \{(a_1, ..., a_n) \in \mathbb{F}^n : a_1 + \cdots + a_n = 1\}$ is clearly not a subspace of $\mathbb{F}^n$ since $\vec{0} \notin W_2$ ($\vec{0} = (0, 0, ..., 0)$). However, note that for the set $W_1 = \{(a_1, ..., a_n) \in \mathbb{F}^n : a_1 + \cdots + a_n = 0\}$, $\vec{0} \in W_1$, and if $x, y \in W_1$ and

$\alpha \in \mathbb{F}$, then:

$$\alpha x + y = \alpha(a_1, ..., a_n) + (b_1, ..., b_n) = (\alpha a_1 + b_1, ..., \alpha a_n + b_n) \in W_1$$

Because the entries in the n-tuple all add to 0 as well. Hence by Theorem 1.3, W_1 is a subspace of the vector space $\mathbb{F}^n$. ■

Problem 1.3.11
The set $W = \{f(x) \in P(\mathbb{F}) : f(x) = 0,$or has degree n $\}$ is not a subspace of $P(\mathbb{F})$ since it is not closed under vector addition. Consider two vectors with the same leading coefficient when subtracted.

Problem 1.3.12
Consider the subset of upper triangular matrices, U, of the vector space $M_{m\times n}(\mathbb{F})$. Let $A, B \in U$, and let $\mu \in \mathbb{F}$. Then $\mu A = (\mu A)_{ij} = \mu A_{ij} = 0$ when $i > j$, so that $\mu A \in U$. Now $A + B = (A + B)_{ij} = A_{ij} + B_{ij} = 0$ when $i < j$, so that $A + B \in U$ as well. Also, trivially $O \in U$, where O is the $m \times n$ zero matrix. Hence, by Theorem 1.3, U is a subspace of $M_{m\times n}(\mathbb{F})$ as desired. ■

Problem 1.3.13
Let $S \neq \emptyset$ and $\mathbb{F}$ be a field. Let $s_0 \in S$, and define the set $\{f \in \mathcal{F}(S, \mathbb{F}) : f(s_0) = 0\}$. Clearly the set is a subset of $\mathcal{F}(S, \mathbb{F})$. Also the zero function is in the set since the zero function maps all elements of S to 0. Now suppose f, g are elements of the set, and $a \in \mathbb{F}$. Then we see that:

$$af(s_0) + g(s_0) = a0 + 0 = 0$$

And hence the set is closed under scalar multiplication and addition. Therefore by Theorem 1.3 this set is a subspace of $\mathcal{F}(S, \mathbb{F})$ as desired. ■

Problem 1.3.14

Problem 1.3.15

Problem 1.3.16

Problem 1.3.17
Let V be a vector space. Suppose that W is a subspace of V. Let $x, y \in W$ and $a \in \mathbb{F}$. Then by definition W must contain the zero vector of V, and so $\vec{0} \in W$. Also, W must be closed under scalar multiplication and addition, hence $ax+y \in W$. Conversely, suppose that $\vec{0} \in W$ and $ax+y \in W$. Then W contains the zero vector and is closed under scalar multiplication and vector addition, and hence by Theorem 1.3 is a subspace of V. ■

Problem 1.3.18
Suppose W is a subspace of a vector space V. Then by Theorem 1.3, we know that $\vec{0} \in W$ and $ax + y \in W$ for all $x, y \in W$ and $a \in \mathbb{F}$. Conversely, suppose that $\vec{0} \in W$, and $ax + y \in W$ for all $x, y \in W$ and $a \in \mathbb{F}$. Then clearly $W \subseteq V$, and it is closed under scalar multiplication and vector addition. Also it is a vector space over the same field and is thus by definition a subspace of V. ■

Problem 1.3.19
Let W_1 and W_2 be subspaces of a vector space V. Suppose $W_1 \cup W_2$ is a subspace of V. Now suppose that $W_1 \nsubseteq W_2$ and $W_2 \nsubseteq W_1$. Then $x \in W_1$ and $y \in W_2$, we have $y = (x + y) - x \in W_2$, a contradiction.

Conversely, suppose that $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$. Without loss of generality, assume $W_1 \subseteq W_2$. Then $W_1 \cup W_2 = W_2$, which is a subspace of V and hence $W_1 \cup W_2$ is a subspace of V. ■

Problem 1.3.20

Problem 1.3.21

Problem 1.3.22

Problem 1.3.23
a.) Let W_1 and W_2 be subspaces of a vector space V. Then $W_1 + W_2 = \{x + y : x \in W_1, y \in W_2\}$. Since $\vec{0} \in W_1$ and $\vec{0} \in W_2$, then $\vec{0} = \vec{0} + \vec{0} \in W_1 + W_2$. Let $x + y, x' + y' \in W_1 + W_2$ and $a \in \mathbb{F}$. Then $(x + y) + (x' + y') = x + y + x' + y' = (x + x') + (y + y') \in W_1 + W_2$. Also, $a(x + y) = ax + ay \in W_1 + W_2$. Hence by Theorem 1.3, the sum $W_1 + W_2$ is a subspace of V, which clearly contains W_1 and W_2. ■

b.) Suppose U is a subspace of V that contains W_1 and W_2. Then for any $x \in W_1$ and $y \in W_2$, we know that $x, y \in U$. Hence $x + y \in U$ by additive closure of the subspace. Since the choice of x and y were arbitrary, clearly $W_1 + W_2 \subseteq U$. ■

Problem 1.3.24
Let $\vec{x} \in \mathbb{F}^n$ and suppose $\vec{x} \in W_1 \cap W_2$, so that $\vec{x} = \{0, 0, \ldots, 0\} = \vec{0}$. Thus $W_1 \cap W_2 = \{\vec{0}\}$. Now suppose $\vec{x} \in \mathbb{F}^n$. Then $\vec{x} = (a_1, a_2, \ldots, a_n)$, with $a_i \in \mathbb{F}$ for $1 \leq i \leq n$. Notice that $\vec{x} = \vec{v} + \vec{w} = (a_1, \ldots, a_{n-1}, 0) + (0, \ldots, 0, a_n)$, and $\vec{v} \in W_1$ and $\vec{w} \in W_2$, so that $\mathbb{F}^n = W_1 + W_2$. Given these two conditions, we can conclude that $W_1 \oplus W_2 = \mathbb{F}^n$. ■

Problem 1.3.25

Problem 1.3.26

Problem 1.3.27

Problem 1.3.28
Let W_1 denote the set of all skew-symmetric $n \times n$ matrices with entries from $\mathbb{F}$. O, the zero matrix is trivially an element of the set since $O^T = -O = O$. Now let A, B be skew-symmetric, and $\alpha \in \mathbb{F}$. Then $(\alpha A + B)^T = \alpha A^T + B^T = -(\alpha A) - B = -(\alpha A + B)$ by the results shown in Problem 1.3.3. Hence $\alpha A + B$ is again a skew-symmetric matrix, and so the set is closed under addition and scalar multiplication. By Theorem 1.3, it is a subspace of $M_{n \times n}(\mathbb{F})$.

Let W_2 denote the subspace of $M_{n \times n}(\mathbb{F})$ consisting of all symmetric matrices. Suppose that $X \in W_1 \cap W_2$. Then $X^T = -X$ because X is skew-symmetric, but also $X^T = X$ since X is symmetric. Therefore $X = O$, the zero matrix. Hence $W_1 \cap W_2 = \{0\}$. Now let $X + Y \in W_1 + W_2$, where $X \in W_1$ and $Y \in W_2$. Then $X + Y \in M_{n \times n}(\mathbb{F})$ trivially, and hence $W_1 + W_2 \subseteq M_{n \times n}(\mathbb{F})$. Now suppose $A \in M_{n \times n}(F)$. Then:

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$$

But note that $\frac{1}{2}(A + A^T)^T = \frac{1}{2}(A^T + A) = \frac{1}{2}(A + A^T \implies \frac{1}{2}(A + A^T) \in W_1$. Additionally, see that $\frac{1}{2}(A - A^T)^T = \frac{1}{2}(-A^T + A) \implies \frac{1}{2}(A - A^T) \in W_2$. Hence any $A \in M_{n \times n}(\mathbb{F})$ can be written as the sum of a skew-symmetric and a symmetric matrix. Hence $M_{n \times n}(\mathbb{F}) \subseteq W_1 + W_2$. Therefore we have shown that $M_{n \times n}(\mathbb{F}) = W_1 \oplus W_2$ as desired. ■

Problem 1.3.29

Problem 1.3.30
Let W_1 and W_2 be subspaces of a vector space V. Suppose that $V = W_1 \oplus W_2$. Then $W_1 \cap W_2 = \{0\}$ and $V = W_1 + W_2$. Let $v \in V$ such that $v \neq \vec{0}$. Then $v = x_1 + x_2$, where $x_1 \in W_1$ and $x_2 \in W_2$. In the case where $x_1 = \vec{0}$, $v \in W_2$. Also if $x_2 = \vec{0}$, then $v \in W_1$. In the case where $x_1 = x_2 \neq \vec{0}$, suppose that $v = x_1 + x_2$ and $v = x_1' + x_2'$. Then without loss of generality, $x_1 = x_1' + x_2' - x_2 \in W_1 \implies x_2 \in W_1$, but $x_2 \in W_2$ and $W_1 \cap W_2 = \{\vec{0}\}$, so $x_2 = \vec{0}$. Hence a contradiction so that the decomposition of v into $x_1 \in W_1$ and $x_2 \in W_2$ is unique.

Conversely, suppose that $v \in V$ can be written uniquely as $x_1 + x_2$, where $x_1 \in W_1$ and $x_2 \in W_2$. Hence $V = W_1 + W_2$. Then $\vec{0} = x_1 + x_2 \implies x_1 = -x_2$, and so without loss of generality $x_2 \in W_1$, but v is written uniquely, and so $x_2 = 0 \implies W_1 \cap W_2 = \{\vec{0}\}$, and so $V = W_1 \oplus W_2$ by definition of direct sum.

■

Problem 1.3.31
Let V be a vector space over $\mathbb{F}$ and let W be a subspace of V.

a.) Suppose that $v + W$ is a subspace of V. Then by definition $\vec{0} \in v + W \implies v + w = \vec{0} \implies v = -w \in W$, hence $v \in W$. Conversely suppose that $v \in W$. Then the set $v+W = W$. Hence $v+W$ is a subsapce of V since W is a subspace of V. ■

b.) Suppose that $v_1 + W = v_2 + W$. Then $v_1 + h = v_2 + k$ for some $h, k \in W$. So then we have $v_1 - v_2 = k - h \in W$, and thus $v_1 - v_2 \in W$. Conversely, suppose that $v_1 - v_2 \in W$. Then $v_1 - v_2 = w$ for some $w \in W$. Then $v_1 = v_2 + w \implies v_1 \in v_2 + W \implies v_1 + W = v_2 + W$. ■

c.) Suppose that $v_1 + W = v_1' + W$ and $v_2 + W = v_2' + W$. Then we see:

$$(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W = (v_1' + v_2') + W = (v_1' + W) + (v_2' + W)$$

So that addition of cosets is well-defined. Now let $a \in \mathbb{F}$. Then $a(v_1 + W) = (av_1) + W = (av_1') + W = a(v_1' + W)$, and hence scalar multiplication of cosets is a well-defined operation as well.

d.) We will prove that the set V/W is a subspace of V. Clearly $V/W \subseteq V$. Let $v_1+W, v_2+W \in V/W$. Then $(v_1+W)+(v_2+W) = (v_1+v_2)+W \in V/W$ since $v_1 + v_2 \in V$. Also $a(v_1 + W) = (av_1) + W \in V/W$ since $av_1 \in V$. Finally, the zero vector of V/W is $\vec{0}_V + W$. Therefore by Theorem 1.3 V/W is a subspace of V, called the quotient space of V modulo W. ■

1.4 Linear Combinations and Systems of Linear Equations

Problem 1.4.1
a.) True.

b.) False. It is defined to be $\{0\}$.

c.) False.

d.) False. (0)

e.) True.

f.) False.

Problem 1.4.2

Problem 1.4.3

Problem 1.4.4

Problem 1.4.5

Problem 1.4.6
Let $S = \{(1,1,0),(1,0,1),(0,1,1)\}$. Clearly $span(S) \subseteq \mathbb{F}^3$. Now let $v \in \mathbb{F}^3$ such that $v = (\alpha, \beta, \gamma)$, with $\alpha, \beta, \gamma \in \mathbb{F}$. Then $a + b = \alpha$, $a + c = \beta$, and $b + c = \gamma$, where $a, b, c \in \mathbb{F}^3$. Then $a(1,1,0) + b(1,0,1) + c(0,1,1) = (a+b, a+c, b+c) = (\alpha, \beta, \gamma) = v$, and since v was arbitrary, we see $\mathbb{F}^3 \subseteq span(S)$. Hence $span(S) = \mathbb{F}^3$ as desired. ■

Problem 1.4.7
By scaling each e_i, $1 \leq i \leq n$, by a scalar from the field and then adding each vector componentwise, we see that:

$$span(\{e_1, ..., e_n\}) = \{(a_1, a_2, ..., a_n) : a_i \in \mathbb{F}\} = \mathbb{F}^n$$

And hence the set $\{e_1, ..., e_n\}$ is a generating set for $\mathbb{F}^n$.

Problem 1.4.8
Scale each vector in the set by a scalar and then add each vector componentwise to find that the span of the set is equal to:

$$span(\{1, x, x^2, ..., x^n\}) = \{a_1 + a_2x + a_3x^2 + \cdots + a_nx^n : a_i \in \mathbb{F}\} = P_n(\mathbb{F})$$

Problem 1.4.9
By scaling each of the matrices presented by a scalar and then adding them via properties of matrix addition, we can clearly see that:

$$span(\{\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\}) = \{\begin{pmatrix}a & b\\b & c\end{pmatrix} \in M_2(\mathbb{F}) : a, b, c, d \in \mathbb{F}\}$$

Note that the set on the right is just the set of 2x2 matrices with coefficients from $\mathbb{F}$. Hence by definition this set of matrices generates $M_2(\mathbb{F})$. ■

Problem 1.4.10
By scaling each matrix by a constant and adding them by matrix addition, We see that:

$$span(\{M_1, M_2, M_3\}) = \{\begin{pmatrix}a & b\\b & c\end{pmatrix} \in M_2(\mathbb{F}) : a, b, c \in \mathbb{F}\}$$

Now let $B \in M_2(\mathbb{F})$ be a symmetric matrix. Then $B^T = B \implies b_{21} = b_{12} \implies B \in span(\{M_1, M_2, M_3\})$. Hence the span of $\{M_1, M_2, M_3\}$ is equal to the set of 2x2 symmetric matrices. ■

Problem 1.4.11
Let $v \in span(\{x\})$. Then v is some linear combination of x, namely a multiple of x. Hence $v \in \{ax : a \in \mathbb{F}\} \implies span(\{x\}) \subseteq \{ax : a \in \mathbb{F}\}$. The converse follows trivially in that the set on the right is the set of all multiples of x, identical to the span of x. ■

Problem 1.4.12
Suppose that $W \subseteq V$ is a subspace of a vector space V. Since W is a subspace of V that contains $W \subseteq V$ trivially, then by Theorem 1.5, $span(W) \subseteq W$. By definition of span, we see that any for any $w \in W$, $w \in span(W)$, so $W \subseteq span(W)$, and hence $span(W) = W$. Conversely, suppose that $span(W) = W$. By Theorem 1.5, the span of any subset of V is a subspace of V. Hence $W \subseteq V$ is a subspace of V as desired. ■

Problem 1.4.13
Let V be a vector space and let $S_1, S_2 \subseteq V$. Suppose that $S_1 \subseteq S_2$. Then, since all of S_1 is contained in S_2, any linear combination of vectors in S_1 are necessarily contained in the set of all linear combinations of S_2, or equivalently, $span(S_1) \subseteq span(S_2)$. ■

In particular, if $span(S_1) = V$, then clearly $span(S_2) = V$ as well.

Problem 1.4.14
Let V be a vector space. Suppose that $S_1, S_2 \subseteq V$. Then the set $span(S_1 \cup S_2)$ contains all linear combinations of vectors in S_1 or S_2 or both. Let $x \in span(S_1 \cup S_2)$. In the case where x is a linear combination of vectors in S_1, we have that $x \in span(S_1)$. By symmetry, the same is true for S_2. In the case where x is a linear combination of vectors in both S_1 and S_2, then we can decompose x into its components in S_1 and S_2. Hence $x = s + t$ where $s \in span(S_1)$ and $t \in span(S_2)$. Therefore $span(S_1 \cup S_2) \subseteq span(S_1) + span(S_2)$. Now suppose that $x \in span(S_1) + span(S_2)$. Then $x = x' \in span(S_1 \cup S_2)$, and hence we see that $span(S_1 \cup S_2) = span(S_1) + span(S_2$. ■

Problem 1.4.15
Let V be a vector space and let $S_1, S_2 \subseteq V$. Let $x \in span(S_1 \cap S_2)$. Then $x = \sum_{i=1}^{n} a_i v_i$, where $a_i \in \mathbb{F}$ and $v_i \in S_1 \cap S_2$. Since $v_i \in S_1 \cap S_2$, then clearly $v_i \in span(S_1)$ and $v_i \in span(S_2)$, and $v_i \in span(S_1) \cap span(S_2)$. Since x is a linear combination of the v_i, then we see that $x \in span(S_1) \cap span(S_2) \implies span(S_1 \cap S_2) \subseteq span(S_1) \cap span(S_2)$ as desired. ■

An example of set equality above would be if we took $S_1 = S_2 = \emptyset$. Then $span(\emptyset) = \{0\} = span(\emptyset) \cap span(\emptyset) = \{0\}$. For a case where they are unequal, consider $S_1 =$

Problem 1.4.16
The subset S in question is simply a linearly independent subset of the vector space V.

Problem 1.4.17
The condition required would be that the subspace in question is finite.

1.5 Linear Dependence and Linear Independence

Problem 1.5.1
a.) False. At least one vector is a linear combination of the others.

b.) True. The zero vector is a linear combination of any vector.

c.) False. Linearly dependent sets must be non-empty.

d.) False. Not always true.

e.) True.

f.) True. Definition of linearly independent set.

Problem 1.5.2
a.) Linearly dependent. Multiply first matrix in the set by -2, obtain the second matrix.

b.) Linearly independent. No scalar makes the first matrix into the second.

Problem 1.5.3
We will show this explicitly.

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

And clearly none of the scalars are zero. Hence the set is linearly dependent in $M_{3\times 2}(\mathbb{F})$ by definition.

Problem 1.5.4
Clearly $\alpha\vec{e}_i \neq \vec{e}_j$ for any $i \neq j$ and all $\alpha \in \mathbb{F}$. Hence no vector in the set is a linear combination of the others. Thus by definition the set is linearly independent in $\mathbb{F}^n$.

Problem 1.5.5
Clearly $1 \neq x^i$ for any $1 \leq i \leq n$. Hence 1 is not a linear combination of the

other vectors in the set. Also $\alpha x^i \neq x^j$ for any $i \neq j$ and all $\alpha \in \mathbb{F}$. Hence no other vectors in the set are linear combinations of the others. Therefore the set is linearly independent in $P_n(\mathbb{F})$.

Problem 1.5.6
Consider the set $\{E^{ij} : 1 \leq i \leq m, 1 \leq j \leq n\} \subseteq M_{m\times n}(\mathbb{F})$. By the notation defined in the problem, we can clearly see $\alpha E^{ij} \neq E^{kh}$ for any $i \neq k, j \neq h$ and all $\alpha \in \mathbb{F}$. Hence no matrix in this set is a linear combination of the others and thus the set is linearly independent in $M_{m\times n}(\mathbb{F})$.

Problem 1.5.7
We aim to find a linearly independet set that generates the subspace of diagonal matrices. Note that:

$$span(\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\}) = \{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \in M_2(\mathbb{F}) : a, b \in \mathbb{F}\}$$

And we know that the set on the right is precisely the set of all diagonal 2x2 matrices. Additionally, the set of two matrices is clearly linearly independent. Hence this is a linearly independet subset that spans the subspace.

Problem 1.5.8
Let $S = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\} \subseteq \mathbb{F}^n$.

a.) Suppose $\mathbb{F} = \mathbb{R}$. Then $a(1, 1, 0) + b(1, 0, 1) + c(0, 1, 1) = (a + b, a + c.b + c)$. Then the set is linearly dependent $\iff$ $a + b = 0, a + c = 0, b + c = 0$. However then we obtain the relations $a = -b$ and $a = b$ concurrently, and hence $a = b = 0 \implies c = 0$. Therefore the only linear combination of the three vectors is the trivial combination and thus the set is linearly indepedent.

b.) Now suppose that $\mathbb{F}$ has characteristic 2. Then we see that $(1, 1, 0) + (1, 0, 1) + (0, 1, 1) = (1 + 1, 1 + 1, 1 + 1) = (0, 0, 0)$ since $1 + 1 = 0$ in $\mathbb{F}$. Hence there is a nontrivial linear combination of the three vectors that constructs the zero vector and thus the set is linearly dependent.

Problem 1.5.9
Let V be a vector space and let $u, v \in V$ be distinct vectors. Suppose that $\{u, v\}$ is linearly dependent. Then by definition $u = \alpha v$ for some $\alpha \in \mathbb{F}$, so that u is a multiple of v. By symmetry, v is also a multiple of u. Conversely, suppose that, without loss of generality, u is a multiple of v. Then $v = \alpha u$ for some $\alpha \in \mathbb{F}$ and hence $\{u, v\}$ is linearly dependent by definition. ■

Problem 1.5.10
Consider the set in $\mathbb{R}^3$, that of $\{(1, 1, 1), (1, 2, 3), (3, 2, 1)\}$. None of the vectors are a multiple of the others and they form a linearly dependent set.

Problem 1.5.11

Let $S = \{u_1, ..., u_n\}$ be a linearly independent subset of a vector space V over the field $\mathbb{Z}/2\mathbb{Z}$. Since $\mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$ (residue classes), the only possible scalars for the vectors in $span(S)$ are 0 and 1. Since there are n total vectors in S, and two options for scalars, we see that there must be 2^n different vectors in $span(S)$.

Problem 1.5.12
a.) Theorem 1.6: Let V be a vector space and let $S_1 \subseteq S_2 \subseteq V$. Suppose S_1 is linearly dependent. Let $S_1 = \{u_1, ..., u_k\}$, so we know that $\sum_{i=1}^{k} a_i u_i = 0$, where not all of the $a_i = 0$. Let $S_2 = \{u_1, ..., u_n\}$. Then all of the u_i for $1 \leq i \leq k$ are still in S_2, and we see that $\sum_{i=1}^{n} a_i u_i = 0$, setting $a_i = 0$ for $k < i \leq n$. Then not all of the scalars are zero and the set is thus linearly dependent. ■

b.) Corollary: Let V be a vector space and let $S_1 \subseteq S_2 \subseteq V$. Suppose that S_2 is linearly independent. Then in particular we know that none of the vectors in S_2 are a linear combination of the other vectors. Thus since $S_1 \subseteq S_2$, none of the vectors in S_1 are a linear combination of the other vectors and hence S_1 is linearly independent by definition. ■

Problem 1.5.13
a.) Let V be a vector space over a field of characteristic not equal to 2. Let u, v be distinct vectors in V. Suppose that $\{u, v\}$ is linearly independent. Then in particular we know that $\vec{u} \neq \mu\vec{v}$ for any $\mu \in \mathbb{F}$. Now suppose by way of contradiction that the set $\{u+v, u-v\}$ is linearly dependent. Then there exists $\mu \in \mathbb{F}$ such that $\vec{u}+\vec{v} = \mu(\vec{u}-\vec{v}) = \mu\vec{u} - \mu\vec{v} \implies (1-\mu)\vec{u} = -(1-\mu)\vec{v} \implies \vec{u} = -\vec{v}$, thus contradicting the fact that $\vec{u} \neq \mu\vec{v}$ for any $\mu \in \mathbb{F}$. Hence the set $\{u+v, u-v\}$ is linearly independent. Conversely, now suppose that the set $\{u+v, u-v\}$ is linearly independent. Then $\vec{u}+\vec{v} \neq \mu(\vec{u}-\vec{v})$ for any $\mu \in \mathbb{F}$. Now suppose by way of contradiction that $\{u, v\}$ is a linearly dependent set. Then $\vec{u} = \lambda\vec{v}$ for some $\lambda \in \mathbb{F}$ by Problem 1.5.9. Then from the hypothesis, $\vec{u}+\vec{v} \neq \mu(\vec{u}-\vec{v}) \implies (\lambda+1)\vec{v} \neq \mu\vec{v}(\lambda-1) \implies \lambda+1 \neq \mu\lambda-\mu \implies \lambda(1-\mu) \neq -(\mu+1) \implies \lambda \neq 1$, so that $\vec{u}$ and $\vec{v}$ are not he same, but can be multiples, thus a contradiction. Hence the set $\{u, v\}$ must also be linearly independent as desired. ■

b.)

Problem 1.5.14
Suppose that S is a linearly dependent set. If $S = \{0\}$, then trivially linearly dependent. If $S \neq \{0\}$, then there exist some distinct vectors $v, u_i \in S$ for $1 \leq i \leq n$ such $v = \sum_{i=1}^{n} \mu_i u_i$ for $\mu_i \in \mathbb{F}$ for $1 \leq i \leq n$. Thus v is by definition a linear combination of the vectors $u_i \in S$. Conversely, suppose that $v \in S$ is a linear combination of $u_1, ..., u_n \in S$. Then by definition of a linearly dependent set, S is linearly dependent. If $S = \{0\}$, then by definitoon any set containing the zero vector is linearly dependent and thus S is linearly dependent. ■

Problem 1.5.15

Let $S = \{u_1, ..., u_n\}$ be a finite set of vectors. Suppose that the set S is linearly dependent. Then by definition, either one of the $u_i = \vec{0}$ for $1 \leq i \leq n$, or some $u_{k+1} \in S$ is a linear combination of $u_1, ..., u_k \implies u_k \in span(\{u_1, ..., u_k\})$. Conversely, suppose that $u_1 = \vec{0}$, then by definition S is a linearly dependent set and we are done. Now consider $u_i \neq \vec{0}$, and suppose $u_{k+1} \in span(\{u_1, ..., u_k\})$ for $1 \leq k < n$. Then u_{k+1} is a linear combination of the vectors $u_1, ..., u_k$ and hence the set S is linearly dependent by definition. ■

Problem 1.5.16 XX
Suppose that a set S is linearly independent. Then by the corollary to Theorem 1.6, any subset of S is linearly independent as well. Therefore necessarily any finite subset of S is also linearly independent. Conversely, suppose that each finite subset of S is linearly independent.

Problem 1.5.17
Let $M \in \{A \in M_n(\mathbb{F}) : a_{ij} = 0 \ \forall \ i > j\}$, a square upper triangular matrix. Then in particular the columns of M are:

$$\begin{pmatrix} \mu_{11} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \mu_{12} \\ \mu_{22} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \mu_{13} \\ \mu_{23} \\ \mu_{33} \\ \vdots \\ 0 \end{pmatrix} \dots$$

Where $\mu_{ij} \in \mathbb{F}$ for all $1 \leq i, j \leq n$. Clearly the set of column vectors are linearly independent, and hence the columns of the matrix M are linearly independent as well.

Problem 1.5.18 XX
Let $S = \{p_1(x), p_2(x), ..., p_n(x)\} \subseteq P(\mathbb{F})$ such that $\deg(p_i(x)) \neq \deg(p_j(x))$ for all $i \neq j$. Suppose by way of contradiction that the set S is linearly dependent. Then there is some vector, say $p_1(x)$ of degree k, such that:

$$\sum_{i=1}^{n} \mu_i p_i(x) = p_1(x)$$

However then the polynomial on the left would have degree k as well,

Problem 1.5.19
Suppose that $\{A_1, ..., A_k\}$ is a linearly independent subset of $M_n(\mathbb{F})$. Then we know that for $\mu_i \in \mathbb{F}$ with $1 \leq i \leq k$,

$$\sum_{i=1}^{k} \mu_i A_i = O \implies (\sum_{i=1}^{k} \mu_i A_i)^T = \sum_{i=1}^{k} \mu_i A_i^T = O^T = O$$

From the properties of the matrix transpose proved in Problem 1.3.3. Since the set $\{A_1, ..., A_k\}$ was linearly independent, then $\mu_i = 0$ for all $1 \leq i \leq k$, and hence the same holds for the set $\{A_1^T, ..., A_k^T\}$. Therefore the set $\{A_1^T, ..., A_k^T\}$ is linearly independent in $M_n(\mathbb{F})$ as desired. $\blacksquare$

Problem 1.5.20
Let $f, g \in \mathcal{F}(\mathbb{R}, \mathbb{R})$ defined by $f(t) = e^{rt}$ and $g(t) = e^{st}$, with $s \neq r$. Suppose by way of contradiction that the set $\{f, g\}$ is linearly dependent. Then $\alpha f(t) = g(t)$ for some $\alpha \in \mathbb{R}$ such that $\alpha \neq 0$. But then $\alpha e^{rt} = e^{st} \implies \alpha e^{rt-st} = \alpha e^{t(r-s)} = 0$. However we know that $e^t \neq 0$ for any $t \in \mathbb{R}$. Therefore $\alpha = 0$, a contradiction. Therefore the set $\{f, g\}$ is linearly independent in $\mathcal{F}(\mathbb{R}, \mathbb{R})$ as desired. $\blacksquare$

1.6 Bases and Dimension

Problem 1.6.1
a.) False. $\{0\}$.

b.) True.

c.) False. Consider $P(\mathbb{F})$.

d.) False. Bases need not be unique.

e.) True.

f.) False. $\dim(P_n(\mathbb{F})) = n + 1$.

g.) False. $\dim(M_{m\times n}(\mathbb{F})) = mn$.

h.) True.

i.) False. S is not linearly independent.

j.) True.

k.) True. These are the trivial subspaces of V.

l.) True.

Problem 1.6.2
a.) This set is a basis for $\mathbb{R}^3$.

b.) This set is not a basis for $\mathbb{R}^3$.

c.) This set is a basis for $\mathbb{R}^3$.

d.) This set is a basis for $\mathbb{R}^3$.

e.) This set is not a basis for $\mathbb{R}^3$.

Problem 1.6.3
a.) False, the set is not a basis for $P_2(\mathbb{R})$.

b.) True, is a basis for $P_2(\mathbb{R})$.

c.) True, is a basis for $P_2(\mathbb{R})$.

d.) True, this set is a basis for $P_2(\mathbb{R})$.

e.) False, it is not a basis for $P_2(\mathbb{R})$.

Problem 1.6.4
We are given three different polynomials and are asked if as a set they generate the vector space $P_3(\mathbb{R})$. Given that $\dim(P_3(\mathbb{R}) = 3+1 = 4$, and the set contains three polynomials, we conclude that the set does not generate the vector space. This is because any finite generating set for $P_3(\mathbb{R})$ contains at least 4 vectors, by Corollary 2 to Theorem 1.10.

Problem 1.6.5
This set contains four vectors, while the dimension of $\dim(\mathbb{R}^3) = 3$. Therefore there are more vectors than the dimension of the vector space and thus this set cannot be linearly independent. Refer again to Corollary 2 of Theorem 1.10.

Problem 1.6.6
Three bases for $\mathbb{F}^2$ are as follows:

$$\beta_1 = \{(1,0),(0,1)\},\ \beta_2 = \{(0,1),(1,-1)\},\ \beta_3 = \{(1,0),(1,1)\}$$

Three bases for $M_2(\mathbb{F})$ are as follows:

$$\beta_1 = \{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\},$$

$$\beta_2 = \{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\},$$

$$\beta_3 = \{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\}$$

Problem 1.6.7
See that $(-4)u_1 = u_3$, so discard u_3 from the set. Those primes in u_4 were evil, let me save you the trouble, discard it! A basis for $\mathbb{R}^3$ is given by the set $\{u_1, u_2, u_5\}$.

Problem 1.6.8

Problem 1.6.9
Let $a, b, c, d \in \mathbb{F}$. Then a linear combination of u_1, u_2, u_3, and u_4 is as follows:

$$a(1,1,1,1) + b(0,1,1,1) + c(0,0,1,1) + d(0,0,0,1) =$$

$$(a, a+b, a+b+c, a+b+c+d)$$

An arbitrary vector of $\mathbb{F}^4$ has a unique representation given by choice of scalars in the above vector.

Problem 1.6.10
a.)

b.)

c.)

d.)

Problem 1.6.11
Let u, v be distinct vectors of a vector space V. Suppose $\{u, v\}$ is a basis for V and let $a, b \in \mathbb{F}$ such that $a, b \neq 0$. Let $x \in V$. Then $x = \alpha u + \beta v$ for some $\alpha, \beta \in \mathbb{F} \implies x = \frac{\alpha}{a} au + \frac{\beta}{b} bv$ and so any vector $x \in V$ can be represented as $x = \gamma(a\vec{u}) + \delta(b\vec{v})$ for some $\gamma, \delta \in \mathbb{F} \implies span(\{a\vec{u}, b\vec{v}\}) = V$. Now since the set $\{u, v\}$ is linearly independent, we know that u is not a multiple of v, and thus any $a\vec{u}$ is not a multiple of $b\vec{v}$. Hence the set $\{a\vec{u}, b\vec{v}\}$ is linearly independent by Problem 1.5.9. Thus $\{a\vec{u}, b\vec{v}\}$ is also a basis for V.

Now consider the set $\{\vec{u}+\vec{v}, a\vec{u}\}$. Let $x \in V$ such that $x = (\alpha + \beta a)\vec{u} + (\alpha)\vec{v} = \alpha\vec{u} + \beta a\vec{u} + \alpha\vec{v} = \alpha(\vec{u}+\vec{v}) + \beta(a\vec{u}) \in span(\{\vec{u}+\vec{v}, a\vec{u}\}) \implies V \subseteq span(\{\vec{u}+\vec{v}, a\vec{u}\})$. The reverse containment holds trivially since $\vec{u}, \vec{v} \in V$. Thus $span(\{\vec{u}+\vec{v}, a\vec{u}\}) = V$. Now suppose by way of contradiction that $a\vec{u}$ is a multiple of $\vec{u}+\vec{v}$. Then for some $\lambda \in \mathbb{F}$, we have:

$$\vec{u} + \vec{v} = \lambda(a\vec{u}) \implies \vec{v} = \lambda a\vec{u} - \vec{u} = \vec{u}(\lambda a - 1)$$

But since $\lambda a - 1 \in \mathbb{F}$, this implies that $\vec{v} = \delta\vec{u}$ for some $\delta = \lambda a - 1$. However by the linear independence of the set $\{\vec{u}, \vec{v}\}$, we know that $\vec{u}$ is not a multiple of $\vec{v}$, and hence a contradiction. Therefore $a\vec{u}$ is not a multiple of $\vec{u} + \vec{v} \iff \{\vec{u}+\vec{v}, a\vec{u}\}$ is a linearly independent set. Hence the set $\{\vec{u}+\vec{v}, a\vec{u}\}$ is also a basis for V. ■

Problem 1.6.12
Let $\vec{u}, \vec{v}, \vec{w}$ be distinct vectors in a vector space V. Suppose that $\{\vec{u}, \vec{v}, \vec{w}\}$ is a

basis for V. Now consider the set $\{\vec{u}+\vec{v}+\vec{w}, \vec{v}+\vec{w}, \vec{w}\}$. Let $\vec{x} \in V$, then for some $\alpha, \beta, \gamma \in \mathbb{F}$,

$$\vec{x} = (\alpha)\vec{u} + (\alpha+\beta)\vec{v} + (\alpha+\beta+\gamma)\vec{w} = \alpha\vec{u} + \alpha\vec{v} + \alpha\vec{w} + \beta\vec{v} + \beta\vec{w} + \gamma\vec{w} =$$

$$\alpha(\vec{u}+\vec{v}+\vec{w}) + \beta(\vec{v}+\vec{w}) + \gamma(\vec{w}) \in span(\{\vec{u}+\vec{v}+\vec{w}, \vec{v}+\vec{w}, \vec{w}\})$$

And hence $V \subseteq span(\{\vec{u}+\vec{v}+\vec{w}, \vec{v}+\vec{w}, \vec{w}\})$, and because the reverse containment holds trivially, $span(\{\vec{u}+\vec{v}+\vec{w}, \vec{v}+\vec{w}, \vec{w}\}) = V$. Since the set $\{\vec{u}, \vec{v}, \vec{w}\}$ is a basis, and by definition linearly independent, we know that no vector is a multiple of the others in the set. Suppose by way of contradiction that the set $\{\vec{u}+\vec{v}+\vec{w}, \vec{v}+\vec{w}, \vec{w}\}$ is linearly dependent. Then we know that for some $\alpha, \beta, \gamma \in \mathbb{F}$:

$$\alpha(\vec{u}+\vec{v}+\vec{w}) + \beta(\vec{v}+\vec{w}) + \gamma(\vec{w}) = \vec{0} \implies$$

$$(\alpha)\vec{u} + (\alpha+\beta)\vec{v} + (\alpha+\beta+\gamma)\vec{w} = \vec{0}$$

But since $\{\vec{u}, \vec{v}, \vec{w}\}$ is linearly independent, there exists no nontrivial solutions to the equation $a\vec{u} + b\vec{v} + c\vec{w} = \vec{0}$. Hence replacing $a = \alpha$, $b = \alpha+\beta$, and $c = \alpha+\beta+\gamma$, we see that the above equation also has no nontrivial solutions. Hence a contradiction to the linear dependence of the set. Therefore the set $\{\vec{u}+\vec{v}+\vec{w}, \vec{v}+\vec{w}, \vec{w}\}$ is linearly independent and spans the vector space and is thus by definition a basis for the vector space V. ■

Problem 1.6.13
From the system of linear equations provided, we obtain the relation $x_1 = x_2 = x_3$. Therefore this subspace of $\mathbb{R}^3$ consists of all vectors with equal component values. Hence the simplest basis for this subspace would be given by the set $\{(1,1,1)\}$ which is clearly linearly independent, and spans the subspace. ■

Problem 1.6.14
Consider the subspace of $\mathbb{F}^5$, $W_1 = \{(a_1, a_2, a_3, a_4, a_5) \in \mathbb{F}^5 : a_1 - a_3 - a_4 = 0\}$. Note that any arbitrary $\vec{x} \in W_1$ has the form:

$$\vec{x} = \begin{pmatrix} a_3+a_4 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} a_2 + \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} a_3 + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} a_4 + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} a_5$$

Where $a_i \in \mathbb{F}$ for $i \in \{2,3,4,5\}$. And the vectors on the right form a linearly independent and spanning set for the subspace. Since there are 4 vectors, $\dim(W_1) = 4$.

Now consider $W_2 = \{(a_1, a_2, a_3, a_4, a_5) \in \mathbb{F}^5 : a_2 = a_3 = a_4,\ a_1 + a_5 = 0\}$.

Note that any arbitrary $\vec{x} \in W_2$ has the form:

$$\vec{x} = \begin{pmatrix} a_1 \\ a_2 \\ a_2 \\ a_2 \\ -a_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} a_1 + \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} a_2$$

With $a_1, a_2 \in \mathbb{F}$. The vectors on the right are clearly linearly independent and also span the subspace W_2. Therefore they form a basis for the subspace and thus $\dim(W_2) = 2$.

Problem 1.6.15
Let $W = \{A \in M_n(\mathbb{F}) : (A) = 0\}$ denote the subspace of $M_n(\mathbb{F})$ of matrices with zero trace.

Problem 1.6.16

Problem 1.6.17

Problem 1.6.18

Problem 1.6.19
Suppose that each $\vec{v} \in V$ can be expressed uniquely as a linear combination of vectors in the set $\beta = \{\vec{u_1}, ..., \vec{u_n}\} \subseteq V$. We will prove that β is then a basis for the vector space V. First, note that the span of any subset of V is a subspace of V, and thus we have $span(\beta) \subseteq V$. For the reverse containment, let $\vec{v} \in V$, so that

$$\vec{v} = \sum_{i=1}^{n} a_i \vec{u}_i$$

for unique $a_i \in \mathbb{F}$ for all $1 \leq i \leq n$. Thus by definition $\vec{v} \in span(\beta)$, and since $\vec{v}$ was arbitrary, $V \subseteq span(\beta) \implies span(\beta) = V$. Now all that is left to prove is that the set β is linearly independent. Since $\vec{0} \in V$, it can be expressed as a unique linear combination of vectors in β, with $\lambda_i \in \mathbb{F}$, but we also know that:

$$\vec{v} + \vec{0} = \sum_{i=1}^{n} a_i \vec{u}_i + \sum_{i=1}^{n} \lambda_i \vec{u}_i = \sum_{i=1}^{n} (a_i + \lambda_i) \vec{u}_i = \vec{v} = \sum_{i=1}^{n} a_i \vec{u}_i$$

But since $a_i + \lambda_i = a_i$ for all $i \iff \lambda_i = 0$ for all i, then we know that since the representation is unique, the only linear combination of vectors in β to create the zero vector is the trivial solution, the set β is linearly independent. Hence β is a basis for the vector space V as desired. $\blacksquare$

Problem 1.6.20

Problem 1.6.21
Let V be a vector space. We will show by contrapositive that if V is infinite dimeensional, then V contains an infinite linearly independent subset. Suppose that V does not contain an infinite linearly independent subset. Then any linearly independent subset of V is necessarily finite. Hence any basis for V is necessarily finite as well, since it is a linearly independent subset. Thus by definition V is a finite dimensional vector space. Hence by contrapositive, if V is an infinite dimensional vector space, then it necessarily contains an infinite linearly independent subset. Conversely, suppose that V contains an infinite linearly independent subset. Now suppose by way of contradiction that a finite subset β is a basis for V. However we see that then the basis contains less vectors than the infinite linearly independent subset, contradicting the linear independence of the infinite subset. Hence V cannot have a finite basis set and therefore V is infinite dimensional. ■

Problem 1.6.22
Let W_1 and W_2 be subspaces of a finite-dimensional vector space V. Suppose that $\dim(W_1 \cap W_2) = \dim(W_1)$. Note that $W_1 \cap W_2$ is a finite dimensional subspace of W_1, which is also finite dimensional, and so by Theorem 1.11, we know that since $\dim(W_1 \cap W_2) = \dim(W_1$, then $W_1 \cap W_2 = W_1 \implies W_1 \subseteq W_2$. Conversely, suppose $W_1 \subseteq W_2$, so then $W_1 \cap W_2 = W_1$, and necessarily $\dim(W_1 \cap W_2) = \dim(W_1)$. Hence the containment of the subspace W_1 in W_2 is the necessary and sufficient condition for the relation $\dim(W_1 \cap W_2) = \dim(W_1)$ as desired. ■

Problem 1.6.23

Problem 1.6.24

Problem 1.6.25

Problem 1.6.26

Problem 1.6.27

Problem 1.6.28
Let V be a vector space over the field $\mathbb{C}$ such that $\dim(V) = n$. Let the subset

$\beta = \{u_1, ..., u_n\}$ be a basis for V. Let $v \in V$, then for $a_j, b_j \in \mathbb{R}$, $1 \le j \le n$,

$$v = \sum_{j=1}^{n}(a_j + b_j i)u_j = \sum_{j=1}^{n} a_j u_j + i\sum_{j=1}^{n} b_j u_j$$

Which follows from the fact that any $z \in \mathbb{C}$ is such that $z = a + bi$ where $a, b \in \mathbb{R}$. Now, suppose we regard V as a vector space over the field $\mathbb{R}$. Then note that we have:

$$v = \sum_{j=1}^{n} a_j u_j + i\sum_{j=1}^{n} b_j u_j = \sum_{j=1}^{n} a_j u_j + \sum_{j=1}^{n} b_j(iu_j)$$

And hence any vector in V is an element of the span of the set $\beta \cup \{iu_1, ..., iu_n\}$. Additionally this set is obviously linearly independent by the linear independence of the set β. Therefore $\beta \cup \{iu_1, ..., iu_n\}$ is a basis for V and clearly $\dim(V) = 2n$. ■

Problem 1.6.29

a.) Suppose that W_1 and W_2 are finite-dimensional subspaces of a vector space V. Let $\{u_1, ..., u_k\}$ be a basis for $W_1 \cap W_2$. Since $W_1 \cap W_2$ is a subspace of W_1, by the corollary to Theorem 1.11 we can extend this basis to a basis for W_1. Let the set $\{u_1, ..., u_k, v_1, ..., v_m\}$ be such an extension. Now since $W_1 \cap W_2$ is also a subspace of W_2, extend the basis for $W_1 \cap W_2$ to a basis for W_2 to obtain the set $\{u_1, ..., u_k, w_1, ..., w_p\}$. Define the set $\beta = \{u_1, ..., u_k, v_1, ..., v_m, w_1, ..., w_p\}$. Now consider the set $W_1 + W_2$. Let $t \in W_1 + W_2$, such that $t = x + y$ where $x \in W_1$ and $y \in W_2$, then observe:

$$t = x + y = (\sum_{i=1}^{k} a_i u_i + \sum_{i=1}^{m} b_i v_i) + (\sum_{i=1}^{k} c_i u_i + \sum_{i=1}^{p} d_i w_i) =$$

$$\sum_{i=1}^{k}(a_i + c_i)u_i + \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{p} d_i w_i \implies t \in span(\beta)$$

Hence $W_1 + W_2 \subseteq span(\beta)$. Now let $s \in span(\beta)$, so we have that:

$$s = \sum_{i=1}^{k} a_i u_i + \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{p} c_i w_i$$

Then let $a_i = \mu_i + \lambda_i$ for $1 \le i \le k$. Then the above equation becomes:

$$s = \sum_{i=1}^{k}(\mu_i + \lambda_i)u_i + \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{p} c_i w_i =$$

$$(\sum_{i=1}^{k} \mu_i u_i + \sum_{i=1}^{m} b_i v_i) + (\sum_{i=1}^{k} \lambda_i u_i + \sum_{i=1}^{p} c_i w_i) \implies s \in W_1 + W_2$$

Meaning $span(\beta) \subseteq W_1 + W_2 \implies span(\beta) = W_1 + W_2$. Now we will show that β is a linearly independent set. We have $\vec{0} \in W_1 + W_2$, by definition of a subspace, and so there exist scalars a_i, $1 \leq i \leq k$, b_i, $1 \leq i \leq m$, and c_i, $1 \leq i \leq p$, all in $\mathbb{F}$, such that:

$$\vec{0} = \sum_{i=1}^{k} a_i u_i + \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{p} c_i w_i \implies$$

$$-\sum_{i=1}^{k} a_i u_i = \sum_{i=1}^{k} (-a_i) u_i = \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{p} c_i w_i$$

However note that each of the vectors u_i for $1 \leq i \leq k$ are not multiples of the vectors v_i for $1 \leq i \leq m$ or the vectors w_i for $1 \leq i \leq p$ due to the linear independence of the bases for W_1 and W_2 previously established. Hence $a_i = b_i = c_i = 0$ for all respective i indices, and hence the set β is a linearly independent set by definition. Therefore β is a basis for the subspace $W_1 + W_2$ containing $k+m+p$ vectors, and thus $W_1 + W_2$ is a finite-dimensional subspace. From the above discussion, we can formulate the following relation:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$$

Which is the desired relation between the sum and intersections of the subspaces of V. ■

b.) Let W_1 and W_2 be finite-dimensional subspaces of a vector space V and let $V = W_1 + W_2$. We know that $V = W_1 \oplus W_2 \iff W_1 \cap W_2 = \{\vec{0}\}$. Which is equivalent to $\dim(W_1 \cap W_2) = 0$. In fact, from the equation derived in the first part of the problem, we have:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$$

Which, if $\dim(W_1 \cap W_2) = 0$, becomes:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2)$$

Hence from the above discussion we can deduce that $V = W_1 \oplus W_2 \iff \dim(W_1 + W_2) = \dim(W_1) + \dim(W_2)$.

Problem 1.6.30

Consider the vector space $M_2(\mathbb{F})$. Let $W_1 = \{\begin{pmatrix} a & b \\ c & a \end{pmatrix} \in M_2(\mathbb{F}) : a, b, c \in \mathbb{F}\}$. Note that the zero matrix $O \in W_1$ trivially. Now let $A, B \in W_1$ such that:

$$A = \begin{pmatrix} a & b \\ c & a \end{pmatrix} \text{ , } B = \begin{pmatrix} a' & b' \\ c' & a' \end{pmatrix}$$

Then, letting $\mu \in \mathbb{F}$, we can see that:

$$\mu A + B = \begin{pmatrix} \mu a & \mu b \\ \mu c & \mu a \end{pmatrix} + \begin{pmatrix} a' & b' \\ c' & a' \end{pmatrix} = \begin{pmatrix} \mu a + a' & \mu b + b' \\ \mu c + c' & \mu a + a' \end{pmatrix} \in W_1$$

Hence W_1 satisfies all conditions for identification as a subspace of $M_2(\mathbb{F})$ as per Theorem 1.3. A basis set for this subspace is given by the following set:

$$\{\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}\} \implies \dim(W_1) = 3$$

Now consider the set $W_2 = \{\begin{pmatrix}0 & a\\-a & b\end{pmatrix} \in M_2(\mathbb{F}) : a, b \in \mathbb{F}\}$. The zero matrix $O \in W_2$ trivially. Now let $A, B \in W_2$ such that:

$$A = \begin{pmatrix}0 & a\\-a & b\end{pmatrix} \ , \ B = \begin{pmatrix}0 & a'\\-a' & b'\end{pmatrix}$$

Then, letting $\mu \in \mathbb{F}$, we can see that:

$$\mu A + B = \begin{pmatrix}0 & \mu a\\-(\mu a) & \mu b\end{pmatrix} + \begin{pmatrix}0 & a'\\-a' & b'\end{pmatrix} = \begin{pmatrix}0 & \mu a + a'\\-(\mu a + a') & \mu b + b'\end{pmatrix} \in W_2$$

And thus again by Theorem 1.3, W_2 is a subspace of $M_2(\mathbb{F})$. We can easily find that a basis for W_2 is given by:

$$\{\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\} \implies \dim(W_2) = 2$$

Now we turn to the question of $\dim(W_1 + W_2)$ and $\dim(W_1 \cap W_2)$. From the results obtained in Problem 1.6.29, if we obtain 1, we can find both of these dimensions. We eventually obtain the set:

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

Which is a basis for the subspace $W_1 \cap W_2$ (find this by requiring conditions of matrices in W_1 and W_2 and then going from there). Hence $\dim(W_1 \cap W_2) = 1$, and thus from the equation derived in Problem 1.6.29, we know that $\dim(W_1 + W_2) = 3 + 2 - 1 = 4$.

Problem 1.6.31
Let W_1 and W_2 be subspaces of a vector space V such that $\dim(W_1) = m$ and $\dim(W_2) = n$, with $m \geq n$.

a.) We know that $\dim(W_1) \geq \dim(W_2)$. Additionally, $W_1 \cap W_2$ is a subspace of both W_1 and W_2, and hence from Theorem 1.11, we know that $\dim(W_1 \cap W_2) \leq \dim(W_2) \implies \dim(W_1 \cap W_2) \leq n$. ■

b.) Consider the formula derived in Problem 1.6.29,

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$$

Which holds true given the finite-dimensionality of the subspaces W_1 and W_2. Now notice that in the case where $\dim(W_1 \cap W_2) = 0$, we have the equality

$\dim(W_1 + W_2) = n + m$. However in the case where $\dim(W_1 \cap W_2) > 0$, we have that $\dim(W_1 + W_2) < m + n$. Hence we see that $\dim(W_1 + W_2) \leq m + n$ as desired. ■

Problem 1.6.32
a.) Consider the subspaces of $\mathbb{R}^3$, $W_1 = \{(a, b, 0) \in \mathbb{R}^3 : a, b \in \mathbb{R}\}$ and $W_2 = \{(a, 0, 0) \in \mathbb{R}^3 : a \in \mathbb{R}\}$. Note that $\dim(W_1) = 2$ and $\dim(W_2) = 1$. Observe that $W_1 \cap W_2 = \{(a, 0, 0) \in \mathbb{R}^3 : a \in \mathbb{R}\} = W_2 \implies \dim(W_1 \cap W_2) = \dim(W_2) = 1$.

b.) Now consider the subspaces of $\mathbb{R}^3$, $W_1 = \{(a, b, 0) \in \mathbb{R}^3 : a, b \in \mathbb{R}\}$ and $W_2 = \{(0, 0, c) \in \mathbb{R}^3 : c \in \mathbb{R}\}$. Note that $\dim(W_1) = 2$ and $\dim(W_2) = 1$. Now observe that $W_1 + W_2 = \{(a, b, c) \in \mathbb{R}^3 : a, b, c \in \mathbb{R}\} = \mathbb{R}^3$. Hence it is easy to see that $\dim(W_1 + W_2) = 2 + 1 = 3$, which is of course the same as the dimension of the vector space $\mathbb{R}^3$.

c.) Now consider the subspaces of $\mathbb{R}^3$, $W_1 = \{(a, 0, c) \in \mathbb{R}^3 : a \in \mathbb{R}\}$ and $W_2 = \{(a, b, 0) \in \mathbb{R}^3 : b \in \mathbb{R}\}$. Note that $\dim(W_1) = 2$, $\dim(W_2) = 2$, which satisfies the requirements given in the problem. Observe that $W_1 \cap W_2 = \{(a, 0, 0) \in \mathbb{R}^3 : a \in \mathbb{R}\}$, and so $\dim(W_1 \cap W_2) = 1 < \dim(W_2) = 2$. Additionally, see that the set $W_1 + W_2 = \{(a, b, c) \in \mathbb{R}^3 : a, b \in \mathbb{R}\}$. Then $\dim(W_1 + W_2) = 3 < 2 + 2 = 4$

Problem 1.6.33
a.) Let W_1 and W_2 be subspaces of a vector space V such that $V = W_1 \oplus W_2$. Suppose that β_1 is a basis for W_1 and β_2 is a basis for W_2. By the direct sum definition, we know that $W_1 \cap W_2 = \{\vec{0}\} \implies \beta_1 \cap \beta_2 = \emptyset$, due to the fact that $\vec{0} \notin \beta_1, \beta_2$ by the definition of a basis, and the fact that $\beta_1 \subseteq W_1$ and $\beta_2 \subseteq W_2$. Since $V = W_1 + W_2$, and $span(\beta_1) = W_1$ and $span(\beta_2) = W_2$, then by Problem 1.5.19, we know that $span(\beta_1) + span(\beta_2) = span(\beta_1 \cup \beta_2) = V$. Since we know that $span(\beta_1) \cap span(\beta_2) = \{\vec{0}\}$, and that basis representations are unique, then $\vec{0} = \vec{0}_{W_1} + \vec{0}_{W_2} \in V$, and so the only linear combination that gives the zero vector is the trivial combination. Hence the set $\beta_1 \cup \beta_2$ is a basis for V. ■

b.) Conversely, let $\beta_1 \cap \beta_2$ be disjoint bases for subspaces W_1 and W_2 of a vector space V. Suppose now that $\beta_1 \cup \beta_2$ is a basis for V. Then by definition $span(\beta_1 \cup \beta_2) = span(\beta_1) + span(\beta_2) = V$ by Problem 1.5.19. Since β_1 is a basis for W_1 and β_2 is a basis for W_2, we have now that $W_1 + W_2 = V$. We can also see that $W_1 \cap W_2 = \{\vec{0}\}$, for if this was not the case, then we would have two unique representations for a vector, contradicting the fact that $\beta_1 \cup \beta_2$ is a basis for V. Therefore we have that $V = W_1 \oplus W_2$ as desired. ■

Problem 1.6.34
a.) Let W_1 be a subspace of a finite-dimensional vector space V. Since V is finite-dimensional, it has a finite basis. Let β_1 be a basis for W_1. Extend β_1 to

a basis for V by adjoining the set β_2, so that $\beta_1 \cup \beta_2$ is a basis for V. Clearly $\beta_1 \cap \beta_2 = \emptyset$, since if not then they would contain duplicate vectors, contradicting the fact that their union is a basis for V. Now note that since $\beta_2 \subseteq \beta_1 \cup \beta_2$, the set β_2 is necessarily linearly independent by the corolloary to Theorem 1.6. Now define $W_2 = span(\beta_2)$, a subspace of V, and β_2 is clearly a basis for W_2. Now by Problem 1.6.33, since $\beta_1 \cup \beta_2 = \emptyset$ and $\beta_1 \cup \beta_2$ is a basis for V, we have that $V = W_1 \oplus W_2$ as desired. ■

b.) Consider the vector space $\mathbb{R}^2$ and the subspace $W_1 = \{(a_1, 0) : a_1 \in \mathbb{R}\}$. The obvious first choice for a subspace W_2 such that $V = W_1 \oplus W_2$ is the subspace $W_2 = \{(0, a_2) : a_2 \in \mathbb{R}\}$. Note that $W_1 + W_2 = \{(a_1, a_2) : a_1, a_2 \in \mathbb{R}\} = \mathbb{R}^2$. For the next choice, consider the subspace $W_2' = \{x(a_1, a_2) : a_1, a_2 \in \mathbb{F}\}$

Problem 1.6.35
Let W be a subspace of a finite-dimensional vector space V. Let $\{u_1, ..., u_k\}$ be a basis for W and let $\{u_1, ..., u_k, u_{k+1}, ..., u_n\}$ be an extension of that basis to a basis of V.

a.) We know that $V/W = \{v + W : v \in V\}$. Now consider the set $\{u_{k+1} + W, u_{k+2} + W, ..., u_n + W\}$. We will show that this set is a basis for V/W. First let $v + W \in V/W$, where $v \notin W$, since if $v \in W$, then $v + W = W$. Then:

$$\sum_{i=k+1}^{n} a_i(u_i + W) = \sum_{i=k+1}^{n} a_i u_i + W = v + W$$

And hence $V/W \subseteq span(\{u_{k+1} + W, ..., u_n + W\}$. Now let $x \in span(\{u_{k+1} + W, ..., u_n + W\}$. Then:

$$x = \sum_{i=k+1}^{n} a_i(u_i + W) = \sum_{i=k+1}^{n} a_i u_i + W = v + W$$

Where $v = \sum_{k+1}^{n} a_i u_i \in V \implies v + W \in V/W$. Thus $span(\{u_{k+1} + W, ..., u_n + W\} \subseteq V/W$. Hence the span of the set generates V/W. Now we must prove that this set is linearly independent.

$$\sum_{i=k+1}^{n} a_i(u_i + W) = \sum_{k+1}^{n} a_i u_i + W = 0 + W \iff \sum_{i=k+1}^{n} a_i u_i \in W$$

However none of the u_i for $k + 1 \leq i \leq n$ are in the basis set for W, and hence all of the $a_i = 0$. Therefore the set is linearly independent. Hence the set $\{u_{k+1} + W, ..., u_n + W\}$ is a basis for the susbpace V/W. ■

b.) We note that $\dim(V) = n$ and that $\dim(W) = k$ from the bases given in the problem. Then we see that $\dim(V/W) = n - k$, and so we can see that:

$$\dim(V/W) = \dim(V) - \dim(W)$$

Is a formula relating the dimension quotient space of V modulo W to the dimension of V and W.

1.7 Maximal Linearly Independent Subsets

Problem 1.7.1
a.) False. The maximal principle states that a family of sets contains a maximal element $\iff$ for each chain in the family there is a member of the family that contains all members of the collection.

b.) False. Not necessarily.

c.) False. Refer to example 5 in the section.

d.) True.

e.) True. Refer to the discussion above Theorem 1.12.

f.) True. Refer to Theorem 1.12.

Problem 1.7.2

Problem 1.7.3
Consider the set $\mathbb{R}$ as a vector space over the field $\mathbb{Q}$. Now suppose by way of contradiction that $\mathbb{R}$ is a finite-dimensional vector space. Now consider $\pi \in \mathbb{R}$. Note that the set $\{1, \pi, \pi^2, ...\}$ is linearly independent over $\mathbb{Q}$, since:

$$\sum_{i=0}^{\infty} q_i \pi^i \neq 0$$

For any $q_i \in \mathbb{Q}$. This is evident since π is a transcendental number and thus is not the root of any polynomials with coefficients in $\mathbb{Q}$. Hence there exists an infinite linearly independent subset of $\mathbb{R}$, contradicting the finite dimensionality previously asserted. Therefore the vector space $\mathbb{R}$ over $\mathbb{Q}$ is infinite-dimensional. $\blacksquare$

Problem 1.7.4
Let W be a subspace of a vector space V. Suppose that β is a basis for W. Then $\beta \subseteq V$, and β is necessarily a linearly independent subset of V. Hence by Theorem 1.13, there exists a maximal linearly independent subset of V that contains β. We know from Theorem 1.12 that this maximal linearly independent subset of V is a basis for V, and hence β is contained in a basis for V as desired. $\blacksquare$

Problem 1.7.5 XX
Let β be a subset of an infinite-dimensional vector space V. Now suppose that β is a basis for V. Let $\beta = \{u_1, u_2, ...\}$ and let $v \in V$, so by definition of a basis,

we have that for $a_i \in \mathbb{F}$:

$$v = \sum_{i=1}^{\infty} a_i u_i$$

Problem 1.7.6

Problem 1.7.7

2 Linear Transformations and Matrices

2.1 Linear Transformations, Null Spaces, and Ranges

Problem 2.1.1
a.) True. This is the definition of a linear transformation.

b.) False. T must also satisfy $T(cx) = cT(x)$ for $c \in \mathbb{K}$.

c.) False. T must also be linear, as to satisfy hypothesis for Theorem 2.4.

d.) True. Refer to Theorem 2.4.

e.) False. It should be $\dim(V)$, since $T : V \to W$ by dimension theorem.

f.) False. Requires T to be injection.

g.) True. Refer to the corollary to Theorem 2.6.

h.) False.

Problem 2.1.2
Consider $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $T(a, b, c) = (a - b, 2c)$. Let $x, y \in \mathbb{R}^3$ such that $x = (a, b, c)$ and $y = (\alpha, \beta, \gamma)$, and let $\mu \in \mathbb{K}$. Then:

$$T(\mu a + \alpha, \mu b + \beta, \mu c + \gamma) = (\mu a + \alpha - \mu b - \beta, 2\mu c + 2\gamma) =$$

$$((\mu a - \mu b) + (\alpha - \beta), 2\mu c + 2\gamma) = (\mu a - \mu b, 2\mu c) + (\alpha - \beta, 2\gamma) =$$

$$\mu(a - b, 2c) + (\alpha - \beta, 2\gamma) = \mu T(x) + T(y)$$

And hence T is a linear transformation. Now note that $T(a, b, c) = \vec{0}$ when $a - b = 0$ and when $2c = 0 \implies a = b$ and $c = 0$. Therefore $N(T) = \{(a, a, 0) \in \mathbb{R}^3 : a \in \mathbb{R}\}$, so a basis for this set is given by $\{(1, 1, 0)\}$. By Theorem 2.2, we see that since $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ is a basis for $\mathbb{R}^3$, we have:

$$R(T) = span(\{(1, 0), (-1, 0), (0, 2)\})$$

And hence a basis for $R(T)$ is given by the set $\{(1, 0), (0, 1)\}$. The dimension theorem holds since $\dim(\mathbb{R}^3) = 3 = \dim(N(T)) + \dim(R(T)) = 1 + 2$. Clearly T is not injective since $N(T) \neq \{\vec{0}\}$. T is surjective, however, since $R(T) = \mathbb{R}^2$.

Problem 2.1.3
Consider $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(a, b) = (a + b, 0, 2a - b)$. Let $x, y \in \mathbb{R}^2$ such that $x = (a, b)$ and $y = (\alpha, \beta)$, and let $\mu \in \mathbb{K}$. Then:

$$T(\mu x + y) = T((\mu a + \alpha, \mu b + \beta)) = (\mu a + \alpha + \mu b + \beta, 0, 2\mu a + 2\alpha - \mu b - \beta) =$$

$$([\mu a + \mu b] + [\alpha + \beta], 0, [2\mu a - \mu b] + [2\alpha - \beta]) = \mu(a + b, 0, 2a - b) + (\alpha + \beta, 0, 2\alpha - \beta) =$$

$$\mu T(x) + T(y)$$

So that T is linear. Note that $T((a,b)) = \vec{0}$ when $a+b=0$ and $2a-b=0$ $\implies a=-b$ and $2a=b$, but then $2(-b)=b \implies -2=1$, and thus $a=b=0$. Hence $N(T)=\{\vec{0}\}$. Using Theorem 2.2, we see that since $\{(1,0),(0,1)\}$ is a basis for $\mathbb{R}^2$, then:

$$R(T) = span(\{(1,0,2),(1,0,-1)\})$$

And hence the set $\{(1,0,2),(1,0,-1)\}$ is a basis for $R(T)$. The dimension theorem is verified in that $\dim(\mathbb{R}^2) = 2 = \dim(N(T)) + \dim(R(T)) = 0+2$. By Theorem 2.4, T is injective. T is not surjective since $R(T) \neq \mathbb{R}^3$.

Problem 2.1.4

Consider $T : M_{2\times 3}(\mathbb{K}) \to M_{2\times 2}(\mathbb{K})$ defined by:

$$T(\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}) = \begin{pmatrix} 2a-b & c+2b \\ 0 & 0 \end{pmatrix}$$

Let $A, B \in M_{2\times 3}(\mathbb{K})$ such that:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}, \; B = \begin{pmatrix} \alpha & \beta & \gamma \\ \delta & \epsilon & \phi \end{pmatrix}$$

And let $\mu \in \mathbb{K}$. Now we will show that T is linear.

$$T(\mu A + B) = T(\begin{pmatrix} \mu a+\alpha & \mu b+\beta & \mu c+\gamma \\ \mu d+\delta & \mu e+\epsilon & \mu f+\phi \end{pmatrix}) =$$

$$\begin{pmatrix} 2\mu a + 2\alpha - \mu b - \beta & \mu c + \gamma + 2\mu b + 2\beta \\ 0 & 0 \end{pmatrix} =$$

$$\begin{pmatrix} [2\mu a - \mu b] + [2\alpha - \beta] & [\mu c + 2\mu b] + [\gamma + 2\beta] \\ 0 & 0 \end{pmatrix}$$

$$= \mu \begin{pmatrix} 2a-b & c+2b \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 2\alpha-\beta & \gamma+2\beta \\ 0 & 0 \end{pmatrix} = \mu T(A) + T(B)$$

And thus T is linear. Note that $T(A) = O$ only when $2a-b=0$ and $c+2b=0$ $\implies 2a=b$ and $2b=-c$. Therefore we can see that:

$$N(T) = \{\begin{pmatrix} a & 2a & -4a \\ 0 & 0 & 0 \end{pmatrix} \in M_{2\times 3}(\mathbb{K}) : a \in \mathbb{K}\}$$

And thus a basis for the null space is easily seen to be:

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

Now we aim to find a basis for $R(T)$. Using Theorem 2.2, we find that:

$$R(T) = span(\{\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\}) =$$

$$= span(\{\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\})$$

So we find that a basis for $R(T)$ is given by the set:

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

Thus we can see that teh dimension theorem is satisfied, as $\dim(M_{2\times 3}(\mathbb{K})) = 6 = \dim(N(T)) + \dim(R(T)) = 4 + 2$. By Theorem 2.3, T is not injective since $\dim(N(T)) \neq 0$. Additionally, T is clearly not surjective as $R(T) \neq M_{2\times 2}(\mathbb{K})$.

Problem 2.1.5 XX
Now consider $T : P_2(\mathbb{R}) \to P_3(\mathbb{R})$ defined by $T(f(x)) = xf(x) + f'(x)$. Let $f(x), g(x) \in P_2(\mathbb{R})$ and $\mu \in \mathbb{K}$. Then:

$$T((\mu f+g)(x)) = x(\mu f+g)(x)+(\mu f+g)'(x) = x\mu f(x)+xg(x)+(\mu f)'(x)+g'(x)$$

$$= \mu(xf(x)) + \mu f'(x) + xg(x) + g'(x) = \mu T(f(x)) + T(g(x))$$

By the linearity properties of the derivative. Hence T is a linear transformation. Now see that $T(f(x)) = 0 \implies xf(x) + f'(x) = 0 \implies xf(x) = -f'(x)$

Problem 2.1.6 XX
Consider $T : M_n(\mathbb{K}) \to \mathbb{K}$ defined by $T(A) = \mathrm{tr}\,(A)$. Let $A, B \in M_n(\mathbb{K})$ and let $\mu \in \mathbb{K}$. Then

$$T(\mu A + B) = \mathrm{tr}\,(\mu A + B) = \mu\,\mathrm{tr}\,(A) + \mathrm{tr}\,(B) = \mu T(A) + T(B)$$

By the results shown in Problem 1.3.6. Hence T is a linear transformation. Now observe that $T(A) = 0 \implies \mathrm{tr}\,(A) = 0$

Problem 2.1.7
Let V, W be vector spaces over $\mathbb{K}$ and let $T : V \to W$ be a linear transformation.

1.) Suppose T is linear. Then since $0_V \in V$, $T(0_V + 0_V) = T(0_V) + T(0_V)$, but also $T(0_V + 0_V) = T(0_V)$, so then we have that $T(0_V) = T(0_V) + T(0_V) \implies 0_W = T(0_V)$ subtracting $T(0_V)$ from both sides of the previous equation. ■

2.) Suppose T is linear. Then $T(cx + y) = T(cx) + T(y) = cT(x) + T(y)$, first using property a.) and then property b.) from the definition of a linear transformation. Conversely, suppose that $T(cx + y) = cT(x) + T(y)$ for all $x, y \in V$ and $c \in \mathbb{K}$. Then choosing $c = 1 \in \mathbb{K}$ as the multiplicative identity, we find that $T(x + y) = T(x) + T(y)$, and so the first property is shown. For the second, choose $y = \vec{0}_V$. Then $T(cx) = cT(x)$ and so the second property is recovered. Hence T is a linear transformation by definition. ■

3.) Suppose T is a linear transformation. Let $x, y \in V$. Then we know that

$T(x+(-y)) = T(x)+T(-y) = T(x)+T((-1)y) = T(x)-T(y)$. Since x, y were arbitrary, this holds for all $x, y \in V$. ■

4.) Suppose T is linear. For our base case, $n = 1$, we see that by definition $T(a_1x_1 + a_2x_2) = a_1T(x_1) + a_2T(x_2)$ for $a_1, a_2 \in \mathbb{K}$ and $x_1, x_2 \in V$. Suppose that the hypothesis holds for the nth case. Then

$$T(\sum_{i=1}^{n} a_i x_i) = \sum_{i=1}^{n} a_i T(x_i)$$

Adding $T(a_{n+1}x_{n+1})$ to both sides, we see that:

$$T(\sum_{i=1}^{n} a_i x_i) + T(a_{n+1}x_i) = \sum_{i=1}^{n+1} a_i T(x_i)$$

And thus the inductive hypothesis holds for the $n + 1$th case. Therefore the relation holds for any $x_1, ..., x_n \in V$ and $a_1, ..., a_n \in \mathbb{K}$. The proof of the sufficiency is trivial and consists of the relation obeying the properties listed in the definition for a linear transformation. ■

Problem 2.1.8
For the transformation in Example 2, we have $T_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T_\theta(a, b) = (a\cos\theta - b\sin\theta, a\sin\theta + b\cos\theta)$. Let $x, y \in \mathbb{R}^2$ such that $x = (a, b)$ and $y = (\alpha, \beta)$, and let $\mu \in \mathbb{K}$. Then:

$$T_\theta(\mu x + y) = T_\theta((\mu a + \alpha, b + \beta)) =$$

$$([\mu a + \alpha]\cos\theta - [b + \beta]\sin\theta, [\mu a + \alpha]\sin\theta + [b + \beta]\cos\theta) =$$

$$\mu(a\cos\theta - b\sin\theta, a\sin\theta + b\cos\theta) + (\alpha\cos\theta - \beta\sin\theta, \alpha\sin\theta + \beta\cos\theta) =$$

$$\mu T_\theta((a, b)) + T_\theta((\alpha, \beta)) = \mu T_\theta(x) + T_\theta(y)$$

And hence the transformation of rotation by θ in Example 2 is linear.

Now for Example 3, we have $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T((a, b)) = (a, -b)$. Let $x, y \in \mathbb{R}^2$ be as above, and also let $\mu \in \mathbb{K}$. Then we see that:

$$T(\mu x + y) = T((\mu a + \alpha, \mu b + \beta)) = (\mu a + \alpha, -[\mu b + \beta]) =$$

$$(\mu a + \alpha, -\mu b - \beta) = \mu(a, -b) + (\alpha, -\beta) = \mu T(x) + T(y)$$

And thus again we find that T in Example 3 is a linear transformation by definition. ■

Problem 2.1.9
a.) $T((0, 0)) = (1, 0) \neq (0, 0)$, and therefore T doesn't map 0 to 0 and cannot be linear.

b.) $T(a+\alpha, b+\beta) = (a+\alpha, [b+\beta]^2) = (a+\alpha, b^2+2b\beta+\beta^2) \neq T(a,b)+T(\alpha,\beta)$, and therefore condition 1 of the definition of a linear transformation does not hold.

c.) $T(a+\alpha, b+\beta) = (\sin a + \alpha, 0) \neq (\sin a, 0) + (\sin \alpha, 0)$. Thus again condition 1 does not hold. (sine function does not have linearity properties).

d.) Choose negative scalar and then note that $T(\mu a, \mu b) = \mu(|a|, b)$ doesn't necessarily hold. Condition 2 fails.

e.) $T(cx) = T(ca, cb) = (ca+1, cb) \neq c(a+1, b)$. Again condition 2 fails to hold.

Problem 2.1.10
Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is linear and $T(1,0) = (1,4)$ and $T(1,1) = (2,5)$. Then note that:

$$T(2,3) = 3T(1,1) - T(1,0) = 3(2,5) - (1,4) = (6-1, 15-4) = (5,11)$$

To find if T is injective, see that $T(2,3) - 2T(1,1) = T(0,1) = (5,11) - 2(2,5) = (5-4, 11-10) = (1,1)$. Using Theorem 2.2, we see that:

$$R(T) = span(\{T(1,0), T(0,1)\} = span(\{(1,4),(1,1)\})$$

And $\{(1,4),(1,1)\}$ is clearly a linearly independent set, which contains 2 vectors and thus is a basis for $\mathbb{R}^2$. Hence we see that $R(T) = \mathbb{R}^2$ and thus by Theorem 2.5, T is injective.

Problem 2.1.11
Note that since $T(1,1) = (1,0,2)$ and $T(2,3) = (1,-1,4)$, we find:

$$T(8,11) = 2T(1,1) + 3T(2,3) = 2(1,0,2) + 3(1,-1,4) =$$

$$(2,0,4) + (3,-3,12) = (5,-3,16)$$

To prove that such a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ exists, note that the set $\{(1,1),(2,3)\}$ is linearly independent and spans $\mathbb{R}^2$. By Theorem 2.6 such a T must exist. $\blacksquare$

Problem 2.1.12
Note that the set $\{(1,0,3),(-2,0,-6)\}$ is linearly dependent and thus cannot be a basis. Therefore we cannot apply Theorem 2.6 to assert the existence of a such a $T : \mathbb{R}^2 \to \mathbb{R}^2$. $\blacksquare$

Problem 2.1.13
Let V and W be vector spaces, let $T : V \to W$ be linear, and let $\{w_1, ..., w_k\}$ be a linearly independent subset of $R(T)$. Suppose that $S = \{v_1, ..., v_k\} \subseteq V$ is chosen such that $T(v_i) = w_i$ for $1 \leq i \leq k$. Then the only linear combination of

the vectors w_i to make the zero vecto is the trivial combination, so $a_i = 0$ for $1 \leq i \leq k$ and we see that since T is linear, so $T(0) = 0$, we see:

$$\vec{0} = \sum_{i=1}^{k} a_i w_i = \sum_{i=1}^{k} a_i T(v_i) = T(\sum_{i=1}^{k} a_i v_i) \implies \sum_{i=1}^{k} a_i v_i = \vec{0}$$

And this is the only representation of the zero vector, since $\{w_1, ..., w_k\}$ is linearly independent. Hence the only linear combination of the set S is the trivial combination, and so by definition S is a linearly independent subset of V as desired. ■

Problem 2.1.14
Let V and W be vector spaces and $T : V \to W$ be linear.

a.) Suppose that T is injective. Let $S = \{v_1, ..., v_k\} \subseteq V$ be a linearly independent set. Then note that, since $T(0) = 0$,

$$\sum_{i=1}^{k} a_i v_i = \vec{0} \implies T(\sum_{i=1}^{k} a_i v_i) = \sum_{i=1}^{k} a_i T(v_i) = \vec{0}$$

So that all $a_i = 0$ for $1 \leq i \leq k$, and hence the set $\{T(v_1), ..., T(v_k)\} \subseteq W$ is linearly independent by definition. Conversely, suppose that T carries linearly independent subsets of V onto linearly independent subsets of W. Now suppose by way of contradiction that $x \in V$, with $x \neq \vec{0}$ such that $T(x) = 0 \in W$. Note that the set $\{x\} \subseteq V$ is trivially linearly independent. Thus $\{T(x)\} = \{\vec{0}\}$ must be linearly independent as per the hypothesis, but this is a contradiction because any set containing the zero vector is linearly dependent. Hence the set $\{x\}$ must be linearly dependent, and the only set containing one vector that is linearly dependent is $\{\vec{0}\}$. Thus $x = \vec{0}$ and we see that $N(T) = \{\vec{0}\}$, so that by Theorem 2.4 T is injective. ■

b.) Let T be injective and $S \subseteq V$. Suppose that S is a linearly independent set. By part a.) of this problem, T carries linearly independent subsets of V onto linearly independent subsets of W. Hence $T(S)$ must also be linearly independent. Conversely, suppose that $T(S)$ is linearly independent. Then again by part a.), since T is injective, S is linearly independent. ■

c.) Suppose that $\beta = \{v_1, ..., v_n\}$ is a basis for V and that T is injective and surjective. By definition of surjectivity, we know that $R(T) = W$. Further, by Theorem 2.2 we have that:

$$R(T) = W = span(T(\beta)) = span(\{T(v_1), ..., T(v_n)\})$$

So that the set $T(\beta)$ generates W. Then, by part b.) of this problem, since T is injective, and β is by definition a linearly independent subset of V, we know that $T(\beta)$ is also a linearly independent set. Therefore the set $T(\beta)$ spans W

and is linearly independent and is thus a basis for W as desired. ■

Problem 2.1.15
Consider the vector space of polynomials with coefficients in the reals. Define $T : P(\mathbb{R}) \to P(\mathbb{R})$ by $T(f(x)) = \int_0^x f(t)dt$. A basis for $P(\mathbb{R})$ is given by the set $\{1, x, x^2, ...\}$ (recall Example 5 of section 1.6). Applying Theorem 2.2, we find that

$$R(T) = span(\{T(1), T(x), T(x^2), ...\}) = span(\{x, x^2/2, ...\})$$

And hence $R(T) \neq P(\mathbb{R})$ since $1 \notin span(T(\beta))$. Therefore T is not surjective. Additionally, since the only polynomial $f(x)$ such that $\int_0^x f(t)dt = 0$ is $f(x) = 0$, the zero polynomial. Hence $N(T) = \{0\} \implies T$ is injective by Theorem 2.4. ■

Problem 2.1.16
Again consider the vector space of polynomials with coefficients in $\mathbb{R}$. Define $T : P(\mathbb{R}) \to P(\mathbb{R})$ by $T(f(x)) = f'(x)$. This transformation is linear by Example 6 in Section 2.1. Let $\beta = \{1, x, x^2, ...\}$ be a basis for $P(\mathbb{R})$. Then applying Theorem 2.2, we see:

$$R(T) = span(T(\beta)) = span(\{T(1), T(x), T(x^2), T(x^3), ...\}) =$$

$$span(\{0, 1, 2x, 3x^2, ...\}) = span(\{1, x, x^2, ..\}) = P(\mathbb{R})$$

Hence T is surjective. Clearly $N(T) \neq \{0\}$ since any constant polynomial maps to 0 under T. Therefore T is not injective. ■

Problem 2.1.17
Let V and W be finite-dimensional vector spaces and $T : V \to W$ be linear.

a.) Suppose that $\dim(V) < \dim(W)$. By the Dimension Theorem, we have then that $\dim(N(T)) + \dim(R(T)) < \dim(W)$. Now suppose that T is surjective, so that by definition $R(T) = W \implies \dim(R(T)) = \dim(W) \implies \dim(N(T)) < 0$ after subtraction from the above equation. This is a contradiction as dimension cannot be negative. Hence T cannot be surjective. ■

b.) Now suppose that $\dim(V) > \dim(W)$. Then again from Theorem 2.3 we have $\dim(N(T)) + \dim(R(T)) > \dim(W)$. Now suppose by way of contradiction that T is injective, so by Theorem 2.4 $\dim(N(T)) = 0$. Then the above equation becomes $\dim(R(T)) > \dim(W)$, but since $R(T)$ is a subspace of W, this is a contradiction as per Theorem 1.11. Hence T cannot be injective. ■

Problem 2.1.18 XX
We would like to find a linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $N(T) = R(T)$. Consider the example of projection onto the x-axis. Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T((a, b)) = (a, 0)$. Note that $N(T) = \{(0, b) \in \mathbb{R}^2 : b \in \mathbb{R}\}$ and that $R(T) = \{(a, 0) \in \mathbb{R}^2 : a \in \mathbb{R}\}$

Problem 2.1.19

Problem 2.1.20
Let V and W be vector spaces with subspaces $V_1 \subseteq V$ and $W_1 \subseteq W$. Suppose now that $T : V \to W$ is linear. Consider the set $T(V_1) = \{T(v) \in W : v \in V_1\} \subseteq W$. Note that $\vec{0}_W \in T(V_1)$ since $\vec{0}_V \in V_1$ by subspace conditions. Now let $T(v), T(u) \in T(V_1)$ and $\mu \in \mathbb{K}$. Then since $\mu v + u \in V_1$ by subspace closure, then necessarily $T(\mu v + u) = \mu T(v) + T(u) \in T(V_1)$ by linearity of the transformation. Hence $T(V_1)$ is closed under vector addition and scalar multiplication and contains the zero vector, and is thus a subspace of W. Now consider the set $T^{-1}(W_1) = \{v \in V : T(x) \in W_1\} \subseteq V$. Since W_1 is a subspace, $\vec{0}_W \in W_1$, and thus $\vec{0}_V \in T^{-1}(W_1)$ trivially. Now let $v, u \in T^{-1}(W_1)$ and $\mu \in \mathbb{K}$. Then $\mu T(v) + T(u) = T(\mu v + u) \in W_1 \implies \mu v + u \in T^{-1}(W_1)$, so that the set satisfies all conditions in Theorem 1.3 for identity as a subspace of V. Thus we find that the image of a subspace under a linear transformation is again a subspace, and that the preimage of a subspace is again a subspace. ■

Problem 2.1.21
Define the transformations $T, U : V \to V$ by:

$$T(a_1, a_2, \ldots) = (a_2, a_3, \ldots) \ , \ U(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots)$$

a.) Let $(a_1, a_2, \ldots), (b_1, b_2, \ldots) \in V$ and $\mu \in \mathbb{K}$. Then we see that, as per coordinatewise addition of sequences defined in Example 5 of section 1.2,

$$T(\mu a_1 + b_1, \mu a_2 + b_2, \ldots) = (\mu a_2 + b_2, \mu a_3 + b_3, \ldots) =$$

$$\mu(a_2, a_3, \ldots) + (b_2, b_3, \ldots) = \mu T(a_1, a_2, \ldots) + T(b_1, b_2, \ldots)$$

And hence the left shift operator T is linear. Now, using the same vector sequences and scalar as above, we find that:

$$U(\mu a_1 + b_1, \mu a_2 + b_2, \ldots) = (0, \mu a_1 + b_1, \mu a_2 + b_2, \ldots) =$$

$$\mu(0, a_1, a_2, \ldots) + (0, b_1, b_2, \ldots) = \mu U(a_1, a_2, \ldots) + U(b_1, b_2, \ldots)$$

And thus the right shift operator U is linear. ■

b.) We will show that T is surjective but not injective. Let the sequences $(a_1, a_2, \ldots), (b_1, b_2, \ldots) \in V$. Suppose that

$$T(a_1, a_2, \ldots) = T(b_1, b_2, \ldots) \implies (a_2, a_3, \ldots) = (b_2, b_3, \ldots)$$

However we can see that $a_1 = b_1$ does not necessarily follow from the above hypothesis. We could easily choose $a_1 \neq b_1$ and then find that their images under T are equal, and hence T cannot be injective. Now for surjectivity, suppose $(c_1, c_2, \ldots) \in V$. Then there exists $(a, c_1, c_2, \ldots) \in V$, where $a \in \mathbb{K}$, such that

$T(a, c_1, c_2, \ldots) = (c_1, c_2, \ldots)$. Therefore T is surjective. ■

c.) We will show that U is injective but not surjective. Let the sequences $(a_1, a_2, \ldots), (b_1, b_2, \ldots) \in V$. Now suppose that:

$$U(a_1, a_2, \ldots) = U(b_1, b_2, \ldots) \implies (0, a_1, a_2, \ldots) = (0, b_1, b_2, \ldots)$$

And from this we can see that $a_i = b_i$ for all i necessarily, which implies that $(a_1, a_2, \ldots) = (b_1, b_2, \ldots) \implies U$ is injective. Now for surjectivity, let $(c_1, c_2, \ldots) \in V$. Then there does not exist a vector $(a_1, a_2, \ldots) \in V$ such that $T(a_1, a_2, \ldots) = (c_1, c_2, \ldots)$ since $T(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots)$. Hence U is not surjective. ■

Problem 2.1.22

Problem 2.1.23

Problem 2.1.24
a.) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$. If T represents projection on the y-axis along the x-axis, then a formula for T is given by: $T(a, b) = (a, 0)$.

b.) If T now represents a projection on the y-axis along the line $L = \{(s, s) : s \in \mathbb{R}\}$, then a formula for T is given by $T(a, b) = (b, b)$.

Problem 2.1.25

Problem 2.1.26

Problem 2.1.27

Problem 2.1.28
Let V be a vector space and W be a subspace of V. Let $T : V \to V$ be linear. Consider the subspace $\{\vec{0}\}$. Then $T(\vec{0}) = \vec{0} \subseteq \{\vec{0}\}$ trivially, and hence the zero subspace is T-invariant. Similarly, for any $v \in V$, we see that $T(v) = v' \in V \implies T(V) \subseteq V$, so that V is T-invariant. Now consider the subspace $R(T) \subseteq V$. Let $x \in R(T)$, so then $x = T(v) \subseteq R(T)$ for some $v \in V$, and so $T(R(T)) \subseteq R(T)$ and $R(T)$ is a T-invariant subspace. Finally, consider the subspace $N(T)$. Let $x \in N(T)$, so that $T(x) = \vec{0} \in V$, and since $\vec{0} \in N(T)$ trivially, $T(v) \in N(T) \implies T(N(T)) \subseteq N(T)$ and so the null space is T-invariant. ■

Problem 2.1.29

Let V be vector space and W a subspace of V, and let $T : V \to V$ be linear. Suppose now that W is T-invariant. Then the restriction of T on W is teh map $T_W : W \to W$ defined by $T_W(x) = T(x)$, for all $x \in W$. We will prove that T_W is linear. First, let $v, w \in W$ and $\lambda \in \mathbb{K}$. Then we see that:

$$T_W(\lambda v + w) = T(\lambda v + w) = \lambda T(v) + T(w)$$

following from the linearity of T. Hence T_W is a linear transformation. ■

Problem 2.1.30
Let V be a vector space with subspace W and $T : V \to V$ be linear. Suppose that T is the projection on W along a subspace $W' \subseteq V$. Then we know by definition that for $x \in V$, $x = x_1 + x_2$ where $x_1 \in W$ and $x_2 \in W'$, we have $T(x) = x_1$. Let $y \in W$. Then $y = y + \vec{0}$, where $\vec{0} \in W'$, and so:

$$T(y) = T(y + \vec{0}) = y \in W$$

And since $y \in W$ was arbitrary, we necessarily have that $T(W) \subseteq W$ so that W is a T-invariant subspace. Notice also that for $T_W : W \to W$, letting $w \in W$, we see that $T_W(w) = T(w) = w$, and hence $T_W = I_W$, T maps all vectors in W to themselves and is hence the identity transformation for W. ■

Problem 2.1.31
Let V be a vector space with subspace W and $T : V \to V$ be linear. Suppose that $V = R(T) \oplus W$ and W is T-invariant.

a.) We know that $R(T) \cap W = \{\vec{0}\}$ by definition of direct sum. Let $x \in W$, then $T(x) = \vec{0}$ necessarily, since the only vector that is an image of a vector in W under T is the zero vector. Hence $x \in N(T)$, and since $x \in W$ was arbitrary, we have that $W \subseteq N(T)$. ■

b.) Now suppose that V is finite-dimensional. By Problem 1.6.29, since we are given that $V = R(T) \oplus W$, then $\dim(V) = \dim(R(T)) + \dim(W)$. However, by the dimension theorem, we know that $\dim(V) = \dim(R(T)) + \dim(N(T))$, and therefore necessarily $\dim(N(T)) = \dim(W)$. But since $W \subseteq N(T)$ by the first part of the problem, and W is a subspace of V, then W is a subspace of $N(T)$, and so by Theorem 1.11, $W = N(T)$ as desired. ■

c.) XXXX

Problem 2.1.32
Let V be a vector space with subspace W and let $T : V \to V$ be linear. Suppose that W is T-invariant. Then we know by Problem 2.1.29 that T_W is linear. Note that $N(T_W) = \{w \in W : T_W(w) = \vec{0}\}$. Let $x \in N(T_W)$, so $T_W(x) = \vec{0}$ and $x \in W \implies x \in N(T) \cap W$. Hence $N(T_W) \subseteq N(T) \cap W$. For the reverse containment, let $x \in N(T) \cap W \implies x \in W$ and $T(x) = \vec{0}$, but $T(x) = T_W(x) = \vec{0}$ since $x \in W$, so $x \in N(T_W)$. Hence $N(T_W) = N(T) \cap W$ as desired. Now note

that the set $R(T_W) = \{T_W(w) \in W : T_W(w) = v \in W\}$. Let $x \in R(T_W) \implies T_W(x) = y \in W \implies x \in T(W)$. Hence $R(T_W) \subseteq T(W)$. Now let $x \in T(W)$, so that $x = T(y)$ for some $y \in W$. But then $x = T_W(y) \implies x \in R(T_W)$. Therefore $R(T_W) = T(W)$ as desired. ■

Problem 2.1.33

Problem 2.1.34

Problem 2.1.35
Let V be a finite-dimensional vector space and $T : V \to V$ be linear.

a.) Suppose that $V = R(T) + N(T)$. Let $x \in R(T) \cap N(T)$, so $x = T(y)$ for some $y \in V$ and $T(x) = \vec{0}$. But then $T(x+y) = T(x) + T(y) = \vec{0} + T(y) = T(y) \implies x + y = y \implies x = \vec{0}$, and hence $R(T) \cap N(T) = \{\vec{0}\}$. Therefore $V = R(T) \oplus N(T)$. ■

b.) Suppose that $R(T) \cap N(T) = \{\vec{0}\}$, so $\dim(R(T) \cap N(T)) = 0$. Then by Problem 1.6.29, followed by the dimension theorem, we know:

$$\dim(R(T) + N(T)) = \dim(R(T)) + \dim(N(T)) - \dim(R(T) \cap N(T)) =$$

$$\dim(R(T) + N(T)) = \dim(R(T)) + \dim(N(T))$$
$$\implies \dim(V) = \dim(R(T) + N(T))$$

And since $R(T)+N(T)$ is a subspace of V by Problem 1.3.22, we can apply Theorem 1.11 and see that $\dim(V) = \dim(R(T) + N(T)) \implies V = R(T) + N(T)$. Hence $V = R(T) + N(T)$ and $R(T) \cap N(T) = \{\vec{0}\} \implies V = R(T) \oplus N(T)$ as desired. ■

Problem 2.1.36

Problem 2.1.37
Let V and W be vector spaces over common field $\mathbb{Q}$. Suppose $T : V \to W$ is an additive mapping. Then we know by the definition of additive function that $T(x + y) = T(x) + T(y)$ for all $x, y \in V$. Now let $q = a/b \in \mathbb{Q}$, and then see that:

$$T(qx) = T(\frac{a}{b}x) = T(\underbrace{\frac{1}{b}x + \cdots + \frac{1}{b}x}_{\times a}) = \underbrace{T(\frac{1}{b}x) + \cdots T(\frac{1}{b}x)}_{\times a} = aT(\frac{1}{b}x) = \frac{a}{b}T(x)$$

And hence $T(qx) = qT(x)$ for all $x \in V$ and $q \in \mathbb{Q}$, and so T satisfies the two conditions for identification as a linear transformation. ■

Problem 2.1.38
Let $T : \mathbb{C} \to \mathbb{C}$ be a function defined by $T(z) = \overline{z}$. Letting $z, w \in \mathbb{C}$, we see that:

$$T(z + w) = \overline{z + w} = \overline{z} + \overline{w} = T(z) + T(w)$$

by properties of complex numbers. Hence by definition T is an additive function. Note however that:

$$T(1 \cdot i) = \overline{1 \cdot i} = \overline{i} \cdot \overline{1} = -i \cdot \overline{1} = -i \cdot -1 = i \neq iT(1) = i \cdot \overline{1} = i \cdot -1 = -i$$

And hence T cannot be a linear transformation because it fails condition 2 regarding scalars. ■

Problem 2.1.39
Consider the vector space $\mathbb{R}$ over the field $\mathbb{Q}$. We know that $\mathbb{R}$ has a basis, since every vector space has a basis by Theorem 1.13. Let β be a basis for $\mathbb{R}$ and let $x, y \in \beta$ such that $x \neq y$. Define a mapping $f : \beta \to V$ by $f(x) = y$ and $f(y) = x$. By the results shown in Problem 2.1.34, we know that there exists a linear transformation $T : V \to V$ such that $T(u) = f(u)$ for all $u \in \beta$. Then note that $T(x+y) = T(x)+T(y) = f(x)+f(y) = x+y$, and so f is an additive function. Consider $c = y/x$, then see that $T(cx) = T(y) \neq cT(x) = y^2/x$. Therefore f is not a linear function because it fails condition 2.

Problem 2.1.40
Let V be a vector space over $\mathbb{K}$ and W a subspace of V. Define the mapping $\eta : V \to V/W$ by $\eta(v) = v + W$ for $v \in V$.

a.) Let $x, y \in V$ and $\lambda \in \mathbb{K}$. Then $\lambda x + y \in V$ and we see that by the rules for addition of cosets shown in Problem 1.3.31, we have:

$$\eta(\lambda x + y) = (\lambda x + y) + W = (\lambda x + W) + (y + W) =$$

$$\lambda(x + W) + (y + W) = \lambda\eta(x) + \eta(y)$$

And hence by definition η is a linear transformation. Now note that the subspace $N(\eta) = \{v \in V : v + W = 0 + W\} = W$ since $v + W = 0 + W = W \iff v \in W$ by Problem 1.3.31. ■

b.) Suppose that V is finite-dimensional. Now apply the dimension theorem, and we find that since $N(\eta) = W \implies \dim(N(\eta)) = \dim(W)$, so:

$$\dim(V) = \dim(N(\eta)) + \dim(R(\eta)) = \dim(W) + \dim(R(\eta))$$

However we note now that η is surjective, since if $v + W \in V/W$, then $v \in V$ necessarily such that $\eta(v) = v + W \implies R(\eta) = V/W \implies \dim(R(\eta)) = \dim(V/W)$. Hence the above equation then becomes:

$$\dim(V) - \dim(W) = \dim(V/W)$$

Which is the same formula obtained in Problem 1.6.35 as desired. ■

c.) The results obtained for the formula relating the dimension of the vector space, its subspace, and the quotient space of this subspace were the same as in Problem 1.6.35. The method, however, varied considerably as this time we approached the situation in regards to the linear transformation, and used properties of that linear transformation to make conclusions about the relationships.

2.2 The Matrix Representation of a Linear Transformation

Problem 2.2.1
a.) True. This follows from the definition on page 82.

b.) True, matrix representations equal implies transformations equal.

c.) False. $[T]_\beta^\gamma$ is an $n \times m$ matrix since $T : V \to W$.

d.) True. This is Theorem 2.8.

e.) True. By definition on page 82.

f.) False. $T : V \to W \neq T : W \to V$.

Problem 2.2.2
a.)

$$[T]_\beta^\gamma = \begin{pmatrix} 2 & -1 \\ 3 & 4 \\ 1 & 0 \end{pmatrix}$$

b.)

$$[T]_\beta^\gamma = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 0 & 1 \end{pmatrix}$$

c.)

$$[T]_\beta^\gamma = \begin{pmatrix} 2 & 1 & -3 \end{pmatrix}$$

d.)

$$[T]_\beta^\gamma = \begin{pmatrix} 0 & 2 & 1 \\ -1 & 4 & 5 \\ 1 & 0 & 1 \end{pmatrix}$$

e.)

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & & \vdots \\ \vdots & & & \ddots & 0 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

f.)

$$[T]_\beta^\gamma = \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 0 & 0 & & \\ \vdots & 1 & & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}$$

(A matrix with 1's on the other diagonal, not the main diagonal)

g.)

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Problem 2.2.3

$$[T]_\beta^\gamma = \begin{pmatrix} -1/3 & -1 \\ 0 & 1 \\ 2/3 & 0 \end{pmatrix}$$

$$[T]_\alpha^\gamma = \begin{pmatrix} -7/3 & -4/3 \\ 2 & 3 \\ 2/3 & 4/3 \end{pmatrix}$$

Problem 2.2.4

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Problem 2.2.5
a.)

$$[T]_\alpha = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

b.)

$$[T]_\beta^\alpha = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

c.)

$$[T]_\alpha^\beta = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}$$

d.)

$$[T]_\beta^\gamma = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix}$$

e.)

$$[A]_\alpha = \begin{pmatrix} 1 \\ -2 \\ 0 \\ 4 \end{pmatrix}$$

f.)

$$[f(x)]_\beta = \begin{pmatrix} 3 \\ -6 \\ 1 \end{pmatrix}$$

g.)

$$[a]_\gamma = \begin{pmatrix} a \end{pmatrix}$$

Problem 2.2.6 XXX
Let V and W be vector spaces over a field $\mathbb{K}$ and let $T, U : V \to W$ be linear transformations, and let $\mu \in \mathbb{K}$. Then note that the zero transformation T_0 plays the role of the zero vector. Let $x \in V$ and now see that:

$$(\mu T + U)(x) = (\mu T)(x) + U(x) = \mu T(x) + U(x)$$

which again is a linear transformation from V to W, by the usual linearity properties of functions. Hence $\mathcal{L}(V, W)$ is closed under function addition and scalar multiplication and is thus a vector space. FINISH

Problem 2.2.7
Let V and W be finite-dimensional vector spaces and $T : V \to W$ be linear. Now let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V and $\gamma = \{w_1, \ldots, w_m\}$ be a basis for W. Let $\lambda \in \mathbb{K}$. There exist unique scalars a_{ij}, with $1 \leq i \leq m$ and $1 \leq j \leq n$ such that:

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \implies (\lambda T)(v_j) = \sum_{i=1}^{m} (\lambda a_{ij}) w_i$$

And hence we see that:

$$([\lambda T]_\beta^\gamma)_{ij} = \lambda a_{ij} = (\lambda [T]_\beta^\gamma)_{ij}$$

And therefore part b.) of Theorem 2.8 is proved. ■

Problem 2.2.8
Let V be a finite-dimensional vector space with ordered basis β. Define the mapping $T : V \to \mathbb{F}^n$ by $T(x) = [x]_\beta$. Let $\mu \in \mathbb{F}$ and $x, y \in V$. Then note that:

$$T(\mu x + y) = [\mu x + y]_\beta = \mu [x]_\beta + [y]_\beta$$

Which follows from a particular case of the results from Theorem 2.8. Hence by definition T is a linear transformation as desired. ■

Problem 2.2.9
Consider the vector space $\mathbb{C}$ over the field $\mathbb{R}$. Define a mapping $T : \mathbb{C} \to \mathbb{C}$ by $T(z) = \overline{z}$. Recall that in Problem 2.1.38, we showed that T is not linear if $\mathbb{C}$ is regarded as a vector space over the field $\mathbb{C}$. Let $z, w \in \mathbb{C}$ and $r \in \mathbb{R}$. Then note that:

$$T(z + w) = \overline{z + w} = \overline{z} + \overline{w} = T(z) + T(w)$$

by the familiar properties of complex numbers. Hence T is an additive mapping.

$$T(rz) = \overline{rz} = r(\overline{z}) = rT(z)$$

Which follows since $r \neq i$ since $i \notin \mathbb{R}$. Hence T is a linear transformation by definition. To find a matrix representation of this transformation, note that $\beta = \{1, i\}$ is a basis for $\mathbb{C}$ over $\mathbb{R}$. Then $T(1) = 1$ and $T(i) = -i$, and so we have:

$$[T]_\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

And so this is the matrix representation of the linear transformation T as required. ■

Problem 2.2.10

Let V be a vector space with ordered basis $\beta = \{v_1, \ldots, v_n\}$. Define $v_0 = \vec{0}$. Then by Theorem 2.6, there exists $T : V \to V$ such that $T(v_j) = v_j + v_{j-1}$ for $1 \le j \le n$. First see that $T(v_1) = v_1 + \vec{0} = v_1$, then $T(v_2) = v_2 + v_1$, $T(v_3) = v_3 + v_2$, and so on. Hence:

$$[T]_\beta = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 & \\ 0 & 1 & 1 & & & \\ 0 & 0 & 1 & & & \\ \vdots & & & \ddots & \vdots & \\ 0 & 0 & 0 & & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Is the matrix representation of T with respect to the ordered basis β.

Problem 2.2.11

Problem 2.2.12

Problem 2.2.13
Let V and W be vector spaces, and let $T, U \in \mathcal{L}(V, W)$ such that $T, U \neq T_0$, the zero tranformation. Suppose that $R(T) \cap R(U) = \{\vec{0}\}$. Now, by way of contradiction, suppose that the set $\{T, U\} \subseteq \mathcal{L}(V, W)$ is linearly dependent. Then we know by definition that for some nonzero $a \in \mathbb{K}$ and $x \in V$ such that $x \neq \vec{0}$,

$$aT(x) = U(x) \implies T(ax) = U(x) \implies a = 0$$

Since the intersection of the image of V under T and under U are $\{\vec{0}\}$, but $x \neq \vec{0}$ so $a = 0$, but this is a contradiction to the linear dependence of the set $\{T, U\}$. Hence the subset $\{T, U\}$ is a linearly independent subset of $\mathcal{L}(V, W)$. ■

Problem 2.2.14
Consider the vector space $P(\mathbb{R})$ and define $T_j(f(x)) = f^j(x)$ for $j \ge 1$. Consider the set $D = \{T_1, T_2, \ldots, T_n\} \subseteq \mathcal{L}(P(\mathbb{R}))$, where $n \in \mathbb{Z}^+$. Suppose by way of contradiction that the set D is linearly dependent. Then arbitrarily choose $T_k \in D$, where $1 \le k \le n$, such that T_k is a linear combination of the other transformations in D. Then we have, for $a_i \neq 0$ for all i,

$$T_k = \sum_{i=1}^{n-1} a_i T_i \implies T_k(x) = \sum_{i=1}^{n-1} a_i T_i(x) \implies kx^{k-1} = \sum_{i=1}^{n-1} a_i (ix^{i-1})$$

But we know that $x^{k-1} \neq ax^{i-1}$ for any $i \neq k$, and any $a \in \mathbb{R}$. Hence T_k cannot be represented as a linear combination of the other transformations, contradicting the linear dependence of the set D. Therefore we conclude that $\{T_1, \ldots, T_n\} \subseteq \mathcal{L}(P(\mathbb{R}))$ is a linearly independent set for any choice of $n \in \mathbb{Z}^+$

as desired. ■

Problem 2.2.15
Let V and W be vector spaces, and let $S \subseteq V$. Now define the set $S^0 = \{T \in \mathcal{L}(V, W) : T(x) = 0 \;\forall\; x \in S\}$

a.) Let $T, U \in S^0$ and let $\mu \in \mathbb{K}$. Note first that the zero transformation $T_0 \in S^0$ trivially, as $T_0(x) = 0$ for all $x \in V$. Now see that:

$$(\mu T + U)(x) = \mu T(x) + U(x) = \mu \cdot 0 + 0 = 0 \implies \mu T + U \in S^0$$

Following from the definition of $\mathcal{L}(V, W)$ as a vector space in Problem 2.2.6. Hence the set $S^0 \subseteq \mathcal{L}(V, W)$ contains the zero transformation, and is closed under transformation addition and scalar multiplication, and is thus a subspace of the vector space $\mathcal{L}(V, W)$ by Theorem 1.3. ■

b.) Supose $S_1, S_2 \subseteq V$ such that $S_1 \subseteq S_2$. Let $T \in S_2^0$. Then $T(x) = 0$ for all $x \in S_2$, but since $S_1 \subseteq S_2$, then $T(x) = 0$ for all $x \in S_1$ as well, so that we have $T \in S_1^0$ necessarily. Hence we find that $S_2^0 \subseteq S_1^0$. ■

c.) Suppose that V_1 and V_2 are subspaces of V. First see that the subspaces $V_1 \cap V_2 \subseteq V_1 + V_2$ trivially, and thus by part b.) of this problem, since both are subsets of V, we have $(V_1 + V_2)^0 \subseteq (V_1 \cap V_2)^0$. Let $T \in (V_1 \cap V_2)^0$. Then for $x \in V_1 \cap V_2$, we have $T(x) = 0$. But also, for $y \in V$, we have $y = x + x \implies y \in V_1 + V_2$ since $x \in V_1$ and $x \in V_2$. Then $T(y) = T(x + x) = T(x) + T(x) = 0 + 0 = 0 \implies T(y) = 0 \implies T \in (V_1 + V_2)^0$. Hence we then have that $(V_1 \cap V_2)^0 \subseteq (V_1 + V_2)^0$, and so we have set equality from the previous containment and $(V_1 + V_2)^0 = (V_1 \cap V_2)^0$ as desired. ■

Problem 2.2.16
Let V and W be vector spaces such that $\dim(V) = \dim(W)$ and let $T : V \to W$ be linear. Let $\beta = \{v_1, \ldots, v_k\}$ be a basis for $N(T)$, and extend this basis to a basis for V as $\beta \cup \{v_{k+1}, \ldots, v_n\}$. Then by Theorem 2.6, there exists $w_j \in W$ such that $T(v_j) = w_j$ for $k + 1 \leq j \leq n$. Then the set $\{w_k, \ldots, w_n\}$ is linearly independent and now we extend it to a basis for W, $\gamma = \{w_1, \ldots, w_k, w_{k+1}, \ldots, w_n\}$. Then:

$$[T]_\beta^\gamma = \begin{pmatrix} O & O \\ O & \mathcal{I}_{n-k} \end{pmatrix}$$

Since the first k vectors map to zero vector, and the rest of v_i are 1-1 scalar multiples of w_i for all relevant i. This is a diagonal matrix as required. ■

2.3 Composition of Linear Transformations and Matrix Multiplication

Problem 2.3.1
a.) False. $[UT]_{\alpha}^{\beta} = [U]_{\beta}^{\gamma}[T]_{\beta}^{\gamma}$ by Theorem 2.11.

b.) True.

c.) False.

d.) True. This is part d.) of Theorem 2.12.

e.) False.

f.) False.

g.) True. This is part d.) of Theorem 2.15.

h.) False.

i.) False. Note that the hypthesis of Theorem 2.15 is not satisfied.

j.) True.

Problem 2.3.2
a.)

$$A(2B+3C) = \begin{pmatrix} 20 & -9 & 18 \\ 5 & 10 & 8 \end{pmatrix}$$

$$(AB)D = \begin{pmatrix} 29 \\ -26 \end{pmatrix} = A(BD)$$

b.)

$$A^T = \begin{pmatrix} 2 & -3 & 4 \\ 5 & 1 & 2 \end{pmatrix}$$

$$A^T B = \begin{pmatrix} 23 & 19 & 0 \\ 26 & -1 & 10 \end{pmatrix}$$

$$BC^T = \begin{pmatrix} 12 \\ 16 \\ 29 \end{pmatrix}$$

$$CB = \begin{pmatrix} 27 & 7 & 9 \end{pmatrix}$$

$$CA = \begin{pmatrix} 20 & 26 \end{pmatrix}$$

Problem 2.3.3
a.) Performing the required computations with the basis vectors β and γ as defined in the problem for the linear transformations $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ and $U : P_2(\mathbb{R}) \to \mathbb{R}^3$, we find:

$$[U]_\beta^\gamma = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix},\ [T]_\beta = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 3 & 6 \\ 0 & 0 & 4 \end{pmatrix},\ [UT]_\beta^\gamma = \begin{pmatrix} 2 & 6 & 6 \\ 0 & 0 & 4 \\ 2 & 0 & -6 \end{pmatrix}$$

And note that the matrix $[UT]_\beta^\gamma$ does satisfy the results of Theorem 2.11.

b.) Let $h(x) = 3 - 2x + x^2$. Then, with β still as the standard ordered basis for $P_2(\mathbb{R})$, we have:

$$[h(x)]_\beta = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$$

And since $U(h(x)) = (1, 0, 2) = 1\vec{\gamma}_1 + 2\vec{\gamma}_2$, we can then easily find that:

$$[U(h(x))]_\gamma = \begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix}$$

Now we can also use Theorem 2.14 to verify our result by using the matrix $[U]_\beta^\gamma$ found in part a.) as follows:

$$[U(h(x))]_\gamma = [U]_\beta^\gamma [h(x)]_\beta = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix}$$

And so we see that the results obtained are indeed verified by Theorem 2.14.

Problem 2.3.4
a.)

$$[T]_\alpha [A]_\alpha = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ -1 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 4 \\ 6 \end{pmatrix}$$

b.)

$$[T(f(x))]_\alpha = [T]_\beta^\alpha [f(x)]_\beta = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ -6 \\ 3 \end{pmatrix} = \begin{pmatrix} -6 \\ 2 \\ 0 \\ 6 \end{pmatrix}$$

c.)

$$[T(A)]_\gamma = [T]_\alpha^\gamma [A]_\alpha = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 2 \\ 4 \end{pmatrix} = 5$$

d.)

$$[T(f(x))]_\gamma = [T]_\beta^\gamma [f(x)]_\beta = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 6 \\ -1 \\ 2 \end{pmatrix} = 12$$

Problem 2.3.5 XX
First we will prove parts b.) and d.) of Theorem 2.12. Let $a \in \mathbb{K}$. Let A be an $n \times n$ matrix and B be an $n \times p$ matrix. Then note that:

$$[a(AB)]_{ij} = \sum_{k=1}^{n} aA_{ik}B_{kj} = \sum_{k=1}^{n} (aA)_{ik}B_{kj} = \sum_{k=1}^{n} A_{ik}(aB)_{kj}$$

And thus $a(AB) = (aA)B = A(aB)$ which proves part b.) of Theorem 2.12. Now suppose that V is a finite-dimensional vector space with $\dim(V) = n$, and β is an ordered basis for V. Note that $I_V \in \mathcal{L}(V)$ is clearly a bijection. Then I_V maps β to β and thus $[I_V]$ is an $n \times n$ diagonal matrix with 1's on the main diagonal, and thus $[I_V]_\beta = I_n$.

Now we will prove the corollary to Theorem 2.12. Let A be an $n \times n$ matrix and $B_1, \ldots, B_k$ be $n \times p$ matrices, $C_1, \ldots, C_k$ be $q \times m$ matrices, and $a_1, \ldots, a_k \in \mathbb{K}$. Then:

Problem 2.3.6

Problem 2.3.7

Problem 2.3.8
First we will prove Theorem 2.10. Let V be a vector space and $T, U_1, U_2 \in \mathcal{L}(V)$. Then note that:

$$(T(U_1 + U_2))(x) = T((U_1 + U_2)(x)) = T(U_1(x) + U_2(x)) =$$

$$T(U_1(x)) + T(U_2(x)) = (TU_1 + TU_2)(x)$$

Problem 2.3.9 XX
We aim to find all linear transformations $U, T : \mathbb{K}^2 \to \mathbb{K}^2$ such that $UT = T_0$

but $TU \neq T_0$. Consider $UT = T_0$. We see that for $(a,b) \in \mathbb{K}^2$, we have $U(T(a,b)) = (0,0)$, which implies that $R(T) \subseteq N(U)$. But $T(U(a,b)) \neq (0,0)$, so $R(U) \nsubseteq N(T)$.

Problem 2.3.10
Let $A \in M_n(\mathbb{K})$. Suppose that A is a diagonal matrix. Then $A_{ij} = 0$ for all $i \neq j$, with $1 \leq i,j \leq n$. By Theorem 2.12, we know that $A = I_n A$, and by definition of the identity matrix, $\delta_{ij} = (I_n)_{ij}$, we have necessarily that $A = \delta_{ij} I_n \implies A_{ij} = \delta_{ij} A_{ij}$. Conversely, suppose we have that $A_{ij} = \delta_{ij} A_{ij}$. Then $A_{ij} = 0 \cdot A_{ij}$ whenever $i \neq j$ by the Kronecker delta. Hence $A_{ij} = A_{ij} \neq 0 \iff i = j$, so $a_{ij} \neq 0$ only on the main diagonal entries of the matrix A, and thus by definition, A is a diagonal matrix. ■

Problem 2.3.11
Let V be a vector space and $T : V \to V$ be linear. Suppose that $T^2 = T_0$. Then for any $x \in V$, $T(x) \in R(T)$ and we have $T(T(x)) = \vec{0} \implies T(x) \in N(T)$. Hence $R(T) \subseteq N(T)$. Conversely, suppose that $R(T) \subseteq N(T)$. Then let $T(x) \in R(T)$ be arbitrary, so we have that $T(T(x)) = \vec{0}$ for all $T(x) \in R(T)$. Thus $T(T(x)) = T^2(x) = T_0(x) \implies T^2 = T_0$ as desired. ■

Problem 2.3.12
Let V, W, Z be vector spaces, and let $T : V \to W$ and $U : W \to Z$ be linear.

a.) Suppose UT is injective. Let $x, y \in V$. Then $U(T(x)) = U(T(y)) \implies x = y$ by definition of injectivity. But if $x = y$, then since T is linear, $T(x-y) = T(\vec{0}_V) = \vec{0}_W \implies T(x) = T(y)$. Hence $U(T(x)) = U(T(y)) \implies T(x) = T(y)$ so we have that T is injective also. The injectivity of U is not required. ■

b.) Suppose that UT is surjective. Then $z \in Z \implies$ there exists $x \in V$ such that $z = U(T(x))$. But also $T(x) \in W$ exists and so U is surjective also. The surjectivity of T is not required.■

c.) Suppose that U and T are bijections. Now let $x, y \in V$ and suppose $U(T(x)) = U(T(y))$. Then by the injectivity of U, $T(x) = T(y)$, and subsequently by the injectivity of T, $x = y$ so that UT is injective as well. Now suppose $z \in Z$. Then by the surjectivity of U, there exists some $w \in W$ such that $z = U(w)$, but also by the surjectivity of T, there exists some $x \in V$ such that $T(x) = w$, and hence $z = U(T(x))$, and UT is surjective. Therefore UT is injective and surjective and is thus a bijection as desired. ■

Problem 2.3.13
Let $A, B \in M_n(\mathbb{K})$. Then since $A_{ij} = A^T_{ji}$ for $1 \leq i \leq n$, we have:

$$\operatorname{tr}(A^T) = \sum_{i=1}^{n} A^T_{ii} = \sum_{i=1}^{n} A_{ii} = \operatorname{tr}(A)$$

Additionally we can find:

$$\operatorname{tr}(AB) = \sum_{i=1}^{n}(AB)_{ii} = \sum_{i=1}^{n}(\sum_{j=1}^{n}(A_{ij}B_{ji})) =$$

$$\sum_{j=1}^{n}(\sum_{i=1}^{n}(B_{ji}A_{ij})) = \sum_{i=1}^{n}(BA)_{ii} = \operatorname{tr}(BA)$$

Which follows from the commutativity of scalar multiplication and the definition given in the section regarding matrix multiplication and the definition of the trace. ■

Problem 2.3.14
a.) Let B be an $n \times p$ matrix. Suppose $\vec{z} \in \mathbb{K}^p$ such that $\vec{z} = (a_1, \dots, a_p)^T$. We will show that $B\vec{z}$ is a linear combination of the columns of B. By Theorem 2.13, we know that for $\vec{v}_j$, the jth column of B, we have $\vec{v}_j = B\vec{e_j}$. Using this and the definition of matrix multiplication, we find:

$$B\vec{z} =$$

b.)

c.)

d.)

Problem 2.3.15

Problem 2.3.16
a.) Let V be a finite-dimensional vector space and $T \in \mathcal{L}(V)$. Suppose that $(T) = (T^2)$. By Problem 2.1.28, $R(T)$ is a T-invariant subspace $\implies T(R(T)) \subseteq R(T) \implies R(T^2) \subseteq R(T)$. Hence $R(T^2)$ is a subset of $R(T)$ that is a subspace, and hence by Theorem 1.11, $T^2 = T$. Now suppose by way of contradiction that $x \in V$ such that $x \neq \vec{0}$ and $x \in R(T^2) \cap N(T) \implies T(T(x)) \in V$ and $T(x) = \vec{0}$.

Problem 2.3.17

Problem 2.3.18 XX
We will prove using the definition of matrix multiplication that matrix multiplication is an associative operation. Let A be an $m \times n$ matrix, B be an $n \times p$ matrix, and C be a $p \times q$ matrix. Then:

$$(AB)C = (AB)_{ij}C_{ij} = (\sum_{k=1}^{n} A_{ik}B_{kj})C_{ij}$$

Problem 2.3.19

Problem 2.3.20

Problem 2.3.21

Problem 2.3.22

Problem 2.3.23

2.4 Invertibility and Isomorphism

Problem 2.4.1
a.) False. $([T]_\alpha^\beta)^{-1} = [T^{-1}]_\beta^\alpha$.

b.) True. Bijection requires surjectivity and injectivity.

c.) False.

d.) False. $\dim(M_{2\times 3}(\mathbb{K}) = 6 \neq \dim(\mathbb{K}^5) = 5$, Theorem 2.19.

e.) True. $\dim(P_n(\mathbb{K})) = \dim(P_m(\mathbb{K})) \iff n = m$, Theorem 2.19.

f.) False.

g.) True. $BA^{-1} = A^{-1}B = I \implies B = A$

h.) True. Recall a.) of Theorem 2.15.

i.) True. Think about in terms of dimensions of transformation.

Problem 2.4.2
a.) We can immediately see that T is not invertible since $\dim(\mathbb{R}^2) \neq \dim(\mathbb{R}^3)$, which follows from the Lemma previous to Theorem 2.18.

b.) T is not invertible, for the same reason as above.

c.) Taking the standard basis for $\mathbb{R}^3$, we find: $T(1,0,0) = (3,0,3)$, $T(0,1,0) = (0,1,4)$, and $T(0,0,1) = (-2,0,0)$, which is clearly a linearly independent set and spans $\mathbb{R}^3$, so T is surjective. Hence by Theorem 2.5, T is also injective, and hence T is invertible.

d.) The vector spaces do not have equal dimensions and thus T cannot be invertible. Refer to Theorem 2.19.

e.) T is not invertible since $\dim(M_2(\mathbb{R})) = 4 \neq \dim(P_2(\mathbb{R})) = 3$.

f.)

$$T(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \; T(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; T(\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}) = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

$$T(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

And this set of transformed basis vectors is clearly linearly independent and spans $M_2(\mathbb{R})$, so T is surjective. Hence by Theorem 2.5, T is also injective and therefore it is invertible.

Problem 2.4.3
a.) No. Clearly $\dim(\mathbb{K}^3) = 3 \neq \dim(P_3(\mathbb{K})) = 4$, so that the vector spaces cannot be isomorphic by Theorem 2.19.

b.) Yes. We have $\dim(\mathbb{K}^4) = 4 = \dim(P_3(\mathbb{K}))$, so by Theorem 2.19, the vector spaces are isomorphic.

c.) Yes. $\dim(M_2(\mathbb{R})) = 4 = \dim(P_3(\mathbb{K}))$, so by Theorem 2.19, these vector spaces are isomorphic.

d.) No. $\dim(V) = 3 \neq \dim(\mathbb{R}^4) = 4$. Recall Problem 1.6.15, and substitute $n = 2$ to find the dimension to be $n^2 - 1 = 4 - 1 = 3$.

Problem 2.4.4
Let A and B be invertible $n \times n$ matrices. Then take the product AB and observe:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$$

So we see that by the definition of invertibility of a matrix, $(AB)^{-1} = B^{-1}A^{-1}$ as desired. ■

Problem 2.4.5

Let A be invertible. Then note, since we know that $(AB)^T = B^T A^T$, by rules of matrix transposes, we can see that:

$$(A^{-1})^T A^T = (AA^{-1})^T = (I_n)^T = I_n$$

$$A^T (A^{-1})^T = (A^{-1}A)^T = (I_n)^T = I_n$$

Hence by the definition of an inverse matrix, we find the desired relation between a tranpose and its inverse: $(A^T)^{-1} = (A^{-1})^T$. ■

Problem 2.4.6
Suppose that the matrix A is invertible and $AB = O$ for some matrix B. Then we can multiply both sides of the equation on the left by the inverse of A, and we find:

$$AB = O \implies A^{-1}(AB) = A^{-1}O \implies (A^{-1}A)B = O$$
$$\implies IB = O \implies B = O$$

So the matrix B must be the zero matrix, which follows since any matrix multiplied by the zero matrix O is necessarily the zero matrix. ■

Problem 2.4.7
a.) Let A be an $n \times n$ matrix. Suppose that $A^2 = O$. Now suppose by way of contradiction that A is invertible. Then we can apply the inverse of A to the relation and we see:

$$A^2 = AA = O \implies A^{-1}AA = A^{-1}O \implies I_n A = O \implies A = O$$

However the zero matrix is not invertible since there does not exist a matrix B such that $OB = BO = I_n$, since any matrix multiplied by the zero matrix is again the zero matrix. Hence a contradiction, and so A is not invertible. ■

b.) Let A be an $n \times n$ matrix. Suppose that $AB = O$ for some nonzero $n \times n$ matrix B. Now suppose that A is invertible, and apply its inverse. Then by the results of Problem 2.4.6, we find $B = O$, a contradiction. Hence A cannot be invertible. ■

Problem 2.4.8
We will prove corollaries 1 and 2 of Theorem 2.18.

Corollary 1: Let V be a finite-dimensional vector space with an ordered basis β, and let $T : V \to V$ be linear. Suppose that T is invertible. Let $\dim(V) = n$, so that $[T]_\beta$ is an $n \times n$ matrix. Now $T^{-1} : V \to V$ satisfies $TT^{-1} = T^{-1}T = I_V$. Thus we have:

$$I_n = [I_V]_\beta = [TT^{-1}]_\beta = [T]_\beta [T^{-1}]_\beta$$

And similarly, we can find $I_n = [T^{-1}]_\beta [T]_\beta$, so that $[T]_\beta$ is invertible by definiton and $([T]_\beta)^{-1} = [T^{-1}]_\beta$. Conversely, suppose that $A = [T]_\beta$ is invertible. Then there exists $n \times n$ matrix B such that $AB = BA = I_n$. By Theorem 2.6 there exists $U \in \mathcal{L}(V)$ such that:

$$U(v_j) = \sum_{i=1}^{n} B_{ij} w_i$$

for $1 \leq j \leq n$, where $\beta = \{v_1, \ldots, v_n\} = \{w_1, \ldots, w_n\}$. Then $[U]_\beta = B$, and we can show $U = T^{-1}$ by seeing:

$$[UT]_\beta = [U]_\beta [T]_\beta = BA = I_n = [I_V]_\beta$$

And similarly that $[TU]_\beta = [I_V]_\beta$. Hence $TU = I_V = UT \implies U = T^{-1}$ so that T is invertible as desired. ■

Corollary 2: Let A be an $n \times n$ matrix. Suppose that A is invertible. Hence there exists some $n \times n$ matrix A^{-1} such that $AA^{-1} = A^{-1}A = I_n$. From Theorem 2.15, we see:

$$L_{I_n} = I_{\mathbb{K}^n} \implies L_{AA^{-1}} = L_{A^{-1}A} = I_{\mathbb{K}^n} \implies L_A L_{A^{-1}} = L_{A^{-1}} L_A = I_{\mathbb{K}^n}$$

Then by definiton L_A is an invertible transformation, and $(L_A)^{-1} = L_{A^{-1}}$. Conversely, suppose that L_A is invertible. Then there exists some transformation L_B such that $L_A L_B = L_B L_A = I_{\mathbb{K}^n}$, but by Theorem 2.15, we know that $L_A L_B = L_{AB}$ and $L_B L_A = L_{BA}$, and also that $L_{I_n} = I_{\mathbb{K}^n}$. Hence:

$$L_{AB} = L_{BA} = I_{\mathbb{K}^n} \implies AB = BA = I_n$$

And this is the definition of A being invertible. ■

Problem 2.4.9
Let A and B be $n \times n$ such that AB is invertible. If AB is invertible, then by Problem 2.4.4, its inverse is given by $B^{-1}A^{-1}$. Hence B^{-1} and A^{-1} exist so that both A and B must also be invertible. ■.

If A and B are arbitrary matrices, meaning not both are $n \times n$, then the above need not hold. For example, consider A being $n \times m$ and B being $m \times n$ matrices. Then the product AB is an $n \times n$ matrix, but both A and B cannot be invertible sicne they are not square matrices.

Problem 2.4.10
Let A and B be $n \times n$ matrices such that $AB = I_n$.

a.) Since the product $AB = I_n$, and I_n is trivially invertible as $I_n^{-1} = I_n$, then since A and B are $n \times n$, by Problem 2.4.9, A and B are invertible. ■

b.) By the first part of the problem, we know that A and B are invertible, and that $AB = I_n$. In particular, since A is invertible, then we can multiply both sides of the equation on the left by A^{-1}. Then:

$$AB = I_n \implies A^{-1}AB = A^{-1}I_n \implies I_n B = A^{-1} \implies B = A^{-1}$$

And also we can similarly find $A = B^{-1}$. Essentially, for square matrices, a one-sided inverse is identical to a two-sided inverse. ■

c.)

Problem 2.4.11
From Example 5, define:

$$T : P_3(\mathbb{R}) \to M_2(\mathbb{R}),\ \ T(f) = \begin{pmatrix} f(1) & f(2) \\ f(3) & f(4) \end{pmatrix}$$

We will show that T is injective by using the Lagrange interpolation formula to show that only one $f(x) \in P_3(\mathbb{R})$ maps to the 2×2 zero matrix. We would like to find a polynomial in $P_3(\mathbb{R})$ with roots at $x = 1, 2, 3, 4$ only. Equivalently, we will construct a polynomial f of degree at most 3 whose graph contains the points $(1, 0), (2, 0), (3, 0), (4, 0)$. However note that, by the discussion at the end of Chapter 1.6 in the text, that if $f \in P_3(\mathbb{R})$ and $f(x) = 0$ for 4 distinct scalars $1, 2, 3, 4 \in \mathbb{R}$, then f is the zero function. Hence $N(T) = \{f\} = \{\vec{0}\}$ and so by Theorem 2.4, T is injective. ■

Problem 2.4.12
We will prove Theorem 2.21. Let V be an arbitrary finite-dimensional vector space with ordered basis β. Let $\dim(V) = n$ and V be over $\mathbb{K}$. Then consider the standard representation of V with respect to β, $\phi_\beta : V \to \mathbb{K}^n$, defined by $\phi_\beta(\vec{x}) = [\vec{x}]_\beta$. We previously proved that ϕ_β is linear. Now suppose that $[x]_\beta = [y]_\beta$ for $x, y \in V$. Let $\beta = \{u_1, \ldots, u_n\}$. Then by definition of coordinate vector relative to β, for $a_1, \ldots, a_n \in \mathbb{K}$:

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = [y]_\beta \implies x = \sum_{i=1}^{n} a_i u_i = y$$

So we necessarily have that $x = y$. Hence ϕ_β is injective and thus by Theorem 2.5, since $\dim(V) = \dim(\mathbb{K}^n)$, ϕ_β is also surjective. Therefore ϕ_β is an isomorphism from V to $\mathbb{K}^n$ as desired. ■

Problem 2.4.13
Let $\cong$ denote "is isomorphic to". Let V be a finite-dimensional vector space. Then $I_V \in \mathcal{L}(V)$ is the identity transformation on V. Since I_V is trivially linear, and is a bijection, $V \cong V$, so that the relation $\cong$ is reflexive.

Now let W be a finite-dimensional vector space. Suppose that $V \cong W$. Then by Theorem 2.19, $\dim(V) = \dim(W)$, and swapping the roles of V and W, we find that $W \cong V$, so that the relation is symmetric.

Now let Z be a finite-dimensional vector space. Suppose that $V \cong W$ and

$W \cong Z$. Then again by Theorem 2.19, $\dim(V) = \dim(W)$, but also $\dim(W) = \dim(Z)$, so that $\dim(V) = \dim(Z)$. Therefore $V \cong Z$ and so the relation is proved to be transitive.

Hence $\cong$ satisfies all of the conditions for an equivalence relation on vector spaces, and is thus an equivalence relation as desired. ■

Problem 2.4.14
Let

$$V = \{\begin{pmatrix} a & a+b \\ 0 & c \end{pmatrix} : a, b, c \in \mathbb{K}\}$$

We will construct an isomorphism from V to $\mathbb{K}^3$. First, note that a basis for V is given by the set:

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

Give β the ordering as denoted above, so that now β is an ordered basis for V. Now apply the standard representation of V with respect to β transformation, $\phi_\beta : V \to \mathbb{K}^3$ by $\phi_\beta(A) = [A]_\beta$ for each $A \in V$. By Theorem 2.21, ϕ_β is the desired isomorphism from V to $\mathbb{K}^3$.

Problem 2.4.15
Let V and W be vector spaces such that $\dim(V) = \dim(W) = n$, and let $T : V \to W$ be linear. Now let β be a basis for V. Suppose that T is an isomorphism. Then by Problem 2.1.14 c., since T is both injective and surjective, we have that $T(\beta)$ is a basis for W. Conversely, suppose that $T(\beta)$ is a basis for W. Since V and W are both finite-dimensional, apply Theorem 2.2 to obtain $R(T) = \operatorname{span}(T(\beta)) \implies R(T) = W$ since $T(\beta)$ is a basis. Hence T is surjective and since $\dim(V) = \dim(W)$, by Theorem 2.5, T is injective also. Hence T is an isomorphism as desired. ■

Problem 2.4.16
Let B be an $n \times n$ invertible matrix. Define the mapping $\Phi : M_n(\mathbb{K}) \to M_n(\mathbb{K})$ by $\Phi(A) = B^{-1}AB$. Let $A, C \in M_n(\mathbb{K})$ and $\mu \in \mathbb{K}$. Then observe that:

$$\Phi(\mu A + C) = B^{-1}(\mu A + C)B = (\mu(B^{-1}A) + B^{-1}C)B =$$

$$\mu B^{-1}AB + B^{-1}CB = \mu\Phi(A) + \Phi(C)$$

By the properties derived in Theorem 2.12 regarding the distributivity of matrices. Hence Φ is linear. Now let $A, C \in M_n(\mathbb{K})$ and suppose that $\Phi(A) = \Phi(C)$. Then we have that:

$$\Phi(A) = \Phi(C) \implies B^{-1}AB = B^{-1}CB \implies (BB^{-1})AB = (BB^{-1})CB$$

$$\implies I_n AB = I_n CB \implies AB = CB \implies A = C$$

where the last action follows from B being invertible and thus $B \neq O$, so that the left cancellation law applies. Hence Φ is injective. Noting that Φ is linear and the dimension of the domain and codomain are equal, by Theorem 2.5, Φ is surjective. Therefore Φ is an isomorphim. $\blacksquare$

Problem 2.4.17
Let V and W be finite-dimensional vector spaces and $T : V \to W$ be an isomorphism. Let V_0 be a subspace of V.

a.) First, since V_0 is a subspace, it contains the zero vector and since T is linear, $T(\vec{0}) = \vec{0} \in T(V_0)$. Let $x, y \in V_0$ and let $T(x), T(y) \in T(V_0)$ and $\mu \in \mathbb{K}$. Then note:

$$\mu T(x) + T(y) = T(\mu x) + T(y) = T(\mu x + y) \in T(V_0)$$

Since $\mu x + y \in V_0$ by definition of a subspace. Hence $T(V_0)$ is closed under vector addition and scalar multiplication and contains the zero vector, and is hence a subspace of W. $\blacksquare$

b.) By Problem 2.4.15, if β is a basis for V_0 then $T(\beta)$ is a basis for $T(V_0)$, as a special case. Clearly these two sets contain the same number of vectors, and hence $\dim(V_0) = \dim(T(V_0))$. $\blacksquare$

Problem 2.4.18
Consider the polynomial $p(x) = 1 + x + 2x^2 + x^3$. Noting that $\phi_\beta(p(x)) = \begin{pmatrix} 1 & 1 & 2 & 1 \end{pmatrix}^T$, we can follow the procedure in Example 7 to find:

$$L_A \phi_\beta(p(x)) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}$$

And since $T(p(x)) = p'(x) = 1 + 4x + 3x^2$, we have:

$$\phi_\gamma T(p(x)) = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}$$

And so we again find that $\phi_\gamma T(p(x)) = L_A \phi_\beta(p(x))$

Problem 2.4.19
From Example 5 of Section 2.1, we have the mapping $T : M_2(\mathbb{R}) \to M_2(\mathbb{R})$, defined by $T(A) = A^T$ for each $A \in M_2(\mathbb{R})$, is linear. Let $\beta = \{E^{11}, E^{12}, E^{21}, E^{22}\}$ be a basis for $M_2(\mathbb{R})$ by Example 3 of Section 1.6.

a.) Note that $T(E^{11}) = E^{11}$, $T(E^{12}) = E^{21}$, $T(E^{21}) = E^{12}$, and $T(E^{22}) = E^{22}$.

Hence we can find that the matrix representation of T with respect to β is:

$$[T]_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Using the methods outlined in Section 2.2.

b.) Consider the matrix $M \in M_2(\mathbb{R})$ such that:

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

Let $A = [T]_\beta$ and then we have:

$$L_A\phi_\beta(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 4 \end{pmatrix}$$

$$\phi_\beta T(\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}) = \phi_\beta(\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}) = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 4 \end{pmatrix}$$

And thus we have verified that $L_A\phi_\beta(M) = \phi_\beta T(M)$.

Problem 2.4.20

Let $T : V \to W$ be linear, and $\dim(V) = n$ and $\dim(W) = m$. Let β be an ordered basis for V and γ be an ordered basis for W. Let $A = [T]_\beta^\gamma$. Since $V \cong \mathbb{K}^n$, by Theorem 2.21, and $N(T)$ is a subspace of V, then $\phi_\beta(N(T))$ is a subspace of $\mathbb{K}^n$ by Problem 2.4.17. However then note that:

$$L_A\phi_\beta(N(T)) = \phi_\gamma(T(N(T)) = \phi_\gamma(\vec{0}) = \vec{0}$$

So that $\phi_\beta(N(T)) = N(L_A)$, and $N(L_A)$ is a subspace of $\mathbb{K}^n$, and so again by Problem 2.4.17, we now have $\dim(N(T)) = \dim(N(L_A)) \implies \text{nullity}(T) = \text{nullity}(L_A)$. Now by applying the dimension theorem, we know that:

$$\dim(V) = \dim(N(T)) + \dim(R(T)) \implies$$

$$\dim(N(T)) + \dim(R(T)) = \dim(N(L_A)) + \dim(R(L_A))$$

Since $\dim(\mathbb{K}^n) = \dim(V) = n$. But since we previously found that $\dim(N(T)) = \dim(N(L_A))$, we can cancel the two from the above equation, and we are left with $\dim(R(T)) = \dim(R(L_A)) \implies \text{rank}(T) = \text{rank}(L_A)$. ■

Problem 2.4.21

Problem 2.4.22
Let $c_0, c_1, c_2, \ldots, c_n$ be distinct scalars in $\mathbb{K}$. Define $T : P_n(\mathbb{K}) \to \mathbb{K}^{n+1}$ by $T(f) = (f(c_0), f(c_1), \ldots, f(c_n))$. Let $f, g \in P_n(\mathbb{K})$ and $\mu \in \mathbb{K}$. Then:

$$T(\mu f + g) = ((\mu f + g)(c_0), \ldots, (\mu f + g)(c_n)) =$$
$$(\mu f(c_0) + g(c_0), \ldots, \mu f(c_n) + g(c_n)) =$$
$$\mu(f(c_0), \ldots, f(c_n)) + (g(c_0), \ldots, g(c_n)) =$$
$$\mu T(f) + T(g)$$

And hence T is linear. Now let $h \in P_n(\mathbb{K})$ and suppose that $h \in N(T)$. Then $T(h) = (h(c_0), \ldots, h(c_n)) = (0, \ldots, 0) = \vec{0}$, and so we have that h has roots at $c_0, c_1, \ldots, c_n$. Then by the discussion at the end of section 1.6, since $h \in P_n(\mathbb{K})$ and $h(c_i) = 0$ for $n+1$ distinct scalars in $\mathbb{K}$, then h is the zero function. Hence $N(T) = \{h\} = \{\vec{0}\}$ and by Theorem 2.4, T is injective. Now by Theorem 2.5, since $\dim(P_n(\mathbb{K})) = \dim(\mathbb{K}^{n+1})$, and T is injective, T is also necessarily surjective, so that T is a bijection. Therefore T is an isomorphism from $P_n(\mathbb{K})$ to $\mathbb{K}^{n+1}$ as desired. ■

Problem 2.4.23
Let V be the vector space of sequences in $\mathbb{K}$ that have only a finite number of nonzero terms. Define $T : V \to P(\mathbb{K})$ by:

$$T(\sigma) = \sum_{i=0}^{n} \sigma(i)x^i$$

where n is the largest integer such that $\sigma(n) \neq 0$. Let $\sigma, \tau \in V$ and $\mu \in \mathbb{K}$.

$$T(\mu\sigma + \tau) = \sum_{i=0}^{n}(\mu\sigma + \tau)(i)x^i = \sum_{i=0}^{n}(\mu\sigma(i)x^i + \tau(i)x^i) =$$
$$\mu\sum_{i=0}^{n}\sigma(i)x^i + \sum_{i=0}^{n}\tau(i)x^i = \mu T(\sigma) + T(\tau)$$

Using the operations for addition and scalar multiplication of sequences defined in Example 5 of section 1.2. This shows that T is linear. Now suppose that $\sigma \in N(T)$. Then $T(\sigma) = \vec{0} \in P(\mathbb{K})$, and so we see:

$$T(\sigma) = \sum_{i=0}^{n}\sigma(i)x^i = \vec{0}$$

However note that the set $\beta = \{1, x, x^2, \ldots\}$ is a basis for $P(\mathbb{K})$, and as such no nontrivial linear combination gives the zero vector. Thus $\sigma(i) = 0$ for all $1 \leq i \leq n$ applicable, and the rest $i > n$ are 0 by construction, so that σ is necessarily the zero sequence. Hence $N(T) = \{\vec{0}\}$ and by Theorem 2.4, T is injective. Now suppose $p(x) \in P(\mathbb{K})$. Then since β is a basis for $P(\mathbb{K})$, we can write, for $a_1, \ldots, a_n \in \mathbb{K}$:

$$p(x) = \sum_{i=0}^{n} a_i x^i = \sum_{i=0}^{n} \sigma(i) x^i = T(\sigma)$$

By constructing a sequence $\sigma \in V$ with $\sigma(i) = a_i$ for $1 \leq i \leq n$. Hence T is surjective, and therefore T is an isomorphism from V to $P(\mathbb{K})$ as desired. $\blacksquare$

Problem 2.4.24
Let V and Z be vector spaces and $T : V \to Z$ linear. Define the mapping $\overline{T} : V/N(T) \to Z$ by $\overline{T}(v + N(T)) = T(v)$ for any coset $v + N(T) \in V/N(T)$.

a.) Let $v, v' \in V$ and suppose that $v + N(T) = v' + N(T)$. Then observe:

$$\overline{T}(v + N(T)) = \overline{T}(v' + N(T)) \implies T(v) = T(v')$$

So we have that $\overline{T}$ is a well-defined mapping. $\blacksquare$

b.) Let $v + N(T), w + N(T) \in V/N(T)$, and $\lambda \in \mathbb{K}$. Then observe:

$$\overline{T}(\lambda[v + N(T)] + [w + N(T)]) = \overline{T}([\lambda v + N(T)] + [w + N(T)]) =$$

$$\overline{T}((\lambda v + w) + N(T)) = T(\lambda v + w) = \lambda T(v) + T(w) = \lambda \overline{T}(v + N(T)) + \overline{T}(w + N(T))$$

Which follows from the presupposed linearity of T. Hence $\overline{T}$ is a linear transformation. $\blacksquare$

c.) Suppose that $T(x) = 0$. Then $x \in N(T)$, and so $x + N(T) = N(T) \in V/N(T) \implies x = \vec{0} + N(T)$. Hence $N(\overline{T}) = \{\vec{0} + N(T)\}$, which is the zero vector in $V/N(T)$. Thus by Theorem 2.5, $\overline{T}$ is injective. Now suppose $z \in Z$. Since T is surjective, there exists $x \in V$ such that $z = T(x)$. But then $T(x) = \overline{T}(x + N(T)) \implies z = \overline{T}(x + N(T))$, and hence $\overline{T}$ is surjective. Therefore $\overline{T}$ is a bijection and hence an isomorphism from $V/N(T)$ to Z. $\blacksquare$

d.) We will prove that $T = \overline{T}\eta$. Let $x \in V$ be arbitrary. Then we observe:

$$(\overline{T}\eta)(x) = \overline{T}(\eta(x)) = \overline{T}(x + N(T)) = T(x) \implies T = \overline{T}\eta$$

Which is the same as if T was applied directly to x. Hence the diagram in Figure 2.3 commutes.

Problem 2.4.25
Let $V \neq 0$ be a vector space over $\mathbb{K}$, and let S be a basis for V. Let $\mathcal{C}(S, \mathbb{K})$ denote the vector space of all functions $f \in \mathcal{F}(S, \mathbb{K})$ such that $f(s) = 0$ for all but a finite number of $s \in S$. Let:

$$\Phi : \mathcal{C}(S, \mathbb{K}) \to V$$

be a mapping defined by $\Phi(f) = 0$ if f is the zero function, and

$$\Phi(f) = \sum_{s \in S,\ f(s) \neq 0} f(s)s$$

otherwise. First we will show that the mapping Φ is linear. Let $f, g \in \mathcal{C}(S, \mathbb{K})$ and $\mu \in \mathbb{K}$. Then observe that:

$$\Phi(\mu f + g) = \sum_{s \in S,\ (\mu f + g)(s) \neq 0} (\mu f + g)(s)s \ = \sum_{s \in S,\ (\mu f + g)(s) \neq 0} (\mu f(s) + g(s))s =$$

$$\mu \sum_{s \in S,\ f(s) \neq 0} f(s)s \ + \sum_{s \in S,\ g(s) \neq 0} g(s)s \ = \mu\Phi(f) + \Phi(g)$$

And hence we have that the mapping is linear. Now we will show that it is injective and surjective. First, let $g \in \mathcal{F}(S, \mathbb{K})$ and suppose $\Phi(g) = 0$. Then we have that:

$$\Phi(g) = \sum_{s \in S,\ g(s) \neq 0} g(s)s \ = \vec{0}$$

But since S is a basis for V, then it cannot contain the zero vector, and hence each $s \neq \vec{0}$. Additionally, for each each $s \in S$, $g(s) \in \mathbb{K}$ and $g(s) \neq 0$, but since $s \in S$, then no linear combination of basis vectors can produce the zero vector. Hence g is the zero function and $N(\Phi) = \{0\}$, so that by Theorem 2.5, the mapping Φ is injective. Now suppose that $v \in V$. If $v = \vec{0}$, then $\Phi(f) = v$, where f is the zero function, as previously found. Suppose $v \neq \vec{0}$. Then since S is a basis for V, v can be represented as a unique linear combination of each of the vectors $s \in S$. So we have, for a finite number of nonzero scalars $a_i \in \mathbb{K}$,

$$v = \sum_{s \in S,\ a_i \neq 0} a_i s$$

Now by the Lagrange interpolation formula, we will construct a polynomial function f such that for each $s \in S$, we have $f(s_i) = a_i$. Letting g_i be the Lagrange polynomials associated with each of the a_i, we have:

$$f = \sum_i s_i g_i \implies f(s_i) = a_i$$

So that from the above representation, we now have:

$$v = \sum_{s \in S,\ f(s) \neq 0} f(s)s \implies v = \Phi(f)$$

And since $v \in V$ was arbitrary, each vector in V has a preimage under Φ, so that Φ is surjective. Hence Φ is a bijection and thus an isomorphism from $\mathcal{C}(S, \mathbb{K})$ to V. In particular, this shows that every nonzero vector space can be viewed as a space of functions. ■

2.5 The Change of Coordinate Matrix

Problem 2.5.1
a.) False. The jth column of Q is $[x_j]_\beta$.

b.) True. From the definition.

c.) True. Rearrange the equation in Theorem 2.23.

d.) False.

e.) True. Note the hypothesis for Theorem 2.23.

Problem 2.5.2
a.)
$$Q = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

b.)
$$Q = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$$

c.)
$$Q = \begin{pmatrix} 3 & -1 \\ 5 & -2 \end{pmatrix}$$

d.)
$$Q = \begin{pmatrix} 2 & -1 \\ 5 & -4 \end{pmatrix}$$

Problem 2.5.3
a.)
$$Q = \begin{pmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_0 & b_0 & c_0 \end{pmatrix}$$

b.)
$$Q = \begin{pmatrix} a_0 & b_0 & c_0 \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$$

c.)

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

d.)

$$Q = \begin{pmatrix} 2 & 1 & 1 \\ 3 & -2 & 1 \\ -1 & 3 & 1 \end{pmatrix}$$

e.)

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

f.)

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

Problem 2.5.4

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 13 \\ -5 & -9 \end{pmatrix}$$

Problem 2.5.5

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & -1/2 \end{pmatrix}$$

Problem 2.5.6
a.)

$$[L_A]_\beta = Q^{-1}AQ = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 11 \\ -2 & -4 \end{pmatrix}$$

b.)

$$[L_A]_\beta = Q^{-1}AQ = \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} -6 & 0 \\ 0 & 2 \end{pmatrix}$$

c.)

$$[L_A]_\beta = Q^{-1}AQ =$$

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

d.)

$$[L_A]_\beta = Q^{-1}AQ =$$

$$\begin{pmatrix} 1/6 & 1/6 & -1/3 \\ 1/2 & -1/2 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 13 & 1 & 4 \\ 1 & 13 & 4 \\ 4 & 4 & 10 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 18 \end{pmatrix}$$

Problem 2.5.7
a.)

b.)

Problem 2.5.8
Let $T : V \to W$ be a linear transformation from a finite-dimensional vector space V to a finite-dimensional vector space W. Let β and β' be ordered bases for V. Let γ and γ' be ordered bases for W. Suppose Q is the change of coordinate matrix that changes β'-coordinates to β-coordinates, and that P is the change of coordinate matrix that changes γ'-coordinates to γ-coordinates. Let I_V be the identity transformation on V and I_W the identity transformation on W. Then $T = I_W T = T I_V$. Hence:

$$P[T]_{\beta'}^{\gamma'} = [I_W]_{\gamma'}^{\gamma}[T]_{\beta'}^{\gamma'} = [I_W T]_{\beta'}^{\gamma} = [T I_V]_{\beta'}^{\gamma} = [T]_{\beta}^{\gamma}[I_V]_{\beta'}^{\beta} = [T]_{\beta}^{\gamma}Q$$

And subsequently, we see that $[T]_{\beta'}^{\gamma'} = P^{-1}[T]_{\beta}^{\gamma}Q$ as desired. ■

Problem 2.5.9
Let $\sim$ denote the relation "is similar to" in regards to elements of $M_n(\mathbb{K})$. Suppose $A \in M_n(\mathbb{K})$. We know that $I_n \in M_n(\mathbb{K})$, and that I_n is trivially invertible, $I_n^{-1} = I_n$. Then we have $A = I_n^{-1}AI_n = I_nAI_n = A$, so that $A \sim A$, and the relation $\sim$ is reflexive.

Now let $A, B \in M_n(\mathbb{K})$. Suppose that $A \sim B$. Then there exists $Q \in M_n(\mathbb{K})$ such that $A = Q^{-1}BQ$. But then applying Q to the left of both sides, followed by an application to the right, we find: $QA = BQ \implies QAQ^{-1} = B$, so that $Q^{-1} \in M_n(\mathbb{K})$ satisfies the condition. Hence $B \sim A$ and the relation $\sim$ is symmetric.

Finally, let $A, B, C \in M_n(\mathbb{K})$. Suppose that $A \sim B$ and that $B \sim C$. Then there exist invertible matrices $P, Q \in M_n(\mathbb{K})$ such that $A = Q^{-1}BQ$ and $B = P^{-1}CP$. But then, substituting, we find:

$$A = Q^{-1}(P^{-1}CP)Q = (Q^{-1}P^{-1})C(PQ) = (PQ)^{-1}C(PQ)$$

Which follows from Problem 2.4.4, since P and Q are invertible, then PQ is invertible and $(PQ)^{-1} = Q^{-1}P^{-1}$. Hence we see that $A \sim C$, so that the relation $\sim$ is transitive. Therefore the relation $\sim$ satisfies all of the conditions for an equivalence relation on $M_n(\mathbb{K})$. ■

Problem 2.5.10

Suppose A and B be are $n \times n$ matrices such that A is similar to B. Then there exists an invertible $n \times n$ matrix Q such that $A = Q^{-1}BQ$. Then, by Problem 2.3.13, we know that since BQ is an $n \times n$ matrix:

$$\operatorname{tr}(A) = \operatorname{tr}(Q^{-1}BQ) = \operatorname{tr}((Q^{-1})(BQ)) = \operatorname{tr}((BQ)(Q^{-1})) =$$

$$\operatorname{tr}(B(QQ^{-1})) = \operatorname{tr}(BI_n) = \operatorname{tr}(B)$$

And since A, B were arbitrary, we have shown that similar matrices have the same trace as desired. ■

Problem 2.5.11

Let V be a finite-dimensional vector space with ordered bases α, β, and γ.

a.) Suppose $Q = [I_V]_\alpha^\beta$ and $R = [I_V]_\beta^\gamma$. Then $RQ = [I_V]_\beta^\gamma [I_V]_\alpha^\beta = [I_V I_V]_\alpha^\gamma = [I_V]_\alpha^\gamma$ by Theorem 2.11. Hence RQ is the change of coordinate matrix that changes α coordinates into γ coordniates. ■

b.) Suppose $Q = [I_V]_\alpha^\beta$. Then by definition of a change of coordinate matrix, Q is invertible, and so we see that $Q^{-1} = ([I_V]_\alpha^\beta)^{-1} = [I_V^{-1}]_\beta^\alpha = [I_V]_\beta^\alpha$. Therefore Q^{-1} is the change of coordinate matrix changing β coordinates to α coordinates. ■

Problem 2.5.12

We will prove the corollary to Theorem 2.23. Let $A \in M_n(\mathbb{K})$ and let γ be an ordered basis for $\mathbb{K}^n$. Let β be the standard ordered basis for $\mathbb{K}^n$. We have the left-multiplication transformation $L_A : \mathbb{K}^n \to \mathbb{K}^n$ as a linear operator, and $[L_A]_\beta = A$ by defintion. Since L_A is a linear operator, and β and γ are ordered bases for $\mathbb{K}^n$, then the hypothesis of Theorem 2.23 is satisfied, and hence we have:

$$[L_A]_\gamma = Q^{-1}[L_A]_\beta Q \implies [L_A]_\gamma = Q^{-1}AQ$$

But also note that $[L_A]_\gamma = Q^{-1}[L_A]_\beta Q = [I_{\mathbb{K}^n}]_\beta^\gamma [L_A]_\beta [_{\mathbb{K}^n}]_\gamma^\beta = [L_A]_\beta^\gamma [I_{\mathbb{K}^n}]_\gamma^\beta = [L_A]_\gamma$, so that Q necessarily has the jth vector of γ in its jth column. ■

Problem 2.5.13
Let V be a finite-dimensional vector space over $\mathbb{K}$, and let $\beta = \{x_1, x_2, \ldots, x_n\}$ be an ordered basis for V. Let $Q \in M_n(\mathbb{K})$ be invertible. Define:

$$x'_j = \sum_{i=1}^{n} Q_{ij} x_i$$

for $1 \le j \le n$. Set $\beta' = \{x'_1, x'_2, \ldots, x'_n\}$. Then suppose that for $a_1, \ldots, a_n \in \mathbb{K}$:

$$\sum_{j=1}^{n} a_j x'_j = \vec{0} \implies \sum_{j=1}^{n} a_j (\sum_{i=1}^{n} Q_{ij} x_i) = \vec{0} \implies \sum_{j=1}^{n} \sum_{i=1}^{n} (a_j Q_{ij}) x_i = \vec{0}$$

Since β is a linearly independent set, the only linear combination of the x_i to give the zero vector is the trivial combination, and hence $a_j Q_{ij} = 0$ for $1 \le i, j \le n$. Therefore the set β' is linearly independent. By corollary 2 to Theorem 1.10, since β' is a linearly independent subset of V with exactly n vectors, and $\dim(V) = n$ by having its basis as β, then β' is a basis for V. Hence we have shown that $Q = [I_V]_{\beta'}^{\beta}$ is the change of coordinate matrix that changes β'-coordinates into β-coordinates. ■

Problem 2.5.14 XX
Let $A, B \in M_{m \times n}(\mathbb{K})$. Suppose there exist $P, Q \in M_{m \times n}(\mathbb{K})$ such that P, Q are invertible and $B = P^{-1}AQ$. Consider $\mathbb{K}^n$ with standard ordered basis β and $\mathbb{K}^m$ with standard ordered basis γ, and the left-multiplication transformation $L_A : \mathbb{K}^n \to \mathbb{K}^m$. Let β' be another ordered basis for $\mathbb{K}^n$ and let γ' be another ordered basis for $\mathbb{K}^m$. By Problem 2.5.13, Q is the change of coordinate matrix changing β' coordinates to β coordinates, $Q = [I_{\mathbb{K}^n}]_{\beta'}^{\beta}$, and similarly $P = [I_{\mathbb{K}^m}]_{\gamma'}^{\gamma}$. Note that we have:

$$B = P^{-1}AQ \implies [L_A]_{\beta'}^{\gamma'} = [I_{\mathbb{K}^m}]_{\gamma}^{\gamma'} [L_A]_{\beta}^{\gamma} [I_{\mathbb{K}^n}]_{\gamma'}^{\gamma}$$

2.6 Dual Spaces

Problem 2.6.1
a.) False. Linear functionals bring vectors to scalars.

b.) True.

c.) True. But does not hold in the infinite-dimensional case.

d.) True.

e.) False. Not necessarily.

f.) True. $([T^T]_{\gamma'}^{\beta'})^T = [T]$.

g.) True. But doesn't hold in the infinite-dimensional case.

h.) False. The derivative of a function is not necessarily a scalar.

Problem 2.6.2
a.) The map f is a linear functional on $P(\mathbb{R})$.

b.) The map f is not a linear functional on $\mathbb{R}^2$.

c.) The map f is a linear functional on $M_2(\mathbb{K})$.

d.) f is not linear, and therefore cannot be a linear functional on $\mathbb{R}^3$.

e.) The map f is a linear functional on $P(\mathbb{R})$.

f.) The map f is a linear functional on $M_2(\mathbb{K})$.

Problem 2.6.3
a.) We will find a dual basis $\beta^* = \{\phi_1, \phi_2, \phi_3\}$. We consider:

$$1 = \phi_1(1, 0, 1) = \phi_1(\vec{e}_1) + \phi_1(\vec{e}_3)$$

$$0 = \phi_1(1, 2, 1) = \phi_1(\vec{e}_1) + 2\phi_1(\vec{e}_2) + \phi_1(\vec{e}_3)$$
$$0 = \phi_1(0, 0, 1) = \phi_1(\vec{e}_3)$$

Thus we can see that $\phi_1(\vec{e}_3) = 0$, $\phi_1(\vec{e}_1) = 1$, and $\phi_1(\vec{e}_2) = -1/2$. Therefore $\phi_1(x_1, x_2, x_3) = x_1 + (-1/2)x_2$. Similarly, we have:

$$0 = \phi_2(1, 0, 1) \implies \phi_2(\vec{e}_1) = -\phi_2(\vec{e}_3)$$

$$1 = \phi_2(1, 2, 1) \implies \phi_2(\vec{e}_2) = 1/2$$
$$0 = \phi_2(0, 0, 1) \implies \phi_2(\vec{e}_3) = 0$$

Thus $\phi_2(x_1, x_2, x_3) = (1/2)x_2$. Finally, we observe:

$$0 = \phi_3(1, 0, 1) \implies \phi_3(\vec{e}_1) = -1$$

$$0 = \phi_3(1, 2, 1) \implies \phi_3(\vec{e}_2) = 0$$
$$1 = \phi_3(0, 0, 1) \implies \phi_3(\vec{e}_3) = 1$$

And so $\phi_3(x_1, x_2, x_3) = x_3 - x_1$. Given these calculations, we can conclude that the dual basis for β is as follows: $\beta^* = \{x - (1/2)y, (1/2)y, -y + z\}$.

b.) Now let $\beta = \{1, x, x^2\}$ be the standard ordered basis for $P_2(\mathbb{R})$. We will find $\beta^* = \{\phi_1, \phi_2, \phi_3\}$. First, we see that:

$$1 = \phi_1(1) \implies \phi_1(\vec{e}_1) = 1$$

$$1 = \phi_2(x) \implies \phi_2(\vec{e}_2) = 1$$

$$1 = \phi_3(x^2) \implies \phi_3(\vec{e}_3) = 1$$

And hence $\beta^* = \{a, b, c\}$, where $f(x) = a + bx + cx^2$ is any element of $P_2(\mathbb{R})$.

Problem 2.6.4

Problem 2.6.5

Problem 2.6.6

Problem 2.6.7

Problem 2.6.8
Consider the plane through the origin $0 = ax + by + cz$, with $a, b, c \in \mathbb{R}$ in the vector space $\mathbb{R}^3$, which can be represented by the vector $(a, b, c) \in \mathbb{R}^3$. Let $\eta \in (\mathbb{R}^3)^*$ such that $\eta(a, b, c) = 0$. Then clearly we see that $(a, b, c) \in N(\eta)$. As such we can identify any plane through the origin in $\mathbb{R}^3$ with the null space of some linear functional in the dual space of $\mathbb{R}^3$. Similarly for $\mathbb{R}^2$, any line passing through the origin can be identified with the null space of a linear functional in $(\mathbb{R}^2)^*$.

Problem 2.6.9
Suppose the mapping $T : \mathbb{K}^n \to \mathbb{K}^m$ is linear. Define $\eta_i(x) =$

Problem 2.6.10

Problem 2.6.11
Let V and W be finite-dimensional vector spaces over $\mathbb{K}$. Consider the isomorphisms $\psi_1 : V \to V^{**}$ and $\psi_2 : W \to W^{**}$ as defined in Theorem 2.26. Let $T : V \to W$ be linear and define $(T^T)^T = T^{TT}$. Let $x \in V$. Then:

$$(\psi_2 T)(x) = \psi_2(T(x)) = \widehat{T(x)}$$

$$(T^{TT}\psi_1)(x) = T^{TT}(\psi_1(x)) = T^{TT}(\widehat{x}) = (T^T)^T(\widehat{x}) = T$$

Problem 2.6.12

Problem 2.6.13
a.) Let V be a finite-dimensional vector space over $\mathbb{K}$ and $S \subseteq V$. Define $S^0 = \{\phi \in V^* : \phi(x) = 0\ \forall x \in V\}$. Clearly the functional that maps all vectors in V to 0 is in S^0, which is the zero element of V^*. Let $\phi, \eta \in S^0$ and $\mu \in \mathbb{K}$. Then note that:

$$(\mu\phi + \eta)(x) = \mu\phi(x) + \eta(x) = 0 + 0 = 0 \implies \mu\phi + \eta \in S^0$$

So that $S^0 \subseteq V^*$ contains the zero vector and is closed under vector addition and scalar multiplication, and is hence a subspace. ■

b.) We have $(S^0)^0 = \{\phi \in V^{**} : \phi(\eta) = 0\ \forall \eta \in S^0\}$. Now let $\phi \in (S^0)^0$. Then $\phi(\eta) = 0$ for any $\eta \in S^0 \implies \phi(\eta(x)) = 0$ for all $x \in S$. But then $\phi \in$

Let $\phi \in \operatorname{span}(\psi(S))$, so that $\phi(\eta) = 0$ for a

c.)

Problem 2.6.14

Problem 2.6.15

Problem 2.6.16

Problem 2.6.17

Problem 2.6.18

Problem 2.6.19

Problem 2.6.20
Let $V, W \neq \{\vec{0}\}$ vector spaces over a field $\mathbb{K}$. Let $T \in \mathcal{L}(V, W)$.

a.) Suppose that T is surjective. Then, in particular, we know that $R(T) = W$. Consider the dual map $T^t : W^* \to V^*$. Let $\phi, \psi \in W^*$. Suppose that $T^t(\phi) = T^t(\psi) \implies \phi \circ T = \psi \circ T \implies \phi = \psi$ by the right cancellation law for T, since it is surjective. Hence, T^t is injective. Conversely, suppose that T^t is in fact injective. Now suppose by way of contradiction that $R(T) \neq W$.

b.)

2.7 Homogenous Linear Differential Equations with Constant Coefficients

3 Elementary Matrix Operations and Systems of Linear Equations

4 Determinants

4.1 Determinants of Order 2

Problem 4.1.1
a.) False. Refer to discussion above Theorem 4.1.

b.) True. This is Theorem 4.1.

c.) False. $\det(A) = 0 \iff A$ not invertible by Theorem 4.2.

d.) False. The area of the parallelogram is equal to the absolute value of the determinant of u, v.

e.) True. Refer to discussion of handedness of coordinate systems.

Problem 4.1.2
a.) $\det(A) = 24 - (-6) = 30$

b.) $\det(B) = -5 - 12 = -17$

c.) $\det(C) = -8 - 0 = -8$

Problem 4.1.3
a.) $\det(A) = -2 + 3i + 2i - 3i^2 - (3 + 2i - 12i - 8i^2) = -10 + 15i$

b.) $\det(B) = 35i - 14i^2 - (-18 + 6i - 12i + 4i^2) = 36 + 41i$

c.) $\det(C) = 12i^2 - 12 = -12 - 12 = -24$

Problem 4.1.4
a.)

$$A = \begin{pmatrix} 3 & 2 \\ -2 & 5 \end{pmatrix} \implies \det(A) = 15 - (-4) = 19$$

b.)

$$A = \begin{pmatrix} 1 & -3 \\ 3 & 1 \end{pmatrix} \implies \det(B) = 1 - (-9) = 10$$

c.)

$$A = \begin{pmatrix} 4 & -6 \\ -1 & -2 \end{pmatrix} \implies \det(A) = -8 - 6 = -14$$

d.)

$$A = \begin{pmatrix} 3 & 2 \\ 4 & -6 \end{pmatrix} \implies \det(A) = -18 - 8 = 10$$

Problem 4.1.5
Let $A, B \in M_2(\mathbb{K})$ such that B is formed by interchanging the rows of A, so that we have now, with $a, b, c, d \in \mathbb{K}$:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$

Note that in terms of the determinant, we now have $\det(A) = ad - bc$ as per the definition, but also observe that:

$$\det(B) = bc - ad = -(-bc + ad) = -(ad - bc) = -\det(A)$$

And hence we have shown the required relation since B is the matrix obtained by interchanging rows of A. ∎

Problem 4.1.6
Let $A \in M_2(\mathbb{K})$ with arbitrary $a, c \in \mathbb{K}$ such that A has two identical columns. Then we have that:

$$A = \begin{pmatrix} a & a \\ c & c \end{pmatrix} \implies \det(A) = ac - ac = 0$$

Hence if A is a 2×2 matrix that has two identical columns then A necessarily has a zero determinant. ∎

Problem 4.1.7
Let $A \in M_2(\mathbb{K})$ be arbitrary, say with $a, b, c, d \in \mathbb{K}$ such that:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

Then we have $\det(A) = ad - bc$. But also $\det(A^T) = ad - bc = \det(A)$. Hence we have shown that 2×2 matrixes have identical determinant to their transposes. ∎

Problem 4.1.8
Let $A \in M_2(\mathbb{K})$ be an upper triangular matrix, and let $a, b, d \in \mathbb{K}$ be arbitrary such that:

$$A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \implies \det(A) = ad - 0(b) = ad$$

Hence the determinant of a 2×2 upper triangular matrix A is the product of the diagonal entries of A, in this case, ad, as desired. ∎

Problem 4.1.9
Let $A, B \in M_2(\mathbb{K})$ such that:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

Note that $\det(A) = ad - bc$ and $\det(B) = \alpha\delta - \beta\gamma$. Then $AB \in M_2(\mathbb{K})$ also, shown as follows:

$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{pmatrix}$$

However then we can see that:

$$\det(AB) = (a\alpha + b\gamma)(c\beta + d\delta) - (a\beta + b\delta)(c\alpha + d\gamma) =$$

$$ac\alpha\beta + ad\alpha\delta + bc\beta\gamma + bd\gamma\delta - ac\alpha\beta - ad\beta\gamma - bc\alpha\delta - bd\gamma\delta) =$$

$$ad\alpha\delta + bc\beta\gamma - ad\beta\gamma - bc\alpha\delta = (ad - bc)(\alpha\delta - \beta\gamma) = \det(A) \cdot \det(B)$$

Hence the determinant of the product of two 2×2 matrices is equal to the product of the determinant of both matrices taken individually as desired. More simply, $\det(AB) = \det(A)\det(B)$ for any $A, B \in M_2(\mathbb{K})$. ■

Problem 4.1.10
a.) Let $A \in M_2(\mathbb{K})$. Let C be the classical adjoint of A as defined in the problem. Then observe that:

$$CA = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} ad - bc & bd - bd \\ -ac + ac & -bc + ad \end{pmatrix} = (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$AC = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & bd - bd \\ -ac + ac & -bc + ad \end{pmatrix} = (ad - bc) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Hence we can see that $CA = AC = (ad - bc)I_2 = [\det(A)]I_2$. ■

b.) Note that $\det(A) = ad - bc$ and $\det(C) = da - bc = ad - bc$, so that necessarily $\det(A) = \det(C)$, so the determinant of a 2×2 matrix is equal to the determinant of its classical adjoint. ■

c.) Let $A \in M_2(\mathbb{K})$ be as above, and notice that:

$$A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \implies C^T = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

Which is indeed the transpose of the classical adjoint of A as seen before. Thus the classical adjoint of A^T is C^T as per the definition given in the problem statement. ■

d.) Let $A \in M_2(\mathbb{K})$. Suppose that A is invertible. Then by Theorem 4.2, we know that a formula for the inverse of A is given by:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{pmatrix} = [\det(A)]^{-1} C$$

By definition of the classical adjoint of A defined in the problem. ■

Problem 4.1.11
Let $\delta : M_2(\mathbb{K}) \to \mathbb{K}$ be a function. It is shown in the text of the section that the function δ is equivalent to the determinant function det.

Problem 4.1.12
Let $\{u, v\}$ be an ordered basis for $\mathbb{R}^2$. Suppose that $O\begin{pmatrix} u \\ v \end{pmatrix} = 1$. Then by definition of orientation, we know that:

$$1 = \frac{\det(\begin{pmatrix} u \\ v \end{pmatrix})}{|\det(\begin{pmatrix} u \\ v \end{pmatrix})|} \implies \det(\begin{pmatrix} u \\ v \end{pmatrix}) > 0$$

And hence since the area of the parallelogram formed by u and v can be associated to the determinant above, then we know that the area of the parallelogram is positive, meaning that both u and v are above the x-axis. Hence the coordinate system is right-handed. Conversely, suppose that the coordinate system is right-handed. Then by the same logic as above the determinant must be positive and thus $O(\begin{pmatrix} u \\ v \end{pmatrix}) = 1$. ■

4.2 Determinants of Order n

Problem 4.2.1
a.) False. Only when the other rows are held constant.

b.) True. This is Theorem 4.4

c.) True. This is the corollary to Theorem 4.4.

d.) True. This is Theorem 4.5.

e.) False. $\det(B) = k\det(A)$ by determinant properties.

f.) False. $\det(B) = \det(A)$, from Theorem 4.6.

g.) False. $\det(A) \neq 0$.

h.) True. Review Example 4 and see Problem 4.2.23.

Problem 4.2.2
We can see that by multiplying each row of the matrix on the right by 3 gives us the matrix on the left. By the properties of determinants, this is equivalent to $k = 3^3 = 27$.

Problem 4.2.3
If we take the matrix on the right and then multiply the first row by 2, the second row by 3, add a multiple (5) of the third row to the second row, and then multiply the second row by 7, we obtain the matrix on the left. By the properties of determinants, we have $k = 2 \cdot 3 \cdot 7 = 42$.

Problem 4.2.4
Take the matrix on the right and interchange the first and second rows. Then add the third row to the first row, followed by adding the first row minus the third row to the second row. Then we obtain the matrix on the left. Hence by the properties of the determinant, $k = -1$.

Problem 4.2.5

$$\det(A) = -12$$

Problem 4.2.6

$$\det(A) = -13$$

Problem 4.2.7

$$\det(A) = -12$$

Problem 4.2.8

$$\det(A) = -13$$

Problem 4.2.9

$$\det(A) = 22$$

Problem 4.2.10

$$\det(A) = 4 + 2i$$

Problem 4.2.11

$$\det(A) = -23$$

Problem 4.2.12
This is a massive calculation, I will not be doing this out. I would advise everyone to not as well, very frustrating problem.

Problem 4.2.13

$$\det(A) = (-1)^{1+3} \cdot 1 \cdot \begin{vmatrix} 0 & 2 \\ 4 & 5 \end{vmatrix} = 0 - 8 = -8$$

Problem 4.2.14

$$\det(A) = (-1)^{3+1} \cdot 7 \cdot \begin{vmatrix} 3 & 4 \\ 6 & 0 \end{vmatrix} = 1 \cdot 7(0 - 24) = 7(-24) = -168$$

Problem 4.2.15

$$\det(A) = 0$$

Problem 4.2.16

$$\det(A) = 36$$

Problem 4.2.17

$$\det(A) = -49$$

Problem 4.2.18

Problem 4.2.19

Problem 4.2.20

Problem 4.2.21

Problem 4.2.22

Problem 4.2.23
Let $A \in M_n(\mathbb{K})$ be upper triangular, specifically:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ 0 & b_2 & b_3 & \cdots & b_n \\ 0 & 0 & c_3 & \cdots & c_n \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & x_n \end{pmatrix}$$

By Theorem 4.4, $\det(A)$ can be evaluated by cofactor expansion along any row, choose the nth row. Then:

$$\det(A) = (-1)^{n+n} x_n \det(A_{nn}) = (-1)^{2n} x_n \det(A_{nn}) = x_n \det(A_{nn})$$

But then note that by removing row n and column n, we then see:

$$\det(A_{nn}) = (-1)^{(n-1)+(n-1)} y_{n-1} \det(A_{(n-1)(n-1)}) =$$

$$(-1)^{2(n-1)} y_{n-1} \det(A_{(n-1)(n-1)}) = y_{n-1} \det(A_{(n-1)(n-1)})$$

So that now we have from the initial equation:

$$\det(A) = x_n(y_{n-1} \det(A_{(n-1)(n-1)}))$$

Iterating this process until we reach the first row of the matrix, we eventually find that:

$$\det(A) = a_1 b_2 c_3 \cdots y_{n-1} x_n$$

Which is simply the product of all of the diagonal entries of the matrix A. Since $A \in M_n(\mathbb{K})$ was arbitrary, we necessarily have that the determinant of an upper triangular matrix is the product of the entries on its main diagonal as desired. ■

Problem 4.2.24
We will prove the corollary to Theorem 4.3. Suppose that $A \in M_n(\mathbb{K})$ has a row consisting entirely of zeros, say row i, $1 \le i \le n$. Then, by Theorem 4.4, choose to evaluate the determinant of A by cofactor expansion across the row i containing all zeros. Clearly we find that:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot 0 \cdot \det(A_{ij}) = 0 + 0 + \cdots + 0 = 0$$

And hence if a square matrix A contains a row of only zeros, it necessarily has a zero determinant. ■

Problem 4.2.25
Let $A \in M_n(\mathbb{K})$. Then let $k \in \mathbb{K}$, and consider kA. We can view this as multiplying each row of the matrix A by k. In terms of the properties of determinants, this is equivalent to multiplying $k \cdot k \cdots k$ exactly n times. Hence $\det(kA) = k^n \det(A)$. ■

Problem 4.2.26
Note that since $A \in M_n(\mathbb{K})$, then $(-A) = (-1)A$, so that we can apply the determinant property with $k = -1$, and find that by Problem 4.2.25, we have:

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

So that whenever the matrix has an even number of rows and columns, then $\det(-A) = \det(A)$. Note that whenever the matrix A has an odd number of rows and columns, then $\det(-A) = -\det(A)$.

Problem 4.2.27
Let $A \in M_n(\mathbb{K})$. Suppose that A has two identical columns. Then the columns of A are necessarily linearly dependent, and thus the rank of A is less than n. By the corollary to Theorem 4.3, we have that $\det(A) = 0$. ■

Problem 4.2.28
Let E_1 denote an elementary matrix formed by interchanging rows of I_n. Then by Theorem 4.5, $\det(E_1) = -\det(I_n) = -1$. Now let E_2 denote an elementary matrix performed by multiplying a row of I_n by a nonzero scalar k. Then by the properties of determinants, we know that $\det(E_2) = k\det(I_n) = k \cdot 1 = k$. Now finally let E_3 denote the elementary matrix obtained by adding a multiple of one row to another row of I_n. Then by Theorem 4.6, $\det(E_3) = \det(I_n) = 1$.

Problem 4.2.29

Problem 4.2.30
Let the rows of $A \in M_n(\mathbb{K})$ be $a_1, a_2, \ldots, a_n$, and let B be the maatrix in which rows are $a_n, a_{n-1}, \ldots, a_1$. So that we have:

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad B = \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_1 \end{pmatrix}$$

Note that we can get from matrix A to matrix B by interchanging rows, specifically by interchanfing the first row with the last, and then the second with the second to last, and so on. Note that if n is even, then there are precisely $n/2$ interchanges. If n is odd, then there will be $(n-1)/2$ interchanges required. Hence in the case where n is even, we find that:

$$\det(B) = (-1)^{n/2}\det(A)$$

While if n is an odd number, then we find:

$$\det(B) = (-1)^{(n-1)/2}\det(A)$$

5 Diagonalization

5.1 Eigenvalues and Eigenvectors

Problem 5.1.1
a.) False. Consider T_θ, as defined in the section.

b.) True. Any multiple of the eigenvector also satisfies the properties of an eigenvector.

c.) True. $O \in M_n(\mathbb{K})$ has no eigenvectors since it has no eigenvalues.

d.) False. Refer to the definition of eigenvalue. Scalars including zero.

e.) False.

f.) False. Best shown using an example.

g.) False. Consider the example of C^∞ given in the section.

h.) True. Since similar matrices are just the same linear operator under a different ordered basis, if a matrix is similar to a diagonal matrix, then by Theorem 5.1, there is a basis of eigenvectors.

i.) True.

j.) False.

k.) False. Best shown using an example.

Problem 5.1.2
a.)

$$[T]_\beta = \begin{pmatrix} 0 & 2 \\ -1 & 0 \end{pmatrix}$$

Also β is not a basis consisting of eigenvectors of T.

b.)

$$[T]_\beta = \begin{pmatrix} -2 & 0 \\ 0 & -3 \end{pmatrix}$$

Also β is indeed a basis consisting of eigenvectors of T.

c.)

$$[T]_\beta = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Also β is indeed a basis consisting of eigenvectors of T.

d.)

$$[T]_\beta = \begin{pmatrix} 4 & 0 & 3 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Also β is not a basis consisting of eigenvectors of T.

e.)

$$[T]_\beta = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Also we find that β is not a basis consisting of eigenvectors of T.

f.)

$$[T]_\beta = \begin{pmatrix} -3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

We find that β is in fact a basis consisting of eigenvectors of T.

Problem 5.1.3
a.) The eigenvalues of the matrix A are $\lambda = 4$ and $\lambda = -1$. For $\lambda = 4$, we have a set of eigenvectors:

$$\{t \begin{pmatrix} 1 \\ -1 \end{pmatrix} : t \in \mathbb{R}\}$$

And for the eigenvalue $\lambda = -1$, we have another set of eigenvectors as follows:

$$\{t \begin{pmatrix} 2 \\ 3 \end{pmatrix} : t \in \mathbb{R}\}$$

Hence a basis β for $\mathbb{R}^2$ is given by $\beta = \{(2,3),(1,-1)\}$. We now see:

$$Q = \begin{pmatrix} 2 & 1 \\ 3 & -1 \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} -1 & -1 \\ -3 & 2 \end{pmatrix}, \quad D = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$$

b.)

c.) The eigenvalues of the matrix A are $\lambda = 1$ and $\lambda = -1$. For $\lambda = 1$, we have a set of eigenvectors:

$$\{t \begin{pmatrix} 1 \\ 1-i \end{pmatrix} : t \in \mathbb{C}\}$$

And for the eigenvalue $\lambda = -1$, we have another set of eigenvectors as follows:

$$\{t \begin{pmatrix} i-1 \\ 2 \end{pmatrix} : t \in \mathbb{C}\}$$

Hence a basis β for $\mathbb{C}^2$ is given by $\beta = \{(1, 1-i), (i-1, 2)\}$. And we now see:

$$Q = \begin{pmatrix} 1 & i-1 \\ 1-i & 2 \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} 2 & 1-i \\ i-1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

d.) The only eigenvalue for the matrix A is $\lambda = 1$. The set of eigenvectors corresponding to this eigenvalue is as follows:

$$\{s \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} : s, t \in \mathbb{R}\}$$

Since A has only 1 eigenvalue, and $\dim(\mathbb{R}^3) = 3$, then there does not exist a basis for $\mathbb{R}^3$ such that

Problem 5.1.4 XXX
a.) to do

Problem 5.1.5
Theorem 5.4: Let T be a linear operator on a vector space V, and let λ be an eigenvalue of T. Suppose $v \in V$ is an eigenvector of T corresponding to λ. By definition of eigenvector, we know that $v \neq \vec{0}$, and:

$$T(v) = \lambda v \implies T(v) - \lambda v = (T - \lambda I)v = \vec{0}$$

And so $v \in V$, when acted on by the mapping $(T - \lambda I)$, maps to the zero vector. Hence by definition $v \in N(T - \lambda I)$. Conversely, suppose that $v \neq \vec{0}$ and $v \in N(T - \lambda I)$. Then $(T - \lambda I)v = \vec{0} \implies T(v) - \lambda v = \vec{0} \implies T(v) = \lambda v$. Thus by definition v is an eigenvector of T corresponding to λ as desired. $\blacksquare$

Problem 5.1.6
Let T be a linear operator on a finite-dimensional vector space V, and let β be an ordered basis for V. Suppose that λ is an eigenvalue of T. Then from the definition of eigenvector, there is a $v \in V$ such that $v \neq \vec{0}$ and $T(v) = \lambda v$. Then since $T(v) = [T]_\beta [v]_\beta$, we have that $[T]_\beta [v]_\beta = \lambda [v]_\beta$, so that by definition λ is an eigenvalue of $[T]_\beta$. Conversely, suppose that λ is an eigenvalue of $[T]_\beta$. Then there is a $v \in \mathbb{K}^n$ such that $v \neq \vec{0}$ and $[T]_\beta [v]_\beta = \lambda [v]_\beta$. But since $[T]_\beta [v]_\beta = T(v) \implies T(v) = \lambda v$ so that λ is an eigenvalue of T as desired. $\blacksquare$

Problem 5.1.7
Let T be a linear operator on a finite-dimensional vector space V. Define

$\det(T) = \det([T]_\beta)$, where β is any ordered basis for V.

a.) We will show that the definition of $\det(T)$ as defined above is independent of choice of basis for V. Let β and γ be two ordered bases for V. Then from the above definition we know that $\det(T) = \det([T]_\beta)$ and $\det(T) = \det([T]_\gamma)$. Hence we can easily see that $\det([T]_\beta) = \det([T]_\gamma)$. $\blacksquare$

b.) Suppose that T is invertible. Let β be a basis for V. Then by Theorem 2.18, since T is invertible we have that $[T]_\beta$ is invertible. By the corollary to Theorem 4.7, we know that $\det([T]_\beta) \neq 0$. Hence from the above definition, we have that $\det([T]_\beta) = \det(T) \neq 0$. Conversely, suppose that $\det(T) \neq 0$. Then we immediately know that $\det([T]_\beta) \neq 0$, so that $[T]_\beta$ is necessarily invertible. Hence by the one-to-one correspondence to T, we can see that T is invertible as well. $\blacksquare$

c.) Suppose T is invertible. By Theorem 2.18 and by the definition of the determinant of a linear operator given in the problem, we see:

$$\det(T^{-1}) = \det([T^{-1}]_\beta) = \det([T]_\beta^{-1})$$

And since $[T]_\beta^{-1}$ is invertible, $\det([T]_\beta)\det([T]_\beta^{-1}) = 1$, and we necessarily have:

$$\det(T^{-1}) = \det([T]_\beta^{-1}) = \frac{1}{\det([T]_\beta)} = \frac{1}{\det(T)} \implies \det(T^{-1}) = [\det(T)]^{-1}$$

d.) Suppose U is another linear operator on V. Then $\det(U) = \det([U]_\beta)$ by the definition given in the problem. Then observe that:

$$\det(TU) = \det([TU]_\beta) = \det([T]_\beta[U]_\beta) =$$

$$\det([T]_\beta)\det([U]_\beta) = \det(T)\det(U)$$

Which follows from the corollary to Theorem 2.11 when applied to U and T, as well as Theorem 4.7, which allows us to separate the matrix product. $\blacksquare$

e.) Let β be an arbitrary ordered basis for V and $\lambda \in \mathbb{K}$. Then observe that:

$$\det(T - \lambda I_V) = \det([T - \lambda I]_\beta) = \det([T]_\beta - \lambda[I_V]_\beta) = \det([T]_\beta - \lambda I)$$

Where the linearity follows from Theorem 2.8 and the fact that $[I_V]_\beta = I$ follows from Theorem 2.12d. $\blacksquare$

Problem 5.1.8

a.) Let T be a linear operator on a finite-dimensional vector space. Suppose T is invertible. Suppose by way of contradiction that 0 is an eigenvalue of T. Then by definition there exists some $v \in V$ such that $v \neq \vec{0}$ and $T(v) = 0v = \vec{0}$. However since we know that $N(T) = \{\vec{0}\}$ since T is a bijection, and so for any $v \in V$, $T(v) = \vec{0} \iff v = \vec{0}$. This is a contradiction to v being an eigenvector and thus 0 is not an eigenvalue of T. Conversely, suppose by way of contrapositive that T is not invertible. Then in particular we know that $N(T) \neq \{\vec{0}\}$, so that there exists some $v \in V$ such that $v \in N(T)$ but $v \neq \vec{0}$. Hence $T(v) = \vec{0} = 0v$, so that by definition v is an eigenvector of T corresponding to the eigenvalue 0. Hence 0 is an eigenvalue of T. Therefore we can conclude by the contrapositive statement that if 0 is not an eigenvalue of T, then T is invertible. ■

b.) Let T be an invertible linear operator. Suppose $\lambda \in \mathbb{K}$ is an eigenvalue of T. Then there exists some $v \in V$ such that $v \neq \vec{0}$ and:

$$T(v) = \lambda v \implies v = T^{-1}(\lambda v) = \lambda T^{-1}(v) \implies \lambda^{-1} v = T^{-1}(v)$$

Which follows from the linearity of T^{-1}. We can arrange the above to see $T^{-1}(v) = \lambda^{-1} v$, so that λ^{-1} is an eigenvalue of T^{-1} by definition. Conversely, suppose that λ^{-1} is an eigenvalue of T^{-1}. There there is some $w \in V$ such that $w \neq \vec{0}$ and:

$$T^{-1}(w) = \lambda^{-1} w \implies w = T(\lambda^{-1} w) = \lambda^{-1} T(w) \implies \lambda w = T(w)$$

And hence we can easily see that since $T(w) = \lambda w$, that λ is also an eigenvalue of T. ■

c.) These analogous results for matrices are easily found by simply substituting T with its matrix representation.

Problem 5.1.9
Let $A \in M_n(\mathbb{K})$ such that A is an upper triangular matrix. By Problem 4.2.23, we know that $\det(A) = \Pi_{i=1}^{n} A_{ii}$. Now consider the matrix $A - \lambda I_n$, which is clearly also an upper triangular matrix. We see:

$$\det(A - \lambda I_n) = \Pi_{i=1}^{n} (A_{ii} - \lambda) = (A_{11} - \lambda)(A_{22} - \lambda) \cdots (A_{nn} - \lambda)$$

And notice that $\det(A - \lambda I_n) = 0$ precisely when $\lambda = A_{ii}$ for any $1 \leq i \leq n$. Hence each diagonal entry A_{ii} is an eigenvalue of A. Therefore the eigenvalues of an upper triangular matrix are the entries on its main diagonal. ■

Problem 5.1.10
Let V be a finite-dimensional vector space and $\lambda \in \mathbb{K}$.

a.) Let β be an ordered basis for V. Then $[\lambda I_V]_\beta = \lambda [I_V]_\beta = \lambda I$. All of which follows from Theorem 2.12. ■

b.) Let $\dim(V) = n$. Using the result of the first part of the problem, the characteristic polynomial of λI_V is defined to be:

$$f(t) = \det(\lambda I_V - tI) = \det(I - tI) = \det((1 - t)I) = (1 - t)^n$$

c.) From the results of the previous part we can easily see that λI_V has one eigenvalue, namely $\lambda = 1$. Let $\beta = \{v_1, \ldots, v_n\}$ be a basis for V. Notice that each of the basis vectors are in fact eigenvectors corresponding to $\lambda = 1$ since $\lambda I_V(v_i) = \lambda v_i$ for any $1 \leq i \leq n$. Hence β is a basis for V consisting of eigenvectors of λI_V, and so by Theorem 5.1, λI_V is diagonalizable. ■

Problem 5.1.11
a.) Let A be a square matrix. Suppose A is similar to a scalar matrix λI. Then there exists an invertible matrix Q such that $A = Q^{-1}(\lambda I)Q$. But by the properties of matrix multiplication shown in Theorem 2.12, we see:

$$A = Q^{-1}(\lambda I)Q = \lambda(Q^{-1}I)Q = \lambda(Q^{-1}Q) = \lambda I$$

Which is the desired result. ■

b.) Suppose A is an $n \times n$ diagonal matrix having only one eigenvalue, say λ. Then by Problem 5.1.9, since diagonal matrices are a subset of upper triangular matrices, then we know that the eigenvalues of A are the entries on the main diagonal. However, λ is the only eigenvalue of A, and thus every entry on the diagonal of A is $\lambda \implies A$ is a scalar matrix. ■

c.) Consider the matrix:

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

The characteristic polynomial of this matrix is easily seen to be $f(t) = (1-t)^2$, so that the only eigenvalue is $\lambda = 1$. However note that the set of eigenvectors corresponding to $\lambda = 1$ is:

$$\{\ t(1,0) : t \in \mathbb{K}\}$$

And clearly we can construct no basis consisting of eigenvectors of the matrix, and so by Theorem 5.1, it is not diagonalizable. ■

Problem 5.1.12
a.) Let A and B be $n \times n$ matrices. Suppose that B is similar to A. Then there exists an invertible matrix Q such that $A = Q^{-1}BQ$. But then since $\det(Q^{-1})\det(Q) = \det(Q^{-1}Q) = \det(I_n) = 1$, we can see:

$$\det(B - tI_n) = \det(Q^{-1})\det(Q)\det(B - tI_n) = \det(Q^{-1})\det(B - tI_n)\det(Q) =$$

$$\det(Q^{-1}(B - tI_n)Q) = \det((Q^{-1}B - tQ^{-1})Q) = \det(Q^{-1}BQ - tQ^{-1}Q) =$$

$$\det(Q^{-1}BQ - tI_n) = \det(A - tI_n)$$

And so we can see that the characteristic polynomials of A and B are in fact equivalent for any choice of t, and thus the same. ■

b.) Explicitly, let β and γ be ordered bases for a finite-dimensional vector space V, and let $T \in \mathcal{L}(V)$. Then:

$$[T]_\beta = Q^{-1}[T]_\gamma Q$$

But letting $A = [T]_\beta$ and $B = [T]_\gamma$ from the first part of the problem, we can immediately see that the two operators have identical characteristic polynomials. Therefore the choice of basis for a linear operator does not affect its characteristic polynomial.

Problem 5.1.13
Let T be a linear operator on a finite-dimensional vector space V over a field $\mathbb{K}$, let β be an ordered basis for V, and let $A = [T]_\beta$.

a.) Suppose $v \in V$ and $\phi_\beta(v)$ is an eigenvector of A corresponding to the eigenvalue λ. Then we know from the commutativity of the diagram that:

$$T(v) = \phi_\beta^{-1}([T]_\beta \phi_\beta(v)) = \phi_\beta^{-1}(\lambda \phi_\beta(v)) = \lambda(\phi_\beta^{-1}\phi_\beta)(v) = \lambda I_V v = \lambda v$$

So that v is an eigenvector of T corresponding to λ. ■

b.) Let λ be an eigenvalue of A. Suppose $y \in \mathbb{K}^n$ is an eigenvector of A corresponding to λ. So we have $Ay = [T]_\beta y = \lambda y$. Subsequently, from the commutativity of the diagram, we have:

$$T(\phi_\beta^{-1}(y)) = \phi_\beta^{-1}(L_A(y)) = \phi_\beta^{-1}(Ay) = \phi_\beta^{-1}(\lambda y) = \lambda \phi_\beta^{-1}(y)$$

Hence $\phi_\beta^{-1}(y)$ is an eigenvector of T corresponding to λ as desired. ■

Problem 5.1.14
Let $A \in M_n(\mathbb{K})$. Then by Theorem 4.8, we know that, in particular:

$$\det(A - \lambda I_n) = \det((A - \lambda I_n)^T) = \det(A^T - \lambda I_n^T) = \det(A^T - \lambda I_n)$$

Which follows from Problem 1.3.3 and the fact that the transpose of the identity matrix is again the identity matrix. Hence we can see that the characteristic polynomial of A is equivalent to the characterisitic polynomial of A^T. Hence a matrix and its transpose have identical eigenvalues as well. ■

Problem 5.1.15
a.) Let T be a linear operator on a finite-dimensional vector space V, and let x be an eigenvector of T corresponding to the eigenvalue λ. For our base case $n = 1$, we have immediately that $T(x) = T^1(x) = \lambda x$. Let $m \in \mathbb{Z}^+$. Suppose that the relation holds for the mth case. Then we have that $T^m(x) = \lambda^m x$. Then we can see that applying T once more:

$$T(T^m(x)) = T(\lambda^m x) = \lambda^m T(x) = \lambda^m \lambda x = \lambda^{m+1} x$$

So we have found that $T^{m+1}(x) = \lambda^{m+1}x$. Thus Hence the relation holds for the $(m+1)$th case and thus by the principle of induction, it holds for all $m \in \mathbb{Z}^+$. Therefore for any $m \in \mathbb{Z}^+$, x is an eigenvector of T^m corresponding to the eigenvalue λ^m. ■

b.)

Problem 5.1.16
a.) Let $A, B \in M_n(\mathbb{K})$. Suppose that A and B are similar. Then there exists an invertible matrix Q such that $A = Q^{-1}BQ$. Then by Problem 2.3.13, we know that:

$$\operatorname{tr}(A) = \operatorname{tr}(Q^{-1}BQ) = \operatorname{tr}((Q^{-1}B)Q) = \operatorname{tr}((BQ^{-1})Q) =$$
$$\operatorname{tr}(B(Q^{-1}Q)) = \operatorname{tr}(BI_n) = \operatorname{tr}(B)$$

And hence similar matrices have the same trace. ■

b.)

Problem 5.1.17
a.)

b.)

c.)

d.)

Problem 5.1.18
a.) Let $A, B \in M_n(\mathbb{C})$. Suppose that B is invertible. Now let $c \in \mathbb{C}$ and consider the matrix $A + cB$.

$$(A + cB)_{ij} = A_{ij} + cB_{ij} = 0 \implies c = -\frac{A_{ij}}{B_{ij}}$$

b.)

Problem 5.1.19

Problem 5.1.20
Let A be an $n \times n$ matrix with characteristic polynomial:

$$f(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

Now evaluate the above polynomial at $t = 0$. Doing so, we find that $f(0) = 0 + \cdots + 0 + a_0 = a_0$. Since the evaluation of the characteristic polynomial at 0 of a square matrix is equivalent to finding the determinant of the matrix, by the definition, we see:

$$f(0) = \det(A) = a_0$$

And by the corollary to Theorem 4.7, we know that A is invertible only when $\det(A) \neq 0 \implies a_0 \neq 0$. Similarly for the converse, if $a_0 \neq 0$ then we know immediately that $\det(A) \neq 0$, and so A is invertible. ■

Problem 5.1.21
a.) Let A and $f(t)$ be as in Problem 5.1.20. Set $n = 2$, so that we have $A \in M_2(\mathbb{K})$, and $f(t) = t^2 + a_1 t + a_0$. Note that we find, by definition of the characteristic polynomial for a matrix:

$$f(t) = t^2 - (A_{11} + A_{22})t + (A_{11}A_{22} - A_{21}A_{12}) =$$

b.)

Problem 5.1.22
a.) Let T be a linear operator on a vector space V over a field $\mathbb{K}$, let $g(t) \in \mathbb{K}[t]$, a polynomial with coefficients in $\mathbb{K}$. Suppose that $x \in V$ is an eigenvector of T corresponding to eigenvalue λ. Then $T(x) = \lambda x$. If $g(t) = \sum_{i=0}^{n} a_i x^i$, then we have then that:

$$g(T)(x) = g(T(x)) = a_0 T^0(x) + a_1 T(x) + a_2 T^2(x) + \cdots + a_n T^n(x) =$$
$$= a_0 x + a_1 T(x) + a_2 T^2(x) + \cdots + a_n T^n(x) =$$
$$= a_0 x + a_1 \lambda x + a_2 \lambda^2 x + \cdots + a_n \lambda^n x = g(\lambda)x$$

Which follows from the linearity property of T. Hence we can see that x satisfies the condition of an eigenvector of $g(T)$ corresponding to eigenvalue $g(\lambda)$. ■

b.) Simply set $A = [T]_\beta$, where β is some ordered basis for V. Done.

c.) From Problem 5.1.3a, we have the matrix $A \in M_2(\mathbb{R})$, where:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$$

Consider the polynomial $g(t) = 2t^2 - t + 1$, eigenvector $x = (2, 3)$, and corresponding eigenvalue $\lambda = 4$. From part b.) of this problem, we have:

$$g(A) = 2\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}^2 - \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 14 & 10 \\ 15 & 19 \end{pmatrix}$$

And note that:

$$g(A)\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 & 10 \\ 15 & 19 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 58 \\ 87 \end{pmatrix} = 29\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Which verifies with what we expected to obtain, as $g(4) = 29$.

Problem 5.1.23

Suppose that $f(t)$ is the characteristic polynomial of a diagonalizable linear operator $T \in \mathcal{L}(V)$. By Theorem 5.1, since T is diagonalizable, there exists a basis $\beta = \{x_1, x_2, \ldots, x_n\}$ for V consisting of eigenvectors of T. From Problem 5.1.22, we know that for any $x_i \in \beta$ with $1 \leq i \leq n$:

$$f(T)(x) = f(\lambda)x = 0x = \vec{0}$$

Where λ is the eigenvalue corresponding to the particular choice of x_i. This follows since any eigenvalue is necessarily the root of the characteristic polynomial of T by Theorem 5.2. Hence by Theorem 2.2, we know that:

$$R(f(T)) = \operatorname{span}(\{f(T)(x_1), f(T)(x_2), \ldots, f(T)(x_n)\}) = \operatorname{span}(\{\vec{0}\}) = \{\vec{0}\}$$

So that we necessarily have $f(T) = T_0$, in other words, $f(T)$ is the zero operator on V. ■

Problem 5.1.24

Problem 5.1.25
This is probably a typo in the text, as Theorem 5.3 is not shown to have any corollaries in the section. I guess we will never know what truths these corollaries hold.

Problem 5.1.26
We know that for each entry in any matrix in $M_2(\mathbb{Z}/2\mathbb{Z})$, there are two options from the underlying field, $\mathbb{Z}/2\mathbb{Z} = \{0, 1\}$. Let $A \in M_2(\mathbb{Z}/2\mathbb{Z})$ be arbitrary. Then:

$$\det(A - \lambda I_2) = \det(\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}) = (a-\lambda)(d-\lambda) - bc$$

And so for our characteristic polynomial for any arbitrary A, we necessarily have the equation:

$$(a-\lambda)(d-\lambda) - bc = ad - a\lambda - d\lambda + \lambda^2 - bc =$$

$$\lambda^2 - (a+d)\lambda + (ad - bc) = 0$$

Hence since each entry $a, b, c, d \in \mathbb{Z}/2\mathbb{Z}$ is one of two options, we are choosing from two options four times. However note that there are repeat polynomials, so that not all of the $2^4 = 16$ possible combinations result in distinct polynomials. Observe that there are 2 distinct options for the λ^1 term and 2 distinct options for the constant term. Hence the number of distinct characteristic polynomials of the matrices in $M_2(\mathbb{Z}/2\mathbb{Z})$ is: $2 \cdot 2 = 4$. These distinct possibilities are $\lambda^2 - \lambda + 1$, $\lambda^2 - \lambda$, $\lambda^2 + 1$, and λ^2.

5.2 Diagonalizability

Problem 5.2.1
a.) False. Consider the identity transformation.

b.) False. Consider an eigenvalue with eigenspace of dimension 2.

c.) False. $\vec{0} \in E_\lambda$.

d.) True. Eigenvectors corresponding to different eigenvalues are linearly independent by Theorem 5.5.

e.) True. This is Theorem 5.1 applied to matrices.

f.) False. Refer to Theorem 5.9, the characteristic polynomial must split.

g.) True.

h.) True. $V = \bigoplus_{i=1}^{k} W_i \implies W_i \cap W_j = \{\vec{0}\}$ by definition.

i.) False. The condition for direct sum is $W_j \cap (\sum_{i \neq j} W_i) = \{\vec{0}\}$.

Problem 5.2.2
a.) We can see that the characteristic polynomial of A is $f(t) = (1-t)^2$, so that indeed it splits. However note that for the single eigenvalue $\lambda = 1$, we find that the eigenspace is of the form $E_\lambda = \{(1,0)t : t \in \mathbb{R}\}$, so that clearly $\dim(E_\lambda) \neq 2$. Hence condition 2 of the diagonalizability test fails. Hence A is not diagonalizable.

b.) Here again, $f(t) = (\lambda - 4)(\lambda + 2)$, so that the characteristic polynomial of A splits. By the corollary to Theorem 5.5, since A has 2 distinct eigenvalues, it is diagonalizable. $E_4 = \{(1,1)t : t \in \mathbb{R}\}$ and $E_{-2} = \{(1,-1)t : t \in \mathbb{R}\}$. We find:

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Where the eigenvectors $(1,1)$ and $(1,-1)$ clearly form a basis for $\mathbb{R}^2$.

c.) $f(t) = (\lambda - 5)(\lambda + 2)$, and since A has two distinct eigenvalues it is diagonalizable. $E_5 = \{(1,1)t : t \in \mathbb{R}\}$ and $E_{-2} = \{(4,-3)t : t \in \mathbb{R}\}$. So we find eventually:

$$Q = \begin{pmatrix} 1 & 4 \\ 1 & -3 \end{pmatrix}$$

Where the eigenvectors $(1,1)$ and $(4,-3)$ clearly form a basis for $\mathbb{R}^2$.

d.)

e.) Note that the characteristic polynomial of A is $f(t) = (\lambda - 1)(\lambda^2 + 1)$, and so it does not split over $\mathbb{R}$. Hence A is not diagonalizable.

f.) For A we find the characteristic polynomial as $f(t) = (3 - \lambda)(1 - \lambda)^2$, so that indeed it splits over the field. Consider the eigenvector $\lambda = 1$. Then see $E_\lambda = \{(1, 0, 0)t : t \in \mathbb{R}\}$, and clearly $(E_\lambda) = 1$ and thus does not equal the multiplicity of $\lambda = 1$. Hence A is not diagonalizable.

g.)

Problem 5.2.3
a.)

b.)

c.)

d.)

e.)

f.)

Problem 5.2.4
Let $A \in M_n(\mathbb{K})$. Suppose that A has n distinct eigenvalues, say $\lambda_1, \ldots, \lambda_n$. For each i, choose an eigenvector v_i corresponding to λ_i. Noting that eigenvalues of A are eigenvalues of L_A, we know that the set $\{v_1, \ldots, v_n\}$ is linearly independent by Theorem 5.5. Since $\dim(\mathbb{K}^n) = n$, this set is clearly a basis for $\mathbb{K}^n$. Thus, by Theorem 5.1, L_A is diagonalizable, and by extension, the matrix A is diagonalizable as well. ■

Problem 5.2.5
Let $A \in M_n(\mathbb{K})$. Suppose that A is diagonalizable. Then there exists an invertible matrix Q and a diagonal matrix D such that $A = Q^{-1}DQ$. Suppose that:

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

and let $f(t)$ be the characteristic polynomial of A. Then:

$$f(t) = \det(A - tI_n) = \det(\begin{pmatrix} \lambda_1 - t & 0 & \cdots & 0 \\ 0 & \lambda_2 - t & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n - t \end{pmatrix}) =$$

$$= (\lambda_1 - t)(\lambda_2 - t)\cdots(\lambda_n - t) = (-1)^n(t - \lambda_1)(t - \lambda_2)\cdots(t - \lambda_n)$$

And clearly we can see that the above polynomial splits over the field $\mathbb{K}$. Hence the characteristic polynomial of any diagonalizable matrix splits. ■

Problem 5.2.6

Problem 5.2.7
The characteristic polynomial of A is easily found to be $f(t) = (t-5)(t+1)$. Therefore eigenvalues of A are $\lambda_1 = 5$ and $\lambda_2 = -1$. Since A has two distinct eigenvalues, then by Problem 5.2.4, A is diagonalizable. Note that we have two eigenspaces:

$$E_{\lambda_1} = \{(1,1)t : t \in \mathbb{R}\}$$

$$E_{\lambda_2} = \{(1,-2)t : t \in \mathbb{R}\}$$

And clearly the set $\{(1,1),(1,-2)\}$ is linearly independent and is thus a basis for $\mathbb{K}^2$. By Theorem 5.1, we can formulate an expression for A as:

$$A = Q^{-1}DQ = \begin{pmatrix} -2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}$$

And then we can see that a formular for A^n, $n \in \mathbb{Z}^+$, is given by:

$$A^n = (Q^{-1}DQ)^n = Q^{-1}D^nQ = \begin{pmatrix} -2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 5^n & 0 \\ 0 & (-1)^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}$$

The same holds true for any diagonalizable matrix A.

Problem 5.2.8
Suppose that $A \in M_n(\mathbb{K})$ has two distinct eigenvalues, λ_1 and λ_2, such that $\dim(E_{\lambda_1}) = n - 1$. Then immediately we know that there are $n - 1$ vectors in a basis for the eigenspace E_{λ_1}. But, note that since λ_2 is an eigenvalue, it has multiplicity greater than or equal to 1. Hence, associating $T = L_A$ in the hypothesis of Theorem 5.7, we see that $1 \leq \dim(E_{\lambda_2})$, so that there is at least one vector in the basis for E_{λ_2}. Let β_1 be a basis for E_{λ_1} and β_2 a basis for E_{λ_2}. Then by Theorem 5.8, $\beta_1 \cup \beta_2$ is a linearly independent set, and contains at least n vectors, and so it either is itself a basis for $\mathbb{K}^n$ or contains a basis for $\mathbb{K}^n$. Hence there must exist an ordered basis for $\mathbb{K}^n$ consisting of eigenvectors of A. So by Theorem 5.1, A is diagonalizable. ■

Problem 5.2.9
a.) Let T be a linear operator on a finite-dimensional vector space V, and suppose there exists an ordered basis β for V such that $[T]_\beta$ is an upper triangular matrix. Let $\dim(V) = n$, so that $[T]_\beta$ is an $n \times n$ matrix. Then, by Problem 4.2.23, we know that $\det([T]_\beta)$ is the product of the diagonal entries. Let a_{ij} denote the entry of $[T]_\beta$ in the ith row and jth column. The characteristic polynomial of $[T]_\beta$ is as follows:

$$f(t) = \det([T]_\beta - tI_n) = \prod_{i=1}^{n}(([T]_\beta)_{ii} - t) = (-1)^n(t - a_{11})(t - a_{22})\cdots(t - a_{nn})$$

So that clearly the characteristic polynomial of T splits over the field. Note also that the characteristic polynomial is independent of choice of basis. ■

b.) The results of the above proof are easily seen if we take $A \in M_n(\mathbb{K})$ and have $A = [T]_\beta$. Importantly, note also that since linear operators under different bases are equivalent to similar matrices, we have necessarily that similar matrices share the same characteristic polynomial. Specifically, if $A, B \in M_n(\mathbb{K})$ such that $f(t)$ is the characteristic polynomial of A, and $A \sim B$, then $f(t)$ is the characteristic polynomial of B.

Problem 5.2.10
Let V be a vector space with $\dim(V) = n$, and let $T \in \mathcal{L}(V)$, with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ and corresponding multiplicities $m_1, m_2, \ldots, m_k$. Now suppose that β is a basis for V such that $[T]_\beta$ is an upper triangular matrix. Let $f(t)$ be the characteristic polynomial of T. From Problem 5.2.9 we know that the characteristic polynomial of T splits. So for $c, a_1, \ldots, a_n \in \mathbb{K}$, we have:

$$f(t) = c(t - a_1)\cdots(t - a_n) = \det(T - tI_n)$$

But since $[T]_\beta - tI_n$ is also an upper triangular matrix, as was seen in Problem 5.2.9, we know that:

$$\det(T - tI_n) = \prod_{i=1}^{n}(([T]_\beta)_{ii} - t) \implies$$

$$f(t) = c(t - a_1)\cdots(t - a_n) = (([T]_\beta)_{11} - t)(([T]_\beta)_{22} - t)\cdots(([T]_\beta)_{nn} - t) =$$
$$= (-1)^n(t - ([T]_\beta)_{11})\cdots((t - ([T]_\beta)_{nn})$$

Which implies that $c = (-1)^n$ and $a_i = ([T]_\beta)_{ii}$ for $1 \leq i \leq n$. So we can see that when $t = ([T]_\beta)_{ii}$ for any $1 \leq i \leq n$, then $\det([T]_\beta - tI_n) = f(t) = 0 \iff ([T]_\beta)_{ii}$ is an eigenvalue of T, via Theorem 5.2. Hence each of the entries on the diagonal of $[T]_\beta$ must be the eigenvalues of T, namely, $\lambda_1, \lambda_2, \ldots, \lambda_k$. But note that if $k < n$, then there must be some λ_i where $1 \leq i \leq k$ that is repeated, for these are the only eigenvalues of T as per the hypothesis. Hence we can easily find that our characteristic polynomial becomes:

$$f(t) = (-1)^n (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k}$$

So that each λ_i appears m_i times for $1 \leq i \leq k$. Therefore the diagonal entries of $[T]_\beta$ are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ repeated according to the values of $m_1, m_2, \ldots, m_k$ respectively. ■

Problem 5.2.11
Let $A \in M_n(\mathbb{K})$ such that A has distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ corresponding to multiplicities $m_1, m_2, \ldots, m_k$. Also, let A be similar to an upper triangular matrix.

a.) Let U be the upper triangular matrix. First, from Problem 5.1.12, we know that A and U have identical characteristic polynomials. Hence A and U possess the same eigenvalues with corresponding multiplicities, so that the upper triangular matrix has distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ corresponding to multiplicities $m_1, m_2, \ldots, m_k$. From Problem 5.2.10, we know that the entries on the main diagonal are precisely those $\lambda_1, \lambda_2, \ldots, \lambda_k$, with each λ_i appearing m_i times for $1 \leq i \leq k$. Thus we see by definition of the trace:

$$\operatorname{tr}(U) = \sum_{i=1}^{k} m_i \lambda_i$$

But note that by Problem 5.1.16, A and U also share the same trace, so that we necessarily have:

$$\operatorname{tr}(A) = \operatorname{tr}(U) = \sum_{i=1}^{k} m_i \lambda_i$$

Which is the desired equality. ■

b.) Recall from Problem 4.2.23 that the determinant of an upper triangular matrix is the product of the entries on its main diagonal. In the first part of the problem we determined that the matrix U has main diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_k$, with each λ_i appearing m_i times for $1 \leq i \leq k$. Therefore:

$$\det(U) = \prod_{i=1}^{n} \lambda_i^{m_i}$$

But note that by definition, if $A \sim U$, then there exists an invertible matrix Q such that $A = Q^{-1}UQ$. However also:

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

$$\det(QAQ^{-1}) = \det(U)$$

By properties of determinants derived in Chapter 4. This shows that similar matrices have the same determinant. Therefore, since A is similar to U, we can see that:

$$\det(A) = \det(U) = \prod_{i=1}^{n} \lambda_i^{m_i}$$

And hence the desired relation is proved to hold. ■

Problem 5.2.12
Let T be an invertible linear operator on a finite-dimensional vector space V.

a.) From Problem 5.1.8(b), we know that if λ is an eigenvalue of T, then λ^{-1} is an eigenvalue of T^{-1}. Note that since T is invertible, then $\lambda \neq 0$ by Problem 5.1.8(a). Consider the subspace $E_\lambda = \{x \in V : T(x) = \lambda x\}$. Now suppose $x \in E_\lambda$. Then we know that:

$$T(x) = \lambda x \implies T^{-1}(T(x)) = T^{-1}(\lambda x) \implies x = \lambda T^{-1}(x)$$
$$\implies T^{-1}(x) = \lambda^{-1} x \implies x \in E_{\lambda^{-1}}$$

And since $x \in E_\lambda$ was arbitrary, we necessarily have $E_\lambda \subseteq E_{\lambda^{-1}}$. The reverse containment is shown in a similar manner. Suppose $y \in E_{\lambda^{-1}}$, so then we have:

$$T^{-1}(y) = \lambda^{-1} y \implies T(T^{-1}(y)) = T(\lambda^{-1} y) \implies y = \lambda^{-1} T(y)$$
$$\implies T(y) = \lambda y \implies y \in E_\lambda$$

And thus $E_{\lambda^{-1}} \subseteq E_\lambda$, so that we now have $E_\lambda = E_{\lambda^{-1}}$ as desired. ■

b.) Suppose that T is diagonalizable. Then by Theorem 5.11, we know that $V = \bigoplus_{i=1}^{k} E_{\lambda_i}$, where $\lambda_1, \ldots, \lambda_k$ are the distinct eigenvalues of T. However, by Problem 5.1.8(b), we note that since T is invertible, then $\lambda_1^{-1}, \ldots, \lambda_k^{-1}$ are eigenvalues of T^{-1} and we then have $E_{\lambda_i} = E_{\lambda_i^{-1}}$ for $1 \leq i \leq k$ by the results of the first part of this problem. Therefore we have:

$$V = \bigoplus_{i=1}^{k} E_{\lambda_i} = \bigoplus_{i=1}^{k} E_{\lambda_i^{-1}}$$

And thus by Theorem 5.11, since V is the direct sum of the eigenspaces of T^{-1}, we have that T^{-1} is diagonalizable linear operator. ■

Problem 5.2.13
a.) Consider $A \in M_2(\mathbb{K})$ such that:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \implies A^T = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Notice that the eigenvalues of A are

b.) Let $A \in M_n(\mathbb{K})$ such that λ is an eigenvalue of A. We know that the dimension of the eigenspace corresponding to λ is equivalent to the nullity of the matrix $A - \lambda I_n$. Hence:

$$\dim(E_\lambda) = \dim(N(A - \lambda I_n))$$

c.) Let $A \in M_n(\mathbb{K})$. Suppose that A is diagonalizable. Then we know that there exists an ordered basis for $\mathbb{K}^n$ consisting of eigenvectors by Theorem 5.1 applied to matrices. By Problem 5.1.14, we know that the eigenvalues of A and A^T are the same. Hence there exists an ordered basis consisting of eigenvectors of A^T. Hence A^T is diagonalizable. ■

Problem 5.2.14

a.)

$$y_g(t) = k_1 e^{2t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + k_2 e^{-2t} \begin{pmatrix} 1 \\ -3 \end{pmatrix}$$

b.)

$$y_g(t) = k_1 e^{3t} \begin{pmatrix} -2 \\ 1 \end{pmatrix} + k_2 e^{-2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

c.)

$$y_g(t) = k_1 e^{t} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + k_2 e^{t} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + k_3 e^{2t} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Problem 5.2.15

Problem 5.2.16

Problem 5.2.17

a.) Suppose that T and U are simultaneously diagonalizable linear operators on a finite-dimensional vector space V. Then, by definition, there exists an ordered basis β for V such that $[T]_\beta$ and $[U]_\beta$ are diagonal matrices. Let γ be another ordered basis for V. Then we know that there exists some invertible matrix $Q = [I_V]_\beta^\gamma$ such that: $[T]_\beta = Q^{-1}[T]_\gamma Q$ and $[U]_\beta = Q^{-1}[U]_\gamma Q$. Clearly both of the matrices on the right are diagonal, and therefore, since the choice of γ was arbitrary, we can conclude that $[T]_\beta$ and $[U]_\beta$ are simultaneously diagonalizable for any choice of basis β for V. ■

b.) Let $A, B \in M_n(\mathbb{K})$. Suppose that A and B are simultaneously diagonalizable matrices. Then by definition there exists an invertible matrix $Q \in M_n(\mathbb{K})$ such that $Q^{-1}AQ$ and $Q^{-1}BQ$ are diagonal matrices. Let γ be the standard ordred basis for $\mathbb{K}^n$ and β another ordered basis, so $Q = [I_{\mathbb{K}^n}]_\beta^\gamma$ Then we know that $[L_A]_\gamma = A$ and $[L_B]_\gamma = B$, so we see:

$$[L_A]_\beta = [I_{\mathbb{K}^n}]_\gamma^\beta [L_A]_\gamma [I_{\mathbb{K}^n}]_\beta^\gamma = Q^{-1}AQ$$

$$[L_B]_\beta = [I_{\mathbb{K}^n}]_\gamma^\beta [L_B]_\gamma [I_{\mathbb{K}^n}]_\beta^\gamma = Q^{-1}BQ$$

And hence $[L_A]_\beta$ and $[L_B]_\beta$ are both diagonal matrices and hence L_A and L_B are simultaneously diagonalizable linear operators. ■

Problem 5.2.18

a.) Let V be a vector space and let $T, U \in \mathcal{L}(V)$. Suppose that T and U are simultaneously diagonalizable linear operators. By Problem 5.2.17(a), $[T]_\beta$ and $[U]_\beta$ are also simultaneously diagonalizable matrices for any ordered basis β, and hence there exists: $Q = [I_V]_\beta^\gamma$ such that $Q^{-1}[T]_\beta Q$ and $Q^{-1}[U]_\beta Q$ are diagonal matrices. Let γ be any ordered basis of V. Then:

b.)

Problem 5.2.19

Problem 5.2.20 XXX

Let $W_1, W_2, \ldots, W_k$ be subspaces of a finite-dimensional vector space V such that $W_1 + \cdots + W_k = V$. Now suppose that:

$$V = \bigoplus_{i=1}^{k} W_i$$

Then we can see that $V = W_1 \oplus (W_2 \oplus \cdots \oplus W_k)$, where $\oplus_{i=2}^k W_i$ is clearly a finite-dimensional subspace. Applying the results of Problem 1.6.29, we can see that:

$$\dim(V) = \dim(W_1) + \dim(\sum_{i=2}^{k} W_i)$$

Then, after applying this result inductively, on the rest of the subspaces, we can eventually find that:

$$\dim(V) = \sum_{i=1}^{k} \dim(W_i)$$

Now for the converse, suppose that $\dim(V) = \dim(W_1) + \cdots + \dim(W_k)$. Then by a similar recursive argument to that of the one above, we can apply the results of Problem 1.6.29, which is an if and only if, once again to find that V is the direct sum of $W_1, W_2, \ldots, W_k$.

Problem 5.2.21

Problem 5.2.22

Problem 5.2.23

5.3 Matrix Limits and Markov Chains

5.4 Invariant Subspaces and the Cayley-Hamilton Theorem

Problem 5.4.1
a.) False. Since by definition $T(\vec{0}) = \vec{0}$ for any linear operator, and $\{\vec{0}\}$ is a T-invariant subspace of any vector space.

b.) True. This is Theorem 5.21.

c.) False. $v = T(w)$ is a possibility.

d.) False.

e.) True. $g(t)$ is the characteristic polynomial of T by the Cayley-Hamilton Theorem.

f.) True.

g.) True.

Problem 5.4.2
a.) Suppose $\vec{v} \in W$ such that $\vec{v} = a_0 + a_1x + a_2x^2$, where $a_0, a_1, a_2 \in \mathbb{R}$. Then note that $T(\vec{v}) = a_1 + (2a_2)x \in P_2(\mathbb{R})$. Since $\vec{v}$ was arbitrary, $T(W) \subseteq W$ and W is a T-invariant subspace of $P_3(\mathbb{R})$.

b.) Suppose $\vec{v} \in W$ such that $\vec{v} = a_0 + a_1x + a_2x^2$, where $a_0, a_1, a_2 \in \mathbb{R}$. Then note that $T(\vec{v}) = (a_0 + a_1x + a_2x^2)x = a_0x + a_1x^2 + a_2x^3 \notin P_2(\mathbb{R})$. Since $\vec{v}$ was arbitrary, $T(W) \nsubseteq W$ and W is thus not a T-invariant subspace of $P(\mathbb{R})$.

c.) Supppose $\vec{v} \in W$ such that $\vec{v} = (a, a, a)$, where $a \in \mathbb{R}$. Then note that $T(\vec{v}) = T(a + a + a, a + a + a, a + a + a) = (3a, 3a, 3a) \in W$. Hence W is a T-invariant subspace of $\mathbb{R}^3$.

d.) Suppose $g(t) \in C([0.1])$ such that $g(t) = \lambda t + \mu$, where $\lambda, \mu \in \mathbb{R}$. Then:

$$T(g(t)) = (\int_0^1 g(t)dt)t = (\int_0^1 \lambda t + \mu dt)t =$$

$$(\lambda \int_0^1 tdt + \mu \int_0^1 1dt)t = (\frac{\lambda}{2} + \mu)t$$

We recognize the image of $g(t)$ as a polynomial of degree one, with coefficient $\lambda/2 + \mu \in \mathbb{R}$. Since polynomials are continuous on the interval $[0, 1]$, we see that $T(g(t)) \in C([0, 1])$. Hence W is a T-invariant subspace of $C([0, 1])$.

e.) Suppose $A \in W$ such that, for $a, b, c \in \mathbb{R}$, we have:

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Note that $A = A^T$. Now observe that:

$$T(A) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} = \begin{pmatrix} b & a \\ c & b \end{pmatrix} \notin W$$

Which follows since $a \neq c$ necessarily. So $T(A)^T \neq T(A)$, and thus $T(W) \not\subseteq W$. Therefore W is not a T-invariant subspace of $M_2(\mathbb{R})$.

Problem 5.4.3
Let T be a linear operator on a finite-dimensional vector space V.

a.) $T(\vec{0}) = \vec{0}$ by definition of a linear transformation. Hence $T(\vec{0}) \in \{\vec{0}\}$, so the zero subspace is a T-invariant subspace of V. Now let $v \in V$. Then $T(v) \in V$ by definition of a linear operator, so $T(v) \in V$ for all $v \in V \implies V$ is a T-invariant subspace of V. ■

b.) Let $x \in N(T)$. Then $T(x) = \vec{0}$, and $\vec{0} \in N(T)$ always $\implies T(N(T)) \subseteq N(T)$ so $N(T)$ is a T-invariant subspace of V. Now let $y \in R(T)$. Then there exists some $x \in V$ such that $T(x) = y$. Then $T(y) = T^2(x) \in R(T)$ and so $R(T)$ is a T-invariant subspace of V. ■

c.) Let λ be an eigenvalue of T. Then we know that E_λ is a subsapce of V. Let $x \in E_\lambda$, so we have that $T(x) = \lambda x \implies T(T(x)) = \lambda T(x) \implies T(x) \in E_\lambda$. Hence E_λ is a T-invariant subspace of V for any eigenvalue λ of T. ■

Problem 5.4.4
Let $T \in \mathcal{L}(V)$, and let W be a T-invariant subspace of V. Let $g(t)$ be any polynomial. Suppose $x \in W$.

$$g(t) = \sum_i a_i t^i \implies g(T) = \sum_i a_i T^i$$

Then we see that $g(T)(x) = \sum_i a_i T^i(x) \in W$ since $T(x) \in W$ for any $x \in W$ by definition of T-invariant subspace. Hence W is a $g(T)$-invariant subspace of V for any polynomial $g(t)$.

Problem 5.4.5 XXX
Let $T \in \mathcal{L}(V)$. Let $\mathcal{C}$ denote a collection of T-invariant subspaces of V. Choose any $W_i \in \mathcal{C}$, and consider the set $\bigcap_i W_i$. Let $x \in \bigcap_i W_i$. Then $T(x) \in \bigcap_i W_i$ since each W_i is T-invariant. ???

Problem 5.4.6
a.)

$$\beta = \{(1,0,0,0),(1,0,1,1),(1,-1,2,2)\}$$

b.)

$$\beta = \{x^3, 6x\}$$

c.)

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

d.)

$$\beta = \{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}\}$$

Problem 5.4.7
Let V ve a vector space and W a T-invariant subspace of V. Let $T : V \to V$ be linear. Then consider $T|_W = T_W : W \to W$. Let $x, y \in W$ and $\mu \in \mathbb{K}$. We know that $\mu x + y \in W$. Then:

$$T_W(\mu x + y) = T(\mu x + y) = \mu T(x) + T(y) = \mu T_W(x) + T_W(y)$$

Which follows from the supposed linearity of T. Therefore the restriction of T to W is a linear operator on W as desired. ■

Problem 5.4.8
Let $T \in \mathcal{L}(V)$ with T-invariant subspace W of V. Suppose v is an eigenvector of T_W with corresponding eigenvalue λ. Then, since $v \in W$, we know that:

$$T_W(v) = \lambda v \implies T_W(v) = T(v) = \lambda v$$

So that necessarily v is also an eigenvector of T with corresponding eigenvalue λ as well. ■

Problem 5.4.9
a.) *By means of Theorem 5.22*:
A basis for W is given by the set $\beta = \{(1,0,0,0),(1,0,1,1),(1,-1,2,2)\}$. Note that $T^3(\vec{z}) = 3\beta_2 - 3\beta_3$. Hence:

$$0\vec{z} + 3T(\vec{z}) - 3T^2(\vec{z}) + T^3(\vec{z}) = \vec{0}$$

And therefore by Theorem 5.22(b),

$$f(t) = (-1)^3(3t - 3t^2 + t^3) = -t(3 - 3t + t^2)$$

By means of determinants: Let β be an ordered basis for W as above.

$$[T_W]_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -3 \\ 0 & 1 & 3 \end{pmatrix} \implies \det([T_W]_\beta - tI_3) = -t(3 - 3t + t^2)$$

Clearly we obtain the same characteristic polynomial for T_W.

b.) *By means of Theorem 5.22*:
A basis for W is given by the set $\beta = \{x^3, 6x\}$. Note that $T^2(\vec{z}) = \vec{0}$. Hence:

$$0\vec{z} + 0T(\vec{z}) + T^2(\vec{z}) = \vec{0}$$

And therefore by Theorem 5.22(b),

$$f(t) = (-1)^2(0 + 0t + t^2) = t^2$$

By means of determinants: Let β be an ordered basis for W as above. Then:

$$[T_W]_\beta = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \implies \det([T_W]_\beta - tI_2) = t^2$$

And so we again obtain the same characteristic polynomial for T_W.

c.) *By means of Theorem 5.22*:
A basis for W is given by the set $\beta = \{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\}$. Note that $T(\vec{z}) = \vec{z}$. Hence:

$$-\vec{z} + T(\vec{z}) = \vec{0}$$

And therefore by Theorem 5.22(b),

$$f(t) = (-1)^1(-1 + t) = -(t - 1) = 1 - t$$

By means of determinants: Let β be an ordered basis for W as above. Then:

$$[T_W]_\beta = \begin{pmatrix} 1 \end{pmatrix} \implies \det([T_W]_\beta - tI_1) = 1 - t$$

And so we again obtain the same characteristic polynomial for T_W.

d.) *By means of Theorem 5.22*:
A basis for W is given by the set $\beta = \{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}\}$. Note that $T^2(\vec{z}) = 3T(\vec{z})$. Hence we have that:

$$0\vec{z} - 3T(\vec{z}) + T^2(\vec{z}) = \vec{0}$$

And therefore by Theorem 5.22(b),

$$f(t) = (-1)^2(0 - 3t + t^2) = t(t-3)$$

By means of determinants: Let β be an ordered basis for W as above. Then:

$$[T_W]_\beta = \begin{pmatrix} 0 & 0 \\ 1 & 3 \end{pmatrix} \implies \det([T_W]_\beta - tI_2) = -t(3-t) = t(t-3)$$

And so we again obtain the same characteristic polynomial for T_W.

Problem 5.4.10

Problem 5.4.11
a.) Let T be a linear operator on a vector space V, let $\vec{v}$ be a nonzero vector in V, and let W be the T-cyclic subspace of V generated by $\vec{v}$. Suppose $\vec{x} \in W$. Then for some particular $a_0, a_1, \ldots, a_k \in \mathbb{K}$ we have:

$$\vec{x} = \sum_{i=0}^{k} a_i T^i(\vec{v}) \implies T(\vec{x}) = T(\sum_{i=0}^{k} a_i T^i(\vec{v})) = \sum_{i=0}^{k} a_i T^{i+1}(\vec{v}) \in W$$

And hence $T(W) \subseteq W \implies$ the T-cyclic subspace W generated by $\vec{v}$ is T-invariant. $\blacksquare$

b.) Let Z be a T-invariant subspace of V. Suppose that $\vec{v} \in Z$. Since Z is T-invariant, then $T(\vec{v}) \in Z$, however then $T^2(\vec{v}) \in Z$. Recursively using this argument, we can see that $T^m(\vec{v}) \in Z$ where $m \in \mathbb{Z}^+$. Since $\{T^i(\vec{v})\} \subseteq Z$ for $0 \leq i \leq m$, by Theorem 1.5, $\operatorname{span}(\{\vec{v}, T(\vec{v}), T^2(\vec{v}), \ldots\}) = W \subseteq Z$. So if Z contains $\vec{v}$, then Z also contains W. $\blacksquare$

Problem 5.4.12
Let T be a linear operator on a finite-dimensional vector space V, and let W be a T-invariant subspace of V. Choose $\gamma = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ as an ordered basis for W, and extend this basis to an ordered basis $\beta = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$ for V. Let $A = [T]_\beta$ and $B_1 = [T_W]_\gamma$. Choose $\vec{v}_i$ with $1 \leq i \leq k$. Then $T(\vec{v}_i) \in W$ because W is T-invariant. Hence $T(\vec{v}_i)$ has a basis representation:

$$T(\vec{v}_i) = \sum_{j=1}^{k} a_j \vec{v}_j = \sum_{j=1}^{k} a_j \vec{v}_j + \sum_{j=k+1}^{n} 0\vec{v}_j \implies$$

$$A = [T]_\beta = \begin{pmatrix} B_1 & B_2 \\ O & B_3 \end{pmatrix}$$

Where $B_1, B_2, B_3 \neq O$. This follows because the β vectors $\vec{v}_{k+1}, \ldots, \vec{v}_n$ have coefficients of 0 in the matrix representation of any vector in W. The rest of the matrices are nonzero since if $\vec{w} \in V$ such that $\vec{w} \notin W$, then:

$$\vec{w} = \sum_{i=1}^{n} a_i \vec{v}_i$$

So that not all of the coefficients are 0. Hence the desired equality is proven for use in proving Theorem 5.21. ■

Problem 5.4.13
Let T be a linear operator on a finite-dimensional vector space V, let $\vec{v} \neq \vec{0}$ $\vec{v} \in V$, and let W be the T-cyclic subspace of V generated by $\vec{v}$. Let $\vec{w} \in V$ be arbitrary. Suppose that $\vec{w} \in W$. Then we know that for $a_0, a_1, \ldots, a_k \in \mathbb{K}$:

$$\vec{w} = \sum_{i=0}^{k} a_i T^i(\vec{v}) = g(T)(\vec{v})$$

So that $\vec{w}$ is the evaluation of the polynomial $g(T)$ at $\vec{v}$. Conversely, suppose that:

$$\vec{w} = g(T)(\vec{v})$$

Where $g(t)$ is some polynomial. Suppose $g(t) = a_0 + a_1 t + \cdots + a_k t^k$. Then we can see that from the hypothesis:

$$\vec{w} = g(T)(\vec{v}) = a_0\vec{v} + a_1 T(\vec{v}) + \cdots + a_k T^k(\vec{v}) \implies \vec{w} \in W$$

Since $W = \operatorname{span}(\{\vec{v}, T(\vec{v}), T^2(\vec{v}), \ldots\})$ and clearly $\vec{w}$ is some linear combination of the vectors in the set. Hence $\vec{w}$ is a vector in the T-cyclic subspace of V generated by $\vec{v}$ as desired. ■

Problem 5.4.14

Problem 5.4.15
The Cayley-Hamilton for matrices is shown by letting $[L_A]_\gamma = A$ for the ordered basis γ of $\mathbb{K}^n$. Then by applying the Cayley-Hamilton theorem for transformations, we obtain the desired result for matrices.

Problem 5.4.16
a.) Let V be a finite-dimensional vector space and $T \in \mathcal{L}(V)$. Suppose that the characteristic polynomial of T splits. Suppose W is any T-invariant subspace of V. Now restrict T to W. We know from Theorem 5.21 that the characteristic polynomial of this restriction divides the characteristic polynomial of T itself. Hence if the characteristic polynomial splits into linear factors, then the characteristic polynomial of $T|_W$ must be some combination of those linear factors in order for it to divide as per the theorem. Hence the characteristic polynomial of $T|_W$ must also split. ■

b.) Let V be a finite-dimensional vector space and $T \in \mathcal{L}(V)$. If we suppose that a T-invariant subspace of V is nontrivial, then we know that it must have dimension greater than or equal to 1. Hence the characteristic polynomial of the restriction of T to this T-invariant subspace must have degree greater than or equal to 1, which directly indicates at least one root for the polynomial. Since these roots are precisely the eigenvalues of the restricted transformation, we can conclude that any nontrivial T-invariant subspace of V contains at least one eigenvector of T. ■

Problem 5.4.17
Let A be an $n \times n$ matrix. Then, from the Cayley-Hamilton theorem applied to matrices, we know that if $f(t)$ is the characteristic polynomial of A, then $f(A) = O$. Consider the set $\text{span}(\{I_n, A, A^2, \ldots\})$. Let $f(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$. Then we know that:

$$f(A) = O = (-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 \implies$$

$$A^n = (1/(-1)^{n+1})[a_{n-1}A^{n-1} + \cdots + a_1 A + a_0]$$

And since we are assuming that $A \neq O$, none of the $A_i \neq O$ for $1 \leq i \leq n$. Hence A^n is a linear combination of the A_i for $1 \leq i \leq n-1$, which means that $\dim(\text{span}(\{I_n, A, A^2, \ldots\})) < n-1$. However, by multiplying both sides of the above equation by A on the right once more, we can see:

$$A^{n+1} = (1/(-1)^{n+1})[a_{n-1}A^n + a_{n-2}A^{n-1} + \cdots + a_1 A + a_0]$$

So that necessarily, A^{n+1} is a linear combination of the matrices $I_n, A, A^2, \ldots, A^n$. Hence $\text{span}(\{I_n, A, A^2, \ldots\}) = \text{span}(\{I_n, A, A^2, \ldots, A^{n-1}, A^n\})$, so that we now clearly have $\dim(\text{span}(\{I_n, A, A^2, \ldots\})) \leq n$. ■

Problem 5.4.18
Let A be an $n \times n$ matrix with characteristic polynomial

$$f(t) = (-1)^n t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

a.) Suppose that A is invertible. Then by Problem 5.1.8(a), we know that 0 is not an eigenvalue of A. Hence $f(0) = a_0 \neq 0$. Conversely, suppose that the constant term $a_0 \neq 0$. Then $\lambda = 0$ cannot be the root of the polynomial $f(t)$, and so again by Problem 5.1.8(a), A is invertible. ■

b.) Suppose that A is invertible. Then, since we know $f(t)$, we know that from the Cayley-Hamilton theorem for matrices,

$$f(A) = (-1)^n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 = O$$

Then, if we apply the inverse of A, A^{-1}, to the right of both sides of the above equation, we obtain:

$$(-1)^n A^n A^{-1} + a_{n-1}A^{n-1}A^{-1} + \cdots + a_1 AA^{-1} + a_0 A^{-1} = O$$

$$\implies (-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n + a_0 A^{-1} = O$$

Now we subtract the left-most term from both sides and then divide by the scalar a_0 to obtain:

$$(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n = (-a_0)A^{-1} \implies$$

$$A^{-1} = (\frac{1}{a_0})[(-1)^n A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I_n]$$

Which is the desired expression for the inverse of A. ■

Problem 5.4.19

Problem 5.4.20

Problem 5.4.21

Problem 5.4.22

Problem 5.4.23

Let V be a finite-dimensional vector space, let $T \in \mathcal{L}(V)$, and let W be a T-invariant subspace of V. Now let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ are eigenvectors of T corresponding to distinct eigenvalues. Suppose that $\vec{v}_1 + \vec{v}_2 + \cdots + \vec{v}_k \in W$. For the base case, choose $k = 1$. Then clearly $\vec{v}_1 \in W$. For the inductive step, suppose that the hypothesis holds for the $k-1$th case. Then $\vec{v}_1 + \vec{v}_2 + \cdots \vec{v}_{k-1} \in W$, and $\vec{v}_i \in W$ for $1 \leq i \leq k-1$. Subsequently, we see that if $\vec{v}_1 + \vec{v}_2 + \cdots \vec{v}_{k-1} + \vec{v}_k \in W$, then by closure of the subspace, $\vec{v}_k \in W$ and thus all $\vec{v}_i \in W$ for $1 \leq i \leq k$. ■

Problem 5.4.24

Let T be a diagonalizable linear operator, and let W be a non-trivial T-invariant subspace.

Problem 5.4.25

Problem 5.4.26

Problem 5.4.27

Let T be a linear operator on a vector space V, and let W be a T-invariant subspace of V. For any $\vec{v} + W \in V/W$, define $\overline{T} : V/W \to V/W$ by:

$$\overline{T}(\vec{v} + W) = T(\vec{v}) + W$$

a.) Suppose $\vec{v}+W=\vec{\mu}+W$. Then we know that $(\vec{v}-\vec{\mu})+W=\vec{0}+W$. Then, by the linearity of T, $\overline{T}((\vec{v}-\vec{\mu})+W)=T(\vec{v}-\vec{\mu})+W=T(\vec{v})-T(\vec{\mu})+W=\vec{0}+W \implies T(\vec{v})+W=T(\vec{\mu})+W$. Therefore the mapping $\overline{T}$ is well-defined as desired. ■

b.) Let $\vec{x},\vec{y}\in V$. Now let $\vec{x}+W,\vec{y}+W\in V/W$ and $\alpha\in\mathbb{K}$. Then:

$$\overline{T}(\alpha\vec{x}+\vec{y}+W)=T(\alpha\vec{x}+\vec{y})+W=\alpha T(\vec{x})+T(\vec{y})+W=$$

$$\alpha(T(\vec{x})+W)+(T(\vec{y})+W)=\alpha\overline{T}(\vec{x}+W)+\overline{T}(\vec{y}+W)$$

And hence the mapping $\overline{T}$ satisfies the conditions for a linear operator on V. ■

c.) Let $\eta:V\to V/W$ defined by $\eta(\vec{v})=\vec{v}+W$ be the mapping defined in Problem 2.1.40. We will show that the diagram Figure 5.6 in the text commutes. Suppose $\vec{x}\in V$. Then observe:

$$(\eta T)(x)=\eta(T(\vec{x}))=T(\vec{x})+W$$

$$(\overline{T}\eta)(\vec{x})=\overline{T}(\eta(\vec{x}))=\overline{T}(\vec{x}+W)=T(\vec{x})+W$$

And hence we can see that since $\vec{x}\in V$ was arbitrary, the diagram commutes, that is, $\eta T=\overline{T}\eta$. ■

Problem 5.4.28
Let $\gamma=\{\vec{v}_1,\ldots,\vec{v}_k\}$ be an ordered basis for W, and extend this basis to an ordered basis $\beta=\{\vec{v}_1,\ldots,\vec{v}_k,\vec{v}_{k+1},\ldots,\vec{v}_n\}$ for V. Then, by the result of Problem 1.6.35(a), we know that $\alpha=\{\vec{v}_{k+1}+W,\ldots,\vec{v}_n+W\}$ is a basis for V/W. Let $B_1=[T]_\gamma$ and let $B_3=[\overline{T}]_\alpha$

Problem 5.4.29

Problem 5.4.30

Problem 5.4.31

Problem 5.4.32

Problem 5.4.33
Let V be a vector space and $T\in\mathcal{L}(V)$. Let $W_1,W_2,\ldots,W_k$ be T-invariant subspaces of V. From Problem 1.3.23(a), $\sum_{i=1}^k W_i$ is a subspace of V. Now let $\vec{x}\in\sum_{i=1}^k W_i$, so that:

$$\vec{x}=\sum_{i=1}^k a_i\vec{w}_i$$

For $\vec{w}_i \in W_i$ for each $1 \leq i \leq k$, and $a_i \in \mathbb{K}$ for all $1 \leq i \leq k$. Now, if we apply T to $\vec{x}$, we find:

$$T(\vec{x}) = T(\sum_{i=1}^{k} a_i\vec{w}_i) = \sum_{i=1}^{k} a_i T(\vec{w}_i)$$

But since each W_i is T-invariant, $T(\vec{w}_i) \in W_i$ for each $1 \leq i \leq k$, and hence $T(\vec{x}) \in \sum_{i=1}^{k} W_i \implies T(\sum_{i=1}^{k} W_i) \subseteq \sum_{i=1}^{k} W_i$. Hence the sum of the W_i is a T-invariant subspace of V as well. ■

Problem 5.4.34

Problem 5.4.35

Problem 5.4.36
This is a simple deduction in the context of the results shown in Theorem 5.11. Since each eigenspace has dimension 1, the span of a single eigenvector, and T is diagonalizable $\iff$ V is the direct sum of the eigenspaces, which are T-invariant.

Problem 5.4.37
Let V be a finite-dimensional vector space and $T \in \mathcal{L}(V)$. Let $W_1, W_2, \ldots, W_k$ be T-invariant subspaces of V such that $V = \bigoplus_{i=1}^{k} W_i$. Let β_i be an ordered basis for W_i for each $1 \leq i \leq k$, and let $\beta = \bigcup_{i=1}^{k} \beta_i$ be an ordered basis for V. Then by Theorem 5.25, we know that:

$$[T]_\beta = \bigoplus_{i=1}^{k} [T_{W_i}]_{\beta_i} \implies \det([T]_\beta) = \prod_{i=1}^{k} \det([T_{W_i}]_{\beta_i})$$

This is the desired relation. ■

Problem 5.4.38
Let V be a finite-dimensional vector space and $T \in \mathcal{L}(V)$, and let $W_1, W_2, \ldots, W_k$ be T-invariant subspaces of V such that $V = \bigoplus_{i=1}^{k} W_i$. Suppose that T is diagonalizable. Then

Problem 5.4.39

Problem 5.4.40

Problem 5.4.41

Problem 5.4.42

6 Inner Product Spaces

6.1 Inner Products and Norms

Problem 6.1.1
a.) True. An inner product takes in two vectors and outputs a scalar.

b.) True. Real or complex inner product spaces.

c.) False. It is linear in first component, but conjugate linear in the second component.

d.) False.

e.) False. Theorem 6.2 does not state that V need be finite-dimensional.

f.) False. Every matrix has a transpose and conjugate.

g.) False. Must hold for all $\vec{x} \in V$.

h.) True. Consider Theorem 6.1(e).

Problem 6.1.2
Let $\vec{x}, \vec{y} \in \mathbb{C}^3$ such that $\vec{x} = (2, 1+i, i)$ and $\vec{y} = (2-i, 2, 1+2i)$. Using the standard inner product, we find:

$$\langle \vec{x}, \vec{y} \rangle = 2(2+i) + (1+i)(2) + i(1-2i) = 8+5i$$

$$||\vec{x}|| = \sqrt{\langle \vec{x}, \vec{x} \rangle} = \sqrt{7}$$

$$||\vec{y}|| = \sqrt{\langle \vec{y}, \vec{y} \rangle} = \sqrt{14}$$

$$||\vec{x}+\vec{y}|| = \sqrt{\langle \vec{x}+\vec{y}, \vec{x}+\vec{y} \rangle} = \sqrt{37}$$

Problem 6.1.3
Let $f, g \in C([0,1])$ such that $f(t) = t$ and $g(t) = e^t$. Then, using the inner product defined in Example 3, we find:

$$\langle f, g \rangle = \int_0^1 te^t dt = 1$$

$$||f|| = \sqrt{\langle f, f \rangle} = \sqrt{\int_0^1 t^2 dt} = \sqrt{\frac{1}{3}} = \frac{1}{\sqrt{3}}$$

$$||g|| = \sqrt{\langle g, g \rangle} = \sqrt{\int_0^1 e^{2t} dt} = \sqrt{\frac{e^2-1}{2}}$$

$$||f+g|| = \sqrt{\langle f+g, f+g \rangle} = \sqrt{\int_0^1 (t+e^t)^2 dt} = \sqrt{\frac{3e^2+11}{6}}$$

Problem 6.1.4

a.) We will prove conditions (b) and (c) of the definition of an inner product for the Frobenius inner product shown in Example 5. First we show condition (c), let $A, B \in M_n(\mathbb{K})$. Then:

$$\overline{\langle A, B\rangle} = \overline{\operatorname{tr}(B^*A)} = \overline{\sum_{i=1}^{n}(B^*A)_{ii}} = \overline{\sum_{i=1}^{n}\sum_{k=1}^{n}(B^*)_{ik}A_{ki}} =$$

$$\overline{\sum_{i=1}^{n}\sum_{k=1}^{n}\overline{B}_{ki}A_{ki}} = \sum_{i=1}^{n}\sum_{k=1}^{n}B_{ki}\overline{A}_{ki} = \sum_{i=1}^{n}\sum_{k=1}^{n}B_{ki}(A^*)_{ik} =$$

$$\sum_{i=1}^{n}\sum_{k=1}^{n}(A^*)_{ik}B_{ki} = \sum_{i=1}^{n}(A^*B)_{ii} = \operatorname{tr}(A^*B) = \langle B, A\rangle$$

And hence condition (c) holds. Now for condition (b), suppose $C, D \in M_n(\mathbb{K})$, and $\mu \in \mathbb{K}$. Then note that:

$$\langle \mu C, D\rangle = \operatorname{tr}(D^*(\mu C)) = \operatorname{tr}(\mu(D^*C)) = \mu \operatorname{tr}(D^*C = \mu\langle C, D\rangle$$

From the linearity properties of the trace derived in Problem 1.3.6. This is sufficient to prove (b), and thus the Frobenius inner product $\langle\cdot,\cdot\rangle$ satisfies all of the conditions for an inner product on the vector space $M_n(\mathbb{K})$. ■

b.)

$$\langle A, B\rangle = \operatorname{tr}(B^*A) = \operatorname{tr}(\begin{pmatrix} 1-4i & 4-i \\ 3i & -1 \end{pmatrix}) = -4i$$

$$||A|| = \sqrt{\langle A, A\rangle} = \sqrt{\operatorname{tr}(\begin{pmatrix} 10 & 2+4i \\ 2-4i & 6 \end{pmatrix})} = \sqrt{16} = 4$$

$$||B|| = \sqrt{\langle B, B\rangle} = \sqrt{\operatorname{tr}(\begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix})} = \sqrt{4} = 2$$

Problem 6.1.5

Let $\vec{x}, \vec{y} \in \mathbb{C}^2$ and let $A \in M_2(\mathbb{C})$ such that:

$$A = \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix}$$

Define $\langle\cdot,\cdot\rangle$ by $\langle \vec{x}, \vec{y}\rangle = \vec{x}A(\vec{y})^*$. We will show that $\langle\cdot,\cdot\rangle$ is an inner product on $\mathbb{C}^2$. First, for condition (a), let $\vec{z} \in \mathbb{C}^2$, and we see:

$$\langle \vec{x}+\vec{z}, \vec{y}\rangle = (\vec{x}+\vec{z})A(\vec{y})^* = (\vec{x}A + \vec{z}A)(\vec{y})^* = \vec{x}A(\vec{y})^* + \vec{z}A(\vec{y})^* = \langle \vec{x}, \vec{y}\rangle + \langle \vec{z}, \vec{y}\rangle$$

Which proves condition (a). Now let $\mu \in \mathbb{C}$. Then observe that:

$$\langle c\vec{x}, \vec{y}\rangle = (c\vec{x})A(\vec{y})^* = (c\vec{x}A)(\vec{y})^* = c(\vec{x}A(\vec{y})^*) = c\langle \vec{x}, \vec{y}\rangle$$

And thus condition (b) is shown to hold. For condition (c), we see:

$$\overline{\langle \vec{x}, \vec{y} \rangle} = \overline{\vec{x}A(\vec{y})^*} = (\vec{x}A(\vec{y})^*)^* = \vec{y}A^*(\vec{x})^* = \vec{y}A(\vec{x})^* = \langle \vec{y}, \vec{x} \rangle$$

Which follows since $A = A^*$. Now for the last condition, condition (d), suppose $\vec{x} \neq \vec{0}$, so we see:

$$\langle \vec{x}, \vec{x} \rangle = \vec{x}A(\vec{x})^* = (a_1, a_2)A(a_1, a_2)^* =$$

$$= \begin{pmatrix} a_1 & a_2 \end{pmatrix} \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix} \begin{pmatrix} \overline{a_1} \\ \overline{a_2} \end{pmatrix} = \begin{pmatrix} a_1 - ia_2 & ia_1 - 2a_2 \end{pmatrix} \begin{pmatrix} \overline{a_1} \\ \overline{a_2} \end{pmatrix} =$$

$$= \overline{a_1}(a_1 - ia_2) + \overline{a_2}(ia_1 - 2a_2) = \overline{a_1}a_1 - i\overline{a_1}a_2 + i\overline{a_2}a_1 - 2\overline{a_2}a_2 =$$

$$= ||a_1||^2 + Re(-i\overline{a_1}a_2 + i\overline{a_2}a_1) + 2||a_2||^2 > 0$$

Since both $a_1 = a_2 \neq 0$. Therefore (d) is shown to hold and hence $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{C}^2$ as desired. ■

Now we will compuite $\langle \vec{x}, \vec{y} \rangle$ for $\vec{x} = (1 - i, 2 + 3i)$ and $\vec{y} = (2 + i, 3 - 2i)$.

$$\langle \vec{x}, \vec{y} \rangle = \begin{pmatrix} 1 - i & 2 + 3i \end{pmatrix} \begin{pmatrix} 1 & i \\ -i & 2 \end{pmatrix} \begin{pmatrix} 2 - i \\ 3 + 2i \end{pmatrix} =$$

$$= \begin{pmatrix} 4 - 3i & 5 + 7i \end{pmatrix} \begin{pmatrix} 2 - i \\ 3 + 2i \end{pmatrix} = 6 + 21i$$

Problem 6.1.6
Theorem 6.1(b): Let $\vec{x}, \vec{y} \in V$ and $c \in \mathbb{K}$.

$$\langle \vec{x}, c\vec{y} \rangle = \overline{\langle c\vec{y}, \vec{x} \rangle} = \overline{c}\overline{\langle \vec{y}, \vec{x} \rangle} = \overline{c}\langle \vec{x}, \vec{y} \rangle$$

By property (b) and (c) of inner products. ■

Theorem 6.1(c): Let $\vec{x} \in V$.

$$\langle \vec{x}, \vec{0} \rangle = \langle \vec{x}, \vec{x} - \vec{x} \rangle = \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, -\vec{x} \rangle =$$

$$= \langle \vec{x}, \vec{x} \rangle + \overline{\langle -\vec{x}, \vec{x} \rangle} = \langle \vec{x}, \vec{x} \rangle + (\overline{-1})\langle \vec{x}, \vec{x} \rangle = \langle \vec{x}, \vec{x} \rangle - \langle \vec{x}, \vec{x} \rangle = 0$$

Which follows from property (a) of the inner product and the results proved in part (a) of Theorem 6.1. Similarly to the above, we can find that $\langle \vec{0}, \vec{x} \rangle = 0$, and hence:

$$\langle \vec{x}, \vec{0} \rangle = \langle \vec{0}, \vec{x} \rangle = 0$$

Which is the desired equality. ■

Theorem 6.1(d): Let $\vec{x} \in V$. Suppose that $\langle \vec{x}, \vec{x} \rangle = 0$. Then we see:

$$\langle \vec{x}, \vec{x} \rangle - \langle \vec{x}, \vec{x} \rangle = \langle \vec{x}, \vec{x} \rangle \implies$$

$$\langle \vec{x} - \vec{x}, \vec{x} \rangle = \langle \vec{x}, \vec{x} \rangle \implies \langle \vec{0}, \vec{x} \rangle = \langle \vec{x}, \vec{0} \rangle = \langle \vec{x}, \vec{x} \rangle$$

Which follows from property (a) of an inner product and the results of Theorem 6.1(c). We necessarily have that $\vec{x} = \vec{0}$. ■

Theorem 6.1(e): Let $\vec{x}, \vec{y}, \vec{z} \in V$. Suppose that $\langle \vec{x}, \vec{y} \rangle = \langle \vec{x}, \vec{z} \rangle$ for all $\vec{x} \in V$. Then:

$$\langle \vec{x}, \vec{y} \rangle - \langle \vec{x}, \vec{z} \rangle = 0 \implies \langle \vec{x} - \vec{x}, \vec{y} - \vec{z} \rangle = \langle \vec{0}, \vec{y} - \vec{z} \rangle = 0$$

Which follows from Theorem 6.1(a) and property (a) of an inner product. However, we know from Theorem 6.1(c) that $\langle \vec{0}, \vec{y} - \vec{z} \rangle = \langle \vec{y} - \vec{z}, \vec{0} \rangle = 0, \implies \vec{y} - \vec{z} = \vec{0} \implies \vec{y} = \vec{z}$, which is the desired result. ■

Problem 6.1.7
Theorem 6.2(a): Let V be an inner product space over $\mathbb{K}$. Let $\vec{x}, \vec{y} \in V$ and $c \in \mathbb{K}$. Then by property (b) of an inner product and Theorem 6.1(b), we have:

$$||c\vec{x}|| = \sqrt{\langle c\vec{x}, c\vec{x} \rangle} = \sqrt{c\bar{c}\langle \vec{x}, \vec{x} \rangle} = \sqrt{c^2 \langle \vec{x}, \vec{x} \rangle} = |c|\sqrt{\langle \vec{x}, \vec{x} \rangle} = |c| \cdot ||\vec{x}||$$

Which is the desired equality. ■

Theorem 6.2(b): Let $\vec{x} \in V$. Suppose that $||\vec{x}|| = 0$. Then:

$$||\vec{x}|| = \sqrt{\langle \vec{x}, \vec{x} \rangle} = 0 \implies \langle \vec{x}, \vec{x} \rangle = 0 \implies \vec{x} = \vec{0}$$

Which follows directly from Theorem 6.1(d). Conversely, suppose that $\vec{x} = \vec{0}$. Then:

$$\vec{x} = \vec{0} \implies ||\vec{x}|| = \sqrt{\langle \vec{0}, \vec{0} \rangle} = \sqrt{0} = 0$$

By the definition of the norm. And so (b) of Theorem 6.2 is shown to hold. ■

Problem 6.1.8
a.) This is not an inner product on $\mathbb{R}^2$ since it does not satisfy condition (d) of the definition. For example, take $(1,2) \in \mathbb{R}^2$. Then $\langle (1,2), (1,2) \rangle = 1 - 4 = -3$.

b.) This is also not an inner product on $M_2(\mathbb{R})$ since it does not satisfy condition (d) of the definition. Take the matrix:

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

Then $\langle A, A \rangle = \operatorname{tr}(A + A) = -4 < 0$ and $A \neq O$.

c.) This is not an inner product on $P(\mathbb{R})$ since it does not satisfy condition (d). For instance, take $f(x) = a$ where $a \in \mathbb{R}$. Then:

$$\langle f(x), f(x) \rangle = \int_0^1 (a)' a dt = \int_0^1 0 a dt = \int_0^1 0 dt = 0$$

But $f(x) \neq \vec{0}$ and clearly $\langle f(x), f(x) \rangle = 0$.

Problem 6.1.9

a.) Let β be a basis for a finite-dimensional inner product space V. Let $\beta = \{\vec{z}_1, \vec{z}_2, \ldots, \vec{z}_n\}$. Suppose $\vec{x} \in V$ such that $\langle \vec{x}, \vec{z}_j \rangle = 0$ for all $\vec{z}_j \in \beta$, with $1 \leq j \leq n$. Then we know that $\vec{x}$ is orthogonal to every basis vector. Since $\vec{x} \in V$, and β is a basis for V, we know that for particular $a_1, a_2, \ldots, a_n \in \mathbb{K}$:

$$\vec{x} = \sum_{i=1}^{n} a_i \vec{z}_i$$

Given this, we can now see that:

$$\langle \vec{x}, \vec{x} \rangle = \langle \vec{x}, \sum_{i=1}^{n} a_i \vec{z}_i \rangle = \sum_{i=1}^{n} \overline{a}_i \langle \vec{x}, \vec{z}_i \rangle = \sum_{i=1}^{n} \overline{a}_i \cdot 0 = 0$$

Since $\langle \vec{x}, \vec{z}_j \rangle = 0$ for all $\vec{z}_j \in \beta$, with $1 \leq j \leq n$ as per the hypothesis. Then we can see that by Theorem 6.1(d), since $\langle \vec{x}, \vec{x} \rangle = 0 \iff \vec{x} = \vec{0}$. ■

b.) Let β be a basis for a finite-dimensional inner product space V as above. Suppose that $\vec{x}, \vec{y} \in V$ such that $\langle \vec{x}, \vec{z} \rangle = \langle \vec{y}, \vec{z} \rangle$ for all $\vec{z} \in \beta$. Now let $\vec{w} \in V$, so we have that for particular $c_1, \ldots, c_n \in \mathbb{K}$:

$$\vec{w} = \sum_{i=1}^{n} c_i \vec{z}_i$$

Now we can see that we have:

$$\langle \vec{w}, \vec{x} \rangle = \langle \sum_{i=1}^{n} c_i \vec{z}_i, \vec{x} \rangle = \sum_{i=1}^{n} c_i \langle \vec{z}_i, \vec{x} \rangle = \sum_{i=1}^{n} c_i \overline{\langle \vec{x}, \vec{z}_i \rangle} = \sum_{i=1}^{n} c_i \overline{\langle \vec{y}, \vec{z}_i \rangle} =$$

$$= \sum_{i=1}^{n} c_i \langle \vec{z}_i, \vec{y} \rangle = \langle \sum_{i=1}^{n} c_i \vec{z}_i, \vec{y} \rangle = \langle \vec{w}, \vec{y} \rangle$$

Then, since the choice of $\vec{w} \in V$ was arbitrary, we can see that the above holds for all $\vec{w} \in V$. Hence by Theorem 6.1(e), we have that $\vec{x} = \vec{y}$ as desired. ■

Problem 6.1.10

Let V be an inner product space, and let $\vec{x}, \vec{y} \in V$. Suppose that $\vec{x}$ and $\vec{y}$ are orthogonal vectors in V. Then, by the definition, we know that $\langle \vec{x}, \vec{y} \rangle = 0$. Then:

$$||\vec{x} + \vec{y}||^2 = \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle =$$

$$||\vec{x}||^2 + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + ||\vec{y}||^2 = ||\vec{x}||^2 + ||\vec{y}||^2$$

Since $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle = 0$ as per the orthogonality of $\vec{x}$ and $\vec{y}$. In the case where $V = \mathbb{R}^2$, then we have the familiar Pythagorean Theorem. ■

Problem 6.1.11
Let V be an inner product space, and let $\vec{x}, \vec{y} \in V$. Then:

$$||\vec{x} - \vec{y}||^2 = \langle \vec{x} - \vec{y}, \vec{x} - \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle - \langle \vec{x}, \vec{y} \rangle - \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle =$$
$$\langle \vec{x}, \vec{x} \rangle - \langle \vec{x}, \vec{y} \rangle - \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle$$

Now consider:

$$||\vec{x} + \vec{y}||^2 = \langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle =$$
$$\langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle$$

So that we necessarily see that:

$$||\vec{x} + \vec{y}||^2 + ||\vec{x} - \vec{y}||^2 = \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle +$$
$$+ \langle \vec{x}, \vec{x} \rangle - \langle \vec{x}, \vec{y} \rangle - \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle = 2\langle \vec{x}, \vec{x} \rangle + 2\langle \vec{y}, \vec{y} \rangle =$$
$$= 2||\vec{x}||^2 + 2||\vec{y}||^2$$

Which is the desired relation. This is the familiar Parallelogram Law when $V = \mathbb{R}^2$. ■

Problem 6.1.12
Let V be a vector space, let $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\} \subseteq V$ be an orthogonal set, and let $a_1, a_2, \ldots, a_k \in \mathbb{K}$. Then we see:

$$||\sum_{i=1}^{k} a_i \vec{v}_i||^2 = \langle \sum_{i=1}^{k} a_i \vec{v}_i, \sum_{i=1}^{k} a_i \vec{v}_i \rangle = \sum_{i=1}^{k} a_i \bar{a}_i \langle \vec{v}_i, \vec{v}_i \rangle = \sum_{i=1}^{k} |a_i|^2 \cdot ||\vec{v}_i||^2$$

Which follows from the definition of the norm, Theorem 6.1(b), and properties of complex numbers. ■

Problem 6.1.13
Suppose that $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on a vector space V over the field $\mathbb{K}$. Consider the function $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_1 + \langle \cdot, \cdot \rangle_2$. Let $\vec{x}, \vec{y}, \vec{z} \in V$ and $c \in \mathbb{K}$. Then:

$$\langle \vec{x} + \vec{z}, \vec{y} \rangle = \langle \vec{x} + \vec{z}, \vec{y} \rangle_1 + \langle \vec{x} + \vec{z}, \vec{y} \rangle_2 =$$
$$= \langle \vec{x}, \vec{y} \rangle_1 + \langle \vec{z}, \vec{y} \rangle_1 + \langle \vec{x}, \vec{y} \rangle_2 + \langle \vec{z}, \vec{y} \rangle_2 = \langle \vec{x}, \vec{y} \rangle + \langle \vec{z}, \vec{y} \rangle$$

So we have that $\langle \cdot, \cdot \rangle$ satisfies condition (a) of an inner product. Now for condition (b), we have:

$$\langle c\vec{x}, \vec{y} \rangle = \langle c\vec{x}, \vec{y} \rangle_1 + \langle c\vec{x}, \vec{y} \rangle_2 = c\langle \vec{x}, \vec{y} \rangle_1 + c\langle \vec{x}, \vec{y} \rangle_2 =$$
$$= c(\langle \vec{x}, \vec{y} \rangle_1 + \langle \vec{x}, \vec{y} \rangle_2) = c\langle \vec{x}, \vec{y} \rangle$$

And thus condition (b) is satisfied. For condition (c), we have:

$$\overline{\langle \vec{x}, \vec{y} \rangle} = \overline{\langle \vec{x}, \vec{y} \rangle_1 + \langle \vec{x}, \vec{y} \rangle_2} = \overline{\langle \vec{x}, \vec{y} \rangle_1} + \overline{\langle \vec{x}, \vec{y} \rangle_2} =$$

$$\langle \vec{y}, \vec{x} \rangle_1 + \langle \vec{y}, \vec{x} \rangle_2 = \langle \vec{y}, \vec{x} \rangle$$

And so this condition is proved to hold. Now for the final condition (d), suppose that $\vec{x} \neq \vec{0}$. Then we know that $\langle \vec{x}, \vec{x} \rangle_1 > 0$ and $\langle \vec{x}, \vec{x} \rangle_2 > 0$. Thus:

$$\langle \vec{x}, \vec{x} \rangle = \langle \vec{x}, \vec{x} \rangle_1 + \langle \vec{x}, \vec{x} \rangle_2 > 0$$

So condition (d) is satisfied, and each of the conditions is satisfied for all vectors in V and scalars in $\mathbb{K}$. Therefore $\langle \cdot, \cdot, \rangle$ is an inner product on V. ■

Problem 6.1.14
Let A and B be $n \times n$ matrices, and let $c \in \mathbb{K}$. Then observe:

$$(A + cB)^* = \overline{(A + cB)^T} = \overline{A^T + cB^T} = \overline{A^T} + \overline{cB^T} = A^* + \overline{c}B^*$$

By the results of Problem 1.3.3 and the definition of the conjugate transpose of a matrix. ■

Problem 6.1.15 XXX
Let V be an inner product space, and let $\vec{x}, \vec{y} \in V$. Suppose $|\langle \vec{x}, \vec{y} \rangle| = ||\vec{x}|| \cdot ||\vec{y}||$. In the case where $\vec{y} = \vec{0}$, then the identity holds trivially. So suppose $\vec{y} \neq \vec{0}$. Let:

$$a = \frac{\langle \vec{x}, \vec{y} \rangle}{||\vec{y}||^2}$$

And let $\vec{z} = \vec{x} - a\vec{y}$. Then we see:

$$\langle \vec{z}, \vec{y} \rangle = \langle \vec{x} - (\frac{\langle \vec{x}, \vec{y} \rangle}{||\vec{y}||^2})\vec{y}, \vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle - (\frac{\langle \vec{x}, \vec{y} \rangle}{||\vec{y}||^2})\langle \vec{y}, \vec{y} \rangle =$$

$$= \langle \vec{x}, \vec{y} \rangle - \frac{\langle \vec{x}, \vec{y} \rangle}{||\vec{y}||^2} \cdot ||\vec{y}||^2 = \langle \vec{x}, \vec{y} \rangle - \langle \vec{x}, \vec{y} \rangle = 0$$

So that $\vec{z}$ and $\vec{y}$ are orthogonal vectors. Now note that:

$$|a| = \frac{|\langle \vec{x}, \vec{y} \rangle|}{||\vec{y}||^2} = \frac{||\vec{x}|| \cdot ||\vec{y}||}{||\vec{y}||^2} = \frac{||\vec{x}||}{||\vec{y}||}$$

Now consider the representation $\vec{x} = \vec{z} - a\vec{y}$. Clearly $||\vec{x}||^2 = ||\vec{z} - a\vec{y}||^2$. From Problem 6.1.10, we know that since $\vec{y}$ and $\vec{z}$ are orthogonal vectors in V, we have:

Problem 6.1.16

Problem 6.1.17
Let T be a linear operator on an inner product space V. Now suppose that $||T(\vec{x})|| = ||\vec{x}||$ for all $\vec{x} \in V$. Suppose that we have $\vec{y} \in N(T)$. Then we know that $T(\vec{y}) = \vec{0}$. But by the hypothesis, we have $||T(\vec{y})|| = ||\vec{0}|| = 0$, so necessarily $||T(\vec{y})|| = \sqrt{\langle \vec{y}, \vec{y} \rangle} = 0$. Squaring both sides, we easily find that $\langle \vec{y}, \vec{y} \rangle = 0$

However by Theorem 6.1(d), we know that this occurs if and only if $\vec{y} = \vec{0}$. Hence $N(T) = \{\vec{0}\}$ and so by Theorem 2.4, T is injective. ■

Problem 6.1.18

Problem 6.1.19
a.) Let V be an inner product space. Let $\vec{x}, \vec{y} \in V$. Then:

$$||\vec{x} + \vec{y}||^2 = \langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle =$$

$$= ||\vec{x}||^2 + 2Re\langle \vec{x}, \vec{y} \rangle + ||\vec{y}|||^2$$

b.)

Problem 6.1.20

Problem 6.1.21

Problem 6.1.22

Problem 6.1.23

Problem 6.1.24

Problem 6.1.25

Problem 6.1.26
Let $|| \cdot ||$ be a norm on a vector space V. Define the scalar $d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||$ for each ordered pair of vectors in V. Let $\vec{x}, \vec{y}, \vec{z} \in V$.

a.) Since $\vec{x} - \vec{y} \in V$ by closure, and condition (1) of the norm definition says that $||\vec{x}|| \geq 0$ for all $\vec{x} \neq 0$, we can easily see that:

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}|| \geq 0$$

b.) Given commutativity of vector addition from the definition of a vector space, we can see that:

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}|| = ||\vec{x} + (-\vec{y})|| = || - \vec{y} + \vec{x}|| = ||(-1)(\vec{y} - \vec{x}|| =$$

$$= |-1| \cdot ||\vec{y} - \vec{x}|| = ||\vec{y} - \vec{x}|| = d(\vec{y}, \vec{x})$$

c.) Consider the vectors $\vec{x} - \vec{z} \in V$ and $\vec{z} - \vec{y} \in V$. Then, by property (3) of the norm, we have:

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}|| = ||(\vec{x} - \vec{z}) + (\vec{z} - \vec{y})|| \leq$$

$$||\vec{x} - \vec{z}|| + ||\vec{z} - \vec{y}|| = d(\vec{x}, \vec{z}) + d(\vec{z}, \vec{y})$$

d.)

$$d(\vec{x}, \vec{x}) = ||\vec{x} - \vec{x}|| = ||\vec{0}|| = 0$$

e.)

$$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}|| \neq 0 = ||\vec{0}|| \implies \vec{x} - \vec{y} \neq \vec{0} \implies \vec{x} \neq \vec{y}$$

The function $d(\vec{x}, \vec{y})$ is the distance between $\vec{x}$ and $\vec{y}$.

Problem 6.1.27

Problem 6.1.28

Let V be a complex inner product space with inner product $\langle \cdot, \cdot \rangle$. Let the function $[\cdot, \cdot] : V \times V \to \mathbb{R}$ be a real-valued function such that $[\vec{x}, \vec{y}]$ is the real part of the complex number $\langle \vec{x}, \vec{y} \rangle$ for all $\vec{x}, \vec{y} \in V$. Suppose that V is regarded over the field $\mathbb{R}$. Then note:

$$[\vec{x} + \vec{z}, \vec{y}] = Re\langle \vec{x} + \vec{z}, \vec{y} \rangle = Re(\langle \vec{x}, \vec{y} \rangle + \langle \vec{z}, \vec{y} \rangle) =$$

$$= Re\langle \vec{x}, \vec{y} \rangle + Re\langle \vec{z}, \vec{y} \rangle = [\vec{x}, \vec{y}] + [\vec{z}, \vec{y}]$$

And so property (a) of an inner product holds. For property (b), let $c \in \mathbb{R}$, so we have:

$$[c\vec{x}, \vec{y}] = Re\langle c\vec{x}, \vec{y} \rangle = Re(c\langle \vec{x}, \vec{y} \rangle) = cRe\langle \vec{x}, \vec{y} \rangle = c[\vec{x}, \vec{y}]$$

And thus property (b) is satisfied. For property (c), we have that:

$$\overline{[\vec{x}, \vec{y}]} = Re\overline{\langle \vec{x}, \vec{y} \rangle} = Re\langle \vec{y}, \vec{x} \rangle = [\vec{y}, \vec{x}]$$

So (c) holds as well. Now for (d), suppose that $\vec{x} \neq \vec{0}$. Then:

$$[\vec{x}, \vec{x}] = Re\langle \vec{x}, \vec{x} \rangle = \langle \vec{x}, \vec{x} \rangle > 0$$

And so property (d) is satisfied. Therefore $[\cdot, \cdot]$ is an inner product on V. Furthermore, we see that:

$$[\vec{x}, i\vec{x}] = \bar{i}[\vec{x}, \vec{x}] = -i[\vec{x}, \vec{x}] \implies Re(-i\langle \vec{x}, \vec{x} \rangle) = 0$$

Which follows since i multiplying the inner product creates a complex number with no real component; a pure imaginary number.■

Problem 6.1.29

Problem 6.1.30

7 Canonical Forms

7.1 The Jordan Canonical Form I

Problem 7.1.1
a.) True. $E_\lambda \subseteq K_\lambda$ by the definition.

b.) False. Refer to discussion below the definition of generalized eigenvectors.

c.) False. The characteristic polynomial of the linear operator must split.

d.) True. This is the corollary to Theorem 7.6.

e.) False.

f.) XTrue. This is the result of Theorem 7.4(b).

g.) True.

h.) XFalse. This only holds true if $\dim(E_\lambda) = n$, not if $\dim(V) = n$.

Problem 7.1.2

Problem 7.1.3

Problem 7.1.4
Let T be a linear operator on a vector space V, and let γ be a cycle of generalized eigenvectors that correspond to eigenvalue λ. Consider the subspace of V $\text{span}(\gamma)$. Let $\vec{x} \in \text{span}(\gamma)$, so that we have:

$$\vec{x} = \sum_{i=0}^{p-1} a_i (T - \lambda I)^i(\vec{x}) \implies$$

$$T(\vec{x}) = T(\sum_{i=0}^{p-1} a_i (T - \lambda I)^i(\vec{x})) = \sum_{i=0}^{p-1} a_i (T - \lambda I)^{i+1}(\vec{x}) \in \text{span}(\gamma)$$

Since $\vec{x} \in \text{span}(\gamma)$ was arbitrary, we can see that $T(\text{span}(\gamma)) \subseteq \text{span}(\gamma)$. ■

Problem 7.1.5
Let $\gamma_1, \gamma_2, \ldots, \gamma_k$ be cycles of generalized eigenvectors of $T \in \mathcal{L}(V)$ corresponding to an eigenvalue λ. Suppose that the initial eigenvectors are distinct. Now suppose by way of contradiction that the cycles are not disjoint. Without loss of generality, suppose $\vec{x} \in \gamma_i \cap \gamma_j$, for some $i \neq j$ in $1 \leq i, j \leq k$. Then by the definition of a cycle of generalized eigenvectors, we know that there exists a smallest

$p \in \mathbb{Z}^+$ such that $(T - \lambda I_V)^p(\vec{x}) = \vec{0}$. However, now note that $(T - \lambda I_V)^{p-1}(\vec{x})$ is an eigenvector of T that is in $\gamma_i \cap \gamma_j$, contradicting the uniqueness of the initial eigenvector of the cycles. Hence the cycles $\gamma_1, \gamma_2, \ldots, \gamma_k$ are disjoint. ■

Problem 7.1.6

a.) Let $T \in \mathcal{L}(V, W)$. Suppose that $\vec{x} \in N(-T)$. Then $(-T)(\vec{x}) = -T(\vec{x}) = \vec{0} \implies T(\vec{x}) = \vec{0}$, so that $\vec{x} \in N(T)$. Hence $N(-T) \subseteq N(T)$. The reverse containment follows trivially, and we have $N(T) = N(-T)$. ■

b.) Let $T \in \mathcal{L}(V, W)$. Let $\vec{x} \in N((-T)^k)$. Then we have $(-T)^k(\vec{x}) = (-1)^k T^k(\vec{x}) = \vec{0} \implies T^k(\vec{x}) = \vec{0}$. Hence $N((-T)^k) \subseteq N(T^k)$. Again, the reverse containment follows trivially in much the same fashion as the first. We then have $N(T^k) = N((-T)^k)$. ■

c.) Now let $T \in \mathcal{L}(V)$, and let λ be an eigenvalue of T. Choose some $k \in \mathbb{Z}^+$. Then $N((T - \lambda I_V)^k) = N((-T + \lambda I_V)^k) = N((\lambda I_V - T)^k)$ by the results of parts (a) and (b) of this problem. ■

Problem 7.1.7

a.) Let $U \in \mathcal{L}(V)$, where $\dim(V) < \infty$. For our base case, we will show that the relation holds for the $k = 2$ case. Suppose $\vec{x} \in N(U)$, so we have $U(\vec{x}) = \vec{0} \implies U^2(\vec{x}) = U(U(\vec{x})) = U(\vec{0}) = \vec{0} \implies \vec{x} \in N(U^2)$. Hence we have $N(U) \subseteq N(U^2)$, so the base case holds. Now suppose that the relations holds for the kth case. We have $N(U) \subseteq N(U^2) \subseteq \cdots \subseteq N(U^k)$. Now suppose $\vec{x} \in N(U^k) \implies U^k(\vec{x}) = \vec{0}$. Applying U once more, we see $U^{k+1}(\vec{x}) = U(U^k(\vec{x})) = U(\vec{0}) = \vec{0}$, so that necessarily $N(U^k) \subseteq N(U^{k+1})$. Thus by induction the relations holds for the $k + 1$th case and we have:

$$N(U) \subseteq N(U^2) \subseteq \cdots N(U^k) \subseteq N(U^{k+1}) \subseteq \cdots$$

Which is the desired relation. ■

b.)

c.)

Problem 7.1.8

Let V be a finite-dimensional vector space and let $T \in \mathcal{L}(V)$ such that the characteristic polynomial of T splits. Consider the representation for any $\vec{x} \in V$ given in Theorem 7.3. If we take the results in Theorem 7.4, in conjunction with Theorem 5.25, we can clearly see that $V = \bigoplus_{i=1}^k K_{\lambda_i}$, where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct eigenvalues of T. Given the unique representation afforded by direct sums, it is clear that the representation of Theorem 7.3, $\vec{x} = \vec{v}_1 + \vec{v}_2 + \cdots + \vec{v}_k$, where $\vec{v}_i \in K_{\lambda_i}$ for $1 \leq i \leq k$, is unique as well. ■

Problem 7.1.9

Let V be a finite-dimensional vector space and let $T \in \mathcal{L}(V)$ such that the characteristic polynomial of T splits.

a.) We will prove Theorem 7.5(b). In addition to the above, let β be a basis for V such that β is the union of disjoint cycles of generalized eigenvectors of T. Then, by (a) and Problem 7.1.4, we know that for each cycle γ in β, span (γ) is T-invariant, and $[T|_{\text{span}\,(\gamma)}]_\gamma$ is a Jordan block. Then $[T]_\beta$ is the direct sum of the Jordan blocks, and so by the definition, β is a Jordan canonical basis for T. ■

b.) Suppose that β is a Jordan canonical basis for T. Let λ be an eigenvalue of T, and $\beta' = \beta \cap K_\lambda$. Since, by definition, β is the disjoint union of cycles of generalized eigenvectors, then if $\vec{v} \in \beta'$, then $\vec{v}$ is a generalized eigenvector of T corresponding to λ. XXX

Problem 7.1.10
Let V be a finite-dimensional vector space and let $T \in \mathcal{L}(V)$ such that the characteristic polynomial of T splits. Let λ be an eigenvalue of T.

a.) Suppose that γ is a basis for K_λ consisting of the union of q disjoint cycles of generalized eigenvectors corresponding to λ. Suppose $\dim(E_\lambda) = n$, so that there are n linearly independent eigenvectors of T corresponding to λ. Take each of the initial eigenvectors in the q disjoint cycles. Since γ is a basis for K_λ, each of the initial eigenvectors is distinct, and thus the set of initial vectors forms a linearly independent set of eigenvectors of T corresponding to λ. Thus $q \leq \dim(E_\lambda)$, for $n = \dim(E_\lambda)$ is the maximum number of linearly independent eigenvectors of T corresponding to λ. ■

b.) Let β be a Jordan canonical basis for T. Suppose that $J = [T]_\beta$ has q Jordan blocks with λ in the diagonal position for each block. Each of the Jordan blocks corresponds to a cycle of generalized eigenvectors of T corresponding to λ. Note that the initial eigenvectors of the cycles correspond to distinct eigenvalues of T. Hence, similar to the argument for (a), if $\dim(E_\lambda) = n$, then $q \leq n$, for there can be at most n linearly independent eigenvectors of T corresponding to λ. This gives us $q \leq \dim(E_\lambda)$. ■

Problem 7.1.11
We will prove Corollary 2 to Theorem 7.7. Let A be an $n \times n$ matrix such that the characteristic polynomial of A splits. Let $A = [L_A]_\gamma$, where γ is the standard ordered basis for $\mathbb{K}^n$. Thus the characteristic polynomial of L_A splits, and by Corollary 1 to Theorem 7.7, L_A has a Jordan canonical form, call it J. L_A also must have a Jordan canonical basis, call it β. Then $J = [L_A]_\beta$. We can then see that:

$$[L_A]_\beta = [I_{\mathbb{K}^n}]_\gamma^\beta [L_A]_\gamma [I_{\mathbb{K}^n}]_\beta^\gamma \implies J = QAQ^{-1}$$

So that A is similar to J as desired, proving the corollary. ■

Problem 7.1.12
Let T be a linear operator on a finite-dimensional vector space V such that the characteristic polynomial of T splits. T has distinct eigenvalues, say $\lambda_1, \ldots, \lambda_k$ with corresponding multiplicities $m_1, \ldots, m_k$. From Theorem 7.7, we know that for each eigenvalue λ_i, $1 \leq i \leq k$, of T, the subspace K_{λ_i} has an ordered basis β_i consisting of a union of disjoint cycles of generalized eigenvectors for each $1 \leq i \leq k$. Then, by Theorem 7.4, we know that the union of the ordered bases β_i of K_{λ_i} form a basis for V. Using the previous fact, Theorem 5.10(d) is satisfied, and this is a sufficient condition for Theorem 5.10(a), which is that $V = \bigoplus_{i=1}^{k} K_{\lambda_i}$. Therefore V is the direct sum of the generalized eigenspaces of T as desired. ■

Problem 7.1.13
Let V be a finite-dimensional vector space, let $T \in \mathcal{L}(V)$ such that the characteristic polynomial of T splits, and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of T. For each i, $1 \leq i \leq k$, let J_i be the Jordan canonical form of $T|_{K_{\lambda_i}}$, the restriction of T to K_{λ_i}. By Theorem 7.1, each K_{λ_i} is a T-invariant subspace of V, and further, from Theorem 7.7 we know that $V = \bigoplus_{i=1}^{k} K_{\lambda_i}$. Additionally, we have that each K_{λ_i} has a Jordan canonical basis, call it β_i for each $1 \leq i \leq k$, since they each possess a Jordan canonical form J_i. Hence, by Theorem 7.4, $\beta = \bigcup_{i=1}^{k} \beta_i$ is an ordered basis for V. Let $J = [T]_\beta$ and let $J_i = [T|_{K_{\lambda_i}}]_{\beta_i}$ for each $1 \leq i \leq k$. Then the hypothesis for Theorem 5.25 is satisfied, and we conclude that $J = \bigoplus_{i=1}^{k} J_i$. ■